LNCS 14750

Klaus Miesenberger
Petr Peňáz
Makoto Kobayashi (Eds.)

Computers Helping People with Special Needs

19th International Conference, ICCHP 2024
Linz, Austria, July 8–12, 2024
Proceedings, Part I

Part I

ICCHP AAATE

Springer

Lecture Notes in Computer Science 14750

The series Lecture Notes in Computer Science (LNCS), including its subseries Lecture Notes in Artificial Intelligence (LNAI) and Lecture Notes in Bioinformatics (LNBI), has established itself as a medium for the publication of new developments in computer science and information technology research, teaching, and education.

LNCS enjoys close cooperation with the computer science R & D community, the series counts many renowned academics among its volume editors and paper authors, and collaborates with prestigious societies. Its mission is to serve this international community by providing an invaluable service, mainly focused on the publication of conference and workshop proceedings and postproceedings. LNCS commenced publication in 1973.

Klaus Miesenberger · Petr Peňáz ·
Makoto Kobayashi
Editors

Computers Helping People with Special Needs

19th International Conference, ICCHP 2024
Linz, Austria, July 8–12, 2024
Proceedings, Part I

Springer

Editors
Klaus Miesenberger
Johannes Kepler University
Linz, Austria

Petr Peňáz
Masaryk University
Brno, Czech Republic

Makoto Kobayashi
Tsukuba University of Technology
Tsukuba, Japan

ISSN 0302-9743 ISSN 1611-3349 (electronic)
Lecture Notes in Computer Science
ISBN 978-3-031-62845-0 ISBN 978-3-031-62846-7 (eBook)
https://doi.org/10.1007/978-3-031-62846-7

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Welcome to the Proceedings of ICCHP 2024!

We are proud to put the proceedings of the 19th edition of the ICCHP conference in your hands. Since its inception in 1989, ICCHP has evolved to become one of the largest scientific conferences on accessibility, assistive technologies and digital technology-enhanced inclusion. Even if the digital revolution is still accelerating, ICCHP keeps its user-centered focus on socio-technical innovation and aims at integrating the rapid technological deployments into a more inclusive, flexible, personalized and adaptable digital society. This allows global technical, political, legal and social developments to be reflected in the frame of the ICCHP topics eAccessibility, assistive technologies and digital inclusion and in relation to such domains as eHealth and eCare, eLearning and eInclusion, eDemocracy and eGovernment, eServices and social inclusion, ambient and assisted active living, accessible traveling and tourism, user-centered design for all, and many other research fields.

We are proud and thankful that the call for contributions was answered by about 200 researchers and experts from more than 50 countries from all continents, with 266 submissions in all. ICCHP ran a highly competitive process for selecting the contributions for publication and presentation. The Program Committee including 140 experts guaranteed that each paper was reviewed by three to five experts in a double-blind process. One member of the panel of 17 conference chairs assessed the review results of each contribution and brought it into discussion at the decision meeting. This two-phase selection procedure guaranteed the high scientific quality, making ICCHP unique in our field. 104 contributions got accepted and you will find them in these proceedings embedded in thematically grouped chapters.

The contributions are grouped in chapters which are based on Special Thematic Sessions (STS) prepared and promoted by researchers and experts in these very specific domains. The introductions to the STSs provide the reader with a comprehensive overview of both the scientific field and the contributions included. With this approach ICCHP provides proceedings which are well readable and understandable to those in but also beyond our field, making them to a key reference for the science, research and practice for the implementation of a more open and accessible mainstream digital society.

ICCHP 2024 was hosted on July 8–12, 2024, at Johannes Kepler University Linz, providing an accessible event open to everyone interested in new and original ways to put technology at the service of people living with a disability. Participants contributed to the conference in multiple ways:

- Scientific contributions, which underwent the rigorous peer review process and are now available in two different publications:
- these *Springer Lecture Notes in Computer Science (LNCS)* volumes, which focus more on the development, engineering and computer science perspective of Assistive Technology and Accessibility, and
- the *Open Access Compendium "Assistive Technology, Accessibility and (e)Inclusion" (OAC)*, which focuses more on research related to policy, education, implementation and the social services perspective.

Papers accepted for both publication channels were presented in integrated thematic sessions, which facilitated a unique cross-disciplinary view and discussion in the specialized domains. The STS introductions included in these LNCS proceedings also reference and discuss the OAC contributions.

- Inclusion Area contributions, which provided a space for discussion, collaborations and participation of multiple stakeholders. Contributions to the Inclusion Area were reviewed by the program committee and were presented at the conference in different formats including workshops, tutorials, seminars, posters and policy sessions, where participants showcased projects, products and services. Inclusion Area contributions are published in the digital Book of Abstracts, which collects short descriptions of all contributions to the conference.
- Students presented their ongoing research at The Young Researchers' Consortium.

We would like to thank the Panel of Chairs, the Young Researchers Committee, the Inclusion Area Committee, the International Scientific Committee, the Organizing Committee and the student volunteers who dedicated precious time and efforts to the organization of this event. We thank all those who helped to put ICCHP in place and thereby supported the AT, digital accessibility and a better quality of life for people with disabilities. Special thanks go to all our supporters and sponsors, displayed at: https://www.icchp.org/sponsors-24.

Moreover, we want to thank all of the participants at the conference for furthering the mission of ICCHP, based on the belief that technology can contribute to breaking barriers, empowering people and enhancing equity and inclusion for people with all abilities.

July 2024

Makoto Kobayashi
Klaus Miesenberger
Petr Peňáz

The original version of this book had a typing error in the name of the third editor. "Makato Kobayashi" should have been "Makoto Kobayashi". This has now been corrected. A correction to this book can be found at https://doi.org/10.1007/978-3-031-62846-7_62

Organization

ICCHP Committees

Our special thanks go to those contributing to put this conference in place:

Conference Chair

Kobayashi, M.	Tsukuba University of Technology, Japan

Publication Committee

Miesenberger, K.	Johannes Kepler University Linz, Austria
Penáz, P.	Masaryk University Brno, Czech Republic

Scientific Committee Chairs

Archambault, D.	Université Paris 8, France
Bühler, C.	Technical University Dortmund, Germany
Coughlan, J.	Smith-Kettlewell Eye Research Institute, USA
Covarrubias Rodriguez, M.	Politecnico di Milano, Italy
Debevc, M.	University of Maribor, Slovenia
Fels, D.	Toronto Metropolitan University, Canada
Kouroupetroglou, G.	University of Athens, Greece
Manduchi, R.	University of California at Santa Cruz, USA
Mavrou, K.	European University Cyprus, Cyprus
Weber, G.	Technical University Dresden, Germany
Zagler, W.	TETRAGON Braille Systems GmbH, Austria

Young Researchers Committee

Archambault, D.	University Paris 8, France
Chen, W.	University of Bergen, Norway
Fels, D.	Toronto Metropolitan University, Canada
Fitzpatrick, D.	National Disability Authority, Ireland

Kobayashi, M.	Tsukuba University of Technology, Japan
Morandell, M.	Smart In Life, Health Univ. of Appl. Sciences Tirol, Austria
Pagliara, S. M.	University of Cagliari, Italy
Pontelli, E.	New Mexico State University, USA
Prazak-Aram, B.	Ferdinand Porsche FERNFH, Austria
Ruh, D.	RuhGlobal, USA
Weber, G.	Technische Universität Dresden, Germany
Zimmermann, G.	Stuttgart Media University, Germany

Innovation Area Committee

Pavlicec, R.	Masaryk University Brno, Czech Republic
Petz, A.	Johannes Kepler University Linz, Austria
Puehretmair, F.	KI-I, Austria

Scientific Committee

Abbott, C.	King's College London, UK
Abou-Zahra, S.	Amazon, Austria
Andreoni, G.	Politecnico di Milano, Italy
Andrich, R.	EASTIN, Italy
Atkinson, M.	Samsung, UK
Augstein, M.	Upper Austria Univ. of Applied Sciences, Austria
Azevedo, L.	Instituto Superior Tecnico, Portugal
Banes, D.	Dave Banes Access, UK
Batusic, M.	Fabasoft, Austria
Bernareggi, C.	Univ. degli Studi di Milano, Italy
Boudy, J.	Télécom SudParis, France
Bu, J.	Zhejiang University, China
Caruso, G.	Politecnico di Milano, Italy
Chamberlain, H.	Helen Chamberlain Consulting, USA
Christensen, L. B.	Sensus, Denmark
Chutimaskul, W.	King Mongkut's Techn. Univ. Thonburi, Thailand
Craddock, G.	Centre for Excellence in Univ. Design, Ireland
Darvishy, A.	ZHAW Univ. of Applied Sciences, Switzerland
De Witte, L.	The Hague Univ. of Applied Sciences, The Netherlands
Debeljak, M.	University of Ljubljana, Slovenia
DeRuyter, F.	Duke University Medical Centre, USA

Desideri, L.	AIAS Bologna Onlus, Italy
Diaz del Campo, R.	Antarq Tecnosoluciones, Mexico
Draffan, E. A.	University of Southampton, UK
Dupire, J.	CNAM, France
Ebling, S.	University of Zurich, Switzerland
Emiliani, P. L.	Institute of Appl. Physics “Nello Carrara”, Italy
Encarnação, P.	Catolica Lisbon, Portugal
Engelen, J.	Katholieke Univ. Leuven, Belgium
Galinski, Ch.	InfoTerm, Austria
Gardner, J.	Oregon State University, USA
Gherardini, A.	AIAS Bologna Onlus, Italy
Hakkinen, M. T.	Educational Testing Service, USA
Heimgaertner, R.	IUIC, Germany
Heumader, P.	Fabasoft, Austria
Hoeckner, K.	Hilfgemeinschaft Blinde & Sehschwache, Austria
Hoogerwerf, E. J.	AIAS Bologna Onlus, Italy
Holone, H.	Østfold University College, Norway
Hosono, N.	NPO Niimaru, Japan
Inoue, T.	Nat. Rehab. Center for PwD, Japan
Iversen, C. M.	U.S. Department of State (Retired), USA
Jitngernmadan, P.	Burapha University, Thailand
Kacorri, H.	University of Maryland, USA
Kalinnikova, L.	University of Gaevle, Sweden
Kiswarday, V.	University of Primorska, Slovenia
Koumpis, A.	University of Passau, Germany
Kozuh, I.	University of Maribor, Slovenia
Kueng, J.	JKU Linz, Austria
Kunz, A.	ETH Zurich, Switzerland
Kushalnagar, R.	Gallaudet University, USA
Leblois, A.	G3ict, USA
Lee, C.	G3ict, IAAP, USA
Leporini, B.	Italian Nat. Research Council (CNR), Italy
Lhotska, L.	Czech Technical University in Prague, Czech Republic
Magnusson, Ch.	Lund University, Sweden
Magnusson, M.	Moscow State University, Russia
Matausch-Mahr, K.	Inst. Integriert Studieren JKU Linz, Austria
Mohamad, Y.	Fraunhofer Institute, Germany
Moon, I.	Dong-Eui University, South Korea
Mrochen, I.	MultiAccess Centre, Poland
Muratet, M.	INS HEA, France
Murillo-Morales, T.	Inst. Integriert Studieren JKU Linz, Austria

Nussbaum, G.	KI-I, Austria
Paciello, M.	Pearson, USA
Panek, P.	Vienna University of Technology, Austria
Petrie, H.	University of York, UK
Petz, A.	Inst. Integriert Studieren, JKU Linz, Austria
Petz, P.	Upper Austria Univ. of Applied Sciences, Austria
Pissaloux, E.	University of Rouen Normandy, France
Rassmus-Groehn, K.	Lund University, Sweden
Raynal, M.	Toulouse University, France
Rice, D.	National Disability Authority, Ireland
Scherer, M.	Inst. Matching Person & Technology, USA
Simsik, D.	University of Kosice, Slovakia
Slavik, P.	Czech Technical University in Prague, Czech Republic
Sloan, D.	Paciello Group, UK
Stiefelhagen, R.	Karlsruhe Institute of Technology, Germany
Stoeger, B.	Johannes Kepler University Linz, Austria
Takahashi, Y.	Toyo University, Japan
Teixeira, A.	Universidade de Aveiro, Portugal
Teshima, Y.	Chiba Institute of Technology, Japan
Tjoa, A. M.	Technical University of Vienna, Austria
Truck, I	Paris 8 University, France
Velleman, E.	Accessibility Foundation, The Netherlands
Vigo, M.	University of Manchester, UK
Vigouroux, N.	IRIT Toulouse, France
Vogler, C.	Gallaudet University, USA
Wada, C.	Kyushu Institute of Technology, Japan
Wagner, G.	Upper Austria Univ. of Applied Sciences, Austria
Wakatsuki, D.	Tsukuba University of Technology, Japan
Waszkielewicz, A.	FRONia, Poland
Watanabe, T.	University of Niigata, Japan
Weber, H.	ITA, RPTU, Germany
Yamaguchi, K.	Nihon University, Japan
Yesilada, Y.	Middle East Technical University, Cyprus

Organization Committee

Bieber, R.	Austrian Computer Society, Austria
Bukovský, T.	Masaryk University Brno, Czech Republic
Cincibus, Z.	Masaryk University Brno, Czech Republic
Feichtenschlager, P.	Institute Integriert Studieren, JKU Linz, Austria

Klug, R.	Institute Integriert Studieren, JKU Linz, Austria
Miesenberger, K.	Institute Integriert Studieren, JKU Linz, Austria
Mück, T.	Austrian Computer Society, Austria
Ondra, S.	Masaryk University Brno, Czech Republic
Pavlíček, R.	Masaryk University Brno, Czech Republic
Petz, A.	Institute Integriert Studieren, JKU Linz, Austria
Schult, C.	Institute Integriert Studieren, JKU Linz, Austria
Wernerová, P.	Masaryk University Brno, Czech Republic

ICCHP Roland Wagner Award Committee

Archambault, D.	Université Paris 8, France
Bühler, C.	TU Dortmund, FTB Volmarstein, Germany
Draffan, E. A.	University of Southampton, UK
Höckner, K.	Hilfsgemeinschaft Blinde & Sehschwache, Austria
Miesenberger, K.	Institute Integriert Studieren, JKU Linz, Austria
Zagler, W.	TETRAGON Braille Systems GmbH, Austria

Contents – Part I

Art Karshmer Lectures in Access to Mathematics, Science and Engineering

Tactile Graphics and 3D Models for Blind People and Shape Recognition by Touch

Contents – Part II

Interaction Techniques for Motor Disabled Users

Augmentative and Alternative Communication Innovations in Products and Services

Advanced Technologies for Innovating Inclusion and Participation in Labour, Education, and Everyday Life

Disability, Inclusion, Service Provision, Policy and Legislation

Software, Web and Document Accessibility

Software, Web and Document Accessibility

Introduction to the Special Thematic Session

Nataša Rajh[1(✉)], Matjaž Debevc[2], Klaus Höckner[3], Reinhard Koutny[1], and Klaus Miesenberger[1]

[1] Institut Integriert Studieren, Johannes Kepler University, Linz, Austria
Natasa.Rajh@partner.jku.at, {Reinhard.Koutny, Klaus.Miesenberger}@jku.at

[2] Institute of Media Communications, Faculty of Electrical Engineering and Computer Science, University of Maribor, Maribor, Slovenia
matjaz.debevc@guest.um.si

[3] Hilfsgemeinschaft der Blinden und Sehschwachen Österreichs, Vienna, Austria
hoeckner@hilfsgemeinschaft.at

https://www.jku.at/institut-integriert-studieren/

Abstract. This STS prolongs the long-lasting series of publications focusing on digital accessibility as one of the key pillars of ICCHP since its start in 1987. This session consists of a wide range of publications, which tackle different dimensions of software and document accessibility. The topics range from as wide as software engineering processes in general to as specific as an example of a concept of an accessible illustrated digital book. The vast array of interests and varying levels of specificity within different topics highlights the broadness of the research area and the ample room for further exploration.

Keywords: Software accessibility · Digital Accessibility · Web accessibility · Web accessibility evaluation · PDF accessibility · eBooks · Assistive technologies

1 Introduction

Due to its multimedia potential for adapting, optimizing, and presenting content for different senses (vision, hearing, haptic) and the multiformity of input and manipulation through a broad range of assistive technologies (AT), digital technology has become the central convergence point for supporting people with disabilities in society The more digital advancements occur due to the digital revolution, the greater the potential for interaction among mainstream digital technology, standardized Human-Computer Interface, and digital assistive technology (AT). This interplay only works if accessibility is considered and respected in all stages of design, development and content generation/provision.

There is an enormous inclusive potential based on adaptability and personalization which no other technology so far could provide. Therefore, the potential and the need

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 3–10, 2024.
https://doi.org/10.1007/978-3-031-62846-7_1

to respect accessibility has been promoted by the stakeholder groups of people with disabilities, started by the independent living movement in parallel to the upcoming digital revolution over the last 50 and more years.

Due to its obvious benefits for the target group, service providers, and the social sector, along with its positive impact on general usability and fostering a more open and democratic society, accessibility has gained recognition as a key socio-technical principle in policy, legislation, and administration. This recognition is best underlined by the UN Convention on the Rights of People with Disabilities (UNCRPD) [1] and the related legal implementations (e.g. the Americans with Disabilities Act [2], the European Disability Act [3], …).

In particular, digital accessibility has been on the agenda of IT R&D and industry over the last decades leading well elaborated standards and techniques, developed in a global cooperation of all stakeholder groups and available at platform-independent level (e.g. W3C/WAI/WCAG [4], EN 301 549 [5], ISO/IEC Guide 71:2014 [6], ISO 14289 (PDF/UA) [7], DAISY Consortium [8], etc.).

Today, when prominent voices such as Jakob Nielsen [9] outline that "Accessibility has failed," we have to and can give a strong answer to emphasize the remarkable progress, to show that these principles and standards are implemented and well-supported across all major platforms for system design, development and implementation. This stands as one of the key outcomes of the global digital accessibility cooperation.[1]

But we have to take these critiques and concerns seriously. It is evident that accessibility has a "last mile problem" in terms of uptake by everybody for any piece of digital content. Very often accessibility is not implemented and part of our day-to-day practice. Even within the sector itself, we can find many examples of non or poor digital accessibility. There is a strong need for further R&D to better support the implementation of accessibility in the design, engineering and content provision process at the level of Integrated Development Environments (IDEs), Content Management Systems (CMS) and content development/provision at web, document or software level in general.

Therefore, the question is no longer whether to implement digital accessibility, but rather how to do so efficiently and effectively in the design, engineering, and content provision process.

2 Overview of the Contributions

The compilation of contributions presented in this STS offers a comprehensive overview of the current state and future directions in accessibility research and practice. With a total of fourteen papers [10–23], the collection is organized into four thematic themes that collectively address the vast spectrum of accessibility concerns within the digital landscape. This organization not only facilitates a focused exploration of each theme but also highlights the interdisciplinary nature of accessibility efforts, underscoring the importance of collaboration across various domains to advance digital inclusivity.

The first theme, "Accessibility in the engineering, design, content provision, and administration process," encompasses papers that delve into the foundational aspects of

[1] ICCHP 24 plans for a keynote and a panel discussion on this topic which is closely related to this STS.

creating accessible digital environments. This group stresses the importance of integrating accessibility from the beginning of engineering and design processes through content provision and administrative practices, advocating for a proactive approach throughout the developmental lifecycle.

The second theme, “Mobile accessibility,” focuses on the challenges and opportunities associated with making mobile technologies accessible to all users, including those with disabilities. Papers within this group explore the usability and accessibility barriers present in mobile applications. The papers highlight the need for inclusive design practices, to ensure mobile technologies bridge, not hinder, digital participation.

“Digital document accessibility” forms the third theme, concentrating on the accessibility of digital documents, with a particular focus on PDFs. This section sheds light on the significant challenges faced in making academic publications and other digital documents accessible. The contributions explore innovative approaches for creating accessible PDFs and highlight the urgency of addressing document accessibility in the broader context of digital inclusivity.

The final theme, “Evaluation of and Education for Accessibility,” addresses the critical role of evaluation tools and educational initiatives. This theme includes studies on developing tools for assessing digital content accessibility and curriculums for educating novices and experts. It emphasizes continuous assessment and improvement in accessibility practices and fostering understanding through hands-on learning.

The contributions in this STS offer valuable insights into digital accessibility, emphasizing its multifaceted nature. Grouping the papers into four thematic areas underscores the interconnectedness of engineering, design, mobile technologies, digital documents, and educational efforts in promoting digital inclusivity. Each theme contributes to a holistic understanding of the challenges and opportunities, reflecting a collective commitment to a more inclusive digital future.

2.1 Accessibility in the Engineering, Design, Content Provision and Administration Process

Many accessibility issues arise due to a lack of consideration for accessibility in the process. Even if accessibility is integrated into all platforms and technically well-supported, it is frequently disregarded during the process. In the landscape of digital innovation and technical infrastructure, ensuring accessibility for all individuals, including persons with disabilities, is paramount. A collection of studies sheds light on the strides being made towards integrating accessibility into various facets of engineering, design, content provision, and administrative processes. These studies not only highlight the challenges faced in making technologies and environments accessible but also propose forward-thinking solutions and frameworks aimed at overcoming these obstacles.

Accessibility within the Software Engineering (SE) Process is a well-researched, complex and broad field. Despite substantial research efforts in the past, several areas remain without definitive frameworks and proposals, warranting further investigation. Rajh et al.’s Systematic Literature Research paper about SE Process and Integrated Development Environments [10] aims to provide an overview of previous research within the field and identify any potential gaps in the literature. By performing a SLR about previous related SLRs and a new SLR, the authors were searching for guidelines, standards, methods, techniques and tools available to support development of accessible software. In addition, they wanted to know how accessibility support is involved in IDEs. While many resources were identified, the literature on accessibility integration in IDEs was found to be sparse. Furthermore, the lack of approaches encompassing all phases of the SE process was also emphasized.

Engaging end-users in the design process isn't merely one strategy; rather, it can significantly enhance the ability to create inclusive products and services. In Gordon et al.'s work "Towards a Framework of Inclusive Software Design Process and Practices" [11] multiple design models and design processes of end-user involvement are explored, to deepen understanding of applying inclusive design in real-world software projects and improve facilitating participatory processes in the software industry.

The paper "Empowering Ethical Assistive Technologies: A User-Centric Consent Control Interface", by Vaidya and Stiernet [12], presents an innovative solution, the User-Centric Consent Control Interface (UCCCI), designed to address privacy concerns and empower users, including those with disabilities, to have control over their data and the technologies they utilize. This initiative is pivotal in fostering trust and reliability in assistive technologies by prioritizing user autonomy, privacy, and transparency. By providing users with intuitive mechanisms for managing consent, the UCCCI exemplifies a significant advancement in the ethical development and deployment of assistive technologies, ensuring they are both accessible and respectful of user rights.

2.2 Mobile Accessibility

The examination of user experience inequalities, particularly within mobile applications, underscores a significant concern: mobile interfaces and applications, while designed to enhance convenience for sighted users, often inadvertently erect barriers for users with disabilities. To tackle those inequalities, Tanwar and Rao [13] performed usability tests, to reveal user experience elements that create barriers for people who rely on screen readers in their paper "Inequality In User Experience: Can Mobile User Interfaces That Help Sighted Users Create Barriers For Visually Challenged People?".

The discussion around the accessibility of mobile apps for students in higher education further amplifies this concern. Mobile applications, especially those aimed at supporting mental wellbeing, are critical tools that can offer support and resources for students facing stress, anxiety, and other mental health challenges. To understand the digital barriers and present recommendations, Simsek and Chen [14] performed a heuristic evaluation and user testing, described in "The Accessibility and Usability of Mobile Apps for Students' Mental Wellbeing in Higher Education".

A promising approach for addressing these accessibility challenges is use of simulation frameworks, as described by Garfias and Namboodiri in "On the use of a Simulation Framework for Studying Accessibility Challenges People with Disabilities within Indoor Environments" [15]. By simulating real-world environments and scenarios, developers and designers can gain insights into the barriers faced by persons with disabilities. Simulation frameworks, facilitating controlled testing of accessibility features, can bridge the gap between current mobile app accessibility and universal design ideals.

2.3 Digital Document Accessibility

The discourse on digital document accessibility, with a particular focus on PDFs, is enriched by the findings from four pivotal studies. They collectively advance our understanding of the barriers and opportunities for document inclusivity enhancements.

Pierrès et al.'s [16] work on PDF accessibility in international academic publishing highlights a significant issue: the systemic inaccessibility of PDF documents disseminated by academic publishers. This study illuminates the challenges faced by users with disabilities, stemming from the lack of structured tagging and metadata that are crucial for screen reader compatibility. This absence not only impedes access to scholarly work but also contradicts the foundational principles of academic inclusivity and knowledge sharing.

Meanwhile, Szentirmai et al.'s [17] reflection on "The Accessibility Paradox" confronts the ironic reality that research dedicated to accessibility often fails to adhere to accessible publishing practices. This paradox not only highlights a gap in the commitment to inclusivity among researchers and publishers but also calls for a cultural shift towards consistent implementation of accessibility standards in academic outputs.

Osthof et al.'s [18] examination of tools designed to remediate PDF accessibility issues complements this by identifying critical usability challenges in leading accessibility tools. The findings, indicating that none of the evaluated tools achieved a satisfactory usability score, underscore the necessity for more intuitive, user-friendly solutions that empower users to create accessible documents without undue burden.

Further enriching this dialogue, Gharbieh et al. [19] in "Flexi Picture eBook (FPB)" introduces an innovative approach to improving digital content accessibility beyond textual data. By focusing on adaptable non-text elements, the study presents a model for enhancing the usability and accessibility of complex content formats within digital documents. This approach not only addresses the needs of users with visual impairments but also sets a new standard for the design of inclusive digital educational materials.

2.4 Evaluation of and Education for Accessibility

Lange et al.'s work on "Tools for Novice and Expert Accessibility Professionals" [20] highlights the pressing need for next-generation web accessibility evaluation tools that cater to a diverse user base. The study underscores the importance of redesigning tool lists to be more intuitive and user-friendly, thus enabling both novice and experienced professionals to efficiently find and utilize the right tools for accessibility evaluation. This initiative aligns with the broader goal of universal digital accessibility, enabling everyone to contribute to accessible digital environments.

Complementing the push for better tools, Ara and Sik-Lanyi [21] introduce a declarative model, a sophisticated approach aimed at refining the accuracy and reliability of automated web accessibility testing tools. By addressing the limitations of existing tools, Ara's model proposes a comprehensive framework that incorporates updated guidelines, user criteria, and a simplified guideline interpretation process. This model is instrumental in generating more transparent and understandable accessibility evaluation reports, making it easier for developers and evaluators to identify and rectify accessibility issues.

Moreover, the "Developing A Model Curriculum On How To Provide Actionable Feedback On Web Accessibility Issues" by Baumann and Dirks [22] emphasizes the crucial role of education in promoting web accessibility. By equipping individuals, especially those with disabilities, with the knowledge and skills to provide actionable feedback on accessibility barriers, this curriculum fosters a participatory approach to digital accessibility. It highlights the necessity of integrating accessibility considerations into the educational programs of Vocational Education and Training (VET) providers and Organizations of Persons with Disabilities (OPD), ensuring that future generations of digital technology users and developers are well-versed in accessibility principles.

Petrie and Darzentas' study on "Introducing Computer Science Students to Designing for Accessibility" [23] further illustrates the importance of hands-on learning accessibility through practical exercises using a low-cost simulation kit. By simulating the experiences of individuals with disabilities and aging, this educational approach nurtures empathy and insight, encouraging future developers to prioritize accessibility in their designs.

3 Discussion

The discussion on "Software, Web, and Document Accessibility" emphasizes the vital role of accessibility across various digital platforms. These studies stress the need for inclusive design practices, advocating for the integration of accessibility considerations into the design and development process. They also highlight the importance of using tools for accessibility evaluation and promoting accessibility awareness through education and training. The interconnectedness of software, web, and document accessibility is evident, with advancements in one area driving progress across the entire digital accessibility landscape.

The synthesis of findings from these studies provides a roadmap for advancing digital inclusivity. By embracing innovative solutions, fostering collaboration among stakeholders, and prioritizing the needs of individuals with disabilities, we can create a digital environment that is accessible, usable, and welcoming to all. The challenge is now to translate these insights into actionable strategies for meaningful change in software, web, and document accessibility, reflecting the diversity and inclusivity of society.

References

1. United Nations: Convention on the Rights of People with Disabilities 2006. https://social.desa.un.org/issues/disability/crpd/convention-on-the-rights-of-persons-with-disabilities-crpd. Accessed 08 Apr 2024
2. ADA.gov: The Americans with Disabilities Act. https://www.ada.gov/. Accessed 08 Apr 2024
3. European Commission: European Accessibility Act (EAA). https://ec.europa.eu/social/main.jsp?catId=1202. Accessed 08 Apr 2024
4. W3C/WAI/WCAG: WCAG 2 Overview. https://www.w3.org/WAI/standards-guidelines/wcag/. Accessed 08 Apr 2024
5. ETSI: EN 301 549 V3 the harmonized European Standard for ICT Accessibility. https://www.etsi.org/human-factors-accessibility/en-301-549-v3-the-harmonized-european-standard-for-ict-accessibility?jjj=1712920351854. Accessed 08 Apr 2024

6. ISO: ISO/IEC Guide 71:2014 Guide for addressing accessibility in standards. https://www.iso.org/standard/57385.html. Accessed 08 Apr 2024
7. ISO: ISO 14289-2:2024: Document management applications: Electronic document file format enhancement for accessibility. https://www.iso.org/standard/82278.html. Accessed 08 Apr 2024
8. The DAISY Consortium. https://daisy.org/. Accessed 08 Apr 2024
9. Jakob Nielsen: Accessibility Has Failed: Try Generative UI = Individualized UX. https://jakobnielsenphd.substack.com/p/accessibility-generative-ui. Accessed 08 Apr 2024
10. Rajh, N., Miesenberger, K., Koutny, R.: Accessibility in the software engineering (SE) process and in integrated development environments (IDEs): a systematic literature review. In: Miesenberger, K., Kobayashi, M., Penaz, P. (eds.) ICCHP 2024. LNCS, pp. 11–18. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-62846-7_2
11. Gordon, D., Curley, A., Murphy, E., Gilligan, J., Becevel, A., O'Sullivan, D.: Towards a framework of inclusive software design process and practices. In: Petz, A., Miesenberger, K. (eds.) ICCHP Open Access Compendium "Future Perspectives on Accessibility, AT and eInclusion"; Linz, Austria, Proceedings. Johannes Kepler University Linz (2024)
12. Vaidya, S., Stiernet, M.M.: Empowering ethical assistive technologies: a user-centric consent control interface. In: Petz, A., Miesenberger, K. (eds.) ICCHP Open Access Compendium "Future Perspectives on Accessibility, AT and eInclusion"; Linz, Austria, Proceedings. Johannes Kepler University Linz (2024)
13. Tanwar, S., Rao, P.V.M.: Inequality in user experience: can mobile user interfaces that help sighted users create barriers for visually challenged people? In: Miesenberger, K., Kobayashi, M., Penaz, P. (eds.) ICCHP 2024. LNCS, pp. 19–30. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-62846-7_3
14. Simsek, H., Chen, W.: The accessibility and usability of mobile apps for students' mental wellbeing in higher education. In: Miesenberger, K., Kobayashi, M., Penaz, P. (eds.) ICCHP 2024. Springer (2024)
15. Garfias, F.J.R., Namboodiri, V.: On the use of a simulation framework for studying accessibility challenges people with disabilities within indoor environments. In: Miesenberger, K., Kobayashi, M., Penaz, P. (eds.) ICCHP 2024. LNCS, pp. 31–37. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-62846-7_4
16. Pierrès, O., Schmitt-Koopmann, F., Darvishy, A.: PDF accessibility in international academic publishers. In: Miesenberger, K., Kobayashi, M., Penaz, P. (eds.) ICCHP 2024. LNCS, pp. 38–46. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-62846-7_5
17. Szentirmai, A.B., Inal, Y., Torkildsby, A.B.: The accessibility paradox: can research articles inspecting accessibility be inaccessible? In: Miesenberger, K., Kobayashi, M., Penaz, P. (eds.) ICCHP 2024. LNCS, pp. 47–54. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-62846-7_6
18. Osthof, L.M., Schwarz, T., Müller, K.: Testing usability of tools for making PDFs accessible: pressing issues and pain points. In: Miesenberger, K., Kobayashi, M., Penaz, P. (eds.) ICCHP 2024. LNCS, pp. 55–62. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-62846-7_7
19. Gharbieh, D., Punz, M., Miesenberger, K., Salinas-Lopez, V.: Flexi Picture eBook (FPB) - adaptable non-text content for better accessibility and usability. In: Miesenberger, K., Kobayashi, M., Penaz, P. (eds.) ICCHP 2024. LNCS, pp. 63–71. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-62846-7_8
20. Lange, V., Swart, C.V., Jeurens, J., Velleman, E.M.: Tools for novice and expert accessibility professionals: requirements for the next generation Web Accessibility Evaluation Tools list. In: Miesenberger, K., Kobayashi, M., Penaz, P. (eds.) ICCHP 2024. LNCS, pp. 72–83. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-62846-7_9

21. Ara, J., Sik-Lanyi, C.: A declarative model for web content accessibility evaluation process. In: Miesenberger, K., Kobayashi, M., Penaz, P. (eds.) ICCHP 2024. LNCS, pp. 84–92. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-62846-7_10
22. Baumann, L., Dirks, S.: Developing a model curriculum on how to provide actionable feedback on web accessibility issues. In: Miesenberger, K., Kobayashi, M., Penaz, P. (eds.) ICCHP 2024. LNCS, pp. 93–99. Springer, Cham (2024)
23. Petrie, H., Darzentas, J.: Introducing computer science students to designing for accessibility - evaluation of practical exercises with a low-cost simulation kit. In: Miesenberger, K., Kobayashi, M., Penaz, P. (eds.) ICCHP 2024. LNCS, pp. 100-108. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-62846-7_12

Accessibility in the Software Engineering (SE) Process and in Integrated Development Environments (IDEs): A Systematic Literature Review

Nataša Rajh(✉), Klaus Miesenberger, and Reinhard Koutny

Institut Integriert Studieren, Johannes Kepler University, Linz, Austria
Natasa.Rajh@partner.jku.at, {Klaus.Miesenberger, Reinhard.Koutny}@jku.at
https://www.jku.at/institut-integriert-studieren/

Abstract. Software accessibility, once relatively unknown, is now recognized as both crucial and legally mandated, accumulating significant research attention in Computer Science. While various guidelines, methods, and tools have been developed, practical implementation still faces challenges, with persistent barriers in many products. Research often focuses on requirements engineering and evaluation, neglecting the implementation phase. While acknowledging the necessity of including accessibility throughout the engineering process, research on comprehensive approaches has been sparse, primarily focusing on requirements engineering and evaluation, neglecting implementation phase. We see integrating accessibility support into the core of design and development as a crucial step for improvements, revealing a need for integrating accessibility evaluation support into integrated development environments (IDEs). This paper aims to present a collection of techniques, methods and tools that support accessibility involvement within different stages of the engineering process, and additionally to provide a theoretical foundation for research and development of incorporating accessibility evaluation support within IDEs. Thereafter, we analyzed existing related Systematic Literature Reviews (SLRs) and complemented the findings with a new SLR. The study provides a solid base for advanced approaches in integrating accessibility into software engineering (SE).

Keywords: Software accessibility · Software engineering · Systematic literature review · Integrated development environments · Engineering assets

1 Introduction

Most software stakeholders nowadays are aware that they need to make their products accessible, but accessibility is still most often considered just as an afterthought in the testing process. Numerous researchers [1–3] have emphasized the necessity of considering accessibility throughout the entire engineering process, primarily focusing on the development life cycle. However, the transition from recognizing the importance of

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 11–18, 2024.
https://doi.org/10.1007/978-3-031-62846-7_2

accessibility to its tangible integration hasn't been put into practice. This gap may be attributed to the limited availability of research addressing the entire engineering process comprehensively. While widely acknowledged, there's a scarcity of approaches that efficiently guide involvement throughout the entire process. Even rarer are studies on procedures, workflows, and technical approaches for integrating accessibility checking—based on check tools, assistive technology, manual checking, and code inspection—into integrated development environments (IDEs), which is the focus of our research.

Accessibility not being efficiently integrated into the Software Engineering process results in software products at risk of including barriers. This remains a major issue, despite extensive research, the availability of many resources, and the expansion of global legal requirements. If we look at websites from the European Union' public body sectors, for example, they are obliged by law to be accessible [4]. Nevertheless, the results from the first monitoring period of Web accessibility Directive showed that only 4 out of 800 websites inspected with the in-depth method were fully accessible [5].

In order to gain a thorough overview of existing research addressing accessibility involvement in SE process, to find potential flaws and needs, we defined the following research questions:

RQ1: What are guidelines, standards, methods, techniques and tools available to support development of accessible software?
RQ2: What are the procedures, workflows and technical approaches for including accessibility checking into integrated development environments (IDEs)?

From our previous research activities, we were aware that multiple SLRs about web and/or software accessibility were performed in the past. For this reason, our approach involved systematically identifying existing relevant SLRs. We aimed to complement existing findings with our supplementary SLR, focusing on the SE process and IDEs.

To our knowledge, no published SLR has been covering years 2022 and 2023, so we decided to only focus on those 2 years for RQ1. Also, none of the existing SLRs focused on IDEs, so for RQ2 we did not have timeframe limitations.

As the paper is an SLR, we aren't including a separate State of the Art chapter.

Due to page limitations, we are not able to include all details of our SLR process nor list all analyzed papers, however they are available at: https://bit.ly/3TTrfNP.

2 Methodology

To perform a SLR, we followed process guidelines proposed by Kitchenham et al. [6].

As part of the planning part, we defined two research questions stated above. In total we performed three searches, with three search queries: searching for existing SLRs related to our topic, searching for answers for RQ1 in years 2022 and 2023, and searching for answers for RQ2.

Search was performed in January 2024. To search for previously done SLRs related to our research topic, we searched in the following databases: IEEE Explore, ACM Digital Library, Science Direct, Springer Link, Scopus, Web of Science and Google Scholar. The selected databases use different syntaxes and advanced searching possibilities, so the search strings had to be adapted for each one. Due to limitations of Springer Link and Google Scholar, resulting in over 13,000 results from the former and approximately 2 million from the latter, we opted not to utilize them for our own SLR addressing RQ1 and RQ2.

2.1 Existing SLRs

Search Query. In their SLR on SLRs, Kitchenham et al. [7] employed 15 search terms, whereas we opted for ‘systematic literature review’ only. As of January 28, 2024, their paper [6] has garnered 1345 citations on Google Scholar, indicating the widespread recognition and usage of the term ‘SLR’ (Systematic Literature Review) in the field. We employed the following query (used for IEEE Explore and adapted for other databases):

(“Document Title”: web accessibility systematic literature review) OR (“Document Title”:software accessibility systematic literature review)

Inclusion Criteria. The title concludes that the paper is an SLR, or similar, related to SE as a whole, a part of the process, or otherwise closely related to the topic.

Exclusion Criteria. Paper is not peer-reviewed, it is not written in English, based on abstract it is irrelevant for our topic.

Search Results. Our search string generated 340 results, which reduced to 31 after applying inclusion and exclusion criteria. Removing duplicates left us with 17 studies, and two more were found through snowballing. In total, we included 19 papers for data extraction.

2.2 New SLR

Search Queries. To define a search query for RQ1, we leaned on some of the queries from existing SLRs and adapted them for our needs. The following queries were used (examples from Web of Science, they were adapted for other DBs):

For RQ1: *web OR software (Title) and accessib* OR disab* OR universal OR impair* (Title) and development OR engineering OR process (Topic) and guideline* OR standard* OR method* OR technique* OR tool* (All Fields)*
For RQ2: *“IDE” OR “IDEs” OR “integrated development environment” OR “integrated development environments” OR “Visual Studio” OR “Eclipse” (Topic) and accessib* OR impair* OR disab* OR universal (Topic)*

Inclusion Criteria. Title concludes that the paper helps with answering either RQ1 or RQ2. If it addresses RQ1, the paper must be published in 2022 or 2023.

Exclusion Criteria. Paper is not peer-reviewed, it is not written in English, based on abstract it is irrelevant for our RQs, reading the full text reveals the paper does not answer any of our RQs.

Search Results. Our search string defined for RQ1 generated 153 results, which reduced to 92 after applying inclusion and exclusion criteria (without reading the full texts). Removing duplicates left us with 63 studies, and after reading the full texts, the final number came to 44 papers.

Our search string defined for RQ2 generated 675 results, which reduced to 12 after applying inclusion and exclusion criteria. Removing duplicates and checking access left us with 5 papers.

3 Results

After data extraction, we grouped identified SLRs into the following thematic sections:

- papers covering the whole SE process (one on model-driven development),
- papers on evaluation methods and automated tools (one focusing on semantic web technologies and one on e-government),
- 2 papers on specific type of impairment (dyslexia and cognitive disabilities),
- 4 papers on university/educational web pages,
- and 3 papers addressing contexts beyond traditional software products (serious web games, multi-device inclusive environments and digital accessibility).

Many SLR studies we analyzed focused on guidelines, standards, methods, techniques and tools, which contribute to answering our **RQ1**. Three of the analyzed SLRs, authored by Cruz-Portilla et al. [8], Barroso Paiva et al. [9] and Ara et al. [10], stand out as particularly relevant for our research, providing an in-depth examination of accessibility within the engineering process and development life cycle, closely aligning with some of our research goals and offering valuable insights. Assets identified in all three studies were: frameworks, methods, (verification) tools, (accessibility) requirements, processes, and techniques. Additionally, the following assets were identified in two or one of the above mentioned SLRs: prototypes, approaches, strategies, tests, verifications, evaluations, architecture designs, models, guidelines, patterns, architectures, ontologies, plug-ins, reports, APIs, vocabularies, repositories, libraries, analyses, systems, assessments, suggestions, improvement directions, explanations, cost, opportunity, experience and knowledge, metrics, module, algorithm and simulator.

Another SLR covering the whole SW engineering process [1], explored model-driven development (MDD) in accessible software engineering. While MDD offers code generation from models, existing methods ignore accessibility standards. The authors found several solutions, including extensions for including accessibility features.

Most SLRs were focusing on accessibility evaluations [11–15] and the results show that the usage of automated tools is the most frequent technique to evaluate web accessibility. We can confirm the same for the years 2022 and 2023, yet it's worth noting the growing interest in user involvement. Many authors in the early stages of their engineering process have expressed intentions to involve users.

Multiple SLRs [3, 12–16] also provide insights into which automated web evaluation tools were used by researchers. AChecker, WAVE, and TAW are the most frequently used. After analyzing the papers from 2022 and 2023, we came to the same results. While the analyzed SLRs had different results when it came to the frequency of usage, we identified TAW as the most commonly used, followed by WAVE.

Dias et al. [17] provide a list of tools not only for web, but also mobile accessibility evaluation, together with a list of features that they extracted from their analyzed papers. Campoverde-Molina et al. [12] provide a list of manual tools or methods that were used in their analyzed papers: JAWS, 2-switch, HCI and questionnaires. Enco-Jáuregui et al. [18] focused on web accessibility for people with dyslexia, but they also identified evaluation methods in their analyzed studies. The methods were: questionnaires, usability tests, interviews, measurements of the eye movements, participant´s reading speed and verification list, showing strong user focus.

Bittencourt et al. [19] analyzed evaluation techniques for multi-device inclusive environments. If we sort them by usage, the most common ones are: case study, experiment and heuristic evaluations. Here we see the main distinguishment from web evaluations, where automated tests are strongly in the lead.

If we summarize findings from our own analysis, for 2022 and 2023, we identified 7 **guidelines and standards**, together with an approach for simplified WCAG representation. The most studies were addressing WCAG 2.1., just one used the newest WCAG 2.2. Despite being outdated, also WCAG 2.0 were frequently used. Most studies did not state WCAG version they used.

We identified numerous **methods**, many of which focus on usability and user involvement. It is worth mentioning that the terminology for these methods does not seem to be standardized yet, as many authors use different terms for the same methods.

We identified 20 **techniques** that contribute to enhancing or evaluating various aspects of software development, such as document enhancement techniques, deep learning, simulators, etc.

The most common **tools** identified were those for the automatic evaluation of websites, which are widely known. In software engineering practices, it is commonly advised to utilize multiple automatic tools, and we also came across a paper proposing a prototype that introduces an automatic ensemble of such tools. Additionally, we encountered tools supporting other processes in software engineering, some of which are already available, while others are still in the conceptual stage. If we mention some: ESALP 2.0 (Educational System for Accessibility Learning Through Paradigms), a tool for representative page sampling and PDF to HTML converter.

Details and complete lists of identified guidelines, standards, methods, techniques, and tools can be accessed via the link provided in the introduction.

If we also take a look at which phases of the engineering process are the most commonly researched, in line with two previous SLRs [9, 10], our findings for 2022 and 2023 also indicate that testing remained the most researched phase.

Regarding the answer to our **RQ2** we observed that research on involving accessibility into integrated development environments is very limited. There are some papers describing how to make integrated development environments accessible for developers with different types of impairment, which were not relevant for our research, but very few describing how to incorporate accessibility checking support into them, to enable development of accessible products. Most are quite outdated and focusing on Eclipse.

The papers we included in our SLR are ranging from the oldest being written in 2006 to the youngest being written in 2015.

Feigenbaum and Squillace [20] introduced an IBM Rule-based Accessibility Validation Environment called RAVEN. RAVEN is an Eclipse-based tool for inspecting and validating Java rich-client GUIs for accessibility. The environment is distributed as a set of Eclipse plugins, and can be utilized in the Eclipse workbench. The authors advocate that accessibility should be an integral part of the development cycle as well as the test cycle and developed RAVEN with this main principle in mind.

De Branco et al. [21] presented another Eclipse plugin, called AccTrace. In contrast to RAVEN, AccTrace is a CASE tool focusing on requirements that were based on the Accessibility Tasks Model (MTA). The tool utilizes an ontology to define the technical implementation of accessibility and advocated the traceability of accessibility requirements from conception through to the coding phases. AccTrace was also identified by Barroso Paiva et. al. [9].

Another IDE plugin, not for Eclipse, but for NetBeans, is called DIAS. Giakoumis et al. [22] presented DIAS, a tool for impairment simulation. UI developers can experience accessibility limitations through visual, hearing, physical and cognitive impairment simulation performed over Java, mobile and web applications. In addition, DIAS integrates two of the most common assistive technologies - screen reader and a magnifier. The tool was developed as a standalone version and as a plugin for NetBeans IDE.

Atterer [23] presented Wusab, a prototype of a model-based automatic usability validator that aims to improve web based UIs. The prototype was implemented as a Java-based web application, however the author stated that the integration of usability validation into web engineering IDEs is one of the aspects that could be inspected in the future.

Another paper involving models, was written by Antonelli et al. [24]. The authors aimed to assist developers in creating accessible web menus through a model-driven approach, without developers having to know technical details about accessibility. A meta-model AMenu (Accessible Menu) that originated the DSL (Domain Specific Language) was created. The authors believe that such automated mechanisms can be included in CASE tools and IDEs, to propagate developments of accessible systems.

4 Discussion

Firstly, we would like to give some thought on the quality of previous SLRs in regards of repeatability. Based on Kitchenham et al. [6, 7] the SLR methodology aims to be auditable and repeatable. The authors themselves not only provided guidelines for performing SLRs [6], but also performed two SLRs of SLRs in the field of SE. More than a decade has passed after those studies, but we similarly observed that many SLRs we included in our study were lacking transparency and unified methodology approaches. Due to the absence of detailed process descriptions, including data synthesis results, many are not reproducible. In the case of conference proceedings or other publications limited in the number of pages, we suggest that authors provide additional details via an additional link, such as Cruz-Portilla et al. [8].

After conducting a Systematic Literature Review (SLR) on previous SLRs in the field of software/web accessibility and conducting our own SLR, it became evident that there are many available guidelines, standards, methods, techniques, and tools that contribute to accessible software. The pool of available assets continues to expand, as numerous researchers propose modifications to existing ones or develop entirely new assets. However, it is crucial to note that most assets are tailored to specific phases within the software engineering (SE) process and do not encompass the entire process. Thus, there remains a scarcity of concepts that comprehensively address the entirety of the SE process.

Our research also focused on integration of accessibility in integrated development environments. We identified that insufficient research has been conducted in this area, or more precisely, to our knowledge very limited research has been undertaken on the integration of accessibility evaluation support in integrated development environments, mostly focusing on some custom-made plugins. We identified three IDE plugins, each serving a different purpose. RAVEN is an Eclipse-plugin for inspection and validation, AccTrace is an Eclipse-plugin for requirements based on MTA, and DIAS is a NetBeans-plugin for impairment simulation. In addition to these plugins already developed, we've identified two concepts or prototypes that the authors deem suitable for integration into IDEs: Wusab, a model-based automatic usability validator and a meta-model AMenu.

Our SLR underlines the strong need for better integrating accessibility into the core of the SE process and into IDEs and provides a solid base for next steps in R&D in this domain.

References

1. Ordoñez, K., Hilera, J., Cueva, S.: Model-driven development of accessible software: a systematic literature review. Univ. Access Inf. Soc. **21**(1), 295–324 (2022)
2. Gaggi, O., Perinello, L.: Improving accessibility of web accessibility rules. In: Proceedings of the 2022 ACM Conference on Information Technology for Social Good, Limassol, Cyprus, pp. 167–174. ACM (2022)
3. Moreno, L., Martínez, P.: Accessibility compliance for e-government websites: laws, standards, and evaluation technology. In: Research Anthology on Digital Transformation, Organizational Change, and the Impact of Remote Work, pp. 354–373. IGI Global (2021)
4. European Commission: Web Accessibility Directive - Standards and Harmonisation. https://digital-strategy.ec.europa.eu/en/policies/web-accessibility-directive-standards-and-harmonisation. Accessed 08 Apr 2024
5. European Commission: Web Accessibility Directive - Monitoring Reports. https://digital-strategy.ec.europa.eu/en/library/web-accessibility-directive-monitoring-reports. Accessed 08 Apr 2024
6. Kitchenham, B., Charters, S.: Guidelines for performing systematic literature reviews in software engineering. Technical Report EBSE-2007–01, School of Computer Science and Mathematics, Keele University (2007)
7. Kitchenham, B., et al.: Systematic literature reviews in software engineering – a tertiary study. Inf. Softw. Technol. **52**(8), 792–805 (2010)
8. Cruz-Portilla, M., Pérez-Arriaga, J.C., Ocharán-Hernández, J.O., Sánchez-García, Á.J.: Accessibility in the software development life cycle: a systematic literature review. In: 2021 9th International Conference in Software Engineering Research and Innovation (CONISOFT), pp. 97–103. IEEE (2021)
9. Barroso Paiva, D.M., Freire, A.P., de Mattos Fortes, R.P.: Accessibility and software engineering processes: a systematic literature review. J. Syst. Softw. **171**, 110819 (2021)

10. Ara, J., Sik-Lanyi, C., Kelemen, A.: Accessibility engineering in web evaluation process: a systematic literature review. Univ. Access Inf. Soc. (2023)
11. Nuñez, A., Moquillaza, A., Paz, F.: Web accessibility evaluation methods: a systematic review. In: Marcus, A., Wang, W. (eds.) HCII 2019. LNCS, vol. 11586, pp. 226–237. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-23535-2_17
12. Campoverde-Molina, M., Lujan-Mora, S., Garcia, L.V.: Empirical studies on web accessibility of educational websites: a systematic literature review. IEEE Access **8**, 91676–91700 (2020)
13. Campoverde-Molina, M., Luján-Mora, S., Valverde, L.: Accessibility of university websites worldwide: a systematic literature review. Univ. Access Inf. Soc. **22**(1), 133–168 (2023)
14. Akram, M., Bt Sulaiman, R.: A systematic literature review to determine the web accessibility issues in Saudi Arabian university and government websites for disabled people. Int. J. Adv. Comput. Sci. Appl. (IJACSA) **8**(6) (2017)
15. Cisneros, D., Monzón, F.H., Paz, F.: Accessibility evaluation of E-government web applications: a systematic review. In: Soares, M.M., Rosenzweig, E., Marcus, A. (eds.) HCII 2021. LNCS, vol. 12780, pp. 210–223. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-78224-5_15
16. Macakoğlu, Ş.S., Peker, S.: Web accessibility performance analysis using web content accessibility guidelines and automated tools: a systematic literature review. In: 2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), pp. 1–8 (2022)
17. Dias, J., Carvalho, D., Paredes, H., Martins, P., Rocha, T., Barroso, J.: Automated evaluation tools for web and mobile accessibility: a systematic literature review. In: Abraham, A., et al. (eds.) Innovations in Bio-inspired Computing and Applications, vol. 419, pp. 447–456. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-96299-9_43
18. Enco-Jáuregui, L., Meneses-Claudio, B., Auccacusi-Kañahuire, M.: Web accessibility for people with dyslexia: a systematic literature review. EAI Endorsed Trans. Perv. Health Tech. **9** (2023)
19. Bittencourt, I.I., Baranauskas, M.C., Pereira, R., Dermeval, D., Isotani, S., Jaques, P.: A systematic review on multi-device inclusive environments. Univ. Access Inf. Soc. **15**(4), 737–772 (2016)
20. Feigenbaum, B., Squillace, M.: Accessibility validation with RAVEN. In: Proceedings of the 2006 International Workshop on Software Quality, Shanghai, China, pp. 27–32. ACM (2006)
21. De Branco, R.G., Cagnin, M.I., Barroso Paiva, D.M.: AccTrace: accessibility in phases of requirements engineering, design, and coding software. In: 2014 14th International Conference on Computational Science and Its Applications (2014)
22. Giakoumis, D., Kaklanis, N., Votis, K., Tzovaras, D.: Enabling user interface developers to experience accessibility limitations through visual, hearing, physical and cognitive impairment simulation. Univ. Access Inf. Soc. **13**(2), 227–248 (2014)
23. Atterer, R.: Model-Based Automatic Usability Validation – A Tool Concept for Improving Web-Based UIs (2008)
24. Antonelli, H.L., da Silva, E.A.N., de Mattos Fortes, R.P.: A model-driven development for creating accessible web menus. Procedia Comput. Sci. **67**, 95–104 (2015)

Inequality in User Experience: Can Mobile User Interfaces that Help Sighted Users Create Barriers for Visually Challenged People?

Sachin Tanwar(✉) and P. V. M. Rao

Department of Design, Indian Institute of Technology Delhi, New Delhi 110016, India
sachin251287@gmail.com

Abstract. Designing mobile interfaces for enhanced usability and user experience has become a standard practice in modern-day app development. However, this approach often prioritizes the needs of sighted users, leading to a compromised experience for people who are visually challenged or blind. This study reveals the user experience elements that make it easier for sighted users to accomplish a task while creating barriers for people who rely on screen readers. Using task-based usability tests of six globally popular mobile apps, the study compares the experiences of 12 sighted and 15 visually challenged (VC) users. The results reveal drastic differences in usability and experience between the two groups, highlighting the gaps and experience compromises despite being accessible by norms. The study highlights how designing interfaces for enhanced usability and user experience for sighted users compromises six prominent aspects of usability for people who are visually challenged or blind, leading to productivity challenges and poor user experiences, calling for a more inclusive and accessible approach to mobile app design. The study suggests investigating technological advancements, such as building screen reader capabilities to understand the designer's intent or creating tools to support the designer in keeping screen reader limitations in mind while designing, to address such issues and provide an equal experience for VC users.

Keywords: Accessible user experience · Assistive Technology (AT) · (e)Accessibility · Design for All and Universal Design

1 Introduction

The growing popularity and prevalence of smartphones necessitates the accessibility of mobile applications (apps) to all users, including those with visual challenges (VC). The user interface and its experience (UX) of an app play a crucial role in determining its usability and accessibility for both sighted and VC users. Specialized websites or apps for specific user groups provide arguably a tailored experience, but are difficult to maintain for different individuals [1].

Currently, there are over 5 million applications for Android and iOS devices, with the majority of apps (more than 77%) being inaccessible or poorly accessible, as indicated

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 19–30, 2024.
https://doi.org/10.1007/978-3-031-62846-7_3

by multiple researchers [2, 3]. Screen readers (talkback on Android or voiceover on iPhone) help VC users use mobile apps or web contents. In general, mobile apps fall into two categories. First, utility-based apps like e-commerce or ticket-booking apps use deliberate interactions to help users achieve their goals more efficiently and conveniently. While the others are engaging or interactive apps that have content and page layouts designed to enhance user experiences and encourage stickyness of use, for example, social media or media apps.

2 Literature Review

While there have been efforts to improve the usability and accessibility of smartphone applications, multiple studies suggest that visually challenged users still face a lack of understanding about interface usability issues that specifically cater to their needs [4–9]. Recent studies have established that even if the apps are WCAG level 2 compliant, they do not ensure a satisfactory user experience for VC users [10–13].

The literature highlights that while accessibility issues are addressed to some extent, usability issues for VC users are not systematically studied or considered [14, 15]. This lack of systematic understanding can result in VC users facing difficulties in accessing and utilizing smartphone applications for day-to-day purposes, potentially excluding them from utilizing the full potential of the apps that they are using or hindering new app exploration due to their current experiences in contrast to their sighted counterparts [7, 16].

It becomes critical to deeply research the factors that simultaneously simplify the user interface element for sighted users but may complicate it for VC users who rely on screen readers. In addition, it is critical to anticipate and test for potential conflicts between screen reader components and application elements and to provide alternatives for VC users in lieu of inaccessible components [17, 18].

A thorough study could help designers understand accessibility issues, adopt universal design principles, and create usable and accessible interfaces for all users. This study advances this field by examining the usability and experience aspects of accessibility.

Studies have been conducted to examine distinct attributes of applications from the perspectives of accessibility, usability, and user experiences. There is no study that looks at major pain points for visually challenged users that may surface due to designs catered primarily to sighted users. Some specific work on a single aspect of usability can be found in the literature. For instance, Alnfiai [19] strategies to enhance the universality of Captchas as an element of the user interface and their input techniques, catering to both sighted and VC users. Conducting a comprehensive study will facilitate designers in gaining a deeper understanding of accessibility obstacles, embracing universal design principles, and developing interfaces that are both usable and accessible to individuals of all abilities [20–22]. The current study is a significant advancement in this direction, looking at multiple aspects of accessibility in terms of usability and user experience.

3 Research Methodology

3.1 Recruitment and Participants

27 volunteers were recruited for this study, which was conducted using six mobile apps. Initially, 12 volunteers with normal vision were studied. The sample included 9 men and 3 women aged 21–34. Four participants used iOS, and eight used Android devices. In the second phase, 15 VC users with 3+ years of mobile screen reader experience—11 males and 4 females of the same age—were recruited.

Thirteen participants used Android devices, while two used iOS devices. Participants were university students or professionals in both cases with good fluency in English reading, writing, and speaking. Participants were offered to choose any location and device for their comfort. An informed consent form for each participant was outlined for the qualitative research using the task-based think-aloud user testing method.

3.2 Mobile Application (App) Selection and Task Based Usability Testing

To effectively address research issues outlined in the previous section, the research was segmented into various elements as shown in Fig. 1 below. Combining observations with user tasks and conversations was deemed crucial for the current study.

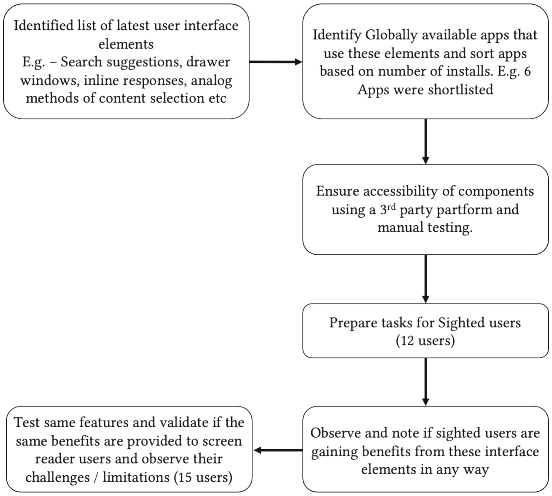


Fig. 1. The block diagram illustrates the process of shortlisting apps for testing. Starting with the selection of common user interface elements, the presence of such interactions in globally popular apps, testing shortlisted apps for WCAG compliance manually using Android and iOS screen readers and a 3rd party tool (accessibility scanner app for Android), and lastly, testing key scenarios with sighted users.

Initially, a pilot test with 7 sighted users identified a list of new user interactions, or UX patterns, before the actual study with 27 users took place. The goal was to identify recently popular UX within mobile apps that makes their work fast, helps them receive quick updates from their friends, boosts their confidence, or helps them maintain privacy and secrecy. Users identified 18 available interaction patterns, such as social media stories from friends, swipe to reply, delivery tracking via map, etc., and most users claimed to be comfortable with these interactions in the initial discussion.

Based on the interactions, relevant apps were identified that were accessible, popular, and had more than one million downloads. The chosen apps also use modern design (UX) patterns and interactions found in other global apps, for example, following material design guidelines, Android best practices, and iOS design standards. An established third-party tool (Accessibility Scanner for Android) and manual testing using Android and iOS screen readers ensured accessibility compliance with WCAG 2.1 conformance criteria using the available WCAG official checklist, which were found to follow AA standards. It was also found that all recruited users were familiar with the selected apps and were using them on a regular basis. The following six applications were chosen (the names of the applications have not been disclosed).

1. An electronic commerce platform that facilitates user exploration of recommended products and facilitates product discovery across many categories.
2. A communication platform that facilitates users to engage in chat conversations with individuals listed in their contact directory. Additionally, it offers the capability to share various types of files and establish group conversations as its primary features.
3. A web-based email service that enables users to establish and manage numerous email accounts. Enables individuals to schedule appointments and receive notifications.
4. A mobile application that enables users to locate and reserve local taxis for immediate or future transportation needs.
5. An online platform that facilitates the creation and consumption of video content by its users.
6. A mobile application that facilitates user interaction, enabling individuals to engage in conversations, connect with friends, and consume various forms of multimedia information, including text, photos, documents, audios, and videos.

Tasks were assigned to people with normal vision to find out if the UX elements helped speed up the process, protect their privacy, or provide more engaging experiences. Twelve normal-sighted participants took the task-based usability test. Post-test observations and conversations were documented. In the second phase, 15 VC volunteers were studied using the same scenarios that had helped sighted users. Accessibility, usability, and user experience issues were noted. This study found dimensions beyond user experience and usability, which are discussed as results.

4 Results

This study examines user interactions, not mobile apps. This study focuses on how screen reader users' unfamiliarity with popular UX elements in mobile apps affects VC users' adoption and usage, not the platforms or apps themselves. Each user performed six subtasks to effectively use the interface, which were evaluated to show different levels of experience (confusion, frustration, restriction of use, and what prevents users from achieving their goals). The study involves conducting qualitative interviews and task-based assessments to evaluate the advantages provided by the user experience (UX) for sighted users while also examining the difficulties faced by VC users in utilizing the straightforward service.

4.1 Confidence vs. Frustration

Scenario: Search an item in the e-commerce app

Task: Conducting a search for a product with a complex nomenclature and seeking a product that has previously been accessed in their respective search history (after observing search history, the product name was announced that was to be searched) (Fig. 2).

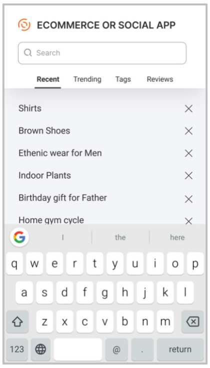


Fig. 2. Indicative screenshot of an e-commerce or social media app that provides trending search suggestions and history as the user starts to type in the search field

Experience of sighted users: Findings suggest that 100% of sighted people use auto-complete suggestions to spell products correctly. Approximately 85% of users used their search history to seamlessly browse this product while typing. Users were confident while using the UX and found it a time saving experience.

Challenges for VC: Virtually no VC user initially considered the search recommendations. They either initiated the inquiry using Mic or by typing the entire text.

Both scenarios required 90% of typing users to enter the entire text, increasing the risk of typing the wrong character.

100% of the task 2 users who selected an item from the search suggestion accidentally triggered a character on the on-screen keyboard panel, resulting in unexpected results and leading to high frustration amongst VC users.

4.2 Multitask vs. Confusion

Scenario: Exploration of User Experience within a video sharing Platform using minimized window interactions.

Tasks for all users: To play a music video and, while it plays, browse other videos from the same channel (Fig. 3).

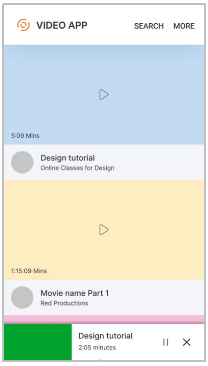


Fig. 3. Indicative screenshot of an app that has a player window minimized to allow users to multi-task

Experience of sighted users: Out of 12 individuals, 11 had a positive experience in terms of comfort and productivity when transitioning between player and album views. The UX was determined to be a user-friendly and convenient solution for these individuals. Simultaneously, users demonstrated the ability to discern between advertisements and the primary video content and were proficient in bypassing the advertisements when they materialized.

Challenges for VC users: All 15 users encountered difficulties optimizing the size of the video player window when minimized. Seven of them ended up shutting down the application and then restarting it because they couldn't maximize the player after it had minimized. Additionally, the lack of clarity regarding the status and position of a minimized window resulted in significant confusion for users. The majority of consumers have difficulty distinguishing between advertisements and the primary video content, particularly when the video is in a reduced view.

4.3 Control vs. No Control

Scenario: Checking out dynamic "Stories" content from friends and brands the user follows.

Tasks for all users: To check out stories of their friends and stop at the most interesting one.

Experience of sighted users: All 12 participants in the study reported a high level of comfort and efficiency in locating an engaging narrative promptly through the utilization of flick gestures. Users had the convenience of pausing at any given moment and swiftly advancing through the dynamic information.

Challenges for VC users: None of the 15 users showed proficiency in effectively organizing "story" content in relation to temporal considerations. The utilization of a screen reader posed significant challenges when attempting to interact with narratives that included auditory elements. The fact that every user was dependent on time to progress automatically to the next story slowed down the process for them. Moreover, upon encountering a captivating story", half of the users failed to pause it at the right moment, resulting in a longer re-visit of the process, which made them feel the UX was challenging to control.

4.4 Quick vs. Slow

Scenario: Live conversation using the messenger app.

Tasks for all users: To locate a collective dialogue wherever one individual has responded to a specific message from another person (in a group chat) (Fig. 4).

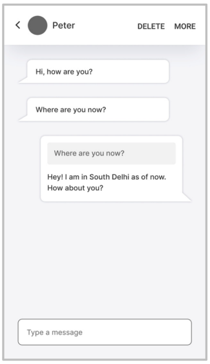


Fig. 4. Indicative screenshot of an app that allows in-line reply for individual chats

Experience of sighted users: All twelve participants demonstrated proficiency in utilizing the respond feature, indicating their prior experience and ease of usage, and found it quick and more immersive. All users were able to comfortably point to the last answered message in a chat session. The overall reaction regarding the experience of this feature was quite positive, since it effectively addresses the issue of accurately referencing specific messages, minimizing any potential confusion.

Challenges for VC users: Initially, users encountered challenges in locating the precise message to which an individual had responded. However, none of the VC users were able to ascertain the message to which the response was directed, as well as its corresponding content. The confusion among users arose from the difficulty in locating a previously responded message, as the screen reader interprets it as two independent messages rather than a cohesive or nested message, resulting in the loss of time and its intended meaning.

4.5 Surety vs. Barrier

Scenario: Planning and booking a taxi or cab from the given app

Tasks for all users: to know if any taxis are present nearby; to reserve a taxi from the current location; to know the direction and estimated time (Fig. 5).

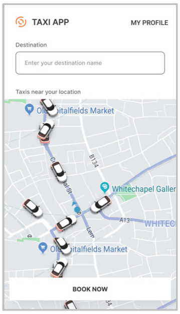


Fig. 5. Indicative screenshot of an app that displays all the taxi's standing nearby current location.

Experience of sighted users: All twelve participants successfully determined the availability of local taxis, providing them with a sense of certainty regarding their ability to secure a cab and the optimal time to make a booking. They effectively communicated the instructions regarding the taxi's direction and projected duration.

Challenges for VC users: All users were unable to obtain information regarding the availability of nearby taxis, resulting in the necessity for each user to navigate the taxi booking form and initiate a ride request in order to ascertain the presence of available taxis. Following the completion of the booking process, all 15 consumers opted to depend on the driver's phone call rather than determining a designated pick-up location based on the anticipated arrival direction of the taxi. This restricts users from obtaining crucial information in intricate scenarios or several designated areas within a closed community.

4.6 Shortcut vs. No Option

Scenario: To send or delete emails from mailing client app

Tasks for all users: To delete two spam emails from email app and restore them (Fig. 6).

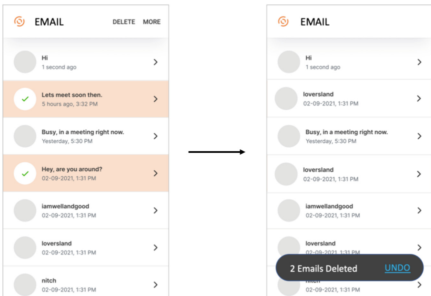


Fig. 6. Indicative screenshot of an app that allows users to undo action that was recently taken

Experience of sighted users: 10 out of 12 users successfully identified the action buttons on the toast notification subsequent to the deletion of emails. Additionally, nine users were able to utilize the undo deletion option, which was available for a brief duration of a few seconds. Further, the process of selecting multiple emails and performing actions on them was more convenient for users.

Challenges for VC users: None of the 15 users were able to locate the toast notification, rendering them unable to "undo" the delete action that the app provided. Nevertheless, all individuals were able to successfully retrieve emails from the Deleted Items section. Furthermore, it was discovered that users experienced limitations in using the application that provides shortcuts on toasts for brief periods in several other instances, such as the cancellation of a sent email or any other condition.

5 Discussion

In this section, further implications that came out of the study are highlighted.

5.1 Informing Users About Time Sensitive Content and Provide Control

Interactions such as stories or toast alerts are temporally constrained. The primary objective of these platforms is to deliver time-sensitive information to their users. However, the utilization of screen readers and user interfaces necessitates that users employ gestures in order to obtain control. There is a need for a more comprehensive approach to address this predicament. One such approach could involve momentarily enabling auto-focus on time-sensitive items as a means to assist users in making timely decisions. In addition to relying solely on artificial intelligence, it is imperative for screen readers to provide support to users in the process of selecting actions. After presenting the dynamic toast information, the screen reader may prompt the user to do a double tap in order to initiate an action or, alternatively, a triple tap to dismiss the toast alert.

5.2 Providing Alternative to Dynamic Visual Pages

Subtitles are a widely recognized method for offering an alternative means of accessing visual information. Future designers and developers may consider exploring a potential method for translating visually immersive, dynamic web pages that include a subsection with informative content, allowing users to keep updated and make autonomous decisions. In situations where users want information regarding the proximity of available taxis, the inclusion of textual information with the visual representation of taxis on a map can enhance user choice and control.

5.3 Visual Clarity and More Control on Screen Readers for On-Screen Content

When relying solely on screen reader technology or visuals, managing on-screen material, such as responding to inline messages or controlling minimized media players that create overlays on other pages, presents challenges. Designers should investigate new methods for visually representing information that offer improved clarity in the separation or connection of related objects. Simultaneously, screen readers could improve accessibility by handling these elements differently using aural cues. At present, the screen reader is capable of audibly rendering all the textual elements included on a webpage. In instances where there are page or content overlays, it would be beneficial for screen readers to incorporate a little variation in audio output. This would enable users to easily discern the overlay contents from the underlying content.

6 Conclusion and Future Work

The primary emphasis of the study was on design elements, namely dynamic panels, micro-interactions, and navigations. These components were initially evaluated with sighted users, who have shown a preference for them when completing assigned tasks. Furthermore, it was observed that these design elements facilitated decision-making and navigation within the application.

As experienced by VC users, this study examined important aspects of mobile interface design from the perspectives of accessibility, usability, and user experience. It was found that most of the VC users did not have good experience with interfaces or interactions that benefit sighted users, even if both user groups were found to use them on a day-to-day basis.

This study focused on primary interactions occurring on landing pages and secondary depth pages. This study emphasizes the need for further investigation of the user interface (UI) components, particularly web interfaces, and their comprehension by screen readers. Exploring if technological advancements, such as enhancing screen reader capabilities, improving webpage design, or a combination of both, can address these issues. The majority of accessibility guidelines advise against using dynamic content or features because they fail to provide a satisfactory user experience for people with visual impairments. Rather than attempting to evade them, it may be worthwhile to conduct further research on how VC individuals can potentially consume digital information in a manner comparable to that of sighted individuals by utilizing screen readers and employing considerate design strategies.

Despite the fact that all tested interfaces adhered to WCAG level 2.1, it was discovered that VC users encounter numerous obstacles when utilizing interface patterns designed to assist sighted individuals or enhance their experiences. Moreover, users tended to terminate the application when they encountered difficulties or obstacles in navigating the menu or seeing unfamiliar information. We mostly attributed this behaviour to their inability to comprehend the interactive elements of the app. We observed that VC users had to memorize important buttons, functions, interactions, and gestures associated with each application to achieve speed. The process of committing to memory every interface element inside each application can significantly deplete users' working memory capacity, resulting in both frustrating user experiences and extended training durations. The combination of page design with screen readers was shown to have a passive effect on users since it did not effectively facilitate communication of the intended design of these interfaces.

Acknowledgements. Authors acknowledge support received from National Centre for Assistive Health Technology (NCAHT), IIT Delhi in carrying out this work.

References

1. Álvarez-Robles, T., Álvarez, F., Orozco, Y., Sandoval-Bringas, J.A.: Analysis of specialized websites in digital libraries: evaluation of UX with blind users. In: Stephanidis, C., et al. (eds.) HCII 2021. LNCS, vol. 13096, pp. 133–146. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-90328-2_9
2. Al-Razgan, M., et al.: A systematic literature review on the usability of mobile applications for visually impaired users. PeerJ Comput. Sci. **7** (2021). https://doi.org/10.7717/PEERJ-CS.771
3. Alshayban, A., Ahmed, I., Malek, S.: Accessibility issues in Android apps: state of affairs, sentiments, and ways forward. 1323–1334 (2020). https://doi.org/10.1145/3377811.3380392
4. dos S. Rosa, J.R., Valentim, N.M.C.: An exploratory study about accessibility, usability and user experience with the visually impaired using mobile applications. In: ACM International Conference Proceeding Series, Article no. 3439998 (2020). https://www.scopus.com/inward/record.uri?eid=2-s2.0-85102882807&doi=10.1145%2f3439961.3439998&partnerID=40&md5=983259e10c8a5c7450cd826ffca4edf5
5. Aizpurua, A., Arrue, M., Vigo, M.: Uncovering the role of expectations on perceived web accessibility. In: Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility - ASSETS 2013, pp. 1–2 (2013). https://doi.org/10.1145/2513383.2513411
6. Stephanidis, C.: Margherita Antona Universal Access in Human-Computer Interaction Methods, Techniques, and Best Practices (2016). http://www.springer.com/series/7409
7. Busaeed, S., Mehmood, R., Katib, I.: Requirements, Challenges, and Use of Digital Devices and Apps for Blind and Visually Impaired (2022). https://doi.org/10.20944/preprints202207.0068.v1
8. Rodrigues, A., Nicolau, H., Montague, K., Guerreiro, J., Guerreiro, T.: Open challenges of blind people using smartphones. Int. J. Hum. Comput. Interact. **36**(17), 1605–1622 (2020). https://doi.org/10.1080/10447318.2020.1768672
9. W3Org. Accessibility, Usability, and Inclusion (2018). https://www.w3.org/WAI/fundamentals/accessibility-usability-inclusion/. Accessed 5 May 2022
10. Alajarmeh, N.: The extent of mobile accessibility coverage in WCAG 2.1: sufficiency of success criteria and appropriateness of relevant conformance levels pertaining to accessibility problems encountered by users who are visually impaired. Univers. Access Inf. Soc. **21**(2), 507–532 (2022). https://doi.org/10.1007/s10209-020-00785-w
11. Power, C., Freire, A., Petrie, H., Swallow, D.: Guidelines are only half of the story. In: Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems - CHI 2012, p. 433 (2012). https://doi.org/10.1145/2207676.2207736
12. Leuthold, S., Bargas-Avila, J.A., Opwis, K.: Beyond web content accessibility guidelines: design of enhanced text user interfaces for blind internet users. Int. J. Hum.-Comput. Stud. **66**, 257–270 (2008). https://doi.org/10.1016/j.ijhcs.2007.10.006
13. Aizpurua, A., Arrue, M., Vigo, M.: Prejudices, memories, expectations and confidence influence experienced accessibility on the Web. Comput. Hum. Behav. **51**, 152–160 (2015). https://doi.org/10.1016/j.chb.2015.04.035
14. Aizpurua, A., Harper, S., Vigo, M.: Exploring the relationship between web accessibility and user experience. Int. J. Hum. Comput. Stud. **91**, 13–23 (2016). https://doi.org/10.1016/j.ijhcs.2016.03.008
15. Griffin-Shirley, N., et al.: A survey on the use of mobile applications for people who are visually impaired. J. Vis. Impairment Blindness **111**, 307–323 (2017). https://doi.org/10.1177/0145482X1711100402
16. Sauer, J., Sonderegger, A., Schmutz, S.: Usability, user experience and accessibility: towards an integrative model. Ergonomics **63**(10), 1207–1220 (2020). https://doi.org/10.1080/00140139.2020.1774080

17. Yan, S., Ramachandran, P.G.: The current status of accessibility in mobile apps. ACM Trans. Accessible Comput. **12**(1), 1–31 (2019). https://doi.org/10.1145/3300176
18. Rodrigues, A., Montague, K., Nicolau, H., Guerreiro, T.: Getting smartphones to talkback: understanding the smartphone adoption process of blind users. In: ASSETS 2015: The 17th International ACM SIGACCESS Conference of Computers and Accessibility, pp. 23–32 (2015). https://doi.org/10.1145/2700648.2809842
19. Alnfiai, M.: Evaluating the accessibility and usability of a universal CAPTCHA based on gestures for smartphones. Univ. Access Inf. Soc. **20**, 817–831 (2021). https://doi.org/10.1007/s10209-020-00730-x
20. Milano, P.: "Usable accessibility" to the Web for blind users "Usable Accessibility" to the Web for blind Users (2016). http://ui4all.ics.forth.gr/workshop2004/files/ui4all_proceedings/adjunct/accessibility/109.pdf. Accessed Jan 2014
21. Zaina, L.A.M., Fortes, R.P.M., Casadei, V., Nozaki, L.S., Paiva, D.M.B.: Preventing accessibility barriers: Guidelines for using user interface design patterns in mobile applications. J. Syst. Softw. **186**, 111213 (2022). https://doi.org/10.1016/j.jss.2021.111213
22. Lee, Y., Lee, J.: A checklist for assessing blind users' usability of educational smartphone applications. Univers. Access Inf. Soc. **18**(2), 343–360 (2019). https://doi.org/10.1007/s10209-017-0585-1

On the Use of a Simulation Framework for Studying Accessibility Challenges Faced by People with Disabilities in Indoor Environments

Francisco Javier Rafful Garfias[1] and Vinod Namboodiri[2](✉)

[1] Wichita State University, Wichita, KS, USA
fjraffulgarfias@shockers.wichita.edu
[2] Lehigh University, Bethlehem, PA, USA
vin423@lehigh.edu

Abstract. Navigating indoor spaces is known to be significantly challenges for individuals with mobility and sensory impairments due to the presence of physical barriers and inadequate accessible signage. Current laws and efforts have not led to meeting diverse needs of these populations. In this work we provide a brief introduction to MABLESim (Mapping for Accessible BuiLt Environments Simulator), a simulation framework for studying indoor space accessibility. MABLESim recreates digital models of indoor environments, allowing for the simulation of diverse mobility scenarios for individuals with varying abilities. MABLESim enables the analysis of critical factors important for efficient mobility in indoor spaces such as route complexity and disability characteristics. Through careful configuration of simulation parameters, MABLESim facilitates the assessment of accessibility challenges in both simple and complex indoor spaces. This framework offers a tool for designers and planners to visualize and address accessibility barriers in built environments.

Keywords: 3D modeling · accessibility · indoor navigation · simulation · algorithms

1 Introduction

Physical obstacles within constructed environments have perpetually presented hurdles for individuals with mobility impairments. Common hindrances encompass narrow entrances, confined passageways, and irregular flooring, further exacerbated by an insufficient provision of essential amenities such as elevators, ramps, and automatic/motorized doors. The complexity of navigation within expansive indoor settings, coupled with inadequate wayfinding signage, amplifies the challenge. Particularly, individuals with sensory disabilities, such as visual impairments, encounter difficulty in utilizing conventional physical wayfinding cues, often struggling to navigate unfamiliar indoor environments, particularly

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 31–37, 2024.
https://doi.org/10.1007/978-3-031-62846-7_4

those of considerable size. This challenge is frequently shared by other demographics, including older adults and individuals with cognitive or intellectual impairments.

To address issues pertaining to physical barriers, legislative measures like the Americans with Disabilities Act [1] and its subsequent amendments [4,5] have been implemented. While such regulations have facilitated considerable advancements over the years, their impact on individuals with non-physical impairments remains limited. Moreover, for those with physical disabilities, the efficacy of these measures is constrained as compliance with stipulated requirements does not invariably ensure spaces that are genuinely accessible in essence, as usability standards are not mandated. Previous methodologies relying on manual assessments and evaluations of environments may not comprehensively encapsulate the complexities of real-world movements and interactions experienced by individuals with disabilities. Furthermore, these manual approaches are often labor-intensive and time-consuming, impeding widespread adoption. Recent endeavors have explored diverse strategies to enhance indoor accessibility for individuals with disabilities, encompassing the integration of visual cues, auditory feedback, tactile mapping, and smartphone-based wayfinding systems. However, these initiatives are hampered by their limited capacity to encompass various disability demographics on a scale substantial enough to yield actionable insights.

In this paper, we introduce the MABLESim (Mapping for Accessible BuiLt Environments Simulator) simulation framework as a scalable solution for assessing indoor space accessibility. MABLESim leverages architectural floor plans of indoor spaces to replicate key features and obstacles pivotal for studying human movements within these environments. By appropriately configuring simulation parameters, a multitude of real-world mobility scenarios can be simulated for individuals with diverse abilities, furnishing invaluable data for comprehending and enhancing the accessibility of examined spaces. As a proof of concept, we present simulation outcomes from two distinct buildings-one simple and compact, the other intricate and expansive-illustrating mobility patterns of individuals with varying abilities (including those with visual impairments, mobility impairments, and no disabilities), and elucidating insights into the accessibility and usability of each building.

2 State of the Art

In the advancement of independent indoor navigation for individuals with visual impairments, significant initiatives such as NavCog have made strides in enhancing accessibility [2]. Concurrently, Karami et al. (2019) conducted a physiological study investigating the effects of blindness on walking and jogging parameters, uncovering notable differences between blind and sighted individuals [9]. Within the domain of agent-based modeling (ABM), Fachada et al. (2015) emphasize the necessity of transparent model descriptions, exemplified by their PPHPC model, which serves as a benchmark for methodological rigor [7]. Additionally, the Flexible Space Subdivision (FSS) framework, introduced by [6], presents a comprehensive and inclusive approach to interior spaces catering to various disabilities. However, all prior human subject studies are constrained by scale and a limited number of participants, hindering a thorough investigation into the impact of various parameters.

A simulation framework tailored for indoor navigation in the context of disability offers manifold advantages that significantly contribute to enhancing accessibility in built environments. One key advantage is the detailed analysis of critical factors such as walking speed, route completion time, source-destination pairs, paths taken, likelihood of getting lost, and route difficulty, all tailored for individuals with disabilities. This comprehensive examination yields a nuanced understanding of the unique challenges faced by people with diverse disabilities, enabling the development of targeted solutions. The simulation framework serves as a valuable tool for designers and planners, providing a virtual environment to simulate and visualize potential barriers and complications that may arise in real-world scenarios.

Furthermore, the simulation framework facilitates iterative testing and refinement of accessibility solutions in a controlled digital environment prior to implementing physical changes to buildings or spaces. This iterative approach ensures that proposed modifications are effective, efficient, and genuinely enhance accessibility, thereby minimizing the need for costly and time-consuming adjustments post-implementation. The ability to customize simulations for various disability profiles adds another dimension of sophistication, enabling a granular understanding of diverse needs and preferences. Such customization ensures that resulting indoor spaces are not only compliant with accessibility standards but also genuinely user-friendly and inclusive, addressing the specific requirements of individuals with different disabilities (Fig. 1).

3 Methodology

MABLESim leverages the Unity engine [11] to convert 2D architectural blueprints, specifically floor plans, into immersive 3D digital models, facilitating detailed simulations that closely resemble real-world environments.

A key strength of MABLESim lies in its meticulous replication of building features and structures from architectural blueprints, ensuring accurate portrayal of dimensions and placements within digital layouts. This attention to detail results in faithful representations of indoor environments, facilitating high-fidelity simulations.

Moreover, MABLESim incorporates contextual information essential for wayfinding, seamlessly integrating details such as room numbers and restroom locations into digital environments. This integration enables the study of navigation and spatial cognition within simulated environments mirroring real-world conditions.

Central to MABLESim's functionality is the integration of navigation capabilities, facilitated by a navigation mesh implemented using C++ scripts within the Unity environment. This mesh allows for intelligent pathfinding, enabling

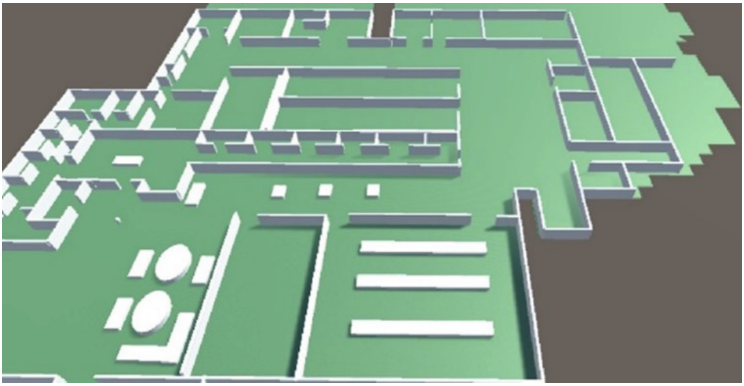


Fig. 1. Screenshot of a simulation scenario created in MABLESim.

simulated entities to navigate complex indoor spaces with ease, mimicking real-world mobility patterns.

Two key configurable parameters within MABLESim are mobility speeds and the wayfinding decision success rate (WDSR), crucial for tailoring simulations to specific scenarios and population categories. Mobility speeds can be adjusted to simulate varying movement capabilities for a population k, as shown in Table 1. The WDSR parameter influences navigation accuracy, with values informed by video traces of Persons With Disabilities (PWDs). For this work, WDSR values were set such that blind/low vision individuals chose the right branch on a route uniformly randomly, whereas all other individual with the benefit of sight chose the correct path all the time. For example, for a branching point on a path with 4 potential options, those with sight impairments chose the correct branch with a probability of 0.25 whereas all others chose the correct path with a probability of 1. Success in navigating an entire route depends on successful choices made at all branching points along the route. For this work, any incorrect choice (applicable to blind/low vision individuals only) incurred a flat penalty of 12 s, while for sighted users, the penalty was reduced to 4 s. After the penalty time elapsed, the user could choose any of the remaining branches to proceed.

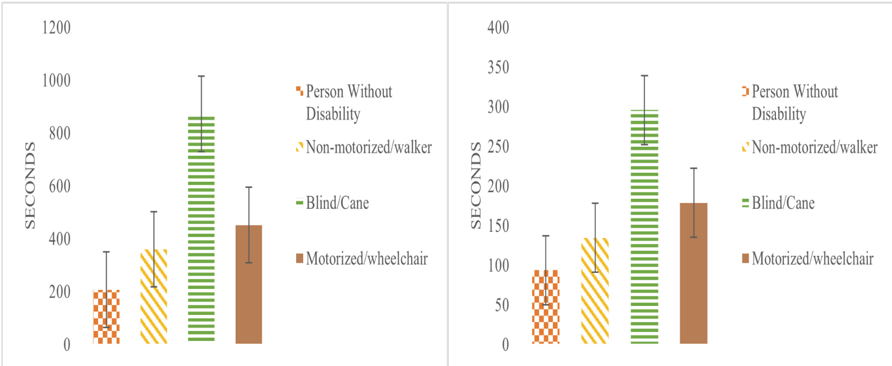
Table 1. Mobility speeds by disability group (Sources: Sharifi et al., 2016; Alves et al.,2020)

Disability Type	Speed (m/s)
No Disability	1.34
Motorized Wheelchair	0.67
Non-motorized Walker	0.98
Blind/Cane	0.78

4 Results

In selecting indoor spaces to demonstrate MABLESim's utility, diverse navigational challenges were considered, including simple and complex routes exemplified by selected buildings. Of the two buildings selected, one of them (building labeled WH) was smaller with only three floors with simpler routes. The other building (labeled HST building) had a larger geographical footprint with four floors and each floor being much larger than the WH building.

Navigational performance across various population groups were evaluated using metrics such as navigation time and efficiency, providing insights into accessibility challenges. The Navigation Time metric combines the impact of route complexity and walking speed to showcase how time efficient a person is when navigating routes. The Navigation Efficiency metric is a measure of the distance walked by a person compared to the actual shortest path that existed between a source destination pair. Results shown are an average across thirty simulation runs for each disability profile. Each run comprises of a random source destination pair within the buildings, with the sources constrained to be one of five entry points. All routes had to be at least 100 ft long to be included (Fig. 3).



(a) WH Building (b) HST Building

Fig. 2. WH and HST Building navigation time results

The results obtained from the simulations not only shed light on the intricate relationship between the complexity of a building and the navigational hurdles faced by its users but also underscore the pronounced impact on individuals with vision impairments. Figure 2a and b vividly depict how individuals with visual impairments encounter significantly prolonged navigation times, especially in more complex and expansive structures. What's noteworthy is that in simpler buildings, the discrepancy in navigation time between a visually impaired individual and someone using a motorized wheelchair is relatively minor. However, as the complexity of routes escalates, so does the magnitude of this discrepancy.

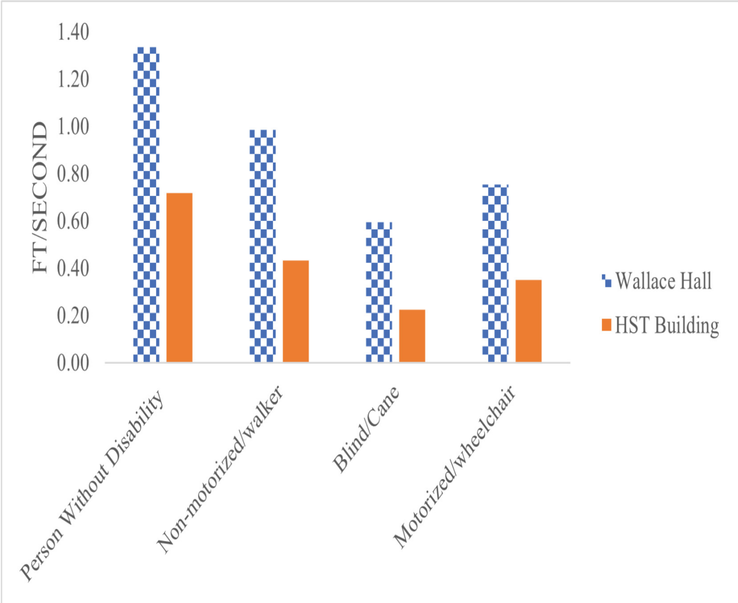


Fig. 3. WH and HST Building navigation efficiency results

Delving deeper into the analysis of route traversal efficiency, those with vision and mobility impairments cover the least distance relative to the time taken, indicative of the profound hurdles they confront in navigating built environments. For these individuals, navigational efficiency experiences a precipitous decline with the amplification of size and complexity.

5 Conclusions and Future Work

This paper introduced a simulation framework called MABLESim and how such a framework can be used to study accessibility challenges of indoor environments. Preliminary evaluation results confirm what is known about populations that face the most challenges in efficiently navigating indoor spaces. Beyond these validating insights, the simulator's principal contribution lies in its ability to serve as a versatile tool for exploring diverse configurations of built environments. Such simulation frameworks can facilitate the evaluation of emerging wayfinding technologies and their potential to enhance accessibility, offering a valuable platform for studying their real-world implications.

Acknowledgements. This work was supported in part by U.S. National Science Foundation awards #2409227, #2340870, #2345057

References

1. ADA: Americans with disabilities act (1990), pub. L. No. 101-336, 104 Stat. 327
2. Ahmetovic, D., Gleason, C., Ruan, C., Kitani, K., Takagi, H., Asakawa, C.: Navcog: a navigational cognitive assistant for the blind. In: Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI 2016), pp. 1–10. ACM (2016). https://doi.org/10.1145/2957265.2961829
3. Alves, F., Cruz, S., Ribeiro, A., Silva, A.B., Martins, J., Cunha, I.: Walkability index for elderly health: a proposal. Sustainability **12**(18), 7360 (2020), https://doi.org/10.3390/su12187360
4. Barry, K.: Toward universalism: what the ada amendments act of 2008 can and can't do for disability rights. Berkeley J. Employm. Labor Law **31**, 203–284 (2010)
5. Blanck, P.: On the importance of the Americans with disabilities act at 30. J. Disability Policy Stud. **34**(3), 176–198 (2023)
6. Diakité, A.A., Zlatanova, S.: Spatial subdivision of complex indoor environments for 3d indoor navigation. J. Location Based Ser. **11**(4), 213–235 (2017). https://doi.org/10.1080/13658816.2017.1376066
7. Fachada, N., Lopes, V.V., Martins, R.C., Rosa, A.C.: Towards a standard model for research in agent-based modeling and simulation. PeerJ Comput. Sci. **1**, e36 (2015). https://doi.org/10.7717/peerj-cs.36
8. Guerreiro, J., Sato, D., Ahmetovic, D., Ohn-Bar, E., Kitani, K.M., Asakawa, C.: Virtual navigation for blind people: transferring route knowledge to the real-world. Inter. J. Hum.-Comput. Stu. **135**, 102369 (2020). https://doi.org/10.1016/j.ijhcs.2019.102369
9. Karami, H., Karami, K., Khafaie, M.A., Zahednejad, S., Arastoo, A.A.: The physiological cost index and some kinematic parameters of walking and jogging in blind and sighted students. Inter. J. Musculoskeletal Pain Prevent. **4**(1), 251–258 (2019). https://doi.org/10.30476/ijms.2019.45386
10. Sharifi, M.S.: Analysis and Modeling of Pedestrian Walking Behaviors Involving Individuals with Disabilities. Ph.D. thesis, All Graduate Theses and Dissertations (2016). https://digitalcommons.usu.edu/etd/4959
11. Unity Technologies: Unity documentation (2023). https://docs.unity.com/

PDF Accessibility in International Academic Publishers

Oriane Pierrès[1,2(✉)], Felix Schmitt-Koopmann[1,2], and Alireza Darvishy[1]

[1] Zurich University of Applied Sciences (ZHAW), Steinberggasse 13, 8400 Winterthur, Switzerland
oriane.pierres@zhaw.ch

[2] University of Zurich, Rämistrasse 7, 8006 Zurich, Switzerland

Abstract. Academic articles are commonly published in Portable Document Format (PDF). However, for many people with visual impairments, PDF formats present significant accessibility issues. This study addresses two research questions: 1) To what extent are PDFs in prominent academic repositories accessible? and 2) To what extent are accessibility issues in academic articles known and addressed by repositories? To answer these questions, 8,000 PDFs from four prominent repositories (Springer, Elsevier, ACM, and Wiley) were retrieved and were automatically analysed according to accessibility criteria based on the Matterhorn Protocol. Additionally, a quantitative content analysis was performed on the submission guidelines of repositories to determine the degree to which accessibility is considered in document creation. Results suggest that most PDFs were not tagged in spite of the fact that some repositories included accessibility in their general author guidelines. This paper concludes with recommendations to improve the accessibility of papers in academic repositories.

Keywords: PDF accessibility · digital accessibility · higher education · persons with visual impairments

1 Introduction

PDF is the predominant format used for scientific publications. Yet scientific PDFs frequently lack enough structural information for proper interpretation by screen readers, making them inaccessible for many visually impaired users [1, 2]. Accessible PDFs require tags and metadata. Tags serve as labels providing semantic information about elements (e.g., heading, figure, formula, table, link, list) in a PDF. Metadata information such as title and default language enable screen-readers to read documents aloud in the correct language and with a meaningful, clearly marked title. Documents lacking tags and metadata can be impractical or nearly unusable for screen-reader users.

This study poses two research questions: 1) To what extent are PDFs in prominent academic repositories accessible? 2) To what extent are accessibility issues in academic articles known and addressed by repositories?

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 38–46, 2024.
https://doi.org/10.1007/978-3-031-62846-7_5

2 Related Work

Previous analyses of international repositories, such as Semantic Scholar and Web of Science, have reported accessibility results ranging from 2 to 15% of papers based on certain minimal accessibility features [1, 2]. Nganji [1] manually analysed articles related to disability, while Wang et al. [2] used Adobe Acrobat to automatically assess PDFs from various scientific fields. Like Wang et al. [2], we took an automated approach to investigate the accessibility of PDFs across subjects. Instead of examining digital archives like Web of Science, we focused on the repositories of prominent publishers (Springer, Elsevier, ACM, and Wiley). In addition to analysing the accessibility of the PDFs within these major repositories, we investigated and compared their respective author guidelines with regard to accessibility requirements.

The existing studies looked at articles published between 2014 and 2018 [1] and 2010 and 2019 [2]. This study provides an assessment of recently published articles, as 86% of our collected PDFs were published in 2023 or 2024.

Wang et al. [2] focused on five accessibility compliance criteria: alternative text (also called "alt text"), table headers, tagged PDFs (metadata), default language, and tab order. This study investigated title metadata and the presence of tags in greater detail by looking at the percentage of tagged content and the presence of semantically meaningful tags for headings, figures, tables, links, and lists. Based on our results as well as our analysis of author guidelines, we have created a set of recommendations to improve the accessibility of PDFs.

3 Methods

To address our two research questions, we conducted two quantitative content analyses: one performed automatically to assess the accessibility of PDFs contained in four prominent repositories, and another conducted manually to identify mentions of accessibility measures in the author submission guidelines from the same repositories.

3.1 Automatic Analysis of PDFs in Repositories

Our analysis focused on four academic repositories: Springer, Elsevier, ACM, and Wiley. Before the study, the top journals from each of Google Scholar's eight subject categories were investigated to assess where research is most frequently published. This short research enabled us to determine which repositories and keywords were more likely to cover a diverse set of subjects.

Although detailed accessibility testing requires manual checks, automated analysis makes it possible to quickly identify inherent issues in a large number of PDFs. Thus, we developed Python crawlers to gather the 50 most recent papers using 40 keywords in each of the four repositories, totalling 8,000 articles. For Elsevier, only open-access articles could be retrieved. Publication dates range from 1971 to 2024, but about 86% of the PDFs were published in 2023 or 2024. We analysed the papers using the PAVE engine [3]. The PAVE engine can generate an accessibility report based on the Matterhorn Protocol's automatically checkable criteria [4]. We logged a selection of 10 criteria from the Matterhorn Protocol which are essential for accessible PDFs. These include 7 tag-based criteria (the proportion of tagged content and the use of tags for headings, figures, formulae, lists, tables, and links) and 3 metadata-based criteria (appropriately marked title data, appropriate language setting, and whether the document is marked as "tagged" in its metadata).

Table 1. Accessibility criteria used to assess the PDFs based on Matterhorn Protocol

Accessibility criteria	Description
Explicit disability mention	The guideline explicitly mentions accessibility or disabilities
Structured content with tagged headings	The guideline requires the use of defined heading styles. For this requirement, the Word templates (if available) were checked to ensure that they encourage the use of correct heading styles
Defined lists	The guideline mentions the need to define list as lists
Meaningful links	The guideline mentions writing meaningful links (i.e. not "https://…, nor "click here", but "Author submission guidelines")
Alt text for images	The guideline requires authors to add alternative texts to their images
Formulae	The guideline mentions how to make math formulae accessible (e.g. including an alt text)
Colour contrast	The guideline requires the use of high colour contrast or using patterns instead of colour-only for all figures, ensuring that they are visible to people with a colour impairment. The criterion was marked as partially fulfilled if the criterion was mentioned briefly without explicit mention of accessibility or disabilities
Font size	The guideline recommends or requires the use of a font size of at least 12 point for body text
Font type	The guideline recommends or requires the use of a sans serif font type
Reading order	The guideline requires authors to check the reading order of their PDF
Tables	The guideline mentions how to make tables accessible, most notably by mentioning table tags and creating simple tables. Adding captions is a plus
Plain/simplified language	The guidelines mention the plain or simplified language

3.2 Manual Analysis of Author Submission Guidelines

To analyse the author submission guidelines from the four repositories, we looked for mentions of twelve important accessibility elements: Structured content with tagged headings, defined lists, meaningful links, alternative text for images, accessible mathematical formulae, colour contrast, font size, accessible font type, proper reading order, accessible tables, plain language, and an explicit mention of accessibility or disability (Table 1). The criteria were based on Web Content Accessibility Guidelines (WCAG) 2.1 [5] and other accessibility recommendations [6, 7].

4 Results

4.1 Tags in PDFs

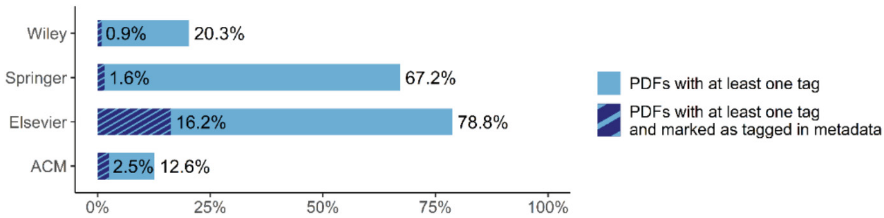


Fig. 1. Distribution of PDFs with at least one tag and marked as tagged (in metadata) in the repositories (N = 8,000).

We considered two criteria to check whether the PDF is tagged: one relates to the metadata and the other to the content. The "tagged PDF" metadata is formal information entered by the document creator that indicates to assistive technologies that the document contains tags. The content criterion is based on whether at least one tag was positively identified by the PAVE engine. Documents collected in Elsevier were more frequently tagged than the other publishers, with about 79% of PDFs containing at least one tag, compared to 67% for Springer, 20% for Wiley, and 13% for ACM (Fig. 1). It should be noted, however, that the presence of just one single tag is very low threshold for accessibility. Across all repositories, about 88% PDFs had more than 80% untagged content, i.e. most documents were completely or partially untagged. Moreover, only a small proportion of PDFs were also marked as "tagged" in their metadata. Only 16.25% of PDFs in Elsevier fulfil both conditions, whereas less than 3% do in other repositories. This finding concurs with the analysis of Wang et al. [2] which found that just 13.4% of the analysed PDFs were marked as "tagged" in metadata.

Figure 2 illustrates the distribution of semantically meaningful tags within the 3,578 PDFs containing at least one tag. None of the PDFs included a tagged mathematical formula, and only two PDFs across all repositories had any tagged tables. In Wiley, only figure tags were identified, while in Springer, less than 3% of PDFs contained a tagged figure, heading, or list. In ACM, a small number of papers contained at least one tagged list (8%), figure (13%), or heading (2%). In Elsevier, about 68% had at least one tagged figure, but only 11% included heading tags. Considering that these PDFs are academic articles, it is unlikely that the tag distribution reflects a real absence of formulae, headings, links, lists, tables, or figures in the documents. This means that most PDFs are not tagged properly.

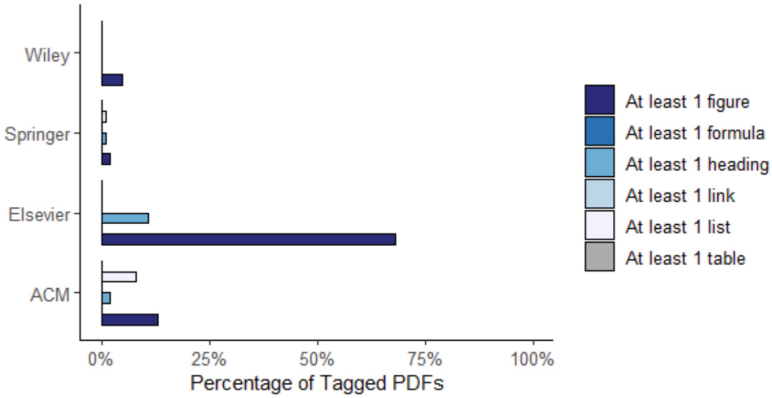


Fig. 2. Distribution of semantically meaningful tags among the 3,578 PDFs with at least one tagged operator.

Among the 1,161 PDFs containing at least one figure, about 77% included an alt text. However, the chosen method of investigation only checks whether the alt text field is empty. A manual analysis suggests that the identified alt texts were largely not meaningful, an issue that was pointed out by Nganji [1]. Moreover, the high proportion of PDFs with an alt text can be explained by the fact that the percentage is calculated based on the smaller subsample of PDFs containing at least one figure. When considering all 8,000 PDFs analysed, the overall percentage of papers containing alt tags drops to 11%, a number similar to Wang et al. with 7.5% [2] and Nganji with 10.5% [1].

4.2 Metadata in PDFs

No PDF claimed to fulfil the PDF UA standard.

All 8,000 PDFs contained a title in their metadata. The information was frequently stored correctly in dc:title in XMP file, except in ACM (Fig. 3). Nevertheless, the PDFs usually did not have their preference view setting on "DisplayDocTitle", which guarantees that screen readers read the title of the document and not the file name, to the exception of ACM that usually fulfilled only the DisplayDocTitle requirement.

Figure 4 indicates that in Springer, Elsevier, and ACM, the PDFs usually have a set language, but this is not systematic as none reach 100%. Wiley lags behind with only 20% of its PDFs with language metadata. In their sample of papers dating from 2010 to 2019, Wang et al. [2] found that this setting was in an upward trend, with "10% compliance in 2010 to more than 25% in 2019" (p.11). This evolution in time has thus continued over time.

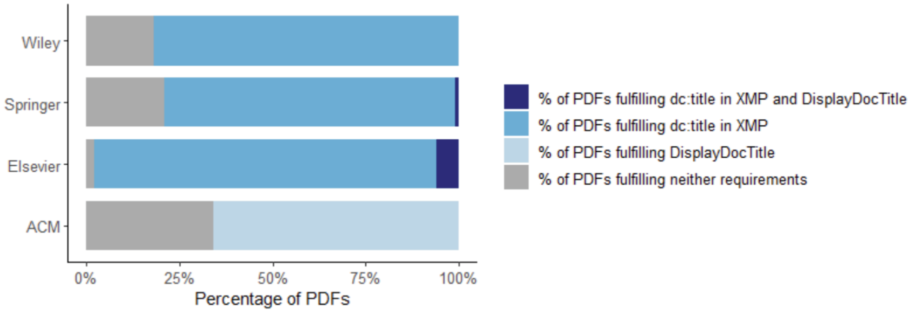


Fig. 3. Distribution of title metadata in repositories (N = 8,000)

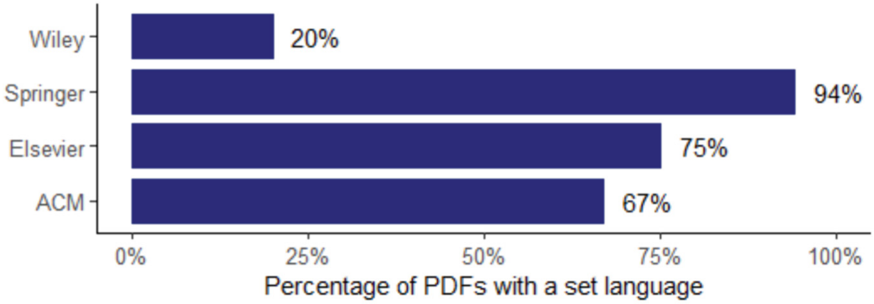


Fig. 4. Distribution of PDFs with a set language metadata in repositories (N = 8,000)

4.3 Repository Author Guidelines

All publishers but Wiley included references to disabilities in their general author submission guidelines (Table 2). Nevertheless, the mentions were usually brief and superficial, with ACM being the exception. ACM had the most extensive coverage of accessibility measures, meeting half of the investigated criteria. However, based on our PDF analysis, their commitment did not translate into more accessible documents. This suggests that the accessibility requirements are not implemented correctly and would require control before publication.

Table 2. Content analysis of the general author guidelines of the four analysed repositories

Repository	Springer	Elsevier	ACM	Wiley
Explicit disability mention	Yes	Yes	Yes	No
Structured content with tagged headings	Yes	No	Yes	No
Defined lists	Yes	No	Yes	No
Meaningful links	No	No	No	No
Alt text for images	Partially	Partially	Yes	Yes
Accessible formulae	No	No	No	No
Colour contrast	Partially	Yes	Yes	Yes
Font size	No	No	No	No
Font type	No	No	No	No
Checking reading order	No	No	No	No
Accessible tables	No	No	Yes	No
Plain / simplified language	No	No	No	No

5 Discussion and Conclusion

Results indicate that even where general author guidelines include accessibility requirements, scientific PDFs are still largely not accessible. In particular, this study highlights that inaccessibility arises from a lack of tagging, which is a fundamental requirement for accessible documents.

Even though all publishers state that authors must follow the guidelines of their specific journals, the publisher's general guidelines set a standard for all journals. For that reason, publishers can make accessibility a requirement for publication. Along with ensuring that articles follow visual standards, publishers could check that accessibility measures are implemented. As recent research has explored the potential of artificial intelligence for document accessibility [8, 9], much of the remediation work could be automated. In particular, a tool like PAVE or PAC3 could enable publishers to request or require authors to check the accessibility of the final version of a PDF before submission. This could be especially interesting for fields that use LaTeX more frequently than Microsoft, as the latter is associated with greater accessibility compliance than the former [2].

Although checking the final version of a PDF is crucial to ensure all tags are included, PDF accessibility must be considered from the beginning, both in templates and in author guidelines, because certain design choices (e.g. tables and colour contrast) cannot be easily remediated. Moreover, the correct use of heading styles can ensure that PDFs are structured correctly from the start. Favouring sans-serif fonts (e.g. Arial), as well as a minimum of 12-point size for body text, are also easy measures to integrate into guidelines. Finally, authors should provide alt text for their images and formulae.

Additionally, this study shows that many PDFs fulfill the default language setting, except in Wiley. Nevertheless, this setting is easy to correct and should be systematic. Similarly, all PDFs included a title metadata, but these were not stored correctly. While correcting metadata is not the most crucial of accessibility measures (unlike tagging), it is an easy step to implement.

5.1 Limitations

This study includes some limitations. First, PDFs were analysed automatically. While this enabled an assessment of a large sample of articles, criteria such as reading order could not be tested. However, as the results indicate that PDFs usually lack tags, a more detailed analysis would not have changed the interpretation of the results significantly. A PDF cannot be correctly tagged if it is never tagged to begin with.

Additionally, this article focused primarily on the accessibility of PDFs for screen-reader users and therefore did not investigate further accessibility criteria such as plain language. Nevertheless, the analysis of the guidelines checked for recommendations regarding plain language and indicates that no publishers made it a requirement or recommendation. Further research could investigate how this is reflected in articles and identify best practices.

Finally, the second analysis focused on author submission guidelines to assess how guidelines could have influenced the accessibility of created PDF documents due to the widespread use of this format. It's important to note that these results do not imply that publishers are completely neglecting accessibility services. Future studies could investigate how frequently alternative formats like HTML or ePub are available and assess their accessibility.

Acknowledgments. This study was funded by the Swiss National Science Foundation under grant 194677 and the two swissuniversities projects P7 and P8.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Nganji, J.T.: An assessment of the accessibility of PDF versions of selected journal articles published in a WCAG 2.0 era (2014–2018). Learned Publishing **31**, 391–401 (2018). https://doi.org/10.1002/leap.1197
2. Wang, L.L., et al.: Improving the accessibility of scientific documents: current state, user needs, and a system solution to enhance scientific PDF accessibility for blind and low vision users. arXiv abs/2105.00076 (2021)
3. Doblies, L., Stolz, D., Darvishy, A., Hutter, H.-P.: PAVE: a Web application to identify and correct accessibility problems in PDF documents. In: Miesenberger, K., Fels, D., Archambault, D., Peňáz, P., Zagler, W. (eds.) ICCHP 2014. LNCS, vol. 8547, pp. 185–192. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-08596-8_29
4. PDF Association: Matterhorn Protocol 1.1 PDF/UA Conformance Testing Model (2021). https://pdfa.org/wp-content/uploads/2021/04/Matterhorn-Protocol-1-1.pdf
5. World Wide Web Consortium (W3C): Web Content Accessibility Guidelines (WCAG) 2.1 (2023). https://www.w3.org/TR/WCAG21/
6. British Dyslexia Association: Dyslexia Style Guide (2023). https://cdn.bdadyslexia.org.uk/uploads/documents/Advice/style-guide/BDA-Style-Guide-2023.pdf?v=1680514568
7. Coolidge, A., Doner, S., Robertson, T., Gray, J.: Accessibility toolkit. BCcampus (2018)
8. Schmitt-Koopmann, F.M., Huang, E.M., Darvishy, A.: Accessible PDFs: applying artificial intelligence for automated remediation of STEM PDFs. In: Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility, Athens, Greece, pp. 1–6. ACM (2022). https://doi.org/10.1145/3517428.3550407

9. Darvishy, A., Nevill, M., Hutter, H.-P.: Automatic paragraph detection for accessible PDF documents. In: Miesenberger, K., Bühler, C., Penaz, P. (eds.) ICCHP 2016. LNCS, vol. 9758, pp. 367–372. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-41264-1_50

The Accessibility Paradox: Can Research Articles Inspecting Accessibility Be Inaccessible?

Attila Bekkvik Szentirmai(✉), Yavuz Inal, and Anne Britt Torkildsby

Department of Design, Norwegian University of Science and Technology, Gjøvik, Norway
attila.b.szentirmai@ntnu.no

Abstract. Relevant literature focusing on the accessibility of electronic documents (e.g., PDFs) has been increasing in recent years. Despite its significance, paradoxically, scientific articles, even those related to accessibility-themed topics, often fail to provide fully accessible content, thus creating a gap between theory and practice. Therefore, we aimed to explore how PDF versions of academic articles on digital accessibility evaluation, published in the past ten years, comply with established accessibility standards. We performed a tool-based evaluation using Adobe Accessibility Checker and a manual evaluation to inspect the articles comprehensively. The results showed that none of the analyzed articles were problem-free. They, however, contained recurring, severe accessibility barriers, making it highly challenging for people who rely on screen readers to access information. Also, the evaluated data showed no discernible pattern of changes over the years.

Keywords: Digital Accessibility · Accessibility Evaluation · Manual Evaluation · Adobe Accessibility Checker · PDF · PDF/UA · WCAG

1 Introduction

Since the early 1990s, Portable Document Format (PDF) has been one of the leading digital formats for sharing and publishing information. PDFs can be read across multiple devices, as they provide the same structure and appearance across different systems, yet this does not guarantee the accessibility of these documents. Inaccessible PDFs exclude readers with visual impairments, amongst others, from accessing their content. The International Organization for Standardization (ISO) provides a standard, ISO 14289, also known as PDF/UA (Universal Accessibility), to ensure that generated PDFs meet accessibility requirements such as alternative text, tags, headings, titles, tables, forms, and links [7, 10]. The standard also limits certain types of content that have the potential to create difficulties in effective interaction for individuals who rely on assistive technologies to live independently - both privately and professionally. However, despite all those efforts, the results of previous studies show that the evaluated PDF documents have considerable accessibility issues - thereby causing severe problems for people who use assistive technologies daily [1, 4, 5, 18].

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 47–54, 2024.
https://doi.org/10.1007/978-3-031-62846-7_6

PDF accessibility is crucial for people with visual impairments to equally access online information and services without undue effort. However, providing accessible documents for everyone - regardless of abilities, age, or skills - seems not to be the primary concern in academia. Just as insufficient awareness among practitioners causes inaccessible digital systems [3, 12], the low awareness of content creators (e.g., publishers, researchers, and educators) prevents individuals with disabilities from effectively accessing the PDF versions of scientific documents, published in journals and conference proceedings. This is the biggest obstacle to the accessibility of online content in conference venues, even related to disabilities, digital accessibility, and universal design. For example, Brady et al. [5] tested the accessibility of 1811 papers from several top conferences in the field of human-computer interaction and accessibility. The authors found that most of the evaluated papers were inaccessible.

Similarly, Nganji conducted studies to evaluate the accessibility of journal articles in the years 2015 and 2018. The author examined a total of 200 articles on disability, selected randomly from different journals by employing tool-based and manual evaluation with a screen reader. In the earlier study [16], articles published between 2009 and 2013 were evaluated, whereas in the follow-up study [17], the evaluation focused on articles published between 2014 and 2018 - notably from the same journals. In the first study, 97% of the articles did not include alternative text for images and 95.5% of the articles needed to be tagged. In the follow-up study, slight progress was observed; that is to say, only 10.5% had an alternative text for images, and 15.5% were tagged.

Some authoring tools were found to potentially improve the quality of documents by making them accessible with high effectiveness and accuracy [18]. However, each tool has shortcomings that negatively affect its performance and prevalence [8]. In short, they are not good enough to create fully accessible documents [19]. Although some tools might help improve the accessibility of PDFs, it is hard to automate the process without hands-on adjustments by content creators [6]. This indicates the significance of having a high level of awareness amongst publishers and researchers, not to mention human insight, especially in academia, to provide inclusivity when generating electronic documents.

In sum, even though PDF documents are widely shared on digital platforms - hence ought to meet accessibility requirements to maximize usability for everyone, previous research shows that not all of them do. Paradoxically, even platforms for scientific articles, including journal and conference proceedings, related to accessibility-themed topics often fail to provide fully accessible content, thus creating a gap between theory and practice and, more importantly, excluding people from accessing the information. Taking this one step further, we explored to what extent academic articles on the accessibility evaluation of a digital system comply with accessibility requirements for everyone.

2 Methods

We explored the current status of accessibility compliance of academic articles, evaluating the accessibility of digital systems (e.g., websites, software, and programs). We analyzed articles published in the last decade using Adobe Accessibility Checker. We also manually evaluated each article for a comprehensive inspection, as the combination of tool-based and manual evaluation is highly suggested to obtain more reliable data and maximize the coverage of accessibility issues [14, 15]. To access the most relevant articles, including journals and conference proceedings, we searched the literature systematically across a wide scientific database, namely Web of Science. The search was limited to articles published between January 2014 and September 2023, and the search terms were based on a combination of keywords: online, web, digital, website, software, program, assessment, evaluation, analysis, and accessibility. Please note that the keywords were combined using the Boolean operators (e.g., OR and AND).

2.1 Search Process

The search process focused on three key areas: accessibility, digital, and evaluation, and accessed relevant articles throughout a four-step process. The initial focus was on 'accessibility' as it is this study's main research subject. Second, keywords related to 'digital' were added to identify articles that evaluated the accessibility of digital systems (e.g., online OR web OR website OR software OR digital OR program). Third, the results were refined to include the keywords related to 'evaluation' (e.g., assessment OR evaluation OR analysis). We included articles that met the following inclusion criteria: focusing on the accessibility evaluation of a digital system and providing empirical evidence about the evaluation. Also, we excluded the articles that met at least one of these criteria: not written in English, published before 2014, not having a full text or published as abstracts, designed as non-empirical research, such as opinion papers, reviews, or editorials.

2.2 Dataset and Procedure

Following the search, we identified relevant articles, considering their titles and abstracts. After removing the duplicates, we applied the inclusion and exclusion criteria to select the most relevant articles for the analysis. This resulted in the involvement of 135 articles. However, we could not access the full text of seven of them, so we included 128 articles. The articles included in the evaluation were distributed between seven major publishers, including Springer, IEEE Digital Library, ScienceDirect, SAGE, ACM, Taylor & Francis Online, Wiley, MDPI, and twenty-five other publishers. Next, we used the Adobe Accessibility Checker (ACC) to evaluate the PDF versions of selected articles. Based on the Web Content Accessibility Guidelines (WCAG) and PDF/UA requirements [2], the checker allows for detailed outcomes of the accessibility inspection. We also manually tested each article for violations that the tool may not have detected, or in case it had delivered false positives. Overall, this method allowed us to systematically evaluate the accessibility of the selected PDF files and, furthermore, identify common violations causing accessibility barriers.

3 Results

All the articles evaluated in the study failed to meet the requirements within the AAC's seven accessibility aspects [2], namely document, page content, forms, tables and lists, alternative text, and headings. We also manually evaluated the logical reading order, color contrast, and navigation links. Details of the tool-based and manual evaluations are presented in the following sections.

3.1 Document

All articles had accessibility permission flags enabled, which ensures content can be copied for assistive technologies. Similarly, none of the articles were image-only PDFs, meaning that the content does not include scanned images and is thus readable by assistive technologies like screen readers. However, the evaluated articles had severe accessibility issues concerning tagged PDF and logical reading order. Only 16 articles (12.5%) had proper structural tags for screen readers and navigation. This results in the fact that most of them (n = 123, 96.1%) need to have content sequenced logically to achieve a meaningful order, as the failure in tagged PDF creates problems with specifying the correct reading order.

The evaluation results indicated that 52 articles (40.6%) correctly set the primary document language, which ensures screen readers pronounce the content appropriately. The lack of title was found to be one of the most violated issues in the articles. In fact, only five of them (3.9%) contained the title in the title bar, which is necessary to improve user experience because it provides a visual cue about the document and its title. On the positive side, most articles provided bookmarks (n = 122, 95.3%) that eased the navigation, and almost all met the contrast requirements (n = 127, 99.2%) with distinguishable contrasts between text color and document background. The articles in question have a traditional black (text) and white (background) format that enhances the legibility and visibility of the documents.

3.2 Page Content

Our evaluation highlighted that only four (3.1%) articles included tagged content, which is important for screen readers to understand the content. Most articles (n = 98, 76.6%) lacked tagged annotations, such as comments and editorial marks (e.g., insert and highlight). The absence of tagged annotations causes further barriers for screen reader users in the attempt to access and perceive content. As regards tab order, it is important in PDFs as tabs are commonly used to navigate a PDF document. However, only ten articles (7.8%) showed a correlation between the visual and structural hierarchy and tab navigation, potentially negatively impacting the documents' navigation and general user experience. Almost half (n = 67, 52.3%) of the articles had reliable character encoding, which ensures that the document's special characters are displayed correctly and consistently.

All evaluated articles passed the tagged multimedia, screen flicker, scripts, and timed responses criteria. As for tagged multimedia, it validates if the content, such as audio and video files, is tagged and identifiable for assistive technologies. Screen flicker protects readers by checking the presence of flashings, animations, and scripts that can cause seizures in people with photosensitive epilepsy. Scripts, conversely, ensure that embedded scripts do not create problems with utilizing assistive technologies, keyboard navigation, and input devices that individuals with disabilities may need to access the documents. Lastly, timed responses check off if readers are not forced to take action in forms and have enough time to interact with the content within a set timeframe. The findings of the study also showed that most documents (n = 89, 69.5%) had non-repetitive navigation links for easier navigation and access to external sources.

3.3 Forms, Tables, and Lists

None of the evaluated articles contained forms, so they were error-free concerning tagged form fields and field descriptions. The former checks structure for interactive forms, such as text field checkboxes; to ensure identification and interaction with the form elements, whereas the latter provides explanatory text for form fields. Only 20 articles (15.6%) had a table structure considering table rows, table headers, and table data - and only nine articles (7%) contained defined header rows in tables, which is essential to understanding the importance of the table and table data. These are critical table elements for people with screen readers to navigate and perceive the table content easily. Furthermore, 111 of the articles' tables (86.7%) had data integrity problems (e.g., rows and column arrangements), which show irregularities and layout formatting issues. Similarly, only 20 articles (15.6%) had ordered format with visual organization, content organization, and navigation (list: list items) and appropriate structure of the list label (LBI) and list body (LBody), which enhances the readability through the document.

3.4 Alternate Text and Headings

Alternative text is crucial for people who need help to visually perceive non-textual content and elements, e.g., images, figures, graphs, and charts in documents. As today, assistive technologies can only interpret and describe the visual content to nonvisual readers with an alternative text. Document headings, however, are crucial to both the visual, hierarchical, and logical structure and provide visual clarity, visual clues, and feeding logical structure and navigation path for assistive technologies. The results of the study showed that most articles (n = 115, 89.8%) failed to provide meaningful descriptions for non-text content in the documents, which might create major accessibility barriers for individuals with visual impairments.

Furthermore, most articles (n = 108, 84.4%) failed to skip an alternative text where it is unnecessary (nested alternative text). This issue is important because redundant information - that is already present in the surrounding text - should not be interpreted with assistive technologies, e.g. if a chart and its results are explained in the text. In other words, it is unnecessary to do the same as an added alternative text to the visual chart. The same number of evaluated articles (n = 108, 84.4%) failed to tag decorative elements (associated with content) and were empty or tagged as decorative instead. Please note that the document should not hide annotation for screen reader users but rather identify and describe the reasons. Only 20 articles (15.6%) passed to hide annotation information (hides annotation) from blind users, making it impossible for them to understand the content. Most of the articles failed to provide essential information for elements beyond images (other elements alternate text). In contrast, only 20 articles (15.6%) had a structured and coherent document hierarchy (appropriate nesting).

4 Discussion and Conclusion

Previous research has shown that partially or completely inaccessible websites, mobile applications, and electronic documents create significant barriers to especially people with disabilities [1, 9, 13]. Although PDFs have been used for sharing and publishing information for years in fields like education, business, health, and academia - and, on top of that, the increasing interest in research on the accessibility of digital systems among researchers - insufficient accessibility remains a persistent challenge, even for accessibility-themed articles published on scientific platforms. Herein lies the paradox: If you want to be exemplary in inclusion and accessibility, you ought to lead by example. In this study, we explored to what extent PDF versions of academic articles concerning the accessibility evaluation of digital systems, such as websites, software, and programs, adhere to established accessibility requirements (i.e., WCAG, PDF/UA). Our findings align with previous research [1, 4, 5, 18], concluding that all evaluated articles were published without sufficient accessibility measures; in turn creating a significant gap between theory and practice. Furthermore, they contained violations in several categories, indicating that they lacked essential formatting, such as appropriate document structure, logical reading order, and alternative text for non-textual content. We conclude that the insufficient accessibility of the accessibility-themed articles evaluated in this study generally hinders access for people with visual impairments or blindness - relying on assistive technologies like screen readers - to navigate and understand digital information effectively.

The ecosystem of scientific publication includes three main pillars: authors, publishers, and authoring tools, each of which has equal importance on the quality of the outcome. We argue that the main challenges in creating accessible PDFs are related to the lack of awareness of publishers and researchers. Besides, efficient authoring tools are needed to notify content creators about errors in their documents before they export. Therefore, raising awareness of diverse user needs and wants is the first step to creating sustainable change [11]. The results of the present study showed that even articles evaluating the accessibility of digital systems fail to meet the established accessibility requirements. Authors' responsibilities, on the other hand, are to generate accessible documents that everyone can access, understand, perceive, and easily interact with.

Web accessibility has been mandatory in many countries for years. The same legal action should be taken to access electronic documents, as they constitute the main parts of sharing and publishing information in the digital society. Furthermore, publishers should make it mandatory for authors before they submit their documents. However, many journals often do not require accessible documents, as accessibility is not a criterion and is furthermore absent from journal templates, which does not support creating inclusive documents. During our manual evaluation, we identified several violations caused by journal templates and their formatting issues, resulting in the fact that on one occasion, a paper failed the color contrast check due to the journal's color theme on tables. Moreover, on almost every occasion, journals' logos lack alternative text, association with content, and/or alternative text, and some journals even generate cover pages containing accessibility issues and broken links. We argue that journals, especially within the themes of inclusion and accessibility, ought to provide accessibility guidelines and support authors to make their documents accessible. Instead, in many journals and conferences, paper submissions are expected to follow accessibility requirements. However, it is up to the authors to create accessible documents.

Authoring tools with accessibility features should promote the creation of accessible documents and inform content creators about their benefits. What is more, they ought to integrate features that comply with established accessibility requirements. Although some research shows that authoring tools can improve the accessibility of documents [18], they have several shortcomings, negatively affecting their efficiency and prevalence [8]. For example, MS Word provides accessibility features, but the accessible formatting can disappear when exporting to PDF. Other tools, such as LaTex, do not have additional plugins for accessible formatting at all, making creating inclusive documents difficult, even for those aware of its necessity.

From this study, we can conclude that the accessibility of PDFs is still an issue, even in research papers on accessibility topics published in the last decade. What is more, our findings suggest that there have not been any significant improvements over the years. Hence, there is an urgent need to raise authors' awareness by making it mandatory for published documents to be fully accessible. Also, we suggest that providing effective and efficient authoring tools might be the first step of change towards a more inclusive and sustainable digital society. This includes publishers who are responsible for the final versions of the articles by providing an accessible template and making the final check before publishing, seeing that authors can't control the accessibility of their articles after submitting the "author version". Future research should explore the potential reasons for insufficient accessible documents from both publishers' and authors' perspectives and furthermore examine the efficiency of existing authoring tools.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Acosta-Vargas, P., Luján-Mora, S., Acosta, T.: Accessibility of portable document format in education repositories. In: 9th International Conference on Education Technology and Computers, pp. 239–242 (2017)
2. Adobe Acrobat XI Pro Accessibility Guide: Best Practices for PDF Accessibility. https://acrobatusers.com/assets/uploads/public_downloads/2190/acrobat-xi-pro-accessibility-best-practice-guide.pdf. Accessed 28 Mar 2024
3. Antonelli, H.L., Rodrigues, S.S., Watanabe, W.M., de Mattos Fortes, R.P.: A survey on accessibility awareness of Brazilian web developers. In: 8th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion, pp. 71–79 (2018)
4. Bigham, J.P., Brady, E.L., Gleason, C., Guo, A., Shamma, D.A.: An uninteresting tour through why our research papers aren't accessible. In: CHI Conference Extended Abstracts on Human Factors in Computing Systems, pp. 621–631 (2016)
5. Brady, E., Zhong, Y., Bigham, J.P.: Creating accessible PDFs for conference proceedings. In: 12th International Web for All Conference, pp. 1–4 (2015)
6. Darvishy, A.: PDF accessibility: tools and challenges. In: Miesenberger, K., Kouroupetroglou, G. (eds.) ICCHP 2018. LNCS, vol. 10896, pp. 113–116. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-94277-3_20

7. Devine, H., Gonzalez, A., Hardy, M.: Making accessible PDF documents. In: 11th ACM Symposium on Document Engineering, pp. 275–276 (2011)
8. Doblies, L., Stolz, D., Darvishy, A., Hutter, H.-P.: PAVE: a Web application to identify and correct accessibility problems in PDF documents. In: Miesenberger, K., Fels, D., Archambault, D., Peňáz, P., Zagler, W. (eds.) Computers Helping People with Special Needs, pp. 185–192. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-08596-8_29
9. Domínguez Vila, T., Alén González, E., Darcy, S.: Accessibility of tourism websites: the level of countries' commitment. Univ. Access Inf. Soc. **19**(2), 331–346 (2020)
10. Drümmer, O., Erle, M.: PDF/UA – a new era for document accessibility: understanding, managing and implementing the ISO standard PDF/UA (universal accessibility): introduction to the special thematic session. In: Miesenberger, K., Karshmer, A., Penaz, P., Zagler, W. (eds.) ICCHP 2012. LNCS, vol. 7382, pp. 585–586. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-31522-0_88
11. Inal, Y.: I feel for you! The effect of empathy exercises on accessibility awareness. Universal Access in the Information Society (2024)
12. Inal, Y., Guribye, F., Rajanen, D., Rajanen, M., Rost, M.: Perspectives and practices of digital accessibility: a survey of user experience professionals in Nordic countries. In: 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society, pp. 1–11 (2020)
13. Inal, Y., Mishra, D., Torkildsby, A.B.: An analysis of web content accessibility of municipality websites for people with disabilities in Norway: Web accessibility of Norwegian municipality websites. In: Nordic Human-Computer Interaction Conference, pp. 1–12 (2022)
14. Inal, Y., Torkildsby, A.B.: Web accessibility in Higher Education in Norway: to what extent are university websites accessible? In: Nocera, J.A., Lárusdóttir, M.K., Petrie, H., Piccinno, A., Winckler, M. (eds.) INTERACT 2023, pp. 111–122. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-42280-5_7
15. Ismailova, R., Inal, Y.: Comparison of online accessibility evaluation tools: an analysis of tool effectiveness. IEEE Access **10**, 58233–58239 (2022)
16. Nganji, J.T.: The Portable Document Format (PDF) accessibility practice of four journal publishers. Libr. Inf. Sci. Res. **37**(3), 254–262 (2015)
17. Nganji, J.T.: An assessment of the accessibility of PDF versions of selected journal articles published in a WCAG 2.0 era (2014–2018). Learned Publishing **31**(4), 391–401 (2018)
18. Pradhan, D., et al.: Development and evaluation of a tool for assisting content creators in making PDF files more accessible. ACM Trans. Accessible Comput. **15**(1), 1–52 (2022)
19. Ribera, M., Pozzobon, R., Sayago, S.: Publishing accessible proceedings: the DSAI 2016 case study. Univ. Access Inf. Soc. **19**(3), 557–569 (2020)

Testing Usability of Tools for Making PDFs Accessible: Pressing Issues and Pain Points

Linus Manuel Osthof, Thorsten Schwarz(✉), and Karin Müller

Karlsruhe Institute of Technology, Karlsruhe, Germany
{thorsten.schwarz,karin.e.mueller}@kit.edu
http://www.access.kit.edu

Abstract. Nowadays, most digital documents are available in the form of PDF files. The main problem, however, is that there are hardly any tools that automatically generate accessible PDFs, which is why they have to be made accessible. In a study with nine participants, we investigate the usability of three tools (Adobe Acrobat Pro, Axes and PAVE), used for making PDFs accessible. Our results show that the SUS score of PAVE is best, followed by AXES and Acrobat. None of the three tools, however, achieves an SUS score above 50 and therefore need improvements. The qualitative analysis points at three main problems: (i) users felt lost and do not know what to do, (ii) there is no or insufficient possibility to track changes and (iii) the interfaces are not easy to use.

Keywords: Accessibility · PDF Documents · User Study

1 Introduction

The Portable Document Format (PDF) format is one of the most widely spread digital document formats in use today, with many people relying on them to transmit information over long distances as well as having a printable document format.

The PDF Association estimates that in 2021 around 90% of documents online were provided in PDF format [7]. Documents like invoices, instruction manuals, but also worksheets, presentations, official documents distributed by government offices etc. are provided to users as PDFs.This means that in order to participate in (digital) society nowadays requires the ability to read PDF documents. Using PDFs for people with sight usually is straight forward because nearly every device has a way to display PDFs and most text editors can generate them. However, most PDFs created by different tools are not accessible for visually impaired readers.

To address this problem lawmakers started to pass laws and regulations to improve the state of accessibility in society. For example, the European Accessibility Act (EAA) forces European countries to implement accessibility in their local laws following general accessibility guidelines.

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 55–62, 2024.
https://doi.org/10.1007/978-3-031-62846-7_7

Adjustments of PDF documents to make them accessible include adding the necessary metadata and tags in order to make it readable with a screenreader. This especially applies for non-linear content (i. e. tables, pictures, graphs) which has to be transformed into linear content (i. e. text). This task is non-trivial in most cases and can also become very complex the more “layers” of nested content has to be translated (e. g. tables with tables inside, or composite images that contain multiple different regions that may in turn contain other complex structures such as tables etc.). Thus, it is essential that there are intuitive tools that support the adaptation of PDFs. In this paper, we investigate the usability of three most known tools that can be used to make PDFs accessible: Adobe Acrobat Pro, Axes and PAVE.

2 Previous Work

The standard for accessible PDFs, PDF/Universal Accessibility (UA) was first published in 2012 and updated in 2014 [6]. Documents created before that date are (most likely) not accessible. While PDF documents generated with Microsoft Word are generally accessible and well structured, there are other programs that do not add accessibility features, such as LaTeX. PDFs generated with LaTeX compilers do not have the appropriate accessibility features enabled by default. There are approaches to tagging PDFs with LaTeX packages or newer LaTeX versions, but these are neither easy to use nor well maintained or only in the process of being developed and coming [4].

There is ample research in the area of technology acceptance in general. For example [11] introduced the Unified Theory of Acceptance and Use of Technology (UTAUT) in 2003. However, there is little research in the application of technology acceptance studies in the context of PDF remediation. The authors of [10] analyse the issues arising when introducing accessibility requirements to PDFs in the context of scientific work. They propose to improve the dissemination of accessible documents by the following steps: (i) Lowering the barriers for authors when trying to submit their documents in accessible form. (ii) Reducing friction in the remediation process itself, by removing unnecessary hurdles from the systems used to remediate documents. (iii) Increasing the general awareness of the importance of accessible documents. Another research area of interest is research on the topic of user assistance systems. As the efficacy of simple help pages and “F1 help” functions has been questioned in the past [9] or even been described as distracting [3], we turn to more integrated approaches. For example, [8] present the concept of Advanced User Assistance Systems, classifying three types: intelligent, interactive and anticipating user assistance systems. Similarly Honold et al. [5] present user assistance systems that provide help to users on the basis of a knowledge base that makes it possible for the system to adapt to users with different levels of expertise.

Looking at these approaches we see ample possibilities to improve the usability of remediation programs by adding assistive functionality that goes beyond a simple instruction page.

In the context of accessible PDF documents, there are strict guidelines that define what a PDF must look like to be recognized as accessible [6] or provides a list for checking PDF/UA conformity (https://www.pdfa.org/). The company axes4 offers the PDF Accessibility Checker (PAC - https://pac.pdf-accessibility.org/de) to check compliance with the PDF/UA standard. There is a comprehensive list of elements that must be handled correctly (correct labeling of elements, alternative texts, correct nesting of headings, marking of heading cells in tables, etc.). Thus, we perform a fundamental study with users to analyze the current status and the various steps towards an accessible PDF using the programs mentioned at the beginning in order to gain new insights for new developments and optimizations.

3 User Study

In this section, we describe a user study in order to get an overview over the most pressing issues and biggest pain points that users have when making PDFs accessible.

3.1 Participants

We selected nine participants based on their novice status with PDF adaptation. Cursory experience with the adaptation of digital documents in general wasn't a disqualifying criterion. Prior to the study, some demographic data (i. e. educational background, occupation) was requested and they were asked to self-assess how well they knew the systems on a Likert scale from 1 to 5 (1 = no experience, and 5 = highly experienced). Knowledge ranged for Axes and Pave from 1–3, Pave (8 times 1, 1 time 3) and Axes (7 times 1, 2 times 2), and for Acrobat experience between 3–4 (7 times 3, 2 times 4). However, subjects showed familiarity with using it to improve accessibility. All of them had academic background (2 B.Sc, 4 M.Sc. and 2 students). The participants work either in academia (2), chemical engineering (2), media industry (1) or industrial engineering (1). Table 1 summarizes the participants and their reported experience with the software tested.

3.2 Method and Procedure

The documents used in our study were created with Libre Office and excerpts from Wikipedia, with minor formatting but otherwise simply copied text with figures. No additional effort was put into making the document more accessible other than configuring Libre Office to export a tag-tree, which is not default behaviour.

Subjects were given a short introduction to the topic consisting of a brief and superficial explanation by the study conductor to simulate a first time encounter with the topic. In a first step, the participants tested two (of our three) tools for PDF adaptation on one of two test documents respectively. To avoid any learning effects, the order of tool and document were randomized. It has to be noted, that

Table 1. List of the participants, their demographic data and reported experience with the 3 tools tested. Note that the reported experience with Adobe Acrobat Pro are related to tasks other than accessibility.

Subject	Education	Profession	Acrobat	Axes	Pave
P1	M.Sc	Academia	3	1	1
P2	HS	Student	3	1	1
P3	B.Sc	Student	3	1	1
P4	B.Sc	Academia	3	1	3
P5	M.Sc	Chemical Engineering	4	2	1
P6	M.Sc	Media Industry	3	1	1
P7	M.Sc	Chemical Engineering	3	2	1
P8	M.Sc	Industrial Engineering	4	1	1
P9	HS	Student	3	1	1

this is already an extra step to be taken by any user trying to make an accessible document in the first place, as it's not the default behaviour. However, this step was necessary, to make sure that all three of the tested programs work with the document to begin with. Afterwards, we conducted a think-aloud session, using the method by [1]. Participants were simply asked to adapt the given document with the given tool. They could choose their own approach as well as decide when they were done with the task. No external help such as googling was allowed.

When subjects declared they were done with adapting the PDF with one tool, they were asked to complete a System Usability Score (SUS) to quantify the usability of the tools. Then, a semi-structured interview was conducted to learn about the issues subjects observed when using the tools. Additionally, we collected the subjects insights into which parts of the tools should be improved, explained better, etc. During the interview, we asked the following eight questions, using each question as a starting point for short discussions:

1. Do you receive support from the tool?
2. How satisfying or frustrating is the usage of the tool?
3. How free and autonomous are you when using the tool?
4. Does the tool give you a feeling of competency? If yes, how much?
5. Accessible documents are important, especially for blind people. Does the tool provide an insight into the problems of people with visual impairments? Does the usage of the tool make you empathize with people with blindness?
6. How do you rate the tools in comparison?
7. Which functionalities have to be improved in your opinion?
8. Do you have additional remarks, suggestions, problems, etc?

We recorded the interviews and video-recorded the think-aloud sessions while interacting with the system. After the sessions, we transcribed the recorded material and analyzed the transcripts by collecting the pain points and further comments.

4 Results

In this section, we report the results and our findings. The participants mentioned besides others the following pain points while using the tools. For example Acrobat Pro is a versatile program providing PDF editing and accessibility features, but users struggled with its complexity. First, the necessity to manually initiate accessibility checks caused frustration. Users found it challenging to add missing metadata due to unclear instructions and were unsure about correct authorship or title information. The programs tested lacked intuitive functionality; users could not grasp the effects of their actions or navigate efficiently through problem areas.

Participants were confused by terms like "metadata" and "tags," showing a gap in understanding fundamental concepts and terminology. This lack of knowledge extended to creating alternative texts for images, with confusion over appropriate length and detail.

Moreover, table handling was problematic; users did not understand how to assign headers correctly. Programs provided little guidance or explanations for errors, hindering effective problem-solving. Crucial functionalities were often hidden or absent, such as zooming into documents in Pave or undoing changes in Axes.

Unexpected program behaviors added to user confusion, particularly when actions led to unforeseen consequences or when click-through from error lists to problematic document elements didn't work as expected. Inconsistencies within the tools' feedback systems further complicated the task.

Participants expressed a need for foundational knowledge on topics like tag naming schemes and document structure rules but received inadequate support from the software tools. They also noted operational difficulties with basic interactions like dragging-and-dropping or scrolling through lists.

P4 expressed it as *"I'm stranded, that is: I can't proceed. Yeah, I'm stuck. I'm missing a manual which tools I have to apply and how"*. Although there might have been some solution to the problems, the participants did not know them. The following paragraph details the problems subjects encountered during their think aloud session as well as the problems identified during the interviews.

4.1 Qualitative Findings

Overall, there was a pervasive sense of frustration and exhaustion due to cumbersome processes and an unclear outcome in terms of accessibility compliance. User experiences highlighted that PDF accessibility software tools have significant usability issues, emphasizing the need for more intuitive design and more comprehensive educational resources. Thus, the investigation of the think-alouds and the interviews revealed three main categories of problems, which we identified as follows:

1. Users feel abandoned, i.e. they don't know
   (a) what to do, how to proceed through the task at hand,
   (b) which information is exactly needed and which inputs are required from them,
   (c) what the terms used in an accessibility context mean (e. g. H1, TD, TR, SPAN etc.)
2. Missing/insufficient possibility to track or validate the changes the underlying data, i.e. there is no way to
   (a) undo a change,
   (b) review and verify changes easily,
   (c) access metadata and other data that already exists (in the background).
3. Confusing/non-intuitive interface, i.e. there are
   (a) many buttons with unclear effect,
   (b) different "working areas" scattered across the User Interface (UI) in a non-coherent manner, and
   (c) unexpected changes in the view-port (e.g. interface elements obstruct vision on others, lists scroll unexpectedly).

4.2 SUS

Looking at the System Usability (SU) scores, we can see that none of the tools tested are well received by users. Table 2 shows the SU scores given by the participants per tool used. Analyzing the SU scores, we can see that none of the tools tested are well received by users. Figure 1 shows how the three systems perform regarding their SU scores. Using the different classifications presented by Bangor et al. [2], we can clearly see that none of the tools tested are well liked by novice users except for Pave by single users (P02,P05,P09). Table 2 shows that there is no uniform picture. However, when we look at the mean of the ratings Acrobat achieved a SU score below 30, AXES between 30 and 40, and Pave between 40 and 50.

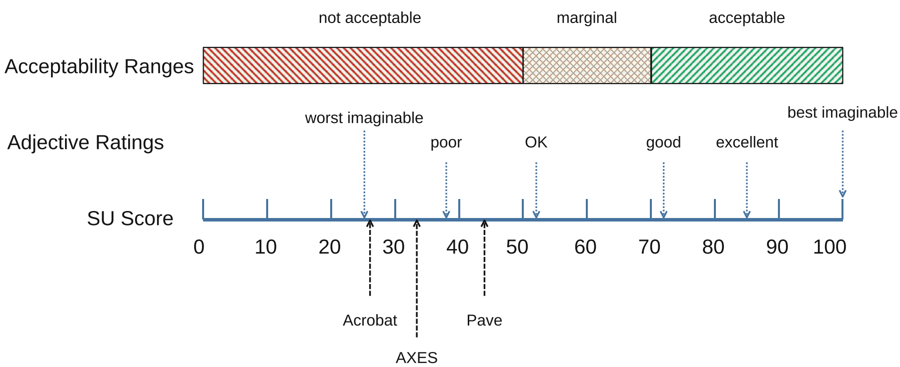


Fig. 1. Acrobat achieved SU score below 30, AXES between 30 and 40 and Pave between 40 and 50. None of them achieved a score over 50.

Table 2. SU scores given to the individual programs by the subjects. An empty cell points out that the subject has not tested the respective tool.

Subject	Acrobat	AXES	Pave
S01	32.5		22.5
S02		30	70
S03	40	52.5	
S04	40	52.5	
S05		17.5	82.5
S06	0	67.5	
S07		15	0
S08		17.5	15
S09	20		60

5 Conclusions

We examined the usability of three tools that make PDF documents accessible with nine participants. Our study suggests that inexperienced users are generally dissatisfied with the PDF customization systems currently available. Various reasons were cited for the difficulty of making PDFs accessible: lack of knowledge about barriers that can occur in PDFs and how to remove them, lack of features in the software and problems with the interface.

We summarize the requirements requested by users as follows. Given the problems identified in the study as well as some related work, a PDF remediation system should fulfill the following conditions:

1. The system must be easy to understand for beginners
2. The system must provide tools that meet all common accessibility requirements
3. The system must guide users through the troubleshooting process
4. The system must provide appropriate help when needed

One of the key findings was that (inexperienced) users need to be guided through the whole process and need a more detailed explanation at certain points. In addition, we realized that this guidance needs to be seamlessly integrated into the system so that the user can directly access explanations and further instructions when needed in order not to interrupt the workflow.

In the future, new tools will be developed to address the above issues by providing guided support, an intuitive interface and trackable changes.

References

1. Alhadreti, O., Mayhew, P.: Rethinking thinking aloud: a comparison of three think-aloud protocols. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, pp. 1–12 (2018)
2. Bangor, A., Kortum, P.T., Miller, J.T.: An empirical evaluation of the system usability scale. Int. J. Human-Comput. Interact. **24**(6), 574–594 (2008)
3. Barrett, L.F., Tugade, M.M., Engle, R.W.: Individual differences in working memory capacity and dual-process theories of the mind. Psychol. Bull. **130**(4), 553 (2004)
4. Frank Mittelbach, U.F.: Latex tagged pdf (Mar 2024). https://www.latex-project.org/publications/indexbytopic/pdf/
5. Honold, F.,et al.: Companion-technology: Towards user- and situation-adaptive functionality of technical systems. In: 2014 International Conference on Intelligent Environments, pp. 378–381 (2014)
6. Document management applications - electronic document file format enhancement for accessibility ISO 14289-1:2014 (2014). https://www.iso.org/standard/64599.html
7. Johnson, D.: Pdf's popularity online - "pdf" versus other digital document formats on the webs (Sept 2021). https://pdfa.org/pdfs-popularity-online/
8. Maedche, A., Morana, S., Schacht, S., Werth, D., Krumeich, J.: Advanced user assistance systems. Bus Inform. Syst. Eng. **58**(5), 367–370 (2016)
9. Sykes, T.: Support structures and their impacts on employee outcomes: a longitudinal field study of an enterprise system implementation. MIS q. 39, 473–495. 10.25300. Tech. rep., MISQ/2015/39.2. 09 (2015)
10. ToBePublished: Marking Up the Wrong Tree - How Our Research Community Should Implement Publication Accessibility. In: Confirmed for the Proceedings of the ACM on Human-Computer Interaction, Issue CSCW (forthcoming)
11. Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D.: User acceptance of information technology: Toward a unified view. MIS quarterly,pp. 425–478 (2003)

Flex Picture eBook Builder - Simplifying the Creation of Accessible eBooks

Danya Gharbieh(✉), Maximilian Punz, Klaus Miesenberger, and Valentín Salinas-Lopez

Institut Integriert Studieren, Altenbergerstr. 69, 4040 Linz, Austria
{danya.gharbieh,maximilian.punz,Klaus.Miesenberger, valentin.salinas_lopez}@jku.at

Abstract. Flex Picture eBooks (FPB) are a format that aims to increase the accessibility of eBooks, by dynamically adapting its contents to the users needs. The project leverages the latest advances in EPUB3 to transform the creation of inclusive digital learning materials. While EPUB3's support for rich media holds immense potential for accessible education, widespread adoption is hindered by a steep learning curve and the time-intensive nature of development. This paper presents FPB's approach to streamlining the authoring of interactive, adaptable EPUB3 publications. FPB empowers the creation of illustrated digital books that dynamically tailor themselves to individual children's needs, fostering equal access to foundational education.

Keywords: Accessible digital publishing · EPUB3 · Accessible digital illustrations

1 Introduction

This paper introduces a novel approach to enhance the accessibility and usability of digital publication through adaptable non-text elements. The EU-funded Flex Picture eBook (FPB) project strives to revolutionize the adaptability of digital publications, particularly illustrated children's books. By testing a prototype publication with children with disabilities, their care providers, and educators, the FPB project aim to improve inclusion in digital publishing.

We present a software suite designed streamlined authoring. It offers innovative tools addressing two core areas:

- Accessible Illustration Creation: Assisting authors in generating illustrations that are inherently designed to accommodate a range of comprehension levels.
- EPUB3 Integration: Providing mechanisms to seamlessly incorporate complex and layered illustrations into compliant EPUB3 publications, ensuring compatibility with common consumer electronic devices and assistive technologies.

By addressing these distinct yet interconnected challenges, the FPB project paves the way for mainstreaming accessibility in the publishing sector.

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 63–71, 2024.
https://doi.org/10.1007/978-3-031-62846-7_8

2 State of the Art

Many word processing tools such as *Google Docs* [3] now support exporting text files to the EPUB format. However, this only allows users to convert static files, and excludes the possibility of handling user input, thus making it unsuited for the creation of Flex Picture eBooks.

Currently, these kinds of eBooks have to be made by hand, which poses a lot of difficulties that could potentially deter authors and publishers from distributing their work in this new accessible format. The main problem is the technical nature of the EPUB file type [1]. Under the hood, they are specially formatted ZIP archive files format, containing Extensible HyperText Markup Language (XHTML) pages. All of the text has to be manually input into these XHTML files and styled through CSS, while the interactive functionalities have to be coded in JavaScript. In some way, each page can be thought of as an individual website that has to be manually programmed. This poses a serious technical challenge for the publishers, the majority of which do not possess a background in programming.

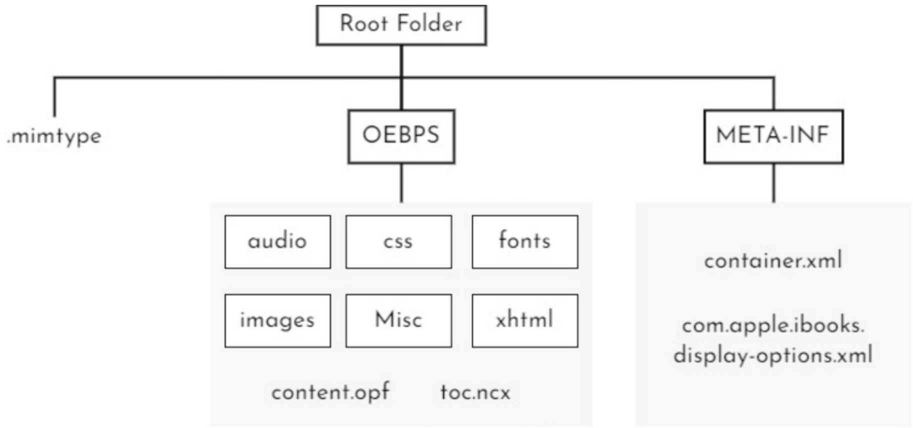


Fig. 1. The file structure of the EPUB Format. The mimetype tells the e-Reader that it is a digital publication, while META INF contains files which point towards the root documents. The OEBPS directory contains the folders in which all of the files inside the book are stored (sorted into the corresponding subfolders).

The EPUB format also has very strict guidelines that have to be fulfilled. If even a single condition is disregarded, or a parameter is altered in an unexpected way, the EPUB cannot be displayed on an otherwise compatible screen reader [8]. This brittleness spans multiple dimensions, such as the way the files are arranged inside the containers. There is a required folder structure that dictates what kind of files are put into which directory (see Fig. 1). All of these file structures have to be created, and even subtle errors such as misspelling a folders name can be detrimental. The format also requires a couple of different files that point the e-Reader to the location of the pages are and help it navigate its contents. These files also have to be written in very particular ways, requiring a number of script tags and specifications to be working properly. Integrated Development Environments (IDEs) and Word Processors, in which these files are usually created, do not tell the user which tags are still missing or what options are allowed, which makes working on these files very opaque and can often times lead to great frustration.

There is specialized software for testing EPUBs, such as the *EPUB Checker* [4], that allows the user to validate their files. Their purpose is to draw attention to the flaws present in the documents and inform them of specific problems. This makes the creation process slightly more transparent, giving the users a list of errors they can fix. But, to make the files compliant to the standard, using a validator still requires knowledge about technical details and can be very time consuming.

Software such as *Sigil* [6], helps navigate this maze somewhat, letting users generate the directory structures and required files. But, using them still is fairly demanding on a technical level, making it inaccessible for people who have no experience working with web technologies. Because, like with the examples of other text processors, the output of these files is static, and the additional Javascript code needed for the creation of Flex Picture eBooks has to be implemented manually. Our proposed software should lower the barrier of entry that comes with the creation of Flex eBooks, by automatically allowing users to generate compliant EPUBs that include the Flex-eBook functionality, with the user having to interact with the coding as little as possible, while giving them access to the full range of functionalities that come with the format.

3 Methodology

In the section below, the methodology behind creating the Flex Picture eBook Builder will be explained. The creation-process was split into two separate, but interconnected parts: The user interface and the system that creates the EPUBS themselves.

3.1 Building an Accessible User Interface

Creating an accessible user interface (UI) is crucial for ensuring that all users, including those with disabilities, can interact with digital products effectively and comfortably. To make the software available to as many people as possible, we decided to use Electron [2] as the development framework, because it allows for the creation of platform independent applications. In the section below, detailed explanations of all of the screens that the user is presented with can be found.

Project. This is the starting point of a FPB publishing journey. Here, the user has the opportunity to manage existing projects or begin new ones. (see Fig. 2)

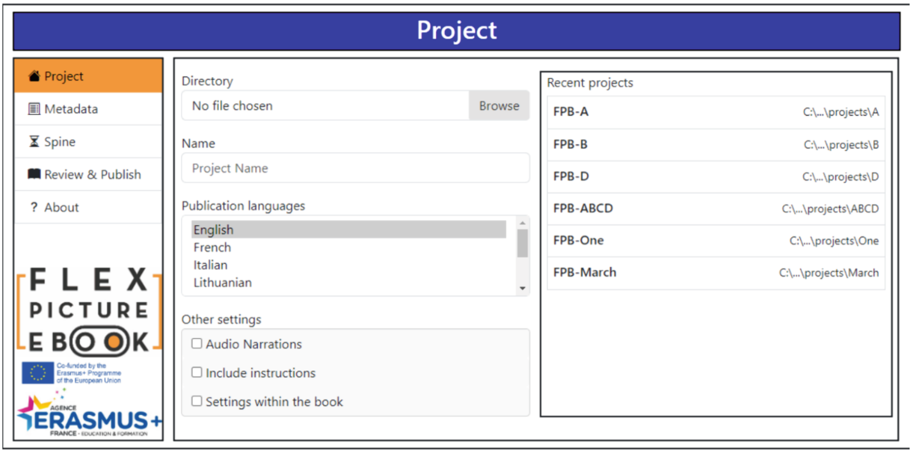


Fig. 2. Project Screen

- Manage Existing Projects: utilize the 'Recent Projects' section located on the right-hand side of the interface. This feature provides convenient access to previous projects.
- Start a New Project: utilize the main space for:
  1. Directory: browse to select the directory to store the project files.
  2. Name: provide a descriptive name for the project.
  3. Publication languages: Specify the languages in which the EPUBs will be created.
  4. Other settings: customize settings for the book, whether it is going to include options like audio narrations, instructional content, and if the settings will be included.

Metadata. Here, the user provides the essential information that describe their FPB. (see Fig. 3)

- Add metadata: choose a type of metadata from 'Available Metadata' on the left, then click "Add" to see the required fields on 'Added Metadata' on the right side.
- Remove metadata: choose item from the 'Added Metadata' list on the right, then click "Remove" to return the fields back to available metadata list on the left side.

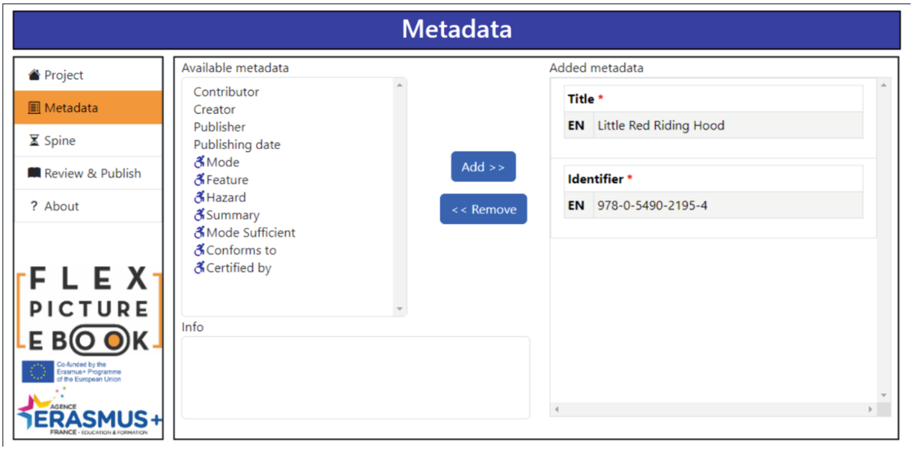


Fig. 3. Metadata Screen

- Added Metadata language-specific types: some types of metadata that need multilingual input (e.g. the title) will have editable table spaces for each of the previously selected publishing languages.
- Required types: The application highlights the required types and prompt the end user to complete them before proceeding.
- Info section: shows information about the selected type of metadata and how to set it properly.

Spine. In this section, all content elements can be added and arrange the pages of the FPB. (see Fig. 4)

- Add and order pages within the publication, excluding the cover (first page) and credit (last page).
- Content section for each page:
  1. Text: enter the page text content in the editable text table for each of the publishing languages.
  2. Narrations: upload audio narration files for each of the publishing languages.
  3. Images and Scripts:
     (a) Image: Browse to select the XHTML file containing the adapted illustration (SVG).
     (b) Style: The application will identify external linked resources such as CSS and JavaScript files and prompt the user to specify their locations for proper integration and rendering within the project.
     (c) Verify linked files and avoid unnecessary duplication of resources used across multiple pages.

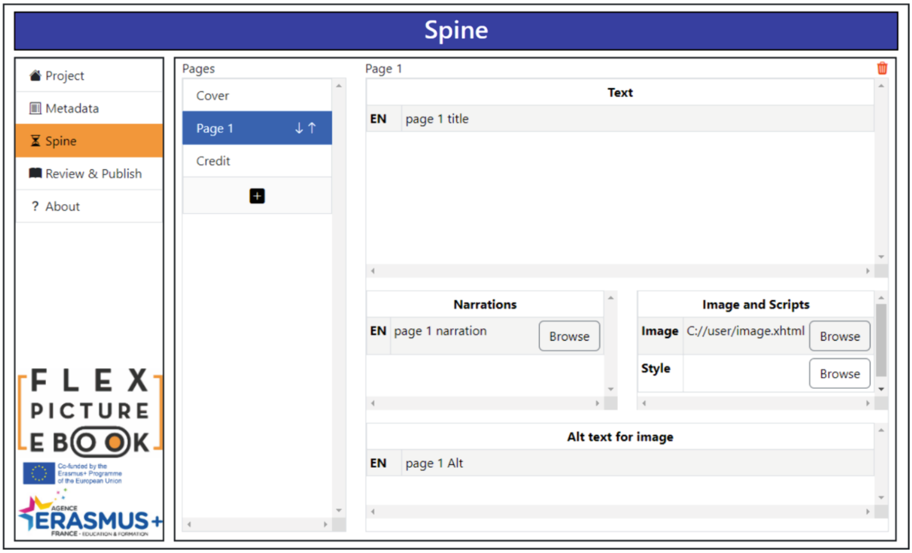


Fig. 4. Spine Screen

4. Alternative text: enter the image alt text in the editable table for each of the publishing languages.
5. Remove: Delete the entire selected page, excluding the cover (first page) and credit (last page).

Review and Publish. In the final section, the FPBs are finalized and prepared for distribution. (see Fig. 5)

– File name: Set the publication file name, By default, it will contain the same value entered in the 'Name' field on the project screen.
– Launch published E-Book: Tick the box to run the eBook immediately (using Thorium).
– Generate: Create the intermediate file structure and the metadata files. All generated content will be placed in a set of subfolders (one per publishing language) in the project root directory. This allows the user to examine and modify the files that will build up the FPBs if needed.
– Review: Validate the generated file structure against EPUB3 standards using EPUB checker tool.
– Publish: Outputs an EPUB from the generated files for each publishing language that is ready for distribution.

3.2 File Management and Automated Flex-EBook Creation

One of the more interesting challenges that we ran into was automatically importing files that are required for the interactivity to work. The images are

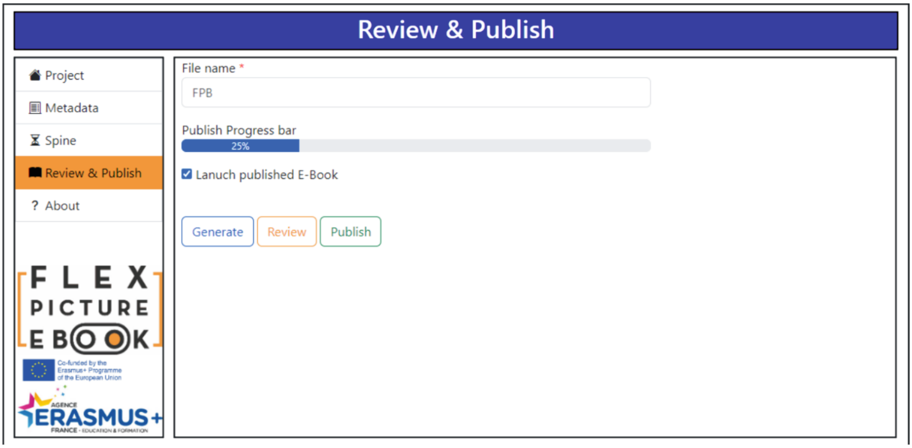


Fig. 5. Review and Publish Screen

saved as SVGs inside XHTML files, and often need additional Javascript code, or styling and formatting from CSS, to be displayed correctly. To make the process as uncomplicated for the user as possible, we decided to automatically import these files into our program. To do this, the XHTML document is scanned line by line for import statements. Using string manipulation, the exact places where these statements point to are then determined. Most of the times these statements specify a location relative to that of the selected file, which means that the absolute position in the file system has to be determined, to import the dependency. After a quick check, the paths are added to the list of documents that is packaged into the EPUB. If multiple files point to the same dependency (meaning that a file has been reused in multiple XHTMLs), the system only imports it once. In case a dependency was not found, the user is prompted to manually select it.

Before the EPUB is created, all of these documents are copied into their corresponding folders. To make everything work properly, the import statements of the XHTMLs need to point to these new locations, which our system solves by changing the old paths, with the new relative ones. Because every EPUB has the same file structure, the creation of the relative paths can be standardized.

There is a number of files that our system generates, such as the menu that is shown at the start of the EPUB, which are dynamically generated from templates. These templates are then filled according to metadata specified in the front end. This allows us to quickly create FPBs in multiple pre-specified languages. One of the issues that are still at play, is the fact that for every language the template has to be manually translated, which is a huge amount of work for us. While most files are language independent, something like the settings menus still requires the names of the buttons to be translated. Most other templates are generated completely dynamically, but the menus are (mostly) pre-written XHTML files, where some content is changed according to the specifications of the user. This is due to the comparatively large size of the menu files, which made this approach more efficient during production. However, this can be a bit of a hurdle in terms of translation, because the text content is entangled in the HTML code. One solution that will be explored, is creating the menus from language agnostic templates and reading the text from files that only contain the names of the items, thus allowing these files to be translated by an automatic translation service like DeepL [5]. These would still have to be checked by natives of that language, but could cut down on a large amount of time and potential costs.

The system is built in such a way, that if every required field was filled in, then the program exports a working EPUB, that can get validated without issues. The output of the program contains two parts: The EPUB and a folder containing all of the raw files, allowing the user to browse them and make sure they contain the right information.
As stated before, the EPUB itself is a ZIP containing all of these documents, with the file extension changed to the *.epub* ending. However, there are some more things that could potentially trip up our program: First, the format requires the ZIP to be in a particular order, namely that the .mimetype file (which tells the eReader that the folder in question is indeed an EPUB), is at the first position. Secondly, the compression inherent in a ZIP File cannot be applied equally on all documents, specifically the .mimetype file has to be left uncompressed and unencrypted [7]. Ordering the elements inside of a ZIP, or compressing elements non-uniformly can be quite challenging without specialized software, so we made sure that the rendering process was as automated as possible.

4 Future Work

Our software should lessen the workload and required technical knowledge to create Flex eBooks and lower the barrier of entry to such a degree that the format becomes more widespread. One of the ways this process could be further simplified, is through the creation of a tool that allows illustrators to export directly to the required format and simplify the process of manually abstracting images. In the next steps, this tool will be implemented as a plugin for Adobe Illustrator, which hopefully makes the task of image-creation easier and more accessible.

Acknowledgments. Flex Picture eBook has received funding from the European Commission's and the Agency Erasmus+ France through the action KA220-SCH Cooperation partnerships in school education. Grant Agreement No.: 2022-1-FR01-KA220-SCH-000088072.

References

1. Daisy Consortium: Epub (2024). https://daisy.org/activities/standards/epub/
2. Electron: Electon framework (2024). https://www.electronjs.org/docs/latest/

3. Google: Google docs (2024). https://docs.google.com/
4. pagina: Epub-checker (2024). https://pagina.gmbh/startseite/leistungen/publishing-softwareloesungen/epub-checker/
5. SE, D.: Deepl (2024). https://www.deepl.com/de/translator
6. Sigil: Sigil - epub editor (2024). https://sigil-ebook.com/
7. W3C: Epub zip file requirements (2023). https://www.w3.org/TR/epub-33/#sec-zip-container-zipreqs
8. W3C EPUB 3 Working Group: Epub 3.3 (2023). https://www.w3.org/TR/epub-33

Tools for Novice and Expert Accessibility Professionals: Requirements for the Next Generation Web Accessibility Evaluation Tools List

Vera Lange, Charlotte Vera Swart, Jasper Jeurens, and Eric Martin Velleman(✉)

HAN University of Applied Sciences, Ruitenberglaan 26, 6826 CC Arnhem, The Netherlands
Eric.Velleman@han.nl

Abstract. The W3C Web Accessibility Evaluation Tools list offers one of the most comprehensive and most referenced lists for accessibility evaluation tools today. The tools in the list can help check or monitor various accessibility issues on websites and apps. This paper describes the re-design of the list to make it more future proof and accessible for a large variety of stakeholders, both accessibility experts, novices and others to find the right tool for their situation and needs. The re-design included defining the audience, information needs, user-stories, design and layout, comprehensive user and technology-oriented filter options, decision support for users, etc. The researchers propose new functionality to guide users to relevant accessibility tools, such as a step-by-step search assistant, new updated filter and search functions, and adding relevant and understandable information for each tool. The end result provides both experienced and novice tool users with options to filter and search for the tool that suits their needs.

Keywords: Accessibility · Tools · Evaluation · Conformance · W3C · web design

1 Introduction

Web accessibility evaluation tools can help identify potential accessibility issues on webpages, apps and in digital documents [1]. There is a wide range of tools available [2, 3]. For example, some tools can check for specific issues, standards versions or technologies, some provide more in-depth feedback that can be used for training, others can be used for deployment or to simulate the experience of a specific user group with a disability. Tools can support manual review or fully automated checks.

The W3C offers the Web Accessibility Evaluation Tools List [2] that includes more than 160 tools and is one of the most comprehensive and most referenced lists for evaluation tools today. It offers an overview of available tools together with some basic information and filtering options to help users find a tool. However, it was showing signs of age. The contents had been maintained regularly, but the user interface had not been revised since 2014 (see Fig. 1).

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 72–83, 2024.
https://doi.org/10.1007/978-3-031-62846-7_9

The Tools list needed a re-design to better satisfy the information needs of today's tool users and to support both experienced and novice users with a basic understanding of web accessibility evaluation. This provided an interesting challenge between the amount of (technical) information available and not overwhelming (new) users. Nielsen describes this balancing act in his 7th heuristic of usability [4], which states that systems should cater to both expert and novice users through flexibility and efficiency of use.

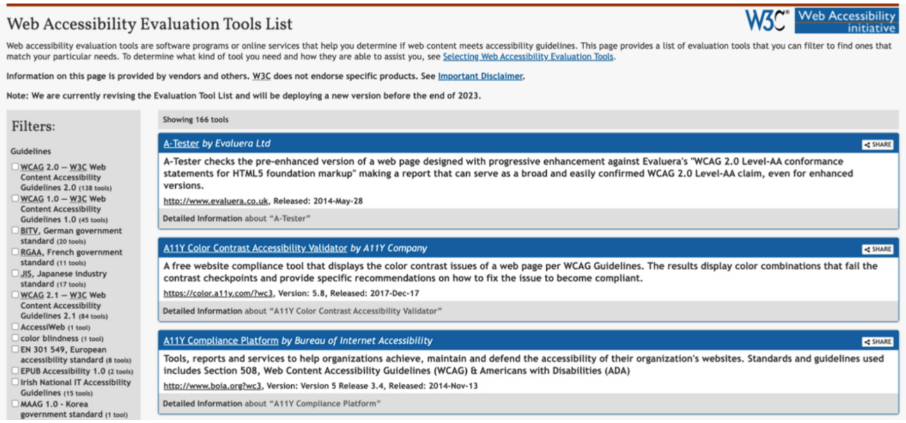


Fig. 1. Screenshot of the old W3C WAI Web Accessibility Evaluation Tools List (31/01/2024)

This paper describes the process of re-designing the accessibility evaluation tools list resource for both expert and new users in the accessibility field and presents the final design results.

2 State of the Art

The adoption of the Web Accessibility Guidelines for public websites and mobile applications has been a priority for many countries for decades [5, 6]. However, European studies show a lag in accessibility implementation due to a variety of reasons including the limited or bad use of tools [7–9]. In an evaluation of Monitoring Reports by Member States, the authors conclude that cross-border provision of accessibility tools is available and can be expected to increase [10]. But an earlier report shows the use of tools for accessibility monitoring is still very limited [11]. The report also shows that tools are not always used correctly. None of the automatic tools that are in the market now are capable of checking all success criteria [1, 12]. Most of them only check for partial aspects of the criteria. This means that automatic tools cannot be used as the only way of monitoring accessibility. They can however support users in finding possible accessibility barriers.

In Europe, monitoring is a legal requirement for Member States [13] and the Accessibility Act requires conformance statements for products and services including how organizations organize and ensure continuous support for accessibility [14]. Finding and using the right tools can support any person working on accessibility in any part of the process. Research suggests a relation between the availability of tools in an organization and the actual implementation level of the guidelines in a website [15].

2.1 Tools Lists

But finding tools and selecting those useful for your specific needs is not as easy as it sounds. Users in our study shared they don't use the W3C Tools list, but instead find tools through co-workers' advice or online curated lists composed by experts [16, 17]. Users state that they prefer having experts select tools for them that are reliable and user-friendly.

Accessibility evaluation tools are also subject of scientific papers, applying them to websites, describing, analyzing and comparing them [18–21]. These publications may support users and give insight into the current state and possibilities of (specific) accessibility evaluation tools.

Another common way of finding tools is by using a search engine. A Google search reveals no similar comprehensive accessibility tools lists that allow users to filter results the way W3C offers, but there are many curated lists of *Top 20 tools* or *Best 5 tools*, sometimes grouped by type or goal. Some lists are composed by tool vendors, promoting their tools alongside alternatives. It is not always clear who is behind the lists, if the tools were tested and if they are biased towards a paid tool.

Many websites that contain large lists of information use search bars and filters to help users find what they need. We studied sites such as large online stores to find design patterns [22, 23]. Large online stores like Amazon and Bol often offer contextual filters that change depending on the product category. Sometimes filters are very technical or detailed, this may create a challenge for people new to the subject. Websites often help these visitors by adding (tooltip) explanations for filters or by offering more in-depth guidance in selecting these filters such as a filter wizard.

These design patterns are currently not used in the designs of accessibility tools lists. Most of these lists contain only a limited number of tools. However, with its extensive number of tools and diverse user groups, the WAI accessibility evaluation tools list would greatly benefit from employing these patterns.

2.2 User Groups and Other Stakeholders

Evaluation tools are useful for a variety of users. Web developers, designers and content authors use tools to evaluate the accessibility of their product during design and development. Web content auditors and QA testers use tools to audit existing websites and web content for accessibility guidelines conformance. And content procurers often search for tools to check accessibility during the procurement process. In coordination with the EOWG, this set of roles was defined as the primary user group for the re-design of the Tools list. The secondary user group includes product managers, policy and decision makers, legal staff, researchers and educators.

Important stakeholder in this project [24] was the W3C Web Accessibility Initiative (WAI) Education and Outreach Working Group (EOWG) [25]. They develop strategies and resources to promote awareness, understanding, and implementation of web accessibility. The Tools list is one of the many resources that EOWG is responsible for. During the design process, the design team coordinated with them to ensure the design iterations stayed within the brief and requirements.

3 Methodology

3.1 Design Thinking Approach

At the start of the design process, there were many possible design directions, but we had little indication of the needs of the different user groups, so an iterative approach was chosen with intermittent testing to involve the potential users and other stakeholders in the redesign process of the tools list. Design thinking prescribes such an approach and is a method often used in projects with vaguely defined goals and contexts. It structures the design iterations in several phases or mindsets: Empathize, Define, Ideate, Prototype, and Test [26].

We started by conducting contextual inquiries to define the requirements. The design was iterated multiple times by prototyping and testing finally leading to a complete, accessible draft website.

The way we involved users was done in accordance with the HAN Research Ethics Committee.

3.2 Working with the W3C Education and Outreach Working Group

The design iterations and research results were presented to the EOWG regularly and surveys were sent out to all members to discuss the findings and gather input on the design and different key aspects. The EOWG currently has 49 participants representing 32 organizations. They include persons with disabilities. The EOWG's work is based on a consensus model - a process in which every member of the working group can add their objection or reservation on an aspect of the design, which then needs to be discussed among all members until a consensus is reached.

Separate discussions were held with members of the EOWG who were knowledgeable about a specific aspect of the tools list or accessibility tools. For example, we spoke with two members who are also tool vendors. They helped us to further refine the filters, making sure they matched the potential features and properties of tools.

3.3 Contextual Inquiries

Contextual inquiries involve aspects of both interviews and observations, striking a balance between the participants being in the lead and leaving room for researchers to step in at key moments to ask for clarification or steer in a particular direction [27]. First, we asked users to talk about their experience with accessibility, evaluation, and tools in their everyday work. Next, we asked users to demonstrate how they search for tools for their own accessibility evaluation goals. Finally, we asked participants to visit the Tools list and evaluate it based on their information and functionality needs.

We conducted online contextual inquiries with 8 users, including current and potential users of the tools list and tool users as well as vendors. When selecting participants, we made sure to recruit users that reflected the diversity of roles within the target user group. We spoke with two junior accessibility advisors, a web editor for government websites, an accessibility lead, one UX lecturer, one UX & accessibility lecturer, and a product manager responsible for accessibility at their company.

3.4 Data Clustering

Another source of insights about the old version of the tool list was the GitHub repository of the site. Here we could find comments and issues about specific elements and use them in further analysis.

Notes collected from the inquiries and the repository were analyzed by clustering using a data-driven approach. Each observation formed the starting point of a potential group, to which similar or related data points were then added. From this, the number of occurrences was counted. Patterns were identified through discussion of the grouped data. Groups were titled as much as possible from the perspective of users (e.g. *"How do I know which tool is good quality?").*

3.5 Requirements

The data from the contextual inquiries provided insights into the user needs and issues with the current tools page. From there we defined requirements for the re-design. Requirements were prioritized as either "necessary" or "suggested". This distinction was made based on the frequency and severity of user comments and observations during use of the tool list, and requirements set by the W3C. Necessary requirements are essential to the usability and effectiveness of the tool list. Requirements marked as "suggested" are valuable for the design process, but are not absolutely required for a well-functioning list, are in conflict with requirements set by the W3C, or were observed in a small number of interviews and thus more research would be needed to estimate their importance. The user needs and requirements are discussed in the results section.

3.6 Design Iterations and Testing

Based on the list of requirements, the research team built upon existing design patterns to generate ideas. We aimed to prototype and test quickly using hallway tests with researchers, software developers and UX professionals at first. These spontaneous and small-scale tests provided the team with quick and critical feedback while the design was in its early stages.

In later iterations, we included expert reviews by consulting the EOWG through surveys and discussion groups. And toward the end of the design process, we performed an in-depth usability test.

3.7 In-Depth Usability Test

Six people participated in a usability test of a mid-fidelity prototype of the tools list. Two design lecturers, two web designers, one accessibility tester and one accessibility manager. Some are experts on web accessibility and others are new to the field.

The goal of the test was to evaluate the usability of the design. We used a thinking aloud protocol [28] to prompt participants to actively voice their thoughts. We asked questions for clarification and to get a better idea of the participants' experience.

The test was structured through multiple tasks. While getting a first look at the page, we asked for the user's first impression and who they think the target user group might be. Then, participants were asked to find a tool based on an accessibility evaluation use case from their daily work. This open task allowed us to observe how users would navigate the page without guidance. Two other tasks involved users finding tools for specific goals. If testers didn't use the Filter Assistant, they were instructed to in an additional task.

Afterwards, we asked questions such as: "Looking at the list of filters, are there any you aren't sure what they mean?", and "What would you like to see differently?".

4 Results

During the research and design process we identified relevant user needs, defined the main requirements, explored potential design solutions and developed the final design.

4.1 User Needs

We noticed a great variety of goals and information needs of the different user groups. For example, some users required additional guidance or information: *"Finding the tools is challenging, knowing what the right criteria are to decide which tool to use is difficult, so you will fail a lot"*. While other users commented on a lack of being able to use the list effectively: *[while attempting to search the page using the ctrl-f shortcut and getting 34 results] "I don't have time for this. People will drop out at this point."*

Three key user needs were defined after evaluating the previous design with users:

1. User need: present me less technical, more user goal-oriented information

   While the filters previously focused on the technical side of tools, the redesigned list should support the intuitive search process of tool users, which is often goal-oriented (e.g. *"I need a tool that can test my website's contrast.").* Users found it difficult to relate to and understand the technical descriptions of tool properties in both the filters and the tool descriptions. The redesigned list needed to lower this barrier of entry and speak the users' language.
2. User need: support me in making the right choice for a tool

   Selecting from a wide range of available tools is seen as ineffective and overwhelming: *"It becomes a big bucket filled with tools and if you're lucky you'll find something."* Instead, users look for outside guidance in making the right choice. This can be in the form of colleagues and experts: *"go to someone who's already involved, who will advise you the tools that work well"*. Or looking for evidence in numbers by choosing the most popular and widely used tools: *"if [the tool] is from a 3rd party I'd say skip it. […] use the big, well-known [tools]"*.
3. User need: let me easily scan the list for relevant information

   Providing a comprehensive list of tools, each with their own technical and diverse features, can overwhelm users. The layout and structure of the list needed to enable all users to easily scan for what's relevant for them, including expert users seeking specific, in-depth information.

4.2 Requirements

Overall, we defined 41 *necessary* requirements and 13 additional *suggested* requirements. The requirements were categorized in 6 themes, as shown in Table 1. The full list of all requirements can be found on the W3C website [29].

Table 1. Requirement categories and example requirements

Category	Example
Finding a tool	The list should allow filtering based on user goals
Composing the tools list	It should be clear how the list is filled and regulated, so users can more easily estimate the quality of the entries
Information needs	The information provided per tool should at least contain: use case, features, date of last update, release date, author/company
Filters	Filters should have plain titles and if necessary, a clear description
Design & layout	The layout of content should aid readability and overview
Adding & updating tools	The site should provide concise and clear information about the submission & validation process

4.3 Design Explorations

Different filtering and sorting solutions were designed and tested to accommodate different users and display the results.

Filter Assistant. Primarily, the filter assistant (Fig. 2) guides new or less experienced users through the list of filters. By answering a few questions regarding their accessibility goal, technology, standard or use case, the system automatically selects the related filter. While the assistant offers valuable guidance for new users, the added steps could be considered cumbersome for more experienced users. Thus, we felt the interaction should not be prescriptive, and instead leave control to individual users to opt in or out whenever needed. Therefore, users are free to skip questions or exit the assistant at any point, returning to the tools list with the filters still selected.

With input from the EOWG, the design team iterated on the order of questions presented in the assistant, creating a logical and contextual sequence in which some answers cause related questions to be added or removed from the queue (e.g. users looking for browser plugins are asked which browser it should be compatible with).

For additional guidance, the assistant also provides descriptions of filters. For example, it explains the difference between automated and manual testing tools. Both new and experienced users found the assistant valuable, stating *“the assistant helped [them] to decide what kind of tool [they] needed.”* and it can *“help understand the filters.”*

Filtering by Role or Goal. Since many users stated that they would like to find tools relevant to their use case/goal and role, we designed the option for users to select their

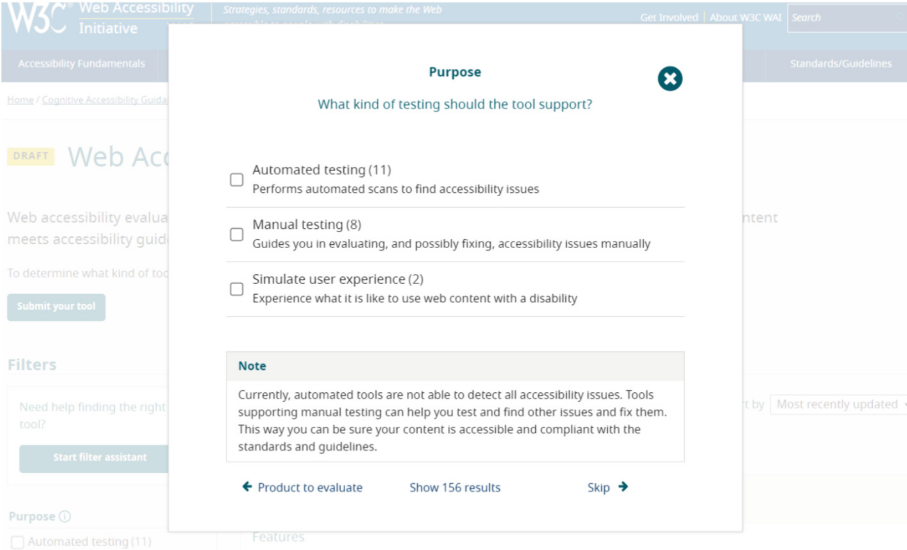


Fig. 2. Screenshot of the second question of the filter assistant (Beta)

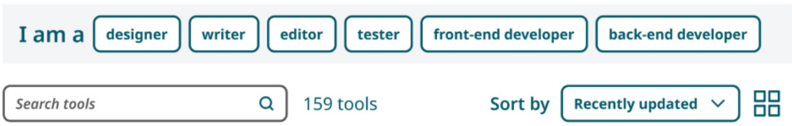


Fig. 3. Screenshot of the top of the page, showing a selection of different roles to choose from.

professional role to quickly find a selection of tools relevant to them. This was prototyped as a row of buttons at the top of the list where users could select their role, such as *I am a designer* (Fig. 3). This would automatically present the tools that fit the general needs of this role.

During user testing participants felt reluctant to use this feature. They didn't want the system to determine which tools were relevant to them, out of fear of missing other relevant tools. Instead, they chose their own way of finding a tool by filtering and searching. Therefore, we decided to not implement this feature.

Selection and Wording of Filters. Since the filters are the primary way users can find a relevant tool, they need to be clear, comprehensive and quick to understand and use. We went through multiple versions of the filter list, re-arranging filters into categories, changing their wording, removing obsolete filters and adding necessary ones. For example, filters previously categorized as *Automatically checks…* are now found under the title *Scope*. New categories of filters include *Product evaluated* (e.g. Website, Document, Source code), *Purpose* (e.g. Manual testing) and *ACT Implementation*. The discussion about the filters is publicly available in the minutes of EOWG.

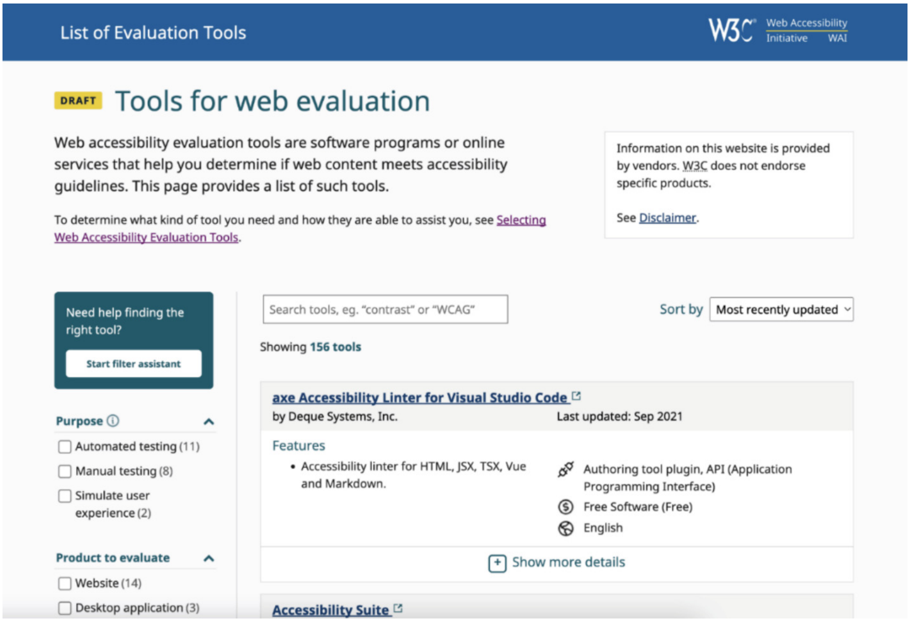


Fig. 4. Screenshot of the re-designed draft of the Web Accessibility Evaluation Tools List.

4.4 Final Design

The final design (Fig. 4) addresses the need to serve both novice and expert users from different professional domains and for multiple uses and technologies. The new filter system forms the main functionality for users to find appropriate tools. This includes the possibility to search by keywords, thereby *custom-filtering* the results in real-time. Finally, the *filter assistant* provides further guidance for new users.

Filter Options. The filter options are designed to allow filtering based on user goal, technology, standards e.g. to check contrast and to filter on key features for tools. Some filters are collapsed, because tests showed that only a few experienced users might need them, while others are expanded for easy access. The filters themselves also inform and inspire new users by exposing them to various topics and practical applications of accessibility. Accessible tooltips were added for those filters worded in more technical or professional jargon.

Tools Description. Tool vendors submit information about their tool to the list, which is evaluated by the EOWG before publishing. This process uses Github for transparency meaning that all input can be followed online. Many tool users do not understand the technical terms vendors use to describe tools. The new tools list bridges that gap and presents tools in a clear, easy-to-scan format and provides information that is relevant for both novice and expert tool users. To match these user needs, tools show key features and a selection of details, such as the type of tool and whether it is paid or free. Additional details are collapsed by default, so the information is available, but it won't overwhelm new users.

5 Conclusion

Our design process led to a redesign that is more future-proof and more accessible to a wider audience. It enables tool vendors to present their products in an informative, less technical way, while providing both experienced and novice tool users with options to filter and search for the tool that suits their needs. Based on the re-design, other resources were also adapted like the pages with more information about selecting tools and the mechanism for tool providers to add and update tools.

6 Limitations and Discussion

Not all features that proved useful during testing made it into the published version of the re-designed list. Some user needs conflicted with the requirements of the W3C. For example, users would prefer a curated list by experts or a possibility to sort by most used or highest rated tools, so they don't have to compare or try different tools. This was not compatible with the W3C aim to maintain a neutral position, offering a comprehensive overview, without endorsing specific tools.

Originally, we had planned a quantitative test of the site when sending out invitations to tool vendors to submit their tool into the updated list. Unfortunately, this test had to be postponed because of a delay in publishing the list. We believe quantitative tests should be conducted to evaluate the usability of the list in its published state.

Acknowledgments. The re-design was part of the Web Accessibility Initiative - Communities of Practice (WAI-CooP) Project, a European Commission (EC) co-funded project, Horizon 2020 Program (101004794).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. White, K., Velleman, E., Hansma, M., Lange, V.: Selecting Web Accessibility Evaluation Tools. https://www.w3.org/WAI/test-evaluate/tools/selecting/. Accessed 08 Apr 2024
2. W3C/WAI: Web Accessibility Evaluation Tools List. https://www.w3.org/WAI/test-evaluate/tools/list/. Accessed 10 Apr 2024
3. Hadadi, S., Sarsenbayeva, Z., Kay, J.: Analysis of W3C's 167 accessibility evaluation tools for the design phase. In: Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility, pp. 1–7. ACM, New York (2023). https://doi.org/10.1145/3597638.3614474
4. Nielsen, J.: Usability heuristics for user interface design. https://www.nngroup.com/articles/ten-usability-heuristics/. Accessed 10 Apr 2024
5. Giannoumis, G.A.: Transnational convergence of public procurement policy. Int. Rev. Law Comput. Technol. **29**, 183–206 (2015). https://doi.org/10.1080/13600869.2015.1055662
6. Laurin, S., Cederbom, A., Martinez-Usero, J., Kubitschke, L., Wynne, R., Cullen, K.: Measures to Improve Accessibility of Public Websites in Europe. NDA, Centre for Excellence in Universal Design, Dublin (2014)

7. Velleman, E.M., Nahuis, I., van der Geest, T.: Factors explaining adoption and implementation processes for web accessibility standards within eGovernment systems and organizations. Univ. Access Inf. Soc. **16**, 173–190 (2017). https://doi.org/10.1007/s10209-015-0449-5
8. Bianchini, D., et al.: Study supporting the review of the application of the Web Accessibility Directive (WAD) VIGIE 2020-0656 Executive summary in English (2022)
9. te Velde, R., et al.: Accessibility of websites and mobile apps (2017)
10. te Velde, R., van der Linden, N., van der Spiegel, R., Engbers, W., le Clercq, C., Enzerink, S.: Study supporting the review of the application of the Web Accessibility Directive (WAD). (European Commission CNECT: Directorate-General for Communications Networks), Brussels (2022). https://doi.org/10.2759/034644
11. Laurin, S., et al.: Monitoring methodologies for web accessibility in the European Union. European Union, Brussels (2016). https://doi.org/10.2759/483693
12. Vigo, M., Brown, J., Conway, V.: Benchmarking Web Accessibility Evaluation Tools: Measuring the Harm of Sole Reliance on Automated Tests. https://doi.org/10.1145/2461121.2461124
13. European Parliament and the Council: Directive (EU) 2016/2102 of the European Parliament and of the Council on the accessibility of the websites and mobile applications of public sector bodies (2016). https://eur-lex.europa.eu/eli/dir/2016/2102/oj
14. European Parliament, Council of the European Union: Directive (EU) 2019/882 of the European Parliament and of the Council of 17 April 2019 on the accessibility requirements for products and services (Text with EEA relevance) (2019). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32019L0882
15. Velleman, E.M.: The implementation of web accessibility standards by Dutch municipalities: Factors of resistance and support (2018)
16. Tools to test for accessibility and quality (Website). https://iacobien.nl/artikelen/tools/. Accessed 06 Apr 2024
17. Baker, K.: 22 Web Accessibility Testing Tools & What They Test For (Website). https://blog.hubspot.com/website/web-accessibility-testing-tools. Accessed 10 Apr 2024
18. Padure, M., Pribeanu, C.: Comparing six free accessibility evaluation tools. Inform. Econ. (2020). https://doi.org/10.24818/issn14531305/24.1.2020.02
19. Timbi-Sisalima, C., Amor, C.I.M., Otón, S., Hilera, J.R., Aguado-Delgado, J.: Comparative analysis of online web accessibility evaluation tools. In: Proceedings of the International Conference on Information Systems Development (ISD). aisel.aisnet.org (2016)
20. Kumar, S., Dv, J., Biswas, P.: Comparing ten WCAG tools for accessibility evaluation of websites. Technol. Disabil. (2021). https://doi.org/10.3233/TAD-210329
21. Ismailova, R., Inal, Y.: Comparison of online accessibility evaluation tools: an analysis of tool effectiveness. IEEE Access **10**, 58233–58239 (2022). https://doi.org/10.1109/ACCESS.2022.3179375
22. Moran, K.: Defining Helpful Filter Categories and Values for Better UX. https://www.nngroup.com/articles/filter-categories-values/. Accessed 10 Apr 2024
23. Holst, C.: Display “Applied Filters” in an Overview. https://baymard.com/blog/how-to-design-applied-filters. Accessed 10 Apr 2024
24. WAI-CooP project. https://www.w3.org/WAI/about/projects/wai-coop/. Accessed 01 Apr 2024
25. Education and Outreach Working Group (EOWG). https://www.w3.org/WAI/about/groups/eowg/. Accessed 09 Apr 2024
26. Meinel, C., Leifer, L., Plattner, H.: Design Thinking: Understand - Improve - Apply. Springer, Berlin (2010). https://doi.org/10.1007/978-3-642-13757-0
27. Beyer, H., Holtzblatt, K.: Contextual design. Interactions **6**, 32–42 (1999). https://doi.org/10.1145/291224.291229

28. Jaaskelainen, R.: Think-Aloud Protocol. Handbook of Translation Studies. John Benjamins Publishing, Amsterdam (2010)
29. W3C/WAI Education and Outreach Working Group: Requirements from interviews with users for the W3C Accessibility Evaluation Tools List. https://www.w3.org/WAI/EO/wiki/Evaluation_tools/Revision_2021#Requirements_from_interviews_with_users. Accessed 11 Apr 2024

A Declarative Model for Web Content Accessibility Evaluation Process

Jinat Ara[1](✉) and Cecilia Sik-Lanyi[1,2]

[1] Department of Electrical Engineering and Information Systems, University of Pannonia, Egyetem u. 10, Veszprem 8200, Hungary
jinat.ara@mik.uni-pannon.hu

[2] Hungarian Research Network, Piarista u. 4, Budapest 1052, Hungary

Abstract. Providing a complete accessibility evaluation report of web platforms is certainly a complex task that requires effort and a certain level of knowledge about web prototypes. Due to the complexity of the web platform, accessibility criteria are frequently ignored by web designers and developers hindering the accessing opportunities of people who have special needs. Thus, the development of an advanced website accessibility evaluation tool is an emerging demand especially to provide complete accessibility reports considering every type of disability. The goal of this work is to develop a web accessibility evaluation tool to implement in real-life cases to evaluate website accessibility. Although, the existing web accessibility evaluation tools are effective, due to some limitations their reported results seem unclear for the end-user which could act as a prime factor to reduce the effectiveness of the tool. By addressing those vulnerabilities, our prime focus is to develop a web content accessibility testing tool focusing on five aspects to improve the limitations of the existing tools such as i) updated guideline implementation, ii) incorporate guideline simplification process, iii) consider user criteria as additional evaluation criteria, iv) categorize the report for textual and non-textual information and v) provide the overall accessibility report in terms of accessibility score of each disability type.

Keywords: Web accessibility · Automated tools · Accessibility evaluation · Disability · Accessibility improvement

1 Introduction

In recent times, web has become the most frequently interacting digital platform for extracting information for every group of people including people with disabilities. According to the current statistics [1], every day a large number of people with disabilities interact with the web platform to navigate information. Unfortunately, according to [2], 81% of users with disabilities face difficulties while they interact with the website due to a lack of accessibility issues considered during website development. Therefore, to improve the experience of web accessing process, especially for this large group of people with disabilities, the improvement of web accessibility issues is significantly

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 84–92, 2024.
https://doi.org/10.1007/978-3-031-62846-7_10

urgent. Currently, among several tools and techniques, incorporating an automated web accessibility testing tool is the most effective process to identify and improve accessibility issues further [3]. In such a manner, the contribution of developing automated tools in the literature is significantly noticeable. Though these tools are competent enough, most of them has lack consideration of a wide array of aspects related to guidelines, user opinion, evaluation report visualization, and overall score computation which could bias the evaluation process and generate inappropriate results that could directly reduce tools' effectiveness, acceptability, reliability, and fairness. For example, from most of the existing tools generated reports, it is unclear which guidelines they have implemented, also, whether the additional information has been considered or not. Besides, they don't determine accessibility issues for textual components and don't provide the website's accessibility percentage for each disability type. However, ensuring tools' effectiveness, acceptability, reliability, and fairness should be the prime concern of any developed tools that consider web accessibility where these four aspects refer to the representation of complete information about the assessed guidelines, success criteria, conformance level, evaluated web features, additional evaluation criteria, evaluable and non-evaluable guidelines, future improvement suggestions, and overall accessibility score with disability type. Therefore, considering these issues, our designed framework considered five aspects such as i) implements an updated version of WCAG 2.2, ii) considers guideline simplification technique to improve the guideline parsing, iii) considers user criteria as additional criteria along with WCAG, iv) provides evaluation report in terms of textual and non-textual objects of the evaluated website and v) generates overall accessibility evaluation report in terms of every disability type. The developed tool has been evaluated considering a comparative evaluation of functional properties of the tool with existing tools, and an experimental analysis with the user testing. The developed framework is able to improve the automated accessibility evaluation process by limiting the current issues of the existing tools.

This paper is structured in the following order: Sect. 2 provides the literature discussion considering similar recent studies. Section 3 describes the methodological aspects of the proposed tool where we explain the structure of the proposed tool with a detailed description of the tool's functional properties. Section 4 discusses the results obtained from the performed experiment with a comparative analysis of user studies and the tools' results. Finally, Sect. 5 concludes the paper through conclusion.

2 State of the Art

From 1996 to now, web accessibility evaluation remains an important task for web practitioners as it aims to increase social inclusion for people with special needs [4]. To investigate and represent the current scenario of the web accessibility evaluation process, a detailed systematic literature review has been conducted that can be found in [5], which directed that in the literature, majority of the work focused on the experimental investigation of current web incorporating several methods including hybrid evaluation process [6–9]. Although this heuristic evaluation process is effective; considering cost and time minimization, automated tools require low cost, provide reports in a short time, and are easier to implement. Therefore, several works have been contributed to develop an automated web accessibility testing tool. Although existing automated tools are effective, their report might be vulnerable due to several issues related to the tool's effectiveness, acceptability, and reliability (described in Sect. 1). Therefore, it is crucial to investigate the existing automated tools to identify their effectiveness as they are mostly implemented or used tools for website investigation. In that manner, we investigated several existing literatures that proposed automated tools in terms of their effectiveness focusing on their properties and functionality, and identified several vulnerabilities in their solutions. For example, Jens Pelzetter [10] proposed an automated tool to evaluate website accessibility by incorporating the Accessibility Conformance Testing (ACT) rule set through ontology modeling. Though the proposed system is effective, ACT rules incorporate ontology modeling increases the high risk of ambiguity during the testing process. In another automated tool proposed by Michailidou et al. [11], they incorporated HTML Document Object Model (DOM) and predicted webpage visual complexity. Also, Shrestha [12] developed a neural network-based automatic evaluation of webpage image descriptions following the National Center for Accessible Media (NCAM) principles. Though these two works accurately predict accessibility issues, they consider few assessment features such as very few checkpoints, and do not provide an overall accessibility score. Also, the followed guidelines are not widely accepted and might not consider every aspect that needs to be focused on during the evaluation process. In another work, Boyalakuntla et al. [13] developed an automated accessibility evaluation tool providing command line and browser plugin facilities. However, it supports issues related to aria, color-contrast, and interaction-related issues, which is not sufficient to address all the accessibility issues, other objects should be taken into consideration. Hilera and Timbi-Sisalima [14] developed a tool to asses website accessibility according to the semantic similarity of multiple source reports. Also, Ingavélez-Guerra et al. [15] incorporated ontology and knowledge modeling perspective to support accessibility analysis and evaluation of website objects. These developments are useful; however, the evaluation result might be deceptive due to the lack of additional requirements and the chance of acceptability to the user might be reduced. Although these approaches are effective, several factors, limit the outcome of the proposed systems. Mostly, we observed issues related to a few assessment features, improper guidelines with limited checkpoints incorporation, elimination of textual feature assessment, and less attention to overall score computation and representation process. Therefore, in this work, we present a declarative model of the web accessibility evaluation process that can consider the addressed issues and improve the accessibility evaluation process of web content.

3 Methodology

Generally, web page structure defines the representation of a set of objects that present the structural and design elements such as form, menu, links, text, color, image, video, etc. The ratio of these objects is not limited or certain for every webpage which indicates that types of web objects might vary from one page to another. Therefore, it is quite difficult to generate a complete and correct evaluation result of a webpage considering every accessibility criterion. However, considering the standard guideline and incorporating success criteria as much as possible in the investigation process could improve the evaluation process. Therefore, in our study, we incorporated much more web page elements compared to other models, rather than the most selective ones. Compared to other similar models/approaches, our proposed model is simple, easy to understand, and able to generate accurate and reliable results based on common aspects of an HTML Document Object Model (DOM).

The architecture of the proposed model is shown in Fig. 1. Referring to the architectural design (Fig. 1), our initial goal is to improve the effectiveness of the developed tool by incorporating appropriate guidelines (WCAG 2.2, the latest version) and user requirements as an additional evaluation criterion. Also, as WCAG needs technical knowledge to understand and implement effectively, thus guidelines have been simplified using the guideline modeling concept, which can be found in one of our previous studies [16]. After initializing guidelines and user requirements (where initialization of guideline and user criteria is the prime concern as the measurement of standard guidelines and additional criteria can act as a crucial factor to improve the evaluation process), web parsing and web objects extraction have been performed according to the user input where the user has to submit the website link that they want to evaluate. We extracted all the web objects from the HTML Document Object Model (DOM) structure including textual and arbitrary objects referring to HTML tags via given URLs and evaluated them according to the selected guidelines and user requirements. After extracting all the website information regarding 23 web objects (such as audio, video, links, forms, titles, paragraphs, header, etc.), we implemented two complexity analysis algorithms to analyze the complexity of the web objects (one for textual information and another for non-textual information analysis) which evaluate the web objects from DOM structure and categorize the result in terms of textual, non-textual complexity and provide the statistical report through graphical representation. To analyze textual components, we used Natural Language Processing (NLP) techniques and to analyze non-textual components, we used a simple rule-based approach with several auxiliary methods. After complexity analysis, we represented our generated textual complexity report, non-textual complexity report, and statistical report to the end user through three separate window views. The textual complexity report provides results of all the tested WCAG guidelines and user criteria related to the textual component. Similarly, the non-textual complexity report provides results of all the tested WCAG guidelines and user criteria related to the non-textual component. Besides, the statistical report visualization window presents the overall computed accessibility score, accessibility score for each disability type (Cognitive, Vision, Learning, Motion, and Hearing), and the number of Passed (P), Failed (F), Not Detected (ND) and Not tested (NT) guideline information. Here, P, F, ND, and NT are four assessment terminologies that have been used in the evaluation process where 'Passed' refers successfully implemented guidelines; 'Fail' refers wrongly implemented guidelines; 'Not detected' refers not implemented guidelines; and 'Not tested' refers guidelines that require the user or expert testing. Figure 1 shows the proposed declarative model architecture in detail. The proposed model is developed using Python programming language in Sublime Text editor using Windows version 10 and Intel COREi7 processor.

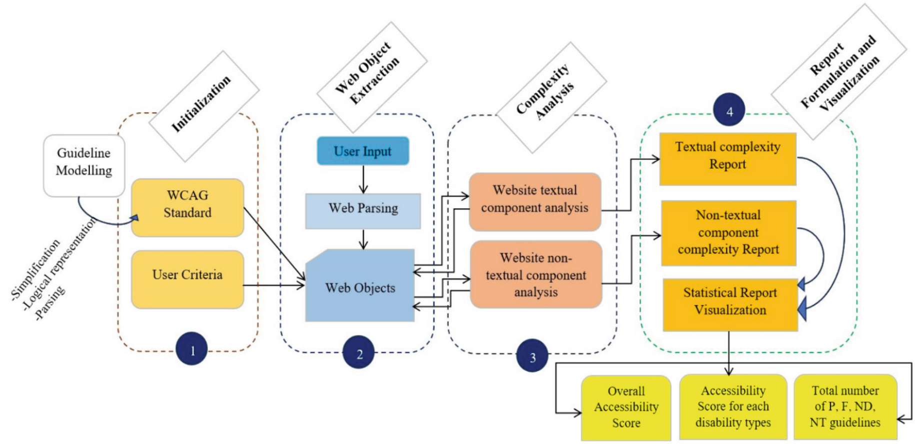


Fig. 1. Architecture of the proposed declarative model for web accessibility evaluation.

4 Results

After developing the proposed model, our immediate plan was to perform software validation by performing a comparative analysis with similar existing models and conducting an experimental analysis. To perform the comparative analysis, we considered several functional properties, and for experimental analysis, we considered a user study to validate the computed result of our proposed tool.

Regarding comparative assessment, we compared our developed tool with the existing six models that have been mentioned in state-of-the-art literature in terms of several functionalities as mentioned in Table 1. Table 1 depicts that for information regarding accessibility evaluation criteria such as guidelines, success criteria, and conformance level, one model provides complete information in this regard (Boyalakuntla et al. (2021)), and two models (Hilera and Timbi-Sisalima (2016); Ingavélez-Guerra et al. (2018)) provides success criteria and conformance level information though they didn't mention which guideline actually they implemented to do the evaluation and other models have no such concern. Besides, for textual and non-textual information and evaluated and not evaluated guideline information, none of the models were found with such concern in their evaluation process. Considering the overall accessibility score, three models were found with such concern though other models didn't compute the overall accessibility score. Finally, focusing on disability types in the accessibility score computation process, none of the models generate an accessibility score for every type of disability. From all of these aspects, the proposed model considers every aspect that has been discussed in Sect. 1 to improve the accessibility evaluation result of a webpage. Therefore, we hypothesize that the proposed model can provide a complete and updated view of webpage accessibility and competent than the other similar approaches or models or tools.

To further evaluate and validate the model, we performed an experimental analysis considering user study. For such, we randomly selected 10 Assistive Technologies webpages and asked 5 participants to evaluate and rate (range 1 to 5) the selected ten

Table 1. Comparative assessment results with existing models.

References	Assessment Functionalities						
Heading level	Criteria-1	Criteria-2	Criteria-3	Criteria-4	Criteria-4	Criteria-6	Criteria-7
Jens Pelzetter (2018)	✗ (No)	✓ (Yes)	✓ (Yes)	✗ (No)	✗ (No)	✓ (Yes)	✗ (No)
Michailidou et al. (2021)	✗ (No)	✗ (No)	✗ (No)	✗ (No)	✗ (No)	✓ (Yes)	✗ (No)
Shrestha (2021)	✗ (No)	✗ (No)	✗ (No)	✗ (No)	✗ (No)	✗ (No)	✗ (No)
Boyalakuntla et al. (2021)	✓ (Yes)	✓ (Yes)	✓ (Yes)	✗ (No)	✗ (No)	✗ (No)	✗ (No)
Hilera and Timbi-Sisalima (2016)	✗ (No)	✓ (Yes)	✓ (Yes)	✗ (No)	✗ (No)	✓ (Yes)	✗ (No)
Ingavélez-Guerra et al. (2018)	✗ (No)	✓ (Yes)	✗ (No)	✗ (No)	✗ (No)	✗ (No)	✗ (No)
Proposed Model	**✓ (Yes)**	**✓ (Yes)**	**✓ (Yes)**	**✓ (Yes)**	**✓ (Yes)**	**✓ (Yes)**	**✓ (Yes)**

webpages in terms of four aspects: interactivity, effectiveness, understandability, and complexity according to their point of view where each participant had to test 2 webpages. All the participants were cognitively stable, had sufficient knowledge about the 'web interface', and were active in the internet platform for study and daily activities. Then the participants' ratings were compared with the overall accessibility score computed by our proposed model for the same selected ten pages, as shown in Table 2. To compare the user-given score and the computed score by the proposed tool, we considered the threshold value approach to classify both scores in order to define their accessibility status. For user given score for a particular webpage, we define a threshold value where if the score $\geq$4.5 (90%), then Completely Accessible; if the score $\geq$3.75 (75%), then Comparatively Accessible; if the score $\geq$2.75 (55%), then Partially Accessible; and if score $<$2.75 ($<$55%), then Slightly Accessible. Besides for computed score by our proposed tool, the stated threshold value is, if the accessibility score $\geq$90%, then Completely Accessible; if the accessibility score $\geq$75%, then Comparatively Accessible; if the accessibility score $\geq$55%, then Partially Accessible; and if accessibility score $<$55%, then Slightly Accessible. According to the stated threshold value, we classified user given score and the computed score by the proposed tool (as shown in Table 2). Table 2 shows that majority of the tested webpages are classified as partially accessible in both users' given score or rating and the computed score. Also, for most of the tested webpages, the accessibility score generated by the proposed tool is significantly related to the participants' or users' ratings in terms of their accessibility status. Therefore, we conclude that the outcome of the computed score by the proposed tool is significantly and positively related to the rating of the participants which depicts the validity of the proposed tool.

Table 2. Experimental results of the tested websites.

Web-ID	Websites URLs	Avg. User Rating/Score	Accessibility Status (according to user rating)	Computed Score by the proposed tool	Accessibility Status (according to the tool's result)
ID-1	https://liveit-project.eu/	2.9	Partially accessible	51.22%	Slightly accessible
ID-2	https://www.easyreading.eu/	3.2	Partially accessible	62.15%	Partially accessible
ID-3	https://www.guidedogs.org.uk/	3.5	Partially accessible	57.5%	Partially accessible
ID-4	https://at4kids.com/	2.8	Partially accessible	40.0%	Slightly accessible
ID-5	https://www.bookshare.org/cms/	3.2	Partially accessible	58.97%	Partially accessible
ID-6	https://veroniiica.com/	2.5	Slightly accessible	40.91%	Slightly accessible
ID-7	https://attoday.co.uk/	3.1	Partially accessible	55.0%	Partially accessible
ID-8	https://www.eastersealscrossroads.org/	2.9	Partially accessible	63.16%	Partially accessible
ID-9	https://cadanat.com/	3.1	Partially accessible	58.97%	Partially accessible
ID-10	https://otswithapps.com/	2.6	Slightly accessible	51.22%	Slightly accessible

5 Conclusion

The web platform is continuously updating and adding new prototypes to facilitate the functionalities to the end user. This updated information is useful though it creates barriers for people with special needs that raise accessibility issues. Therefore, to assess the accessibility of the web platform, an updated and advanced evaluation tool is an emerging need. With this focus, to develop our accessibility evaluation tool, we had to consider several aspects in our mind and incorporate such advanced prototypes in the development to overcome existing shortcomings. However, some challenges had to be faced during the implementation of the proposed solution such as guideline modeling, NLP technique incorporation, and appropriate library selection. Also, we have some limitations of this work, such as we experimented on fewer samples of webpages with fewer users. Therefore, in the future, in the extended version of this work, we plan to incorporate more functional properties into the evaluation report and perform the experiment with more samples and more users.

Acknowledgments. The authors would like to thank the support of the National Research Development and Innovation Office, project no. 2021-1.2.4-TÉT-2021-00007.

Disclosure of Interests. There is no conflict of interest.

References

1. Website Accessibility and How People with Disabilities Use the Web. https://accessly.io/website-accessiblity-truth-how-people-with-disabilities-use-websites/
2. Li, L., Wang, C., Song, S., Yu, Z., Zhou, F., Bu, J.: A task assignment strategy for crowdsourcing-based web accessibility evaluation system. In: Proceedings of the 14th International Web for All Conference, pp. 1–4 (2017)
3. Misra, B.B.: New software tools, databases, and resources in metabolomics: updates from 2020. Metabolomics **17**(5), 49 (2021)
4. Acosta-Vargas, P., Salvador-Ullauri, L.A., Luján-Mora, S.: A heuristic method to evaluate web accessibility for users with low vision. IEEE Access **7**, 125634–125648 (2019)
5. Ara, J., Sik-Lanyi, C., Kelemen, A.: Accessibility engineering in web evaluation process: a systematic literature review. Univ. Access Inf. Soc. 1–34 (2023)
6. Mohamad, Y., Velasco, C.A., Kaklanis, N., Tzovaras, D., Paternò, F.: A holistic decision support environment for web accessibility. In: Miesenberger, K., Kouroupetroglou, G. (eds.) ICCHP 2018. LNCS, vol. 10896, pp. 3–7. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-94277-3_1
7. Li, L., Bu, J., Wang, C., Yu, Z., Wang, W., Wu, Y., Gu, C., Zhou, Q.: CrowdAE: a crowdsourcing system with human inspection quality enhancement for web accessibility evaluation. In: Miesenberger, K., Kouroupetroglou, G. (eds.) ICCHP 2018. LNCS, vol. 10896, pp. 27–30. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-94277-3_5
8. Alahmadi, T.: A multi-method evaluation of university website accessibility: foregrounding user-centred design, mining source code and using a quantitative metric. In: Proceedings of the 14th International Web for All Conference, pp. 1–2 (2017)
9. Song, S., et al.: Web accessibility evaluation in a crowdsourcing-based system with expertise-based decision strategy. In: Proceedings of the 15th International Web for All Conference, pp. 1–4 (2018)
10. Pelzetter, J.: Using ontologies as a foundation for web accessibility tools. In: Proceedings of the 15th International Web for All Conference, pp. 1–2 (2018)
11. Michailidou, E., Eraslan, S., Yesilada, Y., Harper, S.: Automated prediction of visual complexity of web pages: tools and evaluations. Int. J. Hum. Comput. Stud. **145**, 102523 (2021)
12. Shrestha, R.: A neural network model and framework for an automatic evaluation of image descriptions based on NCAM image accessibility guidelines. In: 2021 4th Artificial Intelligence and Cloud Computing Conference, pp. 68–73 (2021)
13. Boyalakuntla, K., Venigalla, A.S.M., Chimalakonda, S.: WAccess–a web accessibility tool based on the latest WCAG 2.2 guidelines. arXiv preprint arXiv:2107.06799 (2021)
14. Hilera, J.R., Otón, S., Martin-Amor, C., Timbi-Sisalima, C.: Towards a service-based architecture for web accessibility federated evaluation. In: 9th International Conference on Advances in Computer-Human Interactions (ACHI 2016), pp. 6–10 (2016)

15. Ingavelez-Guerra, P., et al.: A proposal based on knowledge modeling and ontologies to support the accessibility evaluation process of learning objects. In: 2018 Congreso Argentino de Ciencias de la Informática y Desarrollos de Investigación (CACIDI), pp. 1–5. IEEE (2018)
16. Ara, J., Sik-Lanyi, C.: AccGuideLiner: towards a modelling approach of web accessibility requirements following WCAG 2.2. In: 2023 IEEE International Conference on Smart Information Systems and Technologies (SIST), Astana, Kazakhstan, pp. 10–15 (2023). https://doi.org/10.1109/SIST58284.2023.10223541

How to Provide Actionable Feedback on Web Accessibility Issues – Development of a Model Curriculum and Practical Guidelines

Lukas Baumann(✉) and Susanne Dirks

Department of Rehabilitation Technology, TU Dortmund University, Dortmund, Germany
{lukas.baumann,susanne.dirks}@tu-dortmund.de

Abstract. In the EU, the Web Accessibility Directive (2016, WAD) requires public sector bodies' websites and mobile apps to be accessible to all users and to document and monitor accessibility problems. The WAD also introduced a feedback mechanism that can be used to flag accessibility problems or request information about content that is provided in a non-accessible way. Member state reports from 2021 show that most websites and mobile applications only meet some required demands, although there are hardly any complaints about existing barriers. In this context, the paper presents the validation of a model curriculum and practical guidelines for providing actionable feedback on web accessibility issues. The validation process involved six workshops, a questionnaire and interviews with different stakeholders. In general, the results indicate that the developed resources are valuable for various training contexts and helpful in creating and implementing actual training courses. However, participants stressed the need for more in-depth information on training and partnerships in order to use the material for training. The results will be discussed, focusing on implications for further usage and development of the resources.

Keywords: Digital Participation · Digital Accessibility · Web Accessibility Directive · Feedback Mechanism · Curriculum Development

1 Introduction

Accessibility is a fundamental requirement for the full participation of persons with disabilities [1]. Various national and international laws ensure the right to unrestricted digital participation. In the EU, the Web Accessibility Directive (2016, WAD) requires public sector bodies' (PSB) websites and mobile apps to be accessible to all users, despite their abilities, and to document and monitor accessibility problems. To document and improve digital accessibility, the WAD also introduced three mandatory instruments, including an accessibility statement that PSB must publish for all their websites and mobile apps. A critical part of the accessibility statement is a feedback mechanism that can be used to flag accessibility problems or request information about content that is provided in a non-accessible way.

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 93–99, 2024.
https://doi.org/10.1007/978-3-031-62846-7_11

2 Implementation of the WAD – Actual State

The WAD empowers the European Commission to adopt implementing acts detailing the rules for the practical implementation of many of the Directive's provisions [2]. Every three years, the member states must inform the European Commission about the implementation of the WAD and publish a report on the results of monitoring and enforcement measures [3]. The member state reports from 2021 show that most websites and mobile applications only meet some required demands, although there are hardly any complaints on still existing barriers. Pereira, Matos & Duarte summarise the lack of accessibility statements in most reviewed apps in this reporting period [4]. Results from studies from the European Commission and the European Disability Forum show that accessibility statements were only available on less than half of the analysed websites and that the feedback mechanism was missing on many PSB websites [2, 5]. In their study from 2023 on how healthcare providers in Sweden have applied accessibility statements on their websites, Johnsson et al. showed that 36 of the 37 evaluated healthcare providers published an accessibility statement. Still, none of the healthcare providers fully met the requirements for accessibility statements [6].

Furthermore, results from a survey by the European Commission showed that one-third of the users were not familiar with the existence of the accessibility statement and emphasised the need to improve the visibility and accessibility of the feedback mechanism [7]. The first results from the UPowerWAD ("Users Power the Web Accessibility Directive") project provide results on the usage of accessibility feedback mechanisms, the needs of people with disabilities and the main requirements for a better understanding of existing barriers. Furthermore, the project team published suggestions on how to improve the feedback mechanism [8, 9]. Additionally, the survey results published by Halbach, Fuglerud & Snaprud provide guidance on the design of better feedback mechanisms [10].

3 Development of a Model Curriculum and Practical Guidelines

In the context of the UPowerWAD project and the goal to leverage the knowledge of people with disabilities about the feedback mechanism, the project team developed a "model curriculum on how to train people with disabilities to provide actionable feedback on web accessibility issues" [11] based on the first project results. In general, developing a curriculum describes the process of planning and guiding learning and is central to the teaching and learning process [12]. The main aim is to establish guidance on how to give actionable feedback about accessibility (i.e., the feedback is clear and has the relevant and necessary information that helps the website owner solve the accessibility issue), considering the knowledge and competence of different user groups. The curriculum's target groups are Vocational Education and Training (VET) providers and Organisations of Persons with Disabilities (OPD) across the EU who are interested in providing training courses for their members and the general public.

The consortium also developed "practical guidelines on how to empower people with disabilities to participate in implementing the WAD" [13]. The guidelines are an additional tool to the model curriculum, containing information and practical instruction for creating training courses. They also offer guidance on connecting the curriculum with other project results.

4 Method

As a first step in developing the model curriculum, the most important results of the first project phase (existing barriers, needs, and best-practice-examples; see Baumann et al. [8] for Methodology) were identified. This facilitated the development of an initial draft of the model curriculum. The draft was then validated in six workshops with 135 participants in total (see Table 1). The participants of the workshops were people with disabilities, VET providers, members of OPDs, PSB, and IT suppliers. The workshops focused on the stakeholders' needs, preferences, and expectations regarding the model curriculum. Including these specific stakeholders assured that the "professionals who best know their context and teaching situation can voice their expertise" [14]. By conducting workshops, the researchers aimed to identify new and relevant factors to both the researchers and participants [15]. All workshops took place in May/June 2023 online via Zoom.

Table 1. Validation workshops.

Workshop	Participants	Language
1	30 participants (17 different countries)	English
2	62 participants	Swedish
3	25 participants	French
4	8 participants	German
5	7 participants	German
6	3 participants	German

As shown in Table 1, workshops 1–3 were held in English, Swedish and French. They were held as interactive webinars to raise awareness and receive general feedback on the structure and content of the curriculum. Workshops 4–6 were held with VET providers and representatives of OPDs from Germany. These workshops were realised as focus group meetings [16]. The researchers shared the first draft of the curriculum, and participants actively discussed the project results based on a structured interview guideline.

The results of the workshops confirmed the assumption that the curriculum needs supporting guidelines for the actual implementation of training courses. Subsequently, an initial draft of the practical guidelines was developed. To validate the guidelines, the consortium identified and contacted interested experts (VET trainers, Members of national OPDs and associated partners) from their network. The practical guidelines were sent to the experts along with the other project results. The research method involved a questionnaire and interviews. The validation process took place in January 2024. The validation not only focused on the practical guidelines but also evaluated the full set of the project results. The project partners adjusted the questionnaire/interview guidelines for the different stakeholder groups to get feedback on various aspects. Ten people in total filled out the questionnaire. The six representatives of OPDs operated in five different nations (Sweden, Slovenia, Ireland, Poland, and Hungary). One OPD operates primarily at EU level. The four VET trainers who filled out the questionnaire are working in the ICT sector in Cyprus. Interviews were conducted with four persons based in Germany (2 representatives of OPDs, 1 VET provider and one accessibility expert).

5 Results

The results of the workshops, interviews, and questionnaires to validate the model curriculum and the practical guidelines showed that most participants were willing to develop training courses based on the provided materials, include the content in existing concepts, or further disseminate the project results.

Especially the structure and the content of the curriculum were rated positively. Based on the results, the final model curriculum provides three teaching modules to empower end-users to give feedback on existing barriers:

- Module 1 contains materials to support digital accessibility training. This module focuses on knowledge about digital accessibility and the implementation of the WAD.
- Module 2 addresses the importance of user feedback and highlights specific challenges of using the WAD feedback mechanism.
- Module 3 provides an overview of the necessary basic information for feedback to be actionable.

Each module includes learning units with practical suggestions for content, expected learning outcomes, and suggestions for teaching methods. The modules can be used as a whole or as individual units. The model curriculum was translated into German, French, and Swedish.

The final practical guidelines contain information on various topics related to the curriculum and the teaching itself. They include information on the universal design for learning [17] and other inclusive training methods, relevant target groups and target group's needs, and how to utilise, localise and adapt the model curriculum. Both documents, the model curriculum and the practical guidelines, provide examples of designing different training packages (e.g. singular workshops or detailed training packages).

In general, the validation process results show three main themes relevant to providing training on actionable feedback on web accessibility issues: awareness, empowerment, and practical application (see Fig. 1).

Even though the developed resources seem valuable for various training contexts and are considered helpful for creating and implementing actual training courses, the validation process highlighted that the resources require additional time and effort before implementing training. Participants stressed the need for more in-depth information on training and partnerships in order to use the material for training. This includes the necessity of the local adaptation of the material (e.g. adding specific examples and local resources).

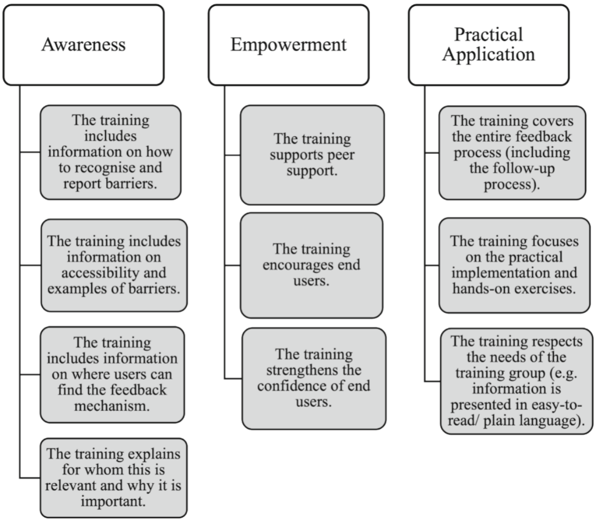


Fig. 1. Main themes relevant for providing training on actionable feedback on web accessibility issues.

6 Discussion

The resources developed in the UPowerWAD project provide a valuable framework for implementing actual training courses. Due to the modular approach, the content of the curriculum can be used in various contexts and allows for adjustments and integration with other courses. VET providers or OPDs providing training courses on web accessibility can either use the curriculum as a whole or integrate single modules into their training. The curriculum and its content can also be used by trainers delivering courses on ICT knowledge. In addition, the curriculum can be adapted to different target groups and contexts and used by advocacy groups, consulting services, or other user-centred approaches. The model curriculum and guidelines must be translated and adapted to national requirements to ensure successful implementation in various EU member states.

The results show that, based on the curriculum, developing further learning material and content that focuses on specific user groups and local conditions might be necessary. Many of these aspects have already been addressed with the development of the practical guidelines, such as the necessity of easy-to-read and plain language for people with intellectual disabilities. However, some suggestions from the workshop participants were not possible to implement during the project. This includes, for example, developing more practical material (e.g. teaching templates, creating a toolbox, etc.) to reduce the time and effort for the trainers to prepare training. These aspects need to be considered in future studies and projects. In this context, developing a train-the-trainer concept needs to be considered. Here, especially OPDs can help VET providers tailor the training to the needs of the participants. This highlights the important role of OPDs in the context of the WAD.

Another critical aspect of the sustainable applicability of the developed resources is their continuous adaptation to current laws, regulations, and guidelines. This includes, for example, integrating WCAG updates (e.g., version 2.2 [18]) and information from the upcoming reporting period into both the curriculum and the guidelines.

7 Conclusion

Both the developed curriculum and the practical guidelines will help VET providers and OPDs conduct courses or workshops to improve users' knowledge, especially users with disabilities, about the websites and apps of PSBs and enable them to provide actionable feedback on accessibility issues. The project results also highlight the need for appropriate and practical resources to develop VET and OPD courses to train people with disabilities to provide actionable feedback on web accessibility issues. The results of the validation of the model curriculum and the accompanying guidelines showed that the translation of the material into the national languages of the member states and the establishment of partnerships between VET providers and OPDs, as well as further adjustments to the developed material are additional vital elements for the success of the adaption of the resources. Finally, it needs to be pointed out that on many websites, the "feedback path" itself is still inaccessible and, therefore, difficult to use for people with disabilities. PSBs need to understand that feedback options must be accessible to all users, and they are obliged to implement them accordingly. In this context, an extension of the curriculum should be considered to be used for website owners' training courses.

Acknowledgments. The project was realised by the Department of Rehabilitation Technology at the TU Dortmund University (Germany) in cooperation with the European Blind Union, Funka (Sweden), and the Synthesis Center for Research and Education (Cyprus). The project received funding from the Erasmus+ Programme of the European Union (Project ID: 2021-1-DE02-KA220-VET-000033176).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. European Commission. https://data.europa.eu/doi/https://doi.org/10.2767/31633. Accessed 14 Apr 2024
2. European Commission. https://digital-strategy.ec.europa.eu/en/library/study-implementation-web-accessibility-directive-monitoring-reporting-and-evaluation. Accessed 14 Apr 2024
3. European Commission. https://digital-strategy.ec.europa.eu/en/library/web-accessibility-directive-monitoring-reports. Accessed 14 Apr 2024
4. Pereira, L., Matos, M., Duarte, C.: Exploring mobile device accessibility: challenges, insights, and recommendations for evaluation methodologies. In: Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), pp. 352–371 (2024)
5. European Disability Forum. https://ec.europa.eu/newsroom/dae/redirection/document/78800. Accessed 14 Apr 2024

6. Jonsson, M., Gustavsson, C., Gulliksen, J., et al.: How have public healthcare providers in Sweden conformed to the European Union's Web Accessibility Directive regarding accessibility statements on their websites?. Univ Access Inf Soc (2023)
7. European Commission. https://digital-strategy.ec.europa.eu/en/library/study-supporting-review-web-accessibility-directive. Accessed 14 Apr 2024
8. Baumann, L., et al.: Involving, empowering, and training end users with disabilities to fully participate in the web accessibility directive objectives first results from the UPowerWAD project. Stud. Health Technol. Inform. **306**, 364–370 (2023)
9. UPowerWAD. https://www.funka.com/en/projekt/upower-wad/. Accessed 14 Apr 2024
10. Halbach, T., Fuglerud, K., Snaprud, M.: User insights for better and more inclusive online public services: a survey study. In: Studies in Health Technology and Informatics, vol. 297, pp. 36–43 (2022)
11. UPowerWAD. https://www.funka.com/en/projekt/upower-wad/main-activities-and-results/curriculum/. Accessed 14 Apr 2024
12. Rogers, A., Taylor, P.: Participatory Curriculum Development in Agricultural Education. A Training Guide. Rome: FAO (1998)
13. UPowerWAD. https://www.funka.com/en/projekt/upower-wad/main-activities-and-results/practical-guidelines/. Accessed 14 Apr 2024
14. Alexander, I., Hjortsø, C.: Sources of complexity in participatory curriculum development: an activity system and stakeholder analysis approach to the analyses of tensions and contradictions. High. Educ. **77**, 1–22 (2019)
15. Ørngreen, R., Levinsen, K.T.: Workshops as a Research Methodology. Electron. J. E Learn. **15**(1), 70–81 (2017)
16. Katz-Buonincontro, J.: How to Interview and Conduct Focus Groups. American Psychological Association (2022)
17. CAST. http://udlguidelines.cast.org. Accessed 14 Apr 2024
18. Campbell, A., Adams, C., Montgomery, R.B., Cooper M., Kirkpatrick, A.: https://www.w3.org/TR/WCAG22. Accessed 14 Apr 2024

Introducing Computer Science Students to Inclusive Design and Accessibility: Evaluation of Practical Exercises with a Low-Cost Simulation Kit

Helen Petrie[1](✉) and Jenny Darzentas[2]

[1] Department of Computer Science, University of York, Deramore Lane, Heslington YO10 5GH, York, UK
helen.petrie@york.ac.uk

[2] Department of Product and Systems Design Engineering, University of the Aegean, 2 Konstantinopouleos Street, GR-84100 Ermoupolis, Syros, Greece
Jennyd@aegean.gr

Abstract. Young developers of digital technologies need to be aware of the principles of inclusive design and accessibility, but it is difficult to teach these concepts in the abstract. Practical simulations of the effects of disability and aging can be engaging for students, but expensive to purchase and are criticised as disrespectful to people with disabilities and older people. We have developed a very low-cost simulation kit and used it in a practical exercise with computer science students, emphasising that this are not the same as being disabled or older. When possible, the practical exercises are complemented by a session with a disabled person who talks about their experience with using digital technologies. The practical exercise gives some insight into the experience of visually disabled and older people in using these technologies. The low-cost nature and tasks we created to undertake with the kit were designed to also instil some fun into the exercise. An evaluation with 63 first year undergraduate computer science students yielded positive ratings of the exercise and many interesting comments from the students.

Keywords: Inclusive design · accessibility · low-cost simulation · evaluation

1 Introduction

Young developers are the ones who usually create new digital technologies to be used by people with disabilities and older people, whether these are mainstream technologies or ones developed specifically to support those user groups. But young developers often have negative attitudes to disabled and older people and little understanding of their lived experiences with digital technologies. To engage them with inclusive design and accessibility principles, they need some understanding of, and curiosity about, these issues. Simulations of disability and ageing have long been, and continue to be, used in education and professional development [9, 10], although they are controversial in how

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 100–108, 2024.
https://doi.org/10.1007/978-3-031-62846-7_12

respectful of disabled and older people they are and how effective they are in changing opinions [5, 8]. Simulation kits can also be very expensive to purchase (e.g. the GERT age simulation suit costs approximately 1400 euros, www.age-simulation-suit.com/; one set of visual impairment glasses approximately 280 euros, vinesimspecs.co.uk/). Therefore, we developed a very low-cost simulation kit and evaluated it with a class of first year undergraduate computer science students, who have three hours of compulsory teaching on inclusive design and accessibility.

In addition to the simulation tasks carried out by the students with the low-cost kit, this module included a lecture from a person with disabilities. For example, Fig. 1 (taken at a lecture which was held during the last semester before the COVID-19 pandemic disrupted face-to-face teaching), shows the lecturer, who is totally blind, demonstrating his braille notetaker to a very attentive group of 1st year computer science students. This demonstration was at the end of the lecture he gave about assistive technology and the needs of disabled people and was very well attended by the students.

Fig. 1. Lecturer who is totally blind demonstrating his braille notetaker to 1st year computer science students. (Copyright: Helen Petrie. Thanks to Kevin Carey, ex-Chair, Royal National Institute of Blind People, for permission to use this photo).

2 State of the Art

It is well established that young people tend to have negative attitudes about older people and many people have negative attitudes and a lack of understanding about people with disabilities. This has been shown across many groups of people, particularly those who should be supporting these groups and in many countries. Ageism has been found extensively in healthcare, for example amongst doctors, nurses and social workers in

Israel [1] and in Poland [4], in care homes in Portugal [12], in the advertising industry in Australia [2] and in studies across a range of workplaces in Canada [7]. Similarly, negative attitudes to people with disabilities are found across many groups of people in many countries [13, 14, 16]. These are just a selection of many studies in many countries, we are not targeting these countries in particular.

Of particular concern to us is ageism amongst young developers, who are those likely to be developing digital technologies that disabled and older people need and want to use, whether they are mainstream technologies or those developed specifically to support them. Wandke et al. [15] noted that there are a number of negative myths about older people and digital technologies which are common in HCI research: that older people are not interested in technology; that they consider them useless and unnecessary; that they lack the physical capabilities to use them; that they cannot understand interactive systems; and that they do not capable of learning to use them. A series of recent studies have shown that young computer science students and young technical developers in a range of countries (Austria, China and the UK) have very ageist attitudes about older people as users of digital technologies [3, 6, 11].

3 Method

3.1 Participants

A class of 180 first year undergraduate Computer Science students at large University in the UK had a one-hour lecture and a two-hour practical exercise about inclusive design and accessibility as part of one of their compulsory courses. The great majority of the students are 18 or 19 years, 84% are men, 15% are women, and approximately 1% identify differently. Attendance figures are recorded, however this information is confidential, but attendance at both activities was high. 63 students answered the final set of questions in the exercise, and different numbers of students answered the questions about each task (due to differing group sizes, see Procedure).

3.2 Materials

To create the low-cost simulation kit, we sourced easy to obtain and the most inexpensive materials we could find.

To simulate visual impairment, four sets of spectacles (or glasses as they are more commonly called in the UK), were created, using very inexpensive low magnification reading spectacles bought in a supermarket (approximately 2.00 euros per set, four sets are required to simulate the four different visual impairment conditions for each group of students, more spectacles could be added to simulate more conditions; all prices are estimates for the UK in 2024). These were then prepared in the following ways to simulate different visual impairments:

- Cataracts: spectacle lenses were smeared with a greasy substance (clear lip gloss) (Fig. 2, Left)
- Diabetic retinopathy: Spots of white corrector fluid/white paint were dotted across the lenses. (Fig. 2, Centre-Left)

- Glaucoma: White correction fluid/white paint was applied around the edge of the lenses. (Fig. 2, Centre-Right)
- Age-related macular degeneration: Black shoe polish was smeared in the middle of the lenses, and a little greasy substance (clear lip gloss) was smeared around the edge of the lenses. (Fig. 2, Right).

The white corrector, lip gloss and shoe polish we already had available from various sources, but might add 6.00 euros to the overall cost if we had bought them specifically for the exercise.

Fig. 2. Modified glasses to simulate conditions (Copyright: Helen Petrie).

To simulate the lack of dexterity associated with ageing (e.g. due to arthritis), students were asked to tape small buttons to the joints of their fingers and then put on two or three thin latex-free protective gloves (to create several layers glove between fingers and objects) (see Fig. 3). Packets of buttons were bought at 2.00 euros each, two backets for each group were needed (they had a tendency to disappear in the exercise sessions). Adhesive tape and a pair of scissors were needed for each group, these came from university stationery supplies. The most expensive item was the gloves (particularly since the pandemic), a box of 100 the appropriate gloves costs about 9.00 euros, but will be sufficient for 30 students.

To simulate the baby, 1.5 L filled bottles of mineral water with a small baby's nappy (diaper) on the base were prepared beforehand. A large bottle of mineral water cost approximately 1.50 euro (designer mineral water not required) and a packet of 28 nappies approximately 7.00 euros. In Fig. 4, a student doing the exercise is shown holding the "baby" and consulting the instructions for the tasks. Note that his standing up from the desk was required by the task, so he could not rest the baby or the phone on the table.

For us, the overall cost for each group of 5 students was approximately 16.10 euros (USD 15.60) (4 sets of spectacles; 2 packets of buttons; 15 gloves, 3 for each student; 1 bottle of mineral water; 1 nappy). We made up nine sets of the kit, as students took the exercise in groups of approximately 30, and we had a spare set for the lecturer and graduate teaching assistants to demonstrate how to attach the buttons to one's fingers (the trick is to stick the button on the adhesive tape and then wind round the finger, not try and balance the button on one's finger and cover with tape). However, these sets of the kit have now lasted for five years of use (they were not used for two academic years during the pandemic). Each year we need more gloves (these are disposed of after use), a small re-supply of buttons and an extra packet of nappies (the adhesive tapes on these disintegrate after robust handling by the students). We also reapply the different coatings to the spectacles if needed.

Fig. 3. Buttons and gloves to simulate dexterity issues (Copyright: Helen Petrie).

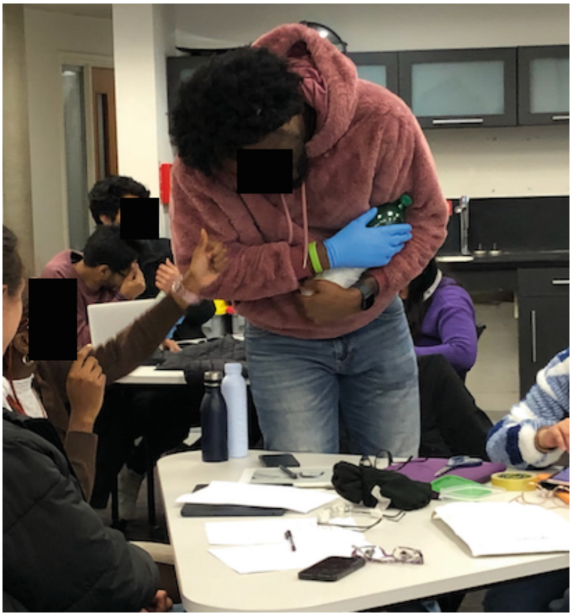


Fig. 4. Simulating holding a baby (Copyright: Helen Petrie).

3.3 Scenarios, Personas and Tasks

Participants were given two different personas and scenarios and four tasks to undertake (they had been studying personas, scenarios and task-based evaluation in their course, so it was emphasised to them that this exercise was building on the standard HCI techniques which they had been learning about). Tasks provided appropriate instructions as needed.

Persona 1 and Scenario 1

Chris is 25 years old and has been partially sighted since birth. They want to meet up with a friend, Helen, to have a coffee.

Task: Text your friend Helen on XXX (lecturer provided her number to make the task as realistic as possible, the students could hear her phone receiving messages during the exercise) asking when she is free.

Instructions for Task 1:

Wear one pair of the glasses provided (your choice) to simulate visual impairment.

Be careful if you wear glasses already not to strain your eyes. (If you wear glasses and the simulation glasses cause strain, just do the task without your own glasses on).

Swap phones with someone else in your group, preferably someone who has a phone you are not familiar with, so you are doing the task with an unfamiliar phone.

Persona 2 and Scenario 2

Sam is a 69-year-old grandparent who is babysitting their six-month-old grandchild. The baby is crying and after having had a look under the nappy, Sam realises that the baby has a bad case of nappy rash.

Task 2: Send a text message to the child's parent's mobile (XXX), asking whether you should call the doctor.

Instructions for Tasks 2 to 4.

Wear a pair of the provided glasses AND the buttons on your joints and two gloves on the hand (three if you can manage it) you use to operate a phone in order to simulate an older person with visual disability and some dexterity problems in their fingers.

Do the task standing up, holding the baby (rocking it to try to stop it crying). Don't put the phone on the table (too easy!), tuck it under your chin.

Task 3: You need to phone the doctor. You don't have their number in your contact list, so you need to dial it manually. Call the doctor on XXX but don't activate the call (it is the lecturer's number)!

Task 4: Imagine you are talking to the doctor on the phone and the doctor asks you to unwrap the baby's nappy enough to check the condition of the rash. Do this while **still** being able to talk on the phone.

3.4 Procedure

Students were given a short introductory talk about simulation of disability and ageing as a potentially useful exercise, emphasising that this is not the same as living with a disability or as an older person They then worked in groups of three to five students, with all students trying the visual impairment simulation spectacles, and then one student in the group trying each of the tasks, assisted and observed by the others. They answered a brief set of questions about each task (one rating item of how easy or difficult it was to do the task, and two open-ended questions about what was difficult and what was surprising about the task) and then a brief set of questions about the overall experience (three rating items, see Table 2, and one open-ended question about what was most useful about the exercise and why was it useful).

4 Results

Table 1 summarises the ratings of ease/difficulty of each task, on a rating scale from 1 = Very easy to 7 = Very difficult. From this we see that Task 1, which involved only wearing the visual impairments glasses, was rated as not significantly different from the midpoint of the scale (which was labelled "Neither easy nor difficult"). However, the three tasks which required students to also wear the dexterity impairment (buttons and gloves on their hand operating the phone) were all rated as significantly more difficult than the midpoint.

Table 1. Ratings of ease/difficulty of the four tasks.

Task	Number of students answering	Median rating (Semi interquartile range)	Wilcoxon one sample test
1	42	3.00 (1.00)	$F = -0.05$, n.s
2	27	4.00 (1.00)	$F = 3.18$, $p < 0.001$
3	21	5.00 (0.50)	$F = 3.29$, $p < 0.001$
4	20	5.00 (1.50)	$F = 2.65$, $p < 0.008$

Table 2 summarises the ratings of the overall experience of the practical exercise. 63 students answered all these questions. From this we can see that the students rated the exercise significantly positive in terms of 1. How it helped them understand more about the difficulties that disabled and older people might encounter with new technologies such as smartphones; 2. That they found the exercise fun; and 3. That they would recommend it to colleagues.

Table 2. Ratings of overall experience of the practical exercise ($N = 63$).

Question	Median rating (SIQR)	Wilcoxon one sample test
To what extent did this practical help you understand more about the difficulties people with disabilities and older people might encounter with smartphones? Rated: Not at all (1) – A great deal (7)	4.00 (1.00)	$F = 5.74$, $p < 0.001$
To what extent was the practical fun? Rated: Not at all (1) – Very much (7)	5.00 (1.00)	$F = 6.43$, $p < 0.001$
Would you recommend trying these simulation techniques to a colleague who was developing an app? Rated: Never (1) – Definitely (7)	5.00 (1.00)	$F = 5.79$, $p < 0.001$

The many answers to the open-ended questions further confirmed the students' engagement and the value of the exercise to them in terms of insights gained from particular difficulties to perform the various tasks. For example, they noted the difficulties with grip when wearing the buttons and gloves, that simulated stiff joints, as well as the difficulties this caused in tasks like typing a message on the phone, further exacerbated by using the glasses. The gloves also simulated the dryer skin of older people leading one student to remark: "I was not previously aware that touchscreens were less responsive for older people". With regard to the spectacles, one remark that summed up many other comments was "the impaired vision … clearly highlighted how important app visibility is". Furthermore, several students noted difficulties in multiple tasking, and doing tasks one handed. Somewhat tongue-in-cheek perhaps, one student commented: "Holding a bottle and using a phone is much harder than I thought it would be".

Finally a good number of comments in addition to noting about the effects of disabling conditions, emphasized expressions of empathy: one student described his experience with the buttons and gloves as "enlightening" in the sense of gaining some "understanding of what my [arthritic] grandad experiences" while another commented with regard to the different glasses, that in his group, some fellow students had vision problems, and that this part of the exercise enabled a "deeper understanding of how our groupmates normally struggle through day to day life". This cultivation of empathy is considered of great value by the authors (one of whom was the organizer for this section of the course for the students) because it demonstrates a shift in awareness about the needs of older and disabled people and hopefully corresponding changes in attitude and behaviour in the approach of younger people to the design and development of products and services for these groups.

5 Discussion and Conclusions

The low-cost simulation kit was easy to put together, but only provides a small range of experiences of visual impairment and ageing. However, it would be easy to extend the types of experiences provided with further low-cost kit and some imagination. It was clear from running the practical exercise that most students were engaged with the exercise, they talked very animatedly, laughed quite a lot and asked many pertinent questions of the lecturer and the graduate teaching assistants who helped to run the exercise. Some of the comments indicated that they understood that this was an imperfect simulation of living with a disability or ageing, but it would have been good to have specifically asked a question about this aspect of the exercise. However, the analysis of the open-ended questions reveals that students gained a number of different insights from participating, so we plan to continue using the exercise and refining it in the future. A long-term follow-up with this cohort of students is also planned for the end of their second year of studying computer science, to see whether the exercises have a lasting effect on their attitudes to older and disabled people and the importance of designing technologies suitable for them.

Acknowledgments. We would like to thank all the students who participated with the exercise and the graduate student assistants, Francesca Foffano and Berk Ozturk.

Disclosure of Interests. The authors declare that they have no competing interests.

References

1. Ben-Harush, A., Shiovitz-Ezra, S., Doron, I., et al.: Ageism among physicians, nurses, and social workers: findings from a qualitative study. Eur. J. Ageing **14**, 39–48 (2017)
2. Brodmerkel, S., Barker, R.: Hitting the 'glass wall': investigating everyday ageism in the advertising industry. Sociol. Rev. **67**(6), 1383–1399 (2019)
3. Chen, Y., Petrie, H.: Ageism and sexism amongst young technicians and older people in China. In: Miesenberger, K., Kouroupetroglou, G., Mavrou, K., Manduchi, R., Covarrubias Rodriguez, M., Penáz, P. (eds.) Computers Helping People with Special Needs. ICCHP-AAATE 2022. LNCS, vol. 13342. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-08645-8_60
4. Dobrowolska, B., Jędrzejkiewicz, B., Pilewska-Kozak, A., et al.: Age discrimination in healthcare institutions perceived by seniors and students. Nurs. Ethics **26**(2), 443–459 (2019)
5. Flower, A., Burns, M.K., Bottsford-Miller, N.A.: Meta-analysis of disability simulation research. Remedial Spec. Educ. **28**(2), 72–79 (2007)
6. Hallewell Haslwanter, J.D., Takacs, C., Gaisch. M.: How age and gender affect the opinions of computing students regarding computer usage and design needs. Informatics **9**(3), 52 (2022)
7. Lagacé, M., Van de Beeck, L., Firzly, N.: Building on intergenerational climate to counter ageism in the workplace? A cross-organizational study. J. Intergenerational Relat. **17**(2), 201–219 (2019)
8. Leo, J., Goodwin, D.: Pedagogical reflections on the use of disability simulations in higher education. J. Teach. Phys. Educ. **33**(2), 460–472 (2013)
9. Nelson, E.E.C., Spence, A.D., Gormley, G.J.: Stepping into the shoes of older people: a scoping review of simulating ageing experiences for healthcare professional students. Age Ageing **52**(12), afad235 (2023)
10. Nicolle, C.A., Maguire, M.C.: Empathic modelling in teaching design for all. In: Proceedings of International conference on human-computer interaction; Universal access in HCI: Inclusive Design in the Information Society, Routledge (2003)
11. Petrie, H.: Ageism and Sexism Amongst Young Computer Scientists. In: Miesenberger, K., Kouroupetroglou, G. (eds.) ICCHP 2018. LNCS, vol. 10897, pp. 421–425. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-94274-2_60
12. de São José J.M.S., Amado C.A.F.: On studying ageism in long-term care: a systematic review of the literature. Int. Psychogeriatrics. **29**(3), 373–387 (2017)
13. Sniatecki, J.L., Perry, H.B., Snell, L.J.: Faculty attitudes and knowledge regarding college students with disabilities. J. Postsecondary Educ. Disabil. **28**(3), 259–275 (2015)
14. VanPuymbrouck, L., Friedman, C., Feldner, H.: Explicit and implicit disability attitudes of healthcare providers. Rehabil. Psychol. **65**(2), 101–112 (2020)
15. Wandke, W., Sengpiel, M., Sönksen, M.: Myths about older people's use of information and communication technology. Gerontology **58**, 564–570 (2016)
16. Zheng, Q., Tian, Q., Hao, C., et al.: Comparison of attitudes toward disability and people with disability among caregivers, the public, and people with disability: findings from a cross-sectional survey. BMC Public Health **16**, 1024 (2016)

Making Entertainment Content More Inclusive

Making Entertainment Content More Inclusive Introduction to Special Thematic Session

Deborah Fels[1] (✉), Rumi Hiraga[2], and Yuhki Shiraishi[2]

[1] Toronto Metropolitan University, Toronto, ON M5B 2K3, Canada
dfels@torontomu.ca

[2] Tsukuba University of Technology, Ibaraki, Japan

Entertainment provides a means to social, cultural and well-being values and factors that are necessary in human societies. Being able to participate in entertainment is crucial to being an active and knowledgeable member of society. Without access, individuals can be left out of social discussions and growth that is dependent on knowing about these values and factors. For examples, children that cannot play the latest video game will be unable to participate in discussions, forums, new ideas and actions that arise from knowing about the game. This can mean that they will be at a disadvantage in social acceptance and peer interactions that are important for social growth. By proving technological solutions and interaction diversity that goes beyond standard input/output techniques, researchers are making important contributions to not only technological innovation but also to the social and cultural integration of individuals who use them.

There are six papers/presentations in this session that have different foci and findings but all have an entertainment context including access to video, mobile and physical games, and live sports and cultural events. There is also a diverse range of users participating in the research such as people who are blind or have low vision, are Deaf or hard of hearing, have upper extremity motor disabilities, and people who are neuro-diverse.

Different inclusion strategies and technologies such as vibrotactile, auditory and multimedia output/displays for multiple types of users as well as novel computer input strategies such as facial gesture and speech recognition, and eye tracking have been developed to provide better access to all of these different types of entertainment applications. The first paper in the session focuses on access to the physical game of darts for people who are blind/have low vision (B/LV). In this paper, a Bluetooth Lite enabled dart board provides the backdrop for an application develop specifically to support B/LV players by speaking a loud where a dart landed on the board, what the point value is and the remaining points needed to win. It keeps track of the scores for up to four players at a time. Vibrotactile support in the form of a 3-D printed, small diameter model of the dart board and a horizontal board with duplicate darts thrown by the player that could be used to tactilely access the dart locations. The researchers trained five B/LV participants on how to play darts and then played games together. Their main findings indicate that B/LV enjoy the game and find it relatively easy to play. They would like better support for aiming their darts before throwing them.

The second paper is a novel input strategy using facial recognition for controlling mobile games by people with upper extremity motor impairments such as those with cerebral palsy as touch screen are often not accessible to these individuals. In this paper,

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 111–112, 2024.
https://doi.org/10.1007/978-3-031-62846-7_13

a detailed description of the facial recognition technique is provided including landmark detection on the face, feature extraction and selection, and classification. The user study involved characterizing 12 real facial expressions from 120 videos in order to populate a facial expression database for use with target users. They obtained a 98% accuracy with a distance threshold of .32.

The third paper describes the process of developing and evaluating an action puzzle game, Tactris, similar to the game of Tetris™ for players who are B/LV. In this game, auditory and tactile output is used to identify the orientation, shape and location of that are dropping for the top of the screen. The objective of the game is to orient differently shaped 2-D blocks so that they form a straight horizontal line with the blocks that have already been placed at the bottom of the screen. Players must move the blocks (rotate the block and move is along the x-axis) to the best-fit location along the horizontal axis to make a flat surface. Sounds are used to indicate that the block is falling, being rotated, or lost and the speed of a block. It is also used to speak aloud the current score and the final score once the game is finished. The tactile display consists of a 48X32 matrix of vertically moving pins. The user study involves six B/LV players who played the game for 3-min. Players report that the system is challenging to use but that the game was enjoyable and they wanted to continue playing it.

The forth paper describes a second type of vibrotactile system, called VIBES, that provides access to the onset of lyrics for D/HoH Karaoke singers. The Apple iPhone's CoreHaptics system that consists of a rotating motor is used to produce the vibrotactile stimuli for system. VIBES was used in two conditions, sound only and sound plus vibration with four different song clips of varying tempo, singer and instrumentation. The main findings suggest that vibrotactile stimuli can improve the detection of song lyrics syllables onset but there is a wide variability that seems to be dependent on the type of song, familiarity with the song from past experience, and the likelihood of interference from other vibrotactile signals produced by the instruments used in a song.

The fifth paper describes a multi-modal, multi-user type, web-based application, called ISeee, designed to provide real-time access to live sporting and cultural events where there is no announcing provided, such as university level sports. It is a timeline-based model where participating users can contribute text, audio, image or video posts about the game in real time. There is also language translation provided by ISeee so that different language users can also participate in real time. Two user studies are described: 1) a museum event, and 2) a live handball game. Six D/HoH and five hearing users visit a museum and participate in a live discussion and description of the museum artifacts. For handball, three B/LV, D/HoH, and four sighted/hearing users watch the game live, produce text descriptions of the game action and score, and collaborate in a question and answer style of conversation while watching it. The System Usability Score (SUS) for the handball study is reported to be above 70 indicated there is good usability. However, the SUS for the museum visit is 55 indicating that that system is not easy to use and need improvement for D/HoH users.

The sixth paper describes the use of eye tracking as an input strategy for game play by people who are neurodiverse. One existing issue with using eye tracking for user input is the need to use dwell time as a selection strategy which results in unwanted or missed selections. Instead, the system users a keyboard activation as the selection

TARGET: Tactile-Audio DaRts for Guiding Enabled Throwing

Kei Kugai[1], Deborah I. Fels[2], and Makoto Kobayashi[1(✉)]

[1] National University Corporation Tsukuba University of Technology, Tsukuba City, Japan
koba@cs.k.tsukuba-tech.ac.jp
[2] Toronto Metropolitan University, Toronto, Canada

Abstract. The game of darts is a popular, simple to learn and difficult to master, social game. The TARGET tool was developed to support access to a BLE-enabled electronic dartboard for people who are Blind/Low Vision. Functionality included aurally announcing the score, current thrown position and remaining score to win. Human researchers tapped a metal rod on the dartboard to assist with aiming the dart, and provided advice on what possible areas on the board should be targeted. Tactile displays consisted of a small diameter replica of the dartboard that could be held in one hand and a second dart board on a horizontal surface that duplicated the dart positions that could be felt by the user. A user study with five blind participants was conducted to learn the game rules, practice throwing the darts and then playing an actual game. Results indicated that blind players enjoyed the game and were willing to play again, were in flow and developed a somewhat high degree of perceived competency over the duration of the study. Future work involves using machine learning technology to improve aiming support.

Keywords: physical games for people who are blind/low vision · darts · self-determination theory

1 Introduction and State of the Art

The modern and standardized game of darts has been a popular social game since 1896 [1]. The method of play and the rules are fairly simple and easy to understand but the game can be difficult to master because of physical accuracy, aim and strategic elements required of players. The dartboard layout contains 20 equally spaced triangular-shaped slices. Each slice contains four segments divided by metal wire, two larger segments, and a smaller inner (closest to the centre) and outer segment that are spaced between the two larger segments. Points are allocated for each segment based on the number assigned to the slice and the segment where the dart lands (see [2]) for an explanation for the dart play and game types). The slice order around the dartboard is standard but not sequential (see Fig. 1).

There are numerous versions of darts with different rules but the most common are the “Zero-One” games [2]. In this type of game, the player begins at numbers such as “301”, “501”, etc. The objective of Zero-One games is to score exactly 0 from either

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 113–119, 2024.
https://doi.org/10.1007/978-3-031-62846-7_14

301 or 501 points [2]. Throwing the dart at segments on the board results in the value of those segments being subtracted from the previous score. The player throws three darts in one turn and the total is then subtracted from the previous score. For example, if the game begins at 301 and a player's darts land on a segment representing the triple 20 (worth 60 points), the 10 segment and the 13 segment, their new score is 301 – (60 + 10 + 13) or 218. If the player's score is 1 or lower than 0 their score is reset to their previous value before that dart was thrown and their turn ends. The next player than throws their three darts and the cycle repeats until a player's score reaches "0".

In order to play darts successfully, then, a player must master two main aspects of the game: 1) aim; and 2) play strategy. Aim involves determining where to throw the dart and with what force/angle in order to achieve the desired points outcome. Where to throw the darts relates to having a complete understanding or mental model of the board layout as described, and holding the dart in the correct position, which [3] outlines should be at 90° to the floor located near the head position at eye level. The dart should be thrown by engaging only the lower arm at the elbow without moving the upper arm and body much [3]. The force used to ensure that the dart stays on the dartboard, rather than falling short or missing the board altogether, requires practice with feedback.

Mastro et al., (2017) [4] asked blind dart learners to learn dart throwing and aiming activities along with sighted experts. Sounds were activated on the dartboard to guide a blind participants' aim, "throw to the tone," and feedback from a sighted person was provided to assist in corrective actions (e.g., move to the right). However, they did not assess dartboard usability, player's engagement in the game or interaction with other players.

The second aspect of darts is play strategy, which refers to the decision's players must make in order to successfully achieve the 0 score before the other player. This involves calculating the points they need based on their performance so far. An example strategy is to hit the highest number of points (e.g., triple-20) at the early stages of the game to reduce the score quickly (e.g., by 60 points at a time). However, just hitting the triple-20 will eventually result in a score of 1 (three triple-20s in the first three throws for the first turn, and a triple-20 for the next two throws of the next turn) which then resets the score to the previous throw, causing an infinite loop that will never complete the game. In this example, and assuming the player can hit the triple 60 successfully in four throws, their score will be 61. The player must then decide what other set of points that they can hit successfully in the remaining two throws of their second turn to reduce the score to 0. For example, hitting a triple-19 (57 points) plus hitting one of the 4-point areas will result in 0, or hitting a triple-18 (54 points) plus hitting one of the 7-point areas will also result in a 0. There are many possible combinations of a two dart throw that could result in a 0 score from 61. Using this strategy, a player could finish the game in six throws (or two turns). Throwing darts consistently and accurately each time is difficult to master and part of what makes the game challenging and competitive [1]. Players must adjust their scoring strategy, using some mental arithmetic, as they hit or miss their target points so that they can reach 0 before the other player.

Self-determination theory suggests that being intrinsically motivated to accomplish tasks requires competency, autonomy and relatedness [5]. The game of darts offers opportunities to achieve all of these factors. An important aspect of being motivated in the game is to develop competency. For people who are blind/low vision (B/LV) learning the board layout and how to throw darts accurately and intentionally to achieve specific points/scores requires time and feedback. The balance between feedback from a sighted player and audio/tactile feedback that can be provided through software in order to gain competency and build a mental model that represents the dartboard remains to be studied.

In order to play darts successfully, then, a player must master aiming and play strategy. Throwing darts consistently and accurately each time is difficult to master and part of what makes the game challenging and competitive [1]. An accessible dart game for people who are B/LV must then support the learning and mastery of aiming and play strategy. We have created an accessible dart version of a bluetooth low energy (BLE) enabled electronic commercial dart board, called GRANBOARD™, in order to examine how B/LV players learn to play the physical game of darts. As such the research questions for an accessible dart game are: 1) What is the usability of the accessible dartboard? 2) How engaging is playing darts using the accessible dartboard? and 3) What is the impact on perceived competence of blind players learning to play darts? In this paper, an accessible dartboard application, called TARGET, and a user study designed to assess the usability, flow, preferences and perceived competency of novice blind dart players are described.

2 Dartboard Hardware/Software Configuration

The GRANBOARD produces unique codes for each individual segment on the board using BLE signals (see https://www.cs.k.tsukuba-tech.ac.jp/labo/koba/2022/11/ble-notify/ for the full listing of codes). The codes are output as 8 or 10 alphanumeric characters (e.g., 332e3540 or 31302e3640) depending on whether the code is below or above 100 respectively. Extracting the second and sixth character of the 8-character block (35 for the 8-character example) and the second, fourth and eighth character for the 10-character block (106 for the 10-character example) provides the unique location code where the dart has landed. These codes, their numeric score equivalent and words associated with each location were listed in a json file (e.g., “35”:[‘outside’, 20, 20, “bottom”]).

The TARGET software was developed to: 1) detect the BLE signals indicating position on the dartboard; 2) speak aloud that position; 3) track and sum the associated values from the dart positions; 4) track the number of throws; and 5) determine the person’s current score and the remaining value needed to win the game based on the three throws of the player’s turn and then speak those values aloud. The software also tracked which player was throwing for up to four players. After the three darts were thrown, a button on the dartboard is pressed and the player’s turn is complete.

Once a dart lands on a segment, the location and value are also spoken. TARGET then says the player’s new score and indicates that it is the next player’s turn. TARGET can track up to four players. See Fig. 1 for a system diagram.

3 Methodology

Following ethics approval, five novice dart learners (two male and three female) were recruited from a department of computer science for B/LV students. They were invited to complete an 8-question pre-study questionnaire that asked demographic questions and

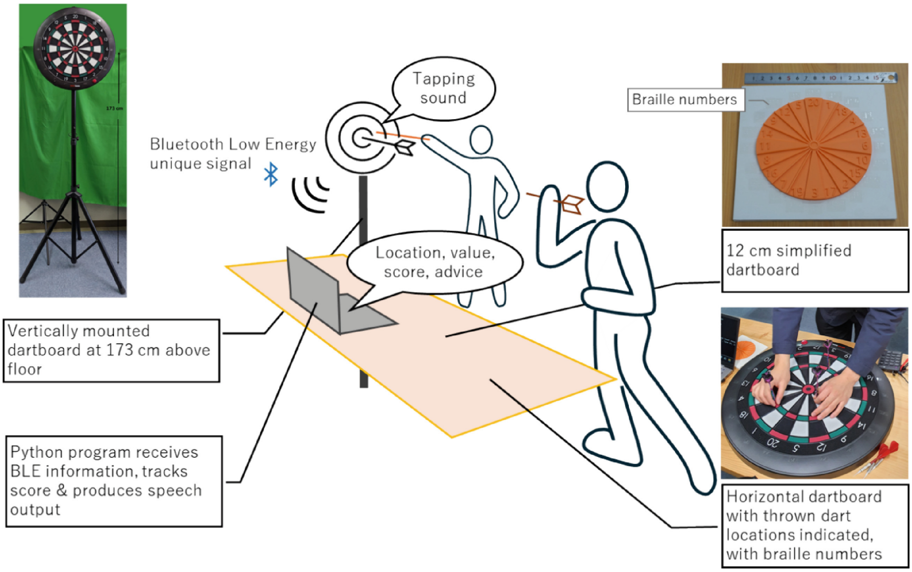


Fig. 1. TARGET system diagram

dart/physical game experience. Participants were provided with eye protection when throwing darts once they completed the pre-study questionnaire. Next, the board layout, how to throw a dart and the game were explained using three dartboards, one mounted at the correct height and location for the game, one 3-d printed 12 cm diameter board for immediate referencing while playing, and one placed horizontally at table height (see Fig. 1). Participants explored all versions of the dartboard tactically. Slice numbers were raised and braille equivalents placed alongside these numbers. Participants were then asked to practice throwing darts for about 10 min with the mounted dart board beginning at about 100 cm, 150 cm and 200 cm horizontally away from the board. Verbal feedback about location and throwing stance was provided by the researcher. Participants also felt their dart's position on the board. Once the participant gained sufficient familiarity with throwing the dart, they were then asked to try and aim for specific areas. The researcher tapped a metal rod on the various segments along with verbal instructions to provide participants with specific slice and segment locations. Participants were then invited to play a game with a novice sighted player who used their non-dominant hand and was not blindfolded. Participants were provided with continuous verbal feedback and help as well as access to tactile feedback using dart placement on the horizontal dartboard and the mini-dartboard. Once the game was completed, the participant was asked to complete a 32 question questionnaire (10 questions from the System Usability Scale, [6]; 10 flow questions [7]; 5 game engagement questions; 5 perceived competency questions [8]; and 2 questions regarding keeping track of the dart location and score; and participate in a 9-question interview regarding preferences and interest in the game. All questions used a 5-point disagree/agree Likert scale where 1 was strongly disagree and 5 was strongly agree.

4 Results and Discussion

The SUS has a mean of 71.5 (SD = 3.79) where a SUS score above 68 is considered above average usability [6]. Comments from the interviews suggested that although a majority thought the game was "easy to enjoy" (P5), some participants thought that a sighted partner would be necessary for providing feedback and collecting darts that fell on the floor, "difficult for totally blind users alone" (P1). For Audio Darts, sighted players are responsible for picking up darts [9]. Sound emitting darts may alleviate this need although mounting additional hardware on a dart will change its shape and weight which may, in turn, affect its throwability properties.

After normalizing the flow data, the highest flow score was 47 and the lowest was 40 out of a total possible score of 50 (M = 43.8, SD = 2.32). These high flow scores indicate that the participants were in flow during gameplay (see Fig. 2a). Comments in the interview suggest that flow could be attributed to the level of challenge and effort required by the game. For example, there were concerns expressed about throwing actions that may have affected flow "it was difficult to throw the darts to the target place" (P3). But, participants suggested that the "difficulties made the game fun" (P4), and that the thinking required in the game also contributed to game enjoyment, "It is fun to play because of using my brain" (P2).

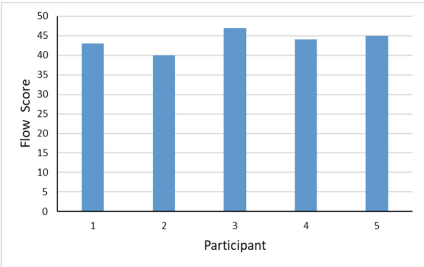


Fig. 2a. Flow scores for each participant

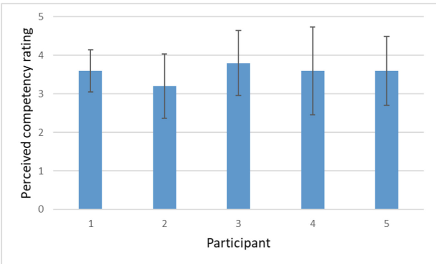


Fig. 2b. Mean and standard deviation of perceived competence rating for each participant

Williams et al. (1998) [8] recommends averaging each person's score to assess perceived competence. Average competency scores thus ranged from 3.2 (SD = .84) to 3.8

(SD = .84) with the maximum rating being 5 or strongly agree (see Fig. 2b). Perceived competence, which follows from flow and engagement [10], showed lower ratings than either flow or usability. While there is insufficient data to carry out statistical correlations, the flow and perceived competence scores seemed to follow a similar pattern as those in the literature (e.g., [10]). For example, the participant with the lowest flow score also had the lowest perceived competence score. Participant's comments related to their perceived competency in learning to play darts mostly related to the difficulty in throwing and aiming the dart "Difficult to aim" (P2) and "difficult to understand the sensation of throwing" (P4). However, they appreciated having the small tactile model that could be used while preparing to throw a dart because they did not need to change their body position. They also suggested several solutions to this difficulty including improving the small tactile model with double and triple areas demarcated, and having light or sound-based technologies that provide aiming information from the players and bullseye point of view.

Game engagement was only measured with five characteristics including fun, boredom, failure, and perception of importance of the game. All participants unanimously rated fun as very high (Mdn = 5, IQR = 0) and boredom as very low (Mdn = 1, IQR = 0). They rated their concern of failure as low (Mdn = 2 or disagree, IQR = 0). They also thought that the game was important to them (Mdn = 4 or agree, IQR = 1). They found that the social aspect of playing with others was enjoyable "It can be enjoyed by multiple players." (P4), and that it did not require physical strength to play.

Finally, there were two questions specifically related to the feedback provided by the software and the tactile representation of the board regarding knowing where a dart landed and the resulting changes in a participant's score. For both questions, participants agreed or strongly agreed that they knew where their dart landed and what their score was (Mdn 5, IQR = 0 and Mdn = 4, IQR = 1 respectively). However, they did mention that when there was much talking or noise in the space (from players cheering and encouragement), it was more difficult to hear the computer feedback indicating where their dart landed. The first participant mentioned that it would be useful to have the dart location duplicated on the horizontally placed board, which was implemented for the remaining participants. Participants acknowledged that having the darts placed on the horizontal board (see Fig. 1) helped them determine their dart positions on the vertical board "It was useful that I could confirm the position with the real dartboard on the table" (P5).

While there were many limitations in this study including a small number of participants, limited time for play, and manually generated aiming feedback, participants were in flow, had a fairly high perceived competency about playing darts likely because the software's features and functions met most of their needs for learning and playing darts. They became confident and competent to play a game with another player after about 40 min of learning and practice using multiple tactile and audio strategies. All participants wanted to play more games even though aiming was challenging. Future work could involve AI-based image processing to analyze hand/arm positions relative to the dartboard and to provide repositioning advice to assist with aiming.

Acknowledgments. We would like to thank all of the participants in the user study. Sander Fels-Leung assisted in developing an initial prototype of TARGET. Funding was generously provided by JSPS KAKENHI Grant Number 23K02587.

References

1. Shaun: Why is Darts So Popular?, https://darthelp.com/articles/why-is-darts-so-popular/. Accessed 23 Jan 2024
2. McNally, W., Walters, P., Vats, K., Wong, A., McPhee, J.: DeepDarts: modeling keypoints as objects for automatic scorekeeping in darts using a single camera. In: Presented at the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2021)
3. Kramer, A.: The Ultimate Book of Darts: A Complete Guide to Games, Gear, Terms, and Rules. Simon and Schuster (2013)
4. Mastro, J.V., Johnston, P.K., Coggins, P.E.: Darters: who needs vision? J. Phys. Educ. Recreation Dance **88**, 37–43 (2017). https://doi.org/10.1080/07303084.2017.1340204
5. Ryan, R.M., Deci, E.L.: Overview of self-determination theory: an organismic dialectical perspective. Handbook Self-Determination Res. **2**, 3–33 (2002)
6. Brooke, J.: SUS: a "quick and dirty" usability scale. In: Usability evaluation in industry, pp. 189–194. Taylor Francis (1996)
7. Engeser, S., Rheinberg, F.: Flow, performance and moderators of challenge-skill balance. Motiv. Emot. **32**, 158–172 (2008). https://doi.org/10.1007/s11031-008-9102-4
8. Williams, G.C., Freedman, Z.R., Deci, E.L., Leone, J.: Perceived competence scales. Diabetes Care **21**, 1644–1651 (1998)
9. Wig, A.: Bullseyes for the blind at the Richfield VFW (2014). https://www.hometownsource.com/sun_current/news/local/bullseyes-for-the-blind-at-the-richfield-vfw/article_2a4eaa58-8f59-52d9-a2f8-6d2a3ac7e053.html
10. Kam, H.-J., Menard, P., Ormond, D., Crossler, R.E.: Cultivating cybersecurity learning: an integration of self-determination and flow. Comput. Secur. **96**, 101875 (2020). https://doi.org/10.1016/j.cose.2020.101875

Personalized Facial Gesture Recognition for Accessible Mobile Gaming

Matteo Manzoni, Dragan Ahmetovic(✉), and Sergio Mascetti

Department of Computer Science, Università degli Studi di Milano, Milan, Italy
matteo.manzoni2@studenti.unimi.it,
{dragan.ahmetovic,sergio.mascetti}@unimi.it

Abstract. For people with upper extremity motor impairments, interaction with mobile devices is challenging because it relies on the use of the touchscreen. Existing assistive solutions replace inaccessible touchscreen interactions with sequences of simpler and accessible ones. However, the resulting sequence takes longer to perform than the original interaction, and therefore it is unsuitable for mobile video games. In this paper, we expand our prior work on accessible interaction substitutions for video games with a new interaction modality: using facial gestures. Our approach allows users to play existing mobile video games using custom facial gestures. The gestures are defined by each user according to their own needs, and the system is trained with a small number of face gesture samples collected from the user. The recorded gestures are then mapped to the touchscreen interactions required to play a target game. Each interaction corresponds to a single face gesture, making this approach suitable for the interaction with video games. We describe the facial gesture recognition pipeline, motivating the implementation choices through preliminary experiments conducted on example videos of face gestures collected by one user without impairments. Preliminary results show that an accurate classification of facial gestures (97%) is possible even with as few as 5 samples of the user.

Keywords: Upper extremity motor impairments · Mobile devices · Video games · Face gestures recognition

1 Introduction

For people with Upper Extremity Motor Impairments (UEMI), the interaction with mobile devices is challenging because it largely relies on the use of the touchscreen interface and therefore on the manual ability of the user [10]. Specific challenges with touchscreen use may also vary based on the actual condition of each user with UEMI. Indeed, some conditions cause difficulties in performing or sustaining precise movements (*e.g.*, cerebral palsy). Other conditions may impair hand strength, and some users may not have any mobility in upper extremities.

Assistive technologies that replace touchscreen interactions with sequences of simpler, more accessible ones have been proposed. However, these sequential

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 120–127, 2024.
https://doi.org/10.1007/978-3-031-62846-7_15

interactions are slow and, therefore, not suitable for time-constrained interaction (*e.g.*, games). In our previous work [3], we propose one-to-one remapping of touchscreen interactions to alternative inputs as a way to enable accessibility of existing games by people with UEMI. As a part of this prior work, we have conducted studies with people with UEMI on the use of external switches and vocal sounds with promising results. However, for those users who cannot access external switches and have a speech impairment (*e.g.*, anarthria), these interactions are still inaccessible.

We propose a new one-to-one interaction substitution method based on personalized Facial Gestures (FGs) recognition. To account for the specific needs of different users with UEMI, our approach relies on few-shot learning to allow the users to define and register their own FGs, with just a few samples of each gesture. The recorded FGs are then mapped to the touchscreen interactions needed to play a target game. Various existing mobile games can be used, including many popular ones that are available to users without disabilities.

In this work, we describe the FG recognition pipeline. In particular, we detail the processes of feature selection, few-shot learning, result aggregation, and fine-tuning. Preliminary experiments on videos of FGs collected by one user without UEMI yield a classification accuracy of 96.99% and the ability to process 8.26 ± 1.55 frames per second on a commodity Android device. As future work, we will conduct a thorough empirical evaluation with representative participants, focusing on confirming the applicability of the proposed approach and measuring its accuracy and appreciation by the target population.

2 Related Work

Mobile device accessibility for people with UEMI [10] is provided through accessibility services (ASs) [1], running in the background that completely replace the default touchscreen interaction paradigm, providing substitutive interactions for all mobile device functionalities. Scanning approaches [2] replace direct selection of a target User Interface (UI) element on the screen with sequential traversal of all UI elements, activating the target once it has been reached. Interactions may be provided through simplified touchscreen gestures, external switch peripherals [2], or FGs [15]. Direct activation of UI elements is also possible through gaze tracking [14] or verbal instructions [16]. However, all these approaches are slower than direct touchscreen access, which makes them unsuitable for mobile gaming [4]. To address this issue, in our previous work we have explored direct remapping of game UI elements to alternative user actions [3]. Specifically, we have proposed vocal interactions or external switch activations as alternative user actions, achieving comparable accuracy and reactivity with respect to touchscreen interaction. However, for users with UEMI who have a speech impairment and cannot access external switches (*e.g.*, cerebral palsy [5] or anarthria [9]), these interactions are still inaccessible. Hence, in this paper, we extend our previous approach [3] with a novel interaction modality based on FG recognition.

Prior literature identifies two key approaches for FG recognition [11]. One frequently used approach is to use deep convolutional neural networks to extract facial features and classify images into FGs. The other common approach first detects geometric facial landmarks, such as the position of the eyes, nose, and mouth. Then it extracts meaningful high-level features, such as distances between the landmarks and areas of polygons defined by sets of them. Finally, simpler machine learning models, like Support Vector Machines (SVMs), are used to classify these features into FGs. We highlight that both approaches are designed to classify a fixed set of predefined FGs, which are same for all the users. Furthermore, to train generalizable classifiers that robustly recognize FGs for different users, a large amount of data is required.

However, for some people with UEMI, none of these methods may be appropriate due to the characteristics of their specific motor impairment that may prevent them from making predefined FGs recognized by the classifier [15]. To address this issue, we propose the use of user-defined gestures. This approach requires user-specific training of the FG recognition model, using only a few examples of FGs that can be collected from a given user. To account for this requirement, our approach combines two machine-learning models: a pre-trained deep-learning model to extract facial landmarks [7] and prototypical networks [13], a few-shot learning approach suitable for FG classification.

3 Methodology

The proposed pipeline is divided into four main steps (see Fig. 1): Landmark detection and normalization (Sect. 3.1), Feature extraction and selection (Sect. 3.2), Classification (Sect. 3.3), and Post-processing (Sect. 3.4).

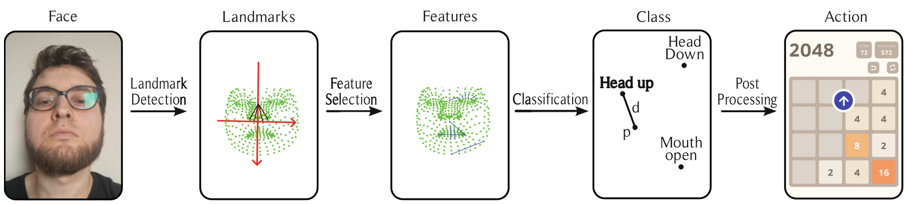


Fig. 1. Steps of the FG recognition pipeline

3.1 Landmark Detection and Normalization

To detect the face landmarks, we used a pre-trained network. Specifically, we used the *Mediapipe Face mesh* library [7] that is capable of real-time detection of 478 facial landmarks, whose position is determined in the image coordinate system. To robustly detect the FGs even if the user moves, the landmark coordinates are then made invariant with respect to the position of the face in the image or its distance from the camera using two approaches. First, we convert the landmarks into a face-centric coordinate system (the axes are the horizontal and the vertical lines shown in red in Fig. 1). Second, the landmark coordinates are normalized with respect to a set of reference distances between pairs of pre-defined landmarks (brown segments connecting the base and the tip of the nose). This set was obtained experimentally by selecting the pairs of points whose distances change the least in various FGs.

3.2 Feature Extraction and Selection

From the normalized landmarks, we extract a set of features to be used as input to the classifier. We considered different possible sets of features, among which we selected the one that yields the most accurate FG recognition (see Sect. 4.1). For our data, the final set of features included head rotation and distances between pairs of landmarks that are close to each other.

Since the detection runs on a resource-constrained device, we also apply the following feature selection procedure. At training time, the features are ordered according to their ability to discriminate different FGs using Fisher score [6]. Then, for each feature starting from the one with the highest score and iterating over the others, the technique computes correlation of that feature with all other features, removing the ones that have a correlation higher than a given threshold with the initial one. An example of features remaining after the selection is shown in Fig. 1 (blue segments above the left eye, on the mouth, and on the chin).

3.3 Classification

During the app setup, the user records a sequence of frames (5 in our tests) for each FG they intend to use to interact with a game. Features are extracted from these frames and used to train a Prototypical Networks model, a machine learning approach suitable for few-shot learning [13]. Specifically, a class prototype is defined for each considered FG using the extracted features as training data.

During inference, input video frames are classified using the trained network. Specifically, considering the features of the FG represented in the input frame, the closest prototype is set as its class. Additionally, we compute the distance value d between the input features and the closest prototype. Clearly, the smaller the value d, the higher the classification confidence with respect to the considered prototype. Based on this intuition, we empirically define three confidence levels (*High*, *Medium*, and *Low*) considering the distance d. If d is greater than the minimum distance between prototypes ($t = 0.8$ in our dataset), the FG is detected with *Low* confidence. Instead, to compute the threshold between *Medium* and *High* confidence, we examined how the classification accuracy varies considering different distance thresholds (see Sect. 4.2), finally selecting the value $T = 0.32$. Both thresholds can be tuned to the user-specific FG dataset.

3.4 Post-processing

The post-processing phase has three objectives: to smooth fluctuating classification results, to convert the classification results into actionable events (start/stop a tap) and to take into account the situation in which the user does not make any recorded FG (the "other" case). The proposed solution is based on two logical components: a filter and a finite-state machine. The filter takes in input the classification result and, if its confidence level is *high*, returns its class. If the confidence level is *medium*, the filter adds the class to a buffer. The buffer has a pre-defined size, which is a system parameter. When more than half of the buffer spaces contain the same class, the filter returns that class. If the confidence level is *low* a "other" class is added to the buffer. The finite state machine defines a state for each class, plus a "other" state. It takes in input the class returned by the filter and changes the state (if needed) to the corresponding class. Upon leaving/entering a state, the system triggers a *end tap*/*start tap* action. The only exception is the "other" state, which does not trigger any action.

4 Results

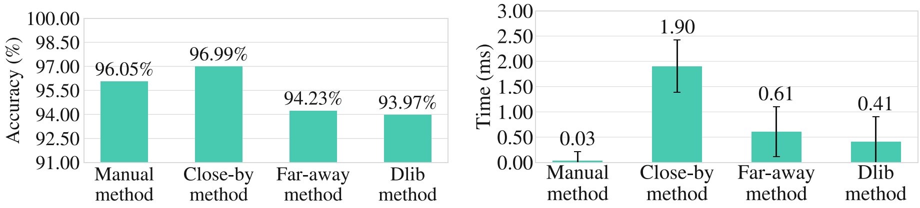
We conducted an extended set of performance tests aimed at tuning the system parameters. The experiments were carried out using a set of 120 videos, each representing a person making one of 12 representative FGs[1] (10 videos for each FG). The videos were collected by one user without UEMI.

4.1 Feature Extraction

We experimented with various sets of features, with the aim of balancing accuracy and the computational cost due to the large number of features. Specifically, we considered: manually selected features based on prior literature [12] including selected distances between landmarks, areas defined by sets of landmarks, and vertical inclinations of segments defined by pairs of landmarks (**Manual method**); distances between pairs of close-by landmarks (**Close-by method**); – distances between pairs of far-away landmarks (**Far-away method**); – distances between pairs of *Dlib* landmarks[2] (***Dlib* method**). Head rotations were considered as an additional feature in all cases. The best overall results (96.99% accuracy) were obtained considering close-by landmarks, with a set of 7098 features (see Fig. 2a), and a processing time overhead of $0.41ms$ (see Fig. 2b). The intuition for using close-by landmarks is that facial expressions are characterized by local changes (*e.g.*, the landmarks around the mouth when it is open or closed). Using these features and implementing the system on a *Samsung Galaxy A53 5G*, the system can process 8.26 ± 1.55 frames per second.

[1] FGs considered in the experiment: smile, open mouth, close left/right eye, curve eybrows, wrinkle nose, turn left/right, incline left/right, raise/lower head.

[2] Subset of *Mediapipe Face Mesh* [7] landmarks that are also defined in *Dlib* [8].



(a) Accuracy (b) Mean processing time

Fig. 2. Feature extraction accuracy and processing time by extraction method

4.2 Classification

Figure 3 shows how classification accuracy varies considering different values of the distance threshold T. For each threshold, the graph shows the percentage of images ("Support") whose distance d falls within that threshold and the accuracy achieved on those images. Based on these results, we select the threshold value $T = 0.32$. This value is data-specific, and can be tuned for each user.

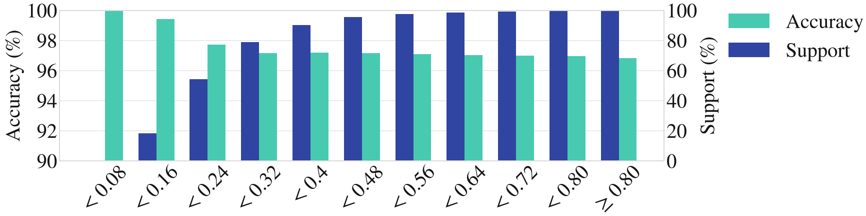


Fig. 3. Aggregate accuracy by intervals of d (distance from closest prototype)

5 Discussion

5.1 Detection Robustness, Personalization, and Generalizability

We designed the FG detection pipeline with three goals in mind: 1) accurate and timely interaction, suitable for mobile gaming, 2) personalization of FGs according to the user's needs, and 3) need for limited training data. The results obtained from the preliminary tests show that our approach satisfies all three criteria. By design, the proposed technique is robust to user movement and distance from the camera. It is able to process more than 8 frames per second on a commodity mobile device, with 97% accuracy. It enables a user to define personalized FGs and use them as interactions to play existing mobile video games. Finally, it only requires a short video of each FG for the training.

These results were computed on a preliminary dataset collected by one user without impairments. The aim was solely to assess the feasibility of the proposed technique and to tune the system parameters during the process. Thus, we cannot consider these results to be representative for users with UEMI. However, since the FGs are personalized and trained separately for each user, the process in itself should apply to users with UEMI without modifications.

5.2 Limitations

Despite the positive results, we acknowledge the technical and practical limitations of our approach. First, the applicability of the approach is limited by the ability to discern facial landmarks in the video frame, and therefore it is not possible in adverse luminosity conditions such as extreme dark or light glare. The presence of others in the camera frame, as well as occlusion of the facial features, may also influence the ability of the system to correctly detect the user's FGs.

Clearly, the applicability of the approach is also limited by the user's ability to make FGs. For users with limited mobility or difficulties in controlling their head and facial features (*e.g.*, people with dystonia), this approach is not suitable. In such cases, other input modalities could be used, such as non-verbal voice interaction or external switches [3].

Finally, the main methodological limitation of our work is the lack of an evaluation by users with UEMI. The overall interaction substitution approach, with other types of input (non-verbal voice input or external switches), has been evaluated with representative users [3]. However, to assess the generalizability of the FG interaction modality, additional user studies with representative users are needed to robustly assess the actual recognition accuracy of the technique.

6 Conclusions and Future Work

This paper describes a pipeline for recognizing personalized facial gestures, selected by the users themselves, to be used as a new interaction approach for making existing mobile games accessible to people with upper extremity motor impairments. After detecting facial landmarks [7] corresponding to a facial gesture, a set of robust features is selected and used to train a Prototypical Network model [13]. Only a short configuration step is required to register the personalized gestures and associate them to the desired game. Afterwards, the system will recognize the registered gestures and trigger the associated game interactions, thus enabling users with UEMI to play the configured games.

As future work we will assess the validity of the proposed approach with representative users, using existing mobile games. In particular, we will assess the accuracy of the approach, the reaction time of the users when using the new interaction modality, and the cognitive and physical load associated to its use. Furthermore we will assess its usability and appreciation by users with UEMI, as well as the applicability of the approach to users with different abilities.

References

1. Android accessibility services. https://developer.android.com/guide/topics/ui/accessibility/service
2. Android accessibility suite. https://play.google.com/store/apps/details?id=com.google.android.marvin.talkback
3. Ahmetovic, D., Riboli, D., Bernareggi, C., Mascetti, S.: RePlay: Touchscreen interaction substitution method for accessible gaming. In: Human Computer Interaction with Mobile Devices and Services. ACM (2021)
4. Bierre, K., Chetwynd, J., Ellis, B., Hinn, D.M., Ludi, S., Westin, T.: Game not over: Accessibility issues in video games. In: Universal Access in Human-Computer Interaction (2005)
5. Bishop, D.V., Brown, B.B., Robson, J.: The relationship between phoneme discrimination, speech production, and language comprehension in cerebral-palsied individuals. J. Speech Lang. Hear. Res. **33**(2), 210–9 (1990)
6. Gu, Q., Li, Z., Han, J.: Generalized fisher score for feature selection. In: Conference on Uncertainty in Artificial Intelligence (2011)
7. Kartynnik, Y., Ablavatski, A., Grishchenko, I., Grundmann, M.: Real-time facial surface geometry from monocular video on mobile GPUs (2019). arXiv preprint arXiv:1907.06724
8. King, D.E.: Dlib-ml: A machine learning toolkit. J. Mach. Learn. Res. **10**, 1755–1758 (2009)
9. Lecours, A., Lhermitte, F.: The “pure form” of the phonetic disintegration syndrome (pure anarthria); anatomo-clinical report of a historical case. Brain Lang. **3**(1), 88–113 (1976)
10. Naftali, M., Findlater, L.: Accessibility in context: understanding the truly mobile experience of smartphone users with motor impairments. In: Computers & Accessibility. ACM (2014)
11. Revina, I.M., Emmanuel, W.S.: A survey on human face expression recognition techniques. J. King Saud Univ.-Comput. Inf. Sci. **33**(6), 619–628 (2021)
12. Rozado, D., Niu, J., Lochner, M.: Fast human-computer interaction by combining gaze pointing and face gestures. Trans. Accessible Comput. **10**(3), 1–18 (2017)
13. Snell, J., Swersky, K., Zemel, R.: Prototypical networks for few-shot learning. In: Advances in Neural Information Processing Systems, vol. 30 (2017)
14. Zhang, X., Kulkarni, H., Morris, M.R.: Smartphone-based gaze gesture communication for people with motor disabilities. In: Human Factors in Computing Systems. ACM (2017)
15. Zhao, X., Fan, M., Han, T.: “i don’t want people to look at me differently” designing user-defined above-the-neck gestures for people with upper body motor impairments. In: Human Factors in Computing Systems (2022)
16. Zhong, Y., Raman, T., Burkhardt, C., Biadsy, F., Bigham, J.P.: JustSpeak: enabling universal voice control on android. In: Web for All. ACM (2014)

Tactris: Inclusive Falling Block Puzzle Game with Audio-Tactile Effects for Visually Impaired People

Masaki Matsuo[1(✉)], Delgerbayar Erdenesambuu[1], Masatsugu Sakajiri[1], Junji Onishi[1], and Takahiro Miura[2]

[1] Tsukuba University of Technology, Kasuga 4-12-7, Tsukuba 305-8521, Ibaraki, Japan
matsuo@cs.k.tsukuba-tech.ac.jp

[2] National Institute of Advanced Industrial Science and Technology (AIST), Kashiwanoha 6-2-3, Kashiwa 277-0882, Chiba, Japan

Abstract. The purpose of this study is to develop an action puzzle game that is easy to play for visually impaired people, and to expand the number of visually impaired people participating in e-sports. To this end, we developed a prototype of *Tactris*, a falling-puzzle game that can be played by using auditory and tactile information, and evaluated the ease of use and playability of the game for six visually impaired people. We have been studying the presentation method, operation method, and game rules necessary to realize a falling object puzzle game that can be played by visually impaired people, and have aimed to develop a falling object puzzle game that combines voice, sound effects, and tactile illustrations. We aimed to research and develop a puzzle game that presents the situation in the game by combining audio, sound effects, and tactile illustrations. Since this interface uses dynamically changing tactile diagrams, it is expected to be used as a content to efficiently promote understanding of tactile perception in education for the visually impaired. It will also help to improve the tactile skills of dynamic tactile diagrams. Six visually impaired people evaluated the ease of use and playability of the game, and although the ease of use remained an issue, there was a strong need to play the game repeatedly.

Keywords: Visual impaired · para e-sports · spatial cognition · audio-tactile cues · falling block puzzle games

1 Background

In recent years e-sports has been gaining momentum, making it easier for visually impaired gamers to become involved. Forza Motorsport now includes a screen reader and driving assist features; Street Fighter 6 includes a sound accessibility feature. This enables the blind and visually impaired who cannot see the screen to participate in e-sports activities. The information on computer games that can be played by visually impaired people is shared on websites, such as Audiogames [1].

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 128–136, 2024.
https://doi.org/10.1007/978-3-031-62846-7_16

Efforts are also underway to develop inclusive games that can be played by both visually impaired and sighted people. Matsuo et al. developed an accessible action role-playing game (RPG) that can be played by anyone regardless of visual condition. RPG was designed to be played using only auditory and tactile senses, with graphic features implemented so that able-bodied people can play the game [1, 2]. However, some game genres are not easily accessible to the visually impaired. The blind and visually impaired play computer games by associating sound effects with situations in the game or by memorising the order and hierarchical structure of items, such as menus, to indirectly grasp screen information. However, action puzzle games, such as falling object puzzles, are difficult for the visually impaired to play because they require the player to make appropriate decisions according to the situation, requiring visual grasp of the entire changing situation on the game screen, including the position and shape of falling blocks and positional relationship of the stacked blocks. This makes it difficult for visually impaired individuals to play the game.

2 Objective

The purpose of this study was to develop an action puzzle game that would be easy for the visually impaired to play, thereby increasing the participation of visually impaired in e-sports. To this end, we developed a prototype of Tactris, a falling-block puzzle game that can be played using auditory and tactile information, evaluating its usability and playability on six visually impaired participants.

We have been studying presentation and control methods, as well as the game rules necessary for realising a falling-object puzzle game accessible to the visually impaired, by combining voice, audio, sound effects, and tactile illustrations. Because this interface uses dynamically changing tactile diagrams, the content is expected to efficiently promote the understanding of tactile perception in education for the visually impaired, which would also help improve the dexterity of dynamic tactile diagrams.

Previous research on action puzzle games for the visually impaired included the Tetris Game for the Visual Impaired Utilising Sound, proposed by Choi et al. [4]. To enable the visually impaired to play the game, 1) the game rules were simplified, 2) pieces were limited to one type, and 3) sound was used to indicate a hole in the puzzle. Suggesting improvements, they stated that the game scenario was simple and could become boring after repeated play, and that the design of additional devices with visually impaired-friendly input was required. The Tactris game we are developing reproduces the complex rules of an action puzzle game. An audio-tactile presentation method is proposed to encourage repeated play to obtain higher scores. We also designed a user interface that combines the output with a tactile display and game controller to make it easy for the visually impaired to operate.

3 Tactris Development

We developed Tactris, a puzzle game played by touching the game screen. Table 1 presents the operating and developmental environments of the system. Figure 1 shows the game screen during play.

Table 1. Development and operating environment of Tactris

Item	Description
Development OS	Windows 11 Pro
Development language	Python 3.10.4
Libraries used	Pygame 2.1.2, pyserial 3.5, accessible-output20.17
System requirements	Windows 10 Pro, Windows 11 Pro
Supported screen readers	JAWS for Windows, NVDA, Window Eyes, PC Talker, Microsoft Speech API
Point diagram output device	KGS Dotview DV-2
Input devices	Keyboard, joystick
Recommended sound presentation device	Stereo speaker, stereo headphones

This game is based on the basic functions of Tetris, the original falling-object puzzle, and is designed to be played by visually impaired users. For the implementation, we referred to the chapter on falling object puzzles in "Python Programming for Fun Learning while Making Games" by Kenichiro Tanaka [5], using Python 3.10.4, pygame 2.1.2, pyserial 3.5, and accessible-output2 0.17 for the development. The text in the game was read by a screen reader. For the game operation, a game controller (JOY-CON, NINTENDO) was used in addition to the PC keyboard.

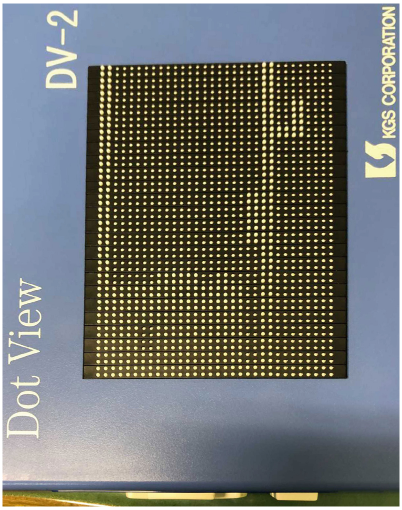


Fig. 1. Tactris play screen displayed on DV-2

The rules of the game consist of moving and rotating seven different types of tetrominoes falling from the top of the playfield to fill a horizontal row without gaps. The

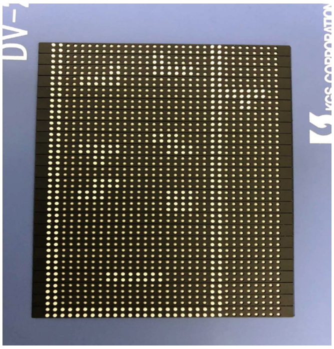


Fig. 2. Seven different types of tetrominoes

types of tetrominoes are shown in Fig. 2. The rows that are filled without gaps are erased while accumulating the score, and the block above the erased row falls. The falling speed of the tetrominoes is accelerated over time. The game is over when the blocks reach the top of the playing field. A DV-2 DotView was used to display tactile information in real time on a computer [6] instead of a PC monitor, The dot view has 48 horizontal pins and 32 vertical pins for a total of 1,536 pins and can present text and graphic information as binary information by moving the pins up and down. The user can play the game by touching the dot-dot display to grasp the shape and positional relationship of the falling tetrominoes and stacked blocks. In addition, sound effects and audio are used to assist in presenting the e-situation in the game. Sound effects for tetrominoes falling, moving, rotating, remaining fixed, and when blocks are lost are supplemental auditory information indicating situational changes. The audio reads the score information using a screen reader. In addition to reading out the increased score when a block disappears and the total score after the game is over, a dedicated button for reading the score can be placed at any time. The research and development efforts were conducted in accordance with an inclusive research methodology in which persons with disabilities were involved in the research process and in research planning, system implementation, evaluation, and improvement [7].

4 Tactris Evaluation

To investigate the usability of Tactris and obtain guidelines for future research and development, an evaluation experiment was conducted on six visually impaired individuals, aged 19 and 29 years, five males, one female, two blind, and four with low vision. Two of the blind participants and one of the low-vision participants had experience using Braille and Braille charts. All participants played computer games for at least two hours per day. Table 2 shows the information on the participants in the experiment.

4.1 Experimental Procedure

First, a questionnaire was administered to the participants asking about their individual disability status and experience using braille and computer games. They were then asked

Table 2. List of participants in the experiment

Participant ID	A	B	C	D	E	F
Gender	Male	Male	Male	Male	Male	Female
Age	20s	20s	20s	20s	20s	10s
Visual impairment status	low vision	low vision	totally blind	low vision	totally blind	low vision
Onset of visual impairment	congenital	congenital	~8 years	~14 years	congenital	congenital
Use of tactile chart per day in the last two weeks	None	None	None	None	None	None
Hours of tactile use per day during most frequent use	None	None	2 h	several times in the past	Used in class	2 h
Hours of game playing per day in the last two weeks	None	5 h	None	1 h	None	8 h
Hours of gameplay per day during the most playing period	2 h	all day	10 h	6 h	2–3 h	all day
Experience in playing Audiogame	Know and play	Know but no experience in playing	Know and play	Know but no experience in playing	Mostly play Audiogame	Know and play

to play Tactris with three patterns of different-sized tetrominoes for three minutes each and were evaluated through usability metric for user experience (UMUX-Lite) for each treatment [8, 9]. The results were converted into an estimated system usability scale (SUS) score based on a formula proposed in [9]. The block was set to fall one level every 0.7 s, increasing in speed every 10 s, and after 3 min, one level every 0.56 s. The

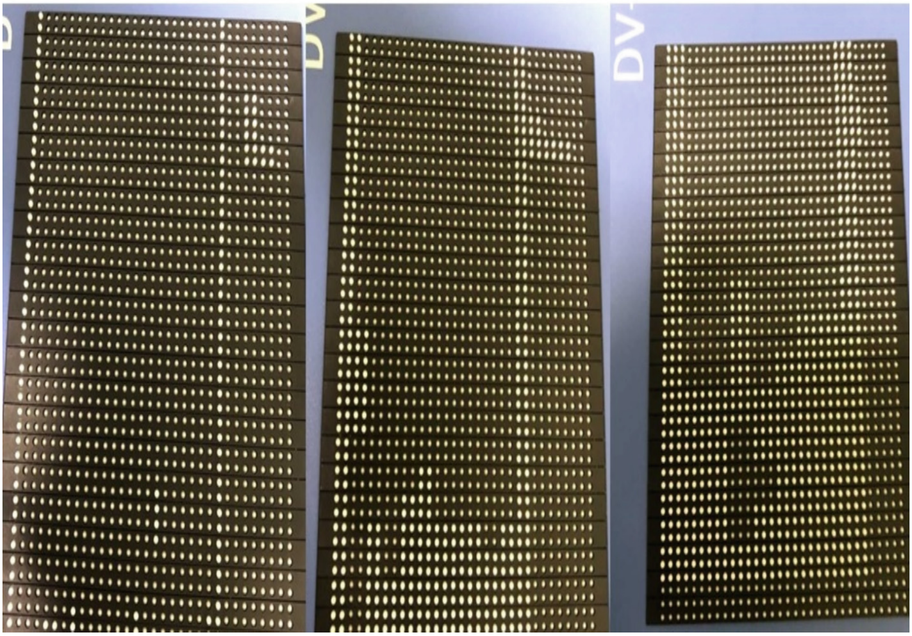


Fig. 3. Example of the three-pattern Tactris used in the experiment. From left to right, one block is represented by 1 × 1, 2 × 2 and 3 × 3

game was controlled using a Joy-Con(R) controller (Nintendo) attached to the bottom of the Dot View DV-2. Finally, a questionnaire was administered to evaluate the game. The respondents were asked to respond on a 7-point Likert scale regarding ease of understanding the game rules, operation, and in-game situation, as well as overall fun factors (Fig. 3).

4.2 Experimental Results and Discussion

The results of the subjective evaluation by the participants are shown in Table 3.

In Pattern 2, three out of six participants succeeded in erasing one level of block during the experiment. In Pattern 3, two out of the six participants succeeded in erasing the block. However, in Pattern 1, no erasures were made. The participants who succeeded in erasing the blocks and obtaining a score played the falling-puzzle game using visual information. This suggests that participants understood the in-game situation in the object-based puzzle game.

The mean SUS score for each treatment was 44.6 for Pattern 1, 60.8 for Pattern 2, and 57.2 for Pattern 3. Based on the above, Pattern 2, which had a block size of 2 × 2, was considered the easiest to use.

The playing time was limited to three minutes in this study, and none of the participants had touched the point diagram within two weeks. If the game were played for a longer period or if the point diagram were used regularly, it could have been rated

Table 3. Results of the evaluation by the participants of the experiment

Participant ID	A	B	C	D	E	F
Enforcement pattern 1 (block size 1 × 1)						
Number of erasures	0	0	0	0	0	0
UMUX-Lite Q1: This game meets my requirements	2	1	6	2	4	5
UMUX-Lite Q2: This game is easy to use	1	1	5	2	4	3
SUS score	28.3	22.9	71.7	33.7	55.4	55.4
Enforcement pattern 2 (block size 2 × 2)						
Number of erasures	0	1	0	0	1	0
UMUX-Lite Q1: This game meets my requirements	3	5	6	5	5	5
UMUX-Lite Q2: This game is easy to use	2	5	7	3	6	4
SUS score	39.2	66.2	82.5	55.4	71.7	50.0
Enforcement pattern 3 (block size 3 × 3)						
Number of erasures	1	0	0	0	0	0
UMUX-Lite Q1: This game meets my requirements	3	6	6	3	3	3
UMUX-Lite Q2: This game is easy to use	1	6	6	4	3	5
SUS score	33.7	77.1	77.1	50.0	44.6	60.8
Questionnaire						
Ease of understanding the rules	6	7	7	7	7	5
Ease of operation	3	7	6	7	7	7
Ease of understanding in-game situations	1	5	5	4	3	7
Interest	4	7	7	5	7	6

as more user-friendly. The speed and acceleration of the falling blocks were constant, although the size of the playfield varied. Therefore, there is a possibility that Patterns 1 and 3 can be improved by adjusting the falling speed to suit the playfield. In the results of the post-play questionnaire, one participant rated the overall fun factor, as 4, 5, and 6 with the remaining three rating it as 7, for an average score of 6. Some participants expressed a strong desire to continue playing the game after the experiment was over. Free feedback included comments such as "I am happy to be able to play again the game that I used to play when I was able to see"; "It was the first time I played the game while touching it, so it was a fresh experience"; "It is good because my eyes do not get tired"; and "The position where the block is dropped is often off by one square." In summary, although we did not obtain a high evaluation for the usability of the game by SUS, the game was found interesting and the users wanted to play it repeatedly.

Based on these improvements, the game interface was improved. Table 4 presents the results of these improvements. Three blind people involved in the development of the improved Tactris tested it and confirmed that the operation errors were reduced using the block fall-point guide. The maximum number of erasure blocks of Tactris at the pre-experimental stage was 6; however, after the system was improved, 126 blocks were successfully erased.

Table 4. Improvements to Tactris based on evaluation results

Addition of new rules	increase in fall speed every time 5 levels are erased
Customisation options	block size, button configurations, game speed
Block fall point guide	Predictive fall point always shown by blinking dot diagram
Voice reading of block shapes	reads out falling and next block
Braille display	Displays operation guide and total score in Braille
Cheering function	Voice cheers the player on when the score is added
Tutorial	Practice mode to feel the type of block and prevent blocks from falling
Easy-to-see screen	Colour scheme based on colour universal design

5 Summary

In this study, we developed a prototype of Tactris, a falling object puzzle game that can be played using auditory information, with the aim of expanding the number of visually impaired participants in e-sports. Six visually impaired people evaluated the ease of use and playability of the game. Although ease of use remained an issue, there was a strong need for the game to be played repeatedly. As a result of system improvements based on the evaluation results, the number of block elimination steps and scores of the visually impaired players improved significantly.

In the future, we plan to analyse the progress of Tactris and improve the interface to make it easier for the visually impaired to play. We are also considering applications in the field of education, such as using this game to acquire touchpoint diagram skills. Currently, Tactris cannot be played without a pointillistic display; therefore, issues remain to be addressed in terms of widespread use. A method for playing games on a general-purpose device using sound and vibration without a point-drawing display is being currently investigated.

Acknowledgments. This study was funded by KAKENHI Grant Numbers 23K16919, 23K17582 and 23H00996.

References

1. AudioGames.net. http://www.audiogames.net/. Accessed 02 Apr 2024
2. Matsuo, M., Miura, T., Onishi, J., Sakajiri, M., Ono, T.: Experience report of blind gamer to develop the accessible action RPG. J. Technol. Persons Disabil. **5**, 172–191 (2017)

3. Matsuo, M., et al.: Inclusive action game presenting real-time multimodal presentations for sighted and blind persons. In: Proceedings of ICMI '21. 2021.10 18–22, pp. 9. Montrial, Canada. ACM, New York, NY, USA
4. Jongmyung, C., Daehee, P., Rang, L., Seonhwa, K.: A Tetris Game for the Visual Impaired Utilizing Sound (2013)
5. Tanaka, K.: Python Programming for Fun Learning While Making Games, Impress R D, ASIN B06XRQS8PN (2017)
6. Dot View DV-2. https://www.kgs-jpn.co.jp/pub/cat/DV2.pdf. Accessed 02 Apr 2024
7. Nind Melanie: What is inclusive research? (2014). https://www.bloomsburycollections.com/book/what-is-inclusive-research/ch1-inclusive-research-defined
8. Brooke, J.: SUS : A Quick and Dirty Usability Scale, Usability Evaluation in Industry, Taylor & Francis, Vol. 1996, pp 189–194
9. Lewis, J.R., Utesch, B.S., Maher, D.E.: UMUX-LITE: When there's no time for the SUS. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 2099–2102. New York, NY: ACM (2013). Author, F.: Article title. Journal **2**(5), 99–110 (2016)

Inclusive Fighting with Mind's Eye: Case Study of a Fighting Game Playing with only Auditory Cues for Sighted and Blind Gamers

Masaki Matsuo[1], Naoya Kitamura[2], Misato Egashira[2], Teruo Hosogai[2], Shunya Hatakeyama[2], Daiki Kato[2], Ko Hayakawa[3], Junji Onishi[1], Masatsugu Sakajiri[1], and Takahiro Miura[1,4](✉)

[1] Faculty of Health Science, Tsukuba University of Technology, 4-12-7 Kasuga, Tsukuba, Ibaraki 305-0821, Japan
{matsuo,ohnishi,sakajiri}@cs.k.tsukuba-tech.ac.jp

[2] ePARA Inc., 1-14-8 Shimomae, Toda, Saitama 335-0016, Japan

[3] Research Center for Advanced Science and Technology, The University of Tokyo, 4-6-1 Komaba, Meguro, Tokyo 153-8904, Japan

[4] Human Augmentation Research Center (HARC), National Institute of Advanced Industrial Science and Technology (AIST), 6-2-3 Kashiwanoha, Kashiwa, Chiba 277-0882, Japan
miura-t@aist.go.jp

Abstract. Visually impaired people frequently find it difficult to enjoy most of these state-of-the-art games, despite the vast amount of accessibility research on content and interfaces. Meanwhile, a growing number of playable games regardless of visual impairment status have been released. However, there are still few games enabling both blind and sighted players to compete against each other. The *Street Fighter 6* (Capcom Co. Ltd.), released in 2023, introduced a sound accessibility feature for the first time in a commercial fighting action game. By clarifying the requirements for sighted and visually impaired players to compete smoothly using this feature, not only can they participate in the game regardless of their disability status, but also new accessible interfaces can be created by using the technology. In this study, the goal is to evaluate the playability of a fighting game with sound accessibility features for visually impaired and sighted groups. This paper reports on the evaluation results of usability and user experience of the sound accessibility features implemented in the *Street Fighter 6*, for two groups.

Keywords: Inclusive game · sound accessibility · visually impaired

1 Introduction

The computer gaming market has been growing worldwide and will expand rapidly in coming years; the Global Games Market Report 2020 estimated that the gaming population would reach 2.7 billion in 2020 and 3 billion in 2023 [2]. The population of computer game players presently encompasses a wider range

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 137–145, 2024.
https://doi.org/10.1007/978-3-031-62846-7_17

of age groups. Moreover, the global pandemic of COVID-19 (coronavirus disease 2019) has led to government *stay-at-home* orders and calls for social distancing in many countries [15], and the demand for computer games became soar.

Amid these situations, computer games have diversified due to advances in hardware performance and software capabilities. Irrespective of whether game platform is stationary or portable, the display density and resolution of screens are increasing, and the information that players need to grasp visually is becoming larger. However, this context has led to the problem that visually impaired people frequently find it difficult to enjoy most of these state-of-the-art games, despite the vast amount of accessibility research on content and interfaces [5,11,13,16,18]. Accordingly, volunteers who are enthusiastic gamers with visual impairments share information on playable and non-playable games on websites that reviews games [1,4]. The fact that such sites exist itself suggests there are only a small number of mainstream games that the visually impaired can play.

Meanwhile, a growing number of playable games regardless of visual impairment status have been released. The 2000s saw the popularity of audio games, and the 2010s saw the proliferation of exergames [12]. Since the mid-2010s, we have been developing inclusive multimodal action RPGs and side-scrolling action games playable regardless of visual impairment status [9]. We have also designed an accessible game IDEs (integrated development environments) operable with a screen reader, which allows for efficient creation of game views for blind developers [9]. In the 2020s, commercial games have been gradually transforming into more accessible ones [10]. In particular, *The Last of Us Part II*® (Naughty Dog, LLC) includes accessibility functions, making it easier for players with visual, auditory, and motor disabilities to play and immerse the game.

In this context, while the number of playable games regardless of disability is increasing, there are still few games enabling both blind and sighted players to compete against each other. The *Street Fighter 6*® (Capcom Co. Ltd.), released in June 2023, introduced a sound accessibility feature for the first time in a commercial fighting action game, based on the advice of a blind professional gamer [3]. By clarifying the requirements for sighted as well as visually impaired players to compete smoothly using this feature, not only can they participate in the game regardless of their disability status, but also new accessible interfaces can be created by using this technology.

In this study, the purpose is to evaluate the playability of a fighting game with sound accessibility features for both visually impaired and sighted groups. This paper reports on the evaluation results of usability and user experience of the sound accessibility implemented in the *Street Fighter 6*® for two groups.

2 Evaluation

2.1 Participants

Twenty sighted persons, seven total blind, and four low vision participated in this evaluation (Age: 31.6±11.8, Range: 9–66). In the sighted participants, there were fifteen men and five women (Age: 31.4±11.8, Range: 9–66). Six of the sighted did not usually play games, five played a few times a week, five played 4–5 times a week, and four played almost every day. In the visually impaired group, there were eight men and three women (Age: 30.8±10.8, Range: 19–54). None of the participants had hearing or physical disabilities.

2.2 Method

For the sighted participants, we exhibited *Street Fighter 6*®, a commercial fighting action game that implemented sound accessibility features for visually impaired players, at a booth where visitors could play the game without any visual information in a two-day game experience event. For the visually impaired, we recruited those who had this game installed in their personal computers with their customized accessibility settings.

The participants were first asked to practice associating visual information with auditory information on a training mode of the game. At that time, we asked the participants to shut off visual information occasionally so that they could pay attention to the audification information provided by the sound accessibility function in addition to the in-game sounds and sound effects. The audification information provided by the sound accessibility function included the distance from the opponent's character in terms of the interaural difference in sound, and the height information in terms of the sound volume. The participants who had their own game controllers or consoles were allowed to use them freely, while those who did not have a game controller were asked to use a PlayStation® controller to operate the game. They were also asked to choose either the classic control mode whose command operation is inherited from the former *Street Fighter*® series or the modern control mode that simplified the special attacks commands mainly for beginners. The opponent of the participants was a blind professional e-sports player who was proficient in sound accessibility. After the match, participants were asked to answer a questionnaire. The questionnaire consisted of the following scales, in addition to items about their basic status such as gameplay frequency, age, gender, and disability status:

- System usability scale (SUS) [7]. We used the overall SUS score and subscale SUS scores of usability and learnability [6]
- User experience questionnaire short version (UEQ-S) [14]. We calculated hedonic, pragmatic, and overall scores from the answers of UEQ-S.

They also answered the 5-point Likert scales of sound accessibility features shown in Table 1.

3 Results and Discussion

3.1 SUS and UEQ-S Scores: Sighted Versus Visually Impaired Participants

First, an analysis of variance (ANOVA) with ART (aligned rank transform) [17] of the overall SUS scores for visual impairment status and gameplay frequency

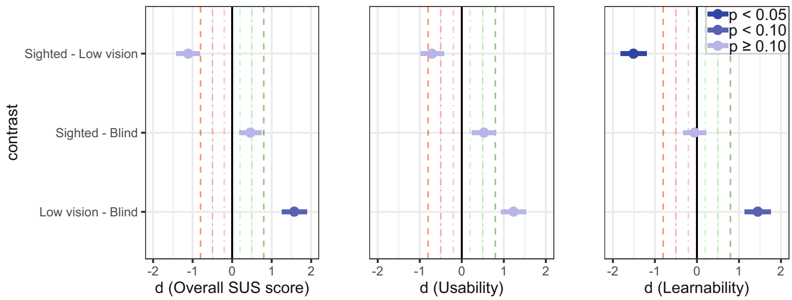


Fig. 1. Effect size of the contrast by vision status on (left) overall SUS, (center) usability, and (right) learnability scores

revealed a main effect with marginal significance for visual impairment status ($p = 0.075 < 0.10$, partial $\eta^2 = 0.24$) and a significant main effect for game playing frequency ($p = 0.006 < 0.01$, partial $\eta^2 = 0.47$). Moreover, the ANOVA using ART with the same factors on usability and learnability showed that there was a significant main effect only for game playing frequency for usability ($p = 0.03 < 0.05$, partial $\eta^2 = 0.36$), whereas there was a significant main effect only for visual impairment for learnability ($p = 0.038 < 0.05$, partial $\eta^2 = 0.29$).

Figure 1 shows the effect sizes Cohen's d of the differences of overall SUS scores and subscale usability and learnability scores across the visual conditions. In the graphs, the further to the right the d-values are, the larger the [A-B] group A on the vertical axis, and the darker the color indicates the magnitude of the p-value. The green and red dashed vertical lines indicate the approximate effect size based on Cohen's criterion [8]. The comparison of overall SUS scores showed that low vision participants tended to rate the game higher than blind users with marginal significance ($p = 0.065 < 0.10$, $d = 1.57$ (large)). However, there was no significant difference in usability score between visual impairment statuses ($p > 0.10$, in all conditions), and such a difference was confirmed for learnability score. In particular, low vision participants tended to rate learnability significantly higher than the sighted ($p = 0.034 < 0.05$, $d = 1.51$ (large)) and higher with marginal significance than the blind ($p = 0.093 < 0.10$, $d = 1.45$ (large)).

Regarding UEQ scores, an ANOVA with ART of the pragmatic score for visual impairment status and game playing frequency confirmed main effects for visual impairment status ($p < 0.001$, partial $\eta^2 = 0.55$) and game playing frequency ($p = 0.015 < 0.05$, partial $\eta^2 = 0.42$). The interaction between visual impairment status and game playing frequency was also significant for the pragmatic score ($p = 0.025 < 0.05$, partial $\eta^2 = 0.50$). An ANOVA with the same factors on overall score indicated a significant main effect for visual impairment status ($p = 0.017 < 0.05$, partial $\eta^2 = 0.34$) and a main effect with significant trend for gameplay frequency ($p = 0.07 < 0.10$, partial $\eta^2 = 0.30$). No significant main effects or interactions were found for the usability.

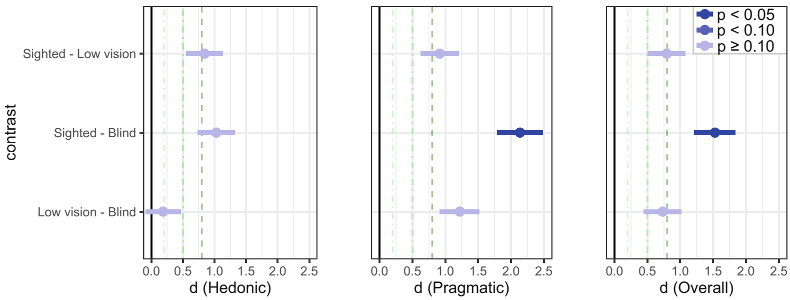


Fig. 2. Effect size of the contrast by vision status on UEQ-S scores: (left) hedonic, (center) pragmatic, and (right) overall scores

Figure 2 illustrates the effect sizes Cohen's d of the differences of UEQ scores across the visual conditions. The comparison of pragmatic and overall scores showed that the sighted tended to rate the game higher than the blind (pragmatic score: $p < 0.001$, $d = 2.14$ (large), overall score: $p = 0.012 < 0.05$, $d = 1.53$ (large)). Though there was no significant difference in hedonic score between visual impairment statuses, effect sizes between the sighted and the visually impaired groups were large generally (sighted − low vision: $p = 0.29 > 0.10$, $d = 0.84$ (large), sighted − blind: $p = 0.105 > 0.10, d = 1.02$ (large)).

These results suggest that, in general, those with low vision rated the leanability of the game significantly higher, while those without visual impairment rated the pragmatic aspects of the game significantly higher. The low vision may have evaluated the game itself as a playable game, in addition to the ability to use both visual and auditory information. The sighted may have valued similar aspects, and they could also felt that inclusiveness could be realized in this way. Meanwhile the blind did not rate either of these aspects highly suggesting that the blind who participated in this study were those who had particular experience in playing games. They made more comments in the free description field than the others. We will discuss this point in the future.

3.2 SUS and UEQ-S Scores: Gameplay Frequency in Sighted Participants

In the sighted group, various differences were observed in the gameplay frequency. Figures 3 and 4 show the overall and partial SUS scores by sighted participants' gameplay frequency. Because SUS scores were non-normally distributed, we plotted the data using box plots. According to the results of ANOVA (analysis of variance) with ART (aligned rank transform) [17], there are significant main effects of gameplay frequency in the overall ($p = 0.02$, partial $\eta^2 = 0.462$) and usability ($p = 0.04$, partial $\eta^2 = 0.390$) SUS scores, but unsignificant main effects in the learnability SUS score ($p = 0.31 > 0.05$, partial $\eta^2 = 0.196$). Among the scores with significant main effects, those who did not play at all significantly or marginally significantly scored higher than those who played every day or 2–3 times a week, according to Figs. 3 and 4.

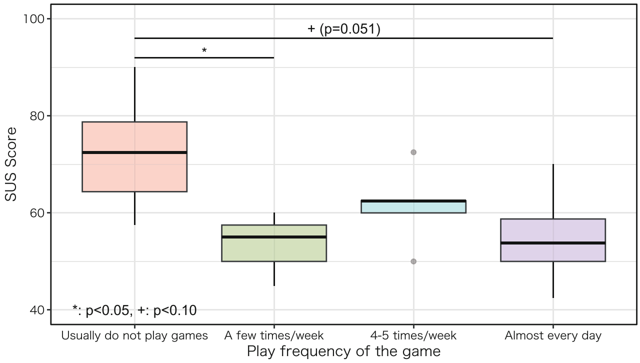


Fig. 3. Overall SUS scores by the gameplay frequency in sighted participants

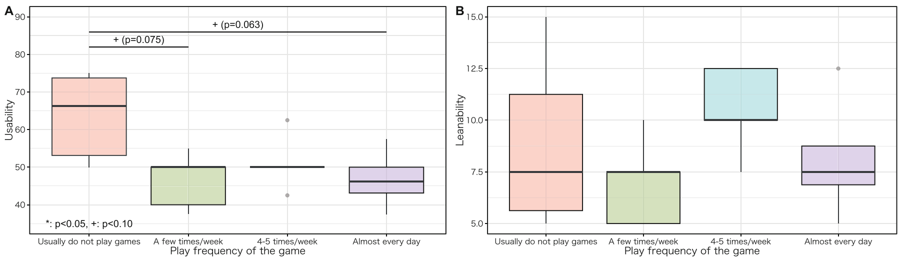


Fig. 4. Usability and learnability scores by the gameplay frequency of the sighted

Figure 5 shows the scores of UEQ-S by the participants' gameplay frequency. According to the results of ANOVA, there are significant main effects of gameplay frequency in the pragmatic scores ($p = 0.049$, partial $\eta^2 = 0.38$). The participants who played games almost every day significantly evaluated lower than those who do not usually play ($p = 0.03$ (Tukey-Kramer test), $d = 1.78$).

These results indicated that the usability of game was highly evaluated by inexperienced players significantly, while the usability of operation learning was independent of playing experience. Moreover, pragmatic scores were generally rated lower by those who played more frequently, while hedonic scores were not significantly different among the player experience. Thus, game play using only the sound accessibility features can be enjoyed to almost the same degree regardless of gameplay frequency (and experience), and that the game experience can potentially provide an inclusive gaming experience regardless of disability or gaming background.

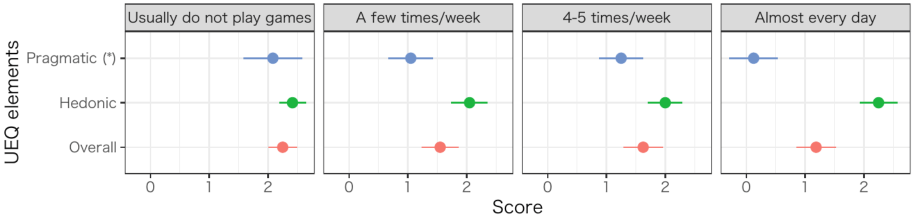


Fig. 5. Scores of UEQ-S by the sighted participants gameplay frequency

3.3 Subjective Evaluations of Sound Accessibility Features

Table 1 shows the subjective impressions of sound accessibility features in visually impaired and sighted participants. These results indicated that the overall evaluation (Q1) was significantly higher than the neutral value (= 3) in both visually impaired and sighted groups.

Table 1. Subjective evaluations of sound accessibility features in visually impaired and sighted participants. The *p* and *d* values were calculated by one-sample *t*-test and Cohen's *d* compared to the neutral value (= 3), respectively.

Question item	Visually impaired (N=11)			Sighted (N=20)		
	Mean (SD)	p	d	Mean (SD)	p	d
Q1. Overall Evaluation of sound accessibility features	4.27 (1.01)	**0.002**	**1.261**	4.57 (0.68)	**<0.001**	**2.324**
Q2. Identification of own and opponent characters	4.09 (1.14)	**0.010**	**0.960**	3.29 (1.31)	0.329	0.218
Q3. Estimating the distance of the opponent	3.73 (1.49)	0.136	0.488	3.57 (1.29)	0.055	0.444
Q4. Estimating where player character (PC) is standing	4.09 (1.38)	**0.025**	0.793	2.95 (1.40)	0.877	0.034
Q5. Estimating the hit distance of PC moves	2.91 (1.22)	0.810	0.074	3.52 (1.29)	0.077	0.406
Q6. Estimating when to hit/guard	3.45 (1.44)	0.320	0.316	3.57 (1.21)	**0.042**	0.473
Q7. Estimating how much drive gauge is left	1.91 (0.94)	**0.003**	**1.156**	2.62 (1.28)	0.189	0.297
Q8. Estimating how much SA gauge is left	2.00 (1.26)	**0.026**	0.791	2.67 (1.39)	0.285	0.240
Q9. Contribution of in-game commentary to situational understanding	2.64 (1.63)	0.476	0.223	3.85 (1.31)	**0.009**	0.649

There were differences between the visually impaired and sighted groups in their perception of the game situation using the sound accessibility features. The visually impaired group could significantly identify their own and opponent characters (Q2) and significantly estimate the positions where their player character was standing (Q4). The sighted could significantly distinguish between hit or guard status (Q6) and understand the situation through the commentary (Q9). They also understand the position of the opponent character (Q3) and the hit distance that their PC moves (Q5) better with marginal significance than the neutral value. Only the visually impaired group mentioned the remaining amount of the drive and SA gauges as significantly difficult to estimate (Q7 and Q8). These points need to be improved in the future.

In the sighted participants, there were no significant differences in the frequency of game playing among the items that showed a significant difference or a marginal significance. Therefore, regardless of the frequency of game playing and experience, even sighted persons can play fighting games using only sound information as cues when the sound accessibility features are sufficiently implemented, and may be able to play together with blind persons in the same situation as blind persons.

Acknowledgments. This work was partially supported by JSPS KAKENHI Grant Numbers JP21K18485 and JP23H03903.

References

1. Audiogame.net. http://www.audiogames.net/. Accessed 12 Apr 2024
2. Newzoo global games market report 2020. https://newzoo.com/insights/trend-reports/newzoo-global-games-market-report-2020-light-version/. Accessed 12 Apr 2024
3. Street fighter 6 with improved sound accessibility designed with ePARA inc. to launch June 2. https://epara.jp/ennews-streetfighter6-230501/. Accessed 12 Apr 2024
4. Summary wiki of accessible games for visually impaired gamers. http://mm-galabo.com/AcGameWiki/. Accessed 12 Apr 2024
5. de Borba Campos, M., Damasio Oliveira, J.: Usability, accessibility and gameplay heuristics to evaluate audiogames for users who are blind. In: Antona, M., Stephanidis, C. (eds.) UAHCI 2016. LNCS, vol. 9737, pp. 38–48. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-40250-5_4
6. Borsci, S., Federici, S., Lauriola, M.: On the dimensionality of the system usability scale: a test of alternative measurement models. Cogn. Process. **10**, 193–197 (2009)
7. Brooke, J.: SUS-A quick and dirty usability scale. Usability Eval. Ind. **189**(194), 4–7 (1996)
8. Cohen, J.: A power primer. Psychol. Bull. **112**(1), 155 (1992)
9. Matsuo, M., Miura, T., Sakajiri, M., Onishi, J., Ono, T.: Audible mapper & ShadowRine: development of map editor using only sound in accessible game for blind users, and accessible action RPG for visually impaired gamers. In: Miesenberger, K., Bühler, C., Penaz, P. (eds.) ICCHP 2016. LNCS, vol. 9758, pp. 537–544. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-41264-1_73
10. Matsuo, M., et al.: Inclusive action game presenting real-time multimodal presentations for sighted and blind persons. In: Proceedings of the 2021 International Conference on Multimodal Interaction, pp. 62–70 (2021)
11. Miesenberger, K., Ossmann, R., Archambault, D., Searle, G., Holzinger, A.: More than just a game: accessibility in computer games. In: Holzinger, A. (ed.) USAB 2008. LNCS, vol. 5298, pp. 247–260. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-89350-9_18
12. Miura, T., Yabu, K.: Narrative review of assistive technologies and sensory substitution in people with visual and hearing impairment. Psychologia **65**(1), 70–99 (2023)
13. Porter, J.R., Kientz, J.A.: An empirical study of issues and barriers to mainstream video game accessibility. In: Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility, p. 3. ACM (2013)
14. Schrepp, M., Hinderks, A., Thomaschewski, J.: Design and evaluation of a short version of the user experience questionnaire (UEQ-s). Int. J. Interact. Multimedia Artif. Intell. **4**(6), 103–108 (2017)

15. Tull, M.T., Edmonds, K.A., Scamaldo, K.M., Richmond, J.R., Rose, J.P., Gratz, K.L.: Psychological outcomes associated with stay-at-home orders and the perceived impact of COVID-19 on daily life. Psychiatry Res. **289**, 113098 (2020)
16. Westin, T., Ku, J.E.J., Dupire, J., Hamilton, I.: Game accessibility guidelines and WCAG 2.0 – a gap analysis. In: Miesenberger, K., Kouroupetroglou, G. (eds.) ICCHP 2018. LNCS, vol. 10896, pp. 270–279. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-94277-3_43
17. Wobbrock, J.O., Findlater, L., Gergle, D., Higgins, J.J.: The aligned rank transform for nonparametric factorial analyses using only anova procedures. In: Proceedings of CHI '11, pp. 143–146 (2011)
18. Yuan, B., Folmer, E., Harris, F.C.: Game accessibility: a survey. Univ. Access Inf. Soc. **10**(1), 81–100 (2011)

Towards Improving the Correct Lyric Detection by Deaf and Hard of Hearing People

Hayato Yamamoto[1](✉), Deborah Fels[2], Keiichi Yasu[1], and Rumi Hiraga[1]

[1] Tsukuba University of Technology, Ibaraki, Japan
{a243103,k-yasu,rhiraga}@a.tsukuba-tech.ac.jp
[2] Toronto Metropolitan University, Toronto, ON, Canada
dfels@torontomu.ca

Abstract. Access to music for people who are Deaf/Hard of Hearing (D/HoH) includes the instrumental portion and the lyrics. While closed captioning can provide the lyrics in text format, it does not necessarily provide accurate timing of the lyrics with the instrumental portion. This study aims to clarify how vibrotactile stimuli affect understanding the onset timing of song lyrics for D/HoH people. To achieve this goal, we developed a system called VIBES: VIBrotactile Engagement for Songs that simultaneously provides music and vibration playback as an iPhone app. Unlike other vibrotactile systems for music, which focus primarily on the percussion/beat or frequencies of the instrumental portions, VIBES presents vibrations for the timing of vocal utterances, syllable by syllable. We conducted a study with 10 participants to determine the system's effectiveness by comparing the understanding of lyric timing with vibration and sound only. Statistically, one of the four songs in the experiment showed significant differences, where understanding lyric timing with vibration is better than in sound-only conditions. Two songs revealed the opposite results, and the other showed no differences between the two conditions. Although these results were unexpected, we obtained several findings to make the next version of VIBES, such as the frequency of vibrations.

Keywords: Deaf and Hard of Hearing · Music · Vibration · VIBES

1 Introduction

Closed captioning is now being added to most video content, including songs and music videos, as soon as they are released on YouTube and other online video services. However, the timing of the captions is often delayed [1], making it difficult to know whether a song's instrumental and lyric portions are synchronized. While this may not be a concern if just "listening" to music, for singing along, such as for karaoke, it causes issues with keeping up with the song.

It is a personal experience of one of the authors as a profoundly deaf person that he is not able to grasp the exact timing of the lyrics with the music, so he is often out of rhythm when singing because of the delays between the onset of the closed captions and music, which makes him enjoy less and prevent achieving high scores in karaoke.

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 146–154, 2024.
https://doi.org/10.1007/978-3-031-62846-7_18

Difficulties grasping rhythm during practice and singing with off-vocal sound sources are the main reasons for feeling uncomfortable at karaoke. Usually, he listens to and practices singing from sound sources, including vocal sounds. However, the background music and singing technique make it difficult to hear the vocal sound, and he often needs to get the singing timing correctly. In addition, since he wants the timing based on the vocal sound, the singing timing must be grasped with something other than an off-vocal sound source. These two types of difficulties (background music interferes with the vocal sounds and off-vocal music gives no lyric timings) need help grasping the correct rhythm and continuing to sing in the proper rhythm.

Although several vibration systems have been developed for D/HoH's music acquisition as Paisa [2] reviewed, these systems don't solve his problems. SoundHug [3] lets D/HoH feel the music as a whole but does not let them understand any specific sound. Ontenna [4] is currently not usable for understanding the lyric timings. So, we developed a system that vibrates the lyrics of a song during playback to eliminate the difficulty of listening to a sound source containing vocal sounds.

This paper describes the system and the results of a user study with ten Deaf/Hard of Hearing (D/HoH) participants. The study compared participants' ability to detect the syllables in lyrics of short-duration segments of four songs with and without the system.

2 Vibes: Vibrotactile Engagement for Songs

VIBES is a singing support tool that provides vibration to indicate syllables to follow the lyric synchronization for D/HoH singers. It can play back an existing song music prepared in advance, either with sound alone (SO) or with a vibration (V) that indicates individual syllables in a song sound and as vibration.

This system was developed as an iOS Swift application using Xcode with two libraries of AVFoundation and CoreHaptics and installed on the iPhone 13. VIBES presents vibrotactile stimuli of syllables for a song that matches the vocal timing of the music. The system also displays lyrics to provide a visual confirmation of the lyric/music timing so that the correct vocal timing can be learned with this system. Figure 1 shows the two screens. Left side of Fig. 1 shows the initial screen display with the titles of the four songs, the singer's name, and the lyrics of segments used in the experiment. By selecting a piece of music, a second screen (see right side of Fig. 1) allows the user to start a song. Currently, VIBES is available on TestFlight as an internal test version. We used the frequency and amplitude of the vibration by iPhone as the default value. The frequency is between 100 to 130 Hz, while the amplitude is unknown.

3 Methodology

3.1 Participants

Ten D/HoH undergraduate students (four female, six male), aged 20–24, interested in music and singing, were invited to participate in a user study. The communication methods used by participants were Japanese sign language and/or speech. Hearing levels ranged from 51 dB Hearing Loss (HL) to 105 dBHL. Eight participants used hearing aids in both ears, a participant used a cochlear implant in one year, and the other was a bimodal user (who used a cochlear implant and a hearing aid).

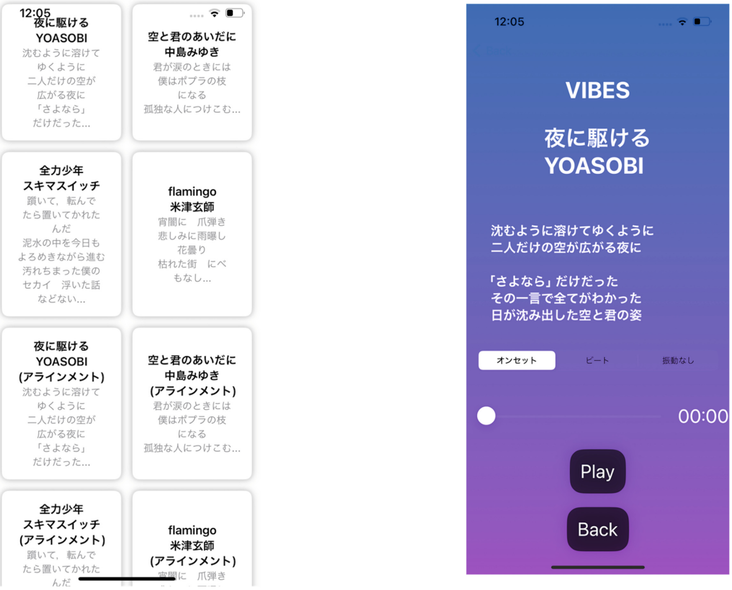


Fig. 1. The initial screen and screen after song selection of VIBES.

Table 1. Songs used in VIBES.

	Sora to Kimi no Aida ni	Zennryoku Shonen	Yoru ni Kakeru	Flamingo
Voice type	Female	Male	Female	Male
Instrument	Piano, drums, guitar, bass guitar (vocals in foreground)	Piano, drums, guitar, buss guitar (vocals in foreground)	Piano, drums, guitar, bass guitar (vocals in foreground)	Piano, drums, guitar, bass guitar (vocals in background)
Tempo	107	132	130	114
Playback time[s]	128.0	102.0	103.0	66.0

Fig. 2. The flow of the study.

3.2 Stimuli

The four Japanese songs were selected from the COSIAN collection [5] (Sora to Kimi no Aida ni, Zennryoku Shonen, Yoru ni Kakeru, and Flamingo). These selections were based on two male and two female singers, one after 2015 and one before for each gender. In all cases, the lyrics are in Japanese only. Each song was downloaded from YouTube (WAV format, sampling frequency of 44,100 Hz). Table 1 shows the length of each segment, the instrument and voice types, and the tempo.

4 Procedure

Participants were asked to complete a pre-study question to gather demographic information such as age and gender, hearing loss, and experience with singing. Following university ethics approval, participants were invited to tap when they noticed the syllable in the song segments. The tapping detection program, Presentation™ by Neurobehavioral Systems, recorded when participants tapped their finger on Surface Pro 7 in milliseconds. It has a resolution of 0.1 ms. This was a within-subjects study, so each participant tapped syllables for all song segments.

After consent, participants were placed in an anechoic room and introduced to VIBES. They were then presented with the sound-only (SO) segment, tapped with sound, were presented with sound and vibration (V), and then tapped with sound (Fig. 2). This procedure was repeated for all four songs. The order of presenting music was randomized.

After completing both conditions for each song, participants were asked to respond to three 5-point Likert-style questions about whether it was easy to hear, could grasp the rhythm, and could determine different parts of the song, such as the introduction and the chorus as well as provide their opinion of the VIBES system. The 5-point scale ranged from strongly disagree (1) to agree strongly (5). We also asked whether they were familiar with the song or had sung it. They were also encouraged to comment on their experience, which was recorded and transcribed. The study took approximately 90 min to complete. Participants received an honorarium of approximately 10 USD.

For each piece of music, the original sound-based lyric timing data for each song was compared to the tapping timing, which included response time. Statistical analyses using SPSS version 29 were based on the time difference between when the participant tapped on the tablet and the actual syllable onset time. The p-value was set to 0.05 for all tests. If a participant tapped 100 ms before the actual syllable onset (undershoot error) or 100 ms after syllable onset (overshoot error), the tap was considered missed, which was not included in the data. However, these missed taps could be regarded as errors for later analysis.

5 Results

5.1 Differences in Lyric Detection Accuracy Between with and Without Vibration

A Shapiro-Wilk test was carried out for the different data for each song and found to be non-normally distributed ($p < 0.05$). Henceforth, non-parametric analyses were used. A Wilcoxon signed rank test showed that there was a significant difference between SO and V conditions for three songs, Sora to Kimi no Aida ni ($W = 24422.50$, $z = -2.90$, $p < 0.05$), Zenryoku Shonen ($W = 14702.50$, $z = -2.30$, $p < 0.05$), and Yoru ni Kakeru ($W = 51281.50$, $z = -3.52$, $p < 0.05$). There was no significant difference between the two conditions for Flamingo. Table 2 and Fig. 3 show each condition and song's means and standard deviations.

Table 2. Means and standard deviation for each song in each study condition, sound only (SO) and vibration (V).

Song title	Mean (SD) – SO (msec)	Mean (SD) – V (msec)
Sora to Kimi no Aida ni	845 (1918)	1491 (3652)
Zenryoku Shonen	376 (220)	342 (290)
Yoru ni Kakeru	622 (1341)	703 (1331)
Flamingo	425 (389)	578 (885)

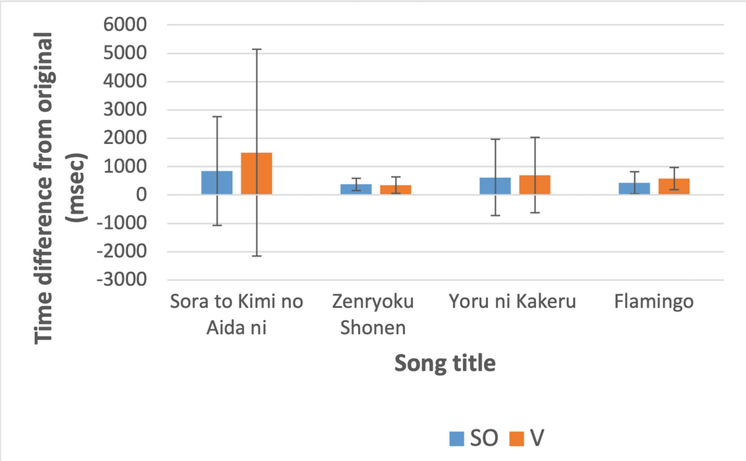


Fig. 3. Mean and standard deviation for each song title and condition (sound only (SO) and sound and vibration (V)).

Friedman's repeated measures test was used to determine whether there were differences in performance between the songs. There was no significant difference between songs for the sound only condition and songs for the vibration condition. Table 2 shows each song's means and standard deviations in the SO condition.

5.2 Subjective Evaluation of the Stimuli

Participants' subjective evaluation in Likert style after the tapping procedures is shown in Table 3 and Fig. 4. Table 3 shows the median and interquartile values of the answers to the questions, and Fig. 4 shows the box plot. More than two participants mentioned that VIBES can tell double consonant and prolonged sounds. These are included in the next version of VIBES.

Table 3. Median and interquartile of the answers for each song.

Song title	Question	Median	IQR
Sora to Kimi no Aida ni	Listening was easy	3.00	2.75
	Understood the rhythm at the introduction	2.00	2.00
	Understood the rhythm	4.00	2.00
Zenryoku Shonen	Listening was easy	4.00	1.00
	Understood the rhythm at the introduction	4.00	2.00
	Understood the rhythm	4.00	3.00
Yoru ni Kakeru	Listening was easy	5.00	0.00
	Understood the rhythm at the introduction	5.00	0.75
	Understood the rhythm	5.00	1.00
Flamingo	Listening was easy	5.00	0.75
	Understood the rhythm at the introduction	4.00	1.00
	Understood the rhythm	4.50	1.00

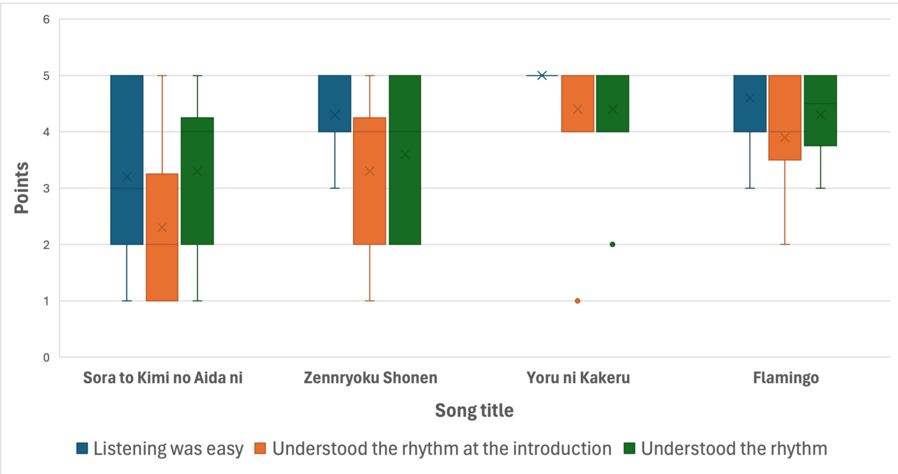


Fig. 4. Distribution of the answers to the three questions

6 Discussion

The mean and standard deviation between SO and V conditions for two songs with significant differences, Sora to Kimi no Aida ni and Yoru ni Kakeru, were unexpected, where the mean was lower for the SO condition, meaning that the tapping time was closer to that of the original timing. Participants were faster at recognizing the onset of the lyric syllables by just hearing it than when the vibrotactile stimuli and sound were present (V) condition.

In addition, for one song, Sora to Kimi no Aida ni, the standard deviation for the SO condition was also considerably smaller than the V condition. This could indicate that vibrotactile and audio/visual stimuli for lyrics delayed the participant's recognition of the lyric onset. Eight of the ten subjects had never heard or sung this song. Even though the vocal was in the foreground of this song, many subjects missed the syllables when moving from the introduction to the melody. For first-time listeners, the SO condition may have led to tapping while searching for the rhythm. Therefore, even if they had missed a syllable, they could respond flexibly in the middle of the tapping. However, in the V condition, after the subjects had grasped the correct rhythm without vibration, if they could not perceive the initial timing with vibration because they had grasped the rhythm, they may have lost it for a specific section of time afterward.

Yoru ni Kakeru is widely known among young people. Nine participants reported that they had heard it before the experiment. Because of the song's familiarity, participants felt the song was easy to listen to and grasp the timings (Fig. 4). Conversely, the V condition generated worse results than the SO condition in tapping. This could be explained by the timing they had previously acquired taking precedence, causing them to become confused and to perform tapping with the vibration.

For the Zennryoku Shonen, the tapping after the V condition was significantly smaller than that for the SO. Five out of ten participants found this piece challenging to listen to. A characteristic of this song is some prolonged sounds where the vocal continues long between a syllable and the next syllable, making it difficult for participants to detect the next syllable in the SO condition. Even if the participants could understand the rhythm so far, once they lost the rhythm, the rhythm in the consecutive part was lost. On the other hand, the subjects could detect the syllable after the long one in the V condition. This made participants perform tapping in the V condition more accurately than in the SO condition.

Flamingo has more distinct instrumental sounds that present beats, such as drums and bass. Therefore, the simultaneous presentation of vibrations by VIBES and natural vibrations by the sound source playback may have affected the perception of accurate timing.

We also care about the frequency range of music and human perception in detecting signals by vibrotactile, visual, and audio stimuli. These differences may have resulted in delays in processing the stimuli and the resultant tap in response, assuming that motor reaction time to produce the tap physically is similar for each participant in both conditions. Some music has a more robust low frequency from the vibrotactile stimuli, and the frequency range of most music is in the middle octave (C4 to C5). Therefore, the frequency value of VIBES is essential for participants to be clear of VIBES stimuli and vibration from low-frequency instrument sound.

When we observed participants during the experiment, they often needed to catch up on the timing of the first word of the lyrics in the SO. When the timings of the actual syllables and the beat onset were minor, the participant possibly tapped along the beat, ignoring the correct timing of the syllables.

In the tapping after the V condition, there was no significant difference in performance between songs. Interviews with the participants revealed that they responded to the difficulties in hearing by listening to the music repeatedly or by increasing the volume. Therefore, tapping after using VIBES in the V did not take advantage of the precise timing presented.

6.1 Limitation

There were several limitations in this early research, including a need for counter-balanced condition presentations and a small number of participants, resulting in a within-subjects design. In addition, participants only had limited exposure to VIBES in that there were only short training sessions followed by short durations of song segments. Additional practice time may assist participants in detecting syllables over longer duration song segments.

6.2 Future Works

The eventual goal of the VIBES system is to assist people in singing songs, such as in Karaoke settings. The current experiment asked participants to tap word onset and not sing. Once the VIBES system has been re-developed to account for the comments and issues, the next step would involve many more songs in VIBES by automatically obtaining the correct onset timing from those songs.

References

1. Butler, J.: Perspectives of deaf and hard of hearing viewers of captions. Am. Ann. Deaf **163**(5), 534–553 (2019). https://www.jstor.org/stable/26663593
2. Paisa, R., Nilsson, N.C., Serafin, S.: Tactile displays for auditory augmentation–A scoping review and reflections on music applications for deaf users. Front. Comput. Sci. **5** (2023). https://www.frontiersin.org/articles/10.3389/fcomp.2023.1085539
3. SundHug. https://pixiedusttech.com/technologies/sound-hug/
4. Honda, T., Baba, T., Okamoto, M.: Ontenna: Design and Social Implementation of Auditory Information Transmission Devices Using Tactile and Visual Senses. In: Miesenberger, K., Kouroupetroglou, G., Mavrou, K., Manduchi, R., Covarrubias Rodriguez, M., Penáz, P. (eds.) Computers Helping People with Special Needs, LNCS, vol. 13342. Springer, Cham. (2022). https://doi.org/10.1007/978-3-031-08645-8_16
5. Yamamoto, Y., Juhan, N., Terasawa, H.: Analysis and detection of singing techniques in repertoires of J-POP solo singers. In: Proceedings of the 23rd International Society for Music Information Retrieval Conference (2022)

Enhancing Accessibility in Sports and Cultural Live Events: A Web Application for Deaf, Hard of Hearing, Blind, Low Vision, and All Individuals

Yuhki Shiraishi[1(✉)], Rumi Hiraga[1], Daisuke Wakatsuki[1], Makoto Kobayashi[2], Takuya Suzuki[1], Ying Zhong[1], and Takeaki Shionome[3]

[1] Faculty of Industrial Technology, Tsukuba University of Technology, 4-3-15 Amakubo, Tsukuba, Ibaraki 3058520, Japan
{yuhkis,rhiraga,waka,suzukit,ying}@a.tsukuba-tech.ac.jp

[2] Faculty of Health Sciences, Tsukuba University of Technology, 4-12-7 Kasuga, Tsukuba, Ibaraki 3058521, Japan
koba@cs.k.tsukuba-tech.ac.jp

[3] Faculty of Science and Engineering, Teikyo University, 1-1 Toyosatodai, Utsunomiya, Tochigi 3208551, Japan
shionome@ics.teikyo-u.ac.jp
https://www.iseee.info

Abstract. This paper addresses the issue in live events, especially in sports viewing, where individuals who are deaf, hard of hearing (DHH), blind, or have low vision (BLV) struggle to access sufficient information. Traditionally, information accessibility has relied on specific professionals or volunteers, which is often inadequate. To tackle this challenge, we propose a mechanism that facilitates information sharing for all individuals, regardless of their abilities. We have developed an inclusive web application tailored to address these needs. This application is beneficial not only for DHH and BLV individuals but also for all audiences. Additionally, the system's applicability extends to other domains, such as museum visits. This paper details the newly developed web application and presents the outcomes of pilot studies conducted in sports viewing and museum settings with DHH and BLV participants. The results of these experiments are analyzed to assess the system's effectiveness and to identify future improvement areas.

Keywords: Information accessibility · Information support · Web application · Deaf and hard of hearing · Blind and low vision

1 Introduction

In live events, particularly during sports viewings, essential information such as game status and in-venue guidance is primarily conveyed through audio. Surveys focusing on support for individuals who are deaf or hard of hearing (DHH) [5]

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 155–163, 2024.
https://doi.org/10.1007/978-3-031-62846-7_19

have shown that the inability to hear announcements and understand detailed game situations significantly detracts from their experience, leading to considerable inconvenience and dissatisfaction. Similarly, for individuals who are blind or have low vision (BLV), the lack of adequate visual information or suitable alternatives can significantly diminish the enjoyment of watching sports, presenting a considerable challenge. Thus, providing accessible information by professionals or volunteers is crucial for DHH and BLV individuals to fully appreciate and enjoy sports events, ensuring they can experience the complete joy of the spectacle.

Due to cost and human resource limitations, providing specialized staff for information support at small-scale sports events in Japan is challenging. Converting Japanese speech to text is more complex than English, as it includes kanji, hiragana, and katakana and has many homophones, which leads to lower accuracy in Japanese speech recognition and a reliance on manual typing, which is slower due to the need for kanji and kana conversion. Collaborative input [10], where multiple staff share typing duties, is used but incurs labor costs. Correcting speech recognition errors also requires multiple staff. While specialized radio terminals with detailed live commentary are available for BLV spectators, such services are limited to specific matches.

Therefore, we propose a mechanism where individuals disseminate and share information regardless of their abilities, providing mutual information support. Specifically, we have developed a real-time timeline posting and viewing system as an accessible web application that allows spectators to share data composed of photos, texts, audio, and videos during sports events. By utilizing a sports viewing timeline aggregated from the data posted by spectators, including the situation of the match and other information, we address the challenges of providing information support to DHH and BLV individuals. Furthermore, the applicability of this system extends to other areas, such as museum visits. This paper details the newly developed web application and presents the results of pilot studies conducted with DHH and BLV participants during sports viewings and museum visits. We analyze the results of these experiments to evaluate the system's effectiveness and identify areas for future improvement.

2 State of the Art

The web application we developed introduces a system that enhances information sharing for everyone, including DHH and BLV individuals. Specifically, it enables various individuals to contribute information leveraging their strengths without requiring specialized skills, thereby facilitating a synergistic exchange of information. We describe this inclusive and collaborative approach to information sharing, adaptable to all abilities, as 'Information Accessibility 2.0'. Essentially, this application is not only advantageous for DHH or BLV individuals but is universally beneficial.

Previous studies have proposed non-expert speech-to-text interpretation methods for DHH, such as automatic speech recognition [2], crowdsourcing for speech-to-text/sign-language-to-text interpretation [3,6], and sign language video-to-text conversion using crowdsourcing [7]. Information support for DHH using crowdsourcing at sports games has also been proposed; prototypes of sports spectator-specific systems for DHH have been evaluated [9]. However, features and accessibility considerations that facilitate ease of use for everyone are limited, and BLV individuals have not been thoroughly evaluated. Despite some proposals for information support for DHH at sports games, such as display systems using Japanese sign language animation synthesis [8], the practices and proposals in this area still need improvement. The importance of captions in sports broadcasting was discussed in [1]. However, issues such as delays in captions and instances where captions obscured essential information were also raised.

Therefore, we aim to develop an inclusive web application to meet these requirements. Our developed application partially conforms to the Web Content Accessibility Guidelines (WCAG) 2.1, targeting at least Level AA [11]. This partial compliance is because we cannot control the content users post. However, to enhance the accessibility of these posts, we have introduced methods such as the automatic generation of alternative content using artificial intelligence (AI), i.e., speech recognition or object recognition, and allowing users other than the poster to input additional information as alternative content.

3 System Overview

Our application provides functionalities such as 'tags' and 'filters', enabling users to access only the information they need selectively and in real-time. This feature dramatically simplifies the process for users with BLV to access vital information, such as scores or player substitutions. Additionally, when used in museums, setting tags based on location or theme allows for easily sharing desired information. The system is designed with flexibility and user-friendliness; full functionality is available for users who create an account, while those without an account can still post limited content through a browser. This approach accommodates a wide range of information literacy levels.

To use the developed web application, users access the server through a web browser on their smartphone or PC and send/receive timeline data. Communication is facilitated using HTTP and WebSocket protocols. Posts are push-delivered to all users in the same room, and the web browser instantly displays the received posts in a timeline format. Importantly, unlike platforms such as "X", reloading is unnecessary to see the latest updates. A snapshot of the screen displayed on the user's web browser is shown in Fig. 1.

Figure 1a depicts the main window. The timeline panel on the left side of the screen displays all posts. Although the original posts are in Japanese, they are automatically translated into English for users who are English speakers. In addition to regular posts with usernames, there are 'flowing' posts (in this example, corresponding to "It's very cold in the venue"). These are messages whose sender cannot be identified, promoting casual posts equivalent to cheers in the venue. On the right, the 'Commentary' panel uses a filter function to display only posts tagged with ‘commentary’. It is also possible to display multiple filter panels simultaneously. Additionally, features such as notifications through sound for posts of pre-specified types, voice reading capabilities, and the ability to share posts on “X” are also available. Furthermore, various shortcuts are provided, enabling operation solely through keyboard commands. Authorized users can remove inappropriate posts or ban users who breach posting guidelines. Defining specific words that prevent posting is also possible, further bolstering the platform’s moderation tools.

Figure 1b displays the posting panel, where users can select any tag when making a post. Furthermore, users can add or remove tags after a post has been made. The panel also supports the posting of images, videos, and audio. If specific tags—‘image analysis’ and ‘audio analysis’—are applied to these media posts, an AI automatically analyzes the content and appends additional information in text form. Figure 1c shows the floor list. As part of the system configuration, multiple ‘rooms’, as shown in Fig. 1d, exist within a ‘floor’. Floors are categorized into three types: those that are visible to all, those that are visible only to those who know the URL of the floor (hidden floors), and those that are accessible only to invited members (members only). The same applies to rooms. In addition, floor members can freely create rooms and invite members to members-only rooms. Opening a room displays the main window, as in Fig. 1a. It is anticipated that a dedicated floor will be created for each event, and a dedicated room will be created for each match.

4 Experiment Design

To evaluate the fundamental effectiveness of the developed system, we conducted demonstration experiments during sports viewing and museum visits. These experiments were carried out with the approval of the research ethics review at the authors’ affiliated institution.

4.1 Sports Viewing Experiment

DHH participants watched handball games while engaging with others by answering questions. On January 14, 2024, an initial analysis was conducted with three DHH university students aged 19–20 using the System Usability Scale (SUS) after two hours of system use. The DHH participants, who had limited experience with sports viewing—two of whom rarely attend sports events and one who attends about 1–2 times a year—spent two hours together. They have an average hearing loss between 95–110 dB in both ears.

For the foundational experiment with BLV individuals, an online environment simulating sports viewing was constructed using Zoom and YouTube live streaming on October 27, 2021. In this experiment, four pairs of BLV participants and one pair of DHH participants watched live-streamed videos of visually impaired bowling from the 17th All Japan Blind Bowling Championship. Each pair was assigned to a breakout room and participated in two sessions, one

(a) Main Window

(b) Posting Panel

(c) Floor List

(d) Room List

Fig. 1. Snapshot of the Developed Web Application

with and one without the developed web application, to share information while watching the videos.

4.2 Museum Visit Experiment

Six DHH university students aged 19–22 participated alongside five actively engaged hearing experiment collaborators on November 29, 2023. The DHH participants spent two hours together with an average hearing loss between 80–105 dB in both ears. Five visited the museum for the first time, while one had visited 3–4 times.

5 Results and Discussion

The sports viewing and museum visit experiments provide valuable insights into the effectiveness of the developed web application in enhancing information accessibility for DHH and BLV individuals.

5.1 Sports Viewing Experiment

The DHH participants achieved an average SUS score of 73.3 (standard deviation of 7.2), rated as 'Good' [4]. The aspects that negatively impacted the score were the 'need for technical support' and perceived 'system inconsistencies', averaging 5.8 out of 10. Regarding the overall sports viewing visit experience facilitated by the system, two participants awarded it the maximum score of 5, and one gave a 4 on a 5-point Likert scale. When asked about their agreement with the research project's efforts, responses were favorable, with one participant giving a 5 and two giving 4 s. While the SUS score indicates the system's potential, further refinements are necessary to improve usability and user experience.

The foundational experiment for BLV individuals demonstrated the web application's potential to support information accessibility in sports viewing scenarios, with more participants reporting obtaining useful information when using the developed system. However, the negative results regarding participants' ability to provide information themselves suggest the need for improvements and user training to encourage active participation. Future research should explore ways to facilitate and encourage active participation from BLV individuals, such as providing accessible input methods and offering user training sessions.

5.2 Museum Visit Experiment

The DHH participants evaluated the system based on SUS, resulting in an average score of 55.0 (standard deviation of 14.1), regarded as 'Poor' [4]. The main factors for the low evaluation were "easy to use" and "most people would learn to use this system very quickly" (both average scores of 3.8 out of 10). When asked if the overall museum visit experience using the developed system was positive, five out of six participants rated it the highest score of 5 on a 5-point

Likert scale, with one participant giving a 4, indicating overall satisfaction with the system. All six participants strongly agreed with the efforts of this research project, giving the highest score of 5. The lower SUS score suggests that initially designed for sports viewing, the system requires optimization for museum settings. The manual switching of tags based on viewing location contributed to the lower usability scores, indicating that exploring automatic tagging methods could enhance the system's usability in museum contexts.

These results suggest that the developed web application is a promising tool for enhancing information accessibility for DHH and BLV individuals in both sports viewing and museum visit contexts. However, the system requires further improvements and context-specific optimizations to fully realize its potential. Future research should focus on addressing the identified usability issues, exploring automatic tagging methods, and conducting larger-scale experiments to validate the system's effectiveness in various settings.

6 Conclusion and Future Works

This paper presents a novel web application to enhance information accessibility for DHH and BLV individuals in live events such as sports viewings and museum visits. The results of the sports viewing and museum visit experiments demonstrate the potential of the developed system in enhancing information accessibility for DHH and BLV individuals. However, the experiments also identified areas for improvement, such as the need for technical support, system inconsistencies, and usability issues in the museum setting.

Future work will prioritize incorporating feedback from DHH and BLV individuals to drive the system's improvement. We will conduct user studies and workshops with DHH and BLV participants to gather insights into their specific needs, preferences, and challenges. This valuable feedback will be used to refine the system's design, features, and user interface, ensuring that it effectively meets the needs of its target users.

In addition to user feedback, we will explore integrating automatic tagging methods, such as indoor positioning technologies and augmented reality, to enhance the system's usability and adaptability to different contexts. We will also conduct larger-scale experiments with diverse participant groups and settings to validate the system's effectiveness and generalizability.

By continuously improving and adapting the web application based on feedback from DHH and BLV individuals, we aim to contribute to developing inclusive technologies that empower individuals with diverse abilities to participate fully in live events and enhance their overall experience.

Acknowledgments. This work was supported by Expense for Strengthening Functions in NTUT Budgetary Request.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Butler, J.: Perspectives of deaf and hard of hearing viewers of captions. Am. Ann. Deaf **163**(5), 534–553 (2019). https://www.jstor.org/stable/26663593
2. Elliot, L., Stinson, M., Easton, D., Bourgeois, J.: College students' learning with c-print's educational software and automatic speech recognition. In: Proceedings of the American Educational Research Association Annual Meeting, pp. 1–22. American Educational Research Association, New York (2008), Paper Presented on March 25, 2008. Funded by the U.S. Department of Education, Office of Special Education and Rehabilitative Services, Model Demonstration Projects for Children with Disabilities, Grant #H324M030070
3. Lasecki, W.S., Miller, C.D., Bigham, J.P.: Warping time for more effective real-time crowdsourcing. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2013, pp. 2033–2036. Association for Computing Machinery, New York (2013). https://doi.org/10.1145/2470654.2466269
4. Lewis, J.R., Sauro, J.: The factor structure of the system usability scale. In: Kurosu, M. (ed.) HCD 2009. LNCS, vol. 5619, pp. 94–103. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-02806-9_12
5. Mizuno, E.: How to provide information for everyone to enjoy sports watching: from the viewpoint of the hearing impaired who are difficult to obtain information at the time of watching [translated from Japanese]. Life Des. Rep. Spring **2017**(4), 29–42 (2017)
6. Shiraishi, Y., Zhang, J., Wakatsuki, D., Kumai, K., Morishima, A.: Crowdsourced real-time captioning of sign language by deaf and hard-of-hearing people. Int. J. Pervasive Comput. Commun. **13**(1), 2–25 (2017). https://doi.org/10.1108/IJPCC-02-2017-0014
7. Tanaka, K., Wakatsuki, D., Minagawa, H.: A study examining a real-time sign language-to-text interpretation system using crowdsourcing. In: Miesenberger, K., Manduchi, R., Covarrubias Rodriguez, M., Peňáz, P. (eds.) ICCHP 2020. LNCS, vol. 12377, pp. 186–194. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-58805-2_22
8. Uchida, T., et al.: Systems for supporting deaf people in viewing sports programs by using sign language animation synthesis. ITE Trans. Media Technol. Appl. **7**(3), 126–133 (2019). https://doi.org/10.3169/mta.7.126
9. Wakatsuki, D., et al.: A study on information support for deaf and hard-of-hearing people using sports game timeline. In: 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 2248–2253 (2018). https://doi.org/10.1109/SMC.2018.00386
10. Wakatsuki, D., Kato, N., Shionome, T., Kawano, S., Nishioka, T., Naito, I.: Development of web-based remote speech-to-text interpretation system captionline. J. Adv. Comput. Intell. Intell. Inform. **21**, 310–320 (2017). https://doi.org/10.20965/jaciii.2017.p0310
11. World Wide Web Consortium: Web content accessibility guidelines (WCAG) 2.1 (2018). https://www.w3.org/TR/WCAG21/. Accessed 02 Feb 2024

Art Karshmer Lectures in Access to Mathematics, Science and Engineering

ChattyBox: Online Accessibility Service Using Fixed-Layout DAISY

Katsuhito Yamaguchi[1,2](✉), Toshihiko Komada[3,5], and Masakazu Suzuki[1,4]

[1] Non-Profit Organization: Science Accessibility Net, Fukuoka, Japan
{eugene,msuzuki}@sciaccess.net
[2] College of Science and Technology, Nihon University, Tokyo, Japan
[3] Junior College Funabashi Campus, Nihon University, Chiba, Japan
komada.toshihiko@nihon-u.ac.jp
[4] Faculty of Mathematics, Kyushu University, Fukuoka, Japan
[5] Non-Profit Organization: Science Accessibility Net, Chiba, Japan
http://www.sciaccess.net/en/

Abstract. A new type of daisy books, "Fixed-layout DAISY" is proposed, in which the whole page is treated as a multi-layer picture, the second layer of which has the same form as the original print document. A DAISY (EPUB3) player can read out any texts on the transparent front layer together with highlighting them. By making use of Fixed-layout DAISY, "ChattyBox" is developed, which provides dyslexic people with various innovative online accessibility services such as automatically converting PDF into Fixed-layout DAISY, playing DAISY books back with a popular browser, etc.

Keywords: Online service · Print disability · Fixed-layout DAISY

1 Introduction

As is well known, DAISY (Digital Accessible Information System) is international standards of accessible e-books [1]. Roughly speaking, there are three different types in DAISY: audio-based, text-based and multimedia types. In Multimedia DAISY, not only text information but also audio files of aloud reading are embedded as media overlay. A DAISY player can play back the audio files with highlighting phrases. In Japan, dyslexic people especially have difficulties in reading Kanji (Chinese characters; how to pronounce them depends on context). In Multimedia DAISY, you can assign correct pronunciations in advance, and it is recognized as the most appropriate format to help dyslexic people with reading Japanese texts.

Accessible EPUB3 has almost same accessibility features as Multimedia DAISY. In addition, to make Multimedia DAISY/accessible EPUB3 contents more useful, we have developed a Windows application named "ChattyBooks" that can convert Multimedia DAISY/accessible EPUB3 into audio-embedded HTML5 with JavaScript ("ChattyBook") [2,3]. Popular browsers such

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 167–174, 2024.
https://doi.org/10.1007/978-3-031-62846-7_20

as Microsoft Edge, Google Chrome, Safari can play back ChattyBook contents which has the almost same functionality and operability of high-quality as the original Multimedia DAISY/accessible EPUB3. A special DAISY/EPUB3 player is no longer needed to play back. Multimedia DAISY, accessible EPUB3 and ChattyBook have almost same accessibility features, and here, we refer to e-books in those formats all together as "daisy books" (non-capital).

In Japan, "The Japanese Society for Rehabilitation of Persons with Disabilities (JSRPD)" [4] has been providing print-disabled students with daisy textbooks since 2008. It now organizes 23 volunteer groups/organizations and produces around 300 titles of daisy textbooks for elementary and junior-high school (the ages of compulsory attendance at school in Japan). Now, around 20,000 print-disabled students (mostly ones with developmental reading disorder, namely, dyslexia) use those textbooks. Unfortunately, however, it is still hard for many other dyslexic people in Japan to obtain daisy books. For instance, concerning senior-high school and higher education, no official organizations including JSRPD do not provide the dyslexic students with daisy textbooks, and many students who do need accessible textbooks due to the severe dyslexia decide to give up the idea of going to senior-high school or college.

As will be discussed later, in all over the world, various online services are operated for print-disabled people to get accessible books. They have been greatly contributed to improve print accessibility. Unfortunately, however, many of them do not necessarily meet some proper demands of dyslexic people. To give a practical/effective solution to this problem, we are developing "Chatty Library" [5] under the Nippon Foundation's grant [6] that is contributed to welfare-aimed innovative projects. It is aimed at providing dyslexic people with not only daisy books themselves but also a variety of other services to meet their demands. "ChattyBox" is a main function in Chatty Library [7]. In it, a user can play back a daisy book in Library with a popular browser. No special application is necessary to read it. By making use of our OCR (optical character recognition) technology, we provide them with an innovative service to convert their-uploaded PDF files into daisy books including a new type of daisy books, "Fixed-layout DAISY." Fixed-layout DAISY is designed so as to meet particular demands of dyslexic people; for instance, it can be produced almost automatically from original PDF.

Here, at first, we review briefly the situation of state-of-the-art online services to provide print-disabled people with accessible books. Next, some problems in those services are discussed, and to meet some demands by dyslexic people, our new approach: Fixed-layout DAISY and ChattyBox are introduced. Finally, some preliminary evaluations are shown.

2 Accessible Books for Dyslexic People

There are many online libraries for print-disabled people in all over the world. For instance, in USA, BookShare is well known as an online service to provide print-disabled people with accessible books. Print-disabled students can read various types of accessible books, more-than-1,200,000 titles, in the possession of BookShare free of charge [8]. Similar libraries are also operated in many other countries. Recently, reading environment for print-disabled people is definitely/remarkably improved.

In Japan, the Japan Braille Library runs such a library named "SAPIE" [9] that provides print-disabled people with more-than-250,000 Braille books and 120,000 DAISY books, most of which are audio-based DAISY (DAISY2.02). It organizes about 480 organizations such as public libraries, schools for the blind, etc., and over-80,000 print-disabled people use this service. Unfortunately, however, daisy books (multimedia type) are not substantial; they possess just around 500 titles, and dyslexic people unnecessarily use this library. While the population of dyslexic people is enormous compared to the visually disabled, there exists serious disparity in available accessible-book titles for each.

As was pointed out, as far as the dyslexic people are concerned, Multimedia DAISY (daisy books) is definitely the best solution as accessible books. However, we have to say that it has certain problems. As is well known, DAISY was originally designed for visually disabled people, and subsequently, dyslexic people have become to use. Thus, DAISY is designed so that various types of print-disabled people can read the same contents. For instance, most popular DAISY/EPUB3 players in the world do not have a function to read out text in a picture. Instead, they would read out an alternative text for the picture if it were available. We have to point out that the picture representation would be rather better for dyslexic people. They can see, and the picture representation should be easier for them to understand if texts in the pictures were read out. They should prefer that to the alternative text. In addition, some other features of Multimedia DAISY such as reflow, automatic aloud-reading by a DAISY player are not absolutely necessary requirements for the dyslexic people.

3 Fixed-Layout DAISY

To meet their demand, we have introduced a layer structure in daisy books. A picture is represented in multiple layers, the second layer of which is an image that appears in the completely same layout as the original print document. On the transparent front layer, text-related information such as characters (code), their coordinate on the page and bounding boxes (a rectangular area being circumscribed on a phrase as a highlighting unit), etc. is stored, and by making use of it, a DAISY (EPUB3) player could read out texts in the picture with a TTS (text-to-speech) voice together with highlighting those phrases. This multi-layer EPUB3 (DAISY) is authenticated by an EPUB3 validator [10], and it should be regarded as a kind of accessible EPUB3. Actually, some EPUB3 (DAISY) players in Japan can read out texts in a multi-layer picture. JSRPD has already adopted this type of daisy textbooks and provided print-disabled students.

The layer structure allows us to develop a new type of accessible e-books named "Fixed-layout DAISY" [11]. In it, the whole page is treated as a multi-layer picture, the second layer of which has the same form as the original print document (the page image). A DAISY (EPUB3) player can read out any texts on the transparent front layer together with highlighting phrases as same as a multi-layer picture put locally on a page (See Fig. 1).

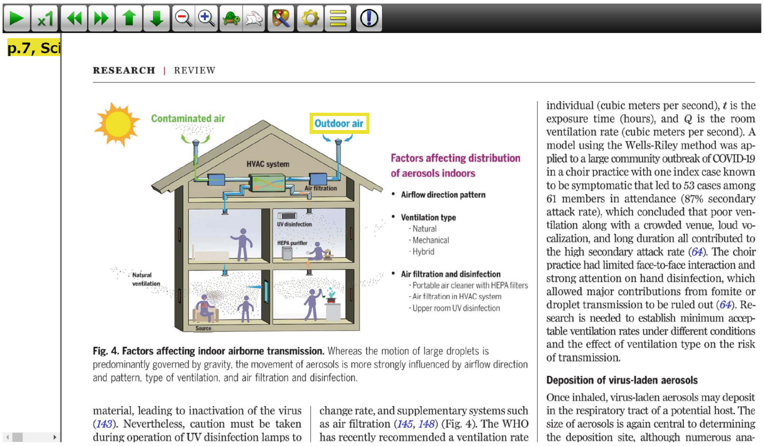


Fig. 1. Sample of Fixed-Layout DAISY: Wang, et al., "Airborne transmission of respiratory viruses," Science, 373 (2021) p. 981.

A main difference of Fixed-layout DAISY from ordinary DAISY is that it does not have the reflow function. The page layout is always kept as same as the original. It does not have information either in which order texts on the page should be read out. These features look disadvantage for print-disabled students to read the content. However, it should be pointed out that Fixed-layout DAISY has a certain advantage. It can be produced at extremely less cost than the reflow-type DAISY. Producing reflow-type DAISY requires you a lot of authoring process so that a DAISY player can read it out properly. As will be discussed later, Fixed-layout DAISY can be produced almost automatically from original PDF.

In Fixed-layout DAISY, its content cannot be read out automatically by a player. You need to click a position on a page where you want to start reading; then, the text is read out up to the end of the text block. Certainly, this mechanism may not work for visually disabled people; however, dyslexic people can see and click a place where they want to read. Fixed-layout DAISY should work for dyslexic people, we believe.

Fixed-layout DAISY should be more appropriate for an exam paper, in which keeping the original layout and enabling random access are usually more important than a textbook. Furthermore, it should be easier to introduce as reasonable accommodation for exams since Fixed-layout DAISY is clearly identical with the original PDF. In addition, it would rather be appropriate to produce the accessible version of a book other than textbooks/exam papers such as comics, picture books, (Pictorial) maps, etc.

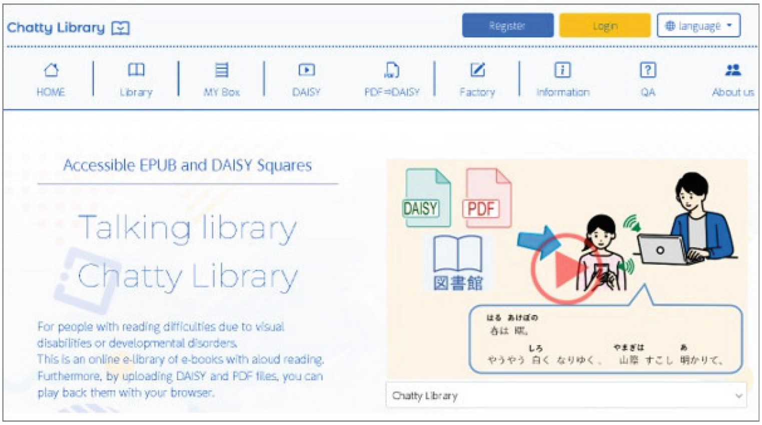


Fig. 2. Chatty Library top page.

4 ChattyBox

Fixed-layout DAISY allows us to develop innovative online services to provide dyslexic people easily with accessible books. As was mentioned, in Chatty Library (See Fig. 2), we realize several accessibility services as the functions of ChattyBox [7].

4.1 OS-Independent Daisy-Book Player

On ChattyBox, you can play back a daisy book with a popular browser. Using a Progressive Web Application (PWA), you can also store a daisy book in cache on your own computer/smart phone to play it back offline. You can upload your own daisy books onto ChattyBox to play them back directly on the web site. It can be used as an OS-independent DAISY player without installing any special (player) application.

By clicking the Upload button on ChattyBox, the dialog for uploading a daisy book appears. By clicking the leftmost "Play Back DAISY, " selecting the daisy book in the Upload Box (or drag-and-dropping it on the box and clicking the Convert button, the file is uploaded and converted into ChattyBook format. After completing those steps, the daisy book (ChattyBook contents) is stored in MyBox, and you can play it back with your browser.

4.2 Conversion of PDF/Text into DAISY

ChattyBox allows you not only to play back a daisy book but also to read your PDF/text file by converting it into a daisy book with audio, which can be played back online. You can also play back text information on the Windows Clipboard as a daisy book with audio as well as a text file.

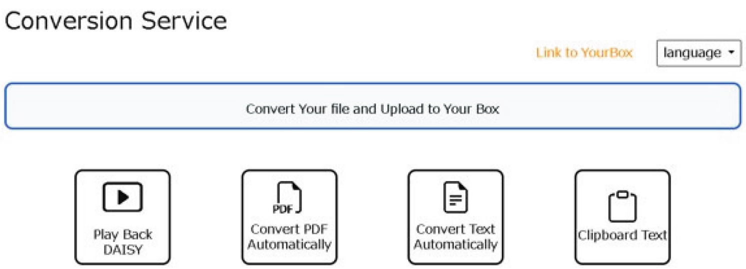


Fig. 3. Conversion Service dialog in ChattyBox.

By clicking the Upload button on ChattyBox and selecting the "Conversion Service," a dialog for conversion appears (See Fig. 3); you need to select either PDF, Text or Clipboard. In each choice, next, the detailed-setting menu for them appears, and you need to set the menu items for conversion. For instance, items in the pop-up menu for PDF conversion are as follows:

– Target page range (a number of pages that can be processed in one go is limited up to 10),
– Voice type (male or female; Japanese or English),
– Daisy type (reflow or fixed-layout),
– Read out characters in a figure or not, etc.

After saving the detailed settings, you can upload PDF in a similar way to the previous. Then, original file is automatically processed with our OCR software, "InftyReader," and a book icon appears on the ChattyBox window as the result. By clicking the Read button, the content is read out with your browser.

InftyReader is our OCR software for STEM (science, technology, engineering and mathematics) document [3,12]. It can recognize not only a text part but also a technical part such as mathematical formulas. All technical parts are automatically converted into accessible form. Mathematical formulas, tables, texts in a figure can be read out as well as text parts with the selected TTS voice. The current version of InftyReader uses "Google Cloud Vision API" [13] as an OCR engine that realizes very accurate recognition/conversion for not only text but also technical parts in a STEM document. Figure 4 shows the system flow in PDF conversion.

If Fixed-layout DAISY were chosen as the destination format, the conversion result usually would be almost perfect. In the case that some errors should be corrected, you can use two applications that can be downloaded from ChattyBox: "ChattyPad" and "LIF Editor". ChattyPad allows you to correct recognition errors in both text and mathematical parts. ChattyPad is the lite version of our STEM-document editor, "ChattyInfty" [12], in which you can edit any mathematical expressions as well as texts in an intuitive manner. If ChattyInfty were available, certainly, you could edit the conversion result with it. ChattyPad/ChattyInfty also allows you to edit contents so that any part of a page is

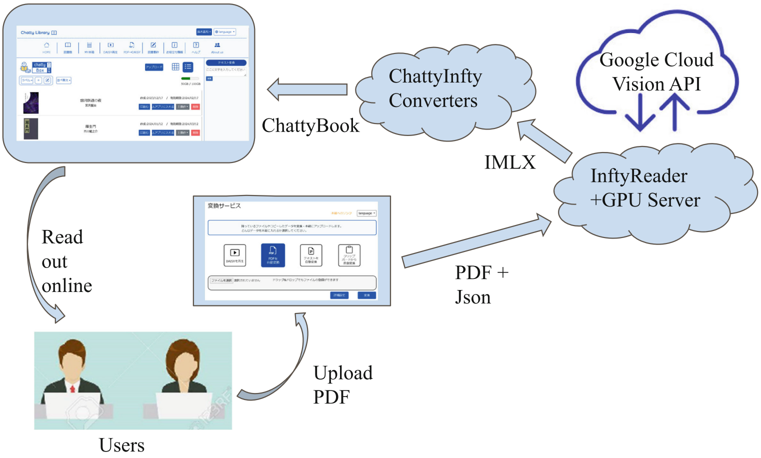


Fig. 4. System flow chart in PDF conversion.

read out in a correct manner with highlighting, by setting a bounding box on the page image and assigning a text to it.

InftyReader performs layout analysis to apply appropriate OCR engines to each block in a page, and the result is recorded as a JSON (JavaScript Object Notation) file. The blocks are classified into three categories: Text (including mathematical expressions), Table and Figure. For instance, if a text part were mis-labeled as Figure, it should give rise to serious recognition errors. Using LIF (layout information file) editor, you can fix errors in layout analysis. By executing the "Operation"-menu item, you can download the file for correction and can correct errors with these applications. By uploading the corrected file, the conversion result on ChattyBox will be revised as well.

As was mentioned, files in other formats besides PDF such as text files and contents on the Windows Clipboard also can be converted into (reflow-type) daisy books. Certainly, text files can be read out with a screen reader, but you cannot customize reading environment as in daisy books such as highlighting phrases, giving correct pronunciations etc.

5 Preliminary Evaluation

At some meetings in Japan, we have recently introduced Fixed-layout DAISY and ChattyBox. At a webinar on 26 February 2023, more-than-200 people (teachers, volunteer-group members, parents and publishers who are associated with education for dyslexic students) participated, and we listened to their opinions. We realized that majority of them gave it strong support. Someone said that Fixed-layout DAISY should be very useful for students with slight dyslexia who were in a transition process to a reader of ordinary print books.

At another webinar on 20 -21 January 2024 for about 20 persons who wanted to be a PC instructor for dyslexic students, we realized that by using ChattyBox, even a computer beginner became able to produce promptly a good daisy book from PDF. We also plan to hold several workshops/seminars/webinars on Fixed-layout DAISY and ChattyBox until the end of this year. Some results obtained in them will be reported in the ICCHP2024 conference.

6 Future Tasks

One of our remaining important tasks is to develop a method to infer correct reading order. We have been repeating one trial and error after another to find a good mechanism to give a correct order to text blocks in a page image; however, unfortunately, not yet. If we found it, truly automatic conversion of PDF into ordinary (reflow type) daisy books would be available on ChattyBox as well. We intend to improve ChattyBox so that a Markdown file and text-based EPUB3 also can be converted into (audio-embedded) daisy books.

Acknowledgement. We greatly appreciate support by the Nippon Foundation to achieve this project.

References

1. DAISY Consortium. http://www.daisy.org/
2. Yamaguchi, K., Kanahori, T., Suzuki, M.: "ChattyBook: Making Math Accessible for Online Education", the 36th CSUN Assistive Technology Conference, online, 09 March (2021)
3. Suzuki, M., Yamaguchi, K.: Conversion of multi-lingual STEM documents in E-Born PDF into Various Accessible E-Formats. In: Proceedings of the 18th International Conference on Computers Helping People with Special Needs (ICCHP-AAATE 2022), Lecco, Springer **LNCS13341** (2022), pp. 7–14. https://doi.org/10.1007/978-3-031-08648-9_2
4. Japanese Society for Rehabilitation of Persons with Disabilities (Japanese only). http://www.dinf.ne.jp/doc/daisy/book/daisytext.html
5. Chatty Library. http://www.chattylib.com/
6. Nippon Foundation. https://www.nippon-foundation.or.jp/en/
7. ChattyBox. http://www.chattylib.com/ChattyBox/
8. BookShare. http://www.bookshare.org/
9. SAPIE (Japanese only). https://www.sapie.or.jp/sapie.shtml
10. pagina EPUB-Checker. https://pagina.gmbh/startseite/leistungen/publishing-softwareloesungen/epub-checker/
11. Suzuki, M., Yamaguchi, K.: On Support for Dyslexic students in Senior-high School and Higher Education. Stud. Health Technol. Inform. **306**, IOS Press (2023), pp. 137–143. https://doi.org/10.3233/SHTI230608
12. Science Accessibility Net. http://www.sciaccess.net/en/
13. Google Cloud Vision API. https://cloud.google.com/vision/

Author Intent: Eliminating Ambiguity in MathML

David Carlisle[1], Paul Libbrecht[2], Moritz Schubotz[3], and Neil Soiffer[4](✉)

[1] Numerical Algorithms Group Ltd., Oxford, UK
david.carlisle@nag.co.uk
[2] IU International University of Applied Science, Erfurt, Germany
paul.libbrecht@iu.org
[3] FIZ Karlsruhe, Berlin, Germany
moritz.schubotz@fiz-karlsruhe.de
[4] Talking Cat Software Inc., Portland, OR, USA
soiffer@alum.mit.edu

Abstract. MathML has been successful in improving the accessibility of mathematical notation on the web. All major screen readers support MathML to generate speech, allow navigation of the math, and generate braille. A troublesome area remains: handling ambiguous notations such as $|x|$. While it is possible to speak this syntactically, anecdotal evidence indicates most people prefer semantic speech such as "absolute value of x" or "determinant of x" instead of "vertical bar x vertical bar" when first hearing an expression. Several heuristics to infer semantics have improved speech, but ultimately, the author is the one who definitively knows how an expression is meant to be spoken. The W3C Math Working Group is in the process of allowing authors to convey their intent in MathML markup via an `intent` attribute. This paper describes that work.

Keywords: Visual Impairment · Assistive Technology · Speech · STEM · Mathematics · Formulas · MathML

1 Background

The W3C first recommended MathML as the method for including mathematical expressions in web documents in 1998. Browser adoption was slow, but by early 2023, all the major browsers supported MathML. Support for MathML by screen readers came along many years before that milestone. The quality of the speech and the number of languages supported (both in speech and in braille) varies. Some math-specific Assistive Technology (AT) software[1] such as SRE [3], MathPlayer [13], and MathCAT [14] put significant effort into inferring

[1] The focus of this paper is on mathematical expressions. Accessibility of expressions involves speech, navigation, and braille. We do not discuss other important accessibility topics such as plots/graphs and diagrams.

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 175–183, 2024.
https://doi.org/10.1007/978-3-031-62846-7_21

x^T	"x to the T" or "x transpose"
x'	"x prime" or "first derivative of x"
$\overline{AB}$	"line segment A B" or "complex conjugate of A times B"
(a, b)	"point a comma b", "open interval from a to b" or "gcd of a and b"
$x\|y$	"x divides y", "x such that y" or "x given y"

Fig. 1. Examples of ambiguous notations

the semantics so that superscripts are not always powers and pairs of vertical bars do not always mean absolute value. Some examples of ambiguity are given in Fig. 1.

One way to avoid ambiguity is to speak the expression syntactically. For example, $|x|$ can be spoken as "vertical bar x vertical bar" and $\overline{AB}$ can be spoken as "start A B end grouping with line above." To our knowledge, there have not been any studies that compare listener preference for syntactic speech vs. semantic speech. It seems likely that semantic speech is preferable because syntactic speech is usually not what people are used to hearing and is also often longer. Comparing the speech for $|x|$ and $\overline{AB}$, syntactic speech requires 9 and 10 syllables, respectively, versus 6 and 5 syllables for semantic speech.

The most common approach to generate semantic speech is to infer what the author means by looking at the notation and its arguments. For example, to distinguish between absolute value and determinant, a single capital letter or a square table as the argument between verticals bars (e.g., $|M|$) would likely be a determinant, not an absolute value [11]. To varying degrees, AT looks at the arguments to generate speech for expressions like x^2 ("x squared"). MathPlayer [13] goes further than most; it has over 800 patterns to improve speech. About 200 of these patterns are only active if the user specifies a subject area (this number includes chemistry speech rules). For example, if the user chooses the subject area "probability & statistics", then $\overline{x}$ is read as "mean of x" instead of "x bar". If "calculus" is chosen, $\times$ is read as "cross product" rather than "times". This helps resolve ambiguity at the cost of having the user inform MathPlayer about the content that is being read. MathPlayer's use of subject area was a motivating factor for the W3C Math Working Group to explore adding "intent". However, it is not rare to see texts that use the same notation for different concepts, e.g. (a, b) for both the open-interval and the coordinates of a point.

Another approach is to use the surrounding textual context to understand the math content [12]. In 2017, we presented a method that extracts definitions for identifiers with an F1 score of 36%. Language models are rapidly improving. In 2023, Bansal, et al. [1] described a matching learning approach to recognize definitions of symbols used in an expression by looking at the immediate surrounding context. For example, their work deduces from "Let $F : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be a C^1-vector field" that F is a C^1 vector field. The paper from 2022 lists an

F1 score of 75% in their data set for finding a definition. However, no numbers indicate how often this is useful for improving speech, which is the final goal of their work.

For large expressions, a number of people have advocated that an overview (or outline or summary) of an expression be given. AsTeR [9] automatically elided subexpressions, but no study was done on its effectiveness. As part of the MathGenie project, a study [6] showed that providing an outline slowed solution time. Nonetheless, outlines were included in MathGenie because the authors felt it would be useful. MathPlayer provides an option to describe an expression rather than read it. In a ClearSpeak navigation study [5] using MathPlayer, user feedback was that outlines were not very useful. The study authors feel part of this is because the implementation was crude relative to other features. As with reading, ambiguity can arise when summarizing expressions.

In addition to speech and braille, the ability to review a portion of an expression is important in expressions that are not simple. Most AT allows for navigation of expressions via a tree-based model, not unlike the MathML representation of the expression. In [7], the authors use a touch/tactile-based approach to navigation for people who are blind but have some small amount of residual sight (enough to resolve light/dark). This allows the users to take advantage of the physical relationships in mathematical notations (e.g., numerators are above denominators) that sighted users take advantage of. They compare their prototype for an iPhone with a tactile grid overlay to JAWS and find statistically significant benefits for touch including less frustration/effort and faster relocation of items. While spatial navigation is more user-friendly, the problem of ambiguous notations is still present.

In [8], the authors note that the use of audio for reading maths textbooks is on the rise, sometimes as an alternative to braille. They point out that the right to learn to read (braille) should be supported. The paper stresses the lack of high-quality studies on the topic of mathematics learning for visually impaired.

While previous work has improved the quality of the generated speech, heuristics can never be perfect and are ultimately guesses as to what the author meant. Furthermore, even when AT knows what the author means, that doesn't necessarily indicate how the author wants a notation pronounced. For example, 1/3 can be read as "one divided by three" or "one third" depending on what is being learned. The work presented here aims to support AT to generate better speech and, in some cases, better braille from MathML expressions.[2]

2 Author Intent in MathML 4

The above approaches significantly improve the understandability of the generated speech. However, they are still heuristics and thus sometimes wrong.

[2] Most braille codes are based on the basic structure of the expression (subscripts, superscripts, ...), not on semantic meaning (index, power, ...). There are a few exceptions to this, e.g. whether ":" is meant to convey a ratio or something else in Nemeth code. "Intent" can also help with braille generation in those special cases.

```
<mrow intent="absolute-value($contents)">
  <mo>|</mo>
  <mi arg="contents">x</mi>
  <mo>|</mo>
</mrow>
```

Fig. 2. Simple intent example

To complement these efforts, the W3C's Math Working Group is updating the MathML standard to allow specification of how an expression should be spoken [2]. Authors can use this standard to correct heuristics, or AI researchers can evaluate the performance of their heuristics. Following the idea of correcting heuristics, the W3C's Math Working Group decided that an approach that uses progressive enhancement is most appropriate: do not require changes; instead, allow for those notations where an author wants to make sure of unambiguity. As a rule of thumb, an author might want to enhance notations in cases when she would explain it in a classroom or at presentation, e.g., when she would write x^T she might say "T means transposed".

2.1 Author Intent Basics

The approach the Math WG settled on is to allow `intent` and `arg` attributes on all MathML elements. The attribute's value has a simple, functional syntax. This syntax allows both the function head (the function name along with its properties) and its arguments to be literals, references to descendant elements, or another function. Literals can be numbers or names. See Section 5.1 of [2] for a full grammar. A simple example for the "absolute-value" concept is shown in Fig. 2.

References are prefixed with `$` and some descendant of the referencing element should have a corresponding `arg` attribute value. In Fig. 2, this is demonstrated with the reference "contents". See Sect. 2.2 for details.

By default, "intent" values should be spoken as functions are spoken, so the expression in Fig. 2 might be spoken as "absolute value of x", but AT is free to use other functional ways of speaking the "intent", such as a terse form ("absolute value x") or a verbose form ("the absolute value of x"). Properties (see Sect. 2.3) allow for other ways of speaking an "intent".

In parallel with the MathML 4 recommendation, the working group is fleshing out a core list of "intent" concepts and "intent" properties with proposed speech hints.[3] This list is a reference for notations/speech that AT implementations should support. The core list is intended to cover most mathematics taught up to the university level. The list includes suggested speech in a few different languages. As a complement, an open list of "intent" concepts and properties is maintained; new notations are constantly created so the open list will never be complete. The open list serves both as a place where people can check to see if

[3] Current working drafts are linked from https://w3c.github.io/mathml-docs/.

```
<msup intent="power($base,$n)">
  <mrow arg="base">
    <mo>(</mo>
    <msup intent="power($base,$n)">
      <mi arg="base">x</mi><mn arg="n">2</mn>
    </msup>
    <mo>+</mo>
    <msup intent="power($base,$n)">
      <mi arg="base">y</mi><mn arg="n">2</mn>
    </msup>
    <mo>)</mo>
 </mrow>
 <mn arg="n">2</mn>
</msup>
```

Fig. 3. Example of nested arguments in “intent”

an “intent” concept has already been thought about and as a source of future additions to the core list. AT is free to implement any concepts or properties in the open lists.

2.2 Intent Concepts

The function name in an “intent” is referred to as the concept name in MathML 4. If the AT knows nothing about the concept name, it should be spoken as written. However, the working group’s “core” list provides names for which AT should be aware of and for which it should provide translations for the languages it supports. Some of these concepts, such as “fraction” have many ways they are spoken depending on the arguments (e.g., “one third”, “one over x”, “one over x all over two over x”, “meters per second”). For someone who is blind, some of these speech patterns may include start/end words or sounds to make it clear where the fraction starts and ends; for others, these extra words or sounds may hinder comprehension. Authors rarely know their readers’ needs, so MathML 4 delegates the exact speech for core concepts to AT. If an author wants to force specific words to be used, a concept name can start with an underscore; no core names start with an underscore.

Concept names are not always a name; they can also be a reference. A reference (either the concept name or an argument) can be any child with an `arg` attribute. References start with a `$` character. The reference does not need to be unique in the document. This allows generating software to reuse templates. The algorithm for finding a reference is to do a depth-first search of the children stopping when a matching `arg` attribute value is found. If the `arg` attribute value matches the reference, the search is done. Otherwise, the element is treated as a leaf and the search continues in the parent. Figure 3 shows an example with nested “intents” for a nested power $\left(x^2 + y^2\right)^2$ that might come from software that uses a template for powers. If an `intent` has illegal syntax or references nonexistent `arg` attributes, the `intent` should be ignored by AT.

2.3 Intent Properties

By default, concept names are spoken in a functional manner, but this is not always appropriate. For example, $x^2\big|_3$ might have the concept name "evaluated-at" and is typically spoken as "x squared evaluated at 3" not as "evaluated-at of x squared and 3". To solve this problem, "intent" can be given a "fixity" property. The allowed values are "function", "silent", "prefix", "infix", and "postfix".[4] Properties begin with ":" and there is no limit to the number of properties that can be attached to a concept name. For "evaluated-at", we might have

```
<mrow intent="evaluated-at:infix($expr, $value)"> ... </mrow>
```

Early on, the Math WG realized that some notations that make use of the `mtable` element are complicated to specify using just concept names. For example, each equation in a system of equations is often divided up into columns to force alignment. To bring each equation back together, it would be necessary to list all the entries in each row as part of an "equation" concept. To remedy this, table properties tell AT how to speak the children. For a system of equations, the table can be marked with the "intent" value "system-of-equations" and AT should ignore the columns and just speak the table as (for example) "2 equations, equation 1 ..., equation 2, ... end equations". The current list of core table properties is "matrix", "piecewise", "system-of-equations", "lines", and "continued-row".

Other properties are used to avoid having generating software know lots of related names. These include properties for chemical elements, units, and roman numerals. There is also a "chemical-equation" property that notifies AT that chemical notation is being used so subscript, superscripts and some operators are spoken appropriately (e.g., "=" is a double bond, not "equals").

2.4 Intent and Content MathML

In MathML, two families of elements are defined: Presentation MathML encodes how expressions are set out typographically with such typical features as subscripts/superscripts, fractions, or bracket-pairs. Content MathML encodes how expressions are understood or interpreted with typical features as function applications, quantifiers, or externally documented symbols.

MathML expressions, i.e., semantics elements, thus can have two trees: a content tree and/or a presentation tree. Linking between content and presentation elements can be done with references or nested semantics elements. Content MathML is made available by and for computing engines. Some translations between presentation and content (with many assumptions) exist with limited scope (e.g., for simple equation expressions). While there may be an interest in

[4] The Math WG is still considering adding other values such as "matchfix" to allow for other speech patterns.

finding on the web expressions with content MathML (e.g., to allow readers to perform computations), they are much less frequent than presentation MathML.

A few experiments led by members of the W3C Math Working Group have shown that content MathML can be used to generate a better accessible presentation of mathematical formulas if the semantic is available. However, the fact that content MathML is less widespread and sometimes less able to encode all mathematical discourse (without adding many symbols) has pushed the W3C Math group to propose a structure that is closer to the speech and that applies to expressions which are directly made in presentation MathML.

3 Evaluating the Success of Intent

While "intent" is aimed at improving speech for end users, the main target for "intent" are authors of documents that contain MathML. The reason tests for end users are not a focus is because "intent" should only improve the accessibility of math and never make it worse: "intent" gives AT the ability to *know* the author's intention of how the notation should be spoken as opposed to having to *guess* the best way to speak the notation. AT has the choice to make use of this information or ignore it (e.g., to produce a syntactic reading rather than a semantic one). Therefore, the way to measure the success of "intent" is two-fold:

1. Are software developers that generate or consume MathML generating or consuming "intent" or planning to generate or consume "intent"?
2. For the authoring software that generates "intent", do users make use of facilities provided so the software can generate "intent"?

It is very early to evaluate either of these criteria given that the MathML 4 recommendation has not even moved to a Candidate Recommendation, let alone an actual recommendation. For the first question, there are definite signs of success as both MathML generators and AT that consumes "intent" have prototypes.

On the generating side, all of the prototype authoring tools that have been developed are text-based and have some resemblance to TEX. In general, the authoring tools use macros or special syntax such as "`\abs`" and '`\det`" to pass along the author's intent when there is ambiguity. Discussions have also included the use of optional macro arguments to pass along intent information. For example: `\times[intent=cross-product]` would produce the following MathML: `<mo intent="cross-product">×</mo>`

Three prototypes have been developed so far:

WikiTexVC: [15] has the option to add valid "intent" syntax with a pseudo TEX macro to add "intent" to the MathML produced by Wikipedia, and other projects using MediaWiki.

UnicodeMath: [10] adds keywords that are Unicode characters when needed to resolve ambiguity and uses those to generate intents. For example, "⒜x" uses the Unicode code point "⒜" (U+249C) to indicate that what follows is the absolute value of x. This is used in speech generation.

SpaceMath: [4] a superset of LaTeX (both text and math) and AsciiMath that includes keywords and macros to generate intents. It will likely become an option for PreTeXt authors.

No WYSIWYG editors that make use of “intent” have been developed yet, but ideas on how they might do this involve the use of specialized templates such as one for binomial that would generate an appropriate “intent” even if using the MathML-code as that of a 2d-vector. Another option is to allow users to select a symbol or expression and provide a menu of options for that symbol or expression. For example, selecting “×” might pop up a menu with the options: “times”, “cartesian-product”, “cross-product”, “direct-product”, and “custom…”.

On the consuming side, both UnicodeMath and MathCAT implement “intent”. Support for “intent” is included in the release version of MathCAT that is used in NVDA and JAWS and several other ATs. To generate speech, MathCAT first produces an intent tree from MathML. This is trivial if “intent” is given. If there is no `intent` attribute, then MathCAT tries to infer the intent using heuristics. The “intent” tree is then used to generate speech in various languages and in different styles of speech. The Math Working Group created a document with many examples comparing MathCAT’s speech with and without “intents”. Among the lessons learned from this exercise was that “intent” properties for tables such as those used for aligned systems of equations greatly simplified MathML generation without complicating the implementation.

At this point in time, we lack information as to users’ willingness to use features that allow generation of “intent”. This is because the number of users of these prototypes is small and mostly includes the software authors and their colleagues. However, all have reported that generating “intent” is relatively straightforward. One of the PreTeXt implementers reported that authors are generally willing to improve the accessibility of their books if it is not much of a burden [D. Farmer, personal communication, 4 April, 2024]. As with many accessibility features (e.g., using headings in documents rather than just changing the font), getting users to use styles/keywords/macros likely requires education as to their benefits. All the prototypes try to minimize any extra work an author needs to do to improve the accessibility of the math.

4 Conclusions

The development of author intents has taught us how flexible mathematical notation can be and how this flexibility is important to mathematicians in their day-to-day practice; all mathematicians we have talked to indicate that inventing a new notation, and being able to exploit it, is common. This flexibility is supported by author intents which make it possible to encode new or existing concepts on any MathML expression which results in speech that is accessible to the reader.

Different communities use different notations. Where these notations overlap, there is ambiguity. It does not seem possible to cleanly partition these communities because they overlap. Because of this, the only way to resolve ambiguity is at the notation level where the author tells the AT what should be said. As an example, the last row of Fig. 1 $x|y$ demonstrates a notation that could be used in a number theory course with the sense of integer division, as such-that in a proof, and of conditional probability all in one paragraph.

It is important to note that "intent" is forward-looking: only new documents can use it. Earlier work using heuristics and textual context are still important to handle legacy documents.

References

1. Bansal, A., Sorge, V., Balakrishnan, M.: Extracting contextual semantic from a concordance containing mathematical definition. Stud. Health Technol. (2023). https://doi.org/10.3233/SHTI230607
2. Carlisle, D.: Mathematical markup language (MathML) version 4.0, editors draft (2023)
3. Cervone, D., Sorge, V.: Adaptable accessibility features for mathematics on the web. In: 16th International Web for All Conference, pp. 1–4 (2019)
4. Farmer, D.: Space Math – A Mathematical Notation System & LaTeX Translator (2024). https://github.com/davidfarmer/SpaceMath
5. Frankel, L., Brownstein, B., Soiffer, N.: Expanding audio access to mathematics expressions by students with visual impairments via MathML. Technical report, ETS (2017). https://doi.org/10.1002/ets2.12132
6. Gillan, D.J., Barraza, P., Karshmer, A.I., Pazuchanics, S.: Cognitive analysis of equation reading: application to the development of the math genie. In: Miesenberger, K., Klaus, J., Zagler, W.L., Burger, D. (eds.) ICCHP 2004. LNCS, vol. 3118, pp. 630–637. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-27817-7_94
7. Howell, J., Quek, F.: Math for those with severe low vision: from the particulars to the gestalt (and back again). In: IEEE Frontiers in Education Conference, pp. 1–9 (2023). https://doi.org/10.1109/FIE58773.2023.10343509
8. Klingenberg, O., Holkesik, A., Augestadt, L.: Digital learning in mathematics for students with severe visual impairment: a systematic review. J. Vis. Impairment **38** (2019). https://doi.org/10.1177/0264619619876975
9. Raman, T.V.: Audio system for technical readings. Ph.D. thesis, Cornell University (1994)
10. Sargent, M.: Welcome to UnicodeMath (2024). https://unicodemath.org/
11. Schubotz, M., et al.: Semantification of identifiers in mathematics for better math information retrieval. In: SIGIR 2016, pp. 135–144 (2016). https://doi.org/10.1145/2911451.2911503
12. Schubotz, M., Krämer, L., Meuschke, N., Hamborg, F., Gipp, B.: Evaluating and improving the extraction of mathematical identifier definitions. In: Jones, G.J.F., et al. (eds.) CLEF 2017. LNCS, vol. 10456, pp. 82–94. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-65813-1_7
13. Soiffer, N.: Mathplayer v2. 1: web-based math accessibility. In: 7th ACM SIGACCESS Conference, pp. 257–258 (2007)
14. Soiffer, N.: MathCAT: Math Capable Assistive Technology (2024). https://nsoiffer.github.io/MathCAT/
15. Stegmuller, J., Schubotz, M.: Wikitexvc: Mediawiki's native latex to MathML converter for Wikipedia (2024). https://doi.org/10.48550/arXiv.2401.16786

PreTeXt as Authoring Format for Accessible Alternative Media

David Austin[1], Robert Beezer[2], Michael Cantino[3], Alexei Kolesnikov[4], Al Maneki[5], and Volker Sorge[6](✉)

[1] Department of Mathematics, Grand Valley State University, Allendale, MI, USA
austind@gvsu.edu
[2] Department of Mathematics and Computer Science, University of Puget Sound, Tacoma, WA, USA
beezer@pugetsound.edu
[3] Northwest Regional Education Service District, Hillsboro, OR, USA
mcantino@nwresd.org
[4] Department of Mathematics, Towson University, Towson, MD, USA
akolesnikov@towson.edu
[5] National Federation of the Blind, Baltimore, MD, USA
[6] School of Computer Science, University of Birmingham, Birmingham, UK
Volker.Sorge@cs.bham.ac.uk

Abstract. We describe an approach for producing accessible mathematical content for blind learners, i.e., learners who primarily use screen readers and tactile materials to access information necessary for learning. Our main focus is on postsecondary education textbooks required for a course of study towards an undergraduate degree in mathematics. The basis for our approach is the PreTeXt authoring language, which combines a rigorous XML document structure with the strength of LaTeX typesetting for formulas and graphics. This not only facilitates the production of formats such as PDF and EPUB but, more importantly, fully accessible HTML and tactile braille output from a single source file. The braille version of the text includes mathematical formulas in Nemeth braille as well as embossable tactile diagrams automatically generated from the source. Similarly, formulas in the HTML version are screen reader accessible, and the graphics can be explored interactively using screen reading and sonification. We applied our software to two large open-source textbooks on abstract algebra and calculus. The quality of the braille transcription was checked by a certified transcriber and the overall readability was checked by a blind mathematician. We have since experimented with the automatic translation of other mathematical textbooks to provide to blind student learners and volunteers.

1 Background

Timely access to high-quality educational materials for blind students in the US has long been a problem [11]. In this paper, we are focusing on accessibility for

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 184–191, 2024.
https://doi.org/10.1007/978-3-031-62846-7_22

learners who primarily use screen readers and tactile materials to access learning materials. The issue is particularly acute for such learners in STEM fields where auditory versions of books are not always sufficient for learning. Students who need a textbook in a technical field often report long lead times required to obtain a braille copy of the book. Compounding the problem, publishers of postsecondary textbooks have increasingly moved content online creating additional discrepancies between the web version, printed version, and braille version of the book.

Transcribing a new textbook is a time-consuming and costly process because it is done manually by a certified specialist. Even if the source file for a mathematics textbook is available, several factors make transcription into braille particularly challenging for mathematical content. We briefly describe the main challenges before explaining how our approach addresses them.

Translation of non-technical "literary" text into braille is quite straightforward using the open-source tool liblouis [9]. Unfortunately, when applied to mathematics, liblouis generally produces faulty Nemeth code. Moreover, the correct integration of Nemeth into literary braille is non-trivial and needs to be carefully done even for simple texts.

Most college-level mathematics books are written and produced using the typesetting system LaTeX, which is designed to create excellent print output, especially for mathematical formulas. Unfortunately, existing LaTeX packages [15, 16] for braille output produce poor results in part because the structure of a LaTeX document does not have sufficient semantic information to produce reliable translation to braille. Consequently, a more structured formal mark-up language is needed.

There have been a number of projects that generate tactile and accessible content from LaTeX. As early as 1998, the Labradoor system [5] specialized in conversion of LaTeX documents to braille, including formulas in Marburg. Later work [2,14] updated and generalized this approach by translating LaTeX formulas into canonicalized MathML. Alternative approaches attempt to directly translate LaTeX into accessible PDF [3] or into web accessible HTML [17]. While the former simply shifts the accessibility issue to the need for making PDF accessible, which is still inherently inaccessible without proprietary technology, the latter relies on transforming LaTeX formulas to MathML, which loses valuable mathematical information.

Consequently, these approaches that first translate formulas to MathML are generally too restrictive as much of their semantic structure is lost, leading to incorrect braille or speech in the context of advanced mathematics. Preserving LaTeX formulas is also important for communication between blind mathematicians and their sighted peers as it is the standard markup format used by mathematicians [13]. In addition, many approaches focus solely on text, document structure and formulas, thereby neglecting the need to translate artefacts such as graphics and code (both pseudo code and program snippets) that are an important part of mathematical and computer science literature. While there have been great strides in making accessible graphical statistics content from R into LaTeX or markdown [18] and graphics [10], these advances are not easily transferable to the conversion of textbooks in other mathematical disciplines. For the inclusion of code, we still see all too often that publishers convert listings elements into graphic figures with uninformative alternative text, effectively obscuring this information. Finally, none of these systems is readily available for a production process or flexible enough to support different languages, braille formats, or output modalities.

The approach taken here uses PreTeXt, an authoring language for structured scholarly documents, which combines a rigorous XML document structure with the strength of LaTeX typesetting for formulas [6], graphics [4], and code [1]. It offers many features designed to support writing STEM textbooks and learning materials with an emphasis on Open Educational Resources (OER). Authoring documents in PreTeXt allows users to integrate computational features such as Sage or GeoGebra and to generate a range of output formats from a single source.

This paper describes our approach employing PreTeXt as the source to generate a number of accessible alternative formats. In particular, we concentrate on the automatic generation of tactile output that includes literary braille, Nemeth code for formulas, and tactile graphics for diagrams as well as on the generation of web accessible content that allows for screen reading and interaction with formulas and diagrams. However, the PreTeXt format is by no means restricted to these and can generate different output formats, languages, braille codes, etc. PreTeXt aims for a holistic approach for translating all components of a scientific document, including formulas, code, graphics, and interactive computational elements, without the need for different technologies. Moreover, it exploits the rich information provided by LaTeX, including numerous specialist packages, without losing information by re-coding into an intermediate format.

2 Background on PreTeXt

PreTeXt is an XML markup language developed to produce technical scholarly texts in a variety of formats. PreTeXt allows users to include interactive features such as Sage, GeoGebra, and WeBWorK into the documents. A short sample of PreTeXt format is provided in Table 1. In addition to the obvious document structure, the document narrative consists of clearly identifiable paragraph text (or literary text) and mathematical formulas in LaTeX, enclosed in tags `<m>` and `<me>`. PreTeXt vocabulary supports more than 300 different elements, described fully in [21].

Since PreTeXt is a markup language, the author specifies the document structure and content, which are in turn rendered by the software in any of the specified formats. There are available translations to (1) print PDF, (2) electronic PDF, (3) accessible HTML, (4) EPUB, (5) Jupyter notebooks, and (6) braille, either with pagination for embossing or without pagination for screen readers. We stress that the conversion is applied to the same source material, without additional efforts to re-style the source. We refer the reader to [6] for more information on the vocabulary and for examples of different outputs of PreTeXt.

Table 1. Sample of a PreTeXt source file.

```
<?xml version="1.0" encoding="UTF-8" ?>
<pretext>
  <article xml:id="final-exam">
    <title>Precalculus Final Exam</title>
    <section xml:id="p1">
      <title>Problem 1</title>
      <p> The height <m>s</m>, in feet, of a water balloon that is thrown
      vertically out of a window, <m>t</m> seconds after it is thrown is
      given by
      <me> s(t) = -16t^2 + 16t + 32. </me>
      </p>
      ...
</pretext>
```

The approach of PreTeXt developers, described in [6], is well-suited for the production of accessible versions of texts in technical fields. In particular, this approach fits well with the task of translating a technical document into braille, as we explain in the following section.

We briefly describe accessibility features available in PreTeXt's HTML output. Some of these accessibility features are available automatically, without any input from the author. These features include screen-reader friendly HTML elements and attributes; keyboard-accessible navigation of the produced HTML document; and screen-readable mathematics, provided by the Speech Rule Engine (SRE) [19]. The SRE produces speech strings for screen-readers from mathematical expressions in LaTeX and implements a full set of Mathspeak and Clearspeak rules. For each mathematical expression, the SRE constructs an internal semantical representation, which enables interactive user hierarchical exploration of complicated expressions in mathematically meaningful order. In a similar manner, SRE produces linearized Braille output that can be made available to readers via a Braille display during exploration.

Additional accessibility features of PreTeXt allow the authors to provide information to make the documents more accessible. One of the main such features is the ability to supply descriptions of images beyond alt-text. The content of these longer descriptions is picked up by screen readers and will be rendered as a transcriber note when the document is translated into braille.

3 Transforming PreTeXt to Braille

To produce a correct braille copy of a document, (human) braille transcribers have to assess the semantics of the translated documents. For example, they need to know not only that a certain phrase appears in bold in the printed version, but also *why* it appears in bold. The rich semantic structure of a PreTeXt document supplies this information leading to the more reliable production of a braille version. In addition, the capability to produce a braille translation from the same source document removes the cost of producing an accessible copy of the textbook, makes it possible to produce an accessible version of a technical document in a matter of minutes, and solves the problem of version lag between different output formats.

So how is the translation to braille achieved? The source file is processed by a Python script. Literary text is transcribed into braille using liblouis, while mathematical formulas are transcribed by the Speech Rule Engine (SRE). The outputs of liblouis and the SRE are then integrated by the script into a single file in BRF format. Files in this format can be used in any embosser or in a braille display. The script ensures that the document is formatted according to the BANA rules and that Nemeth braille produced by the SRE is integrated properly into the text.

The challenge of accurately and automatically translating mathematics formulas deserves a more detailed discussion. Sometimes, the same print mathematical symbol may need to be transcribed differently into braille depending on the context. This raises the possibility that "fixing" the translation of a symbol in one context may break the correct translation of the same symbol in other contexts. A quality assurance mechanism was developed to identify these situations and to ensure that software updates do not create problems in already proof-read and correct material. The XML structure of PreTeXt sources makes it possible to easily create a comprehensive list of all the mathematical expressions in a PreTeXt document and to pair expressions with corresponding links into the web version of the document. With these links in place, the context of each expression can be easily examined, which aids transcribers in identifying mistakes in the automatic transcription. As a by-product of this work, we identified mathematical symbols and fonts that are missing in Nemeth and worked with BANA to address this. As a result, several missing features are being added to Nemeth (for example, the double-struck font that is used to denote the sets of integers, real numbers, and complex numbers; as well as symbols used in mathematical logic and computer science).

4 Generating Accessible Graphics

PreTeXt allows authors to specify graphics either in TikZ [20], the prevalent tool for authoring diagrams in mathematical literature, or in its own XML language. The latter provides a relatively small, but intuitive, vocabulary to declaratively specify diagrams. This language forms the basis for translating graphics into tactile output or web accessible SVG. The basic vocabulary, which is designed to mirror mathematical vocabulary, can roughly be divided into five functional groups: (1) definitions, control, and grouping, (2) basic grids and axes, (3) one-dimensional objects, such as graphs and lines, (4) two-dimensional objects, such as areas and polygons, and (5) labels.

This allows the generation of SVG graphics that can be rendered as tactile output, where any label or contained formula is pre-transcribed into the braille format of the overall text and its corresponding braille code for mathematics.

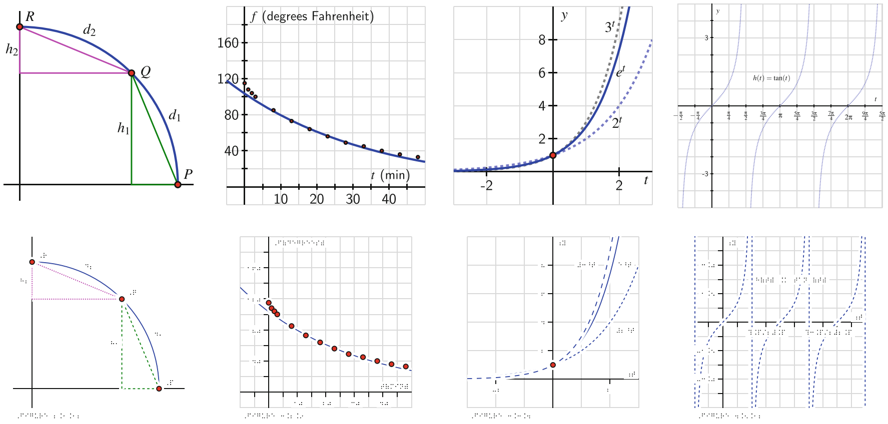


Fig. 1. Examples of mathematical diagrams from [7] and their translation to braille.

Figure 1 gives some examples from a Precalculus textbook. The tactile versions of the graphs are scaled to fit a standard-size braille page. Graphics are not restricted to xy-graphs; complex diagrams like geometric examples or depictions of networks are supported as well.

In addition, the PreTeXt specifications allow annotations to add speech descriptions to graphical elements. These are either added automatically for common elements, such as the grid and axes, or can be customised by authors. Annotations are arranged hierarchically to describe elements starting from an overview to more and more fine-grained details. When translating the diagram into SVG, these elements are embedded into the SVG to enable accessibility support in web browsers, such as dark mode support or magnification and highlighting of semantically connected structures. In particular, when using screen reading in the browser, the embedded annotations provide the speech output to explain elements of the graphic. Moreover, a diagram can be interactively explored using the hiearchical nature of the annotations as a navigational path. This allows readers to obtain an overview of the entire graph before diving into details of particular elements, such as the axes, graphs, or labels and formulas. All mathematical formulas in the diagram are pre-rendered with MathJax including aria-labels to provide speech output. Similarly, graphs can be sonified using the browser's web audio API.

5 Experiments and Usage

We created and tested a set of open-source tools to automate the production of braille versions of college-level mathematics texts. To measure the progress, two open-source textbooks were used to test the quality of translation: "Abstract Algebra: Theory and Applications" [12] and "Active Calculus" [8]. The screen-readable web versions and the braille versions of the textbooks were made available to blind learners of mathematics. In addition, we worked with one particular blind learner using the translation of a textbook "Active Precalculus" [7] into a tactile version.

The learner was provided with the version of the textbook on the web (and used a screen reader to interact with the textbook). We also provided the learner, who is familiar with Nemeth braille, with an electronic braille file in BRF format and with hard copies of embossed graphics, including those given in Fig. 1.

Over the course of four months, the learner worked through a semester-long course based on the textbook, doing all the exercises at the end of each section. We then offered the learner a final exam, based on a final exam routinely given in a Precalculus course. All the problems on the final exam were solved correctly by the learner.

The learner reported that the braille file was initially necessary for the more complex formulas, where the screen reader is "too fast" or provides ambiguous results (for example, "square root of x plus one" could stand for $\sqrt{x}+1$ or for $\sqrt{x+1}$). To save money on the printing costs, the learner used a refreshable display to render the formulas. After gaining more experience with the interactive exploration of complex formulas provided by the SRE, the learner was primarily using the braille display to learn braille codes for new mathematical symbols.

To provide a completely accessible copy of a mathematical textbook, a number of challenges remain. While the learner was able to use WeBWorK to do some of the exercises, several problems that involved WeBWorK-generated graphics were not accessible. Additional work is needed to ensure that all components supported by PreTeXt have robust accessibility solutions.

References

1. Ali, Z., Austin, D., Sorge, V.: Accessible papers from source: the quest for the holy grail. In: 5th International Workshop on Digitization and E-Inclusion in Mathematics and Science (2024)
2. Archambault, D., Fitzpatrick, D., Gupta, G., Karshmer, A.I., Miesenberger, K., Pontelli, E.: Towards a universal maths conversion library. In: Miesenberger, K., Klaus, J., Zagler, W.L., Burger, D. (eds.) ICCHP 2004. LNCS, vol. 3118, pp. 664–669. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-27817-7_99
3. Armano, T., Capietto, A., Coriasco, S., Murru, N., Ruighi, A., Taranto, E.: An automatized method based on LaTeX for the realization of accessible PDF documents containing formulae. In: Miesenberger, K., Kouroupetroglou, G. (eds.) ICCHP 2018. LNCS, vol. 10896, pp. 583–589. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-94277-3_91
4. Austin, D., Sorge, V.: Authoring web-accessible mathematical diagrams. In: Proceedings of the 20th International Web for All Conference, pp. 148–152 (2023)
5. Batusic, M., Miesenberger, K., Stöger, B.: Labradoor: a contribution to making mathematics accessible for the blind (1998)
6. Beezer, R.: PreTeXt: an XML vocabulary for scholarly documents. In: Proceedings of Balisage: The Markup Conference 2018. Balisage Series on Markup Technologies, vol. 21 (2018)
7. Boelkins, M.: Active Prelude to Calculus (2020). http://activecalculus.org

8. Boelkins, M., Austin, D., Schlicker, S.: Active Calculus 2.1. Grand Valley State University Libraries, 2018
9. Egli, C.: Liblouis–a universal solution for braille transcription services. In: Proceedings of Daisy 2009 Conference (2009)
10. Fitzpatrick, D., Godfrey, A.J.R., Sorge, V.: Producing accessible statistics diagrams in R. In: Proceedings of the 14th Web for All Conference. ACM (2017)
11. Gilson, C.L., Dymond, S.K., Chadsey, J.G., Hsu, S.Y.F.: Gaining access to textbooks for postsecondary students with visual impairments. J. Postsecond. Educ. Disabil. **20**(1), 28–39 (2007)
12. Judson, T.: Abstract algebra: theory and applications. Virginia Commonwealth University Mathematics, 2009
13. Maneki, A., Jeans, A.: Latex: What is it and why do we need it? Futur. Reflect. **31**(2) (2012)
14. Murillo-Morales, T., Miesenberger, K., Ruemer, R.: A LaTeX to braille conversion tool for creating accessible schoolbooks in Austria. In: Miesenberger, K., Bühler, C., Penaz, P. (eds.) ICCHP 2016. LNCS, vol. 9758, pp. 397–400. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-41264-1_54
15. Papasalouros, A., Tsolomitis, A.: LaTeX package: Latex2nemeth (2022). https://ctan.org/pkg/latex2nemeth
16. Park, W., Unruh, D.: LaTeX package: braille (2010). https://ctan.org/pkg/braille
17. Polewczak, J.: Latex, mathml, and tex4ht: tools for creating accessible documents. Mathematics, CSUN, Northridge, p. 24 (2005)
18. JooYoung Seo, Sean McCurry, and Accessibility Team: Latex is not easy: creating accessible scientific documents with r markdown. J. Technol. Pers. Disabil. **7**, 157–171 (2019)
19. Sorge, V.: Speech rule engine, v4, 2022. http://speechruleengine.org
20. Tantau, T.: The TikZ and PGF Packages, 2023. https://tikz.dev/
21. XML Schema Documentation for PreTeXt. https://pretextbook.org/doc/schema/. Accessed 10 Apr 2024

Mathematics Accessibility in Primary Education: Enhancing Mathematics Learning Skills and Overcoming Barriers for Visually Impaired Primary Students

Muhammad Shoaib[1](✉), Rosane Minghim[1], Donal Fitzpatrick[2], and Ian Pitt[1]

[1] School of Computer Science and Information Technology, University College Cork, Cork, Ireland
muhammad.shoaib@cs.ucc.ie

[2] Centre for Excellence in Universal Design, National Disability Authority, Dublin, Ireland

Abstract. This paper describes a qualitative study concerning the experiences of blind and visually impaired students learning mathematics in primary schools in Ireland, along with a thematic analysis of the findings. It investigates the instructional challenges and teaching techniques that facilitate an inclusive learning environment through flexible assessment plans. Key focus areas include adapting class collaboration activities, using digital and non-digital materials, and deploying specialized strategies to boost learning skills. This study identified significant barriers, including technology accessibility, classroom mobility, and the need for supportive educational aids such as tactile materials and Braille resources. This paper also discusses the importance of support and assistance, including teacher training, parental involvement, and peer support processes. Based on expert opinions, the study identifies gaps in current solutions and suggests future research and development directions. These include the advancement of User-Centred Design (UCD), the integration of multimodal techniques, and the improvement of collaborative and interactive learning solutions. Furthermore, this study recommends reforming educational policy and comprehensive training programs for educators to ensure that visually impaired students receive an equitable mathematics education. This study lays the groundwork for future advancements in teaching methodologies and assistive technologies, aiming to remove educational barriers and foster academic excellence within blind and visually impaired students.

Keywords: Mathematic · Barriers · Primary · Accessibility · Blind · Visually-Impaired

1 Introduction and Background

Educational inclusion is still a cornerstone of equal schooling. The unique academic requirements of blind and visually-impaired children are often not adequately met during primary education. This study presents a thematic analysis that investigates three main areas: the mathematical learning experiences of blind and visually-impaired students,

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 192–202, 2024.
https://doi.org/10.1007/978-3-031-62846-7_23

the pedagogical approaches of teachers, and expert viewpoints on accessible educational solutions in primary schools.

1.1 Mathematical Learning Experiences of Blind and Visually-Impaired Students

The first area of attention concerns the unique challenges and techniques of mathematics learning for blind and visually impaired students. Previous research includes many studies focusing on specific teaching strategies for these students and identifying several key aspects of mathematics learning. Maćkowski & Brzoza introduced a tutoring platform developed for visually impaired children. The platform utilizes audio-tactile visuals to enhance the learning process for mathematics topics. The efficiency of this platform was examined through quantitative and qualitative research with students from primary and secondary schools. The study attempts to improve learning effectiveness through tailored audio explanations and exercises, catering specifically to the educational needs of visually impaired students [1]. Hayes et al. examine the issues faced by students with visual impairments during standard scientific and mathematics instruction. They stress the significance of classroom adjustments, such as hands-on experiences and extra instructional time, to promote their learning. They also highlighted findings concerning educational policies and practices, i.e., effective adaptation of teaching materials and helping students acquire skills that enhance their independent learning [2]. Steinbach addressed the obstacles students with visual impairments face in obtaining and understanding the mathematics curriculum. It examines several techniques of conveying mathematical information to these children, such as auditory, tactile, and multi-sensory approaches. This study also underlines the need to adjust instructional strategies and resources to suit the needs of visually impaired students, concentrating on boosting their engagement and understanding of mathematical ideas. The study highlights the need for inclusive educational approaches and the development of appropriate multi-sensory solutions to help the mathematical education of visually impaired children [3]. Drijvers & Sinclair reviewed how digital technologies were employed in mathematics instruction. They created a framework encompassing five aims for employing DTs: increasing learning, understanding learning processes, designing for learning, guaranteeing equal access, and influencing curriculum and teaching practices. They address various theoretical foundations, methodologies, and outcomes relevant to these aims, illustrating the evolving nature of research on this topic and the rising relationship between digital technology and mathematical education. The article underlines the significant impact of DTs on the nature of mathematics education and learning [4]. Bobomurod focused on the educational obstacles and methods of teaching mathematics to visually impaired students. This study highlights the need to understand children's attitudes toward mathematics learning from an early age to prevent the development of negative opinions about this subject. Moreover, it examines the content and objectives of the mathematics course, highlighting the need for corrective and developmental conditions, particular techniques, and cognitive activity development. It underlines the significance of adjusting educational methodologies to accommodate the specific demands of visually impaired students, trying for their optimal growth and social adaptation [5].

1.2 Investigation of Teachers' Experiences While Teaching Mathematics to Blind and Visually Impaired Students

The second area of attention is teachers' experiences when educating blind and visually impaired students. The literature shows that teachers are essential in designing inclusive education. They often experience unique pedagogical problems and lack proper preparation. Rushahu conducted a study focusing on the issues and skills of lecturers in Tanzanian higher education institutions to serve visually impaired pupils. A qualitative method was employed with interviews and focus group discussions among 40 participants, including students with visual impairments and professors. The survey demonstrates that most professors lack the abilities and knowledge to identify and address the unique learning demands of these students. It underscores the need for redesigned university curricula, adequate training for professors, and more significant state funding for inclusive education [6]. A thesis by Amanda Esmeralda Mungunda, finished in April 2023, explores the specific obstacles that secondary school mathematics teachers confront when educating learners with vision impairment. Focusing on the Khomas region in Namibia, the study adopts Bruner's idea of constructivism as its theoretical base. Utilizing qualitative research methodologies, it comprises observations and semi-structured interviews with a selected sample of teachers and visually impaired students. The study intends to improve visually impaired students' teaching and learning experiences and offers successful teaching practices, including the need for assistive technologies and instructional materials [7]. Baykaldı explored the experiences of high school mathematics teachers in teaching students with visual impairments. They investigate the challenges faced by the teachers, along with their teaching styles, curriculum preparation, assessment strategies, and ideas regarding inclusive education. The study, conducted through interviews with eight instructors, highlighted a lack of proper training and resources for inclusive education, although they demonstrated 'high teaching efficacy' generally (e.g., with non-visually-impaired students) and had high ideals regarding teaching (e.g., were highly motivated to help visually impaired students. The findings underline the need for increased teacher training and resource development to facilitate inclusive education [8]. Mbulaheni Maguvhe proposed a study focusing on the problems and opportunities in teaching science and mathematics to visually impaired children. The study is based on a case study approach, underlining the critical role of instructors and the necessity for specialized resources and training. It highlights the need to empower instructors to serve visually impaired students in various courses successfully and recognizes the vital role of inclusive education methods. The research gives valuable insights for increasing the learning abilities of visually impaired students in science and mathematics, helping the broader conversation on inclusive education [9].

1.3 Expert Viewpoints on Accessible Educational Solutions in Primary Schools

Thirdly, the study included professionals to obtain suggestions for potential solutions for these children. Integrating expert perspectives is vital in bridging the gap between theoretical knowledge and actual application. Previously, several studies were conducted without the opinion of experts or professionals. IrisMath is a web program developed to aid visually impaired people in executing mathematical operations, notably for engineering students. Inspired by Jupyter Notebooks, this system supports multiple output formats such as LaTex, CMathML, JSON, and audio, assuring accessibility and comprehensibility for visually impaired users. The paper explores the architecture, development, and evaluation of IrisMath, showing its potential as a helpful tool for inclusive education in technical and engineering subjects [10]. Gözde et al. investigate the experiences of visually impairment children in inclusive math courses in Turkey. It focuses on the obstacles these students experience, their opinions on learning methods, and the usefulness of the resources adopted. The study underlines the gap between these children's social and academic demands, underlining a shortage of teacher preparation and reservations towards inclusive education. It emphasizes the need for increased teacher training and creating resources to help visually impaired children in mathematics [11]. Karshmer and Gupta conducted a study on building tools to help visually challenged children learn mathematics. They address the issue of writing and reading complicated mathematical formulas for children. The project at New Mexico State University includes developing logic programming-based systems and interfaces to improve the utilization of mathematical information. This includes building a method for converting Nemeth Code to LaTeX and vice versa, boosting collaboration among blind students and sighted lecturers. The research also discusses employing encoded musical notes and signals of sound to express mathematical concepts, seeking to complete the education cycle with accessible resources for visually impaired children [12].

In this article, we have explored the challenges of blind and visually impaired primary school students. A study was conducted with teachers intended to add depth to this knowledge by examining the lived experiences of teaching in primary schools. Find out the viewpoints of professionals and expert persons. Their thoughts can help researchers and developers build more prosperous, inclusive educational platforms for blind and visually impaired students. This thematic analysis contributes to the ongoing conversation about inclusive education. It strives not only to expose the existing obstacles and opportunities in teaching mathematics to blind and visually impaired primary students but also to provide actionable ideas for improvement, pulling from both academic and practical domains. The remaining article is organized as follows: Sect. 2 explains the methodology; Sect. 3 illustrates the results in the form of extracted themes; Sect. 4 concludes this article and provides recommendations for future work.

2 Methodology

The research employed a qualitative study design, employing semi-structured interviews for data collection. Thematic analysis was used to analyze the gathered data.

2.1 Participants

This study involved nineteen participants divided into three categories: students, teachers, and domain experts. Firstly, this study included five visually-impaired participants from various primary schools, two males and three females. These individuals were recruited from different primary schools. Their ages ranged from 9 to 12 years, with a mean age of 10.80 (Std. Deviation = 1.3 and Variance = 1.7). Student participants in the study met the criteria for legal blindness (visual acuity 20/400 or less). Among the participants, three were completely blind (P1, P2, and P4), and two participants (P3 and P5) had low vision. Most students identified as Republic of Ireland nationals (4), and EU national (1), Secondly, nine primary school teachers (P6-P14) took part in this study. They have diverse roles within the primary school context, including class teachers (1), Support teachers (2), and Special Educational Needs teachers (6), boasting experience ranging from 3 to 22 years. The age range of participants spanned from 26 to 54 years, with a notable gender difference (7 females, 2 males). The majority of participants identified as Republic of Ireland nationals (4), EU National (2), Asian (2), or United Kingdom (1), aligning with broader demographic trends in the teaching profession. Thirdly, five focus group persons (P15-P19) from several organizations, i.e., National Disability Authority (1), National Council for Special Education (1), Saint John of God (1), and VisitEDUfinn Ltd (2), participated in this study. The age range of participants spanned from 40 to 58 years, with a notable gender difference (3 females, 2 males). Their professional experience ranges from 5 to 20 years. The majority of participants identified as Republic of Ireland nationals (2) and EU National (2), United Kingdom (1), aligning with broader demographic trends in their professional careers. Overall, teachers and focus group participants ages ranged from 26 to 58 years, with a mean age of 43.64 (Std. Deviation = 9.44 and Variance = 89.02). Their experience ranges from 3 to 22 years, with a mean experience of 13 (Std. Deviation = 6.28 and Variance = 39.39). For additional demographic details, refer to Tables 1 and 2. Table 1 provides participant details, including participant ID, age, gender, nationality, a description of vision impairment, blind or low vision, and Educational arrangements. Table 2 provides participant details, including participant ID, age, gender, nationality, and experience in number of years.

2.2 Data Analysis

One technique for finding, examining, and summarizing patterns in qualitative data is thematic analysis. Using this methodology, data can be transformed into rich, in-depth knowledge through text analysis. Thematic analysis is a tool for identifying and explaining facts and capturing crucial details related to the data and research question. Thematic analysis is an iterative and distinctive approach to text evaluation that offers a specific understanding and interpretation of the study issue, according to researchers [14]. Thus, in the current study, the concerns of visually impaired students and teachers were identified and extracted using thematic analysis.

The data was examined using the inductive thematic analysis procedure outlined by Braun and Clarke [15]. Thematic analysis was chosen for its adaptability and capacity to yield insightful, interpretative comments [16]. This approach is not theory-driven but instead data-driven, enabling the researcher to uncover unforeseen insights. The analysis adhered to Braun and Clarke's six-step method. Firstly, the researchers read and familiarized themselves with the data. Secondly, codes were manually generated on the transcripts' margins based on the entire dataset. The same unit of text could be encompassed in more than one code, and relevant data for each code were compiled. Thirdly, the coded data were systematically reviewed, cross-checked, and organized

Table 1. Detailed information of student participants

Participants ID	Age	Gender	Nationality	Description of vision-impaired	Educational Arrangments
P1	12	Female	EU Country	Blind	Special education class
P2	10	Male	EU Country	Blind	General education along with special education
P3	11	Female	EU Country	Low Vision	General education along with special education
P4	9	Male	Republic of Ireland	Blind	General education along with special education
P5	12	Female	Republic of Ireland	Vow Vision	General education along with special education

Table 2. Detailed information of Teachers and Focus group participants

Participants ID	Age	Gender	Nationality	Experience in No. of Years
P6	52	Female	Asian	20
P7	42	Female	EU Country	19
P8	26	Female	Asian	3
P9	29	Female	EU Country	18
P10	45	Male	United Kingdom	10
P11	37	Male	Republic of Ireland	14
P12	43	Female	Republic of Ireland	22
P13	54	Female	Republic of Ireland	7
P14	44	Female	Republic of Ireland	10
P15	43	Female	EU Country	13
P16	58	Male	EU Country	15
P17	40	Male	United Kingdom	5
P18	57	Female	Republic of Ireland	20
P19	41	Female	Republic of Ireland	6

into themes. Fourthly, themes were cross-checked against the coded data and the entire dataset and concluded by creating a thematic ‘map’. Fifthly, themes were defined and labeled. Finally, the themes underwent in-depth analysis, with selected data extracts examined in relation to the research question, existing literature, and the subsequent report.

3 Results

3.1 A Thematic Analysis of Mathematics Learning of Blind and Visually Impaired Students in Primary Schools Perceptions of Learning Challenges, Strategies, Support, Resources

After examining the five transcripts, five prominent themes concerning students' viewpoints were found. Figure 1 presents a comprehensive visual representation of the factors affecting the learning process of visually impaired students in mathematics. The diagram also represents a thematic analysis related to the mathematics learning of blind and visually impaired students in primary schools. It is structured around a central theme, "Investigating Mathematics learning of Blind and visually impaired students in primary schools", with related factors branching out into specific categories such as: "Mathematics Topics", "Class Collaboration Activities", "Challenges, Obstacles, and Issues", "Techniques and Methods to Fulfill Learning Needs", and "Assistance and Support".

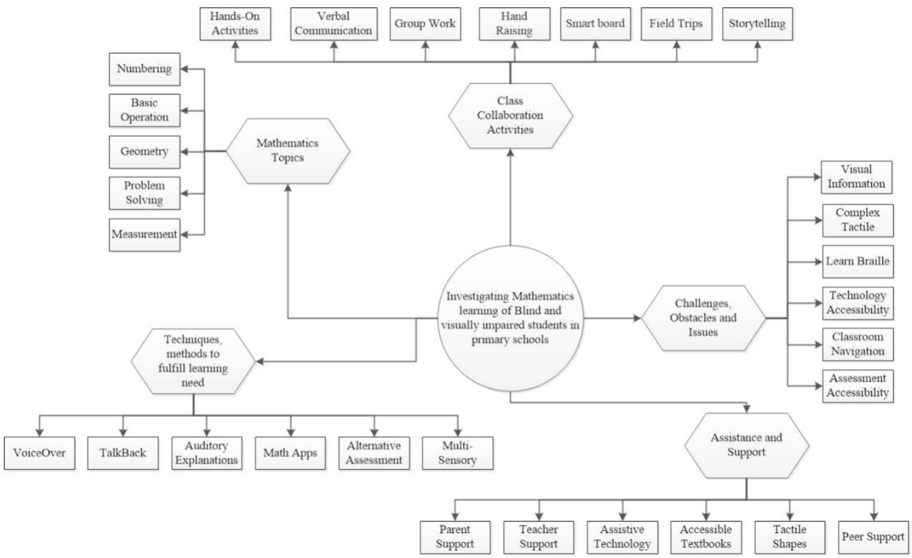


Fig. 1. A thematic analysis of Mathematics learning of Blind and visually impaired students

3.2 Investigation of Teacher Experiences While Teaching Mathematics to Blind and Visually Impaired Students in Primary Schools

An analysis of nine transcripts revealed four leading themes concerning teachers' viewpoints. Figure 2 presents a comprehensive analysis of the various factors affecting the teaching process and the potential strategies to overcome challenges, facilitating a better learning experience for visually impaired students in primary schools. This diagram also represents the thematic analysis of teacher experiences while teaching mathematics to blind and visually impaired students in primary schools. Here's an explanation of its structure: The central theme of the diagram is the "Investigation of teacher experiences while teaching mathematics to Blind and visually impaired students in primary schools." Branching out from this central theme are four main categories: "Barriers, obstacles, and Challenges faced by teachers", "Digital resources", "Non-digital resources", and "Ways to make learning mathematics easy".

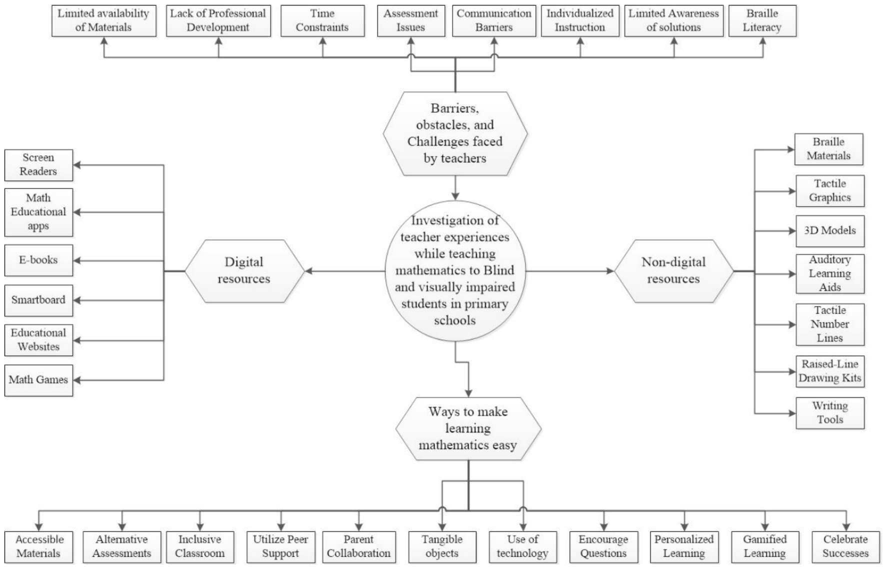


Fig. 2. A thematic analysis of teachers' experiences

3.3 Opinions of Expert Persons to Introduce Accessible Solutions for Blind and Visually Impaired Students

After the exploration of five transcripts disclosed four leading themes concerning focus group persons' viewpoints, Fig. 3 serves as a comprehensive representation of expert perspectives on creating a more inclusive and accessible learning environment for blind and visually impaired students, particularly in the context of mathematics education. It suggests a multi-faceted approach that includes both the refinement of existing solutions and the development of new strategies and tools. This diagram also represents a thematic analysis of the opinions of experts on introducing accessible solutions for blind and visually impaired students. It organizes these opinions into several categories, named: "Valuable strategies for learning mathematics", "Missing features from current solutions", "Creative approaches for designing mathematics solutions", and "Guidelines/methods for developers".

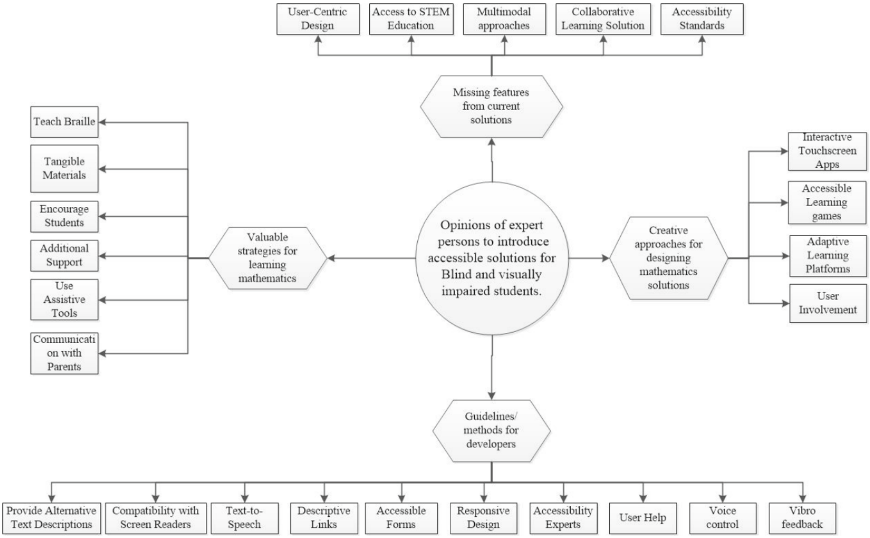


Fig. 3. A thematic analysis of professional experts to introduce accessible solutions

4 Conclusion and Future Work

The thematic analysis provided in this study emphasizes the challenging task of offering an accessible educational setting for blind and visually impaired primary school learners, especially in mathematics. Our research has identified several significant areas of attention, such as the material available for mathematics education, assistance and support for blind and visually impaired students, mathematics topics that require attention, class collaboration activities that are useful for learning, pedagogical methodologies, challenges/obstacles and issues encountered by students and teachers, available essential technologies and approaches to fulfilling learning need, ways to make learning mathematics easier, guidelines/methods for developers, more effective strategies for learning mathematics, missing features from current solutions, creative approaches for designing mathematics solutions and an overall support system encompassing both digital and non-digital resources. Through comprehensive investigation, it is evident that while there are multiple challenges (ranging from the availability of resources and technology integration to the necessity for ongoing professional development of teachers), there are also many approaches that might reduce these challenges and obstacles. Effective strategies such as using tactile resources, Braille understanding, and using assistive technology have emerged as solid enablers in learning.

Moreover, this study has underlined the importance of multimodel approaches in educational tools and the demand to develop UCD-based accessible solutions. The insights from expert perspectives have supported the notion that a team effort is essential to build inventive and flexible math solutions that are interactive and enjoyable for blind and visually impaired students. Due to the ongoing progress of technology, the concept of assistive technology is evolving over time. Because of this, a key component in improving mathematics education for visually impaired pupils is determining how best to adopt and integrate assistive technology in schools in order to provide access to the curriculum. Considering our analysis, we recommend the implementation of extensive training initiatives for teachers that provide them with the knowledge and abilities essential to effectively implement assistive technology-based solutions in classrooms. For the training purpose, attention to policy reform is necessary to ensure that the current frameworks guarantee the training of teachers for inclusive learning experiences to better support blind and visually impaired students. The collaborative activities inside classrooms, together with personalized instruction and assistance of parents and classmates, comprise a comprehensive strategy that may significantly improve the educational experience of visually impaired students. Implementing these findings needs an intense effort that includes continual development and evaluation of methods of instruction, resources, and assistive technologies.

This research has created a platform for further investigation into the improvement of mathematics teaching for blind and visually impaired children. Additional studies should attempt to verify the impact of the suggested strategies on student learning skills. We intend to expand on these findings by introducing an inclusive educational platform with innovative digital learning features. The primary objective of this platform is obvious: minimize challenges and promote a learning atmosphere in which blind and visually impaired students can succeed in mathematics education. This can ensure that these learners are enabled and fully integrated into the structure of educational excellence. In this educational platform, researchers and developers should use touch gestures, audio feedback, and vibro feedback to assist blind and visually impaired students. The development of interactive learning games could allow blind and visually impaired students to learn more easily, just as the use of text-to-speech functionality could allow users to hear mathematical content and resources. Developers should provide customizable content with accessible features, catering to individual needs.

Funding. This publication has emanated from research conducted with the financial support of Science Foundation Ireland under Grant number 18/CRT/6222. For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

References

1. Maćkowski, M., Brzoza, P.: Accessible tutoring platform using audio-tactile graphics adapted for visually impaired people. Sensors **22**(22), 8753 (2022)
2. Hayes, C., Proulx, M.J.: Turning a blind eye? Removing barriers to science and mathematics education for students with visual impairments. Br. J. Vis. Impairment **42**, 544–556 (2022)
3. Steinbach, S.: Looking beyond vision: supports for students who are blind or visually impaired in mathematics. NWEA (2022)
4. Drijvers, P., Sinclair, N.: The role of digital technologies in mathematics education: purposes and perspectives. ZDM Math. Educ. 1–10 (2023). https://doi.org/10.1007/s11858-023-01535-x
5. Bobomurod, S.R.S.A.: Teaching mathematics to blind and visually impaired teenagers. Int. J. Adv. Res. Educ. Technol. Manage. **2**(2) (2023)
6. Rushahu, B.G.: Lecturers' ability to support inclusive learning to students with visual impairment in higher learning institutions in Tanzania. Tanzania J. Popul. Stud. Dev. **30**(1), 41–55 (2023)

7. Mungunda, A.E.: Challenges faced by mathematics teachers in teaching secondary school learners with visual impairment in the Khomas region. Doctoral dissertation, University of Namibia (2023)
8. Baykaldı, G., Corlu, M.S., Yabaş, D.: An investigation into high school mathematics teachers and inclusive education for students with visual impairments. Br. J. Vis. Impairment **42**, 124–213 (2023)
9. Maguvhe, M.: Teaching science and mathematics to students with visual impairments: reflections of a visually impaired technician. Afr. J. Disabil. **4**(1), 1–6 (2015)
10. Zambrano, A., Pilacuan, D., Salvador, M.N., Grijalva, F., Garzón, N.O., Mora, H.C.: IrisMath: a blind-friendly web-based computer algebra system. IEEE Access **11**, 71766–71776 (2023). https://doi.org/10.1109/ACCESS.2023.3281761
11. Bayram, G.İ, Corlu, M.S., Aydın, E., Ortaçtepe, D., Alapala, B.: An exploratory study of visually impaired students' perceptions of inclusive mathematics education. Br. J. Vis. Impair. **33**(3), 212–219 (2015)
12. Karshmer, A.I., Pontelli, E., Gupta, G.: Helping visually impaired students in the study of mathematics. In: FIE'99 Frontiers in Education. 29th Annual Frontiers in Education Conference. Designing the Future of Science and Engineering Education. Conference Proceedings (IEEE Cat. No. 99CH37011), vol. 2, pp. 12C4–5. IEEE (1999)
13. Bogdan, R.C., Biklen, S.K.: Qualitative Research for Education: An Introduction to Theory and Methods, 3rd edn. Allyn and Bacon, Boston (1998)
14. Brooks, J., Horrocks, C., King, N.: Interviews in qualitative research 1–360 (2018)
15. Braun, V., Clarke, V.: Using thematic analysis in psychology. Qual. Res. Psychol. **3**, 77–101 (2006)
16. Boyatzis, R.E.: Transforming Qualitative Information: Thematic Analysis and Code Development. Sage (1998)

Using a Chatbot Integrated with Video Prompting to Enhance Mathematical Word Problem-Solving Skills in Primary School Students with Intellectual Disabilities

Chu-Lung Wu(✉)

National Taichung University of Education, Taichung, Taiwan
wulongd@mail.ntcu.edu.tw

Abstract. The purpose of this study was to investigate the effectiveness of using a chatbot combined with video prompting instructions in improving the mathematical word problem-solving abilities of primary school students with intellectual disabilities. This study adopted a single-subject design with multiple probes across participants. The independent variable was the intervention mode of using a chatbot combined with video prompting instructions, while the dependent variable was the performance of the participants in solving mathematical word problems. The participants of this study were three primary school students with mild intellectual disabilities in grades 4–5. During the intervention period, the teaching activities using the chatbot combined with video prompting instructions were conducted. The intervention period consisted of a total of 12 sessions, with each teaching session lasting for 30 min. In the data analysis, a combination of visual analysis and the Tau-U test was used to analyze the results. Additionally, qualitative data from interviews with the participants and teachers were collected and utilized. The findings revealed a significant improvement in the participants' accuracy of problem-solving after the instructional intervention. The Tau-U values exhibited a large effect from the baseline period to the treatment period, indicating immediate learning effects. Furthermore, all three participants and their resource class teacher expressed a positive attitude towards the effectiveness of the instructional approach. This study provides specific recommendations and examples for future special education teachers on how to apply the integration of chatbots and video prompting strategies in resource classroom teaching. These suggestions can serve as a reference for developing relevant teaching materials, lesson plans, and conducting instruction.

Keywords: Chatbot · Mathematical word problem-solving abilities · Primary school students with intellectual disabilities · Video prompting strategy

1 Introduction

The ability to solve mathematical word problems was a crucial means for students to tackle and resolve unforeseen issues in daily life [1]. Students with intellectual disabilities often encountered challenges when confronted with mathematical word problems, as

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 203–210, 2024.
https://doi.org/10.1007/978-3-031-62846-7_24

their problem-solving abilities were often insufficient [7]. Despite possessing adequate reading and mathematical computation skills, they still struggled with problem-solving in applied scenarios. In recent years, there had been a growing recognition of the importance of using technology to assist special education students in their learning process. Technological interventions had gradually gained attention, as they had the potential to make learning more enjoyable and effective for students.

Video modeling (VM), which provided positive demonstrations, aligned with the principles of social learning theory, suggesting that learning occurred through observation and imitation. Students could acquire targeted skills by observing and imitating the content presented in videos [10]. Video prompting (VP) instruction, a variation of VM, differed in that it broke the video into segments, playing one small portion at a time, reducing students' cognitive load and enhancing their learning outcomes [4]. VM and VP benefited from the development of handheld device technology in recent years. The research in this area increased, and these instructional approaches demonstrated positive teaching outcomes, thus becoming evidence-based teaching strategies [5, 12]. However, in the instructional process using VM or VP, students were often constrained by issues related to video playback, limiting the potential for self-directed learning. Therefore, selecting a suitable video playback interface was an important consideration [11].

In recent years, there was an increasing focus on exploring the impact of chatbots in education [3]. Both retrospective articles from 2021 indicated that research on chatbots in the field of education was being given significant attention, with a noticeable rise in the number of published articles in the past three years [6, 9]. The majority of research on chatbots in education primarily focused on language learning and communication training because these studies were most directly related to the functionalities of chatbots [6]. For example, studies utilized chatbots to assist teachers in delivering instruction in multimedia courses, and the results revealed that chatbots could enhance learners' performance [8]. In the domain of special education, research on chatbots was still in its early stages. A literature review from 2020 on chatbot assessments for students with special needs pointed out that out of 192 relevant articles on chatbot assessments, only 15 were related to individuals with disabilities or special needs [2]. Currently, there is a lack of systematic approaches and sufficient research to evaluate the impact of chatbots on students with disabilities.

Therefore, the purpose of this study was to explore the effectiveness of using a chatbot combined with video prompting (VP) instructions in improving the mathematical word problem-solving abilities of primary school students with mild intellectual disabilities.

2 Methodology

2.1 Participants

The participants consisted of 3 elementary school students with mild intellectual disabilities in grades 4–5. Table 1 profiles the participants. They were all placed in resource classes to receive special education services and received full separate instruction for Mandarin Chinese and Mathematics in the resource class.

Table 1. Participants' information and characteristics.

Pseudonym	Andy	Bella	Charlie
Grade (age)	Fourth grade (10 years old)	Fourth grade (10 years old)	Fifth grade (11 years old)
Gender	Boy	Girl	Boy
Disability	Intellectual disability (mild)	Intellectual disability (mild)	Intellectual disability (mild)
IQ test	WISC-V: 66	WISC-IV: 70	WISC-IV: 69

Note. WISC-V = Wechsler Intelligence Scale for Children-5th Edition; WISC-IV = Wechsler Intelligence Scale for Children-4th Edition

2.2 Procedure

This study employed a single-subject design known as the multiple-probe across participants design. The independent variable was the intervention mode of combining a chatbot with VP, while the dependent variable was the performance of the participants in solving mathematical word problems, specifically measured by the accuracy of their answers on the assessment sheets. During the baseline period of the study, no instructional strategies were implemented, and data collection for the baseline was conducted using math word problem assessments developed by the researcher. The intervention period involved engaging in instructional activities utilizing the chatbot combined with visual prompts. The intervention was implemented a total of 12 times, with each session lasting for 30 min. Two weeks after the final instructional activity, the maintenance phase assessment was conducted using the same evaluation methods as the baseline and intervention periods.

2.3 Materials

The teaching materials were presented through a LINE chatbot interface, which displayed VP for each step of solving the word problems. The teaching materials using the LINE chatbot interface are illustrated in Fig. 1. VP for solving the word problems included 6 steps: reading the problem, highlighting key information, drawing a diagram, writing an equation, solving the problem, and checking the solution. Participants could simply click on the corresponding icon buttons without having to input text, and the corresponding video demonstrating each step were played. The instructional process focused on allowing the participants to interact with the chatbot independently shown as Fig. 2. The teacher provided least to most prompts and clarifications when the students provided incorrect answers or did not respond for more than 10 s.

Fig. 1. The LINE chatbot interface displays six steps for problem-solving in the application.

2.4 Interobserver Agreement

We randomly selected three test results from each participant to compare inter-rater reliability, and the comparison results showed 100% agreement.

2.5 Data Analysis

In the data analysis section, visual analysis and the Tau-U test were employed for data analysis. From visual analysis, we could observe the performance trends of the participants across different phases. Tau-U test, through testing the effect size, measures the effect of the independent variable on the dependent variable.

Fig. 2. The participants were able to independently operate the LINE chatbot.

3 Results

The single-subject graphs for the three participants are shown in Fig. 3.

For Andy, the average correct answer rate during the baseline period was 40.3%. After the instructional intervention, the average correct answer rate increased to 60.2%. There was a noticeable upward trend during the intervention period, followed by a stable performance. There was a significant improvement in learning outcomes from the baseline to the intervention period, with a Tau-U value of 1 (CI90% [0.435, 1], $p = 0.2963$). During the maintenance phase, the average correct answer rate was 81.5%, indicating maintenance of learning effects. The Tau-U value from the intervention to the maintenance period was 0.861, showing a moderate effect (CI90% [0.228, 1], $p = 0.0253$).

Bella showed a significant improvement in learning outcomes from the baseline to the intervention period, with their average correct answer rate increasing from 2.8% to 38.4%. There was an unstable upward trend observed during the intervention period. This improvement was indicated by a Tau-U value of 1 (CI90% [0.367, 1], $p = 0.0094$). During the maintenance phase, although the average correct answer rate decreased to 22.2%, it remained higher than the baseline average, demonstrating the maintenance of learning effects. The Tau-U value from the intervention to the maintenance period was −0.67, indicating a moderate effect (CI90% [−1, −0.034], $p = 0.0833$).

Charlie exhibited a significant improvement in learning outcomes from the baseline to the intervention period, with their average correct answer rate increasing from 19.4% to 60.2%. During the intervention period, there was an unstable upward trend observed. This improvement was supported by a Tau-U value of 1 (CI90% [0.435, 1], $p = 0.0036$). During the maintenance phase, although the average correct answer rate decreased to 37.0%, it remained higher than the baseline average, indicating the maintenance of

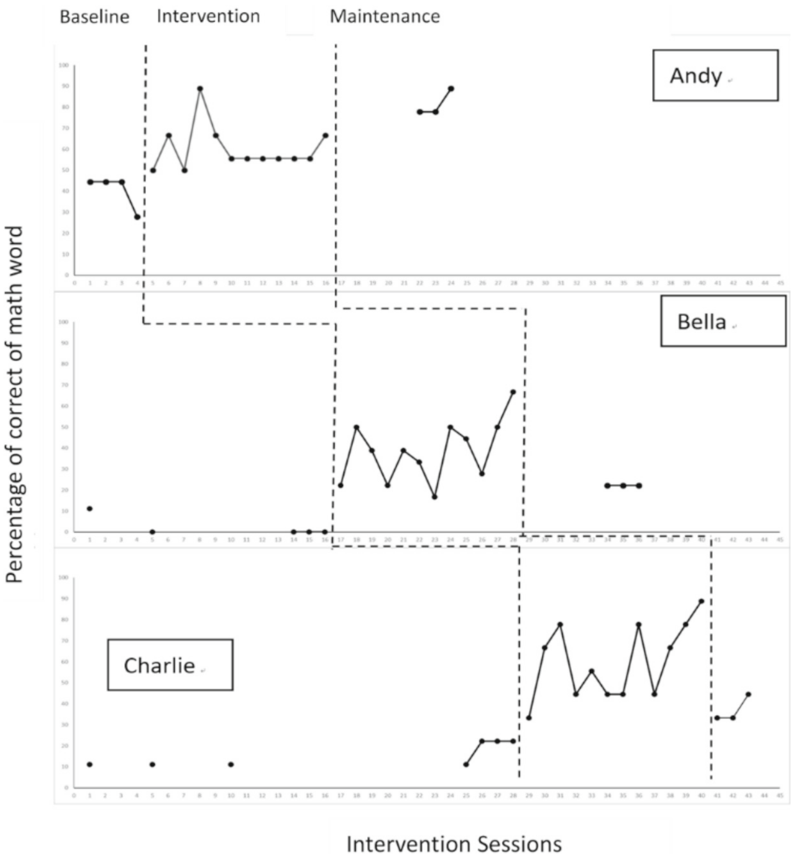


Fig. 3. Accuracy rates of the three participants in answering mathematical word problems at different stages.

learning effects. The Tau-U value from the intervention to the maintenance period was −0.778, suggesting a moderate effect (CI90% [−1, −0.145], p = 0.0433).

4 Discussion

The three participants showed a high level of acceptance towards the combination of LINE chatbot and VP in their instruction. They all enjoyed the learning approach of watching videos through the LINE chatbot interface, and they felt more focused during the learning process. This finding is similar to the results of previous research that utilized VM combined with augmented reality (AR). By reducing the cognitive load associated with operating interfaces, students can concentrate more on the learning content presented in VM [11].

In this study, we utilized the LINE chatbot as the interface for the participants to engage in learning through repeated viewing of VP. During the learning process, students followed the steps demonstrated in the videos and performed the operations accordingly.

If they encountered difficulties or did not understand certain parts, they had the option to replay the videos until they completed all the steps. In these situations, the teacher's approach was to guide the students to rewatch the videos and highlight the key points for them to focus on, rather than directly intervening in the teaching process. The results indicated that although all three students showed immediate improvement, only one of them maintained the effectiveness over time. This could be attributed to the fact that the process of solving application problems requires more than just following structured problem-solving steps; it also involves timely discussions and clarifications to help students grasp the underlying meaning and internalize the problem-solving approach. The chatbot used in this study was limited to providing responses to closed-ended questions and did not possess the real-time feedback and problem-solving capabilities of an Artificial Intelligence Generated Content (AIGC). Future research that integrates AIGC is expected to enhance students' learning outcomes even further.

To conclude, based on the observations during the research process and the experimental results, it can be concluded that using a combination of a LINE chatbot and VP provides a conducive environment for student-driven learning and enhances their ability to engage in independent learning. This study offers concrete suggestions and examples for integrating chatbot technology with video prompts in resource class teaching, serving as a valuable reference for future special education teachers when developing instructional materials and lesson plans.

Acknowledgments. This work was supported by the National Science Council of the Republic of China (Taiwan) for the funding support (grant numbers MOST 112–2410-H-142–007).

Disclosure of Interests. The author has no competing interests to declare that are relevant to the content of this article.

References

1. Bottge, B., Hasselbring, T.: Comparison of two approaches for teaching complex, authentic mathematics problems to adolescents in remedial math classes. Except. Child. **59**, 556–566 (1993)
2. Federici, S., et al.: Inside pandora's box: a systematic review of the assessment of the perceived quality of chatbots for people with disabilities or special needs. Disabil. Rehabil. Assist. Technol. **15**, 832–837 (2020)
3. Ferrell, O.C., Ferrell, L.: Technology challenges and opportunities facing marketing education. Mark. Educ. Rev. **30**(1), 3–14 (2020). https://doi.org/10.1080/10528008.2020.1718510
4. Gardner, S., Wolfe, P.: Use of video modeling and video prompting interventions for teaching daily living skills to individuals with autism spectrum disorders: a review. Res. Pract. Persons Severe Disabil. **38**(2), 73–87 (2013)
5. Hughes, E.M.: Point of view video modeling to teach simplifying fractions to middle school students with mathematical learning disabilities. Learn. Disabil. **17**(1), 41–57 (2019)
6. Hwang, G.-J., Chang, C.-Y.: A review of opportunities and challenges of chatbots in education. Interact. Learn. Environ. **31**, 4099–4112 (2021). https://doi.org/10.1080/10494820.2021.1952615

7. Montague, M., Bos, C.: Cognitive and metacognitive characteristics of eighth-grade students' mathematical problem solving. Learn. Individ. Differ. **2**, 371–388 (1990)
8. Tegos, S., Demetriadis, S., Karakostas, A.: Promoting academically productive talk with conversational agent interventions in collaborative learning settings. Comput. Educ. **87**, 309–325 (2015). https://doi.org/10.1016/j.compedu.2015.07.014
9. Wollny, S., Schneider, J., Di Mitri, D., Weidlich, J., Rittberger, M., Drachsler, H.: Are we there yet? - a systematic literature review on chatbots in education. Front. Artif. Intell. **4**, 654924 (2021). https://doi.org/10.3389/frai.2021.654924
10. Wright, J.C., Knight, V.F., Barton, E.E.: A review of video modeling to teach STEM to students with autism and intellectual disability. Res. Autism Spectr. Disord. **70**, 101476 (2020). https://doi.org/10.1016/j.rasd.2019.101476
11. Wu, C.L.: Using video modeling with augmented reality to teach students with developmental disabilities to solve mathematical word problems. J. Dev. Phys. Disabil. **35**(3), 487–507 (2023). https://doi.org/10.1007/s10882-022-09862-9
12. Yakubova, G., Defayette, M.A., Chen, B.B., Proulx, A.L.: The use of augmented reality interventions to provide academic instruction for children with autism, intellectual, and developmental disabilities: an evidence-based systematic review. Rev. J. Autism Dev. Disord. (2021). https://doi.org/10.1007/s40489-021-00287-2

Audible Charts of Mathematical Functions

Krzysztof Dobosz(✉) and Dawid Hanak

Department of Algorithmics and Software, Silesian University of Technology, Akademicka 16, Gliwice, Poland

Krzysztof.Dobosz@polsl.pl, dawihan807@student.polsl.pl

Abstract. This article describes the implementation of a prototype educational mobile application for visually disabled students, prepared in the form of a research tool. This tool allows students to explore the features of polynomial functions by audibly analyzing their graphs. The core of the tool is to explore the graph of the drawn function. It is implemented using audio feedback that reflects the position of the indicator relative to the function graph line and the value of the nearest point on the graph. This type of sonification is performed through sound waves synthesized in real time with an amplitude that is a function of the distance of the indicator from the graph line. The research goal is to verify which of the feature parameters of the audio graph are most suitable for graph exploration. The following features of the graph sonifying oscillator can be analyzed: type (sinusoidal, square, triangular, sawtoothed), base tone, sound sampling frequency, buffer size, amplitude decay threshold. The above-mentioned parameters can be properly configured, adapting the sound of the charts to the features of the mobile device's touch screen and the student's hearing capabilities.

Keywords: Sonification · Chart · Education · Mobile Application

1 Introduction

Accessibility of mobile applications for people with visual disabilities can be provided by mechanisms built into the devices' operating systems. These include a whole set of accessibility features, such as a screen reader, audio confirmation of entered characters, and high screen contrast settings for people who are only partially blind. It can be easily seen that the main way of informing the user about what is happening in the application is sound. There are different degrees of visual impairment, but Text-To-Speech (TTS) is an integral part of any mobile app accessibility solution for people with visual disabilities. However, TTS is problematic when exploring an area that is not a set of discrete objects. Such a case is the linear graph of a mathematical function. Our research goal is to check what way of auditory representation of a mathematical function is best understood by students with visual disabilities.

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 211–216, 2024.
https://doi.org/10.1007/978-3-031-62846-7_25

2 State of Art

Auditory rendering of mathematical function graphs in mobile applications is a burgeoning field aimed at enhancing accessibility and user experience. These applications utilize sonification techniques to convert mathematical data into sound representations, enabling users to perceive and interact with graphs through auditory feedback.

Ibanez and Mateos [1] provide an extensive review of sonification methods tailored for mobile platforms, emphasizing their potential to facilitate understanding of complex mathematical concepts. Similarly, Gnesdilow and Kobylkina [2] explore specific approaches for representing mathematical functions in mobile apps, with a focus on usability for individuals with visual impairments. Additionally, Sánchez, Gómez, and Pérez [3] investigate practical implementations of audio representation in mobile applications, highlighting its effectiveness in making mathematical data accessible to visually impaired users. In the article entitled "Mobile Audio Games Accessibility Evaluation for Users Who Are Blind" [4], Araújo et al. provide examples of the use of multi-channel audio to achieve a three-dimensional effect and indicate the position of important gameplay elements and specific sound signals for individual actions to easily distinguish occurring situations. A completely different approach to creating a game accessible to people with visual disabilities was presented by Leporini and others in "A Mobile Educational Game Accessible to All, Including Screen Reading Users on a Touch-Screen Device" [5]. The mobile educational game presented in the article had an appropriate graphic and sound background referring to the real world, which was supposed to make the subsequent stages of the game more diverse and interesting. When discussing voice interfaces, we must also mention the Speech-To-Text (STT) technology, which has developed very dynamically over the last few years, but is still not without problems. Users using this mechanism often encounter frustration, which comes down to trying to cope with the problem by changing the articulation, tone or volume of the voice data entered. These types of issues were discussed in article titled "Patterns for How Users Overcome Obstacles in Voice User Interfaces" published by Mayers et al. [6].

Advancements in mobile technology have facilitated the development of sophisticated auditory graph drawing apps with high-quality sound rendering and responsive user interfaces. Despite challenges such as ensuring accuracy and clarity of auditory representations, ongoing research and development efforts aim to address these issues and further improve the usability and effectiveness of auditory graph drawing applications. During our review, we highlighted some existing solutions. Certainly the most interesting and adequate solution in the field under consideration is the *Desmos* graphing calculator. It is available as a mobile and web application available at desmos.com. Its accessibility features were described by Brauner [7]. A more stripped-down solution in the field is *Wolfram Alpha*, available as a mobile or web application at wolframalpha.com [8]. It does not have the same degree of interactivity in terms of sonification and reading of chart properties as *Desmos*, but the creators made an effort to specify sonically characteristic places and shapes of the charts in question. As for the solutions already discussed, *GeoGebra*, available as a mobile and web application at [9], is in the middle in terms of the provided functionalities. Functionally, it is almost the same as the database provided by the *Desmos* graphing calculator, but it does not have the ability to sonify charts and does not use the TTS mechanism, but it also allows for drawing three-dimensional charts and solving geometric problems. In the context of graphing calculators, the only documented and used means ever attempted to sonify function graphs was TTS and replaying a short, previously generated signal. Sample experiments are documented in the article "Math Graphs for the Visually Impaired: Audio Presentation of Elements of Mathematical Graphs" by Jeongyeon et al. [10].

3 Methodology

We started our research work with a review of the state of the field, taking into account existing solutions in the form of mobile and web applications. The next step was to gather information about math teachers' expectations for a fully accessible mobile math app that demonstrates audible function graphs. They were used to define functional requirements. Based on them, a prototype was implemented.

All features of the sound representing the mathematical function and the graph area are made available in the form of configurable parameters in the application configuration. Thanks to this, we obtained a useful research tool. The tool uses the Text-To-Speech engine to emit voice information about special points on the chart (e.g. minimum, maximum, zero places). In addition, the Speech-To-Text mechanism was used to issue voice commands so that the student could operate the application without the teacher's help. Then, it was planned to define coefficients describing the degree of sound in the graph of a mathematical function.

The last step will be to verify the usability of the mobile application and experimentally determine threshold values for selected sound parameters of the graphs.

4 Implementation of Audible Mobile Application

The following functionalities of the mobile application have been implemented:

- drawing graphs of polynomial functions with integer coefficients and degrees
- navigation and providing function patterns and visibility intervals via voice using STT mechanism, using a predefined set of commands. Some of them require the user to enter parameter values sequentially
- "discovering" function graphs using sound by exploring the screen with a pointer - sonification by using the TTS mechanism to read the position of the indicated point and the digital oscillator
- configuration of every aspect of the application:

  sound options: type of oscillator sonifying the graphs, base oscillator tone, oscillator sound sampling frequency, buffer size for the oscillator, amplitude extinction threshold, length of TTS intervals between phrases
  chart options: density of points on the graph, background, axis, line and label colors, thickness of the graph lines, selection of automatic or manual range of values on the y axis, graph grid density.
  graph exploration options: tolerance of the distance of the indicator from the graph line for sonication with an oscillator, tolerance of the distance of the pointer from the graph line for reading properties point via TTS.

When analyzing possible use cases, it was assumed that the user is an assisted student, i.e. one who can use the application fluently, using accessibility settings or teacher assistance. In the example scenario, the user wants to draw the graph of the function $f(x) = x^2 - 2x + 2$. To do this, he first enters the command in the form of the sentence "Draw the function." The key word "function" is recognized and the user is sequentially asked to enter the formula for a polynomial function in general form. Says the phrase "x squared minus two x plus two." The graph of the function shown in Fig. 1 is drawn, and TTS repeats the equation entered by the user. The visibility range remains the default for now - from x −10 to 10, y −100 to 100.

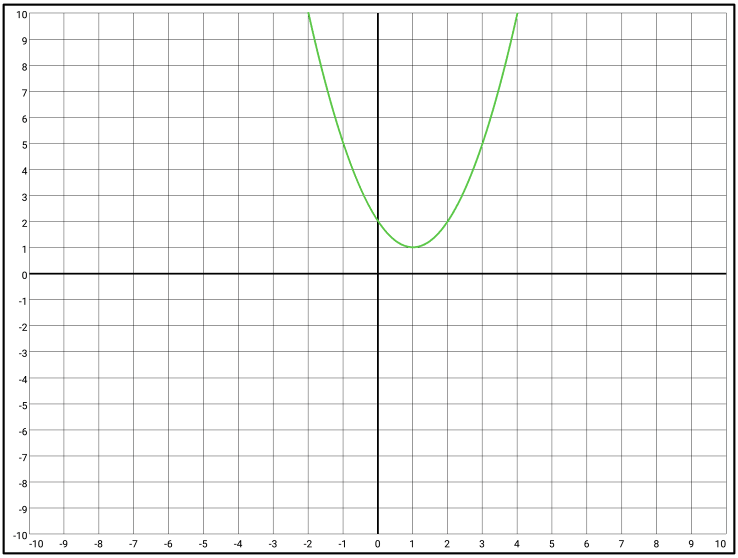


Fig. 1. Graph of the function $f(x) = x^2 - 2x + 2$.

Next, the user wants to change the visibility range of the function so that only positive y values from 0 to 10 in the range x 0 to 4 are visible. He enters the command in the form of the sentence "Select range." The key word "range" is recognized and the user is asked to enter four arguments, respectively, representing the minimum and maximum values of the x and y ranges. The user says "zero", "four", "zero", "ten". After each number spoken, TTS comments on the entered value and the chart is redrawn, taking into account the new range. This allows it to easily deselect a range by entering a return command. If the user wants to recognize the location of the chart axes, he or she enters the command in the form of the sentence "Describe the axes." The key word "axes" is recognized and the previously set range is read to the user.

To accurately recognize the shape and location of the chart, the user uses the sonification function of chart lines and axes. To do this as quickly and conveniently as possible, he decides to turn on the mode of reading chart points. Enters a command in the form of a sentence "Enable points mode." The keyword "point" is recognized and from now on the user is able to receive voice information about the location of each indicated point.

If the user has drawn a function for which he or she is curious about the location of the zeros. In order to find them easily, it introduces the command "Enable zero places mode." The "place" keyword is recognized and from now on, oscillator sonification is performed only for x-intercepts.

5 Audio Graph Parameters

The application defines many parameters of the mathematical function graph mentioned in the previous chapter in the description of configurable elements. Some of them are related to the sound characteristics, which allow us to experiment with different values. It is also possible to select the oscillator type: (sinusoidal, square, triangular, sawtoothed). As a consequence, we are able to calculate the overall sound fill factor of the graph. For any point P located in the available area of the Cartesian coordinate system, we can determine the sound amplitude value depending on the distance from the line representing the mathematical function and the characteristics of the generated signal. Then it is possible to develop a map of points on the device's screen surface, for which we calculate the amplitude. The average value of the amplitude for all points will be called the sound fill factor S. It can be calculated as a definite integral of a function of two variables, which is called a double integral.

6 Conclusions

While verifying the operation of our prototype graphing calculator, we found that for large areas of the touch screen, where $S = 0$, a visually impaired user may have trouble locating the line presenting the function while exploring the screen. A completely blind user has no such chance at all. On the other hand, when a sound is emitted at every point on the screen, searching for the graph lines is very tedious. Our conclusion and recommendation is therefore that graphing calculators should automatically scale the graph of the mathematical function so that the saturation of the area with sound ensures ease of finding information (areas with audible sound are easily found), while not exceeding the maximum threshold in order to avoid excessive information noise.

The next step in our research will be testing the values of individual parameters of audible graphs (including the saturation coefficient) with the participation of students with visual impairments. The results will allow us to clarify guidelines for designers of mobile mathematical applications with audio function graphs.

References

1. Ibanez, M.B., Mateos, P.: Sonification of mathematical functions: a review. J. Multimodal User Interfaces **11**(3), 209–222 (2017)
2. Gnesdilow, D.S., Kobylkina, L.M.: Approaches to representation of mathematical functions for people with visual impairments. Procedia Comput. Sci. **151**, 675–682 (2019)
3. Sánchez, J., Gómez, E., Pérez, R.: Sonification and audio representation of mathematical functions for people with visual disabilities. J. Vis. Impairment Blindness **109**(1), 43–53 (2015)

4. Araújo, M.C.C., Façanha, A.R., Darin, T.G.R., Sánchez, J., Andrade, R.M.C., Viana, W.: Mobile audio games accessibility evaluation for users who are blind. In: Antona, M., Stephanidis, C. (eds.) UAHCI 2017. LNCS, vol. 10278, pp. 242–259. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-58703-5_18
5. Leporini, B., Palmucci, E.: A mobile educational game accessible to all, including screen reading users on a touch-screen device. In: Proceedings of the 16th World Conference on Mobile and Contextual Learning. pp. 1–4 (2017)
6. Myers, C., Furqan, A., Nebolsky, J., Caro, K., Zhu, J.: Patterns for how users overcome obstacles in voice user interfaces. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, pp. 1–7 (2018)
7. Brauner, D.: Graphing for blind students tutorial (2020). https://www.perkins.org/resource/graphing-blind-students-tutorial/. Accessed 25 Dec 2023
8. WolframAlfa. https://www.wolframalpha.com/. Accessed 25 Dec 2023
9. GeoGebra. https://www.geogebra.org/. Accessed 25 Dec 2023
10. Kim, J., Lee, Y., Seo, I.: Math graphs for the visually impaired: audio presentation of elements of mathematical graphs. In: Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems, pp. 1–6 (2019)

Analyzing the Programming Process with Tangible Tools Using Co-occurrence Matrix

Issei Takida, Tatsuo Motoyoshi(✉), Yotaro Fuse, Bilguunmaa Myagmardulam, Noboru Takagi, Kei Sawai, and Hiroyuki Masuta

Department of Intelligent Robotics Engineering, Toyama Prefectural University, Toyama 939-0398, Japan
u354011@st.pu-toyama.ac.jp

Abstract. We developed a programing education tool "P-CUBE3", which uses blocks with tangible information as its language. P-CUBE3 can control audio output by placing HIRAGANA (Japanese character) blocks on a "program mat", and also implements an operation history acquisition system that can record the type of block, its position, and operation time. We conducted a hands-on programming lesson for visually impaired and sighted people, and analyzed the operation history data when they worked on a task program in the lesson. Participants in the lesson were classified into four groups according to whether they had programming experience or not and whether they had visual impairment or not. We classified the programming process into several patterns, and analyzed the relationship between the patterns using a co-occurrence matrix. We report on the characteristics of the programming process for the visually impaired and sighted.

Keywords: tangible user interface · programming learning · programming process · co-occurrence matrix · visually impaired

1 Introduction

The emergence of new technologies and services such as big data and the IoT is expected to advance the sophistication and diversification of IT utilization [1]. Consequently, it is anticipated that the demand for IT will continue to increase, posing a renewed challenge of shortage of IT personnel. Therefore, there are now more and more opportunities for children to learn programming. Many tools used in programming education rely on PCs or tablets, and there is concern that users unfamiliar with operating information devices or those with visual impairments may develop aversions due to difficulties in handling equipment or applications. Against this backdrop, we have been developing a series of tangible block-based programming tools called the P-CUBE series [2,3] targeting programming beginners, including individuals with visual impairments. The system has been improved so that users can learn not only "sequential", "if", and "loop" but also "subroutine" concepts.

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 217–224, 2024.
https://doi.org/10.1007/978-3-031-62846-7_26

In the conventional P-CUBE series, user comprehension was assessed through a paper-based test that were conducted, neglecting an evaluation of the programming process. Moreover, the paper-based test itself is difficult to implement for the visually impaired. Therefore, we developed a function to record the history of block operations and named the implemented version P-CUBE3. The present study aims to construct an analysis method for the programming process for individuals using operation history data for P-CUBE3.

2 P-CUBE3

2.1 System Configuration

Figure 1 shows the system configuration for P-CUBE3. This tool is capable of building programs and controlling audio output by arranging physical programming blocks on a “program mat”. Information regarding the programming blocks can be obtained from radio-frequency identification (RFID) tags attached to the blocks, which are read by RFID readers attached to the program mat.

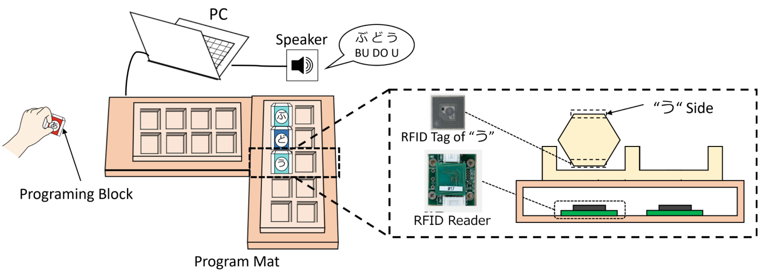
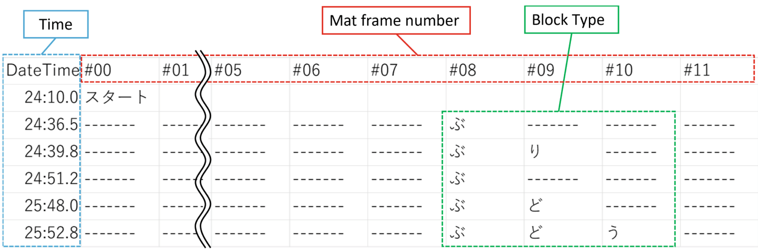


Fig. 1. P-CUBE3 system configuration

2.2 Operation History Acquisition System

P-CUBE3 is equipped with an operation history acquisition system using RFID readers. Figure 2 shows an example CSV file containing operation history data. Each time a block is placed or removed, this information is recorded, together with the time of the operation, the mat frame number, and the block type.



DateTime	#00	#01	#05	#06	#07	#08	#09	#10	#11
24:10.0	スタート								
24:36.5	-------	----	-------	-------	-------	ぶ	-------	-------	-------
24:39.8	-------	----	-------	-------	-------	ぶ	り	-------	-------
24:51.2	-------	----	-------	-------	-------	ぶ	-------	-------	-------
25:48.0	-------	----	-------	-------	-------	ぶ	ど	-------	-------
25:52.8	-------	----	-------	-------	-------	ぶ	ど	う	-------

Fig. 2. Example of operation history data

3 Hands-On Programming Lesson

3.1 Lesson Content

We conducted a hands-on programming lesson using P-CUBE3 for visually impaired and sighted children and obtained operation history data. We asked the participants to work on tasks T1-T3 after explaining how to use P-CUBE3. Figure 3 shows an example of block placement for T1-T3. In Japanese, watermelon is "SU I KA" and melon is "ME RO N".

T1: Say the word "SU I KA" repeatedly.
T2: Say "SU I KA" if switch is ON and "ME RO N" if switch is OFF
T3: Repeatedly say "SU I KA" if switch is ON and "ME RO N" if switch is OFF

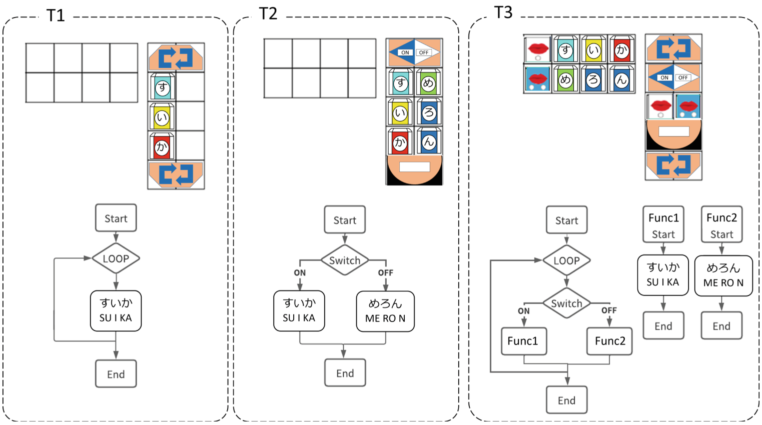


Fig. 3. Example of block placement for T1-T3

3.2 Lesson Participants

Seven visually impaired persons and 10 sighted persons participated in the experiment. The participants were classified according to their programming experience (P-NP) and visual impairment (V-S) as follows. The school grade (as an indicator of programming experience) and the visual abilities of the participants are shown in Table 1.

PV: Visually impaired with programming experience
PS: Sighted with programming experience
NPV: Visually impaired with no programming experience
NPS: Sighted with no programming experience

Table 1. Lesson participants

Participant	Grade	Visual symptom	Participant	Grade	Visual symptom
PV_1	12th	weak sighted	PS_1	4th	sighted
PV_2	7th	weak sighted	PS_2	4th	sighted
PV_3	10th	blind	PS_3	3rd	sighted
NPV_1	10th	weak sighted	NPS_1	2nd	sighted
NPV_2	9th	weak sighted	NPS_2	2nd	sighted
NPV_3	9th	blind	NPS_3	2nd	sighted
NPV_4	2nd	blind	NPS_4	1st	sighted
			NPS_5	1st	sighted
			NPS_6	1st	sighted
			NPS_7	1st	sighted

4 Pre-processing of Operation History Data

4.1 Parameters for Evaluating Programing Process

The parameters used to analyze the operation history data are:. ***Bd*** (Block Edit Distance) is a measure of the workload, and refers to the Levenshtein distance [4], ***Ne*** is the number of necessary blocks for correct placement, and ***Ex*** is the number of extra blocks for correct placement [3].

In order to analyze the programming process, it is necessary to capture the parameter changes at each step of block operation. Since for a single step, each parameter can have a value of “Down“, “Constant”, or “Up”, there are 27 possible patterns, as shown in Table 2. The gray rows indicate patterns that occur in a single block operation. These are designated P1-P6 (from top to bottom) and are described below. Operations P1-P3 are correct operations, and P4-P6 are incorrect. In addition, a special block operation called the “transfer block” is required to execute the program. This operation is designated P0 and is not included in the parameter calculations. Figure 4 shows the parameters and patterns for correct block placement for the word BUDOU (grape in Japanese).

Table 2. Parameter change sequence

Pattern	Bd	Ne	Ex	Pattern	Bd	Ne	Ex	Pattern	Bd	Ne	Ex
1			Down	10			Down	19			Down
2		Down	Const	11		Down	Const	20		Down	Const
3			Up	12			Up	21			Up
4			Down	13			Down	22			Down
5	Down	Const	Const	14	Const	Const	Const	23	Up	Const	Const
6			Up	15			Up	24			Up
7			Down	16			Down	25			Down
8		Up	Const	17		Up	Const	26		Up	Const
9			Up	18			Up	27			Up

P0 Transfer block operation
(program execution)
P1 Placing a block in the correct sequence
(Bd: Down, Ne: Down, Ex: Const)
P2 Selecting an extra block
(Bd: Down, Ne: Const, Ex: Down)
P3 Selecting a block placed in the wrong sequence
(Bd: Down, Ne: Up, Ex: Const)
P4 Placing a block in the wrong sequence
(Bd: Up, Ne: Down, Ex: Const)
P5 Placing an extra block
(Bd: Up, Ne: Const, Ex: Up)
P6 Selecting a block placed in the correct sequence
(Bd: Up, Ne: Up, Ex: Const)

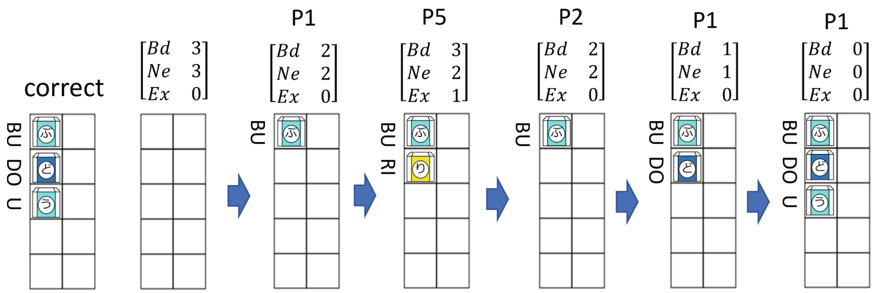


Fig. 4. Parameters for block operation

5 Analysis Methodology for Programming Process

5.1 Co–occurrence Matrix

The co-occurrence matrix represents the number of times multiple elements (usually words or terms) occur simultaneously, and can be used to visualize the relationship between elements.

In the present study, we use a co-occurrence matrix to analyze the relationships between the patterns described in Sect. 3.1. Table 3 shows the co-occurrence matrix for the data acquired during the hands-on programming class using P-CUBE3. By referring to the elements shown in Table 1, the co-occurrence matrix shows the frequency of relationships between patterns, such as "P1 followed by P3 77 times". In this manner, the co-occurrence matrix can be used to analyze the trend of pattern changes.

Table 3. Co-occurrence matrix for patterns

n+1 n	P0	P1	P2	P3	P4	P5	P6
P0	55	12	4	24	9	0	28
P1	120	697	11	77	106	20	219
P2	1	30	5	3	5	5	4
P3	3	144	4	108	80	7	77
P4	7	80	1	146	119	7	43
P5	1	19	25	4	5	5	0
P6	1	269	3	17	59	2	74

5.2 Pattern Breakdown by User Group

Figure 5 shows pie charts indicating the breakdown of correct and incorrect operation patterns for the user groups described in Sect. 3.2. Four different patterns are considered:

CC: Correct operation followed by correct operation.
CI: Correct operation followed by incorrect operation.
IC: Incorrect operation followed by correct operation.
II: Incorrect operation followed by incorrect operation.

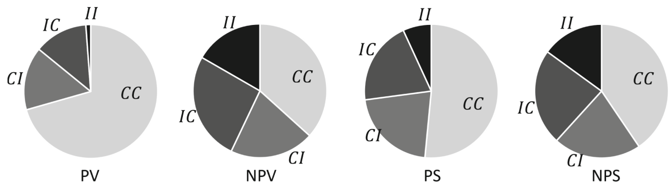


Fig. 5. Breakdown of operation patterns for different user groups

5.3 Chi-Squared Test

In order to determine the relationship between operation patterns and user characteristics, comparisons were made between groups of users with and without programming experience (PV-NPV, PS-NPS) and visually impaired and sighted users (PV-PS, NPV-NPS). The results are shown in Table 4. Test 1 compared the relative number of CC to CI operations, to give an indication of the likelihood of continuously correct operations for each user group. Test 2 compared the relative number of IC to II operations, to give an indication of the likelihood of continuously incorrect operations.

We would expect a higher level of programming proficiency to result in a larger percentage of CC and a smaller percentage of II. Therefore, it is expected that the percentage of CC would be larger and the percentage of II would be smaller for those with programming experience. In addition, visually impaired persons are expected to perform unintended operations due to limited visual information, and are thus more likely to perform incorrect operations. Therefore, it is expected that the percentages of CI and II will be larger for visually impaired persons

In Table 4, results for which significant differences are found are shown in red. It can be seen that for Test 1, significant differences are found between the PV-NPV and PV-PS groups, whereas for Test 2, significant differences are found between the PV-NPV and PS-NPS groups.

Table 4. Chi-squared test results

	PV-NPV	PS-NPS	PV-PS	NPV-NPS
Test 1	0.0001	0.184	0.015	0.666
Test 2	0.003	0.023	0.77	0.931

5.4 Discussion

The results of Test 1 gave a relatively larger number of CC operations for the PV group than for the NPV group, indicating that experienced programmers tended to perform more continuously correct operations. The relative number of

CC operations for the PV group was actually larger than that for the PS group, indicating a high number of continuously correct operations even for the visually impaired. However, the participants in the PV group were in a higher school grade than those in the other groups, which raises concerns that differences may have arisen due to age.

The results of Test 2 showed significant differences when comparing the PV-NPV groups and the PS-NPS groups. Since both comparisons involved participants with and without programming experience, the results indicated that the likelihood of a series of erroneous operations was higher when the user did not have programming experience.

6 Conclusion

We conducted a hands-on class using P-CUBE3 and analyzed the programming process based on the operation history data obtained during the class. Parameters for evaluating block operations were defined, and the changes in these parameters during user programming were classified as patterns. Patterns were classified as correct or incorrect, and a co-occurrence matrix was used to compare the continuity of correct or incorrect operations with or without programming experience, and with or without visual impairment. The results indicated that individuals with programming experience tended to perform more continuously correct operations, whereas individuals with no programming experience showed a higher tendency to perform a sequence of incorrect operations. No significant differences were found between visually impaired and sighted subjects, suggesting the possibility of age-related differences.

References

1. Ministry of Economy, Trade and Industry: Survey Results on the Latest Trends and Future Estimates of IT Human Resources. www.meti.go.jp/shingikai/economy/daiyojisangyoskill/pdf/001s03_00.pdf. Accessed 30 Jan 2024 (in Japanese)
2. Motoyoshi, T., Kakehashi, S., Masuta, H., Koyanagi, K., Oshima, T., Kawakami, H.: Validation of usefulness at the early stage of learning using P-CUBE which block type programming tool. J. Japan Soc. Fuzzy Theory Intell. Inf. **27**(6), 909–920 (2015)
3. Motoyoshi, T., Takida, I., Sawai, K., Masuta, H., Takagi, N.: Implementation and evaluation of a hands-on lesson using P-CUBE3 a tangible programming tool that includes visually impaired users as target users. J. Hum. Interface Soc. **25**(3), 177–188 (2023)
4. Levenshtein, V.: Binary codes capable of correcting deletions. Insertions Reversals Soviet Phys. Doklady **10**, 707–710 (1966)

Tactile Graphics and 3D Models for Blind People and Shape Recognition by Touch

Tactile Graphics and Models for Blind People and Recognition of Shapes by Touch

Introduction to the Special Thematic Session

Tetsuya Watanabe[1](✉), Yoshinori Teshima[2], and Kazunori Minatani[3]

[1] University of Niigata, Ikarashi 2-8050, Nishi-Ku, Niigata 950-2181, Japan
t2.nabe@eng.niigata-u.ac.jp

[2] Chiba Institute of Technology, Tsudanuma 2-17-1, Narashino 275-0016, Chiba, Japan

[3] National Center for University Entrance Examinations, Komaba 2-19-23, Meguro-Ku, Tokyo 153-8501, Japan

Abstract. In the last few decades, remarkable progress has been made in the development of exact 3D solid models. This is mainly attributable to modern technology and techniques such as information technology, shape measurement, and digital manufacturing. 3D models enable blind people to experience the world through touch and are useful as tactile teaching materials not only at schools but also in lifelong learning. In this introduction, we will review previous work on tactile graphics and 3D models and then introduce the presentations in the current STS.

Keywords: Tactile Graphic · 3D Model · Blind and Visually Impaired People

1 Introduction

Research on the use of tactile sense by blind and visually impaired people falls into the fields such as tactile perception, braille, tactile graphics, tactile symbols, three-dimensional (3D) models, and so on. Of these, this STS mainly focuses on tactile graphics and 3D models.

Research on tactile graphics has become active in the ICCHP since the transition of computer interfaces from CUI (character-based UI) to GUI (Graphical UI) and the rise of multimedia computers [1]. These changes have made it easier to handle visual information on computers, and the amount of visual information that people see has increased tremendously. Thus, the means to make visual information accessible to blind and visually impaired people became necessary. In this context, the development of graphic tactile displays to represent dynamically changing GUI screen has begun.

Similarly the recent spread of 3D printers has led to an increase in research on 3D models for blind and visually impaired people. Before the rise of 3D printers, only skilled crafts could make elaborate models. Now, thanks to 3D printers, anyone can make such elaborate models only if 3D data are available. By making most of these state-of-the-art machines, many attempts are being made to increase the opportunities for blind and visually impaired people to touch 3D models.

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 227–231, 2024.
https://doi.org/10.1007/978-3-031-62846-7_27

2 Tactile Graphics and Computers

The first dynamic tactile graphic display would be Dot Matrix Display DMD 12060 (Metec, Germany). In the 2000s, a tactile display using piezoelectric mechanism, the same technology as used in refreshable braille displays, was released from KGS, Japan. However, because of the low pin resolution with inter-pin spaces of 3 mm and the high price, tactile graphic displays have not been widely used yet. In recent years, Orbit Research, the US, has newly released interactive tactile and braille computer, Graphiti (R). Also, American Printing House for the Blind is preparing to release multiline braille display, Monarch, which can render tactile graphics as well. They have not only developed the device but also are involving teachers for visually impaired students and collecting funds for its dissemination in education.

Systems that automatically create tactile graphics from images or numerical data have also been studied extensively. When the image of an object is converted into tactile, important information of it is the shape. To extract the shape data out of the image, basic image processing techniques such as gray-scaling, binarization, and edge detection have been used [2]. However, images contain shadows and color information, which makes object shape recognition difficult. Furthermore, images are the converted form of the real world, which is intrinsically 3D, into 2D, and this also poses cognitive difficulties in imagining a 3D shape from 2D raised line graphics.

It is easier to automatically generate deformed and symbolized tactile graphics, such as tactile graphs and maps [3], from numerical data than converting from images such as pictures. As for these tactile graphics, research is focused on the detectability and distinguishability of tactile symbols on the graphics, and the effect of shape and size of the symbols.

In addition to tactile graphics, audio description is another means of conveying 2D information, and many studies have attempted to explain graphs with audio [4].

3 3D Models for Blind People

In the attempt of supplying 3D models to blind and visually impaired people, some aspects are the same as and some others are different from research on 2D tactile graphics. The difference is that tactile graphic displays have been developed solely for blind and visually impaired people whereas 3D printers have been developed for the general public. For this reason, 3D printers and photogrammetry technology, which creates 3D data from images of an object, are being improved by many companies in the general market steadily and significantly each year. Therefore, the research on and practice of the use of 3D technology for blind and visually impaired people have been focused on verification of its usability, and research on the advantages of 3D models over tactile graphics is of particular interest, especially in production of tactile maps [5, 6]. For easily understand the 3D models by touch, attempts to clarify the requirements of production and explanation methods, and to create guidelines on this have been made [7]. This movement is similar to the creation of guidelines for tactile graphics such as the BANA guideline [8].

4 Presentations in the STS

In the current STS we have eight presentations. Three of them deals with tactile graphics and the other five 3D models.

4.1 Tactile Graphics

A new tool has been developed to automatically generate tactile graphic images of Chinese characters with an embosser. A few characters were made tangible in different font types and the readability of each font by touch were compared [9].

An accessible brainstorming tool has been developed, which is used in meetings and enables blind and visually impaired people to access to nonverbal and spatial information through computers, smartphones, and smartwatches [10].

As we mentioned above, color has been omitted in conventional tactile graphics creating methods. RainbowTact tried to address this issue by automatically translate color information to discernable tactile patterns. In its pilot study, participants were able to distinguish chromatic from achromatic colors well [11].

4.2 3D Models

Two presentations focus on education of blind and visually impaired students using 3D models. ShapeSpace kit has been developed for visually impaired students to learn geometry by folding 2D paper into corresponding 3D shapes [12].

One practical issue with the use of 3D models is how to attach labels to 3D models, which, unlike tactile graphics, are difficult to paste braille labels due to their complex shapes and rough surfaces. Point-and-Tap Interaction recognizes the point where the finger is pointed with the smartphone's camera and reads aloud a description corresponding to the point or object [13]. The number of repeated tapping changes the level of the information to be read.

3D models of internal organs for anatomy classes has been improved. The user test demonstrated the effectiveness of the improvements [14].

The use of 3D models is progressing outside of education as well. Tactile panoramic relief of the city is an attempt to make the landscape 3D. It is not a simple map of the city but mimicking how the human eye perceive the world [15]. The first and third authors report the 3D model distribution service in Japan. How to continue this service after the end of the project that is public-funded is discussed [16].

5 Future Works

In printing 3D models, blind and visually impaired people face the challenges of finding 3D data on the Internet databases, creating 3D data, and operating the 3D printer (it is true that some blind and visually impaired people like the third author of this manuscript do these things by themselves, though). We think that these challenges will be the future research themes in this field.

Acknowledgments. This study was funded by JST RISTEX Japan, Grant Number JPMJRX21I5, and JSPS KAKENHI 20H01705 and 24H00162.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Emiliani, P.L.: Overview of the GUIB project. In: IEE Colloquium on Information Access for People with Disability, pp.11/1–11/3. London, UK (1993)
2. Way, T.P., Barner, K.E.: Automatic visual to tactile translation. I. Human factors, access methods and image manipulation. IEEE Trans. Rehabil. Eng. **5**(1), 81–94 (1997). https://doi.org/10.1109/86.559353
3. Mukhiddinov, M., Kim, S.Y.: A systematic literature review on the automatic creation of tactile graphics for the blind and visually impaired. Processes **9**(10), 1726 (2021). https://doi.org/10.3390/pr9101726
4. Oppegaard, B., Xu, Q., Hurtut, T.: Audio describing tables and charts: a re-visualization process for sharing complex data sets with people who are blind or who have low-vision. Intercom **70**(5), 26–30 (2023)
5. Brittell, M.E., Lobben, A.K., Lawrence, M.M.: Usability evaluation of tactile map symbols across three production technologies. J. Vis. Impairment Blindness **112**(6), 745–758 (2018). https://doi.org/10.1177/0145482X1811200609
6. Holloway, L., Marriott, K., Butler, M.: Accessible maps for the blind: comparing 3D printed models with tactile graphics. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, vol. 198, pp.1–13 (2018). https://doi.org/10.1145/3173574.3173772
7. Leporini, B., Rossetti, V., Furfari, F., Pelagatti, S., Quarta, A.: Design guidelines for an interactive 3D model as a supporting tool for exploring a cultural site by visually impaired and sighted people. ACM Trans. on Accessible Comput. **13**(3), 1–39 (2020). https://doi.org/10.1145/3399679
8. Guidelines and Standards for Tactile Graphics (2022). https://www.brailleauthority.org/sites/default/files/tg/Tactile%20Graphics%20Standards%20and%20Guidelines%202022.pdf. Accessed 11 Apr 2024
9. Sonobe, S., Tanaka, H., Fujiyoshi, A.: Automatic generation of tactile graphics of characters to support handwriting of blind individuals. In: Miesenberger, K., Kobayashi, M., Penaz, P. (eds.) ICCHP 2024, Computers Helping People with Special Needs, 19th International Conference; Linz, Austria, Proceedings. Springer (2024)
10. Koutny, R.: Exploring space: User interfaces for blind and visually impaired people for spatial and non-verbal information. In: Miesenberger, K., Kobayashi, M., Penaz, P. (eds.) ICCHP 2024, Computers Helping People with Special Needs, 19th International Conference; Linz, Austria, Proceedings. Springer (2024)
11. Kim, R., Ka, H.W.: RainbowTact: an automatic tactile graphics translation technique that brings the full spectrum of color to the visually impaired. In: Miesenberger, K., Kobayashi, M., Penaz, P. (eds.) ICCHP 2024, Computers Helping People with Special Needs, 19th International Conference; Linz, Austria, Proceedings. Springer (2024)
12. Aggarwal, M., et al.: Development and assessment of inclusive tangible shape nets to bridge geometry gaps for the visually impaired. In: Miesenberger, K., Kobayashi, M., Penaz, P. (eds.) ICCHP 2024, Computers Helping People with Special Needs, 19th International Conference; Linz, Austria, Proceedings. Springer (2024)

13. Narcisi, A., Shen, H., Ahmetovic, D., Mascetti, S., Coughlan, J.: Accessible point-and-tap interaction for acquiring detailed information about tactile graphics and 3D models. In: Miesenberger, K., Kobayashi, M., Penaz, P. (eds.) ICCHP 2024, Computers Helping People with Special Needs, 19th International Conference; Linz, Austria, Proceedings. Springer (2024)
14. Shinoda, M., Teraguchi, S., Teshima, Y.: Improvements of tabletop model of internal organs for anatomy learning of the visually impaired. In: Miesenberger, K., Kobayashi, M., Penaz, P. (eds.) ICCHP 2024, Computers Helping People with Special Needs, 19th International Conference; Linz, Austria, Proceedings. Springer (2024)
15. Reichinger, A., Rist, F.: Designing an inclusive tactile panoramic relief of the city of Graz. In: Miesenberger, K., Kobayashi, M., Penaz, P. (eds.) ICCHP 2024, Computers Helping People with Special Needs, 19th International Conference; Linz, Austria, Proceedings. Springer (2024)
16. Watanabe, T., Minatani, K.: 3D model distribution service for blind and visually impaired people. In: ICCHP 2024 Open Access Compendium (2024)

Designing an Inclusive Tactile Panoramic Relief of the City of Graz

Andreas Reichinger[1](✉), Florian Rist[2], Georg Regal[3], and Irene Sudy[4]

[1] VRVis Forschungs-GmbH, Donau-City-Strreet 11, 1220 Wien, Austria
reichinger@vrvis.at
[2] TU Wien, Karlsplatz 13, 1040 Wien, Austria
[3] AIT Austrian Institute of Technology GmbH, Center for Technology Experience, Giefinggasse 2, 1210 Vienna, Austria
[4] Hilfsgemeinschaft der Blinden und Sehschwachen Österreichs, Wien, Austria

Abstract. In this paper, we report on the practical experience of designing a tactile representation of the view over a city and landscape from the elevated position of a hill, which was digitally created from a combination of various geographic data sources and photographs. The tactile relief is part of a permanent museum exhibition and is mounted on an outdoor balcony enabling all museum visitors to experience the breathtaking view. Unlike existing works, we intended to create a faithful tactile representation of the view, mimicking many aspects of how the human eye perceives the world. This includes correct panoramic projection, three-dimensional representation of all buildings with realistic surface textures, correct depth layering, and plausible foreshortening not only in image space but also in depth. As perceived by the human eye, near objects are not only larger, but also have a more pronounced depth. We describe the entire process, from design considerations, focus groups, data acquisition, a depth-aware projection mapping algorithm, texture generation, the inclusion of markers and a legend pointing out important buildings, production, mounting, and final evaluation.

Keyword: Design for All and Universal Design · Assistive Technology (AT) · (e)Accessibility · Tactile Relief

1 Introduction

In 2020, the *Graz Museum* has opened a new branch in the former *Stall- and Cannon Bastion* on the iconic *Schlossberg*. In addition to being an important part of Graz's history, the location also offers an outstanding view of the city. As the museum is committed to being a museum for all, they intended to make this special view accessible to everyone. Based on talks with the city's representative for inclusion, a site visit and two workshops with BVI people and representatives, the curators specified the basic concept as *"a translation of the city view into a tactile model that makes the urban landscape with architecture, river, trees and horizon tangible for everyone, especially BVI people, conveying essential points of orientation."*

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 232–239, 2024.
https://doi.org/10.1007/978-3-031-62846-7_28

2 Related Work and Hypothesis

Trying to find references on how to realize the aforementioned concept, we searched the literature, and discussed our findings with the curators, design team and representatives of the target groups in aforementioned workshops.

A large body of work focuses on tactile maps, i.e. top-down plans of streets and buildings. A good summary of recent developments has been collected by Cole [2]. These range from overview plans for limited regions, such as a campus, to multi-page atlases for entire cities or tactile globes [8]. While these are great for wayfinding and orientation, we wanted a more realistic, three-dimensional representation than line drawings.

Examples of more realism are tactile city models, which are already available in many public places, mostly hand-modeled and cast in bronze (e.g. presented on the website [7]). These are good for getting an overview of a particular space, while also having detailed models of the different buildings. However, we wanted a representation that would convey the view from that specific point over the city and way beyond, as far as the mountains over 60 km away, which is not feasible with a scale model.

If everything, up to the horizon should be included, we need something similar to the panoramic images often found at scenic overlooks, presented as prints or etchings, usually with names of points of interest. In our research, we found only a few panoramic examples that had been made accessible to BVI people. One example is a panorama by the artist Adolf Held mounted next to the *Kaiserburg* in *Nuremberg*, Germany, which combines a floor plan of the castle with a view of the city, consisting of rough outlines of rooftops with a few buildings highlighted as flat surfaces together with their names in Braille and raised type [4]. We wanted more realism, detail and a sense of depth, similar to an example by the artist Ádám Farkas in Budapest [7], which renders the view of the *Buda Castle and Castle Hill* in a diorama-like style. While this example is comparatively simple, from the ground looking up, where the hill naturally ends the visibility, our view is from the top of the *Schlossberg* hill looking down, with thousands of buildings and many kilometers of landscape in free sight.

After discussions with the whole team, we posed the hypothesis, that a detailed panoramic relief, which would transform the exact panoramic view into a 3D tactile version, would be interesting for all people, and mounting it on the balcony would allow BVI visitors to experience the same view, alongside seeing visitors. It should not become a simple map of the city, but rather translate the visual qualities, the spatial proximity of the sights, the depth of perspective, etc. in a tactile way, mimicking how the human eye perceives the world. The remaining sections detail the design process, and gives an evaluation about how our hypothesis held in real-world tests.

3 Data Acquisition

As a reference for the relief design, we took a sequence of 32 photos with a Nikon D810 DSLR camera and a 105 mm lens, standing at the future mounting point of the relief, and stitched them into a panoramic image (see Fig. 1a). In the tradition of panoramas mounted at lookout points, we initially chose a cylindrical projection map and cropped it to a region that included the main landmarks. After evaluating different projection options for the vertical direction with all involved, we chose an *equirectangular* projection. This projection has an even spatial distribution proportional to the viewing angle, similar to what a human would observe when scanning the view with their eye movements. Although this projection produces some horizontal squeezing at the top and bottom, it is more compact than the Mercator projection and allows us to show all the important areas on a relief of $1260 \times 450\ \text{mm}^2$, corresponding to a viewing angle of $88° \times 31.5°$ centered on the main viewing direction. This was approved by all involved.

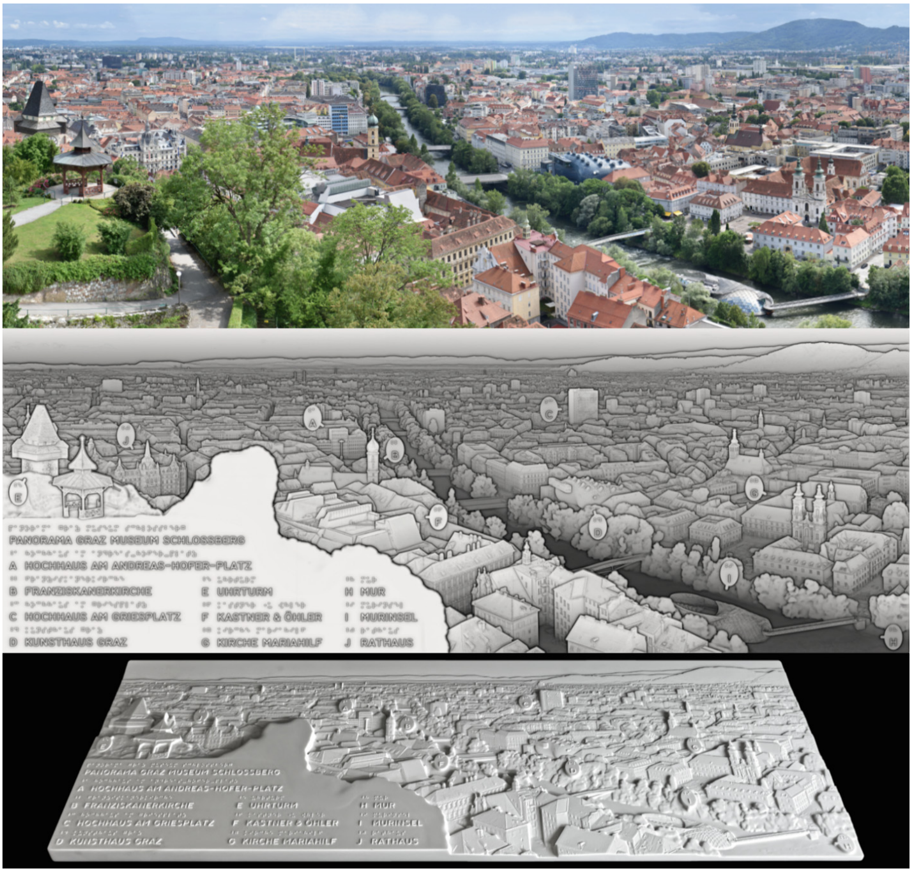


Fig. 1. Stages of the relief development. a) Panorama photo as basis for the tactile version, taken at the mounting point. b) Rendering of the final CAD model. Such previews were used to discuss progress. c) The finished $1260 \times 450 \times 32\ \text{mm}^3$ relief was CNC machined from DuPont Corian®.

In contrast to tactile reliefs made from images only, where the missing depth information has to be created (e.g. [1, 5]), in this case, the 3D geometry of the world is readily available. The *City Surveying Office of Graz* provided us with spatial data of the visible region in a variety of formats, most importantly digital elevation maps, building models derived from roof data, more detailed models of special buildings, and vegetation maps.

Distant regions – where a high resolution is less important – were taken from publicly available elevation data. All sources were individually brought into a common Euclidean coordinate system.

4 Projection-Mapping

In order to bring the world space geometry into the relief space, a suitable projection method had to be developed. To be independent of the capabilities of the rendering software, we performed the panoramic projection mapping directly on the geometry. First, a virtual camera was matched to the panoramic photo and optimized using corresponding points marked in the photo and the 3D scene. The camera-space vertex coordinates (X_c, Y_c, Z_c), where Z_c is the distance along the main axis of the camera, were transformed into image-space coordinates (x_i, y_i, z_i) using Eq. (1), and appropriately scaled and offset to match the pixel coordinates of the reference photo.

$$x_i = \tan^{-1}\frac{X_c}{Z_c},\ y_i = \tan^{-1}\frac{Y_c}{\sqrt{{X_c}^2 + {Z_c}^2}},\ z_i = \frac{1}{\sqrt{{X_c}^2 + {Z_c}^2}} \tag{1}$$

The transforms for x_i and y_i correspond to the standard equirectangular projection to match the projection of the panoramic image. As perceived by the human eye, near objects are not only larger, but also have a more pronounced depth. To mimic this observation in the depth z_i of the relief, we chose the binocular *disparity* used in 3D computer vision. It measures the difference in the projection of a point on two horizontally spaced cameras, like the human eye, which is inversely proportional to distance. The difference in depth of near objects results in a greater difference in disparity, allowing the eyes to infer more detail in depth, than the same difference of far objects. This is how disparity creates the desired effect in the relief. As can be seen in Eq. (1), we ignore the up-axis (Y_c) in the distance calculation so that the distance of a building remains constant over its entire height, resulting in a constant depth in the relief.

The result is a nicely layered diorama-like relief, but it would have to be very thick for individual buildings to have enough depth to feel their shape and layering to adjacent buildings. Similar to the skew transformation performed in [6], we fit a plane to the ground, render its depth, and subtract it from the rendered geometry. This effectively compresses the relief depth, with the ground parallel to the relief plane, but preserves the local depth of the buildings, which are now tilted forward. This gives a good impression of depth all the way to the horizon and maximizes the sense of depth by utilizing the full height of the relief across the whole relief. In addition, important buildings were manually scaled in depth, as much as adjacent buildings would allow, for a more three-dimensional look and feel. The resulting geometry was rendered into a depth map and further processed in this domain (see Fig. 1b for the final depth map).

5 Additional Details, Points of Interest, and Legend

Vegetation and water were specially filtered to create organic, soft shapes that contrast with the edgy buildings and are easily recognizable. Areas of vegetation were detected in the color image using color-based segmentation in an appropriate range of greens and browns. In these areas, the depth of the projected elevation map was filtered using a morphological opening filter with an elliptical structuring element that created a bumpy surface reminiscent of bushes or tree branches.

Nearby buildings, such as the city's landmark *Uhrturm* and the *Chinese Pavilion* were modeled in great detail. Bushes were virtually trimmed to free up views of the city and the Uhrturm's unique clock hands. Many details and over 60 chimneys were added to the coarse building models to match the actual view. Thin features such as church spires were reinforced and slightly tilted backwards for mechanical stability.

Fine details were extracted from the color photo using a series of filters and superimposed on the geometry as slight surface variations. In this way, even the roof shingles, windows and inscriptions can be felt. Cranes were removed from the photo and current construction sites were digitally completed to make the relief timeless.

Tactile markers (see Fig. 1b and c) were designed to point out important landmarks without obstructing the view. The raised round shape is easy to recognize. A line points to the building. A Braille and raised letter refers to the legend, provided by blind experts. The lower left relief area would represent part of the *Schlossberg* and trees. As insufficient geometry was available, it has been flattened to accommodate the legend that names the major landmarks. Blind experts approved this idea, as this avoids long hand movements during tactile exploration, as is often the case with a separate legend.

6 Production and Mounting

The finished relief (see Fig. 1c) measures 1260 × 450 mm^2 and is 32 mm high at its highest point. It was CNC machined from DuPont Corian® solid surfaces, a material that is weather resistant, pleasant to the touch, does not heat up excessively in direct sunlight, and can be machined to a high level of detail. Since the material is only available in thicknesses of up to 19 mm, two boards were glued together using Corian glue, which is almost indistinguishable from the solid material. In December 2020, the relief was mounted outdoors on the balcony railing with a metal structure that allowed for thermal expansion. According to regulations, it is mounted at a height of approximately 80 cm, is fully wheelchair accessible and at a comfortable height for adults while still being reachable for children (see Fig. 2).

Fig. 2. The tactile relief mounted at the balcony, allows perceiving the view even in bad weather.

7 Evaluation

In the 3.3 years, since the installment of the relief end of 2020 until the end of March 2024, the museum has counted 172,732 visitors, of which most have seen the relief. Three main evaluation rounds were conducted. *The first evaluation round* was conducted as a workshop with five BVI persons on June 31st 2021. Four persons were legally blind, and one person was visually impaired with 20% visual sight. The feedback was provided during a guided museum inspection including the relief, which also resulted in an audio-guide for the whole museum [3]. *The second evaluation round* comprised feedback from around 50 to 60 BVI persons who participated in special guided tours through the entire exhibition including the relief and occasional feedback from other visitors and museum staff. Since the installment of the relief, 3–4 special guided tours each year with around 3–10 BVI persons for each tour have been conducted so far. The feedback of those persons was gathered by two different museum guides who then forwarded the given feedback to us. The feedback conducted in the wild was less structured than the feedback of the workshop round, but it is of high value as it gives us important feedback on the relief in the real world and daily usage irrespective of weather conditions and daytime since its installment. *The third evaluation round* was conducted as a group session at the museum in form of structured interviews deepening on feedback gathered so far, with three fully blind persons, one visually impaired person and one orientation and mobility trainer on April 8th 2024. The feedback of the three evaluation rounds was structured into *(a) overall impression*, *(b) location/placement*, *(c) comprehension of perspective, (d) haptics*, and, *(e) contrast*.

The *(a) overall impression* was perceived positively by the participating persons, especially as none of them had ever seen a relief with these specific features before.

The evaluating participants liked the *(b) location and the placement* of the relief at the exact spot where the panoramic photo was taken. BVI people could then simultaneously understand what sighted people see from there. The outdoor placement of the relief was sometimes perceived as disturbing. Some people did not want to touch the relief because of pollution caused by environmental exposure. Dirt or traces of rainwater were an unpleasant surprise and negatively influenced the experience. For booked tours for BVI people, this was not a problem, as the surface was cleaned before the tour arrived. In order to improve the situation, draining holes could be made retrospectively in places where water collects and a canopy could be installed.

The idea of a *(c) perspective visualization* was perceived as very good in itself, as most BVI persons are not yet familiar with models of this kind and it therefore enables a completely new perspective. However, people who are blind from birth have no concept of perspective and have never learnt how to imagine a view. Those people could not understand the relief very well without help. Blind people usually perceive maps or models from a bird's eye view. It therefore took them some time to get used to this impression and to grasp it. In general, many blind people have difficulties with two-dimensional views of objects. Therefore, a tactile model that depicts the buildings and surroundings in three dimensions would be good for understanding the relations. In combination with explanations from the tour guides or an audio guide, however, they managed to get a sense of it. The audio guide could be improved, by giving more details to selected buildings, and could offer an introduction to perspective for people who are new to that concept. Nevertheless, when asked specifically, people could perceive nearer objects as larger, could detect prominent buildings like church towers and found it interesting to learn how the human eye perceives the world.

The *(d) haptics* of the material was perceived as pleasant, compared to metal models in particular. The evaluating BVI persons realized the different haptic surfaces and what they stand for only with visual assistance and guided help. It would be ideal if buildings had a different surface structure than trees and shrubs. We specifically made the vegetation more rounded in design and left the houses more angular; however, feedback showed that this was not enough. In order to be able to work these out independently by BVI persons, a corresponding legend with the respective tactile section and its respective meaning would be necessary, for example "smooth undulating surface" for vegetation along a river. Some types of visualization in the tactile panorama also differ from those usually used in tactile maps, e.g. ripples for water.

Visually impaired people negatively mentioned the missing *(e) contrast*, the monochromatic design and the white surface of the relief. It would be desirable to have a better color contrast on the raised text, and, e.g. individual buildings highlighted in different colors, especially those buildings that appear in the legend. Indeed, a multi-coloring of the relief was discussed during design, but then discarded for budget reasons. The problem was somehow tackled by offering support by an audio guide, but coloring is indeed an important aspect to consider for future work. We are also evaluating ways to add color subsequently.

8 Conclusion and Future Work

We presented a method for creating a viewpoint for BVI people that aims to faithfully translate the visual experience into a tactile relief based on a combination of photographs and 3D GIS data. The resulting relief is now part of the permanent exhibition and has already been examined by many visitors, including groups of BVI people. Initial feedback from museum staff and visitors has been very positive. Blind and sighted visitors can experience it side-by-side, and it fits the inclusive direction of the exhibition. It can be read like similar non-tactile panoramas, but adds a sense of shape and depth. The evaluation indicates that there is a general interest in tactile reliefs of such views; however, many aspects can be improved. People not familiar with the concept of perspective would need special learning material, different objects need a more clear differentiation, e.g. by texture, and contrasting colors are important for visually impaired people. Comprehension could be further improved by an in-depth audio guide. In a follow-up work, we intend to improve these issues, and investigate ways to transform this procedure into an automatic workflow that could generate a similar relief from any viewpoint on Earth, as long as elevation data of the visible region are available in sufficient detail.

Acknowledgments. This work was a collaborative effort with Graz Museum under curator Martina Zerovnik, together with Ingrid Holzschuh, Buero41a, Studio WG3, and all participants of the focus groups, with data kindly provided by Stadtvermessungsamt Graz, and produced by the model workshop at Institut für Kunst und Gestaltung, TU Wien. The research was partly funded by the project BeauCoup, which has received funding from the AAL Joint Programme under grant agreement No AAL-2021–8-156-CP.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Buonamici, F., Furferi, R., Governi, L., Puggelli, L., Volpe, Y.: Tactile reproduction of paintings: the experience of the department of industrial engineering of florence. IOP Conf. Ser. Mater. Sci. Eng. **364**, 012101 (2018). https://doi.org/10.1088/1757-899X/364/1/012101
2. Cole, H.: Tactile cartography in the digital age: a review and research agenda. Prog. Hum. Geogr. **45**(4), 834–854 (2021)
3. Graz Museum Schlossberg, Audio-Guide, Tactile Panorama. https://audioguide.grazmuseum.at/graz-museum-schlossberg/chapters/easy_english/6. Accessed 12 Apr 2024
4. Kalb R. 2009. Mit dem Finger über Dächer. https://www.nordbayern.de/region/nuernberg/mit-dem-finger-uber-dacher-1.599843. Accessed 12 Apr 2024
5. Reichinger, A., Neumüller, M., Rist, F., Maierhofer, S., Purgathofer, W.: Computer-aided design of tactile models - taxonomy and case studies. In: Miesenberger, K., Karshmer, A., Penaz, P., Zagler, W. (eds.) Computers Helping People with Special Needs. Lecture Notes in Computer Science, vol. 7383, pp. 497–504. Springer, Berlin / Heidelberg (2012)
6. Reichinger, A., Carrizosa, H.G., Travnicek, C.: Designing an interactive tactile relief of the meissen table fountain. In: Miesenberger, K., Kouroupetroglou, G. (eds.) Computers Helping People with Special Needs, pp. 209–216. Springer International Publishing, Cham (2018). https://doi.org/10.1007/978-3-319-94274-2_28
7. Tapintható-Láthatatlan. 2008. Budapest, panoramic view of Buda. https://tapinthatolathatatlan.hu/budapest_panoramic_view. Accessed 12 Apr 2024
8. Teshima, Y., et al.: Development of tactile globe by additive manufacturing. In: Miesenberger, K., Manduchi, R., Rodriguez, M.C., Peňáz, P. (eds.) Computers Helping People with Special Needs: 17th International Conference, ICCHP 2020, Lecco, Italy, September 9–11, 2020, Proceedings, Part I, pp. 419–426. Springer International Publishing, Cham (2020). https://doi.org/10.1007/978-3-030-58796-3_49

Improvements of Tabletop Model of Internal Organs for Anatomy Learning of the Visually Impaired

Moeka Shinoda[1], Sayaka Teraguchi[2], and Yoshinori Teshima[3](✉)

[1] Master's Program in Innovative Mechanical and Electronic Engineering, Chiba Institute of Technology, 2-17-1 Tsudanuma, Narashino 275-0016, Chiba, Japan
[2] Hiroshima Central Special Support School, 2-1-4 Hesakasenzoku, Higashiku, Hiroshima 732-0009, Hiroshima, Japan
[3] Department of Innovative Mechanical and Electronic Engineering, Chiba Institute of Technology, 2-17-1 Tsudanuma, Narashino 275-0016, Chiba, Japan
yoshinori.teshima@it-chiba.ac.jp

Abstract. This study established an enhanced tabletop model of internal organs designed to anatomy learning for students with visual impairments. We implemented several modifications to the model presented in 2022. These included adjusting the color scheme of the respiratory system components, correcting the positional relationship between the pancreas and stomach, and modifying the shape of the large intestine. Unlike the previous model, which did not incorporate magnets to connect or fixate organs, this updated model integrates this feature. As a result of these modifications, we have developed three variations of the model: Model-A, which does not use magnets to connect and fixate organs; Model-B, which employs magnets solely for fixation purposes and not for connectivity; and Model-C, which uses magnets for both organ connection and fixation. The results of the evaluation experiment revealed that Model-B serves as a superior instructional tool in terms of operability.

Keywords: Internal Organ Models · Anatomy Teaching Materials · Tactile Learning · Shape Simplification · 3D Printing

1 Introduction

For individuals with visual impairment to comprehend shapes, it is necessary to use appropriate three-dimensional (3D) models as instructional tools, as understanding shapes solely from diagrams on paper is impossible. For instance, various services facilitating the creation and use of tactile maps have been made available for the visually impaired and their caregivers [1–3]. Previous studies have yielded numerous 3D teaching materials tailored for the visually impaired [4], encompassing models depicting mathematical curved surfaces [5], enlarged models of plankton skeletons [6], anatomical heart models [7], and tactile globes [8]. Within the physical therapy department of institution for the visually impaired, students aspiring to become licensed masseurs, acupuncture

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 240–251, 2024.
https://doi.org/10.1007/978-3-031-62846-7_29

practitioners, or moxibustion therapists undergo anatomical studies to understand the structure and function of the human body. The developed models were actively utilized by visually impaired students at the school for the blind. In this study, we developed a tabletop model of internal organs tailored specifically for visually impaired individuals, enabling them to learn about organ configuration and shape through tactile exploration. The model, presented in 2022 [9, 10] is shown in Fig. 1 [10]. Subsequently, a visually impaired educator from the school for the blind, who used the model, provided feedback requesting three key enhancements. First, there was a need to differentiate the respiratory system (lungs and trachea) and the circulatory system (heart) by utilizing two different colors. Second, the original model inaccurately placed the stomach and pancreas on the same plane, disregarding the actual position of the pancreas behind the stomach. Therefore, the positional relationship between these organs required correction. Third, the narrow operating space for inserting the small intestine component into the large intestine necessitated a redesign of the large intestine segment. In this study, we implemented these improvements and additionally introduced magnets to facilitate the connection and fixation of organ parts.

Fig. 1. Tabletop model of internal organs presented in 2022 [10]

2 Method

Commercially available organ models are often unsuitable for visually impaired individuals beginning to learn. While these models accurately replicate the shape of actual organs, their complex shapes can make it feel difficult beginners. Moreover, the crowded arrangement of organs in these models makes it challenging for beginners to understand their interconnections. Regarding size, commercially available organ models are typically life-size, making them impractical for individual student use. To address this issue, we resized the model to a more manageable size that can be easily manipulated on a desk. Furthermore, the intricate shapes of organs in commercially available models contain complex details that, although reflects accurate details of the organ, is too difficult for beginners. Thus, we simplified model of organs appropriately. The degree of simplification was not determined was determined through collaborative discussions between the model developer and the teacher of the school for the blind. In reality, organs are closely packed within the body, but in this model, we intentionally introduced gaps between organs to enhance tactile differentiation during palpation. In this study, we incorporated the three key improvements mentioned earlier to the previous model [10]. Furthermore, we introduced magnets to connect and fixate the components of the model.

2.1 The Color Changing of the Respiratory System

The color of the respiratory system components has been changed in the improved model. In the previous model, all three organ parts (lungs, trachea, and heart) were red (Fig. 2(a)). However, in response to feedback and to enhance differentiation or both partially sighted and sighted individuals, the color of the lungs and trachea was changed from red to green. The updated model is shown in the Fig. 2(b).

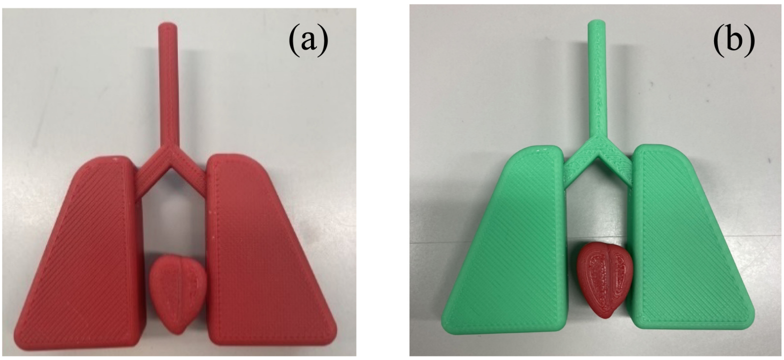


Fig. 2. (a) Respiratory system and heart component before improvement [10] (all three parts are red). (b) Respiratory system and heart component after improvement (respiratory system: green, heart component: red) (Color figure online)

2.2 Improvement of the Positional Relationship Between the Stomach and Pancreas Parts

The shape of the duodenum has been improved to more accurately position the pancreas in the model. The previous model, shown in Fig. 3(a), showcases the stomach, pancreas, and duodenum all located on the same plane, pancreas as illustrated within the red circle. However, given that the pancreas is a retroperitoneal organ, it is positioned dorsal to the stomach. Therefore, the position of the pancreas was adjusted to overlap with the stomach in the improved model. In order to achieve this anatomically correct positioning while maintaining contact between the curved sections of the pancreas and duodenum, the duodenum part was redesigned into a 3D shape. Specifically, the duodenum now extends dorsally from the stomach to the pancreas, and then curves towards the ventral side to connect with small intestine. The resultant model reflecting the improvement is shown in Fig. 3(b).

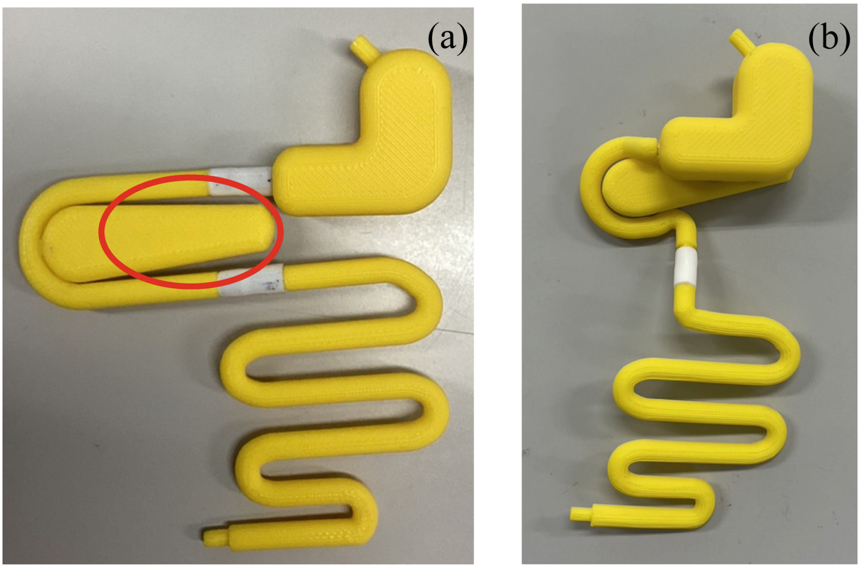


Fig. 3. (a) Organ segments before improvement [10] (around the pancreas). (b) Organ segments after improvement (around the pancreas)

2.3 The Shape Modification of the Large Intestine Part

The shape of the large intestine was modified to address the issue of narrow operating space during the insertion of the small intestine component into the large intestine in the previous model. Figure 4 (left) shows the previous model and Fig. 4 (right) shows the model after improvement. The initial model presented with difficulties in inserting the small intestine into the large intestine because of the constrained space highlighted within the red circle in the left side of Fig. 4. Therefore, we modified the shape of the tail end of the large intestine to ensure sufficient operating space, as illustrated within the green circle on the right side of Fig. 4.

2.4 Utilization of Magnets

In this study, we introduced magnets into the model, a differentiation from our previous model which did not incorporate magnets. We developed three prototype models, each employing different methods for connecting and fixing organ parts. Internal organ model

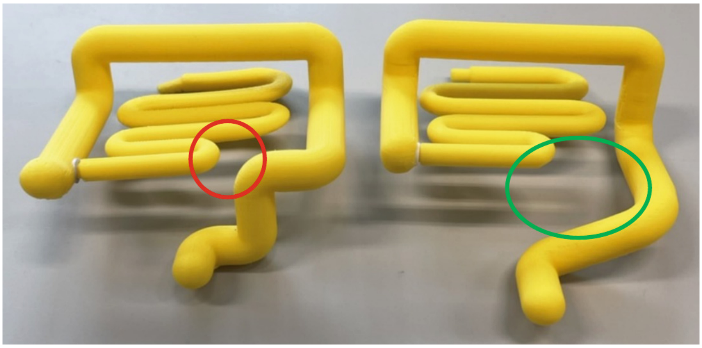


Fig. 4. Large intestine and small intestine (left: before improvement [10], right: after improvement)

A (Model-A) uses pins and rubber tubes exclusively for connecting and securing each organ part, without the use of magnets. Internal organ model B (Model-B) is equipped with pins, rubber tubes, and magnets. Pins and rubber tubes were used to connect between organ parts where actual connections exist in human anatomy. However, as our models do not include the vascular system, connection between certain organs such as the lungs and heart, and the stomach and spleen were omitted. Therefore, magnets were used to conveniently fixate the organs in three locations where no natural connection exist, namely between the left lung and heart, between the stomach and spleen, and between the stomach and pancreas. This facilitates convenient fixation of organs, allowing students to distinguish between parts with actual connections (without magnets) and those without (fixed with a magnet). Internal organ model C (Model-C) uses magnets for both the connection and fixation of organ parts.

2.5 3D Printing

Each organ was designed using 3D-CAD software (Autodesk Fusion 360) according to the improvement policy described above. After designing, we used 3D printers (Afinia3D H800 and Afinia3D H + 1), which employ Fused Filament Fabrication (FFF), to create the model. The filament material was acrylonitrile-butadiene-styrene (ABS) resin. The internal organ model developed in this study features parts where organs are connected using pins. The pins have a diameter of 3 mm across all organs, necessitating a selection of infill density of 65% to increase its strength. Given the substantial size of the pedestal for placing organ parts, an infill density of 20% was selected. The resulting organ parts of Model-B are shown from Figs. 5, 6, 7, 8, 9, 10, 11 and 12.

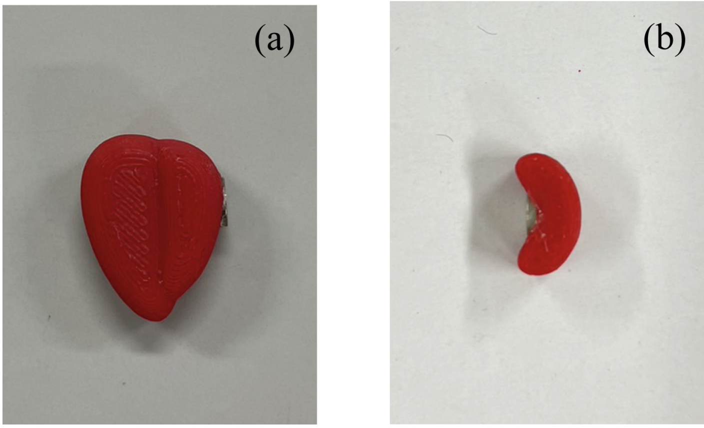


Fig. 5. (a) Heart component: A magnet was embedded in the heart to fixate it to the left lung. (b) Spleen component: A magnet was embedded in the spleen to fixate it to the stomach.

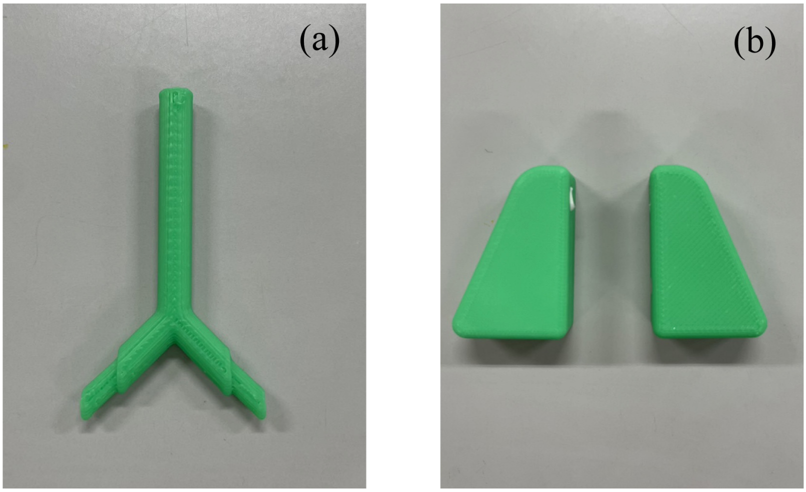


Fig. 6. (a) Trachea component: Insert the pins at the ends of each of the two branches into the holes in both lungs. (b) Two lung segments: Each lung had a hole on the inner surface for connecting to the trachea. The left lung had a magnet to fixate it to the heart.

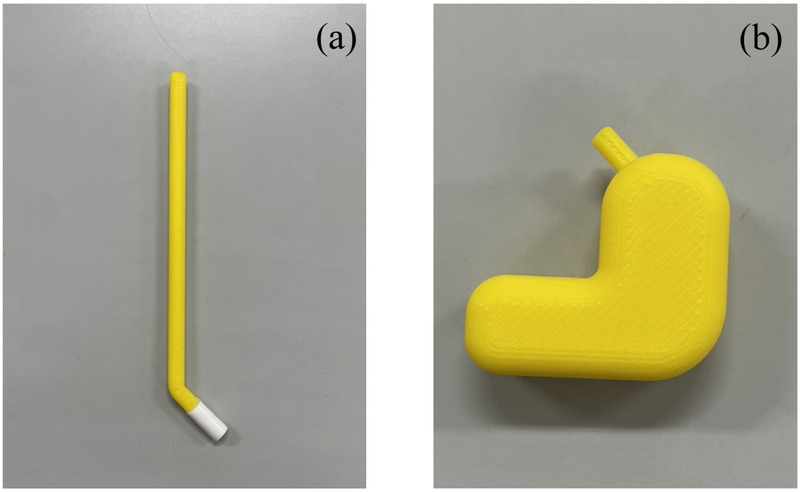


Fig. 7. (a) Esophagus component: A rubber tube was installed at the bottom end to connect the entrance pin of the stomach. (b) Stomach component: An entrance pin (upper) connects to the esophagus and a pin of the duodenum inserts to a hole (left lower). There is a magnet embedded in the upper right curved portion to fixate the spleen.

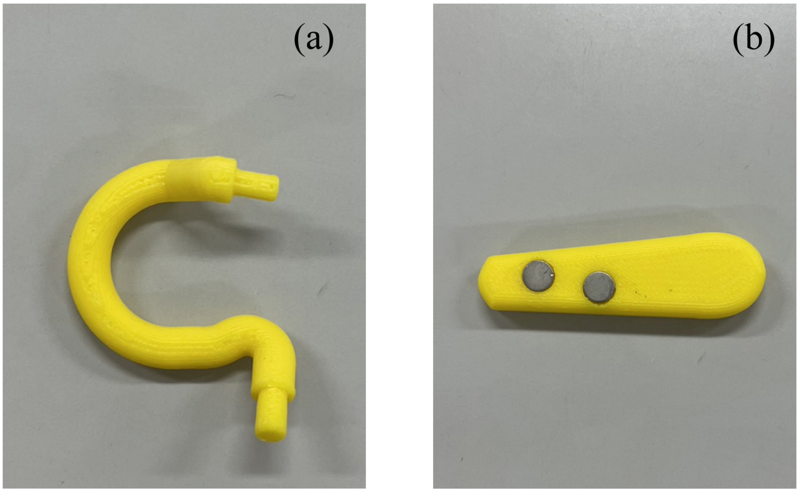


Fig. 8. (a) Duodenum component: The pin at the upper right end connects to the hole of the stomach, and the pin at the lower end connects to the rubber tube of the entrance of small intestine. (b) Pancreas component: The two magnets embedded in pancreas fixate it to the stomach.

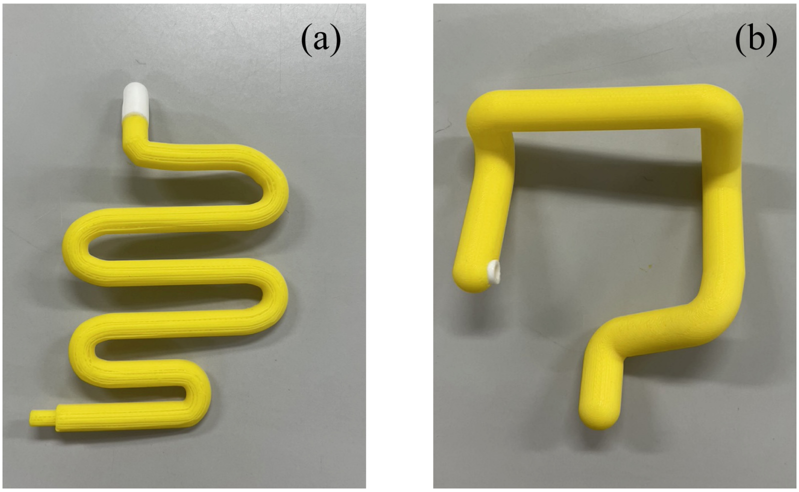


Fig. 9. (a) Small intestine component: An entrance rubber tube (upper end) connects to the duodenum and an exit pin (lower end) connects to the large intestine. (b) Large intestine component: A hole for connecting an exit pin of the small intestine was installed in front of the left side (rubber tube embedded).

Fig. 10. Liver component: The left-lobe part (right side of the figure) is thin. The right-lobe part is thick and fits into the concave portion of the pedestal.

3 Evaluation Experiment

Evaluation experiments were conducted to assess the effectiveness of the internal organ models developed in this study. We compared two models (Model-A and Model-B) in the experiment. Model-C, which utilized magnets for all organ connections and fixation, presented issues during preliminary testing, including unintended rotation of organ parts after connection and inadvertent connections due to magnetic forces. Therefore, Model-C was excluded in the evaluation experiment. The experiment involved a total of nine participants, comprising six sighted individuals and three visually impaired individuals (one completely blind and two partially sighted). To ensure consistency, all sighted participants wore eye masks throughout the experiment. The flow of the investigation was as follows: First, each participant was introduced by tester to the shape of the internal organ parts and how to connect and fixate them. In this experiment, “connect” referred

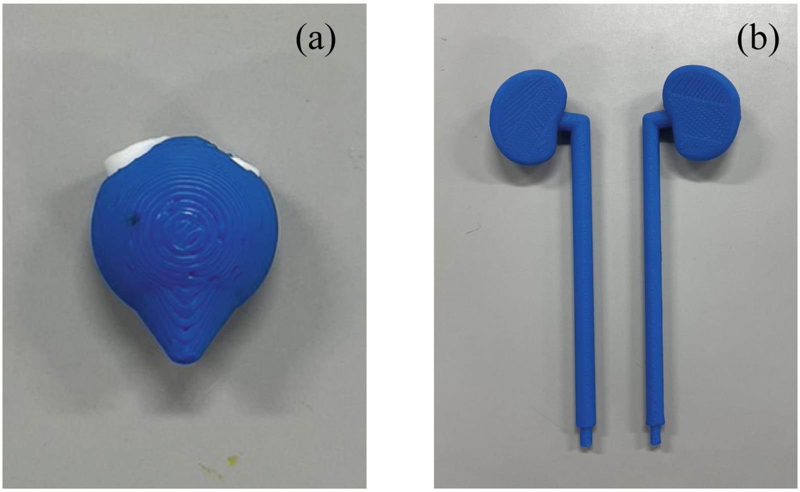


Fig. 11. (a) Bladder component: There are two holes on the upper side of the bladder model for connecting the two ureters. The exserted part on the lower side represents the urine outlet. (b) Kidney and ureter segments: The model of the kidney and ureter were pre-combined. A pin was placed at the end of each ureter to connect it to the bladder.

to actual anatomical connections between two organs, while "fix" indicated areas where no anatomical connection existed. Organ parts were handed over individually to each participant, accompanied by verbal explanations. The experiment sequence followed the anatomical hierarchy from upper to lower, beginning from the respiratory system, moving through the heart, digestive system, spleen, and concluding with the urinary system. Subsequently, the participants learned about the pedestal, which featured concave portions designed to accommodate the organ parts. Participants were then instructed about the positions of the concave portions of the pedestal, progressing from upper to lower sections with hands-on guidance. Subsequently, the participants placed the organ parts on the pedestal, applying their understanding.

The tester then asked four questions to each subject.

Q1 Which model was easier to connect or fixate the organs together?
Q2 Which model was more stable after the organs were connected or fixed?
Q3 Which model was more stable when organ parts were placed on the pedestal?
Q4 After understanding the usage of this teaching material, which model do you think can be used without assistance?

Subjects were provided with four answer options for each question:

- Model-A is more applicable
- Model-B is more applicable
- Both Model-A and Model-B are applicable
- Neither Model-A nor Model-B is applicable

Summary of of response for each question is shown in Table 1.

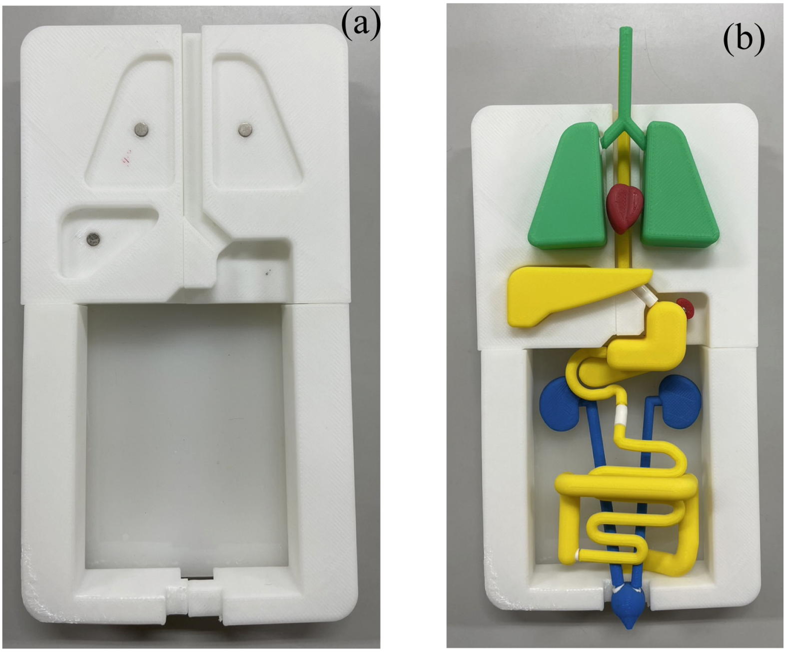


Fig. 12. A pedestal of Model-B. No internal organ part is placed in the pedestal. The size of the pedestal is 215 mm (length) × 120 mm (width) × 35 mm (height). (b) A pedestal with all internal organ components.

Table 1. Summary of response for each question (combined results for sighted and visually impaired individuals)

	Q1	Q2	Q3	Q4
Model-A	0%	33%	0%	0%
Model-B	78%	56%	89%	78%
Both A and B	22%	0%	0%	11%
Neither A nor B	0%	11%	11%	11%

4 Discussion

The response rate for internal organ Model-B was consistently higher across all questions, indicating that Model-B was perceived as easier to connect, more stable, and easier to handle than Model-A. For Question-1, 78% of participants answered Model-B, likely due to the use of magnets in Model-B, which facilitated easy handling of organ parts. For Question-2, 56% of participants answered Model-B, which was slightly lower than the response rate for the other questions. This could be attributed to instances where the first organ part fixed with a magnet came off during the connection process. In particular, there were observations of the spleen detaching from the stomach during the experiment. For

Question-3, 89% of participants answered that Model-B was more stable when placed on a pedestal. This stability can be attributed to the magnets embedded in the back of the liver (Fig. 13 (left)) and each lung (Fig. 13 (right)) of Model-B, as well as in the concave portions of the pedestal, ensuring fixation of the organs. Conversely, Model-A used no magnets. Therefore, it is believed that many participants answered that Model-B was more stable after being placed on the pedestal. For Question-4, 78% of participants answered Model-B. Feedback from three sighted participants highlighted difficulties in placing the ureter and the terminal part of the large intestine on the pedestal, a common challenge noted for both Model-A and B.

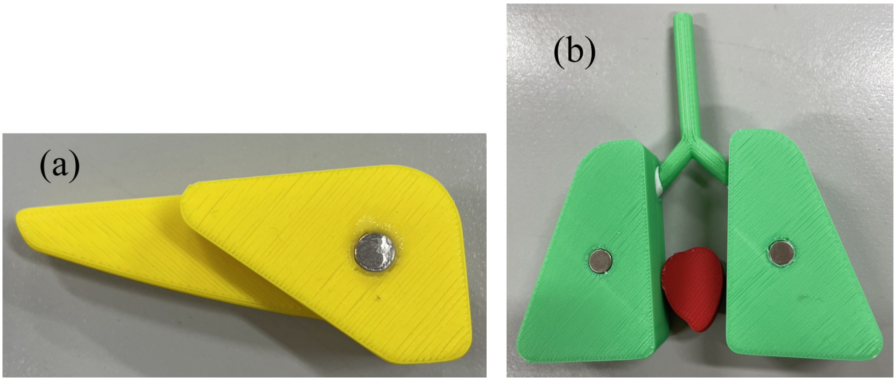


Fig. 13. (a) A magnet was embedded in the back of the liver part. (b) A magnet was embedded in the back of each lung segment.

5 Conclusions

We improved a tabletop model [10] of internal organs for visually impaired individuals by implementing the following improvements. Color adjustment of the respiratory system for better distinction, configuration refinement of the stomach, pancreas, and duodenum to accurately reflect anatomical positioning, modification of the shape of the large intestine to facilitate smoother operation, and introduction of magnets to enhance the operability of the model. Evaluation experiments revealed that internal organ Model-B, which used pins, rubber tubes, and magnets, outperformed the other model in terms of operability, making it a superior teaching aid for visually impaired individuals.

Acknowledgments. This study was partially supported by JSPS KAKENHI Grant Number 24H00162. We would like to thank Editage (www.editage.jp) for English language editing.

References

1. Minatani, K., et al.: Tactile map automated creation system to enhance the mobility of blind persons - its design concept and evaluation through experiment. In: Miesenberger, K., Klaus, J., Zagler, W., Karshmer, A. (eds.) ICCHP 2010, LNCS, vol. 6180, pp. 534–540.Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-14100-3_80
2. Watanabe, T., Yamaguchi, T., Koda, S., Minatani, K.: Tactile map automated creation system using openstreetmap. In: Miesenberger, K., Fels, D., Archambault, D., Peňáz, P., Zagler, W. (eds.) ICCHP 2014. LNCS, vol. 8548, pp. 42–49. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-08599-9_7
3. Watanabe, T., Yamaguchi, T.: Six-and-a-half-year practice of tactile map creation service. Stud. Health Technol. Inform. **242**, 687–694 (2017)
4. Teshima, Y.: Three-dimensional tactile models for blind people and recognition of 3D objects by touch. In: Miesenberger, K., Klaus, J., Zagler, W., Karshmer, A. (eds.) ICCHP 2010, LNCS, vol. 6180, pp. 513–514. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-14100-3_76
5. Teshima, Y., et al.: Models of mathematically defined curved surfaces for tactile learning. In: Miesenberger, K., Klaus, J., Zagler, W., Karshmer, A. (eds.) ICCHP 2010. LNCS, vol. 6180, pp. 515–522. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-14100-3_77
6. Teshima, Y., et al.: Enlarged skelton models of plankton for tactile teaching. In: Miesenberger, K., Klaus, J., Zagler, W., Karshmer, A. (eds.) ICCHP 2010, LNCS, vol. 6180, pp. 523–526. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-14100-3_78
7. Yamazawa, K., et al.: Three-dimensional model fabricated by layered manufacturing for visually handicapped persons to trace heart shape. In: Miesenberger, K., Karshmer, A., Penaz, P., Zagler, W. (eds.) ICCHP 2012, LNCS, vol. 7383, pp. 505–508. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-31534-3_74
8. Teshima, Y., et al.: Development of tactile globe by additive manufacturing. In: Miesenberger, K., Manduchi, R., Rodriguez, M.C., Peňáz, P. (eds.) ICCHP 2020. LNCS, vol. 12376, pp. 419–426. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-58796-3_49
9. Shinoda, M., Koike, A., Teraguchi, S., Teshima, Y.: Development of tabletop models of internal organs for anatomy learning of the visually impaired. In: Miesenberger, K., Kouroupetroglou, G., Mavrou, K., Manduchi, R., Covarrubias Rodriguez, M., Penáz, P. (eds.) Proceedings of Symposium of the Society for Science on Form, Japan, vol.7, No.1, pp. 52–53. Springer, Heidelberg (2022). https://doi.org/10.1007/978-3-031-08648-9_30
10. Shinoda, M., Koike, A., Teraguchi, S., Teshima, Y.: Development of tabletop models of internal organs for anatomy learning of the visually impaired. In: Miesenberger, K., Kouroupetroglou, G., Mavrou, K., Manduchi, R., Covarrubias Rodriguez, M., Penáz, P. (eds.) ICCHP 2022, LNCS, vol. 13341, pp. 261–269. Springer, Heidelberg (2022). https://doi.org/10.1007/978-3-031-08648-9_30

Accessible Point-and-Tap Interaction for Acquiring Detailed Information About Tactile Graphics and 3D Models

Andrea Narcisi[1], Huiying Shen[2], Dragan Ahmetovic[1], Sergio Mascetti[1], and James M. Coughlan[2](✉)

[1] Department of Computer Science, Università degli studi di Milano, Milan, Italy
andrea.narcisi@studenti.unimi.it,
{dragan.ahmetovic,sergio.mascetti}@unimi.it
[2] Smith-Kettlewell Eye Research Institute, San Francisco, CA, USA
{hshen,coughlan}@ski.org

Abstract. We have devised a novel "Point-and-Tap" interface that enables people who are blind or visually impaired (BVI) to easily acquire multiple levels of information about tactile graphics and 3D models. The interface uses an iPhone's depth and color cameras to track the user's hands while they interact with a model. When the user points to a feature of interest on the model with their index finger, the system reads aloud basic information about that feature. For additional information, the user lifts their index finger and taps the feature again. This process can be repeated multiple times to access additional levels of information. For instance, tapping once on a region in a tactile map could trigger the name of the region, with subsequent taps eliciting the population, area, climate, etc. No audio labels are triggered unless the user makes a pointing gesture, which allows the user to explore the model freely with one or both hands. Multiple taps can be used to skip through information levels quickly, with each tap interrupting the current utterance. This allows users to reach the desired level of information more quickly than listening to all levels in sequence. Experiments with six BVI participants demonstrate that the approach is practical, easy to learn and effective.

Keywords: Tactile graphics · 3D models · Blindness · Low vision · Audio labels

1 State of the Art and Related Technology

Tactile graphics and 3D models are indispensable tools for people who are blind or visually impaired (BVI) to access information [12]. While the shapes and structures of such materials, which include tactile maps, relief maps and educational models (e.g., a molecule, skeleton or biological cell), can be accessed tactilely even without vision, other important information such as color, visual texture and printed information is often inaccessible to BVI users. Braille is

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 252–259, 2024.
https://doi.org/10.1007/978-3-031-62846-7_30

often used to label important features on these materials, but there is usually only space for short braille abbreviations (which require a separate braille key defining the abbreviations), and many BVI people don't read braille [11].

A powerful solution to address the inaccessibility of many tactile materials is to make them interactive using a touch-based interface [1]. For instance, the T3 Tactile Tablet[1] from Touch Graphics allows the user to overlay a tactile graphic on a large Android tablet; the device's touchscreen senses finger contact with the graphic and issues a specific audio label for each tactile feature on the graphic that the user touches. In this way BVI users can explore tactile graphics and access semantic information about them whether or not they read braille, using a touch-based interface that is natural in part because of its similarity to the standard touchscreen interface already available on mobile devices.

Computer vision approaches have the potential to make virtually any tactile graphic or 3D model interactive, and are not restricted to 2D graphics that can be overlaid on a tablet touchscreen (and are thin enough for the tablet to sense finger touches). Past work on this approach has been implemented using special cameras with depth sensors [8,10] or other special hardware[2], standard webcams [9] and as smartphone apps [4,5,13]. Some of these approaches have the user point to features by holding a special pointing tool [4], or require the user to augment their fingertip with some sort of visual feature [9]. However, the increasing availability of powerful hand tracking software has enabled the implementation of touch-based interfaces[3] [10,13]; these natural interfaces are both convenient and facilitate natural exploration of a tactile graphic or 3D model, which is often done using many fingers on one or both hands.

Our research continues the development of natural interfaces, motivated by our experience with the CamIO system [4], in which BVI study participants were enthusiastic about the ability to access audio labels but indicated they would prefer not to have to hold a stylus. Leveraging the availability of high-quality depth sensors in high-end iOS devices, we combine depth and color image data to create a novel "Point-and-Tap" interface that allows BVI users to access detailed information about tactile graphics and 3D models that is organized hierarchically in multiple levels. The benefit of the depth sensor is that it allows our interface to sense when the user is touching the surface of a tactile model as opposed to hovering above it; this determination is key to the interface because it enables the ability to detect touch events. We note that this interface is a significant advance on an earlier interface devised for CamIO [3], in which the user holds the tip of a hand-held stylus to the surface of a feature on a 3D map to hear a basic audio label about the feature, and hovers the stylus tip roughly 15 cm above the feature to hear higher-level information about the feature. By contrast, the "Point-and-Tap" interface is more natural and intuitive, and also supports easy access to several levels of information.

[1] https://www.touchgraphics.com/store/t3-t3-books.
[2] https://www.tactonom.com/en/.
[3] https://tactileimages.org/en/reader-app/.

In experiments with BVI participants we demonstrate that our approach is practical, easy to learn and effective. Finally, while we acknowledge that the advanced depth-sensing iOS devices that enable our approach are expensive, many BVI users enjoy the convenience of a multi-function smartphone and often use smartphone apps to accomplish a variety of daily activities [7], which is arguably more economical and practical than having to purchase multiple devices, each for a specific task.

2 The Point-and-Tap System

Our approach (see[4] for a video demonstration) uses a high-end iOS device with depth sensors (in our case, the iPhone 14 Pro) rigidly mounted above a tactile graphic (Fig. 2) so that the camera and depth sensors capture the entire graphic, as well as some space around the graphic. The space around the graphic is needed to facilitate hand tracking, which is performed using the MediaPipe Hands library[5], and which requires a clear view of the entire hand; note that part of the hand may rest outside the graphic when it points to a feature of interest. An iOS app was written that tracks the user's fingers, and recognizes a pointing gesture (Fig. 2) when it occurs, which is defined as a gesture in which the index finger of a specific hand (left or right, as specified by the user) is pointed straight while the other fingers are bent or closed.

The iOS app passes image frames from the iPhone's color camera stream to MediaPipe for hand tracking, and simultaneously acquires depth images from the iPhone's TrueDepth sensor[6], which specifies the depth estimated across a lower-resolution image having the same field of view as the color (RGB) image. When the app is first launched, reference RGB and depth images are acquired of the scene, which includes the tactile graphic without the user's hands being present. A simple color segmentation algorithm is performed on the reference RGB image to determine which pixels in the image belong to which features of interest (e.g., each region in the British Isles map shown in Fig. 2); for the purposes of this algorithm, we used five colors of cardstock to create the tactile graphics for our experiments. Each feature on a tactile graphic has a unique color, and it is associated with five text labels: the first is the name of the feature, and the four remaining labels specify additional information about the feature. To create a new tactile graphic model, the five colors of cardstock are cut out and pasted to a sheet of white cardstock, and the desired text labels for the entire model are entered as text strings in the app. (See Sect. 5 for a discussion of future work to eliminate our dependence on color cues.)

When a pointing gesture is recognized, the reference depth image is used to determine whether the pointing fingertip is touching the graphic or is above it. The first time the user points to a feature on the tactile graphic, the first text

[4] https://www.ski.org/projects/camio-hands.
[5] https://developers.google.com/mediapipe/solutions/vision/hand_landmarker.
[6] https://developer.apple.com/documentation/avfoundation/additional_data_capture/streaming_depth_data_from_the_truedepth_camera.

label corresponding to that feature is read aloud using text-to-speech (TTS). Each time the fingertip is raised up above a certain height threshold and then lowered again to touch the feature, the text label at the next level is read aloud. (Repeated taps cycle the levels from one through five and then back again to one.) An utterance in progress is halted whenever the fingertip is lifted off the feature, and this behavior allows the user to cycle rapidly through multiple levels to arrive quickly at the level they desire. (This way of eliciting multiple levels of feedback from repeated finger taps was inspired by the user interface in the T3 Talking Tablet.)

To minimize the impact of noise in the depth sensor measurements (see Fig. 1) and in MediaPipe's finger tracking, we devised an algorithm that integrates information over a short time interval. In this algorithm, we kept track of the difference between the reference depth image and the current depth image at the location of the pointing fingertip (MediaPipe's hand landmark L8) over time. This sequence of depth differences is stored in a ring buffer (first-in-first-out) containing values for twenty consecutive camera frames (spanning under half a second of video images); a "touch" event is declared whenever three conditions are satisfied: (a) the average depth difference in the buffer is less than 2 cm; (b) the current depth difference is less than 1.5 cm; and (c) the current depth difference at the location of MediaPipe's hand landmark L5 (the knuckle of the index finger) must be less than 12 cm. The third condition rules out instances when the user's hand is moving rapidly, causing the index finger to disappear from the depth map (even though MediaPipe is still tracking it accurately), thereby making the apparent depth difference at the pointing fingertip close to zero even when the fingertip is nowhere near the surface of the tactile graphic; fortunately, under these conditions, the depth map value at L5 is likely to be accurate since the knuckle is part of the main hand structure (not just the narrow fingertip) and remains visible in the depth map.

Since TrueDepth is intended primarily for face recognition, with a recommended distance from the target of roughly 25–50 cm[7], in our setup we mount the iPhone 14 Pro about 50 cm above the tactile graphic. The limited field of view of the camera imposes a constraint on how large the tactile graphic can be and still be entirely captured in the image (Fig. 2); in the future we will explore the possibility of using iOS LiDAR to accommodate larger models. (While LiDAR has a maximum range of about 5 m[8], empirically we have found that the minimum range of LiDAR is roughly 70 cm, and we note that it may be challenging to resolve the fingertip clearly in depth at longer distances.)

3 User Studies

We iteratively tested and refined our Point and Tap interface in a series of pilot experiments, the first with a sighted participant and the next with two BVI

[7] https://support.apple.com/en-us/102381.

[8] https://www.apple.com/newsroom/2020/03/apple-unveils-new-ipad-pro-with-lidar-scanner-and-trackpad-support-in-ipados/.

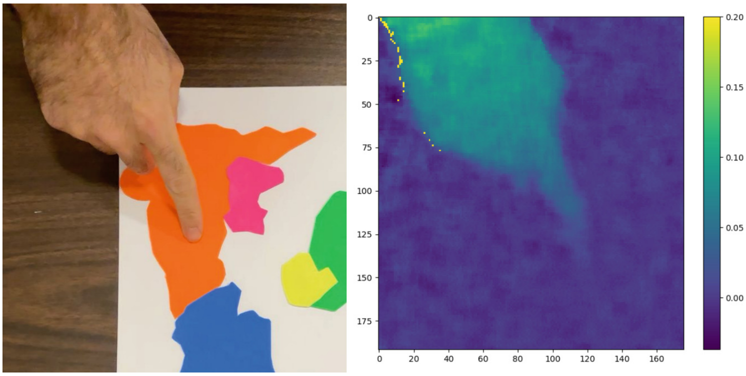


Fig. 1. Left: color image. Right: depth map difference from TrueDepth camera. Color indicates depth differences in meters. The depth difference declines from higher values on the knuckle of the pointing finger to a value at or near zero at the fingertip, but depth noise is clearly visible.

participants. We used the feedback from these pilot experiments to iteratively debug and improve our interface. Notably, we relaxed our definition of the pointing gesture so that it admitted a pointing gesture in which the non-index fingers are not curled as tightly as shown in Fig. 2. The interface and experimental protocol were fixed after these pilot experiments. We then conducted formal experiments with six more BVI participants, P1-P6, ranging in age from 36 to 74 years old (3 female/2 male); we excluded a seventh BVI participant from our analysis because of inadvertent deviations from the experimental protocol. All experiments were conducted under an approved IRB protocol.

Fig. 2. Left: Experimental setup shows iPhone 14 Pro mounted above the British Isles tactile map in cellphone holder with the user making a pointing gesture on the map. Right: close-up of pointing gesture. The five levels of text labels for the region being pointed to are: "Republic of Ireland"; "2: Capital is Dublin"; "3: Population 4.9 million", "4: Area 27133 square miles", "5: Limerick is another city in Ireland".

Each formal experiment included the following: (1) A demographic survey. (2) A 15-minute training session in which the experimenter demonstrated how to use the interface on a simple tactile graphic containing four features (a circle, square, triangle and diamond), and helped the participant learn to use the system. (3) Experiment E1, in which the participant was presented with a tactile map of the sun with the inner planets (Mercury, Venus, Earth and Mars) and was asked to access specific information (e.g., "Mars, level 2") in 50 trials. For each trial, the experimenter gave two scores for each of the participant's responses to evaluate the following: (a) E1a: Did the system do what the user wanted it to do? (b) E1b: Was the user able to elicit the correct information by the end of the trial? (4) Experiment E2, in which the participant was asked to access information that answers high-level questions about two additional tactile models, one containing five rockets and the other a map of the five regions of the British Isles (England, Scotland, Ireland, Northern Ireland and Wales). (5) A System Usability Survey (SUS) [2]. (6) Both specific and open-ended feedback about the system.

4 Experimental Results and Discussion

We report the results of the study in Table 1, with E1a and E1b scores reported as percentages out of 50 trials. We note an important caveat for the data for participant P6: during Experiment E2, despite good performance on the five rockets model, shortly after beginning to work with the British Isles model, she was unable to elicit any feedback from the system because she forgot to make the correct pointing gesture (Fig. 2), unsuccessfully attempting to point while holding her non-index fingers nearly straight. After witnessing this problem for several minutes, the experimenters intervened by reminding her how to make the pointing gesture. Her performance returned to normal after this reminder.

The E1 scores demonstrate that, while the system sometimes (the average E1a score was 86.3%) misinterpreted participants' tap gestures (most often because the user failed to lift their fingertip up high enough above the model before tapping it again, or because they failed to make the correct pointing gesture), in all but one trial for P5 the participants were nevertheless able to get the desired information. The SUS scores demonstrate the usability of the system, with an overall average score of 88.3.

The results of experiment E2 provide additional evidence of the effectiveness of the interface: participants were able to answer all the questions correctly. Positive feedback about the interface included its simplicity and usefulness for acquiring detailed information about a model; some negative feedback was reported about the incidence of errors when the interface failed to recognize a tap gesture. Some participants suggested that the interface should be expanded to recognize additional gestures, such as left/right swipe gestures to jump down/up several levels of audio labels.

Overall, the experiments with BVI participants demonstrate that the approach is practical, easy to learn and effective. Our experiences with the experiments highlight two main areas of improvement. First, the training session should have included a component in which the participants were asked to practice more Point-and-Tap tasks, with the experimenters making specific recommendations based on their performance. Second, the system could have provided periodic (e.g., no more than once every several seconds) audio feedback whenever the pointing fingertip was stationary but no pointing gesture was recognized, which would remind the user to make the pointing gesture if they want feedback.

Table 1. Data for all participants in formal experiments. (See the text for an important caveat about participant P6.) The choices for self-reported Experience scales included None, Low, Medium, High.

P#	Age	Sex	Perception of		Experience with		E1a	E1b	SUS
			light	form	tactile graphics	touchscreens			
P1	46	F	N	N	High	High	82%	100%	75
P2	43	M	N	N	Medium	High	92%	100%	97.5
P3	74	F	Y	Y	Low	Medium	86%	98%	92.5
P4	74	F	N	N	Low	High	76%	100%	77.5
P5	36	M	Y	N	Medium	High	98%	100%	90
P6	62	F	N	N	Low	Medium	84%	100%	97.5

5 Conclusion

We have devised a novel "Point-and-Tap" interface that allows BVI users to acquire detailed information about tactile graphics and other models. User studies with six BVI participants demonstrate that the interface is practical, easy to learn and effective.

We have recently improved the system so that it recognizes the pointing gesture with either the left or right index finger, so the user doesn't need to specify which hand they are pointing with; an audio alert is sounded if the system detects the pointing gesture in both hands. We plan to release the system as a free app on the Apple App Store which will allow non-developers to create and label their own models using colored cardstock. Soon we will add 3D geometric reasoning to our system as used in CamIO, which will eliminate the need for color segmentation, and will allow the app to support both 2D and 3D models, including existing tactile graphics and 3D models. Finally, in the future we will explore methods such as those used in [6] to allow our system to more precisely distinguish between fingertip touch and hover events.

Acknowledgments. HS and JMC were supported by NEI/NIH grant number 2R01EY025332 and NIDILRR RERC grant number 90REGE0018.

Disclosure of Interests. The authors have no competing interests.

References

1. Brock, A.M., Truillet, P., Oriola, B., Picard, D., Jouffrais, C.: Interactivity improves usability of geographic maps for visually impaired people. Hum. Comput. Interact. **30**, 156–194 (2015)
2. Brooke, J., et al.: SUS-a quick and dirty usability scale. Usability Eval. Indu. **189**, 4–7 (1996)
3. Coughlan, J.M., Biggs, B., Shen, H.: Non-visual access to an interactive 3D map. In: Miesenberger, K., Kouroupetroglou, G., Mavrou, K., Manduchi, R., Covarrubias Rodriguez, M., Penáz, P. (eds.) ICCHP-AAATE 2022. LNCS, vol. 13341, pp. 253–260. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-08648-9_29
4. Coughlan, J.M., Shen, H., Biggs, B.: Towards accessible audio labeling of 3d objects. In: Journal on technology and persons with disabilities. In: Annual International Technology and Persons with Disabilities Conference. CSUN (2020)
5. Fusco, G., Morash, V.S.: The tactile graphics helper: providing audio clarification for tactile graphics using machine vision. In: International ACM SIGACCESS Conference on Computers & Accessibility. ACM (2015)
6. Goussies, N.A., et al.: Learning to detect touches on cluttered tables. arXiv preprint arXiv:2304.04687 (2023)
7. Griffin-Shirley, N., et al.: A survey on the use of mobile applications for people who are visually impaired. J. Vis. Impairment Blindness **111**, 307–323 (2017)
8. Shen, H., Edwards, O., Miele, J., Coughlan, J.M.: Camio: a 3D computer vision system enabling audio/haptic interaction with physical objects by blind users. In: International ACM SIGACCESS Conference on Computers & Accessibility. ACM (2013)
9. Shi, L., Zhao, Y., Azenkot, S.: Markit and talkit: a low-barrier toolkit to augment 3D printed models with audio annotations. In: Annual ACM Symposium on User Interface Software and Technology. ACM (2017)
10. Wang, X., Kayukawa, S., Takagi, H., Asakawa, C.: Touchpilot: designing a guidance system that assists blind people in learning complex 3d structures. In: International ACM SIGACCESS Conference on Computers and Accessibility. ACM (2023)
11. Wiazowski, J.: Can braille be revived? a possible impact of high-end braille and mainstream technology on the revival of tactile literacy medium. Assistive Technol. **26**, 227–230 (2014)
12. Zebehazy, K.T., Wilton, A.P.: Straight from the source: perceptions of students with visual impairments about graphic use. J. Vis. Impairment Blindness **108**, 275–286 (2014)
13. Zeinullin, M., Hersh, M.: Tactile audio responsive intelligent system. IEEE Access **10**, 122074–122091 (2022)

Automatic Generation of Tactile Graphics of Characters to Support Handwriting of Blind Individuals

Shunnosuke Sonobe[1], Hitoshi Tanaka[2], and Akio Fujiyoshi[1(✉)]

[1] Ibaraki University, Hitachi 316-8511, Japan
{24nm735r,akio.fujiyoshi.cs}@vc.ibaraki.ac.jp
[2] Tsukuba University of Technology, Tsukuba 305-8521, Japan
htanaka@k.tsukuba-tech.ac.jp

Abstract. For blind individuals who wish to live independently, this study develops an automatic generation tool for tactile graphics of characters. The tool is developed by using the tactile graphics production system BPLOT. The tool generates character images from fonts installed on Windows, extracts the contours of these images, and outputs a figure-drawing program for BPLOT from the coordinates of those contours. In addition to choosing a typeface of characters, character spacing and arrangement, users can also select whether to apply the thinning processing to character images. The evaluation of the tactile graphics of characters was conducted with blind university students as experimental participants. All participants rated the material with Kaisho (regular script) typeface subjected to the thinning processing as having the highest readability.

Keywords: Tactile graphics · Automatic generation · Blind · Handwriting

1 Introduction

Handwriting is still considered as an important human activity. People enjoy calligraphy in many countries, and even among the digital generation, the culture of exchanging handwritten letters continues. Naturally, among the blind, especially those who have become blind later in life, there are those who have the desire to write characters. In Japan, however, the variety of characters is very large, with 92 types of syllabic characters known as “Hiragana” and “Katakana”, in addition to 2,136 types of ideographic characters known as “Joyo Kanji,” which were adopted from China. If a sighted supporter is present, it is possible to convey the shape of a character by guiding the forefinger of a blind person to write on his or her palm. This method is often used for communication with individuals who are both blind and deaf. Until research on writing navigation systems with HCI (Human-Computer Interaction) devices [1,2] advances, leading to widespread utilization and making the devices affordable, it is difficult for blind individuals

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 260–266, 2024.
https://doi.org/10.1007/978-3-031-62846-7_31

to verify the shapes of characters without assistance of a sighted person. Therefore, tools for the confirmation of shapes of characters through tactile graphics has been sought.

For blind individuals who wish to live independently, this study develops an automatic generation tool for tactile graphics of characters. The project began with a phone call from an elderly person who is totally blind. The elderly person had inquired with many publishers of braille materials and manufacturers of assistive technology devices, and it was a braille printer manufacturer where he heard the name and phone number of one of the authors.

The tool is developed by using the tactile graphics production system BPLOT [3–6]. The tool generates character images from fonts installed on Windows, extracts the contours of these images, and outputs a figure-drawing program for BPLOT from the coordinates of those contours. Then, through BPLOT, the final tactile graphics are produced from a braille plotter printer.

Users of the tool can choose a typeface of tactile graphics of characters from the four choices: Mincho (serif font), Gothic (sans-serif font), Kaisho (regular script font). To improve the quality of final tactile graphics, appropriate image processing is applied to character images before extracting the contours, depending on a chosen typeface. In addition to choosing character spacing and arrangement, users can also select whether to apply the thinning processing to character images. For the development of automatic generation of tactile graphics of characters, we referred to the methods used for the automatic generation of tactile maps [7].

The evaluation of the tactile graphics of characters was conducted with four blind university students as experimental participants at Tsukuba University of Technology. As experimental materials, the following four types of tactile graphics are prepared: (1) Mincho typeface, (2) Gothic typeface, (3) Kaisho typeface, (4) Kaisho typeface subjected to the thinning processing. As a result, though the material (1) got the lowest score, all tactile graphics were evaluated as being legible. All participants rated the material (4) as having the highest readability. After the experiment, one participant, who is congenitally blind, expressed gratitude for being able to notice for the first time today that there are such differences in the typefaces of characters.

The authors received a message from the elderly person, stating he was very moved upon touching the tactile graphics of characters. Visiting the elderly person's home to install the tool on his computer and teach him how to use the tool is a challenge for the future.

2 Automatic Generation Tool for Tactile Graphics of Characters

The automatic generation tool for tactile graphics of characters is an application software running on Windows, written in the C# language. It was developed by using the tactile graphics production system BPLOT [3–6].

Fig. 1. The user interface of the tool

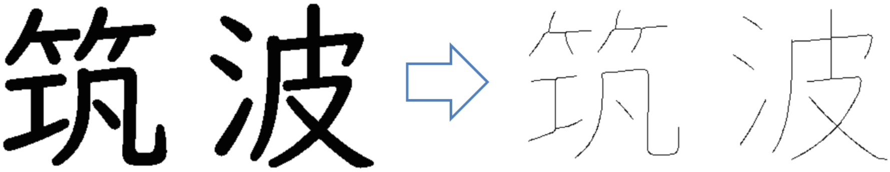


Fig. 2. The thinning processing to a character image

2.1 User Interface

Figure 1 shows the user interface of the tool. Users of the tool can choose a typeface of tactile graphics of characters from the four choices: Mincho (serif font), Gothic (sans-serif font), Kaisho (regular script font). To improve the quality of final tactile graphics, appropriate image processing is applied to character images before extracting the contours, depending on a chosen typeface. In addition to choosing character spacing and arrangement, users can also select whether to apply the thinning processing to character images. An example of the thinning processing to a character image is presented in Fig. 2. Furthermore, the user interface allows users to preview the final tactile graphics.

2.2 Outline of Generation Process

The outline of the generation process of tactile graphics of characters is illustrated in Fig. 3. First, the tool generates a character image from fonts installed on Windows. Users can choose their preferred typeface from the following options: Mincho (serif font), Gothic (sans-serif font), Kaisho (regular script font). Users can also select whether to apply the thinning processing to the character image.

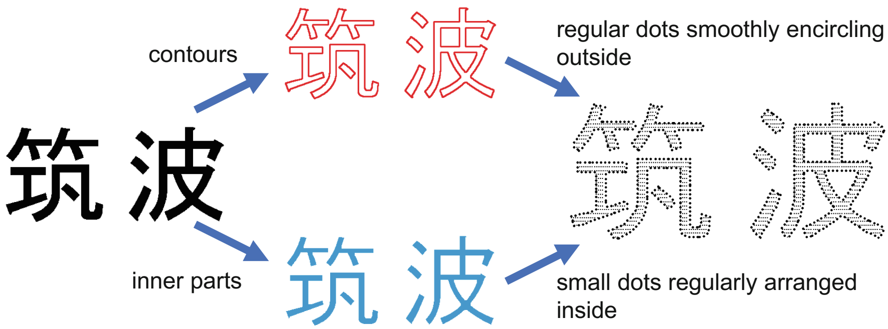


Fig. 3. Outline of the generation process of tactile graphics of characters

Next, the tool extracts the contours of the character image and converts them into a sequence of regular dots smoothly encircling the characters. Simultaneously, the inner parts of the character image are converted into regularly arranged small dots. Finally, the tool outputs a figure-drawing program for BPLOT from the converted dots for tactile graphics, and then, through BPLOT, the final tactile graphics are produced from a braille plotter printer.

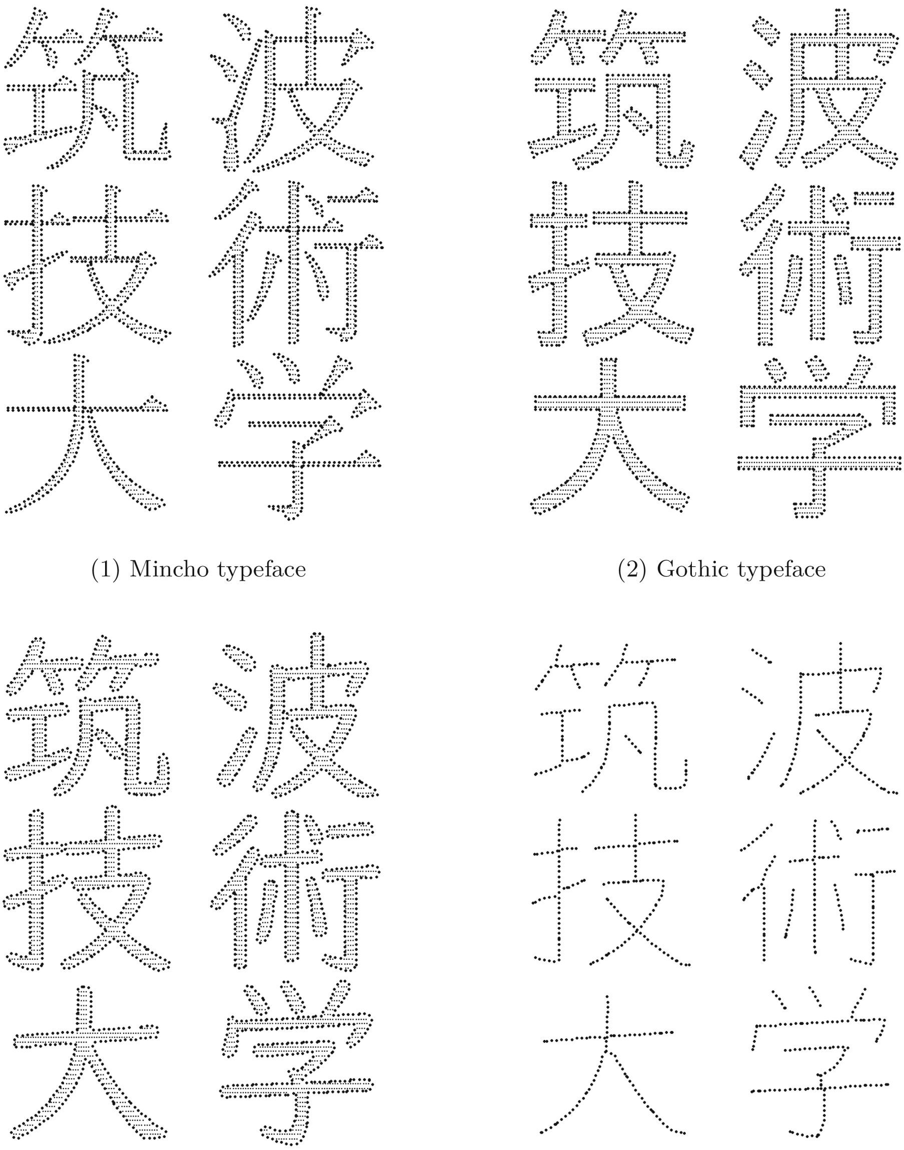
2.3 BPLOT

BPLOT is the first tactile graphics production system for the blind that enables the blind to produce tactile graphics by themselves. Existing tactile graphics design applications such as BES [8] and EDEL-plus [9] have sophisticated GUI (Graphical User Interface) and enable the sighted to produce tactile graphics easily. However, because of the necessity for mouse operations, the blind cannot use these applications by themselves. BPLOT produces tactile graphics from a source text file written in specially designed plotter control language for BPLOT. Because a source file for BPLOT is a text file, it is editable with any text editors by any person who has learned the plotter control language. Therefore, BPLOT enables the blind to produce tactile graphics by themselves.

The figure-drawing program output by the tool consists of a main program that defines the size and position of the characters, and subprograms that define the shape of each character. Since the main program can be easily edited with a text editor, a blind user easily modify the size and position of the characters. By adding some commands to the main program, it is possible to include braille descriptions alongside the tactile graphics.

3 Experimental Evaluation

The evaluation of the tactile graphics of characters was conducted with four blind university students as experimental participants at Tsukuba University of Technology.



(1) Mincho typeface

(2) Gothic typeface

(3) Kaisho typeface

(4) thinned Kaisho typeface

Fig. 4. The four types of tactile graphics of characters

Figure 4 shows the four types of tactile graphics prepared as experimental materials: (1) Mincho typeface, (2) Gothic typeface, (3) Kaisho typeface, (4) Kaisho typeface subjected to the thinning processing.

3.1 Result

The experimental participants were given ample time to read the tactile graphics of characters, and afterwards, they were asked to answer a questionnaire about the ease of reading their shapes using a 5-point Likert scale. The result of questionnaires is shown in Table 1.

Table 1. The ease of reading the shapes using a 5-point Likert scale

experimental material	part. A	part. B	part. C	part. D	average
(1) Mincho	2	4	3	3	3.0
(2) Gothic	3	3	4	4	3.5
(3) Kaisho	3	3	4	4	3.5
(4) thinned Kaisho	4	5	5	5	4.5

Though the material (1) got the lowest score, all tactile graphics were evaluated as being legible. All participants rated the material (4) as having the highest readability.

4 Conclusion and Future Works

This study develops an automatic generation tool for tactile graphics of characters. By refining the user interface to suit each user, it becomes possible for blind individuals to autonomously create tactile graphics of characters.

The recent advancements in generative AI are remarkable, fostering innovation and creativity across various sectors of society. Among the many generative AI applications, OpenAI's ChatGPT has attracted attention because it has ability to facilitate dialogue with users through its remarkable natural language processing capabilities. The situation has been progressed even further since, in September 2023, OpenAI announced that "ChatGPT can now see, hear, and speak" [10]. As future works, we want to enable blind individuals to read and understand charts, diagrams, and graphs through tactile and auditory means by taking advantage of ChatGPT.

References

1. Plimmer, B., Reid, P., Blagojevic, R., Crossan, A., Brewster, S.: Signing on the tactile line: a multimodal system for teaching handwriting to blind children. ACM Trans. Comput. Hum. Interact. (TOCHI) **18**(3), 1–29 (2011)
2. Liu, H., et al.: Calligraphy navigation system for blind people based on visual prosthesis on waist belt. In 2020 IEEE International Conference on Multimedia Expo Workshops (ICMEW), pp. 1–2. IEEE (2020)

3. Fujiyoshi, M., Fujiyoshi, A., Ohtake, N., Yamaguchi, K., Teshima, Y.: The development of a universal design tactile graphics production system BPLOT2. In: Miesenberger, K., Klaus, J., Zagler, W., Karshmer, A. (eds.) ICCHP 2008. LNCS, vol. 5105, pp. 938–945. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-70540-6_141
4. Fujiyoshi, M., et al.: Development of tactile graphics production software for three-dimensional projections. In: Miesenberger, K., Klaus, J., Zagler, W., Karshmer, A. (eds.) ICCHP 2010. LNCS, vol. 6180, pp. 541–547. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-14100-3_81
5. Fujiyoshi, M., Fujiyoshi, A., Osawa, A., Kuroda, Y., Sasaki, Y.: Development of synchronized CUI and GUI for universal design tactile graphics production system BPLOT3. In: Miesenberger, K., Fels, D., Archambault, D., Peňáz, P., Zagler, W. (eds.) ICCHP 2014. LNCS, vol. 8548, pp. 18–25. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-08599-9_4
6. Fujiyoshi, M., Fujiyoshi, A., Tanaka, H., Ishida, T.: Universal design tactile graphics production system BPLOT4 for blind teachers and blind staffs to produce tactile graphics and ink print graphics of high quality. In: Miesenberger, K., Kouroupetroglou, G. (eds.) ICCHP 2018. LNCS, vol. 10897, pp. 167–176. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-94274-2_23
7. Watanabe, T., Yamaguchi, T., Koda, S., Minatani, K.: Tactile map automated creation system using OpenStreetMap. In: Miesenberger, K., Fels, D., Archambault, D., Peňáz, P., Zagler, W. (eds.) ICCHP 2014. LNCS, vol. 8548, pp. 42–49. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-08599-9_7
8. BES (in Japanese). http://www.ttools.co.jp/product/eyes/BES/
9. EDEL-plus (in Japanese). http://www7a.biglobe.ne.jp/~EDEL-plus/
10. OpenAI Blog, "ChatGPT can now see, hear, and speak". https://openai.com/blog/chatgpt-can-now-see-hear-and-speak

Exploring Space: User Interfaces for Blind and Visually Impaired People for Spatial and Non-verbal Information

Reinhard Koutny(✉)

Institut Integriert Studieren, Johannes Kepler University Linz, Altenbergerstraße 69, 4040 Linz, Austria
reinhard.koutny@jku.at

Abstract. Meetings play an important role in today's work environment. Unfortunately, blind and visually impaired people often encounter difficulties in participating fully and equally. The main reasons are twofold: Firstly, visual aids such as whiteboards, flipcharts, or projectors are often used, which present information in a 2D format. Secondly, nonverbal communication plays an essential role, with frequent use of deictic gestures, like pointing gestures, incorporating spatial information into the conversation. These factors lead to a disadvantage for blind and visually impaired individuals in the workplace. As part of the research project MAPVI [5], an accessible brainstorming tool for meetings has been developed, capable of storing nonverbal and spatial information, which can be connected to various devices such as computers, smartphones, and smartwatches. This provides a solid foundation for successfully addressing these issues through innovative user interaction concepts, making spatial information accessible and understandable for blind and visually impaired individuals. Prototypes of these user interface concepts have been developed and tested with the target group in a staged and iterative manner, with implementations running in the browser, on a smartphone, and on a smartwatch. This paper will outline the development and testing procedures, as well as the corresponding test results.

Keywords: Blind and Visually Impaired People · User Centered Design and User Participation · Perception of Spatial Information · Accessibility · Usability

1 Introduction

One cornerstone of a just society are equal job opportunities. Depending on the type of job, this requires accessibility in many different areas: physical access, wayfinding and many more. In the context of office work, business meetings are particularly critical. One special case is brainstorming meetings; a scenario selected for this work due to the manifold nature of social interaction, including non-verbal communication, and various sources of information. Tool support in this area has shifted over the last years from analog means, like flipcharts, to digital ones including mind map tools, group decision making tools or brainstorming tools. This creates the possibility of making

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 267–274, 2024.
https://doi.org/10.1007/978-3-031-62846-7_32

this type of content fully accessible to a broader range of participants, including blind and visually impaired people (BVIP). However, one essential part remains inaccessible; spatial aspects of this context often in combination with non-verbal information. During conversations, people heavily make use of gestures, postures and facial expressions to give necessary context to other participants. Meeting rooms are usually equipped with whiteboards and screens, and people refer to these objects or even pieces of information on them with deictic gestures like pointing gestures. For visually impaired and especially for blind people, this information is not accessible, excluding them from a major part of the conversation. Additionally, information is generated in parallel on multiple levels, resulting in a multi-dimensional information space. In contrast, common techniques of BVIPs to access digital information are all linear; for example, audio output via screenreaders or refreshable braille displays. Two different information spaces especially important to BVIP in our scenario have been identified together with the target group in interviews and questionnaires: Spatial information on whiteboards, where the position and spatial arrangement of notes hold a certain meaning depending on the context; and spatial information of the meeting room including objects and participants including their mimic and gestures.

2 State of the Art

For decades, researchers have developed approaches to solve or at least mitigate the above-mentioned issues. Earlier approaches used mice and keyboards, some even optically tracked pens [3] and PDAs [10] to feed information to the support system and share it on an electronic whiteboard. Researchers have recently developed an approach facilitating a back-projected interactive table and interactive vertical screens. A more comprehensive overview is provided here: [14, 16]. Still, visual aids of business meetings are only one part of the equation. Spatial aspects of the meeting and the venue, including the position of physical objects and participants offer a much greater challenge to present and give important context to a conversation understandably. Research efforts have been concerned with world exploration techniques for BVIP [1, 4, 6, 12, 13, 15]. One issue that has been repeatedly identified is that the auditory channel of a BVIP can be overloaded quickly and therefore, it is of utmost importance for a user interface (UI) to limit the amount of information transmitted via this channel. Otherwise, the user ends up choosing to either operating the user interface or listening to the conversation of the business meeting. In this context, this is barely an option. Consequently, other approaches were explored. One direction is the adaptation of regular braille displays, which usually display one-dimensional text in a haptic manner. Hyperbraille offers in contrast a two-dimensional array of braille elements allowing to display 2D content. It was facilitated in this context [11] and even extended with audio notifications by adopting tangible interaction concepts [9]. This device, however, comes with some drawbacks - most notably the relatively low resolution, the size and weight, and the high price, which lead to the low adoption rate of the community.

Therefore, a complete UI concept was developed [8], UIs based on it and functional prototypes only using off-the-shelf hardware co-developed and tested with the target group, tackling these shortcomings and allowing BVIPs to access spatial, visual and non-verbal information of business meetings. Note that only the user interface concept is topic of this paper and not the retrieval of location data, gestures or mimic of persons.

3 User Interface Concepts

Parts of the outlined information of business meetings can be presented textually for BVIP accessibility. However, conveying implicit information from entity arrangement in the two spaces, the meeting room and the digital 2D space of whiteboards and similar aids remains a challenge. Hence, this approach extends traditional UI concepts. Three UIs have been developed to grant meeting participants, including BVIP, access to all meeting information and enable equal contribution.

- **Web Interface:** This UI displays a visual whiteboard for sighted users while providing BVIP access to all textual information. Additionally, it offers access to real-world meeting aspects like participant gestures and room objects. It complies with WCAG, ensuring full accessibility for BVIP and other users. The interface is responsive and usable on tablets, smartphones, or screens with high zoom levels.
- **Smartphone:** While the web interface is fully accessible, certain spatial information remains challenging for BVIP. Hence, a separate UI uses off-the-shelf smartphones to create a haptic virtual reality representation of the whiteboard. Users can explore and manipulate notes intuitively, elevating BVIP to active contributors.
- **Smartwatch:** Another UI utilizes off-the-shelf smartwatches, enabling BVIP to explore the meeting room. By pointing and swiping the hand across the room, users receive vibration feedback indicating the presence of persons or objects. Gestures can be used to retrieve information about these objects. Searching for specific items triggers a mode switch, aiding navigation toward the target facilitating vibrators in the smartwatch (in addition to sound); if the user is pointing with the finger closer toward the right direction, the vibration gets stronger. It's also feasible to obtain a summary of the room's contents, including people or objects, and receive approximated directions using clock bearings: 12 o'clock = directly ahead, 1 slightly to the right, 3 o'clock = directly to the right, 6 o'clock means directly behind.

4 Methodology

The development and user testing process for three user interface prototypes corresponding to three user interface concepts involved iterative testing with the target user group using various devices such as browsers, smartphones, and smartwatches. The methodology included three main phases:

- **User Interface Prototype Development through Peer Research:** The initial prototypes of the user interfaces were tested and refined with the help of two peer researchers. This phase involved testing and improving the web interface, smartphone app, and smartwatch app to enhance accessibility for blind users.
- **Iterative User Studies on Early Prototypes in Small Groups:** The refined prototypes underwent iterative testing with a small group (3–5 individuals) of BVIP. The objective was to further improve usability and accessibility while avoiding over-optimization for specific users. Feedback from each study iteration was used to refine the interfaces for subsequent tests.

- **User Study in a Larger Group:** The final phase involved conducting a comprehensive user study with a larger group of end-users to gather quantitative and qualitative insights into the effectiveness of the interfaces. Peer researchers were involved in supporting these studies. The user test consisted of 9 tests, divided into 2 sessions per participant, with 3 tests conducted per user interface. Each participant completed both SUS and NASA-TLX questionnaires for each interface for insights in the subjective user satisfaction.

  – System Usability Scale (SUS) scores: SUS questionnaires are commonly employed across industries to measure satisfaction with system usability.
  – NASA-TLX scores: As an addition to SUS scores, which evaluates one user interface as whole, NASA-TLX scores are employed to assess and compare the perceived workload of the three specific tasks per user interface. User were asked to perform the following tasks:

    o Web interface:

      - Create a new note (and explore the user interface by doing so)
      - Find a note and retrieve its additional information
      - Create a group of notes

    o Smartwatch interface (persons and objects within the environment were simulated for the purpose of the user test):

      - Explore the meeting room and tell what/who is in the room
      - Retrieve a summary of what is in the room, including clock bearings
      - Find a specific person with sound and vibration

    o Smartphone interface:

      - Initialize the virtual whiteboard and explore notes on the whiteboard
      - Retrieve additional information of a note
      - Edit the position of a note on the virtual whiteboard

Users also had the opportunity to provide additional remarks on their preferences, dislikes and others. Demographically, the user study included 10 users, comprising of 6 female and 4 male participants, with 5 being blind and 5 being severely visually impaired. The age range of participants spanned from 18 to 67 years old, with an average age of 47.4 years.

5 User Involvement and Test Results

The test results showed significant improvements in both usability and accessibility of the user interfaces throughout the development and user involvement. Initially, the first prototypes, developed with support of peer researchers, were met with some criticism due to their clunkiness and lack of intuitiveness (e.g. misleading labels, bloated navigation). Additionally, certain features were found to be missing, as users envisioned using them in slightly different ways than the peer researchers (e.g. perceivable boundaries of the virtual whiteboard). However, these issues were largely resolved by the second stage of development involving groups (3–5) of other BVIPs. Numerous smaller and bigger usability improvements have been implemented, including the addition of haptic and acoustic borders to enhance interaction of smartphone whiteboard application. For the smartwatch app, an alternative to hand-flick gestures, which some users reported to be unreliable for them, has been introduced by incorporating dwelling functionality. Furthermore, the smartwatch application has been ported to work on smartphones as well, acknowledging diverse user preferences regarding wearable devices.

Subsequent quantitative and qualitative user testing with a larger group (10) demonstrated that the prototypes achieved good levels of accessibility and usability. Participants were asked to fill out both SUS and NASA-TLX questionnaires during the tests.

The System Usability Scale (SUS) is a widely used questionnaire-based tool for assessing the perceived usability of a system or interface [2]. It consists of 10 questions designed to determine the user's subjective experience, with five options ranging from Strongly Agree" (numerical value: 1) to "Strongly Disagree" (numerical value: 5), which was slightly adapted to match the Austrian grading system for easier understandability. The SUS score is calculated by tallying scores from each item, with even items scored by subtracting 1 from their numerical value and odd items scored by subtracting their value from 5. The total score is then multiplied by 2.5 to yield the participant's SUS score. Scores between 60 and 70 are deemed average, with 100 representing the ideal score. Figure 1 shows the combined scores of all participants for each question (Q1 to Q10).

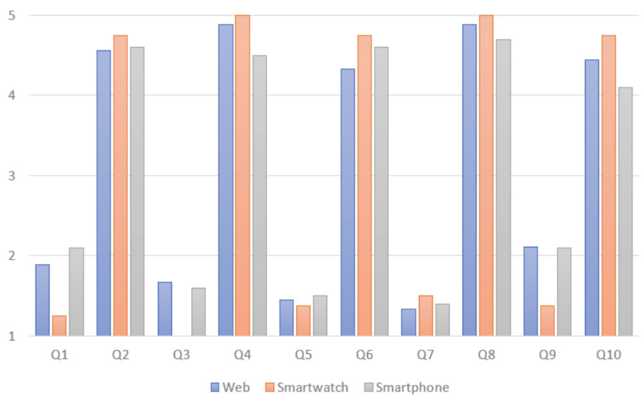


Fig. 1. SUS results of all three user interface concepts per question

The calculated SUS scores for the three interfaces yielded highly satisfactory results:

- Web interface: 87
- Smartwatch interface: 94
- Smartphone interface: 85

In addition to the SUS score which looked at whole user interfaces, NASA-TLX questions were completed by users for each of three tasks for each of the three user interfaces. The NASA Task Load Index (NASA-TLX) is a comprehensive scale used to assess workload during or after task completion. It encompasses six sub-scales, including Mental, Physical, and Temporal Demands, Frustration, Effort, and Performance, to gauge the user's workload effectively. Widely adopted across various domains, NASA-TLX serves as a benchmark for evaluating the effectiveness of measures, theories, or models, and is recommended for use in diverse settings such as aircraft certification, operating rooms, and website design [7]. Note that for this work the Physical sub-scale was removed as the goal was to evaluate the user interface concept and not the hardware which was used. Figures 2, 3 and 4 show the NASA-TLX results of user interface.

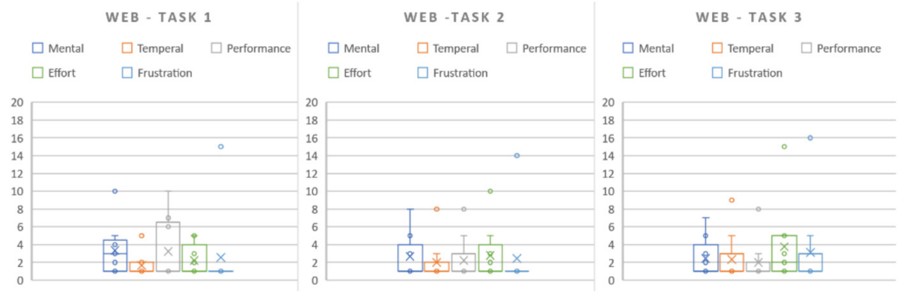


Fig. 2. NASA-TLX scores of web interface

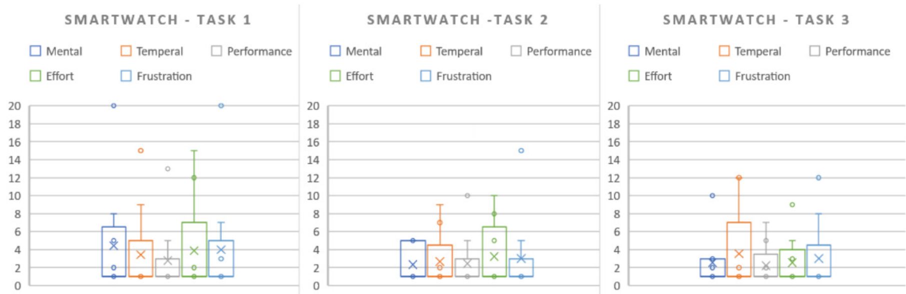


Fig. 3. NASA-TLX scores of smartwatch interface

NASA-TLX scores indicate that the perceived workload was manageable for the majority of participants across the tasks. Moreover, the analysis suggests that, for the tasks assigned, there was minimal disparity observed between the more traditional user interface concept found in the web interface and the two novel ones incorporating smartwatches and smartphones for spatial information. This is especially promising, considering participants' initial exposure to these concepts during the user tests.

In general, these findings, SUS and NASA-TLX scores, indicate that most users experienced a high satisfaction with the interfaces, and many independently suggested

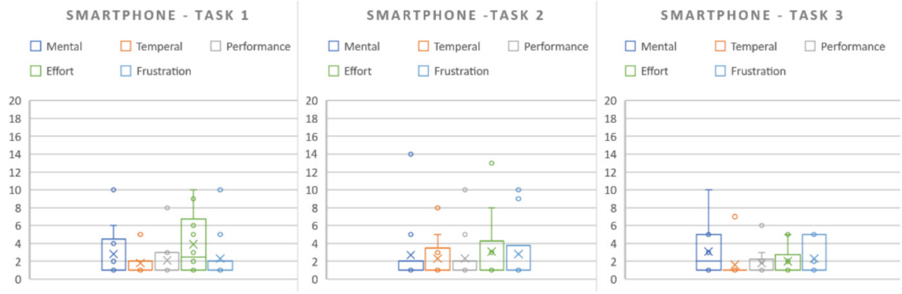


Fig. 4. NASA-TLX scores of smartphone interface

the application of the smartphone and smartwatch UIs in other contexts. Education and navigation were commonly cited as areas where these concepts could be beneficially applied or further developed.

6 Conclusion

This work addresses the challenges faced by blind and visually impaired people (BVIP) in participating fully in workplace meetings due to the reliance on visual aids and nonverbal communication. On the basis of an accessible brainstorming tool capable of storing and conveying spatial information across various devices, user interface (UI) prototypes based on novel UI concepts, including web interfaces, smartphones, and smartwatches have been created and tested together with the target group to level the playing field and promote equal participation. The iterative and inclusive development process, coupled with a staged testing procedure with the target group, ensured that the resulting UI concepts effectively address the needs of BVIP to promote equal opportunities in a professional context. User involvement and tests have shown great interest of the target group in the overall approach and users independently proposed the transfer of the smartphone and smartwatch UI to other contexts. Especially navigation and the educational domain with use cases like teaching geographic maps, the composition of molecules or trigonometry was mentioned. If adapted to these domains, not only brainstorming meetings or business meetings could be made more accessible, but also education, which ultimately affects a person's whole life.

Acknowledgments. This project (MAPVI) including this publication was funded by the Austrian Science Fund (FWF): I 3741-N31.

References

1. Bolt, R.A.: "Put-that-there": Voice and gesture at the graphics interface. ACM SIGGRAPH Comput. Graph. **14**(3), 262–270 (1980). https://doi.org/10.1145/965105.807503
2. Brooke, J.: SUS: a retrospective. J. Usability Stud. **8**, 29–40 (2013)

3. Elrod, S., et al.: Liveboard: a large interactive display supporting group meetings, presentations, and remote collaboration. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 599–607. Association for Computing Machinery, New York, NY, USA (1992). https://doi.org/10.1145/142750.143052
4. Geronazzo, M., et al.: Interactive spatial sonification for non-visual exploration of virtual maps. Int. J. Hum.-Comput. Stud. **85**, 4–15 (2016). https://doi.org/10.1016/j.ijhcs.2015.08.004
5. Gunther, S. et al.: MAPVI: meeting accessibility for persons with visual impairments. In: Proceedings of the 12th ACM International Conference on PErvasive Technologies Related to Assistive Environments, pp. 343–352. Association for Computing Machinery, New York, NY, USA (2019). https://doi.org/10.1145/3316782.3322747
6. Guo, A. et al.: VizLens: a robust and interactive screen reader for interfaces in the real world. In: Proceedings of the 29th Annual Symposium on User Interface Software and Technology, pp. 651–664. Association for Computing Machinery, New York, NY, USA (2016). https://doi.org/10.1145/2984511.2984518
7. Hart, S.G.: Nasa-task load index (NASA-TLX); 20 years later. Proc. Hum. Factors Ergon. Soc. Annu. Meet. **50**(9), 904–908 (2006). https://doi.org/10.1177/154193120605000909
8. Koutny, R., Miesenberger, K.: Accessible user interface concept for business meeting tool support including spatial and non-verbal information for blind and visually impaired people. In: Miesenberger, K. et al. (eds.) Computers Helping People with Special Needs, pp. 321–328. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-08648-9_37
9. Kunz, A. et al.: Accessibility of brainstorming sessions for blind people. In: Miesenberger, K. et al. (eds.) Computers Helping People with Special Needs, pp. 237–244. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-08596-8_38
10. Magerkurth, C., Prante, T.: Metaplan für die Westentasche: Mobile Computerunterstützung für Kreativitätssitzungen. In: Oberquelle, H. et al. (eds.) Mensch & Computer 2001: 1. Fachübergreifende Konferenz, pp. 163–171. Vieweg+Teubner Verlag, Wiesbaden (2001). https://doi.org/10.1007/978-3-322-80108-1_18
11. Pölzer, S., Miesenberger, K.: A tactile presentation method of mind maps in co-located meetings. In: Presented at the TacTT@ITS (2014)
12. Wacker, P., et al.: ARPen: mid-air object manipulation techniques for a bimanual AR system with pen & smartphone. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, pp. 1–12. Association for Computing Machinery, New York, NY, USA (2019). https://doi.org/10.1145/3290605.3300849
13. Willis, S., Helal, S.: RFID information grid for blind navigation and wayfinding. In: Ninth IEEE International Symposium on Wearable Computers (ISWC 2005), pp. 34–37 (2005). https://doi.org/10.1109/ISWC.2005.46
14. Designing user friendly augmented work environments. Comput. Support. Coop. Work (2010)
15. Supporting blind navigation using depth sensing and sonification. In: Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication, https://doi.org/10.1145/2494091.2494173, Accessed 10 Feb 2024
16. The design space of augmented and virtual reality applications for assistive environments in manufacturing. In: Proceedings of the 10th International Conference on PErvasive Technologies Related to Assistive Environments, https://doi.org/10.1145/3056540.3076193, Accessed 10 Feb 2024

RainbowTact: An Automatic Tactile Graphics Translation Technique that Brings the Full Spectrum of Color to the Visually Impaired

Hyun W. Ka(✉) and Rachel Kim

Korea Advanced Institute of Science and Technology, Daejeon 34141, Korea
hyun.ka@kaist.ac.kr

Abstract. Tactile graphics enable visually impaired people to access visual information via touch. Despite extensive guidelines and solutions, effectively representing color, vital for visual communication, remains a key challenge. This research aims to address this gap by developing RainbowTact, inspired by the correlation between visible colors and light wavelengths, to translate colors into intuitive tactile representations. RainbowTact distinguishes chromatic and achromatic colors using tactile wave and dot patterns depicting hues, saturations and brightness. Achromatic shades are shown by dot density and size. RainbowTact meets key design criteria: covering the full color space with omnidirectional, orientation-independent patterns aligning with standards. A software prototype automates conversion. A pilot study evaluated RainbowTact's effectiveness and usability through quantitative and qualitative analyses. Results showed decent color identification success and increasing task efficiency. Users strongly favored RainbowTact, highlighting benefits like pattern differentiation and non-directionality. While initial learning ease scored lower, participants expressed overall positive inclination. This demonstrates RainbowTact's potential to effectively convey color information via intuitive tactile representations to advance tactile graphics capabilities.

Keywords: Tactile Graphics · Color to Tactile · Sensory Substitution

1 Introduction

Tactile Graphics (TG) are essential tools for the visually impaired, providing a means to access visual information through touch. Many researchers and expert groups have proposed and studied various TG representation techniques and production methods [1–6]. Recognizing the need for standardized guidelines for TG to ensure consistency and quality in TG production, the Braille Authority of North America (BANA) compiled and released the Guidelines and Standards for Tactile Graphics [7], and it has served as the gold standard for TG. Despite their great contributions, the current guidelines and most of the existing solutions for TG have largely focused on shape and layout (e.g., sizing of tactile elements, symbols, braille labels), but do not adequately address the representation of color. As a result, the effective representation of color in TG remains a significant challenge. Color, as an important aspect of visual communication, is often lost

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 275–282, 2024.
https://doi.org/10.1007/978-3-031-62846-7_33

in translation to tactile formats, leaving a gap in the information conveyed. This limitation hinders the full potential of TG as an information tool for the visually impaired.

This research aims at closing the color representation gap in TG. We designed and implemented an alternative method, called RainbowTact, for translating color into intuitive and recognizable tactile representations. RainbowTact was inspired by the following two basic truths: first, that color, as perceived by the human eye, corresponds to specific wavelengths within the visible light spectrum; and second, that achromatic colors, which include black, white, and gray, do not correspond to specific wavelengths of light, but result from the absence, complete reflection, or varying intensities of all wavelengths of visible light.

2 State of the Art

In order to successfully convert visual information into TG, it is crucial to understand the tactile sense of the visually impaired. Boven and his colleagues revealed that blind individuals, particularly Braille readers, have enhanced tactile spatial resolution compared to sighted individuals [8]. Using a grating orientation task to measure spatial acuity, the study found that blind participants had a lower orientation threshold (1.04 mm) than sighted participants (1.46 mm), with the primary Braille-reading finger of blind subjects showing even more acute sensitivity (0.80 mm). This suggests that the tactile discrimination of blind individuals is more refined, allowing for the perception of finer details.

The research on converting color information into touchable forms for the blind has evolved over the years, aiming to provide the visually impaired with a means to perceive and appreciate color through tactile sensations. Rataj presented an algorithm that converts visual images into tactile forms by grouping similar colors and translating them into textured lines through image tracing [4]. These are then shaped into 3D patterns using modeling software. The method, though innovative, faces challenges in capturing subtle color nuances, relies on arbitrary tactile patterns, and the complexity of converting detailed images into simplified tactile representations.

Shin et al. developed a 3D tactile system enabling the visually impaired to discern colors through texture, aligning the HSV color model with tactile line patterns for art appreciation [9]. The system correlates hue with line orientation, saturation with width, and brightness with spacing. Although the approach is innovative, it excludes achromatic colors and doesn't account for directional independence, crucial for consistent tactile interpretation from any orientation, which could restrict its practicality.

Isumi et al. presented a method to represent colors in graphics through black-and-white hatching patterns [10], enhancing differentiation with line segments, zigzags, and waves. Although innovative, the method primarily addresses hue, neglecting the full color spectrum and the contrast between colored and grayscale tones. It also overlooks the directional independence of the patterns, which is crucial for universal tactile comprehension. Additionally, the complexity of zigzag and wave patterns may lead to misinterpretations since the study's emphasis was on developing low-cost black and white printing solutions rather than optimized tactile graphics for the visually impaired.

Cho and his colleagues introduced Tactile Color Pictograms (TCPs), which use raised geometric shapes to help visually impaired people identify colors through touch [11], based on the Munsell system for hue, intensity, and lightness representation. While TCPs offer a tactile means to distinguish 29 colors, including six hues, they face limitations due to the Munsell system's incomplete color range, the demanding learning curve for users, potential confusion from orientation-based interpretation, and production challenges posed by complex patterns.

The goal of this research is to develop RainbowTact, a different method of translating color into more intuitive tactile forms, to address the above issues and limitations. RainbowTact is based on two principles: first, that perceivable colors correlate with specific wavelengths in the visible spectrum; and second, that achromatic colors—black, white, and gray—come from the absence, full reflection, or mixed intensities of these wavelengths.

3 Methodology

3.1 Design Criteria and Implementation

To achieve our research goal, we have established the following design criteria (DC) and are proceeding with development accordingly.

- DC1: RainbowTact should distinguish between chromatic and achromatic colors. The color space is divided into two distinct subspaces based on the presence of hue: the chromatic color space and the achromatic color space. This criterion was met by using 2 different tactile patterns: waves of different wavelengths for the chromatic colors and dots of different density for the achromatic colors.
- DC2: RainbowTact should cover the entire chromatic color space and support the color attributes like hue, saturation, and brightness. To satisfy this, we have adopted the following strategy: Hues are indicated by varying line wavelengths, saturations by line widths, and brightness by the spacing between lines, ensuring clear identification and distinction of color attributes.
- DC3: RainbowTact should be omnidirectional. We have taken great care to ensure that the tactile patterns created can be interpreted regardless of their orientation, so that they can be understood from any direction. In addition, we made every effort to optimize the unit size of the lines and dots so that users can identify colors with minimal movement of their fingertips.
- DC4: RainbowTact should provide consistency and quality in the learning, communication, and production of TG. To meet this criterion, we considered compatibility with the existing guidelines [7]. The generated tactile patterns should adhere to the specifications outlined in it, such as line width, spacing, and texture differentiation.
- DC5: RainbowTact should be able to automate the color to TG conversion process. We engineered a three-step color-to-TG algorithm. First, image preprocessing with k-means clustering simplifies colors, grouping pixels by their HSV values and merging similar hues based on a set threshold to reduce complexity. Next, it segments the image into chromatic and achromatic regions. The algorithm then generates tactile patterns—using dot patterns for achromatic areas, where dot size and density adjust to grayscale intensity (larger, denser dots for darker areas, smaller, sparser dots for lighter ones), and wave patterns for chromatic regions, which are systematically varied

according to HSV attributes. Wavelength (λ) is mapped linearly from the hue value range (0 to 180) to correspond with specific hues, saturation is linearly mapped to wave amplitude (A), and brightness levels are depicted by line width—thicker lines for darker colors, in accordance with the BANA guidelines [7]. The following equation was used to generate the different wave patterns.

$$y = A\, sin\, sin\left(\frac{2\pi}{\lambda}x\right)$$

This method ensures that tactile patterns are omnidirectional and can be discerned from any orientation. Figure 1 shows an exemplary conceptual illustration of RainbowTact. Our prototype software, developed in Python, supports the design criteria above, producing outputs suitable for various tactile printing technologies, including thermal swell paper, thermoforming, Braille embossing, and 3D printing.

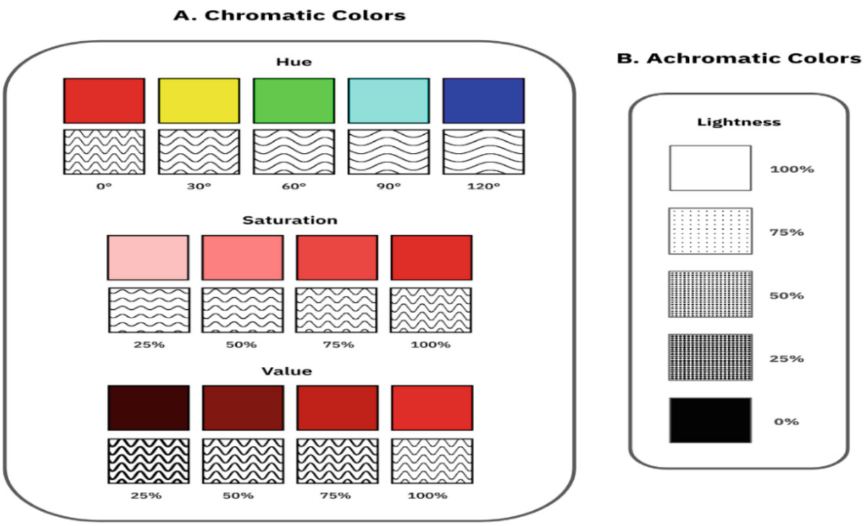


Fig. 1. Conceptual Illustrations of RainbowTact

3.2 Pilot Study

To evaluate the effectiveness and user satisfaction of RainbowTact, we conducted a small-scale pilot study. Three participants (1 female, 2 male) were recruited through a community of visually impaired people. All participants were blind and had Braille literacy and prior experience with TG. Before getting into the main test session, participants were introduced to RainbowTact and were given time (5 min) to familiarize themselves to the concepts and tactile patterns used like wavelength, line interval and width.

The main user test consisted of two sessions (quantitative and qualitative). In the quantitative session, each participant was given 3 tactile maps (Fig. 2, excluding the color images) with features like water area, parks, and buildings, using RainbowTact. They were asked to identify and count the number of color-coded features on each map. The task success rate and completion time were measured.

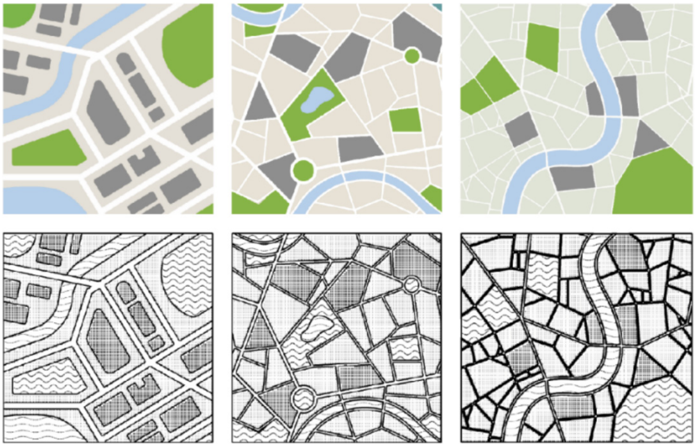


Fig. 2. User study materials for the quantitative session

In the qualitative session, user preference for RainbowTact compared to the traditional grating method was assessed. Participants were given 4 pairs of TG printouts. Figure 3 shows one of the given printouts each presented using either RainbowTact or the traditional grating method. They studied the given TG materials for 15 min. While exploring the two different TG methods, they were encouraged to think aloud, communicating with the research investigator. After completing both sessions, to understand the participants' preference for RainbowTact, we had a semi-structured interview session to get more information like their interpretation strategies and their overall satisfaction, using some relevant survey items from System Usability Scale[12] and open-ended questions.

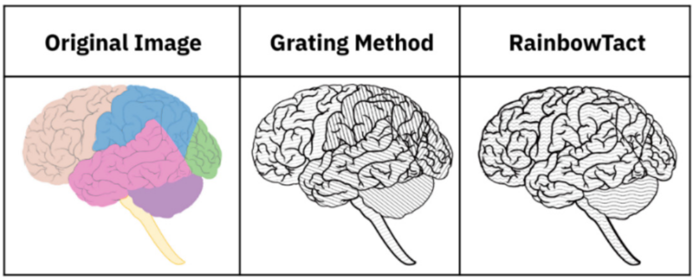


Fig. 3. an example user study material used for the qualitative session

4 Results

The success rates for each participant are shown in Table 1. While there were some variations among the participants, Overall, about two-thirds showed the task success rate. Given the relatively very short training time, the results were quite decent. Participants were able to distinguish chromatic from achromatic colors well, but showed relatively more errors with chromatic colors.

Table 1. Task Success Rate

Participant #	Map 1	Map 2	Map 3	Total
P1	1/3	2/3	2/3	5/9
P2	2/3	2/3	2/3	6/9
P3	2/3	3/3	2/3	7/9

The line chart presented in Fig. 4 indicates a clear trend of increasing efficiency among all participants in task completion times. This consistent progression can be interpreted as evidence that RainbowTact being used allows users to quickly adapt and improve their performance. In terms of task completion time as well, achromatic color identification showed a much faster performance than chromatic one.

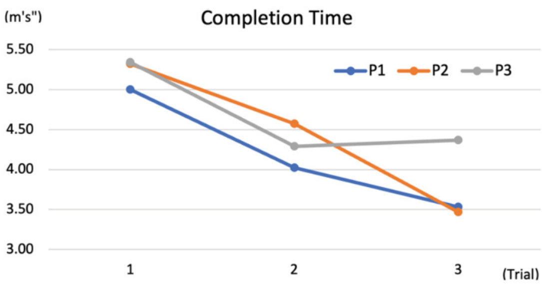


Fig. 4. Task Completion Time

In the qualitative evaluation session, participants uniformly favored the wave patterns generated by RainbowTact over traditional grating methods. Participant #2 remarked on the tactile differentiation, stating, "The degree of waviness is definitely noticeable… But comparing the angle of the slopes is too difficult." Participant #3 emphasized the non-directional advantage of RainbowTact's wave patterns, noting, "No matter which direction my hand is in, I can recognize the wavelengths," highlighting wave patterns' ease of differentiation. Participant #1 pointed out the grating method's limitations: "I can't tell if I am feeling the slope because it actually is a slope, or if it appears to be a slope because the outline is tilted," indicating challenges in perceiving grating cues.

The survey responses, measured on a seven-point Likert scale, further corroborated these qualitative insights. Participants expressed a positive inclination towards future use of RainbowTact, with an average score of 5.33, and rated its color translation features as consistent and intuitive, achieving an average score of 5.67. Nonetheless, the ease of learning RainbowTact was rated moderately, at an average of 2.67, indicating some challenges in initial adaptation. Participants reported a wide range of comfort and confidence levels with RainbowTact, with scores spanning from 2 to 7.

5 Discussion

This research aimed to address the critical gap in effectively representing color information in tactile graphics for the visually impaired. Our proposed method, RainbowTact, demonstrates strong potential in translating colors into intuitive tactile patterns. The pilot study results and user feedback suggest RainbowTact's wave and dot patterns can enable improved color identification and differentiation over traditional methods. Key benefits like omnidirectionality, alignment with existing guidelines, and the ability to discern attributes like hue, saturation and brightness also indicate its promise. With decent success rates achieved even with minimal training, RainbowTact shows the capacity to provide visually impaired users a more holistic access to visual information.

A key insight from our exploration was the innovative correlation of tactile sensations with the visible light spectrum's wavelengths. This inventive mapping, which aligns shorter tactile wavelengths with longer visible wavelengths (and vice versa), endeavors to replicate the inverse relationship observed in light physics. This choice, while based on a scientific rationale, invites further investigation and debate to ensure that the tactile-color association is both intuitive for users and grounded in perceptual science.

The consistency of tactile patterns across different production methods is important [13]. In terms of this, RainbowTact highlighted a significant technical limitation. The demand for higher-resolution tactile representations may exceed the capabilities of standard tactile printing devices, which typically operate at around 20 DPI. This challenge points to a gap in current tactile graphic production technologies, emphasizing the need for advancements that can accommodate the detailed granularity required by RainbowTact's representations. Our ongoing efforts aim to identify the optimal levels of tactile resolution and height necessary to better convey color information through touch.

Future research directions will focus on addressing the identified issues, with an emphasis on refining the tactile-color correlation to enhance RainbowTact's effectiveness and user satisfaction. Engaging with the visually impaired community will remain a cornerstone of our approach, ensuring that subsequent iterations of RainbowTact are informed by user feedback and are tailored to meet their needs and preferences more effectively. Conducting larger-scale studies and real-world applications will also be crucial in validating the efficacy and user satisfaction of RainbowTact, with the ultimate goal of advancing the field of tactile graphics and enriching the informational accessibility for the visually impaired.

Acknowledgments. This study was funded by Ministry of Education of the Republic of Korea and the National Research Foundation of Korea. (#NRF-2021S1A3A2A01087325).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Gardner, J.A.: Tactile graphics: an overview and resource guide. Inf. Technol. Disabil. **3**(4) (1996)
2. Ina, S.: Computer graphics for the blind. ACM Sigcaph Comput. Physically Handicapped **55**, 16–23 (1996)
3. Crombie, D., Lenoir, R., McKenzie, N., Ioannidis, G.: The bigger picture: automated production tools for tactile graphics. In: Miesenberger, K., Klaus, J., Zagler, W.L., Burger, D. (eds.) Computers Helping People with Special Needs, ICCHP 2004, LNCS, vol. 3118, pp. 713–720.Springer, Berlin (2004). https://doi.org/10.1007/978-3-540-27817-7_106
4. Rataj, A.: An algorithm for transforming color images into tactile graphics, ArXiv, vol. cs.GR/0404022 (2004)
5. Ladner, M.I.R., et al.: Automating tactile graphics translation, pp. 150–157 (2005)
6. Krufka, K.B.S.E.: Automatic production of tactile graphics from scalable vector graphics, pp. 166–172 (2005)
7. BANA. Guidelines and Standards for Tactile Graphics, https://www.brailleauthority.org/tg/
8. Van Boven, R.W., Hamilton, R.H., Kauffman, T., Keenan, J.P., Pascual-Leone, A.: Tactile spatial resolution in blind Braille readers. Neurology **54**(12), 2230–2236 (2000)
9. Shin, J., Cho, J., Lee, S.: Please touch color: Tactile-color texture design for the visually impaired, pp. 1–7 (2020)
10. Isumi, R., Yamamoto, K., Noma, T.: Color2Hatch: conversion of color to hatching for low-cost printing. Vis. Comput. **37**(12), 3103–3113 (2021)
11. Cho, J.D., Quero, L.C., Bartolome, J.I., Lee, D.W., Oh, U., Lee, I.: Tactile colour pictogram to improve artwork appreciation of people with visual impairments. Color. Res. Appl. **46**(1), 103–116 (2021)
12. Lewis, J.R.: The system usability scale: past, present, and future. Int. J. Hum.-Comput. Interact. **34**(7), 577–590 (2018)
13. Prescher, D., Bornschein, J., Weber, G.: Consistency of a tactile pattern set. ACM Trans. Accessible Comput. (TACCESS) **10**(2), 1–29 (2017)

New Methods for Creating Accessible Material in Higher Education

STS New Methods for Creating Accessible Material in Higher Education
Introduction to the Special Thematic Session

Michaela Hanousková[1], Boris Janča[1], Lukáš Másilko[1], Karin Müller[2](✉), Svatoslav Ondra[1], Radek Pavlicek[1], Petr Penaz[1], Andrea Petz[3], Thorsten Schwarz[2], and Rainer Stiefelhagen[2]

[1] Masaryk University, Brno, Czech Republic
{hanouskova,janca,masilko,ondra,pavlicek,penaz}@teiresias.muni.cz
[2] Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
{karin.e.mueller,thorsten.schwarz,rainer.stiefelhagen}@kit.edu
[3] Integriert Studieren, Johannes Kepler University, Linz, Austria
andrea.petz@jku.at
https://www.springer.com/gp/computer-science/lncs ,
https://www.access.kit.edu/english/ ,
https://www.jku.at/institut-integriert-studieren/

Abstract. This paper introduces the Special Thematic Session (STS) on *New Methods for Creating Accessible Material in Higher Education* [1]. Inclusive education, as mandated by the UN Convention on the Rights of Persons with Disabilities, demands universities to provide accessible study conditions. However, many study materials lack accessibility especially for visual content such as graphics and there are only few universities that offer a transfer service. Therefore, most students have to create solutions themselves. It is therefore necessary to develop new standardized methods that make it easier for students with visual impairments to access visual content of teaching material and exams e.g. by generating captions of charts, providing standardized descriptions and making graphics available tactilely.

Keywords: Accessibility · Higher Education · Inclusive Education · Disability · STEM Accessibility

1 Introduction

Access to education is crucial for all people, regardless of their individual circumstances or limitations. This principle is not only a matter of equality, but also a cornerstone of an inclusive society that values and promotes the diversity and uniqueness of each individual. In the pursuit of inclusive education, as mandated by Article 24 of the UN Convention on the Rights of Persons with Disabilities, access to higher education is recognized as a fundamental right. This applies not only to physical accessibility, but also to the area of teaching materials; universities and other institutions are obliged to provide barrier-free study conditions.

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 285–290, 2024.
https://doi.org/10.1007/978-3-031-62846-7_34

Thus, it is crucial that institutions in higher education make their teaching and learning materials accessible to ensure that all students can benefit equally from educational opportunities.

A UNESCO report [2] emphasizes the importance of accessible educational materials and highlights that they can help make education systems more inclusive and promote participation for all students. Furthermore, UNESCO stresses that the provision of accessible materials is not only an ethical imperative, but also an important prerequisite for achieving the Sustainable Development Goals.

The benefits of providing accessible materials are manifold. It's not just about meeting legal requirements, but also about creating an inclusive learning environment where all students can reach their full potential. Accessible materials not only help people with disabilities, but also other students who may benefit from alternative ways of presenting or accessing learning.

However, many educational institutions do not provide fully accessible teaching and learning materials or are even not aware of their obligation to provide all students with access to teaching material. It is mostly up to the students to decide how they obtain the teaching materials and how they are made accessible, for example by hiring assistants to transfer literature, teaching them how to make documents accessible or asking lecturers to make their material available digitally so that they at least have access to the content. This has far-reaching implications for how people with disabilities access education. A lack of accessible materials can lead to students with disabilities or other health impairments being excluded from full participation in academic life.

A possible solution involves new methods like artificial intelligence or LLMs (Large Language Models) that show potential in e.g. automatically describing complex structures like tables or mixed/non text elements like pictures, graphs, charts etc. However, first examples show that this process only works semi-automated in the moment - a human supervision is needed to check e.g. the descriptions on completeness but also on a logically appropriate cue, a red thread through the picture or even chart described. Often, pictures and other mixed/non text alternatives are used additionally to text - and a verbose description of an element that is fully explained in one of the paragraphs above does keep the readers from efficiently grabbing the idea, use the provided content and results in using up the readers' time and attention. In the moment, the powerful algorithms of an LLM or AI are not able to link mixed/non text elements to information / positions in a text or understand the text around non textual content and are also not able to decide whether the information in it must be explained and represented or can easily be discarded. So a certain amount of research, definition and standardisation work needs to be done before these technologies can reach their full potential and find their way into making up complex structures and mixed/non textual content accessibly and usable instead building up new barriers.

This Special Thematic Session (STS) seeks to explore the multifaceted challenges associated with creating accessible teaching and learning materials in higher education for students with visual impairments. By examining tools, methodologies, and best practices, this session endeavors to equip educators and institutions with the knowledge and resources necessary to foster truly inclusive learning environments.

2 Current Solutions to Improve Access to Teaching Material

Teaching material includes a wide range of document types, from handwritten and digital notes to scientific papers, slides, books, and assignments. In addition, they frequently comprise graphical content in two and three dimensions, which are essential for class participation, lab activities, and examinations. However, despite these mandates, a significant portion of educational materials remains inaccessible to students with visual impairments. Addressing this accessibility gap necessitates a comprehensive approach, tailored to the unique needs and working methods of individual students.

In the following paragraphs, we present four approaches that developed new solutions on how to improve the accessibility of charts by standardize descriptions, generate them automatically and provide tactile graphics. Another solution addresses the improvement of the accessibility of exams.

2.1 Accessibility of Graphics

While information is often presented visually by pictures, schemes, or charts to summarize complex content, it is not accessible for blind persons. There are currently three main possibilities to make them accessible: (i) creating a description, (ii) substituting the graphical content with another form (e.g., replacing charts with tables), and/or (iii) providing a tactile graphic that allows blind users to explore graphical content via touch. Zebehazy et. al. [8] found that the best way to access graphics for blind students at school is to provide a tactile version and a description. They regard exploring a tactile graphic, discovering information, and answering questions independently is regarded as a fundamental part of the learning process.

Until recently, descriptions were mainly created manually by experts who decided which content of a graphic was most essential and then described it in words. It is usually up to the experts how they describe and structure a description. Now, gradually, it is possible to involve AI-based tools in the process of recognizing and describing visual information, e.g. using a digital visual assistant powered by OpenAI's ChatGPT-4 vision model.

Additionally to descriptions, tactile graphics can add extra information. Moreover, for some purposes, the tactile representation is still the only possible one, e.g. maps. However, creating tactile graphics usually involves the manual work of experts and needs a lot of time. Even here, however, it is possible to use tools that replace some of the manual work - whether it be map generators or conversion tools for preparing tactile charts (see below).

Accessibility of Charts. As screen reader users generally increase the reading speed of their screen reader to access information more quickly, they appreciate structured information. Thus, [6] addresses the problem of describing function graphs. Recognizing the lack of standardized descriptions for mathematical function graphs, particularly for blind individuals, they propose to develop a comprehensive description standard in a user-centered design process. By structuring information according to the needs of blind readers, this approach aims at enhancing accessibility and comprehension in university-level mathematics education.

Nowadays, more and more deep learning methods are being used to generate such descriptions. Alt-Texts are obligatory according to WCAG [9] and are a critical component for enhancing accessibility for blind and visually impaired individuals. To streamline the creation of alternative texts (Alt-Text) for chart images, [3] present a method to support experts in creating these texts with AI-based methods. By harnessing deep learning models and a benchmark dataset, the system facilitates the generation of high-quality Alt-Text, thus improving the accessibility of graphical content in educational materials.

An alternative approach of [5] focuses on developing a real-time chart captioning system optimized for visually impaired individuals in response to the increasing reliance on visual data in online classes and video meetings. By leveraging AI for interpretation and captioning of charts, this system aims to bridge the accessibility gap in dynamic and interactive remote environments, thereby fostering a more inclusive digital learning landscape.

While tactile graphics offer benefits to blind students, they can be challenging to prepare. Thus, [4] addresses the challenge of making chart images tactilely accessible by ChartFormer. It presents an innovative approach leveraging a large vision language model. Through the conversion of raster chart images into tactile-accessible Scalable Vector Graphics (SVGs), this approach aims to facilitate data interpretation for individuals with visual impairments, thereby promoting inclusivity in educational settings.

2.2 Accessibility of Texts

Since people who rely on modern assistive technologies are largely dependent on the existence of digital versions of learning materials, one would expect that in today's digital age, text accessibility would be a simple and solved issue. While the easy availability of digital tools for text creation and editing, text editors, has brought with it the existence of a massive number of digital documents, paradoxically, it has not yet ensured their greater accessibility. The emphasis on the visual aspect, but the omission of technical correctness of the typesetting, often brings difficulties not only to users of assistive technologies. Moreover, some types of documents are traditionally delivered to students in a non-editable form (e.g. tests).

Accessibility of Exams. At university, students often prepare for their own exams with the help of past exams. However, these documents are often only available in a printed format and thus constitute a critical barrier for visually impaired students. [7] addresses this challenge through a semi-automatic AI-based approach, guiding sighted individuals in efficiently converting exam documents into accessible formats. By incorporating necessary accessibility features, this initiative aims to promote equitable access to exam materials and enhance the learning experience for all students.

3 Conclusion

In summary, the contributions of this Special Thematic Session stress the importance of enhancing accessibility and inclusivity in higher education. The presented approaches apply innovative tools, methods, and develop standards that enable educators and institutions to improve equitable access to learning materials for all students, regardless of their abilities. There are still many research questions which have not been answered yet. The new approaches are promising that materials can be made accessible more quickly in the future by overcoming barriers (semi-)automatically and access to information is facilitated for all.

References

1. Müller, K., et al.: STS new methods for creating accessible material in higher education. In: Miesenberger, K., et al. (eds.): ICCHP 2024, LNCS, pp. 285–290. Springer (2024). https://doi.org/10.1007/978-3-031-62846-7_34
2. UNESCO. Every learner matters and matters equally: making education inclusive; background document, UNESCO, Paris. https://unesdoc.unesco.org/ark:/48223/pf0000388991.locale=en
3. Moured, O., Ali Farooqui, S., Müller, K., Fadaei, S., Schwarz, T., M. Javed, R. Stiefelhagen, R.: Alt4Blind: A User Interface to Simplify Charts Alt-Text Creation. In: Miesenberger, K., et al. (eds.): ICCHP 2024, LNCS, pp. 291–298. Springer (2024). https://doi.org/10.1007/978-3-031-62846-7_35
4. Moured, O., Alzalabny, S., Osman, A., Müller, K., Schwarz S., Stiefelhagen, R. ChartFormer: A large vision language model for converting chart images into tactile accessible SVGs. In: Miesenberger, K., et al. (eds.): ICCHP 2024, LNCS, pp. 299–305. Springer (2024). https://doi.org/10.1007/978-3-031-62846-7_36
5. Jeong, S., Song, Y., Lim, S.-B., Cho, W., Park, J. A real-time chart explanation system for visually impaired individuals. In: Miesenberger, K., et al. (eds.): ICCHP 2024, LNCS, pp. 306–312. Springer (2024). https://doi.org/10.1007/978-3-031-62846-7_37
6. Sobota, Z. E., Freese, A.-C., Schwarz, T., Müller, K. Towards establishing a description standard of mathematical function graphs for people with blindness. In: Miesenberger, K., et al. (eds.): ICCHP 2024, LNCS, pp. 313–320. Springer (2024). https://doi.org/10.1007/978-3-031-62846-7_38

7. Moured, O., Wilkening. D., Müller, K., Schwarz, T., Stiefelhagen, R. ACCSAMS: Automatic Conversion of Exam Documents to Accessible Learning Material for Blind and Visually Impaired. In: In: Miesenberger, K., et al. (eds.): ICCHP 2024, LNCS, pp. 321–329. Springer (2024). https://doi.org/10.1007/978-3-031-62846-7_39
8. Zebehazy, K.T., Wilton, A.P.: Quality, importance, and instruction: the perspectives of teachers of students with visual impairments on graphics use by students. J. Visual Impairment Blindness **2014**, 275–286 (2014)
9. Caldwell, B., et al.: Web content accessibility guidelines (WCAG) 2.0. In: WWW Consortium (W3C), vol. 290, 1-34, pp 5–12 (2008)

Alt4Blind: A User Interface to Simplify Charts Alt-Text Creation

Omar Moured[1,2(✉)], Shahid Ali Farooqui[3], Karin Müller[2], Sharifeh Fadaeijouybari[4], Thorsten Schwarz[2], Mohammed Javed[3], and Rainer Stiefelhagen[1,2]

[1] CV:HCI@KIT, Karlsruhe Institute of Technology, Karlsruhe, Germany
moured.omar@gmail.com
[2] ACCESS@KIT, Karlsruhe Institute of Technology, Karlsruhe, Germany
[3] Computer Vision and Biometrics Lab, Indian Institute of Information Technology, Allahabad, India
[4] University of Trier, Trier, Germany
https://cvhci.anthropomatik.kit.edu/, https://www.access.kit.edu/, https://cvbl.iiita.ac.in/

Abstract. Alternative Texts (Alt-Text) for chart images are essential for making graphics accessible to people with blindness and visual impairments. Traditionally, Alt-Text is manually written by authors but often encounters issues such as oversimplification or complication. Recent trends have seen the use of AI for Alt-Text generation. However, existing models are susceptible to producing inaccurate or misleading information. We address this challenge by retrieving high-quality alt-texts from similar chart images, serving as a reference for the user when creating alt-texts. Our three contributions are as follows: (1) we introduce a new benchmark comprising 5,000 real images with semantically labeled high-quality Alt-Texts, collected from Human Computer Interaction venues. (2) We developed a deep learning-based model to rank and retrieve similar chart images that share the same visual and textual semantics. (3) We designed a user interface (UI) to facilitate the alt-text creation process. Our preliminary interviews and investigations highlight the usability of our UI. For the dataset and further details, please refer to our project page: https://moured.github.io/alt4blind/.

Keywords: Alt-Text · Image Retrieval · CLIP Model

1 Introduction

Making charts accessible to a broader audience is essential to ensure equal access. Part of this audience includes people with blindness and visual impairments individuals who access data through different modalities: tactile and auditory. The choice between these modalities often depends on the available technologies

O. Moured and S. A. Farooqui—Both authors contributed equally to this work.

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 291–298, 2024.
https://doi.org/10.1007/978-3-031-62846-7_35

and the preferences of people with blindness and visual impairments. Although technologies like sonification and tactile materials are available, *Alt-Text* remains the primary method for BVI individuals to acquire information about images through screen readers [4]. In accordance by the WCAG[1], an alt text is required for a graphic to be accessible. These summaries are embedded as hidden tags within documents or web pages. A study [2] revealed that over 80% of web pages often neglect to include alt text. Furthermore, even when available, it often doesn't comply with recommended guidelines (e.g. WCAG) and standards (e.g. W3C).

Chart alt-text can originate from two primary sources: manually created by humans or (semi-)automatically generated by AI systems. Human-generated descriptions are often accurate but typically require expertise. Recent studies have highlighted a significant Alt-text deficiency in publications [1]. On the other hand, automatically generated captions are quicker to obtain, and require no expertise but may suffer from information inaccuracy. Thus, crafting a high-quality chart alt-text is not a trivial task [6].

Instead, this work investigates the role of using AI in aiding both experts and non-expert users to write high-quality alt-text by presenting high-quality, similar charts with alt-text as references for them. To achieve this, we built a comprehensive dataset composing 5,000 chart images sourced from venues providing alt-text with their publications. We then trained an AI model to categorize images based on their visual and textual similarities. Subsequently, the most similar images are retrieved and presented to the user through our *Alt4Blind* interface (Fig. 1).

2 Related Work

In this section, we first examine existing datasets in the field of chart analysis, followed by a discussion on the available models and tools for creating textual descriptions of charts.

2.1 Datasets

Chart-2-Text [5] is a popular dataset in computer vision. It comprises a large number of chart images of various types, with each image accompanied by summary and bounding box annotations. However, it is not tailored for people with blindness and visual impairments, the summaries are not structured and do not adhere to the accessibility guidelines. In contrast, Vistext [10] comprises 8822 synthesized images, each manually assigned alt-texts, representing a significant effort. However, it features synthetic charts, restricted to line and bar charts, and the alt-texts were not authored by the original image creators, potentially omitting crucial context. Another dataset, HCI Alt-Text [1], collected chart figures from HCI conferences and aims to analyse the quality of provided alt-texts. However, it consists of 511 chart images, making it less suitable for deep-learning models.

[1] https://www.w3.org/WAI/alt/.

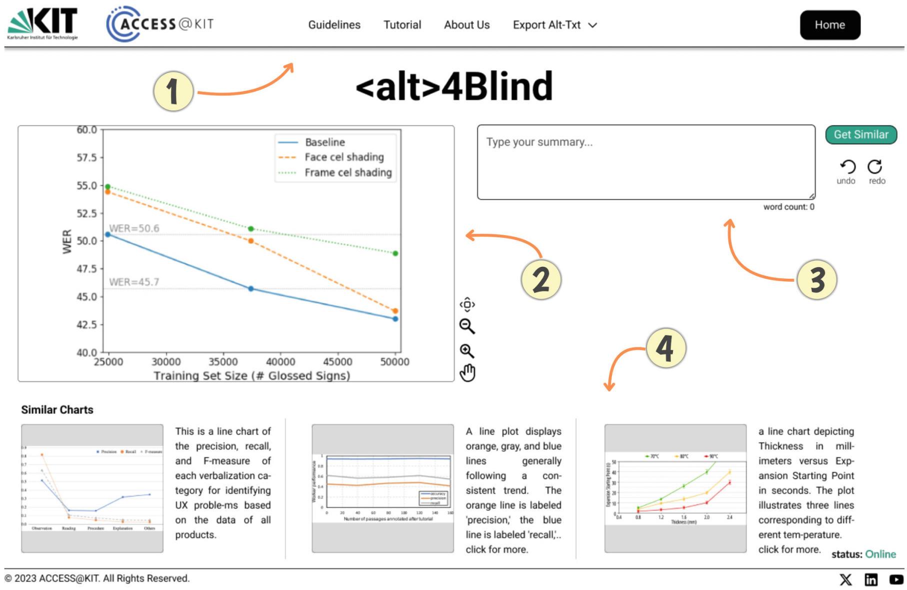


Fig. 1. *Alt4Blind* UI: (1) Menu bar offering access to guidelines and a tutorial. (2) Space for uploaded images featuring a function bar (zoom, move, fit). (3) Text field for user input, accompanied by a button to update the retrieved image. (4) Retrieved charts based on the uploaded image, can be further enhanced with text query.

2.2 Alt-Text Creation Models

Chart Analysis is an emerging field in computer vision, featuring recent works such as ChartAssistant [8] and ChartLLama [3]. These models are trained for various chart-related tasks, including description generation, chart2table, and chart2code, among others. Despite being trained on a vast chart corpus, these models fall short of addressing accessibility concerns.

2.3 Tools

K. Mack [7] analyzed various interface designs to facilitate the creation of high-quality alt-texts. Their results concluded that **interfaces assisting authors in deciding what to include in the alt-text were well-received and led to higher-quality descriptions**. However, there are very few tools that adopt this approach. WATAA [4] is one such tool, which presents users with AI-generated descriptions as a reference. Nevertheless, AI descriptions are prone to hallucinations.

In this paper, we focus on addressing both limitations: (i) creating a database with high-quality alt-texts, and (ii) developing an intelligent tool that aids people with and without accessibility knowledge to author high-quality alt-text.

3 Exploratory Interviews and Findings

With the help of the Center for Digital Accessibility and Assistive Technologies in Karlsruhe Institute of Technologies[2], we recruited six participants with varying levels of accessibility expertise, from none to advanced (working in the field of accessibility). We conducted exploratory interviews with these participants (experts (P1-P3), non-experts (P4-P6)) to identify necessary features for the development of an intuitive user interface, the *Alt4Blind* app. Participants were not compensated.

3.1 Procedure

The interviews took place in a one-hour face-to-face session following the steps:

1. Initially, we introduced a basic line chart featuring three distinct lines, see the illustration in the upper left of Fig. 1–2.
2. Participants were then instructed to write a description for a blind person.
3. Following this, we inquired about the challenges faced and their perspective on essential features for a web tool designed for this task.

No compensation was offered for participation in this study. Furthermore, the user study received approval from the KIT Internal Ethical Committee.

3.2 Qualitative Results

During the interview, we took notes that we analyzed after the sessions. The analysis revealed the following observations: P1-P3, maintained a coherent reading order, starting with visual semantics (e.g., chart type, title, and axis labels) followed by contextual information (e.g., line trends, and comparisons). In response to the follow-up questions, P1 noted that they usually find Multivariate and Panel charts challenging. P4-P6 discussed how they were unfamiliar with alt-texts as it's often hidden in tags, and suggested the need for an initial guideline before start writing. Based on these insights, we identified as a feature that an initial suggestion for user-uploaded images is necessary. Thus, we chose to **(1) display a real initial suggestion**, i.e. displaying real chart images with human-authored high-quality alt-texts. **(2) the images should be visually or contextually relevant to the user chart.**

[2] www.access.kit.edu.

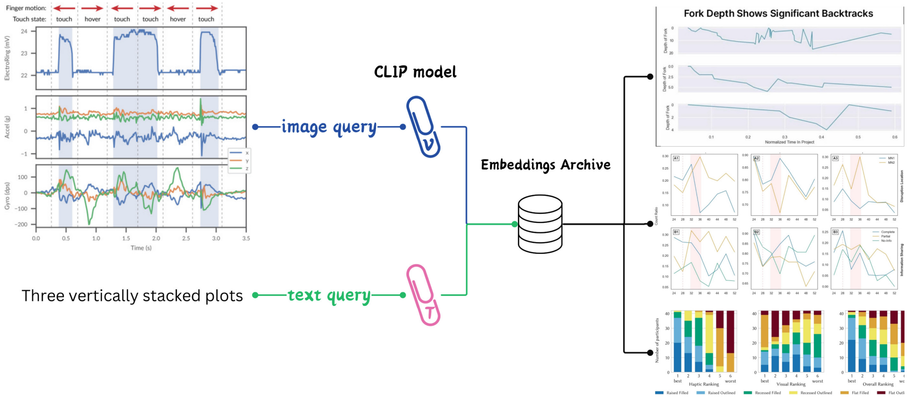


Fig. 2. Our retrieval system leverages both the text and image encoder modules of the fine-tuned CLIP model. This ensures similarity at both visual and contextual levels.

4 Alt4Blind

In this section, we describe (i) our dataset, (ii) the AI-based model for ranking and retrieving similar charts with alt-texts, and (iii) the creation of the Alt4Blind User Interface.

4.1 Creation of Our Dataset

We first crawled 25,000 images with alt-text from publicly available HCI-related conferences spanning the last 10 years to expand on the HCI Alt-Text dataset [1]. In the next step, we filtered out images that are not charts, have a short alt-text or lack semantics, so we end up with 5,000 chart images, marking a tenfold increase from the previous dataset.

Our dataset contains a wide variety of chart types, from familiar forms like Line, Bar, Area, Pie, and Scatter charts to more unique visualizations not present in previous works such as multi-variate and panel charts. We used the 4-level semantic model by Lundgard et al. [6] to rate alt-text quality. Initial scoring was conducted using ChatGPT-4 by assessing the number of semantics and levels present in each alt-text. The richest samples were subsequently manually reviewed.

4.2 Chart Image Retrieval

Image retrieval involves searching for images similar to a given query image or text, with a focus on ensuring the top results include similar chart semantics. For this purpose, we employed CLIP ViT-B/32 [9] as our baseline model. The

CLIP architecture encompasses both a text and an image encoder, each producing a 512-dimensional feature vector. The CLIP model uses the unsupervised contrastive pre-training approach to cluster the samples in the latent space. Hence, we fine-tuned the model on our dataset, with a batch size of 16 for 10 epochs. Next, during the inference phase, Fig. 2, we input the image, extract the encoders' representations and compare it to the samples from the datasets. The top three candidates with the highest cosine similarity scores are then retrieved. Our model achieved 92% in P@3 and 85% in R@3, demonstrating high capabilities in Precision (P@3) and Recall (R@3) for displaying similar chart images within the top three results, respectively.

4.3 User Interface

For our prototype, we utilized React JS[3], a JavaScript library ideal for building user interfaces. It includes a landing page for users to drag-and-drop chart images and a tutorial for first-time use. Once an image is uploaded, the backend model display the top three similar candidates on the main page (see Fig. 1). Users can enlarge to view the full alt-text. The "Get Similar" button allows users to refresh the candidates when they type their summary in the text field.

5 Preliminary Evaluation

We conducted tests with the same participants, P1-P6, to assess our user interface. Participants were presented with Panel and Multivariate charts, which were unfamiliar to all. The samples are depicted in Fig. 1, and participants were instructed to:

1. Review the guidelines and participate in the tutorial chart session first.
2. Use Alt4Blind to upload the chart and create alternative text for the image provided.
3. Describe their experience in detail.

During the session, we tracked mouse movements and recorded user interaction behaviors using pen and paper. All participants produced detailed descriptions. P4-P6 especially benefited from the feature that allows copying sentences from various similar charts to craft their descriptions. Expert users appreciated this feature and recommended adding chart captions to further complement the alternative text. P1 suggested increasing access to more than three similar charts, whereas P3 proposed a feature to replace irrelevant images among the selected charts.

[3] https://react.dev/.

6 Discussion and Limitations

In this section, we shortly discuss our results and the limitations of this work.

Intelligent Features: Our tool has demonstrated efficiency in enabling both inexperienced and experienced users to author high-quality alt-texts. However, it could further benefit from additional intelligent functionalities, such as LLM-based feedback and descriptions, which are currently under development.

Captions: While the current implementation assists users in creating alt-text, we believe that integrating captions into the tool is essential, as captions and alt-texts often complement each other.

Interactions: Current UI does not allow users to have control over the similar chart section. Future iterations should offer users control, allowing them to omit, replace, or view more charts.

User Study: Our initial investigations, though encouraging, were conducted with a limited number of users and lacked comprehensive control measures. Future studies should engage a larger and more diverse group of participants and should involve people with blindness and visual impairments.

7 Conclusion

We presented an online, open-source tool to enhance alt-text writing by using AI image retrieval, making it more engaging. Our shared dataset invites further development by the NLP and Computer Vision communities to advance chart summarization tasks with a focus on people with blindness and visual impairments. Currently, while our work primarily provides examples of alternative texts, it lays the groundwork for future research incorporating Vision-Language models to address accessibility concerns.

Acknowledgments. The authors would like to acknowledge the help of P. Venkatesh for his support in developing the UI. We would also like to thank the HoreKa computing cluster at KIT for the computing resources used to conduct this research.

Disclosure of Interests. This research was funded by the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreements no. 861166.

References

1. Chintalapati, S.S., Bragg, J., Wang, L.L.: A dataset of alt texts from HCI publications: analyses and uses towards producing more descriptive alt texts of data visualizations in scientific papers. In: Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility, pp. 1–12 (2022)
2. Commission, D.R.: The Web: Access and Inclusion for Disabled People; A Formal Investigation. The Stationery Office (2004)

3. Han, Y., et al.: Chartllama: A multimodal LLM for chart understanding and generation (2023)
4. Jeong, H., Chun, M., Lee, H., Oh, S.Y., Jung, H.: Wataa: web alternative text authoring assistant for improving web content accessibility. In: Companion Proceedings of the 28th International Conference on Intelligent User Interfaces, pp. 41–45 (2023)
5. Kantharaj, S., et al.: Chart-to-text: a large-scale benchmark for chart summarization. arXiv preprint arXiv:2203.06486 (2022)
6. Lundgard, A., Satyanarayan, A.: Accessible visualization via natural language descriptions: a four-level model of semantic content. IEEE Trans. Visual Comput. Graphics **28**(1), 1073–1083 (2021)
7. Mack, K., Cutrell, E., Lee, B., Morris, M.R.: Designing tools for high-quality alt text authoring. In: Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility, pp. 1–14 (2021)
8. Meng, F., et al.: Chartassisstant: a universal chart multimodal language model via chart-to-table pre-training and multitask instruction tuning (2024)
9. Radford, A., et al.: Learning transferable visual models from natural language supervision. In: International Conference on Machine Learning, pp. 8748–8763. PMLR (2021)
10. Tang, B.J., Boggust, A., Satyanarayan, A.: Vistext: a benchmark for semantically rich chart captioning. arXiv preprint arXiv:2307.05356 (2023)

ChartFormer: A Large Vision Language Model for Converting Chart Images into Tactile Accessible SVGs

Omar Moured[1,2](✉), Sara Alzalabny[3], Anas Osman[4], Thorsten Schwarz[2], Karin Müller[2], and Rainer Stiefelhagen[1,2]

[1] CV:HCI@KIT, Karlsruhe Institute of Technology, Karlsruhe, Germany
moured.omar@gmail.com
[2] ACCESS@KIT, Karlsruhe Institute of Technology, Karlsruhe, Germany
[3] NeptunLab, University of Freiburg, Freiburg im Breisgau, Germany
[4] Middle East Technical University, Ankara, Turkey

Abstract. Visualizations, such as charts, are crucial for interpreting complex data. However, they are often provided as raster images, which are not compatible with assistive technologies for people with blindness and visual impairments, such as embossed papers or tactile displays. At the same time, creating accessible vector graphics requires a skilled sighted person and is time-intensive. In this work, we leverage advancements in the field of chart analysis to generate tactile charts in an end-to-end manner. Our three key contributions are as follows: (1) introducing the *ChartFormer* model trained to convert raster chart images into tactile-accessible SVGs, (2) training this model on the *Chart2Tactile* dataset, a synthetic chart dataset we created following accessibility standards, and (3) evaluating the effectiveness of our SVGs through a pilot user study with an refreshable two-dimensional tactile display. Our work is publicly available at https://github.com/nsothman/ChartFormer.

Keywords: Vision-Language Models · Chart Analysis · Tactile Charts

1 Introduction

Charts are essential for communicating complex information to a broad and diverse audience. People with blindness and visual impairments access these charts primarily through text-based descriptions (e.g., screen readers), and tactile representations (e.g., embossed paper, tactile displays). The use of tactile charts in educational setting has been shown to enhance the development of people with blindness and visual impairments' analytical skills [11]. These tactile charts are often created using drawing software for vector graphics (e.g., Inkscape, LibreOffice Draw, etc.) and saved in the Scalable Vector Graphics (SVG) format. SVGs are XML-based files that store geometrical shapes using mathematical formulas in a hierarchical structure. This format presents several advantages for creating accessible graphics [1], including (1) each element in an

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 299–305, 2024.
https://doi.org/10.1007/978-3-031-62846-7_36

SVG can be assigned different styles, translating into distinctive textures in the tactile version. (2) SVG files can hold supplementary textual descriptions which enhances interactivity when used with screen readers or tactile displays. (3) ability to be resized without causing blurring or distortion, making them ideal for varying paper sizes or zooming on tactile displays [7]. Creating an SVG chart from a raster image requires *careful simplification of both textual and visual content to support tactile formats while also preserving the integrity of the chart information*. This includes reducing textual content, such as limiting the number of axis labels and focusing on only crucial visual elements, like emphasizing significant scatter points in a scatter plot. Due to these complexities, crafting vector graphics is not a trivial task. Nevertheless, AI-based chart analysis models have demonstrated emerging capabilities in image analysis. In light of these developments, we outline three main contributions in our work: (1) We introduce a transformer-based model that extracts key information and assigns styles for the SVG file. (2) We present the Chart2Tactile dataset, created in adherence to accessibility guidelines. (3) We share insights from a pilot user study involving four blind participants who evaluated SVG graphics generated by our model on a 2D tactile display.

2 Related Work

Very few studies have explored automated heuristics for making raster charts tactically accessible. Most prior research follows a two-step methodology to convert charts, beginning with metadata extraction and followed by conversion into tactile format. In this section, we discuss these methods and also the available benchmarks.

2.1 Metadata Extraction

A recent work, ChartLLama [3], trained a large Vision-Language Model (VLM) on synthetically generated images across 10 chart types. One of its tasks is converting charts into Matplotlib Python codes. While this model demonstrates the potential of VLMs in chart comprehension, it overlooks accessibility concerns. ChartDetective [5] provides a UI for processing SVG chart documents, allowing data extraction through simple drag-and-drop interactions. This tool is useful when SVGs are available, but additional steps are still required for redrawing the data into accessible tactile materials.

2.2 Tactile Charts

A limited number of studies have proposed systems capable of both extracting data and generating tactile materials. Engel et al. presented the *SVGPlott* system [2], a GUI that comprises six steps to convert bitmap images into printable SVG format. This system streamlines the process through step-by-step guidance;

however, it still requires manual extraction of metadata, which may not be feasible for dense charts. A more recent development, Chart4Blind [8], introduced an end-to-end AI-assisted user interface designed to convert charts into tactile formats accessible for the visually impaired. Nevertheless, it currently only supports line charts, highlighting the need for broader support of other types of visualizations.

2.3 Available Datasets

ChartAssistants [6], a recent work, stands out for its comprehensive collection of chart images, each paired with detailed metadata. Although it lays a solid groundwork for converting charts to code, it cannot generate accessible visualizations. In contrast, datasets oriented towards accessibility, like VisText [12], have focused on making visualizations accessible through chart summarization tasks, but none have considered the tactile modality. **To our knowledge, we propose the first dataset for the task of *Chart2Tactile* conversion.**

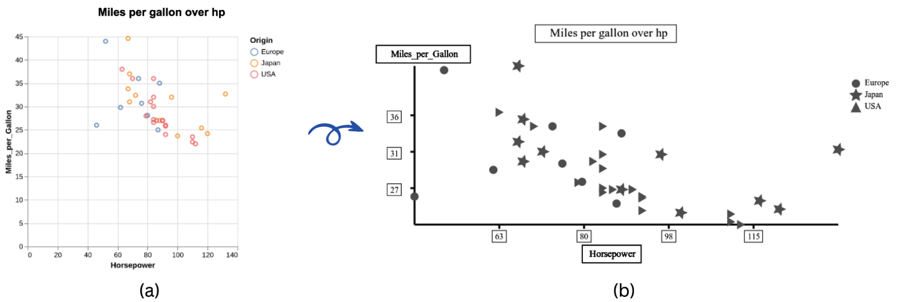


Fig. 1. A scatter plot sample: (a) the original synthesized raster image; (b) the tactile version following accessibility guidelines.

3 Chart2Tactile Dataset

Our dataset comprises 10,000 tactile chart images, spanning 4 categories (line, bar, scatter and error-bar charts) each accompanied by time series data and a raster version. Below, we describe the formation of the *Chart2Tactile* dataset.

3.1 Metadata Collection

To create a visualization, time-series data is essential. We aimed to select a source that would be augmented by our contributions. We identified the VisText [12] and ChartX [13] datasets as the most suitable choices. VisText offers 8,822 images, complete with their data tables and accessible summarizations, featuring univariate time series. ChartX contains 48K chart data covering 22 topics, 18 chart types, with each chart including four modalities: image, CSV, Python code, and text description. A sample raster image is presented in Fig. 1-(a).

3.2 SVG Guidelines

Rendering the metadata as tactile charts necessitates adherence to established guidelines to ensure that the charts are accessible by individuals with visual impairments. This process involves not only the translation of visual information into a tactile format but also the thoughtful consideration of how various elements can be differentiated by touch. We followed various tactile printing guidelines [2,10] to create accessible SVGs. The key requirements we adhered to are summarized as follows:

1. Elements should be distinguishable by touch, using varying thicknesses or symbol types such as dotted or dashed patterns.
2. Thin elements should be avoided.
3. Text in tactile illustrations should be in Braille, oriented horizontally.

Additionally, we collaborated with an expert from the Center for Digital Accessibility and Assistive Technology[1] at Karlsruhe Institue of Technology, specializing in converting educational materials for people with blindness and visual impairments. Their feedback included the following recommendations:

1. Enclose text content with a bounding box for better exploration and distinguishing separate texts more effectively.
2. For dense charts such as scatter plots, only significant, non-overlapping points should be drawn to avoid clutter.
3. Embed description tags for both text and visual elements to enable accessibility via screen readers.

3.3 SVG Creation

For transforming metadata into SVGs, we used the *svgwrite* Python package[2]. For each time series, we synthesized an SVG template and rendered a raster image using *Vega-Lite*[3]. To ensure accuracy, we manually selected samples from each category of the data and conducted a thorough verification process. A tactile sample is illustrated in Fig. 1-(b).

4 ChartFormer Model

We selected LLaVA-1.5 [4] as our baseline model, comprising a vision encoder for image input and a language model for text output decoding. We used the baseline weights from ChartLLama [3], adopted the same hyperparameters for training, and then fine-tuned the model for 10 epochs using our dataset. The model is trained to analyze x-y raster chart images as input and extract the simplified metadata with the styling for the *svgwrite* code (see Fig. 2). More specifically, the following information is extracted:

[1] https://www.access.kit.edu/english/index.php.
[2] https://svgwrite.readthedocs.io/en/latest/.
[3] https://vega.github.io/vega-lite/.

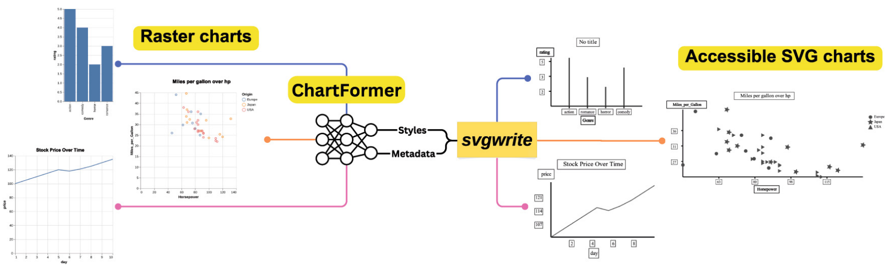


Fig. 2. The ChartFormer takes a raster x-y plot as an input. The essential metadata and styles are extracted, which are then used to populate the svgwrite templates. For better viewing resolution, please visit our project page.

1. The x-y chart type, and titles, including plot, axes and legend titles.
2. Axes range as 3 or 4 labels covering the whole period. The labels should adhere to the encodings (e.g., int, float, fraction, date/time and text).
3. Extract the time-series data for drawing.

The extracted data are then rendered using predefined svgwrite code templates for each chart category. It is important to note that for the scatter plot, we draw 10 points per label unit while separating overlapping points, in adherence to the SVG Guidelines 3.2.

5 Pilot User Study

We conducted a pilot user study with four people with blindness and visual impairments (three males and one female) to evaluate the effectiveness of the generated SVGs on a HyperBraille 2D tactile display[4].

5.1 Procedure

The user study images were randomly sampled from the LG dataset [9], which includes real charts. At the beginning of the session, a test graph was provided to introduce the participants to the available interactions. Afterwards, the 3 line charts shown in Fig. 3 were displayed and the participants were asked to explain and identify the key elements, titles, labels and legends and count lines, in each graph, as well as name few points intersection and line trend.

5.2 Results and Discussion

All participants successfully completed tasks related to charts (A) and (B), which involved identifying intersections and counting lines. They could also accurately describe the line trend as increasing, decreasing, or constant. However,

4 https://metec-ag.de/en/produkte-graphik-display.php.

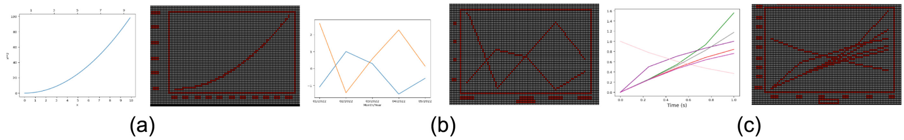


Fig. 3. SVG-formatted line charts used in the user study, showcasing varying complexities: (A) a single line; (B) two lines; (C) six lines. For better viewing resolution, please visit our project page.

in chart (C), participants encountered difficulties in counting all intersections, likely due to the chart's high density. Two participants used zoom features on the tactile display to discern closely positioned elements and intersections more clearly. They also appreciated the audio descriptions, which facilitated access to the chart's textual elements. A common suggestion from all participants was regarding SVG rendering, specifically to address the staircasing effect in the tactile output. The need for smoother line rendering to avoid jagged or stair-like appearances was emphasized.

5.3 Limitations

Although our system has been positively received by the participating people with blindness and visual impairments and collaborators, we believe there is still significant room for improvement: (1) Our system mainly targets x-y plots with two axes and charts of a single type. Future implementations could encompass other chart types. (2) Adding an interface to our system could allow sighted individuals to modify the chart before exporting it, ensuring that textual and visual details are accurately represented. (3) Conducting a larger, formal user study is necessary to assess the performance and furthermore, to experiment with different types of charts beyond just line charts.

6 Conclusion

In this work, we have showcased the potential of vision-language models for enhancing accessibility through the creation of tactile graphics and descriptive content. Current VL models fail to meet the special requirements that different people may need, such as those with visual impairments. It is important for future research to address it by refining models to better serve these diverse groups. We believe that available models such as ChatGPT or Gemini can be utilized to address these tasks with few- or zero-shot tuning, hence minimal computational demands, offering an easier available solution for schools and higher education institutes. We have also introduced the first dataset for tactile visualizations that complies with accessibility guidelines. Despite some limitations, our contributions aim to encourage further research into supporting accessible graphics production.

Acknowledgments. The authors would like to thank the HoreKa computing cluster at KIT for the computing resources used to conduct this research.

Disclosure of Interests. This research was funded by the European Research Council (ERC) grant and the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreements no. 816006 and 861166 respectively.

References

1. Dürnegger, B., Feilmayr, C., Wöß, W.: Guided generation and evaluation of accessible scalable vector graphics. In: Miesenberger, K., Klaus, J., Zagler, W., Karshmer, A. (eds.) ICCHP 2010. LNCS, vol. 6179, pp. 27–34. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-14097-6_5
2. Engel, C., Müller, E.F., Weber, G.: SVGPlott: an accessible tool to generate highly adaptable, accessible audio-tactile charts for and from blind and visually impaired people. In: PETRA 2019 (2019)
3. Han, Y., et al.: ChartLlama: a multimodal LLM for chart understanding and generation (2023)
4. Liu, H., Li, C., Li, Y., Lee, Y.J.: Improved baselines with visual instruction tuning (2023)
5. Masson, D., Malacria, S., Vogel, D., Lank, E., Casiez, G.: Chartdetective: easy and accurate interactive data extraction from complex vector charts. In: CHI 2023 (2023)
6. Meng, F., et al.: Chartassisstant: a universal chart multimodal language model via chart-to-table pre-training and multitask instruction tuning (2024)
7. Moured, O., Alzalabny, S., Schwarz, T., Rapp, B., Stiefelhagen, R.: Accessible document layout: An interface for 2D tactile displays. In: Proceedings of the 16th International Conference on PErvasive Technologies Related to Assistive Environments, pp. 265–271 (2023)
8. Moured, O., Baumgarten-Egemole, M., Roitberg, A., Muller, K., Schwarz, T., Stiefelhagen, R.: Chart4blind: an intelligent interface for chart accessibility conversion. arXiv preprint arXiv:2403.06693 (2024)
9. Moured, O., Zhang, J., Roitberg, A., Schwarz, T., Stiefelhagen, R.: Line graphics digitization: a step towards full automation. In: Fink, G.A., Jain, R., Kise, K., Zanibbi, R. (eds.) ICDAR 2023. LNCS, vol. 14191, pp. 438–453. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-41734-4_27
10. of North America, B.A.: Guidelines and standards for tactile graphics (2010). https://www.brailleauthority.org/tg/web-manual/index.html
11. Paterson, M.: Seeing with the hands': Blindness, touch and the enlightenment spatial imaginary. Br. J. Vis. Impairment, 52–59 (2006)
12. Tang, B.J., Boggust, A., Satyanarayan, A.: Vistext: a benchmark for semantically rich chart captioning. arXiv preprint arXiv:2307.05356 (2023)
13. Xia, R., et al.: Chartx & chartvlm: a versatile benchmark and foundation model for complicated chart reasoning (2024)

A Real-Time Chart Explanation System for Visually Impaired Individuals

YooJeong Song[1], SungHeon Jeong[2], WooJin Cho[1], Soon-Bum Lim[3], and Joo Hyun Park[3](✉)

[1] Semyung University, Jecheon-si, Korea
[2] University of California, Irvine, CA 92617, USA
[3] ICT Convergence Research Institute, Sookmyung Women's University, Seoul, Korea
park.joohyun@sookmyung.ac.kr

Abstract. This research addresses the critical need for real-time chart captioning systems optimized for visually impaired individuals in the evolving era of remote communication. The surge in online classes and video meetings has increased the reliance on visual data, such as charts, which poses a substantial challenge for those with visual impairments. Our study concentrates on the development and evaluation of an AI model tailored for real-time interpretation and captioning of charts. This model aims to enhance the accessibility and comprehension of visual data for visually impaired users in live settings. By focusing on real-time performance, the research endeavors to bridge the accessibility gap in dynamic and interactive remote environments. The effectiveness of the AI model is assessed in practical scenarios to ensure it meets the requirements of immediacy and accuracy essential for real-time applications. Our work represents a significant contribution to creating a more inclusive digital environment, particularly in addressing the challenges posed by the non-face-to-face era.

Keywords: Accessibility · Chart Explanation · Real-time architecture

1 Introduction

With the increasing trend towards remote interaction, such as video conferencing and online education, the use of visual data like charts has become ubiquitous. However, this shift poses unique challenges for visually impaired individuals, who are often excluded from real-time visual data interactions. The central focus of our research is the optimization of a real-time chart captioning system designed to address these challenges. Our approach involves developing an AI model capable of not just interpreting and captioning charts but doing so in a real-time context, thus ensuring timely and accessible information delivery. This paper outlines the development process, the challenges in optimizing for real-time application, and the evaluation methods employed to assess the system's

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 306–312, 2024.
https://doi.org/10.1007/978-3-031-62846-7_37

effectiveness. We emphasize the importance of real-time processing to provide immediate accessibility and comprehension of visual data, a crucial aspect in interactive online settings. Our research is a step towards ensuring that visually impaired individuals can actively participate in and benefit from the digital revolution in communication and education.

2 Related Work

There are two primary methodologies for chart interpretation and analysis. The first approach leverages Question Answering (QA) datasets, such as ChartQA [11], PlotQA [12], and FigureQA [6]. In this method, detailed information about charts is obtained through a QA format, enabling selective extraction of information based on user queries. Models like Deplot [9] and MatCha [10] optimize this process, utilizing QA learning algorithms to effectively interpret these datasets. The second approach employs traditional machine learning captioning techniques. This involves the use of image-text pair datasets such as CoCo [8], Chart-to-Text [7], and VisText [14]. For the interpretation of charts, this method integrates Optical Character Recognition (OCR) models with transformer-based models, facilitating a comprehensive understanding and description of chart contents.

Object detection have been achieved significant advancements after the introduction of models such as YOLO [13] and Fast R-CNN [2]. Notably, the YOLO family has seen continuous evolution to support multi-task operations including segmentation, depth estimation, and object tracking. These enhancements have solidified the utility of YOLO-based models in a wide array of applications, primarily due to their superior processing speed compared to models like Faster R-CNN. Despite their widespread adoption, there exists a growing necessity for these models to support specific tasks, such as captioning, to broaden their applicability. This requirement underscores the need for ongoing research and development within the YOLO framework to extend its functionality beyond traditional object detection tasks.

3 Proposed Method

In this research, we have developed a captioning model that necessitates extensive parallel processing due to its complex encoder and decoder structure. This complexity poses challenges for effective utilization on personal laptops or desktops without high-performance computing capabilities. To address this, we propose the adoption of a client-server architecture, utilizing cloud servers for real-life applications. This architecture is designed to be suitable for real-world usage scenarios. Specifically, we adopt a dual-structure approach, termed the 'detection-captioning' structure. In this setup, a detection model is placed ahead of the captioning model. The detection model first identifies charts before activating the captioning model Fig. 1. This approach not only prevents overloading the cloud server but also increases the probability of accurately responding to

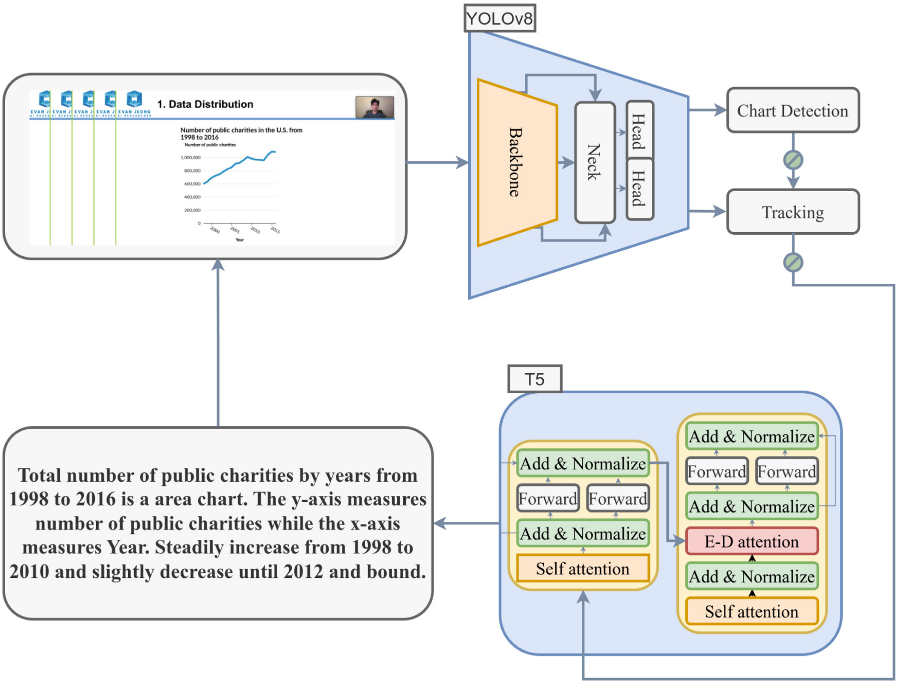


Fig. 1. Overall system architecture: series of frame-based images are inputted, they undergo detection and tracking processes within the YOLO architecture. Following this, a switch activates the T5 model, which then generates captions for the frames. These captions are subsequently displayed on the frames.

actual charts. By reducing unnecessary captioning tasks through the detection model, server resources are utilized more efficiently, providing faster and more accurate results to users. This structure is expected to play a significant role in the practical and effective implementation of a cloud-based captioning system regardless of local computer.

3.1 Detection

The captioning model lacks the ability to discern Out-of-Distribution information, necessitating unnecessary computational costs. This is particularly problematic when using large Language and Vision models, which have significant computational costs, thus failing to guarantee real-time features. Furthermore, given that charts constitute less than 10% of the content in online lectures, a two-step 'detection-captioning' approach can reduce computational costs by nearly tenfold.

3.2 Tracking

Continuously generating captions for detected charts leads to wasteful computational efforts. Therefore, a method that avoids regenerating captions for already processed charts is necessary. This is achieved by tracking videos and generating unique IDs for each chart. The tracking model shares its backbone with the detection model and uses a fine-tuned YOLOv8 [3].

3.3 Captioning

For the captioning component, we have selected the T5 [1] family's VisuaLanguagelT5 model, which operates by extracting features through the visualization embedding of images and generating captions. This model has achieved high performance by directly enhancing the understanding of images through additional inputs like scene graphs and tables, beyond raster images. However, despite its high performance, the model shows weaknesses in real-world applications due to its large size, leading to low computational efficiency.

This system is designed to optimize for the visually impaired, focusing on converting visual information into text. It adopts traditional machine learning methodologies for captioning, emphasizing real-time application and accessibility. The system's architecture, comprising detection, captioning, and tracking components, is tailored to efficiently process and interpret visual data, particularly in scenarios with limited computational resources.

4 Evaluation

4.1 Dataset

The Chart-to-Text and VisText datasets, as previously mentioned, are composed of image-text pairs, consisting of charts (images) and their corresponding captions (texts). These datasets are organized into various sub-categories of charts, such as bar charts and line charts, which influences the balance of data. Notably, categories that are more frequently utilized tend to dominate these datasets, potentially introducing bias. Additionally, the theme of each dataset varies across data points, posing potential challenges for encoders in the data extraction process. Given these characteristics, it is imperative to employ both datasets in the training process. This approach ensures the diversification of data and minimizes bias, thereby optimizing the performance of encoders. Our findings underscore the importance of leveraging multiple datasets to enhance the robustness and accuracy of models designed for chart interpretation and captioning.

4.2 Experiment

In the context of chart detection and tracking, we utilized the VisText dataset to generate bounding boxes, thereby creating a dataset tailored for object detection. This dataset was subsequently employed to train a YOLOv8 model, with a learning rate of 1e-3 and for a duration of 5 epochs. Post-training, charts were identified using a confidence threshold of 0.25 and an Intersection over Union (IoU) threshold of 0.7 to generate bounding boxes. Each detected object was then assigned a unique identifier to facilitate tracking. Our study presents the potential for real-time system capabilities by improving system efficiency through the integration of object detection and tracking, thereby reducing the total computational cost. This was experimentally validated using videos from references [5] and [4]. The computational cost incurred by employing the captioning model in isolation was approximately tenfold higher than that of the system augmented with the detection model. Furthermore, the inclusion of the tracking model in the system demonstrated an approximate eightfold improvement in efficiency compared to the detection-only system. These findings underscore the significant enhancements in system performance and efficiency that can be achieved through the strategic incorporation of object detection and tracking modules, highlighting the feasibility of deploying such a system in real-time applications (Table 1).

Table 1. Total Inference Time(s) for videos with difference system architecture

	Captioning	Detection + Captioning	Tracking + Detection + Captioning
Video- [5]	4219	391	29
Video- [4]	8103	803	120

4.3 Extrinsic Evaluation

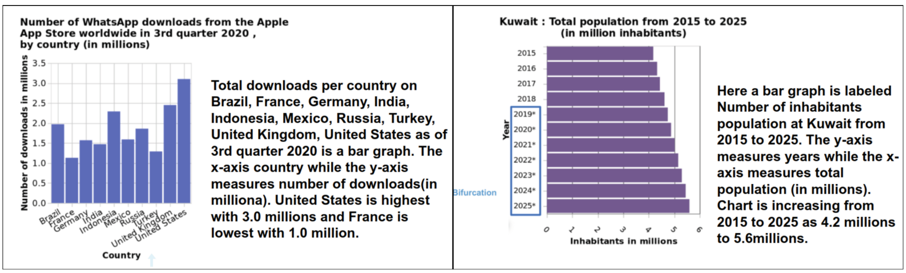


Fig. 2. Example of captioning by model in ChartQA dataset

The comprehensive nature of explanations generated by current systems Fig. 2 may inadvertently pose challenges for visually impaired individuals in real-time applications by details. This observation suggests that a system's effectiveness is not solely determined by the granularity of its explanations. This characteristic necessitates the development of systems that are inherently user-centric, taking into account the diverse needs and preferences of their users. Such an approach could involve strategies like emphasizing key points to enhance the understandability and practicality of captions for all users, including those with visual impairments. This shift towards simplicity and emphasis on critical information could significantly improve accessibility and user experience in assistive technologies.

5 Conclusion

The development of systems that not only focus on the performance of artificial intelligence models but also consider practical aspects for real-world application is crucial across various domains. This is particularly pertinent in the context of platforms designed for visually impaired users, underscoring the necessity for a deep understanding and thoughtful application of their unique requirements. Our research represents a significant step in this direction. By fundamentally comprehending the needs of such platforms and employing state-of-the-art technologies, we have successfully constructed a system that operates efficiently through the implementation of streamlined algorithms. This approach demonstrates the potential of combining advanced AI with user-centric design to create solutions that are both technologically sophisticated and practically applicable in enhancing the lives of visually impaired individuals.

However, our research was evaluated through quantitative metrics. Considering our model as foundational, we leave the inclusion of qualitative tests and the enhancement of features such as voice output as potential directions for future research.

Acknowledgments. This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT)(RS-2023-00211205, RS-2022-00165818).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Colin, R.: Exploring the limits of transfer learning with a unified text-to-text transformer. JMLR **21**(140), 1 (2020)
2. Girshick, R.: Fast r-cnn. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1440–1448 (2015)
3. Jocher, G., Chaurasia, A., Qiu, J.: Yolo by ultralytics (version 8.0. 0)[computer software]. YOLO by Ultralytics (Version 8.0. 0)[Computer software] (2023)
4. with Josh Starmer, S.: A gentle introduction to machine learning (2019). https://youtu.be/Gv9_4yMHFhI?si=3XFGt3M2lRWj8pnb

5. with Josh Starmer, S.: Machine learning fundamentals: Cross validation (2019). https://www.youtube.com/watch?v=fSytzGwwBVw&list=PLblh5JKOoLUICTaGLRoHQDuF_7q2GfuJF&index=2
6. Kahou, S.E., Michalski, V., Atkinson, A., Kádár, Á., Trischler, A., Bengio, Y.: Figureqa: An annotated figure dataset for visual reasoning. arXiv preprint arXiv:1710.07300 (2017)
7. Kantharaj, S., et al.: Chart-to-text: a large-scale benchmark for chart summarization. arXiv preprint arXiv:2203.06486 (2022)
8. Lin, T.-Y., et al.: Microsoft COCO: common objects in context. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) ECCV 2014 Part V. LNCS, vol. 8693, pp. 740–755. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-10602-1_48
9. Liu, F., et al.: Deplot: one-shot visual language reasoning by plot-to-table translation. arXiv preprint arXiv:2212.10505 (2022)
10. Liu, F., et al.: Matcha: enhancing visual language pretraining with math reasoning and chart derendering. arXiv preprint arXiv:2212.09662 (2022)
11. Masry, A., Long, D.X., Tan, J.Q., Joty, S., Hoque, E.: Chartqa: a benchmark for question answering about charts with visual and logical reasoning. arXiv preprint arXiv:2203.10244 (2022)
12. Methani, N., Ganguly, P., Khapra, M.M., Kumar, P.: Plotqa: reasoning over scientific plots. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pp. 1527–1536 (2020)
13. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: unified, real-time object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 779–788 (2016)
14. Tang, B.J., Boggust, A., Satyanarayan, A.: Vistext: a benchmark for semantically rich chart captioning. arXiv preprint arXiv:2307.05356 (2023)

Towards Establishing a Description Standard of Mathematical Function Graphs for People with Blindness

Zoé Esmeralda Sobota, Ann-Christin Freese, Thorsten Schwarz(✉), and Karin Müller

Center for Digital Accessibility and Assistive Technologies, Karlsruhe Institute of Technology, Karlsruhe, Germany
{ann-christin.freese,thorsten.schwarz,karin.e.mueller}@kit.edu
https://www.access.kit.edu

Abstract. Visual representations are widely used for conveying information for sighted people. However, these visualizations are often inaccessible to blind people, requiring special adaptations. Especially in higher education, it is vital to support the learning process how to explore graphics. Although the description of graphics is a solution to make them accessible, there are only few standards for descriptions for this target group. In this paper, we develop a standard for describing function graphs at the university level in a user-centered multi-stage design process, focusing on structuring the information for blind readers and guiding authors to include necessary details to avoid errors.

Keywords: Mathematical functions · Graphs · Description standard · Accessible content

1 Introduction

The simplest way to make images accessible to blind people is by descriptions, requiring no special aids such as tactile graphics or haptic devices. While superficial descriptions suffice for everyday graphics, precise descriptions are essential for mathematical content. Mathematical concepts often use visual elements, but in many cases, alternative representations like accessible tables can convey the same information without graphs. However, functional correlations, crucial in mathematics, pose a challenge for blind people. While tactile graphs and sonification are possible solutions, the latter has limitations [4]. Descriptive approaches offer a potential solution, but there is a lack of standardization, especially in the context of function graphs at university level. In this paper, we address this research gap by exploring existing approaches and defining recommendations for people who create textual descriptions of function graphs for blind people.

2 Related Work

It is important that descriptions follow a predefined pattern. This way, people who create the descriptions know what information is relevant and in what order

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 313–321, 2024.
https://doi.org/10.1007/978-3-031-62846-7_38

it should be listed. Those who need this information can find it again without further delay, which makes their work easier. A good pattern also makes it easier to navigate through the document. A literature search showed that only a few sources have dealt with the structuring and order of information in graph descriptions. Two examples are presented below. The VISCH [1] project of the German Institute for the Study of the Blind and the Federal Competence Centre for Accessibility (Deutsche Blindenstudienanstalt e.V. and Bundeskompetenzzentrums Barrierefreiheit e.V.), published guidelines in 2012, offering a method for describing graphics in textbooks, including function graphs. It emphasizes the importance of a fixed order in descriptions but lacks specific guidelines for more complex function graphs, commonly used in university contexts.

Similarly, the DIAGRAM Center's guidelines [6], developed in collaboration with the Carl and Ruth Shapiro Family National Center for Accessible Media, provide limited assistance for describing intricate graphics, including function graphs. The focus is on covering a variety of graphic types, offering brief instructions rather than comprehensive guidance.

As graphs vary in complexity, with different functions, legends, and axis labels, the absence of standardized regulations, coupled with the diverse nature of function graphs, poses challenges for effective descriptions. [2] presented the i-Graph-LITE system that can be used to explore graphs. They designed the system in a user-centered approach by incorporating blind and non-blind experts (statisticians) to identify questions that could be asked about line graphs. They evaluated the usefulness of the system in two cycles. In their studies, they have already identified useful information to help blind people to understand graphs from descriptions. However, they did not answer the question how students at university level discover and understand graphs.

[9] pointed out that the best way to allow a non-visual access to graphics is to provide both a description and a tactile graphic to support the learning process. Since there is no guideline how to describe graphs for students, people who make them accessible have to decide themselves in which order they describe elements, which leads to inconsistent placement of information and possible omissions.

3 Collecting Requirements for Descriptions

To address the current absence of guidelines, we aim at establishing a standardized approach for describing function graphs. This standard should be versatile and cover different types of function graphs, especially those used in higher education. Tactile graphics of course offer further added value to descriptions, but are not considered in this study, as a description should be provided in addition if tactile graphics are provided. It is therefore also relevant here to standardize the descriptions. Furthermore, the capabilities for generating tactile graphics are not always available. This in turn increases the relevance of the descriptions.

In order to understand and gain more experience on how graphs can be described in a teaching context, an iterative study was conducted starting with describing different mathematical graphs. This was conducted with eight sighted participants. Subsequently, descriptions were compared and an initial order was determined. This was tested and improved for navigability with three blind and one visually impaired person. Within this framework, instructions for writing descriptions were developed iteratively. The instructions were then run through again with four sighted participants who deal with image descriptions in general. Finally, the improved guideline was evaluated with three blind participants.

3.1 Exploratory Study for Describing Functions

In a study, we aimed at collecting requirements for graphic descriptions by (1) investigating how participants describe graphs and (2) testing how useful these descriptions are. We recruited eight sighted participants, aged 20 to 25, including students of mathematics, natural science, and from other disciplines, to ensure diverse perspectives. We divided the participants into two groups (Group A and Group B). Each group was given a graphic to warm up and 3 graphics with different levels of difficulty to work on for the study. The difficulty of the illustrations was measured by the complexity of the function graphs shown, the number of functions included, the presence of a legend or axis labels.

Group A and B were asked to describe the graphics assigned to them and then, using the descriptions of the other group, to redraw the respective graphics of the other group based on their description. Figure 1 displays the three graphics that were selected for group A of the study [3,7,8], and Fig. 2 illustrates two drawings of the function graph in Fig. 1(b) based on different descriptions. No specific medium was prescribed for drawing.

3.2 Results and Development of a Standard

In a further analysis, the accuracy of the drawings and the completeness of the information in the 32 descriptions provided by the participants was reviewed. Additionally, 12 other descriptions from literature and examples provided by our service center were analyzed. The analysis showed that a description following the subsequent fixed structure would be suitable:

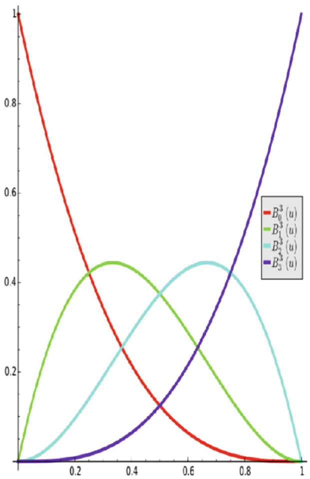
1. metadata, 2. coordinate system, 3. environment, 4. function names, 5. description of the functions and *6. other notable characteristics.*

Metadata contains information about the type of graphic and, if available, its title and caption.

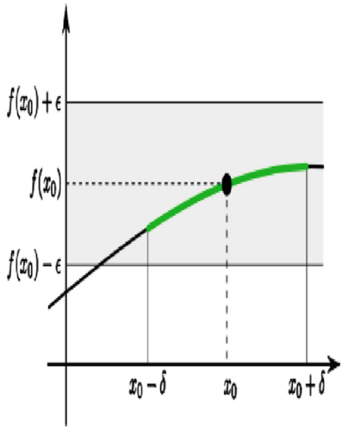
Coordinate system describes the depicted area and provides information about the axes.

Environment is only required if the graphic uses a function to represent a theoretical, mathematical concept. Objects placed around the function can be described there.

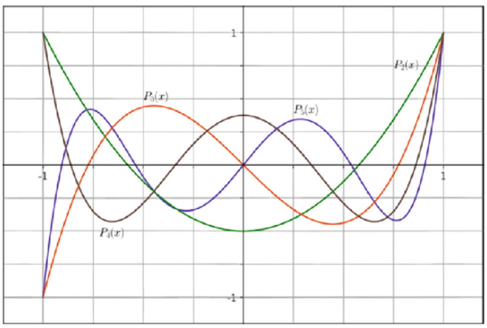
The names of the functions should be listed under **function names** and, if they are very similar, a distinguishing feature of the function, such as color, should also be mentioned. The function should then be described from left to right in the **description of the functions**. Finally, information that has not yet been mentioned can be described under **other notable characteristics**.



(a) 4 Bernstein polynomials [3]

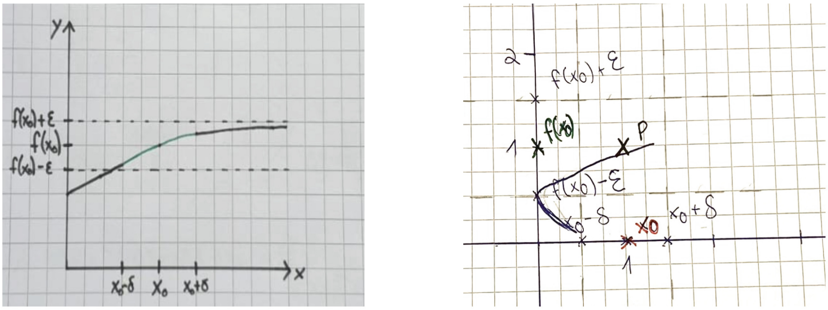


(b) Epsilon delta criterion [7]



(c) 4 Legendre polynomials [8]

Fig. 1. Graphics in group A that increase in difficulty of description from (a) to (c)



(a) Drawing that corresponds to the original **(b)** Drawing that differs from the original

Fig. 2. Drawings of the Epsilon delta criterion based on two descriptions

4 First Evaluation of the Description Standard

Following the establishment of a suitable order, we conducted an evaluation with one visually impaired and three blind persons to enhance standardized descriptions, focusing on improving accessibility through navigation. The description was structured into chapters with specific categories as headings to facilitate quick information retrieval or navigation between essential description elements.

For efficient location of important coordinates, a separate list of coordinates for each function was suggested, including details like maxima or zeros. Subheadings for each function and its coordinate list were recommended for easy reference. Emphasis was placed on mentioning if coordinates were estimated and ensuring consistency, especially in cases of axis symmetry. To enhance the usability of the description, a guide emphasizing the use of appropriate, factual language was developed, providing detailed explanations of the standardization notes.

4.1 Example

The following example shows the developed compact standardized description of Fig. 3 based on the structure suggested in 3.2 [5]. **3. environment** is omitted because the graphic has scaled axes. This means that concrete coordinates can be created using the axis values and that it is not necessary to use relations to other objects for the description. If we consider points **4. function names** and **6. other notable characteristics** from the explanation under 3.2, it becomes clear that these are also redundant, as the function depicted is already precisely described by points **1. metadata**, **2. coordinate system** and **5. description of the function** and no aspects worth mentioning remain open.

1. metadata:
The figure is a function plot.

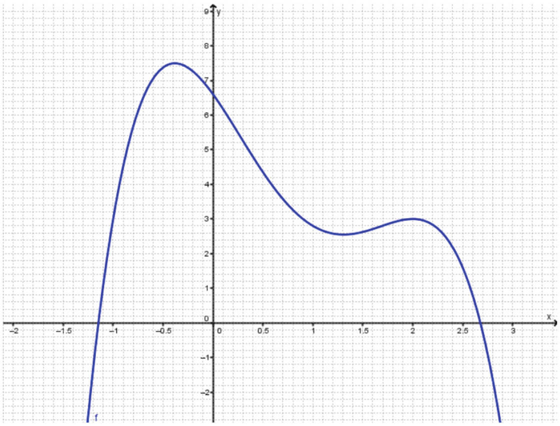


Fig. 3. Function graph in a 2D coordinate system. The x-axis is shown from −2 to 3 and the y-axis from −2 to 9. A blue function $f(x)$ =-0.96$x^4 + 3.8x^3 - 2.7x^2 - 3.9x + 6.6$ is depicted. It has intersections with the x-axis at approximately −1.15 and 2.71 and a y-intercept of 6.6. The global peak of this function is at (−0.38|7.46). (Color figure online)

2. coordinate system:
All four quadrants are shown. The x-axis represents values from −2 to 3 and the y-axis from −2 to 9.

5. description of the function:
A function f is shown.
function f: The function f starts at approximately (−1.3|2.4). It rises steeply, intersects the x-axis for the first time at approximately (−1.15|0) and reaches a global maximum at approximately (−0.4|7.5). Then it falls monotonically, intersects the y-axis at approximately (0|6.6) and has a minimum at approximately (1.3|2.6). From then on, the function rises a little, has another maximum at (2|3) and intersects the x-axis one last time at approximately (2.7|0) before it ends at around (2.9|−2.4).
function f coordinates: The following coordinates are estimated:
- starting point (−1.3|2.4)
- intersection with x-axis (−1.15|0)
- global maximum (−0.4|7.5)
- intersection with y-axis (0|6.6)
- minimum (1.3|2.6)
- maximum (2|3)
- intersection with x-axis (2.7|0)
- end point (2.9|−2.4)

4.2 More Complex Examples

We also investigated the question if the developed standard can be applied to more complex examples as shown in Fig. 1.

Figures 1 (a) and (c) show multiple functions. In order to be able to work with the large number of functions, the names of the functions should first be listed in **4. function names**. Then, in **5. description of the functions**, each function has its own section, consisting of a continuous text description and a subsequent list of coordinates, as in the example in the previous section. Under **6. other notable characteristics**, the symmetry can be noted for Fig. 1(c) or a note can be given that all functions pass through the origin.

When analyzing Fig. 1(b), we found that due to the lack of labeling the axes (steps, numbers), **3. environment** can be used to describe the elements in relation to each other.

1. metadata:
The illustration shows a function plot.
2. coordinate system:
Mainly the first quadrant is shown, but parts of the other quadrants are also visible. No units are shown on the axes. The positive x-axis contains the equidistant points $x_0 - \delta$, x_0 and $x_0 + \delta$. The equidistant points $f(x_0) - \epsilon$, $f(x_0)$ and $f(x_0) + \epsilon$ are shown on the positive y-axis.
3. environment:
Starting from the y-values $f(x_0) - \epsilon$ and $f(x_0) + \epsilon$ there is a horizontal line in each case. The area between these two lines is highlighted in gray.
5. description of the function:
A nameless function is shown.
function: The function starts in the second quadrant and rises strictly monotonically in a slight right-hand curve. It intersects the y-axis below the value $f(x_0) - \epsilon$, continues to rise and enters the area highlighted by the y-values $f(x_0) - \epsilon$ and $f(x_0) + \epsilon$ to the left of $x_0 - \delta$. The function continues to increase strictly monotonically up to the point $(x_0|f(x_0))$. This point is marked by a circle on the function graph. Dashed lines, one vertical from the x-value x_0 and one horizontal from the y-value $f(x_0)$, lead to this point. To the right of this point, the function increases monotonically and remains in the area highlighted by the y-values. The function ends there after the x-value $x_0 + \delta$. Starting from the values $x_0 - \delta$ and $x_0 + \delta$ on the x-axis, there are two vertical lines that rise up to the function graph. The section of the function graph that lies between these x-values is highlighted in color.
function coordinates:

- Starting point: in the second quadrant
- Highlighted point: $(x_0|f(x_0))$
- End point: in the first quadrant, within the area highlighted by the y-values $f(x_0) - \epsilon$ and $f(x_0) + \epsilon$

6. other notable characteristics:
The progression of the function is reminiscent of the progression of a logarithm.

5 Final Evaluation of the Developed Standard

In a second study, the guide for creating descriptions was evaluated with four sighted participants and the description standard with three blind participants. The sighted participants already had experience in creating descriptions. They coped well with the guidelines and could imagine working with them in the future. However, they would have liked information on levels for headings. The blind participants were between 22 and 27 years old and studying a STEM subject at university. The naming of important characteristics, the order and the type of information were positively noted. In graphics that contain several functions, the sequential description of the individual functions was found to be easy to understand. The separation of the important points as a list was particularly positively noted, as this avoids having to laboriously search for information. In addition, the list allows the structure of the function to be recalled after listening to the continuous text description.

6 Conclusions

In this work, we collected requirements for describing graphs from STEM students in an exploratory user study. We then developed a first version of a description standard, which was evaluated with the target group in two iterations. The contribution of this work is a description standard for compact graph descriptions created in a user-centered design process.

In order to standardize function graph descriptions, it is essential that the results are used in practice and further specified, improved and harmonized in the way they are formulated. Therefore, in the next step, the evaluations are to be extended further and the opinions of other people from a wide range of educational and training areas are to be taken into consideration in order to achieve further harmonization and standardization of the descriptions in detail. We also aim at transferring the results to other graphs.

Ultimately, it is also important to use the standard. In this way, a standard can become established and provide added value for users and also a good basis for future artificial intelligence (AI) solutions and the standardized descriptions they create. Using the standard will make it possible to identify more precisely which formats are suitable for creating a successful description and how these can be adapted more precisely to the respective or broader needs. Regular use also means that people with visual impairments know how information is structured and can work more efficiently and quickly with mathematical graphs thanks to the recognition factor. This also efficiently increases the learning effect. As mentioned at the beginning, this work focuses on the benefits of function graph descriptions. One of the purposes is to use this information as a supplement to tactile graphics, but also to find relevant descriptions more quickly and easily within the respective literature. It is not intended to replace tactile graphics, but a combination of both variants is one of the most efficient ways to understand and work with mathematical graphs.

Future work could include automatic generation of function graph descriptions by AI. Unfortunately, image recognition with AI is not yet advanced enough to always describe function graphs correctly. Instead of a full automation, a partial automation would also be possible. For example, a description template could be created that defines an order and the required information. A sighted user could be used to fill in this information and check generated information for correctness. Partial automation would be easier to implement and would already provide added value, as the person creating the description would be supported in writing it.

References

1. Deutsche Blindenstudienanstalt e.V., Bundeskompetenzzentrum Barrierefreiheit e.V.: VISCH - Visualisierte Informationen in Schulbüchern zugänglich machen (2012). https://www.blista.de/sites/default/files/visch-leitfaden.pdf
2. Ferres, L., Lindgaard, G., Sumegi, L., Tsuji, B.: Evaluating a tool for improving accessibility to charts and graphs. ACM Trans. Comput. Hum. Interact. (TACCESS) **20**(5), 1–32 (2013)
3. Kilom691: https://commons.wikimedia.org/wiki/File:Bernstein_poly_3.svg. Accessed 26 Jan 2024
4. Kim, J., Lee, Y., Seo, I.: Math graphs for the visually impaired: audio presentation of elements of mathematical graphs. In: Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems, pp. 1–6 (2019). https://doi.org/10.1145/3290607.3308452
5. Lachner, B.: Funktionswerte aus funktionsgraph ablesen. https://www.geogebra.org/m/a3vwtswa
6. NCAM, DIAGRAM Center: http://diagramcenter.org/table-of-contents-2.html. Accessed 14 Sept 2023
7. Serlo Education e.V.: Analysis Eins - Mathe für Nicht-Freaks. Serlo Education e.V. (2017). https://mathe-builds.serlo.org/analysis1/analysis1.pdf
8. Shishirdasika: https://commons.wikimedia.org/wiki/File:Legendre_Polynomials.svg. Accessed 26 Jan 2024
9. Zebehazy, K.T., Wilton, A.P.: Quality, importance, and instruction: the perspectives of teachers of students with visual impairments on graphics use by students. J. Visual Impairment Blindness **108**(1), 5–16 (2014)

ACCSAMS: Automatic Conversion of Exam Documents to Accessible Learning Material for Blind and Visually Impaired

David Wilkening[1], Omar Moured[1,2](✉), Thorsten Schwarz[1], Karin Müller[1], and Rainer Stiefelhagen[1,2]

[1] ACCESS@KIT, Karlsruhe Institute of Technology, Karlsruhe, Germany
mail@david-wilkening.de, moured.omar@gmail.com
[2] CV:HCI@KIT, Karlsruhe Institute of Technology, Karlsruhe, Germany
https://gitlab.access.kit.edu/David/accsams/

Abstract. Exam documents are essential educational materials for exam preparation. However, they pose a significant academic barrier for blind and visually impaired students, as they are often created without accessibility considerations. Typically, these documents are incompatible with screen readers, contain excessive white space, and lack alternative text for visual elements. This situation frequently requires intervention by experienced sighted individuals to modify the format and content for accessibility. We propose ACCSAMS, a semi-automatic system designed to enhance the accessibility of exam documents. Our system offers three key contributions: (1) creating an accessible layout and removing unnecessary white space, (2) adding navigational structures, and (3) incorporating alternative text for visual elements that were previously missing. Additionally, we present the first multilingual manually annotated dataset, comprising 1,293 German and 900 English exam documents which could serve as a good training source for deep learning models.

Keywords: Document Analysis · Accessible Exams · Layout Segmentation

1 Introduction

The availability of accessible course material is often a critical barrier for students with blindness and visual impairment (BVI) [4]. An important type of learning material is past exams, which give students the opportunity to understand the level of knowledge expected in an exam and practice their acquired knowledge [6]. However, it is often not possible for BVI students to directly access these exam documents, often requiring sighted individuals to carefully modify the format and content to ensure compatibility with screen readers.

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 322–330, 2024.
https://doi.org/10.1007/978-3-031-62846-7_39

Documents are accessed primarily by BVI individuals using various assistive technologies, depending on their preference and availability [5]. Such documents need to follow guidelines [2] to ensure accessibility. However, **the majority of exams are not created as part of an ongoing course in which BVI students participate. Therefore, they are usually created without considering accessibility.** To address these challenges, we developed a semi-automatic AI-based method that aids in converting exam materials into an accessible format. This involves identifying key content, solutions, adding navigational structures, and including missing alternative texts. Our contributions include: (1) Identifying a preferred exam format for BVI students through interviews, (2) Developing a system for efficient document conversion, and (3) Creating the first multilingual dataset of 1,293 German and 900 English exams.

2 Related Work

Making already existing documents accessible after their creation is ongoing research with many subtasks. *SciA11y* [7] is a system specifically developed to convert *PDFs* of scientific papers into accessible *HTML* formats. It transforms multicolumn documents into a linear, one-dimensional layout that screen readers can easily process by creating *HTML* header tags for each section title. Furthermore, *SciA11y* generates accessible hyperlinks for citations and references, facilitating seamless navigation for BVI users. *Nougat* [1] is a transformer encoder-decoder model that converts document images into markdowns in an end-to-end manner. Contrary to *SciA11y* it is also capable of transforming math equations and simple tables. However, (1) both are only tailored for scientific publications and (2) do not address the challenges associated with providing alt-text for all types of content blocks.

Inspired by *SciA11y*, we built a pipeline to convert existing documents into an accessible format. We made modifications to take into account the domain specifics of exam documents and their target use as learning materials. We also address some shortcomings, such as support for mathematical equations in text.

3 Need-Finding Interviews

The adaptation of exam materials into different formats is not just for accessibility but also practiced by KIT's departmental student representatives. A survey[1] among KIT's math and computer science students has revealed that 79.9% of the respondents prefer a reconfiguration of document layouts to eliminate superfluous white space, thus reducing both the cost and paper consumption associated with printing (n = 369). 76.8% of the participants prefer to relocate the sets of solutions from their original placement within the questions to a consolidated section at the end of the document (n = 362).

[1] https://www.fsmi.uni-karlsruhe.de/Fachschaft/Umfrage/.

We conducted interviews to identify how students with BVI use exam documents as learning materials. These interviews were instrumental in determining the type of semantic information that ACCSAMS needs to extract from document files and how the results should be presented to serve as effective learning resources for students with BVI. They served to expand on our knowledge from our experience as a Disability Support Office (DSO) in making documents accessible and remixing exam documents for students with vision in the departmental student representations.

3.1 Participants

The interviews were conducted with two blind male graduate students in computer science and physics. Both make use of the document conversion service for BVI at KIT.

3.2 Procedure

We asked participants the following three questions: (1) Which exam format do you prefer (e.g., *PDF, HTML, WORD*, Markdown) and why? (2) How do you typically interact with exam documents for the purpose of exam preparation? By annotating a separate document or directly editing the file? (3) What layout find most effective for navigating the questions and answers? Each question followed immediately by its answer, or all questions first followed by all answers?

3.3 Findings

Both participants preferred inline editing formats like Word and Markdown over tagged PDFs. They reported difficulties with navigation due to the lack of screen reader-friendly hierarchy in exam documents, which limits quick movement between questions and answers, and noted the absence of alt-text for figures and equations. Regarding layout, opinions differed: one favoured separate documents for solutions, while the other preferred solutions under each question. In response, we outlined design steps to enhance exam accessibility in our UI, including (1) analyzing document layout to identify content blocks and solutions, (2) extracting document hierarchy to support screen reader navigation, (3) adding alt-text to visual elements, and (4) ending with the option to export the document in either Markdown or Word format with the option to provide alternative reading orders by repositioning solutions.

4 Creation of the Exam Documents Dataset

We collected exam materials from the Common Crawl project[2], a public archive of web data. We analyzed eight datasets, released between mid 2022 and the

[2] https://commoncrawl.org/.

end of 2023, extracting *PDFs*, which include “exam” or “klausur” as a keyword in the URL. Positive matches were downloaded. The text layer of the locally available *PDFs* was utilized to refine the data set by iteratively developing a list of common false positive keywords (such as “exam schedule”). In addition, a keyword search was used to suggest a study field for each document.

Subsequently, the remaining documents were manually filtered. Furthermore, we identified whether an exam originated from higher education (university, college or similar) or another source. Unique types of exams (such as entrance exams) were included if they shared a similar format as other exams. As a result, we obtained a collection of 1,293 German (361 computer science, 337 physics, 216 mathematics, 175 economics, 67 law, 33 electrical engineering, 32 mechanical engineering) and 900 English exams (341 computer science, 239 chemistry, 150 math, 124 economics) in native *PDF* format, originating from 281 different domains.

5 The ACCSAMS System

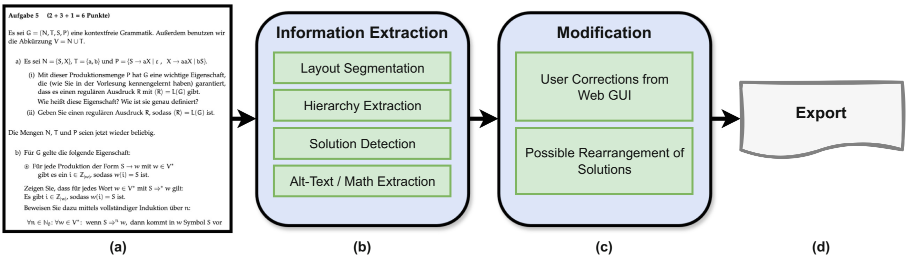


Fig. 1. The ACCSAMS pipeline: (a) Exam page. (b) Automated processing (c) User application of corrections and repositioning of solutions (d) Output to accessible markdown or conversion to WORD format.

The system is based on 3 steps: (1) segmenting the layout of every page, extracting all relevant content blocks, while establishing if they are specific to a solution, (2) determining the reading order and hierarchy between the content blocks, and (3) generating alt-texts automatically or requesting them from the user for figures and tables through our Web GUI.

5.1 Layout Segmentation

The layout segmentation module analyzes the document and splits it into content blocks using YOLOv8[3] as the AI detection model. We first pre-trained it on DocLayNet [3] a dataset of a diverse collection of 80,863 pages from various

[3] https://github.com/ultralytics/ultralytics.

domains. We then fine-tune it on our own exam dataset. For this, we annotated 1917 pages and split them into a training and evaluation corpus (80/20 ratio). The classes and the model results are in Table 1.

5.2 Hierarchy Extraction

To enhance navigation accessibility for screen reader users, extracting and organizing the document's hierarchy is essential. We achieve this by structuring content blocks into a coherent tree structure, applying two heuristic rules. The first rule involves sorting content blocks by their location within the document, using page number and vertical and horizontal coordinates as sorting criteria. The second rule focuses on categorizing list items and headings. This categorization is based on the content's textual characteristics and enumeration styles, such as Roman, numeric, alphabetic, or heading levels. By implementing these strategies, we construct a tree structure that accurately represents the hierarchy of the document.

5.3 Solution Detection

In order to facilitate the relocation of solutions to different parts of a document or to insert hints for screen readers, it is necessary to identify the solutions accordingly. To achieve this, we follow two rules. First, if a content block is a heading and includes keywords such as "solution" or "answer", both the heading itself and all its descendants in the subtree are marked as solution specific content blocks. Our layout segmentation model is trained to not incorrectly identify text elements such as "your answer: ______" as headings. Second, if a content block contains color, it is also identified as a solution.

5.4 User Interface

The web application was implemented to assist users in individual steps to enhance the accessibility of exam documents. These steps include: (1) Upload an exam, (2–5) verify the extracted content blocks, question/solution assignment, alt text, and hierarchy identified by the pipeline, and finally, (6) export the exam to the desired layout and file format.

6 Evaluation

To evaluate the detection of relevant content blocks, we have annotated 1917 pages from our exam dataset and split them into a training (1,533) and evaluation (384 pages) corpus (80/20 ratio). The results can be seen in Table 1.

6.1 System Evaluation

For reading order and hierarchy extraction, we manually annotated 26 different documents originating from different domains to ensure a variety of authors and styles. The evaluation of the reading order was done using the Average Relative Distance (ARD) metric according to [8]. We achieve an average AVD (per complete document) of 6.23 with a standard deviation of 10.05. The hierarchy is evaluated on the same 26 documents by comparing the hierarchy level of each content block. We measure this with an average Euclidean distance of 0.25 (standard deviation 0.43) and for their relative (between two consecutive content blocks in reading order) distance of 0.09 (standard deviation 0.08).

Table 1. Detection model performance evaluated on 384 test pages.

Classes	# Instances	Precision	Recall	mAP50	mAP50-95
Headings	446	0.884	0.948	0.956	0.86
Paragraphs	3475	0.944	0.919	0.965	0.875
List symbols	1933	0.99	0.936	0.987	0.747
Figures	246	0.815	0.878	0.898	0.814
Formulas	271	0.8	0.764	0.848	0.774
Tables	75	0.811	0.88	0.878	0.852
All	6446	**0.874**	**0.887**	**0.922**	**0.821**

6.2 User Study of User Interface

To evaluate the end-user web application of ACCSAMS, we successfully enlisted 4 people with prior experience in transforming examination documents into study resources. Two of these participants are employed at the ACCESS@KIT literature conversion service. The remaining two are members of the joint departmental student representation for the mathematics and computer science faculties.

Procedure. Upon completion of the consent form and the provision of demographic data, the participants were invited to participate in individual online think-aloud sessions. Using their own personal devices, they participated in the sessions using the Firefox or Chrome web browser. An elucidation of the purpose of the application, accompanied by a demonstration of the system, was presented. Subsequently, the participants received a simple two-page math examination document with questions as an initial warm-up exercise. This was followed by a physics examination, comprising 19 non-empty pages, with different documents of questions and solutions. The session culminated with the opportunity for participants to engage in open-ended questions, allowing them to reflect on their experiences and draw parallels with existing tools with which they are familiar for document conversion. In conclusion, participants received a System Usability Scale (SUS) questionnaire.

Results. The system received an average SUS score of 84. All participants agreed that the system is an improvement over the tools they are currently using for document conversion.

All participants expressed the wish for immediate feedback on the impact of their actions. One participant requested a comprehensive view of all subtask predictions, such as layout segmentation, solution detection, hierarchy extraction, and alt texts, with the option to correct inaccuracies in the specific sub task editor. A trio of users shared a common observation that a meticulous review of the entire exam content is required at each stage to validate the predictions, an exercise that rendered the process somewhat daunting and repetitive.

Two participants were surprised by the need to manually input metadata such as the title, duration, and date of the exam. They felt that grading points are significant in layout segmentation. One participant typically places these points in the task's heading, aiding screen reader users in navigating the table of contents. Both participants were unsure about how to handle points for specific anticipated keywords in a student's response. ACCSAMS does not dictate how to handle this, allowing users to incorporate points into `heading` or `text` content blocks as they see fit. One user chose not to label these blocks, excluding them from the output.

7 Discussion and Limitations

Intelligent Features: Using predictive algorithms to divide document conversion into subtasks reduces active editing and simplifies the conversion process. However, this method requires a thorough review of the document for each subtask. Each subtask editor increases cognitive load, leading to user fatigue. Our application's framework mirrors the ACCSAMS pipeline's steps. Improving the repetitive "review and correct" process in the user interface could enhance user satisfaction in future versions.

Handling Grading Points. Grading points were not seen as special elements of examination documents. Their information is seamlessly integrated within the concept of `heading` and `text` content blocks without the need for specialized processing. However, users identified these as distinct elements and experienced uncertainty in dealing with them, given the lack of specialized recognition in the user interface.

Specific Question Formats. ACCSAMS does not explicitly cater to question formats such as *multiple choice* or *fill in the blanks*. Although the generic nature of ACCSAMS and its user interface allows the conversion of these documents, a more structured and automated methodology would be advantageous to guide users.

Need-finding User Study. Our initial investigations, although encouraging, were conducted with a limited number of users and lacked comprehensive control measures. Future studies should involve a larger and more diverse group of BVI individuals. Particularly in light of representing grading points, and dealing with specific question formats such as multiple choice and fill-in texts, there is a need for additional research besides applying generic guidelines for document accessibility.

8 Conclusion

In this work, we introduce ACCSAMS, a pipeline and user interface designed to facilitate efficient transfer of exams to accessible learning materials for individuals with or without BVI. By leveraging AI models along with heuristic rules, we empower sighted individuals to seamlessly enhance the accessibility of exam documents. We believe that this initiative is a first step to encourage other researchers to explore additional features aimed at improving the conversion process. Future work involves further integration of automatic alt-text generation for figures and tables in the pipeline, as well as specialized alt-text editors into the GUI for the various content blocks.

Acknowledgments. The authors thank Viola Buck Cabrera and Iva Andreeva for the implementation of the end-user web application. In addition, we thank the HoreKa computing cluster at KIT for the computing resources used to conduct this research.

Disclosure of Interests. This research was funded by the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreements no. 861166.

References

1. Blecher, L., Cucurull, G., Scialom, T., Stojnic, R.: Nougat: neural optical understanding for academic documents (2023). https://doi.org/10.48550/arXiv.2308.13418
2. Cooper, M., Adams, C., Campbell, A., Montgomery, R.B., Kirkpatrick, A.: Web content accessibility guidelines (WCAG) 2.2, October 2023. https://www.w3.org/TR/2023/REC-WCAG22-20231005/
3. Pfitzmann, B., Auer, C., Dolfi, M., Nassar, A.S., Staar, P.W.J.: DoclayNet: a large human-annotated dataset for document-layout analysis (2022). https://doi.org/10.1145/3534678.353904
4. Reed, M., Curtis, K.: Experiences of students with visual impairments in Canadian higher education. J. Vis. Impairment Blindness **106**(7), 414–425 (2012). https://doi.org/10.1177/0145482X1210600704
5. Singleton, K.J., Neuber, K.S.: Examining how students with visual impairments navigate accessible documents. J. Vis. Impairment Blindness **114**(5), 393–405 (2020). https://doi.org/10.1177/0145482X20953268

6. Van Etten, S., Freebern, G., Pressley, M.: College students' beliefs about exam preparation. Contemp. Educ. Psychol. **22**(2), 192–212 (1997). https://doi.org/10.1006/ceps.1997.0933
7. Wang, L.L., et al.: Scia11y: Converting scientific papers to accessible html. In: Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility. ASSETS 2021, Association for Computing Machinery, New York, NY, USA (2021). https://doi.org/10.1145/3441852.3476545
8. Wang, Z., Xu, Y., Cui, L., Shang, J., Wei, F.: Layoutreader: pre-training of text and layout for reading order detection. CoRR (2021). https://doi.org/10.48550/arXiv.2108.11591

ICT to Support Inclusive Education - Universal Learning Design (ULD)

ICT to Support Inclusive Education - Universal Learning Design (ULD)

Introduction to the Special Thematic Session

Marion Hersh[1](✉) and Barbara Leporini[2]

[1] Biomedical Engineering, University of Glasgow, Glasgow, Scotland
marion.hersh@glasgow.ac.uk
[2] Department of Computer Science, University of Pisa and ISTI - CNR, Pisa, Italy
barbara.leporini@unipi.it

Abstract. This short paper presents 12 papers discussing different features of and approaches to inclusive education and give a flavour of the research in the area and the potential of different types of technology. They can be divided into three main groups, covering curriculum accessibility, the attitudes and experience of students and teachers, and the use of particular technologies. They are introduced by a brief discussion of the importance of inclusive education and the role of accessibility, usability and the social model of disability in supporting it.

Keywords: Inclusive Education · Curriculum Accessibility · Barriers · Attitudes · Technology

1 Introduction

Education is of great value to both individuals for personal development and improving their employment opportunities and society as a whole. A postsecondary education qualification significantly increases a person's chances of getting a job e.g. [1] and this may be even more important for disabled people [2]. However, disabled people are under-represented in postsecondary education [3] and obtain poorer degree results, despite comparable entry qualifications [4]. As a result disabled people have fewer employment opportunities than non-disabled ones [5, 6], and attitudinal and other barriers further reduce their chances [7, 8].

The value of inclusive education in mainstream schools is being increasingly recognised, including through legislation. For instance, 97% of all learners are legally required to be included in mainstream schools in Denmark [9]. However moves to educating disabled students in mainstream schools are taking place at different rates in different countries and there is still not universal access to education. For instance, there are 250 children globally without any education, another 1.7 million primary school teachers are required and primary school enrolment needs to almost triple [10]. Inclusive approaches are best supported by an understanding of disability based on the social model [11] which focuses on society's responsibility for removing the barriers disabled and neurodivergent people would otherwise experience.

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 333–339, 2024.
https://doi.org/10.1007/978-3-031-62846-7_40

ICT (information and communication technologies) have the potential to contribute to doing this, including in the area of education, and thereby support inclusion by providing different ways of representing information, expressing knowledge and engaging in learning, including assessment. This has the further advantages of teaching ICT skills and drawing on the increasing popularity and motivating effects of using ICT, particularly amongst young people.

However, successful ICT use requires appropriate inclusive pedagogical strategies and teacher education. ICT also needs to be fully accessible and usable, so everyone, including assistive technology users, can use it [12] and it find it user friendly, intuitive and satisfying to use. Combining accessibility and usability both enable disabled and other people to use particular systems and make it easy for them to do this [13]. Otherwise learners may experience unnecessary barriers that distract them or even prevent learning [14].

2 Session Papers

The session contains 12 papers, seven of which are published in this volume. A brief summary is presented in Table 1 below. They consider learners from all education stages from pre-school to MSc and also include teachers, cover 13 different countries, including two outside Europe, and consider several different technologies as well as different approaches to accessibility.

2.1 Paper Descriptions

The 12 papers in the session are very varied. Three papers consider design approaches to curriculum accessibility or accessibility evaluation and a further paper discusses the role of an assistive technology centre in providing support to disabled students. Three papers consider the knowledge, attitudes and experiences of students and teachers. Five papers consider the development and evaluation of particular technologies to support groups of disabled students.

The three papers considering design approaches and curricula accessibility include an evaluation of curricula accessibility in different fields in Austria, Czech Republic, Cyprus and Spain (Nuppenau et al.), an evaluation of persona profiles to represent three disabled students with different impairments to be used in curriculum design (Heitmeier et al.); and three case studies of the use of information and communication technologies over a 10 year period to support inclusive bachelors and masters level education at a university in Austria (Freudenthaler-Mayrhofer and Wagner).

In their discussion of curricula accessibility in ‘Education Gap: Exploring the Integration of Accessibility and Universal Design in Higher Education Curricula’, Nuppenau et al. identified a focus on students with visual, hearing and motor impairments and differences across academic fields, institutions and countries. For instance, information and communication technologies, engineering and health sciences focused on technology-enabled inclusion, whereas education, social sciences, humanities, business and law often integrated accessibility and universal design with other legal and social considerations in a context of diversity, human rights and social inclusion.

Table 1. Paper overview

Authors	Type of Paper	Target Learners	Technology	Number	Country
Alé et al.*	Study of disabled students' experiences	School students in hospital	Robotics – Lego mindstorm	31	Chile
Carruba and Covarrubias*	App development	18–25 year olds with cognitive impairments	Virtual reality	N/a	Italy
Coughlan et al.**	Survey of disabled students' experiences	Disabled HE students	Various	50	UK
Covarrubias et al.*	App for 3D printing	19–35 year olds with cognitive or developmental impairments	Extended reality	10	Italy
Freudenthaler-Mayrhofer et al.*	Case studies of inclusive education	HE students	ICT	N/a	Austria
Heitmeier et al.**	Evaluating accessibility personas	HE teaching staff	Personas of disabled students	39	Germany
Hellesnes et al. *	Study of text adaptation	Students with ADHD	Artificial intelligence	20	Norway
Hsieh*	Surveys of teachers	School students	Online education	396	Taiwan
Mrochen*	Study of psychology students and digital accessibility	Psychology students	Digital accessibility	363	Poland
Pagliara et al.*	Study of assistive technology centres	Disabled students	AT centres	N/a	Italy
Nuppenau et al.**	Study of curricula	HE students	Accessibility and universal design	N/a	Austria, Cyprus, Czech Republic,Spain
Szabó and Lanyi*	Evaluation of games	Diabetic preschool	Serious quiz games	3 children, 9 adults	Hungary

* Miesenberger, K.; Kobayashi, M.; Penaz, P. (eds.): ICCHP 2024, Computers Helping People with Special Needs, 19th International Conference; Linz, Austria, Proceedings. Springer (2024)
** ICCHP Open Access Compendium * Future Perspectives on Accessibility, AT and eInclusion*; Linz, Austria, Proceedings. Johannes Kepler University Linz (2024)

Freudenthaler-Mayrhofer and Wagne present three case studies in 'Enhancing Inclusive Education through ICT – Lessons From 10 Years of Supporting Students with Different Challenges'. The case studies cover the use of learning videos and an asynchronous individually defined learning process to support students with diverse prior knowledge on a management accounting course; the use of online training to allow students without German language skills to study in German; and hybrid teaching and online collaboration using MS Teams to support students unable to attend in person. The analysis of the case studies shows the need for an adaptive approach involving a variety of different tools and scenarios.

The profiles of disabled student in 'Evaluating Interactive Accessibility Personas on the BlindDate Website' by Heitmeier et al. were developed using co-design with feedback from disabled people and experts in inclusive education. The evaluation involved eight closed questions from the persona perception scale covering presentation of all relevant information, realism and understanding the personas' feelings and several open questions on higher education teaching staff's view the personas. The profiles were generally received very positively, though it was noted that a single profile could not represent everyone with a particular impairment. In addition, individuals' experiences of learning are affected by a variety of factors such as culture, native language, gender and age, and it would be useful to include these factors in the personas.

In their paper 'A Pedagogical Model for In-Situ Training Interventions through the Support of GLIC Assistance Centers for the Ministerial "Sussidi" Grant.' Pagliara et al. discuss the impact of the support provided by an assistive technology and training centre in Italy. This involved working collaboratively to obtain information from the centre. The data shows a change in the nature of support provided with a reduction of nearly a quarter in requests for computers, tablets and other computing devices and an increase of just over a quarter in requests for specialised assistive technologies and cost per request from 2021–22 to 2023–24 and a very slight decline in the total number of requests over this period. It would be useful to identify when these trends started and whether they are continuing. Analysis of the data also raises a number of questions which require further research about using assistive technology to personalise learning, continuing training, measuring impacts and outcomes and stakeholder collaboration.

The three papers on the experiences, attitudes and knowledge of teachers and students cover disabled students' experiences of barriers and their resolution or otherwise (Coughlan et al.); the impact of the Covid pandemic on teachers' acceptance of and difficulties with distance education (Hsieh et al.); and psychology students views of their digital accessibility skills and training needs (Mrochen).

The most common unresolved barriers identified in Couchlan et al. in 'Accessible by Design? Exploring How Barriers faced by Disabled Students are Resolved in Online and Distance Learning' occurred in online tutorials, including audio quality and pedagogy, forums, including difficulties with instructions, and rich media, including in captions. They also investigated the use of accessible by design elements and found that students were most interested in flexibility to study at any time, recordings of tutorials and extensions to assignment deadlines, with a slightly lower number wanting downloadable versions of web-based materials. While the work indicates some useful accessibility features, it should be noted that those fewer students were interested in could be of great value to those students.

In 'General and Special Education Teachers' Perspectives on Online Inclusive Education in Post-Pandemic Taiwan' Hsieh presents the results of the surveys they carried out during and after the Covid pandemic. They found that teachers with experience of distance education, particularly those using blended learning, had the greatest acceptance of distance education during the pandemic. However, these teachers' attitudes changed to become more negative post-pandemic, whereas the attitudes of teachers without experience of distance education remained the same. Interviews with teachers to investigate the responses found that students fell behind during the pandemic and that distance learning increased teachers' workloads and responsibilities. However, it would be useful to differentiate the impact of factors associated with the pandemic and the restrictions it put on learning from the impact of distance learning per se.

In their paper 'Accessibility and Digital Competencies of Psychology Students - New Perspectives', Mrochen reports the results of a questionnaire to psychology students on their skills, training needs and interest in training on digital accessibility and the barriers to producing accessible digital documents. In particular they investigated the use of digital tools to make MS Office more accessible and facilitate psychology students' communication, collaboration and professional development.

There are five papers on the use of particular technologies, one about the use of artificial intelligence (AI) to support the legal requirement of indivualised adaptation of learning in Norway (Hellesnes et al.) and four about using specific technologies with particular groups of disabled students.

Hellesnes et al.'s three-phase study in 'The Effective Use of Generative AI for Adapted Education: An Exploratory Study in the Norwegian Context' involved interviews with teachers; comparison of text produced by teachers and AI for 12 year olds with reading comprehension challenges due to ADHD; and interviews with teachers and teacher evaluation of the adapted text. Unfortunately, the authors did not consider the student perspective or obtain the views of students with ADHD on computer-generated texts.

The four papers relating to particular groups of student cover the use of donated LEGO Mindstorm robotics in hospital classrooms by multi-year 11–17 year olds attending regular classes (Alé and Sánchez); three quiz games designed to teach newly diagnosed type 1 diabetic pre-school children about permitted foods and carbohydrates (Szabó and Lanyi); and the use of virtual reality (Corruba and Covarrubias) and extended reality (Covarrubias et al.) with students with cognitive (and other) impairments.

In 'Environmental Robotics for Educational Revival in Hospital Classrooms'.

Alé and Sánchez report on project-based learning in hospitals. The projects involved designing efficient solar panels using LEGO sensors and monitors to track the sun's path; and designing catchers to obtain water from fog. Students completed questionnaires prior to the experiments and at three random times during them. On average students had a high rating for situational interest when working on science projects. They generally found interpreting data, working in groups and asking questions more interesting than other activities. Scientific practices such as asking questions and interpreting data were most strongly correlated with situational interest in learning than listening, explaining phenomena, working in groups, writing and performing calculations despite strong interest in some of these activities. Limitations included the convenience sample.

Children, their parents and diabetes experts were involved in the design process of Szabó and Lanyi's three quiz games in their paper 'User-friendly Serious Game Design for Diabetic Preschool Children'. Parents can customise the games to better meet their children's needs using an editor. The games were evaluated by three children and nine adults using usability scales with a different scale for children and adults and both groups gave the games a high score.

In 'Virtual Reality (VR) in Special Education: Cooking Food App to Improve Manual Skills and Cognitive Training for SEN Students using UDL and ICF approaches' Corruba and Covarrubias present a virtual reality app to support students with cognitive or physical impairments learning cooking skills. However, they did not discuss the translation of the cooking skills acquired virtually to the real world. Covarrubias et al. present an extended reality app to support use of a 3D printer in 'Extended Reality for Special Educational Needs: From the Design Process to Real Products through 3D Printing'. The app was initially developed for desktop and Android tablet and subsequently for a multi-reality device allowing interaction with both the virtual and real worlds. The two apps are potentially useful and were developed together with an organisation of parents of children with cognitive impairments. However, the authors unfortunately have a deficit based approach to disabled students focusing on diagnosis bodily functions rather than accessibility requirements and overcoming barriers.

3 Conclusions

While by no means covering everything, the 12 papers give a good impression of research in the area, the potential of different technologies, the progress that has been made and the considerable work remaining to be done. Several of the papers involved co-production approaches, including with children. We consider this very positive and the involvement of disabled learners in particular to have the potential to both empower them and improve learning outcomes.

References

1. Hutcheon, E.J., Wolbring, G.: Voices of "disabled" post secondary students: examining higher education "disability" policy using an ableism lens. J. Divers. High. Educ. **5**(1), 39 (2012)
2. Burgstahler, S.: The role of technology in preparing youth with disabilities for postsecondary education and employment. J. Spec. Educ. Technol. **18**(4), 7–19 (2003)
3. Konur, O.: Teaching disabled students in higher education. Teach. High. Educ. **11**(3), 351–363 (2006)
4. Fuller, M., Bradley, A., Healey, M.: Incorporating disabled students within an inclusive higher education environment. Disabil. Soc. **19**(5), 455–468 (2004)
5. Barnes, H., Thornton, P., Maynard Campbell, S.: Disabled People and Employment: a review of research and development work, York: York Publishing Services (1998)
6. Szeto, A.Y.J.: Assistive technology and rehabilitation engineering. In: Assistive Technologies: Concepts, Methodologies, Tools, and Applications, pp. 277–331. Information Science Reference, Hershey, PA (2014)
7. Daone, L., Scott, R.: Ready, willing, and disabled: survey of UK employers. London, Scope (2003). https://www.scope.org.uk/Scope/media/Images/Publication%20Directory/Ready-willing-and-disabled.pdf?ext=.pdf. Accessed 11 Sep 2017
8. Roberts, S., Heaver, C., Hill, K., et al.: Disability in the workplace: employers' and service providers' responses to the Disability Discrimination Act in 2003 and preparation for 2004 changes. London, Department of Work and Pensions Research Summary (2004)

9. Andersen, H.V., Sorensen, E.K.: Technology as a vehicle for inclusion of learners with attention deficits in mainstream schools. Eur. J. Open Distance E-learning **19**(2), 720–730 (2015)
10. Petrie, H., Kheir, O.: The relationship between accessibility and usability of websites. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 397–406, April 2007
11. UNESCO (2023)
12. Barnes (1991)
13. Leporini, B., Paternò, F.: Applying web usability criteria for vision-impaired users: does it really improve task performance? Int. J. Hum. Comput. Interact. **24**(1), 17–47 (2008)
14. Vlachogianni, P., Tselios, N.: Perceived usability evaluation of educational technology using the System Usability Scale (SUS): a systematic review. J. Res. Technol. Educ. 1–18 (2021)

General and Special Education Teachers' Perspectives on Distance Teaching in Post-pandemic Taiwan

Chia-Yen Hsieh(✉)

Department of Early Childhood Education, National Pingtung University, No. 4-18, Minsheng Road, Pingtun, Pingtung 900391, Taiwan
banduna@nptu.edu.tw

Abstract. This study investigates the impact of the COVID-19 pandemic on the acceptance of distance education among teachers in Taiwan, as well as the attitudes and challenges faced by general and special education teachers in implementing online inclusive education. Adopting a comprehensive research design, 396 participants, including elementary school special education teachers, were surveyed. Questionnaire items based on the Technology Acceptance Model (TAM) were employed to measure perceived usefulness, perceived ease of use, and satisfaction. Structural Equation Modeling (SEM) was utilized for data analysis, revealing significant changes in teachers' attitudes towards online learning. The study underscores the challenges educators encounter in adapting to distance education and emphasizes the necessity of providing ongoing support and training to enhance teachers' professional skills in effectively managing online classrooms. Additionally, the study examines the challenges faced by general and special education teachers in implementing online inclusive education during and after the COVID-19 period. Results indicate that special education teachers encountered more challenges and exhibited lower acceptance of online inclusive education, highlighting the importance of providing targeted support and resources to ensure their effective response to the demands of online.

Keywords: Distance education · Crisis distance education (CDE) · Reduced inequalities · Quality education

1 Introduction

Two years ago, during the COVID-19 pandemic, we collected data on the potential issues and proposed solutions teachers might encounter in implementing online inclusive education (see Hsieh, 2023). Unexpectedly, as the pandemic subsided, schools returned to traditional face-to-face teaching. Nonetheless, we continued to collect data for ten months after the pandemic's conclusion. Consequently, we compiled data on the acceptance of distance education among general and special education teachers in regular and special education classes in Taiwanese elementary schools (using the Technology Acceptance Model, TAM). In essence, this study aims to delve into the differences in

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 340–347, 2024.
https://doi.org/10.1007/978-3-031-62846-7_41

acceptance of distance education among general and special education teachers during and after the COVID-19 pandemic period. Amid the global upheaval caused by the COVID-19 pandemic, various countries, including Taiwan, underwent significant transformations in their educational landscape. The Taiwanese government responded promptly by implementing Crisis Distance Education (CDE), marking an unprecedented departure from traditional face-to-face instructional methods. Even after all schools in Taiwan returned to traditional classroom teaching post-2023, the enduring impact of CDE on instructional practices and the educational landscape remains evident. Existing research on teacher attitudes and challenges during CDE provides valuable insights for developing strategies to support educators in navigating the complexities of distance education. However, effective teaching methods remain crucial for online inclusive education, with challenges identified in teachers' online operational and instructional skills. The effectiveness of online inclusive education is closely linked to collaboration between teachers and parents in creating a conducive and inclusive environment for students with special needs. This intertwined narrative underscores the evolving pattern of education, where experiences of Crisis Distance Education and online inclusive education converge, necessitating continuous research and strategic development to address the challenges and opportunities brought about by these transformative shifts.

The study objectives were as follows.

1. To investigate the variance in acceptance of distance education among primary school general education teachers between the period of the COVID-19 pandemic and the post-pandemic era.
2. To examine the disparities in acceptance of distance education among primary school special education teachers between the period of the COVID-19 pandemic and the post-pandemic era.
3. To analyze the variation in acceptance of distance education among teachers during and after the COVID-19 pandemic in relation to demographic characteristics
4. To employ qualitative analysis to assist in understanding the variance in acceptance of distance education among teachers.

2 Methods

2.1 Participants

In 2022, Taiwan had 98,000 elementary school teachers, with the majority working in metropolitan areas (Department of Statistics, 2023). Therefore, this study conducted a stratified random sampling based on demographic characteristics using data obtained from the 2021 Taiwan Social Change Survey (Round 8, Year 2). Utilizing demographic data, the study calculated the percentage and absolute number of teachers in each township category and employed these calculated values for stratified sampling. Considering that special education teachers in Taiwan's elementary schools accounted for approximately 10% of the total teacher population in 2022 (Department of Statistics, 2023), the study also included around 10% of special education teachers for participation. The first wave of printed and electronic questionnaires (pretest) was administered between May and June 2021, while the second wave of online questionnaires (posttest) was distributed in February 2022, precisely six months after the conclusion of the COVID-19 pandemic. In May 2021, Taiwan experienced a rapid surge in COVID-19 cases, marking the peak period for schools at all levels to implement distance education for online inclusive education. During this period, schools across Taiwan were compelled to utilize distance education due to the impact of the pandemic. The first wave saw a total of 495 questionnaires distributed, with 441 returned (an 89% response rate), and 396 questionnaires were deemed valid. In the second wave, 307 questionnaires were considered valid and included in the final analysis. Only data from participants who engaged in both the pretest and posttest were included, resulting in a total of 396 participants in this study. Specifically, 40 of the participants were elementary school special education teachers.

2.2 Data Collection and Measurement

Questionnaire Items on Teachers' Acceptance of Distance Education. This study used items formulated by Agarwal and Karahanna (2000) and Venkatesh and Davis (2000) to measure the TAM variables perceived usefulness and perceived ease of use; additionally, the items formulated by Oliver (1977) were modified to measure the TAM variable of satisfaction in the context of Taiwan. A first draft of the questionnaire comprising 27 items was subject to content validity analysis performed by three experts. Principal component analysis was then conducted with direct oblique rotation to obtain three dimensions; the cumulative explained variance was 61.996%. The final questionnaire comprised 12 items, of which 3, 4, and 5 measured perceived usefulness, perceived ease of use, and satisfaction, respectively. Each dimension had reliability of greater than .75, and the overall reliability of the scale was .819.

Questionnaire Items on Teaching Difficulties in Distance Education. Questionnaire items on teaching difficulties in distance education were formulated on the basis of semistructured interviews conducted by Hsieh (2023). A first draft of a questionnaire comprising 40 items was subjected to content validity analysis involving three experts. Principal component analysis was then conducted with direct oblique rotation to obtain seven dimensions (namely, teacher professionalism, classroom management, course-related factors, material factors, software-related factors, hardware-related factors, and administrative factors); the cumulative explained variance was 61.272%. The final questionnaire comprised 28 items, of which 5, 7, 3, 3, 3, 4, and 3 measured teacher professionalism, classroom management, course-related factors, material factors, software-related factors, hardware-related factors, and administrative factors, respectively. Each dimension had reliability of greater than .695, and the overall reliability of the scale was .899.

2.3 Data Analysis

Quantitative Data Analysis. The Structural Equation Model (SEM) was utilized to analyze the differences in teachers' acceptance of distance education between the period of the COVID-19 pandemic and the post-pandemic period. Specifically, covariance-based SEMs and the maximum likelihood approach were used. Attitude toward distance education (the latent variable) was divided into three observable variables, namely perceived ease of use, perceived usefulness, and satisfaction, per Davis (1989).

Qualitative Data Analysis. We conducted a qualitative analysis of the open-ended responses provided in the questionnaire. In the absence of any pre-existing theory, we employed thematic analysis to extract patterns from the data (Braun & Clarke, 2006). The qualitative data were exported to Excel for analysis. Coding and analysis were implemented according to the recommendations of Braun and Clarke (2006).

3 Results

3.1 Descriptive Analysis

Of the 396 teachers who completed the survey, 68.7% were female ($n = 272$) and 31.3% were male ($n = 124$). The average participant was 42.12 years old (range: 20 to 57 years old) and had 18.28 years of teaching experience (range: < 6 to > 31 years).

3.2 Pretest-to-Posttest Change in Teachers' Acceptance of Distance Education

The SEM of pretest-to-posttest change in teacher attitude toward distance education satisfied the criteria of model fit proposed by Hair et al. (1998). Specifically, the model's $\chi 2$ value was 9.09 (degrees of freedom [df] = 1, $p < .001$); furthermore, GFI was .994, and AGFI was .937, both of which were greater than the requisite minimum of .90. In addition, the SRMR was .029, which was greater than the requisite minimum of 0.5, and RMSEA was .086, which was slightly greater than the requisite minimum of 0.80. Overall, the proposed SEM had good fit (Hu & Bentler, 1999). The score for teachers' acceptance of distance education was significantly lower in the posttest than in the pretest ($p < .001$; loading $= -.156$).

3.3 Associations of Demographic Variables with Pretest-to-Posttest Difference in Teachers' Acceptance of Distance Education

Only teaching category (special education teachers and general education teachers) and instructional strategy (for distance education) were found to be significantly associated with pretest-to-posttest change in attitude. For brevity, the results for only these two variables are presented in this paper.

Association Between Teaching Category and Change in Teachers' Acceptance of Distance Education. The model describing the relationship between teaching category (general education teachers = 0, special education teachers = 1) and pretest-to-posttest change in teachers' acceptance of distance education had good fit. Relative to special education teachers, general education teachers had a significantly higher teachers' acceptance of distance education score ($p < .001$; difference estimate $= -.220$) in the pretest and a nonsignificantly different teachers' acceptance of distance education ($p = .646$; difference estimate $= .048$) in the posttest.

Association Between Instructional Strategy and Change in Teachers' Acceptance of Distance Education. Instructional strategies were divided into synchronous education, asynchronous education, blended education, and inexperience (inexperience was considered a category because a lack of distance education experience implies that a teacher has not adopted any method to deliver distance education). First, analysis of instructional strategies was conducted by setting inexperience as 0, synchronous education as 1, asynchronous education as 2, and blended education as 3 for model testing. The model describing the relationship between instructional strategy and pretest-to-posttest change in teachers' acceptance of distance education had good fit. The association between instructional strategy and pretest is significant ($p = .006$; magnitude = .175). Thus, teachers who had implemented distance education to a greater degree had teachers' acceptance of distance education that were more positive in the posttest (relative to the pretest). The association between instructional strategy and posttest attitude was nonsignificant ($p = .947$; magnitude = .025). Figure 10 presents the estimated marginal means of attitude toward distance education (Instructional Strategy × Pretest–Posttest Difference). Specifically, the average attitude toward distance education was more negative in the posttest than in the pretest for the participants who had experience of synchronous education, asynchronous education, and blended education. Attitude toward distance education did not change in teachers who had no experience with distance education. In addition, it is also worth noting that on the posttest, teachers with no experience in distance education had higher attitudes toward distance education than other experienced teachers (but the reverse was true on the pretest).

In the pretest, teachers who had experience in distance education had more actively accept distance education than those who did not; this difference was greatest between teachers with experience of blended education and teachers who had no experience with distance education. In other words, during the pandemic, teachers who utilized blended distance education for online inclusive education exhibited significantly higher acceptance of distance education compared to teachers with no prior experience in distance education. However, in the posttest, teachers who had experience of distance education had more negative accept distance education than those who did not; this difference was greatest between teachers with experience of asynchronous education and teachers with no experience adopting distance education.

3.4 Qualitative Data Analysis Results

In addition to the quantitative data analysis mentioned above, we also conducted interviews to collect qualitative data. We conducted interviews from three perspectives, including views on implementing online inclusive education through distance education, opinions on the continued use of distance education after the pandemic, and self-assessment of information technology capabilities along with perceptions of challenges in implementing distance education teaching. Following data encoding analysis, we identified the following perspectives.

Factors of Physical. Our research findings confirm that students' physical considerations influence teachers' willingness to continue using distance education after the pandemic. Almost all interviewed teachers mentioned that physical considerations influence their willingness to adopt distance education. Some teachers emphasized that the lack of face-to-face interaction with students during distance education was a challenge in conducting online integrated education.

Factors of Educational Equity. Many teachers believe that distance education exacerbates educational inequity. Some teachers perceive unfairness in student assignment submissions, stating. Additionally, the "digital divide" is cited as one of the reasons many teachers decide not to continue using distance education post-pandemic.

4 Discussion and Conclusions

4.1 Discussion

This paper makes three major contributions. 1. In-depth understanding of the adaptations to online inclusive education during the CDE period. 2. Formulation of future crisis response strategies. 3. Promotion of educational technology integration. First of all, during the COVID-19 pandemic, schools were compelled to shift to distance education, leading to significant transformations for educators and students. Comparing data between the COVID-19 pandemic period and the post-pandemic era provides profound insights into the adaptive measures taken by schools during the CDE period. Second, through such a comparative analysis, effective strategies for future crisis responses can be formulated. Third, the CDE period witnessed a substantial increase in the application of educational technology in teaching. A comparative examination of data from these two periods reveals teachers' willingness to use educational technology in the educational system and the challenges they face in incorporating it into their teaching practices.

Contrary to our expectations, teachers held a more negative attitude toward distance education after the end of the COVID-19 pandemic. The qualitative results reveal that the majority of elementary school teachers had previously undergone training centered on integrating Information and Communication Technology (ICT) prior to the COVID-19 pandemic. These training courses, implemented for approximately a decade in Taiwan, ensured that virtually all in-service teachers had received training in ICT. However, during the COVID-19 pandemic, when teachers were compelled to implement online inclusive education through distance teaching, they encountered numerous challenges in utilizing this approach. The qualitative results revealed the following challenges. First, many students fell behind in the curriculum during CDE, whether general students or special educational needs (SEN) students, and this was highly worrying for parents and teachers (Authors, 2023). Thus, many teachers called for the curriculum to be adjusted during CDE. In particular, many teachers stated that distance education led to poor learning outcomes and inadequate student–teacher interactions (partly due to a lack of physical eye contact in the online setting). This was also partially due to an inability to enforce punishments against unfinished homework in the distance education setting.

Second, distance education led to unfairness in the classroom and precluded the use of some assessment methods due to concerns regarding cheating. Nonetheless, one teacher noted that distance education may be comparable with traditional learning in terms of student cognition and learning outcomes despite the lower quality of student–teacher interaction. In general, distance education may deprive students of the social and emotional learning that they receive from physical interactions in a traditional classroom, which are as much a part of education as the material on the syllabus (Colao et al., 2020; Toscu, 2023). Experience with distance education was also associated with more positive attitudes in the pretest but more negative attitudes in the posttest. Wang et al. (2023) identified several problems with asynchronous education, such as a lack of opportunities for student participation and low levels of classroom engagement. Notably, the attitude of teachers with no experience in distance education was found to be unchanged in the posttest.

This change in attitudes may have been due to teachers concluding, after their implementation of CDE, that distance education failed to live up to its promise. Their use of asynchronous education mainly involved students accessing and uploading assignments online during the holidays. Therefore, most teachers approved of distance education prior to the implementation of CDE because of the convenience that it seemed to offer. However, problems with distance education began to surface after its implementation. In particular, according to the qualitative interview data collected in this study, some teachers expressed worries regarding a lack of safeguards in distance education platforms against students plagiarizing assignments, having others complete their assignments for them, or looking up answers on Google. In addition, one teacher who was interviewed noted that distance education led to an increase in their workload and in the number of responsibilities that they had to juggle. Teachers also cited a need for greater support from school management. Another teacher noted the lack of emotional and social connections between teachers, students, and parents in distance education that would have otherwise been present in a traditional education setting (Rovai, 2002).

Additionally, based on quantitative analysis results, it was found that compared to general education teachers, special education teachers exhibited lower acceptance of implementing online inclusive education through distance teaching, both during the COVID-19 online distance education period and after the COVID-19 pandemic. Through qualitative analysis, it was revealed that special education teachers perceive the implementation of online inclusive education for s SEN students as requiring more time and collaboration with parents and general education teachers, thereby increasing their workload. For instance, in this study, one interviewed teacher mentioned: "In conducting online inclusive education for visually impaired students, the parents of the student need to prepare specialized mouse and keyboard for the student, which necessitates prior communication with the student's parents." Furthermore, in order to ensure that SEN students feel included in the class, when conducting online inclusive education for SEN students, it is necessary for the student to participate alongside other classmates, requiring thorough communication and collaboration with general education teachers, thus adding a significant amount of workload. Lastly, as special education teachers typically develop Individualized Education Programs (IEPs) for SEN students at the beginning of the semester, implementing online inclusive education would require the reformation of IEPs for these students, further increasing the instructional workload for special education teachers. Therefore, the majority of special education teachers are not willing to continue distance education after the end of the COVID-19 pandemic.

4.2 Conclusions and Implications

Our study contributes to longstanding efforts by the Taiwanese government to implement distance education and incorporate ICT into classrooms since the launch of the 2019

Curriculum Guidelines of K-12 Basic Education (NAER, 2023). Specifically, by administering questionnaires before and after the major outbreak of COVID-19 in Taiwan, this study determined whether the pandemic was a watershed in changing attitudes toward and overcoming difficulties in teachers' implementation of distance education. In particular, the proportion of teachers who had experience of distance education increased from 34.1% before the pandemic to 96% after the pandemic.

We found that the teachers in our sample had more negative attitudes toward distance education after the shift toward CDE due to hiccups in its implementation. These hiccups pertained to poor learning outcomes, the low quality of student–teacher interactions, and the problem of cheating in online assessments. Thus, education authorities should exercise caution with regard to making distance education compulsory.

Acknowledgments. The authors would like to express their gratitude to the school participants in the study.

Disclosure of Interests. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Agarwal, R., Karahanna, E.: Time flies when you're having fun: cognitive absorption and beliefs about information technology usage. MIS Q. **24**(4), 665–694 (2000)
2. Florian, L.: The heart of inclusive education is collaboration. Pedagogika **126**(2), 248–253 (2017)
3. Hsieh, C.H.: Research on online inclusive education in Taiwan in the era of COVID-19. Int. J. Inclusive Educ. (2023)
4. Rice, M.F., Carter, R.A., Jr.: When we talk about compliance, it's because we lived it' – Online educators' roles in supporting students with disabilities. Online Learning **19**(5) (2015)
5. Survey Research Data Archive [URDA.]. https://srda.sinica.edu.tw/datasearch_detail.php?id=3459. Accessed 15 Sept 2023
6. Venkatesh, V., Davis, F.D.: A theoretical extension of the technology acceptance model: four longitudinal field studies. Manage. Sci. **46**(2), 186–204 (2000)
7. For detailed references, please contact the author for further information

Environmental Robotics for Educational Revival in Hospital Classrooms

Jhon Alé-Silva(✉) and Jaime Sánchez

University of Chile, 1045 Santiago, Región Metropolitana, Chile
jhon.ale@ug.uchile.cl

Abstract. The COVID-19 pandemic has had a significant impact on classrooms in Latin America, especially in Chile, where the prolonged suspension of classes has resulted in learning losses, mental health issues, and student disconnection, especially those in more disadvantaged sectors. This study focuses on addressing these challenges by employing strategies aimed at reactivating post-pandemic education among students in hospital classrooms, belonging to a historically marginalized sector in Chile. With the aim of increasing interest in the natural sciences, two didactic modules with environmental robotics were co-designed with our students from hospital classrooms. Thirty-one students, aged between 12–17, who attend classes at the hospital, participated in the study. Surveys were used to measure students' initial personal interest, while the Experience Sampling Method (ESM) assessed the situational interest during the implementation of the modules. The results obtained through Multilevel Analysis revealed that factors such as gender, group work, and the use of conventional teaching methods, such as writing and calculations, had a negative impact on students' interest. Conversely, the use of scientific practices, such as asking questions and interpreting data, contributed positively to increase the student interest. This study provides evidence on pedagogical strategies with technologies that may be more effective in increasing interest in hospital classrooms.

Keywords: Post-pandemic Education · Hospital Classrooms · Educational Robotics · Environmental Education · Inclusive Education

1 Introduction

During the COVID-19 pandemic, education systems worldwide revealed great adaptation and flexibility in education, implementing strategies to preserve educational continuity [1]. However, most governments prioritized ensuring learning continuity over the socio-emotional well-being and mental health of students, resulting in an unequal participation in education and an increase in educational gaps, particularly among socioeconomically disadvantaged groups [1–3].

Chile has also experienced these effects including learning losses affecting disadvantaged contexts and lower socioeconomic sectors more significantly. These learning losses have been accompanied by declining mental and socio-emotional health in all

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 348–358, 2024.
https://doi.org/10.1007/978-3-031-62846-7_42

communities, but particularly among students, resulting in a loss of interest in their education [4].

Due to the pandemic, supporting socio-emotional aspects such as interest in education has become a significant concern for current educational policy. In recent years, there have been various projects aimed at increasing interest in science [5, 6]. However, these efforts still appear to be insufficient [7], partly due to the lack of support for their development [8]. Given this situation, there is a recommended need for more and new studies focused on examining students' interest during teaching practices [9], rather than through retrospective measurements of the phenomenon [10].

Increasing the interest of students in hospital classrooms is an even greater challenge. Due to isolation, school discontinuity, and academic loss, students who lack access to traditional school activities suffer from significant decrease in interest, leading to extreme cases of dropping out of the education system [11]. In the hospital environment, playing, attending school, participating in sports, going out to interesting places, social gatherings, and emotional relationships are limited, affecting their motivation and interest [12, 13]. Faced this scenario, teaching activities in hospital environments require an approach that promotes greater emotional well-being and an increase in interest in education. This can be achieved by incorporating technologies that encourage interaction, participation, and access to learning [12], such as robotics [14].

2 Materials and Methods

The research study was focused on Educational Design [15] conducted in hospital classrooms in a low-income area of Santiago, Chile. Two Environmental Robotics modules were co-designed and implemented in two groups of hospital classrooms. A total of 31 students from multi-grade levels (11 to 17 years) who are attending regular classes participated in the study.

LEGO type technologies were used in the study, considering that these kits had been previously donated to the hospital. It's important to acknowledge that surveys and interviews possess inherent limitations as they rely on students' retrospective self-reports of their interests or personal interests [10].

2.1 Instrument to Data Collection

The study employed two distinct instruments to measure interest during the study situations: a. A pretest questionnaire for students was administered to gauge their personal interest in studies and career choices prior to the interventions, b. An instrument based on the Experience Sampling Method was used [16], enabling students to actively assess their situational interest at three random intervals during each class.

The pretest questionnaire on personal interest applied to students includes questions and statements about personal interest based on the Student Questionnaire PISA 2006 [17] and is based on the adaptation to Spanish carried out by [16], which has also been recognized and validated in different educational contexts [18]. The questionnaire's items aim to obtain indications related to the value of personal interest in learning and studies on a four-point Likert scale [19].

Additionally, assuming that situational interest is expressed during teaching situations, the Experience Sampling Method [20] was used, which seeks to identify students' situational interest at random and near moments during teaching activities. To capture this interest, participating students received messages on their cell phones with pre-recorded questions to answer three times and randomly in each lesson, after a "beep-sound".

The questions encompass various aspects of students' actions and their interactions, including listening, writing, calculating, thinking, among others. Additionally, there was a question that inquiries about the specific activity they were engaged in when the "beep-sound" occurred, such as asking questions, planning investigations, interpreting data, and more. Furthermore, situational interest was assessed through a question regarding the level of interest in the activity, rated on a four-point Likert scale.

2.2 Environmental Robotics Modules

The Environmental Robotics modules were developed to align with the Learning Objectives of the Science Curriculum in Chile, focusing on regional environmental problems and fostering technological solutions. Rooted in socio-constructionism principles [21, 22] and employing project-based learning [23], these modules integrate LEGO Mindstorm EV3 robotics kits previously donated to the hospital classroom. The implementation spanned approximately two months, with students working in groups under teacher guidance during 60 to 90-min sessions. The methodology followed four stages: piloting and introduction, launch, sustained inquiry, and public exhibition.

The pilot phase introduced students to robotics technology, exploring sensors, actuators, and block-based programming. The launch phase-initiated discussions on environmental issues, specifically desertification and renewable energy in Chile. Students were tasked with designing robotic solutions to address these challenges. Practical work commenced with the construction of initial prototypes within an inquiry circuit, structured to guide students through computational problem-solving aligned with scientific inquiry stages.

The first inquiry guide [24] focused on designing solar panels for efficient energy harnessing, utilizing LEGO sensors and motors to track the Sun's path (see Fig. 1).

In the second guide, students addressed water resource management by designing fog catchers, incorporating micro-python programming to optimize water capture considering wind direction. The culmination of these inquiries led to a public exhibition where students communicated their findings and proposals through informative pamphlets, assuming various roles to effectively convey their environmental robotics projects.

The exhibition, held during regular class hours at the hospital, engaged both internal and external stakeholders, showcasing the students' innovative solutions aimed at promoting clean energy and combating climate change in Chile.

Fig. 1. The figure shows activities from the first inquiry guide on solar panels and students engaging in environmental robotics activities.

3 Results and Analysis

3.1 Personal Interest in Science Education

A dimension reduction was carried out through factor analysis of responses to the key components of the personal interest survey, retaining two factors. Principal component analysis and Varimax rotation of the main axis were performed, allowing a maximum of 100 iterations to achieve data convergence. Two rotated components with high loadings for both groups of students (0.573–0.861) were considered. These factors were labeled as "personal interests in science for future activities". Factor analysis allows for data reduction and the prediction of situational interest through regression analysis. A t-student test for independent samples was conducted on the questions related to this factor to assess significant differences between the genders of the total students. Table 1 present results by gender.

According to the results compared by the t-student test for both groups (see Table 1), it can be observed that, for the most part, there were no significant gender differences,

Table 1. Personal Interests in Science for Future Activities.

	Male (n = 15)		Female (n = 16)			
Questions Groups 1 and 2	Mean	SD	Mean	SD	t	p
Q3. I will have many opportunities to use Science when I finish my school studies	2,93	,704	3,38	,500	−2,02	0,817
Q4. The effort I put into science classes will help me in future work or activities I want to pursue	3,80	,414	3,50	,632	1,55	,011
Q5. What I learn in science classes is important to me when considering what I plan to study after leaving school	3,40	,507	3,44	,512	−,205	,690
Q7. It is worth studying in science classes because it will help me in my future career options	3,47	,743	3,31	,704	,593	,751
Q8. What I learn in science classes will help me find a job	3,07	,594	3,13	,957	−,202	,037
Q15. I am interested in science	3,13	,640	3,44	,629	−1,33	,422

except for Questions 4 and 8 (Q4 and Q8), which obtained a p-value less than 0.05. To simplify the analysis of scientific practices observed during classes, the model adopted by [16] was used.

3.2 Situational Interest in Science Education

To simplify the analysis of scientific practices observed during classes, the model adopted by [16] was followed, and the practices were grouped into three categories based on the PISA Framework Scientific Literacy competency model. These categories are:

1. Asking questions and designing scientific investigations (including asking questions, planning investigations, and conducting investigations).
2. Interpreting data and scientific evidence (involving developing models, analyzing data, constructing explanations, and arguing).
3. Explaining phenomena scientifically (encompassing the use of models, problem-solving, and information assessment).

Additionally, other variables were considered in the analysis, such as (4) collaborating in small groups, (5) listening, (6) writing, and (7) calculating, as shown in Table 2.

Table 2. Situational Interests in Different Learning Situations.

	F	Interest (1–4)	SD
Group 1			
Scientific Practices Situations			
1. Asking questions and designing scientific investigations	56	3,55	0,62
2. Interpreting data and evidence scientifically	64	3,62	0,52
3. Explaining phenomena scientifically	32	3,22	0,76
4. Collaborating in a small group	112	3,45	0,65
Traditional Science Teaching			
5. Listening	60	2,68	0,72
6. Writing	72	3,18	0,81
7. Calculating	28	2,85	0,80
Total	424		
Group 2			
Scientific Practices Situations			
1. Asking questions and designing scientific investigations	68	3,33	0,79
2. Interpreting data and evidence scientifically	96	3,73	0,48
3. Explaining phenomena scientifically	76	3,50	0,60
4. Collaborating in a small group	104	3,56	0,62
Traditional Science Teaching			
5. Listening	80	2,83	0,83
6. Writing	60	2,58	1,14
7. Calculating	32	2,92	0,90
Total	516		

Concerning the situational interest responses, it is evident that in both groups, the responses were also quite high, with clear trends towards options 3 and 4 on the scale from 1 to 4. These results indicate the reported frequency of such experiences and the level of situational interest aroused during each school scientific activity. Although the length of the activities in both classrooms was the same, the total number of activities recorded by frequency is slightly higher in the classroom of group 2 ($f = 516$). It should be noted that students can record and be involved in different activities simultaneously, so the recorded frequencies do not necessarily correspond to the total number of observations on the Likert scale ($N = 279$). In general, the second group reported that during the robotics workshops, they ask questions, interpret data and evidence, explain phenomena, listen, and calculate more frequently than group 1. Conversely, students from group 1 reported that in the workshops, they engage in writing activities and small group collaboration more frequently than group 2. Additionally, regarding the total observations, it was noticed that, overall, traditional teaching activities scored lower than situations involving scientific practices.

Moreover, to forecast situational interest, a multilevel regression analysis featuring mixed effects and random intercepts was applied. This analysis incorporated "personal interests in future science-related activities" and self-reported gender as covariates. The mixed-effects model enabled the dissection of interest variance into two discrete components: one that pertains to inter-individual variability. The estimation of parameters for the entire student cohort (N = 31) is detailed in Table 3.

Table 3. Estimated Situational Interest for the Total Set of Students.

Parameter	Estimation	Error	gl	t	sig	IC 95%	
						Lim inf	Lim sup
Intercept	1,954	,353	114	5,531	<,001	1,254	2,654
Gender	−,074	,132	114	−,565	,573	−,336	,187
Personal Interest	,001	,136	114	,004	,997	−,268	,269
Asking Questions	,141	,032	114	4,345	<,001	,077	,205
Interpreting Data	,135	,046	114	2,939	,004	,044	,227
Explaining Phenomena	,053	,033	114	1,589	,115	−,013	,119
Working in Groups	−,029	,035	114	−,815	,417	−,099	,041
Listening	,064	,032	114	1,994	,049	,000	,127
Writing	−,001	,021	114	−,026	,979	−,042	,041
Calculating	−,038	,028	114	−1,346	,181	−,094	,018

As the data is hierarchically nested within students, intra-class correlations (ICC) were calculated for the Random Effects on Situational Interest Students: $\sigma^2 = 0{,}125$, Adjusted ICC/Conditional ICC = 0,790/0,707, N = 31, Observations = 279, Marginal (R^2)/Conditional (R^2) = 0,106/0,813.

Regarding the estimates and the ability to predict situational interest based on environmental robotics teaching activities, reported gender, and students' personal interests, it can be observed in Table 3 that the most powerful predictors in both groups were two of the scientific practices situations: asking questions ($p < 0.001$), interpreting data ($p < 0.005$), and one of the traditional science teaching situations: actions related to listening ($p < 0.05$). Furthermore, for both groups, reported gender, group work, and traditional teaching situations like writing and calculating were negative predictors of situational interest.

Finally, the results of the Random Effects show that the intra-class correlation (ICC) indices indicated that multi-grade levels of students and their disparity explain a significant portion of the variance in the dependent variable, which is greater than 10%. A very high conditional R^2 (0.813) was also reported, measuring the variance represented by the combined fixed and random effects.

4 Discussion

This study examined the relationship between reported situational interest during various teaching activities and factors such as gender and personal interest. The objective of this study was to examine this relationship in specific contexts of multi-grade hospital classrooms in Chile. The research design followed the methodological structure proposed by [16], with the modification that environmental robotics teaching modules were designed based on available resources and considering only the curriculum plans and programs of Chile. The design of the environmental robotics teaching modules aimed to achieve the learning objectives of the Chilean curriculum and considered evidence that would trigger or sustain situational interest.

Based on the results and evidence obtained, it is possible to affirm that there are factors associated with situational interest, either from the content being developed during teaching situations or from the implementation.

When working with scientific practices, students in hospital classrooms, on average, rated their situational interest above 3 (on a scale of 1 to 4). In general, students in hospital classrooms reported that situations associated with interpreting data, working in groups, and asking questions were more interesting than other activities. This was more notably observed in group 2 and is consistent with some previous research [16]. Particularly, activities related to calculations were not associated with situational interest, which contradicts previous research [16, 25].

In the context of predicting situational interest in learning, our multilevel regression analysis allowed for a direct comparison of different predictors. The results from the regression models not only supported our findings but also underscored that scientific practices such as asking questions and interpreting data exhibit a stronger association with situational interest compared to conventional teaching scenarios such as listening activities involving students explaining phenomena, working in groups, writing and performing calculations did not significantly predict situational interest, which contrasts with students' positive perceptions of some of these activities.

This incongruity concerning the predictive value of scientific practices, particularly those involving the scientific explanation of phenomena and group work, contradicts the findings reported by [16, 26]. Thus, it suggests the presence of unexplored contextual variables that may influence the development of situational interest in hospital classrooms [27]. Nonetheless, our results confirm that the utilization of other scientific practices remains a pivotal factor associated with situational interest in the context of natural science learning, aligning with prior research that has consistently demonstrated the link between scientific practices and student engagement [17, 28].

5 Conclusion

This study, by its nature, required collaboration with participating teachers in research and implementation activities. This entailed a significant amount of extra time not initially considered in the work plan, along with the development of support resources to facilitate the work.

The sample was convenience-based; therefore, the results cannot be generalized to other populations in hospital classrooms. However, this limitation can be overcome if similarities in the results are observed, especially those related to how the use of scientific practices in lessons served to predict students' situational interest.

Key evidence regarding situational interest in terms of using scientific practices was provided only by students' responses. In the case of hospital classrooms, the classrooms are multi-grade (with students of various ages), which resulted in students' records containing varying amounts of information and depth. Some situational interest records had more details, especially in older age groups. This can be addressed by focusing interventions on more homogeneous age groups.

While situational interest measurements were conducted sequentially in hierarchical order, observations demonstrated that practices can be simultaneously in different phases or engaging in different activities, making teacher scaffolding a challenging task.

When measuring situational interest with younger students, distractions may occur. Although some records by [16] suggest that this only happens at the beginning of activities, this was a persistent difficulty with hospital students, mainly due to the multi-grade nature of the courses with a much broader age range (12 to 17 years).

Lastly, it should be considered that the teaching practices observed do not allow for comparing increases in interest because associated with the COVID-19 pandemic. Instead, they predict increases in students' situational interest for different teaching practices in relation to their gender and personal interest within the current 2023 period. This raises limitations because it was not possible to know how students' interest in hospital classrooms was before the pandemic and its relationship with the interest examined.

Acknowledgments. This work has been developed with the support of the Basal Project FB0003, funding for Centers of Excellence in Research, Program of Associative Research of ANID, Chile.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Reimers, F., Schleicher, A., Ansah, G.: Schooling Disrupted, Schooling Rethought How the Covid-19 Pandemic is Changing Education. OECD Publishing (2020)
2. World Bank. Learning in Crisis: Prioritizing Education and Effective Policies to Recover Lost Learning (2022)
3. Reimers, F., Schleicher, A.: A Framework to Guide an Education Response to the COVID-19 Pandemic of 2020. OECD Publishing (2020)
4. Ministry of Education Chile. Plan de Reactivación Educativa 2023. Mineduc (2023)

5. Zoechling, S., Hopf, M., Woithe, J., Schmeling, S.: Students' interest in particle physics: conceptualisation, instrument development, and evaluation using Rasch theory and analysis. Int. J. Sci. Educ. **44**, 2353 (2022)
6. Union, E.: Horizon 2020 Work Programme 2016–2017. Science with and for Society, European Commission (2016)
7. DeWitt, J., Archer, L.: Who aspires to a science career? A comparison of survey responses from primary and secondary school students. Int. J. Sci. Educ. **37**, 2170–2192 (2022)
8. Potvin, P., Hasni, A.: Interest, motivation and attitude towards science and technology at K-12 levels: a systematic review of 12 years of educational research. Stud. Sci. Educ. **50**(1), 85–129 (2014)
9. Cheung, D.: The key factors affecting students' individual interest in school science lessons. Int. J. Sci. Educ. **40**, 1–23 (2018)
10. Tuominen, H., Salmela-Aro, K.: Schoolwork engagement and burnout among Finnish high school students and young adults: profiles, progressions, and educational outcomes. Dev. Psychol. **50**(3), 649–662 (2014)
11. Arredondo, T.: La intervención de la escuela hospitalaria en el proceso de escolarización del estudiante en situación de enfermedad. Octaedro (2021)
12. Teijeiro, Y., Rodríguez, J.: Materiales y recursos didácticos en Pedagogía Hospitalaria. Octaedro (2022)
13. Lizasoaín, O.: La Pedagogía Hospitalaria Como un Punto de Partida Para la Orientación de Alumnos con Trastornos del Neurodesarrollo, Eunate, pp. 16–36 (2016)
14. Castiglioni, M., Zappa, G., Pepe, A.: The recreational-educational use of robots in pediatrics. An exploratory research on the well-being of hospitalized children. Mondo Digitale **17**(75) (2019)
15. Plomp, T., Nieveen, N.: An introduction to Educational Design Research. SLO-Netherlands Institute for Curriculum Development (2010)
16. Lavonen, J., et al.: Upper secondary students' situational interest in physics learning in Finland and Chile. Int. J. Sci. Educ. **43**, 2577–2596 (2021)
17. OECD. Student Questionnaire for PISA 2006: Main Study. OECD Publishing, (2015)
18. Toma, R.: Revisión sistemática de instrumentos de actitudes hacia la ciencia (2004–2016). Enseñanza de las Ciencias **38**(3), 143–159 (2020)
19. Wigfield, A., Eccles, J.: Expectancy–value theory of achievement motivation. Contemp. Educ. Psychol. **25**(1), 68–81 (2000)
20. Zirkel, S., Garcia, J., Murphy, M.: Experience sampling research methods and their potential for education research. Educ. Res. **44**(1), 7–16 (2015)
21. Papert, S., Harel, I.: Situating constructionism. In: Constructionism. Ablex Publishing Corporation (1991)
22. Larmer, J., Mergendoller, J., Boss, S.: Setting the Standard for Project Based Learning: A Proven Approach to Rigorous Classroom Instruction. Association for Supervision & Curriculum Development (2015)
23. Sánchez, J.: Integración curricular de TICs concepto y modelos. Revista enfoques educacionales **5**(1), 51–65 (2003)
24. Rokos, L., Zavodska, R.: Efficacy of inquiry-based and "cookbook" labs at human physiology lessons at university level–Is there an impact in relation to acquirement of new knowledge and skills? Eurasia J. Math. Sci. Technol. Educ. **16**(12), 1–13 (2020)
25. Inkinen, J., et al.: High school students' situational engagement associated with scientific practices in designed science learning situations. Sci. Educ. 1–26 (2020)
26. Renninger, K., Kensey, C., Stevens, S., Lehman, D.: Perceptions of science and their role in the development of interest. In: Renninger, K.A., Nieswandt, M., Hidi, S. (eds.), Interest in Mathematics and Science Learning, pp. 93–110. American Educational Research Association (2015)

27. Kanter, D., Konstantopoulos, S.: The impact of a project-based science curriculum on minority student achievement, attitudes, and careers: the effects of teacher content and pedagogical content knowledge and inquiry-based practices. Sci. Educ. **94**(5), 855–887 (2010)
28. Grabau, L.J., Ma, X.: Science engagement and science achievement in the context of science instruction: a multilevel analysis of U.S. students and schools. Int. J. Sci. Educ. **39**(8), 1045–1068 (2017)

Virtual Reality (VR) in Special Education: Cooking Food App to Improve Manual Skills and Cognitive Training for SEN Students Using UDL and ICF Approaches

Maria Concetta Carruba[1](✉) and Mario Covarrubias[2]

[1] Universita' Telematica Pegaso, Naples, Italy
mariaconcetta.carruba@unipegaso.it
[2] Mechanical Engineering Department, Politecnico di Milano, Milan, Italy

Abstract. Virtual reality (VR) enters educational processes today as a tool capable of promoting immersive learning experiences and facilitating the engagement and participation of all students, including those with Special Educational Needs (SEN). This paper aims to present a case study concerning the use of VR to improve manual skills and promote immersive and enjoyable cognitive training for students with disabilities. A VR cooking food preparation app is introduced to illustrate how VR serves as a tool for skill training on one hand and as a genuine digital learning environment on the other. In addition to presenting the technical features of the app, the International Classification of Functioning (ICF) approach by the World Health Organization (WHO) will be discussed. This approach is useful for identifying the best strategies to promote learning for students with cognitive disabilities who participated in this case study. Furthermore, the perspective of universal learning design (UDL), also known as Universal Design for Learning (UDL) by CAST, will be explored to guide teachers and trainers in designing digital and innovative learning activities that can accommodate all students.

Keywords: Virtual Reality · Training · Special Education · Neurodevelopmental Disorder · Down Syndrome · Autism · Inclusive Education · Special Educational Needs

1 Introduction

Education today faces real challenges: on the one hand, the world is becoming increasingly digital and digitized; on the other hand, social and ethical approaches also need to become clearer. In the race for modernization and digitalization, a perspective focusing on human needs is important. Technology can serve as a "bridge" between the complexities of our historical era and the opportunities provided by new forms of culture, thought, production, and participation. Technological developments such as VR, AR, ER, and artificial intelligence present both challenges and opportunities. Tools that simulate, extend, use, or

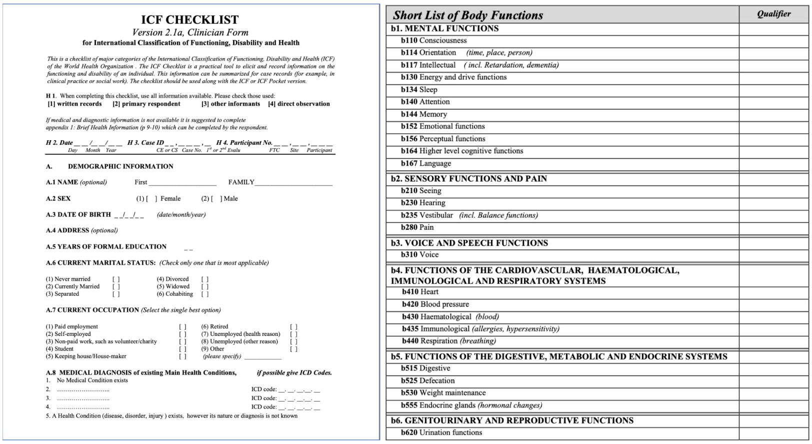
K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 359–366, 2024.
https://doi.org/10.1007/978-3-031-62846-7_43

transcend reality represent potential resources for equity and inclusion when used with a multidisciplinary approach that combines technology with teaching knowledge to support special education. For many disabled people, technology has always been a useful tool that can assist or compensate for one's abilities in areas of perceived disadvantage. Initially considered a "service", today, thanks to new resources, we can boast of a tool that not only pays and helps but also enhances user experience. From specialized equipment to technology for everyone: this is the real innovation required for equity and equality. Disabled individuals can now have technology that is both practical, useful, and entertaining. This paper presents a best practice for using VR technology, specifically a VR app to prepare some tipical Food, to improve motor skills and cognitive training for SEN students involved in the ASPOCollege project. The study cases are designed using UDL approaches to project activities and the ICF checklist to identify students' barriers and facilitators to learning processes.

2 ICF Checklist and UDL Approaches

These activities are based on a preliminary assessment of students according to the World Health Organization model's ICF (International Classification of Functioning) checklist to identify students' strengths, weaknesses, disabilities, and resources needed to support students throughout the program. It is necessary to use standard language and tests to describe job, skill and study. To achieve this goal, WHO created the International Classification of Functioning, Disability and Health (ICF), which complements the International Classification of Diseases (ICD). The ICF was recognized by the World Health Organization (WHO) in 2001 as a classification of health and well-being, including physical patterns and activities, work, activities, participation and environment [1,2]. Additionally, UN recommendations recognize the ICF as a useful tool for protecting the rights of people with disabilities. Unlike traditional biological methods, ICF represents holistic knowledge of biopsychosocial work. The biological approach involves limiting the individual's normal functioning, which can be altered by treatment or injury [3, 4]. The biopsychosocial model defines work difficulties as a conflict between one's health and the demands placed on one's real life [5,6].

Organizing activities from an ICF perspective allows us to start with a good understanding of students and plan activities that they can control. It is important to use the checklist to evaluate. This specific assessment, based on the ICF by WHO, allows for better identification and individualization and differentiation according to the student's abilities. It is important to distinguish between performance and ability in the ICF paradigm: performance information that the student has not yet acquired; This ability allows the student to determine what is most likely to be supportive and successful for the student. In the first part of the checklist (based on WHO model available online for free), there is a range of information about the student and his or her health condition. This is followed by a section on bodily functions; finally, barriers and facilitators are also identified concerning the learning context and environment. This section is essential to then be able to set up the application in a way that is suitable for the student, respecting his or her characteristics but keeping motivation high with a proposal that is nonetheless interesting. The goal is to always work to generate a growth curve.



ICF CHECKLIST

Version 2.1a, Clinician Form

for International Classification of Functioning, Disability and Health

This is a checklist of major categories of the International Classification of Functioning, Disability and Health (ICF) of the World Health Organization . The ICF Checklist is a practical tool to elicit and record information on the functioning and disability of an individual. This information can be summarized for case records (for example, in clinical practice or social work). The checklist should be used along with the ICF or ICF Pocket version.

H 1. When completing this checklist, use all information available. Please check those used:
[1] written records [2] primary respondent [3] other informants [4] direct observation

If medical and diagnostic information is not available it is suggested to complete appendix 1: Brief Health Information (p 9-10) which can be completed by the respondent.

H 2. Date __ __/__ __/__ __ (*Day Month Year*) ***H 3. Case ID*** _ _ , _ _ _ _ , _ (*CE or CS Case No. 1st or 2nd Evalu*) ***H 4. Participant No.*** _ _ , _ _ , _ _ _ (*FTC Site Participant*)

A. DEMOGRAPHIC INFORMATION

A.1 NAME *(optional)* First ____________ FAMILY____________

A.2 SEX (1) [ ] Female (2) [ ] Male

A.3 DATE OF BIRTH _ _/_ _/_ _ *(date/month/year)*

A.4 ADDRESS *(optional)*

A.5 YEARS OF FORMAL EDUCATION _ _

A.6 CURRENT MARITAL STATUS: *(Check only one that is most applicable)*

(1) Never married [ ] (4) Divorced [ ]
(2) Currently Married [ ] (5) Widowed [ ]
(3) Separated [ ] (6) Cohabiting [ ]

A.7 CURRENT OCCUPATION *(Select the single best option)*

(1) Paid employment [ ] (6) Retired [ ]
(2) Self-employed [ ] (7) Unemployed (health reason) [ ]
(3) Non-paid work, such as volunteer/charity [ ] (8) Unemployed (other reason) [ ]
(4) Student [ ] (9) Other [ ]
(5) Keeping house/House-maker [ ] *(please specify)* __________

A.8 MEDICAL DIAGNOSIS of existing Main Health Conditions, ***if possible give ICD Codes.***
1. No Medical Condition exists
2. …………………….. ICD code: __. __. __.__. __
3. …………………….. ICD code: __. __. __.__. __
4. …………………….. ICD code: __. __. __.__. __
5. A Health Condition (disease, disorder, injury) exists, however its nature or diagnosis is not known

Short List of Body Functions	***Qualifier***
b1. MENTAL FUNCTIONS	
b110 Consciousness	
b114 Orientation *(time, place, person)*	
b117 Intellectual *(incl. Retardation, dementia)*	
b130 Energy and drive functions	
b134 Sleep	
b140 Attention	
b144 Memory	
b152 Emotional functions	
b156 Perceptual functions	
b164 Higher level cognitive functions	
b167 Language	
b2. SENSORY FUNCTIONS AND PAIN	
b210 Seeing	
b230 Hearing	
b235 Vestibular *(incl. Balance functions)*	
b280 Pain	
b3. VOICE AND SPEECH FUNCTIONS	
b310 Voice	
b4. FUNCTIONS OF THE CARDIOVASCULAR, HAEMATOLOGICAL, IMMUNOLOGICAL AND RESPIRATORY SYSTEMS	
b410 Heart	
b420 Blood pressure	
b430 Haematological *(blood)*	
b435 Immunological *(allergies, hypersensitivity)*	
b440 Respiration *(breathing)*	
b5. FUNCTIONS OF THE DIGESTIVE, METABOLIC AND ENDOCRINE SYSTEMS	
b515 Digestive	
b525 Defecation	
b530 Weight maintenance	
b555 Endocrine glands *(hormonal changes)*	
b6. GENITOURINARY AND REPRODUCTIVE FUNCTIONS	
b620 Urination functions	

Fig. 1. ICF Checklist

The checklist (Fig. 1) is compiled for each student involved in the project and used for pre-assessment and assessment. In light of the findings, it is then possible to proceed with the design of the training activities. Using the check list from an ICF perspective, it was possible to identify specific codes to work on to enable each student to bring performance as close as possible to attainable abilities given his or her diagnosis. Figure 2 shows how, from December to May (2022), thanks to the facilitators identified and the reduction of learning barriers, students have moved closer and closer to the degree of attainable ability for them and each.

The design of the activities was managed based on the Universal Design for Learning (UDL) approach. All activities were designed by providing for multiple forms of representation, multiple forms of action, and multiple forms of engagement (Fig. 3).

This approach helps make the learning process more real for students with cognitive disabilities and fragility who can thus more easily be guided to work on the how, why and what of learning.

2.1 Users and Special Needs

This case study was conducted as part of the Aspocollege project involving a group of students with cognitive fragility (down syndrome, autism spectrum disorder, intellectual retardation) aged 18 to 25. Some Italian parents, in the

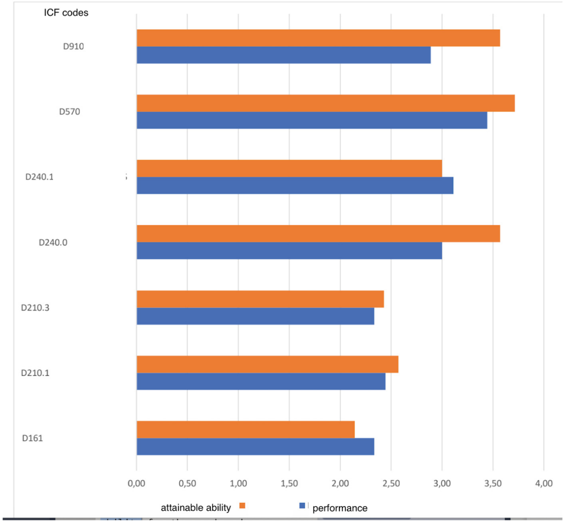


Fig. 2. ICF Results

province of Lecco (Lombardy), who have children with cognitive problems created the Aspoc Association. Aspoc, which stands for Association for the Development of Cognitive Potential in Italian, was established in 2004. Some ASPOC students, based on some specific requirements in terms of diagnosis, residual skills and age group, participate in the ASPOCollege project to continue after school cognitive training. These students participate in the case study. The diagnostic situation of each is different; they share a common difficulty in learning, generalizing, and proceduralizing activities. There is a rigidity of thinking that often makes the process of acquiring new knowledge, skills and competencies laborious. These students have, within the project, the opportunity to experiment with innovative technologies in collaboration with the Milan Polytechnic. This allows a multidisciplinary approach: pedagogical expertise to detect needs and design inclusive activities; technical expertise to design, plan, and implement digital tools useful for achieving the set goals.

2.2 Extended Reality in Special Education

Data shows that the use of AR in general education opens new learning opportunities [4]. Astronomy, for example, is taught using AR tiles, virtual 3D modeling equipment, webcams, projectors, and whiteboards in place of manual methods. AR, air flow, magnetic fields, molecules, etc. It can be used to visualize abstract concepts such as It can also be used to create AR games that encourage student participation. One example given in the article is that students with Android tablets can collect simulated electricity bills from the perspective of a school near a nuclear power plant. Also in [3] the benefit of using this method is seen as a way to represent abstract ideas whose outcome is too small, too large, too fast or too slow to be seen by the human eye. To summarize the main points from

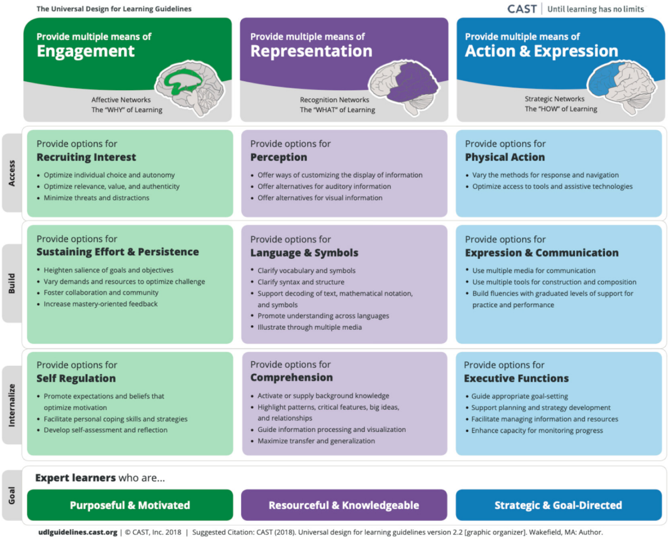


Fig. 3. UDL Guidelines

the examples given, AR focuses on (i) 3D educational content, (ii) collaboration; (iii) student presence, presence and understanding, (iv) seeing the unseen, and (v) connecting formal and informal learning.

AR can also improve students' knowledge and practical skills that are difficult to acquire through traditional education. The main difference is that AR offers students new learning experiences, improves their interest in the content by placing them directly in the content, and improves psychomotor-related knowledge and skills. These results were obtained by Matt Bower and al [2] during their research. They pointed out some aspects of teaching that AR can support: (1) Constructivist learning: AR enables students to be more curious about and engage with content more deeply, (2) Live learning: AR can bring the real world into the classroom, (3) Game-based learning: Using game-based interfaces helps students learn some specific concepts, such as the examples included in [4], (4) Research-based studies based on electronic data collection for future analysis [1].

3 Case Study: Cooking Food App

Cooking food in virtual reality (VR) presents a promising avenue for enhancing the culinary experience, particularly for individuals with disabilities. By leveraging VR technology, people with physical limitations or cognitive impairments can engage in cooking activities in a simulated environment that accommodates their specific needs. VR platforms can offer customizable interfaces, intuitive controls, and sensory feedback, enabling users to participate in various cooking tasks regardless of their physical abilities. Moreover, virtual cooking experiences can provide a safe and accessible space for individuals to practice culinary skills, experiment with recipes, and gain confidence in the kitchen without facing the potential hazards associated with traditional cooking methods. This innovative approach not only promotes inclusivity but also fosters independence and empowerment among people with disabilities, empowering them to explore the joys of cooking and savor the fruits of their culinary endeavors.

4 The APP

The virtual reality cooking app offers an immersive culinary experience with its selection of Italian, Mexican, and Japanese cuisines. Users can delve into the flavors and techniques of each culinary tradition, enhancing their cooking skills while exploring diverse cultural dishes. Additionally, the app's customizable levels cater to users of varying proficiency levels, ensuring accessibility for beginners and providing challenges for experienced cooks. Whether users seek to master the art of pasta-making, perfect the spice blends of Mexican cuisine, or refine the delicate balance of Japanese flavors, the app offers a dynamic platform for culinary exploration and skill development in the immersive realm of virtual reality.

4.1 Cooking in Virtual Reality (VR) Offers Several Advantages for People with Disabilities

Fig. 4. User's while testing.

Accessibility: VR cooking experiences can be designed with accessibility features such as customizable interfaces, intuitive controls, and adjustable difficulty levels, making it easier for individuals with physical disabilities to engage in cooking activities (Fig. 4).
Safety: Virtual cooking eliminates many of the physical hazards associated with traditional cooking methods, such as handling sharp knives, hot surfaces, or

heavy pots and pans, reducing the risk of accidents and injuries for individuals with mobility or dexterity impairments.
Independence: Virtual cooking empowers people with disabilities to cook meals independently, without relying on assistance from others. This fosters a sense of autonomy and self-reliance, enhancing their confidence and quality of life.
Social interaction: Virtual cooking experiences can be shared with friends, family, or online communities, fostering social connections and opportunities for collaboration and peer support among individuals with disabilities.

5 Results

The System Usability Scale (SUS) protocol has been implemented to assess the usability and user experience of the virtual reality (VR) application. By utilizing the SUS, researchers and developers can gather valuable feedback from users regarding the app's interface, functionality, and overall usability. This standardized questionnaire consists of a series of statements designed to measure participants' perceptions of the system's usability, with responses ranging from strongly agree to strongly disagree. Through the SUS protocol, quantitative data can be collected and analyzed to identify areas of improvement and inform iterative design changes. By systematically evaluating the VR app using the SUS, developers can refine and optimize the user interface and enhance the overall user experience, ensuring that the application meets the needs and expectations of its target audience effectively (Fig. 5).

For students with intellectual disabilities or pervasive disorders, learning through concrete experience represents an important facilitator. Theory meets practice and makes it tangible. This also trains students to internalize the processes of both Design and printing through an enjoyable experience that lowers the cognitive load and allows for a fruitful learning experience.

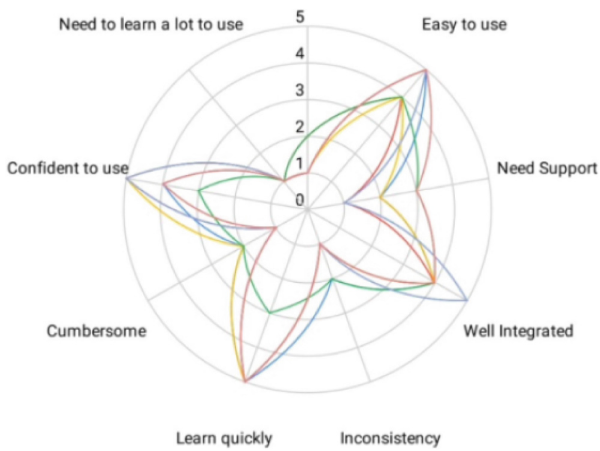


Fig. 5. SUS results

6 Conclusion

In this case study, it emerges how VR could be, when used with UDL approach and based on ICF by WHO observations, a valuable support to promote inclusive education. For students with cognitive difficulties learning is an even more complex process. Using tools that make the learning experience more realistic and immersive can be a valuable aid. The experience demonstrates a high degree of student satisfaction, during all phases of the activity, even the more complex ones or frustration (when you get the steps wrong, you have to start the activity again from the beginning). Students, with these activities carried out on the App, achieved the following goals: - Improved motors skills; - improved and trained oculo-manual coordination; - increased time spent on the activity; - proceduralized the steps for creating the selected food; - divided the task, into more easily manageable subtasks; - improved performance; - improved group relations and experimented with peer tutoring. We can say, therefore, that ER used with this approach promotes Inclusive Education and makes learning an immersive and more enjoyable process for everyone, no one excluded.

References

1. Akçayir, M., Akçayir, G.: Advantages and challenges associated with augmented reality for education: a systematic review of the literature. Educ. Res. Rev. **20**, 1–11 (2017). https://doi.org/10.1016/j.edurev.2016.11.002
2. Bower, M., Howe, C., McCredie, N., Robinson, A., Grover, D.: Augmented reality in education - cases, places and potentials. Educ. Media Int. **51**(1), 1–15 (2014). https://doi.org/10.1080/09523987.2014.889400
3. Nor Farhah Saidin, N.D.A.H., Yahaya., N.: A review of research on augmented reality in education: advantages and applications. Int. Educ. Stud. **8**(13) (2015). https://doi.org/10.1016/j.edurev.2016.11.002, https://www.sciencedirect.com/science/article/pii/S1747938X16300616
4. Wu, H.K., Lee, S.W.Y., Chang, H.Y., Liang, J.C.: Current status, opportunities and challenges of augmented reality in education. Comput. Educ. **62**, 41–49 (2013). https://doi.org/10.1016/j.compedu.2012.10.024, https://www.sciencedirect.com/science/article/pii/S0360131512002527

Accessibility and Digital Competencies of Psychology Students – New Perspectives

Izabela Mrochen(✉)

MultiAccess Centre Izabela Mrochen, Panewnicka 327a/2, 40-773 Katowice, Poland
ibmrochen@gmail.com

Abstract. In the 21st century the digital skills and competencies seem to be a fundamental requirement for many employees. It can be said that willingness to use digital technologies is not very common among therapists or psychologists. Those professionals are perceived as employees with non-digital approaches. However, the global pandemia changed the way doctors, nurses, therapists or psychologists have worked. This is why, it should be underlined that the process of training psychology students should be changed. Generally, they are expected to identify mental problems among their would be patients or clients, while they also ought to break down any technological barriers to improve digital inclusion. As a result, in the context of the academic syllabus, psychology students should have better digital literacy skills enabling them to recognize digital accessibility barriers e.g. found in digital documents. The need for training and education in digital accessibility has increased significantly for some time, which was the reason of modification of the syllabus framework.

This paper aims to explore the effectiveness students study how to design accessible digital documents so as to merge their skills and accessibility issues. A set of proposed changes to ensure the syllabi would cover accessibility issues was implemented in October 2022. To gather full-time and part-time students' feedback on proposed changes, a questionnaire was organized from January 2023 to February 2024. As a result, participants were asked to offer their qualitative comments in relation to the digital training.

Keywords: Digital Accessibility · Digital Documents · Syllabus · Psychology Students · Curriculum

1 Introduction

1.1 New Technologies vs. Traditional Perspectives on Psychology Academic Courses

The advent of new technologies has revolutionized various fields in the labour market. In line with the impact of technology advancement, professional psychologists have faced a range of challenges, which occurred during Covid. They started to conduct therapy sessions remotely, making therapy more accessible for people who were not able to attend in-person sessions. In this case, psychologists were able to use a wide range of

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 367–376, 2024.
https://doi.org/10.1007/978-3-031-62846-7_44

tools like online surveys, mobile apps to track patients' mood, their conduct or social interactions in real time. Generally, psychology programs at universities typically include a spectrum of subjects to provide comprehensive understanding of human behaviour and mental processes without adding very little information on how to use new technologies. The most common subjects include general psychology, cognitive psychology, developmental psychology, social psychology, clinical psychology, and so on, which offer the introductory courses that provide an overview of various topics in psychology covering emotions, feelings, social behaviour, collecting and analyzing data, interpretation of findings, treatment of mental, emotional or behavioural disorders [1]. Thus, the whole idea of valuable insights for therapy and research has been the foundation of psychology students' education. According to the APA Guidelines (2011) [2] the course content standards should explain the complexities of human thought and behaviour as well as ability to apply psychology-based problem-solving skills (2012) [3].

However, it should be underlined that faculty members at psychology departments are expected to incorporate diversity, science, equity, fairness, and inclusion in their course content [4]. As a result, academic curriculum ought to support inclusive learning environment and digital accessibility issues so as to make psychology students familiar with the emotional and psychological challenges associated with individuals who face a spectrum of disabilities, impairments or limitations of activities. Therefore, ongoing efforts are needed to implement awareness and understanding of digital accessibility standards in the field of psychology in the education landscape of psychology departments.

1.2 The Primary Objectives

This paper explores the critical issue of digital accessibility in the curriculum of the psychology department. It underscores the importance of integrating digital tools and resources to enhance the learning experience. The study delves into the perceptions of psychology students towards digital accessibility, and proposes strategies for its improvement. The aim is to foster the educational journey of psychology students at the Department of Psychology at Akademia Górnośląska (the Upper Silesia Academy) in Katowice, Poland. This paper serves as a stepping stone towards a curriculum that includes more information about digital accessibility in higher education so as to improve students' knowledge breaking down any digital barriers.

Despite significant advancements in psychology, it's challenging to find core elements related to accessibility incorporated into the curriculum. Given that the course material in psychology departments had largely stayed the same, this paper seeks to assess the influence and efficacy of accessibility concerns within the academic course titled Applying Information Technology To Psychology (abbr. AITTP), which was offered from October 2022 to February 2023 and again from October 2023 to February 2024. The document presents the students' perspectives on the importance of integrating digital accessibility concerns into the curriculum. The depicted results can inform higher-education curriculum developers how to design successful course content to promote digital inclusion among psychology students.

Section 1 Introduction covers the general aspects of activities that are performed by professional psychologists and therapists. What is more, a set of digital challenges during Covid-19 made psychologists use digital devices to communicate and collaborate with others, in particular with the disabled and people with impairments. As a result, the issues of digital accessibility have been implemented into the syllabus of psychology students in 2022. Section 2 Methodology describes a range of groups of psychology students who took part in the study as well as their qualitative opinions on studying Web Content Accessibility Guidelines (WCAG) and designing accessible digital documents in a range of formats (MS Word, MS Excel, MS PowerPoint, PDF) so as to make the therapeutic digital environment more inclusive in their future workplace. Furthermore, a quantitative analysis is incorporated to examine the variations in student perspectives in Sect. 3 Results. The discussion focuses on the qualitative and quantitative analysis of 366 questionnaires. Section 4 Recommendations offers suggestions and deliberations on a novel method that establishes a beneficial structure for conducting digital accessibility training courses in psychology departments. Section 5 References.

1.3 Related Work

The majority of academic curricula emphasize the fact that inclusive pedagogic strategies are developed to support all students “to gain access to the same learning opportunities at university” (GICEP 2021–2022) [5]. To illuminate those practices, it is worth noticing that psychologists discover new ways in which technology can make collaborating with others more effectively. Unfortunately, it is difficult to find any academic research in which it is possible to find any attempts to test the level of technological competence among psychology students. In addition, the topic of digital accessibility is hardly included at all in curriculum in psychology departments. With this regards, new technologies should bridge theory and practice, which can fill the void in psychology foundations (Chye and Chua 2023) [6]. In May 2020 a report commissioned by London Digital IAPT on Digital Competencies for Psychological Practitioners in IAPT Services was published by Royal Holloway University of London [7]. The report discusses various digital competencies that need to be implemented and developed among psychological practitioners, psychologists and trainers.

What is more, the Topol report [8] in 2019 emphasizes that services and training programmes need to develop digital literacy among all NHS trainees, which could improve the knowledge and skills if the curricula would be redesigned to support healthcare graduates. In 2021, the high demand for psychological services during the global Covid-19 was a catalyst to conduct a cross-sectional online survey to report varying degrees of competence across the digital tasks. The conducted online survey with Aotearoa New Zealand Registered Psychologists only covered the insights into the range of the digital skills as well as digital competencies of professional psychologists in New Zealand (Dobson et al. 2022) [9]. A direct consequence of the acceleration to online working as a result of the Covid-19 pandemic has made professional psychologists aware of the need for implementing the digital issues into the training area. Furthermore, it can be said that this pandemic period became a great springboard for further research in the field of psychology students’ education.

1.4 Procedures and Goals

It is important to emphasize that the integration of digital tools may facilitate communication, collaboration and professional development among would be psychologists [10], which will have a positive effect on a diversified group of patients and clients. The aim of the research was to acquire the psychology students' opinions on implementing digital accessibility issues into the syllabus. Unfortunately, any aspects of digital accessibility had not been included in the AITTP course before 2022. The range of digital issues included in the syllabus covered only general digital competencies such as editing MsOffice documents, broadening knowledge on the Internet history as well as computer operation. Nevertheless, the syllabus lacked external details regarding digital inclusion, disability, or intricate tactics related to the validation of digital documents aimed at enhancing content to be accessible.

Therefore, in September 2022 the syllabus was modified to include a number of strategies for making MS Office documents more accessible. Furthermore, general information on good practices, accessibility standards (e.g. Web Content Accessibility Guidelines), the use of plain language and making digital content clear, concise and well-organized were added to be discussed. After reviewing the psychology students' qualitative comments in February 2023, a set of proposed changes were implemented into the syllabus in October 2023. The primary objectives of the syllabus design were focused on enhancing the accessibility of digital documents (such as MS Word, MS Excel, MS PowerPoint). Consequently, in October 2023, a series of methods pertaining to metadata and PDF file validation were introduced, replacing topics like History of the Internet and the Architecture of Computer Hardware. The focus group members showed greater interest in enhancing the accessibility of digital documents rather than in learning about computer hardware.

The findings from this discussion will inform course instructors, curriculum designers, and institutions about the effectiveness of accessibility measures. Recommendations can guide future improvements to enhance inclusivity in the professional development as well as psychology students' perception. Based on the results of the survey, it can be assumed that students are interested in digital practice skills to be more confident in working digitally as future psychological practitioners in a variety of services.

2 Methodology

2.1 Study Design

The primary objective of this paper is to evaluate the effectiveness and impact of accessibility issues within an academic course in psychology department. Specifically, we aim to understand how well the course accommodates diverse student expectations. As there is no framework of formal digital competencies related to digital accessibility for psychology students, and no research has been conducted in this area so far, the descriptive survey was conducted with first-year full-time (FTS) and part-time (PTS) psychology students. The participants of the survey took part in the AITTP course from 1st October 2022 to 28th February 2023 and from 1st October 2023 to 29th February 2024 at the Department of Psychology. A total of 366 students completed the survey.

2.2 Survey Design and Scope

The study involved an anonymous questionnaire which incorporated closed and open-ended questions to capture the psychology students' opinions about their digital competence, problems connected with digital accessibility as well as their opinions on implementing digital accessibility issues into the syllabus. The questionnaire was designed in paper format. It was composed of three parts: Part 1 Bio Data, Part 2 A set of questions and Part 3 Conclusions. The average time to complete the closed and open-ended questions was about 15 min. The set of questions was identical for all students and the primary inquiries encompassed the following:

- **digital accessibility in prior education** - this point is about whether the students were exposed to issues related to digital accessibility before they started studying psychology,
- **digital accessibility in computer science classes in pre-university education** - this point asks if the students had computer science classes during their elementary, middle, or high school education,
- **inclusion of WCAG standard in pre-university education** – this point asks the students' opinion on whether topics on digital and social inclusion as well as use of the WCAG standard should be discussed before starting higher education,
- **areas of interest** – this aspect requires students to pinpoint a minimum of three subjects (which they found particularly engaging) related to digital accessibility.

Between January and February 2023, 195 freshmen psychology students filled out the survey. In the period from 26 January to 11 February 2024, another 171 students completed the same survey. The respondents represent about 1.26% (N = 366) of all psychology students (N = 29,000 in 2021) [11] and the total student population (which has grown to 1,223,600 since the 2020/2021 academic year) [12] in Poland for the year 2023.

3 Qualitative and Quantitative Analysis

3.1 Respondents

The sample comprised of 366 psychology students. There were 61 (31.28%) full-time students (FTS) and 134 (68.72%) part-time students (PTS) who took part in the questionnaire in the 2022/23 academic year. While in 2023/24 there were 58 (33.92%) FTS and 113 (66.08%) PTS. See Fig. 1 for breakdown of participants' ages.

3.2 Digital Experience

There was variability in students' experience connected with the digital accessibility area. The group of students who took part in the courses that included digital accessibility issues in pre-university education consisted of both full-time students (N = 34, 34.4%) and part-time students (N = 65, 65.7%). However, a significant percentage of students, did not have any knowledge about digital accessibility. There were as many as 85 FTS (approx. 32%) and 181 PTS (68%). Out of the 366 students who participated in the

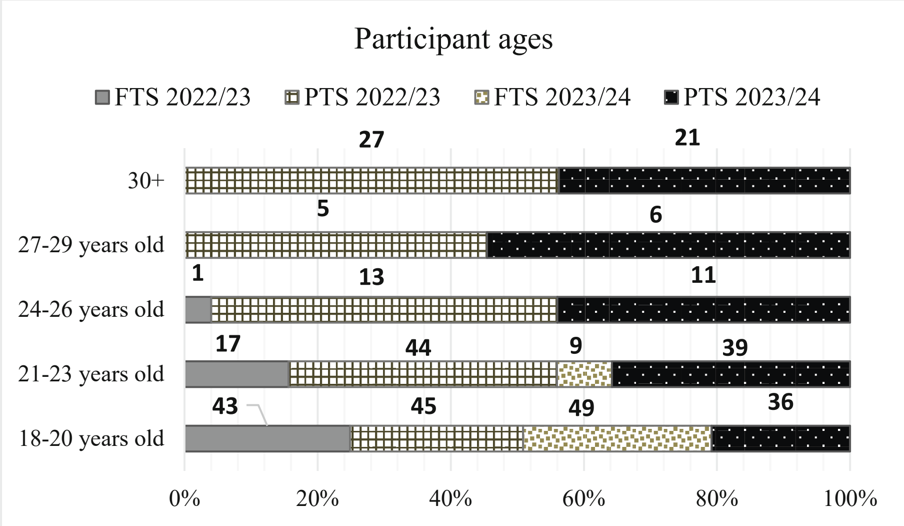


Fig. 1. Participant ages.

Table 1. Digital accessibility in prior education (FTS stands for full-time students and PTS stands for part-time students)

Answer	FTS 2022/2023	PTS 2022/2023	FTS 2023/2024	PTS 2023/2024	Total No
Yes	19	34	15	31	99
No	42	99	43	82	266

survey, 365 responded to the question. See Table 1 for implemented digital accessibility issues in the prior education curriculum.

Then, the participants provided their feedback on discussing digital accessibility issues during their computer classes in pre-university education. Out of 366 respondents, 323 students had computer science classes before studying psychology. A minority of survey participants had digital accessibility issues included in their previous computer science classes (N = 26, 8%). The reminder of participants had no previous digital accessibility issues in their computer science syllabus (N = 297, 92%). See Table 2 for discussing various aspects of digital accessibility during computer classes in pre-university education.

Across all psychology students, there was variability in topics that covered accessibility elements implemented in their computer classes. Majority of the students (92%) had never taken part in any courses on digital accessibility before starting their university education in psychology department. Even if the course on computing was conducted in the secondary school, the syllabus did not include any accessibility issues. Generally,

Table 2. Digital accessibility discussed during computer classes in pre-university education (FTS stands for full-time students and PTS stands for part-time students).

Answer	FTS 2022/2023	PTS 2022/2023	FTS 2023/2024	PTS 2023/2024	Total No
Yes	5	11	4	6	26
No	14	122	54	107	297

they explored typography and its impact on readability, delved into the unique challenges faced by disabled users, and considered the specific needs of users with hearing impairments and the deaf community. The curriculum also covered the theory of color and the implications of colorblindness on user experience. This holistic approach to understanding digital accessibility underscores their commitment to creating inclusive digital environments; yet, this does not suffice to claim a substantial understanding of digital accessibility.

3.3 Feedback About Digital Accessibility

This section presents the collected students' opinions on best practices on how to make MS Office documents clear, concise and well organized. Furthermore, their description of the most interesting accessibility issues that were discussed during the lectures and practical training from 2022 to 2024. In particular, the topics covering types of disabilities, document formatting, describing images, adding alternative texts, plain language guidelines, designing diagrams and tables, etc. were the most interesting for the respondents. In the year 2022/23, a significant 86.15% (N = 168 out of 195) of respondents highlighted the most compelling topics related to digital accessibility. Similarly, in 2023/24, approximately the same proportion of respondents, 83.62% (N = 143 out of 171), identified the most intriguing areas in digital accessibility. See Table 3 for areas of interest presented by 311 (N = 366, 85%) participants of the survey. The table presents the subjects in digital accessibility, ranked from the most captivating to the least regarded ones.

The results revealed that a significant number of students, specifically 29.26% (N = 91) individuals, expressed a keen interest in adding alternative texts to non-textual elements. Moreover, a considerable number of students, precisely 20.58% (N = 64), found the theory of colour and colour blindness to be intriguing. Another noteworthy finding from the survey pertained to the interest shown by 7.07% (N = 22) students in calculating color contrast ratios and utilizing various color contrast applications. Furthermore, when it came to document editing, a considerable number of students, 21.5% (N = 67) to be exact, focused on using styles to create proper heading structures, incorporating descriptive links, and designing accessible tables, graphs, and charts with textured or shaped lines.

Overall, the survey results shed light on the psychology students' awareness and interest in enhancing digital document accessibility. It is evident that a significant portion

Table 3. The participants' feedback on the most interesting areas discussed during the aca-demic course from 2022 to 2024 (FTS stands for full-time students and PTS stands for part-time students).

Areas of Interest	2022/2023 N = 168	%	Areas of Interest	2023/2024 N = 143	%
The disabled users and implementing accessibility	90	53.6%	The disabled users and implementing accessibility	52	36.4%
Alternative text	64	38.1%	Theory of colour (colourblindness)	27	18.9%
Editing Word Docs	41	24.4%	Alternative text	26	18%
Theory of colour (colourblindness)	39	23%	Editing MsOffice Docs (PPT, Excel)	25	17.5%
Editing MsOffice Docs (PPT, Excel)	38	22.7%	Assistive technologies	24	16.8%
WCAG guidelines	34	20.2%	The colour contrast ratio	22	15.4%
Assistive technologies	26	15.5%	Editing Word Docs	18	12.6%
Accessible Websites	17	10%	WCAG guidelines	13	9.1%
The colour contrast ratio	15	8.9%	Accessible Websites	7	4.9%

of the student body is committed to making digital content more inclusive by incorporating alternative texts, considering color choices, calculating contrast ratios, and utilizing various accessibility features during the document editing process. These findings emphasize the importance of prioritizing accessibility in digital document creation and provide valuable insights for further improvements in this area.

4 Recommendations

The feedback analysis highlighted that a spectrum of significant competencies were not developed among freshmen students in pre-university education. A wide range of changes both in new technologies and psychology competencies have had implications for designing the inclusive curriculum to facilitate and recognize the importance of digital accessibility. The new content of syllabus and the practical guidelines implemented into the psychology students' training can raise knowledge on the strategies how to unlock document content to everyone, including people with disabilities or impairments. The new approach to curriculum content can break down digital barriers that can exclude any potential users from accessing the MS Office content that can bridge the gap between psychologists, therapists and potential patients. It should be underlined that the new approach provides a valuable scaffolding of training courses on digital accessibility at psychology departments.

What is more, it is worth noticing that psychology students are expected to know more about digital accessibility because it allows them to understand and cater to the diverse needs of individuals with disabilities (APA 2023) [2], which is a key aspect in the field of psychology. Knowledge in digital accessibility can help them create inclusive digital content, whether for research, therapy, or education. Furthermore, it equips them with the skills to advocate for digital inclusivity, ensuring that psychological services and resources are accessible to all. This not only broadens their professional competence but also aligns with the ethical responsibility of psychologists to promote accessibility and inclusivity.

Disclosure of Interests. The author has no competing interests to declare that are relevant to the content of this article.

References

1. QS World University Rankings by Subject 2023: Psychology. https://www.topuniversities.com/university-subject-rankings/psychology?countries=at. Accessed 12 Apr 2024
2. American Psychological Association (APA). Principles for Quality Undergraduate Education in Psychology (February 2023). https://www.apa.org/education-career/undergrad/principles. Accessed 12 Apr 2024
3. American Psychological Association (APA). Guidelines for Preparing High School Psychology Teachers: Course-Based and Standards-Based Approaches, pp. 5–8 (2011). https://www.apa.org/education-career/undergrad/principles. Accessed 11 Apr 2024
4. American Psychological Association (APA). Principles for Quality Undergraduate Education in Psychology (February 2023), pp. 1–6. https://www.apa.org/education-career/undergrad/principles. Accessed 11 Apr 2024
5. A Guide to Inclusive Curriculum and Educational Practice (GICEP)-Queen Mary Academy (qmul.ac.uk). https://www.qmul.ac.uk/queenmaryacademy/educators/resources/inclusive-curriculum/inclusive-curriculum-guide/. Accessed 05 Apr 2024
6. Yen Leng Chye, S., Leng Chua, G.: Pedagogy and Psychology in Digital Education (2023). https://doi.org/10.1007/978-981-99-2107-2. Accessed 10 Apr 2024
7. Pote, H., Moulton-Perkins, A.: Digital Competencies for Psychological Practitioners in IAPT Services, pp. 4–5 (2020). https://onlinepsychologyandcounselling.com/images/20.05.22_IAPT_Digital_Competences_Framework_Report_May_2020_Final_003.pdf. Accessed 10 Mar 2024
8. Topol Review: preparing the healthcare workforce to deliver the digital future. Report 2019. In: Digital Competencies for Psychological Practitioners in IAPT Services, p. 5
9. Dobson R., Variava R., Douglas M., Reynolds L.M.: Digital competency of Psychologists in Aotearoa New Zealand: A cross-sectional survey. In: Front. Digit. Health, vol. 4 (2022). https://doi.org/10.3389/fdgth.2022.951366/full. Accessed 10 Mar 2024
10. Cummings, J.A.: Technology in the practice of school psychology: the future is past tense. In: Bray, M.A., Kehle, T.J. (eds.) The Oxford Handbook of School Psychology, Chap. 42, pp. 832–859 (2012). https://academic.oup.com/edited-volume/41251/chapter-abstract/350798968?RedirectedFrom=fulltext. Accessed 10 Apr 2024

11. Raport WIB i ZBP. pp. 9–10. https://www.wib.org.pl/wp-content/uploads/2021/09/raport-portfel-studenta-2021-wrzesien-wib-zbp-bde-aktywny-student.pdf?fbclid=IwAR3MN3oR38RaSlUAzbCS_6Kwi15niU7jnyHgPE gaggu26Og2LxBOpTyqgfg
12. Higher Education in the 2022/23 Academic Year – Preliminary Data. Statistics Poland. https://stat.gov.pl/obszary-tematyczne/edukacja/edukacja/szkolnictwo-wyzsze-w-roku-akademickim-20222023-wyniki-wstepne.8.9.html

Extended Reality for Special Educational Needs: From the Design Process to Real Products Through 3D Printing

Mario Covarrubias[1(✉)], Maria Concetta Carruba[2], Paolo Tasca[3], Paolo Perego[4], and Marta Mondellini[5]

[1] Mechanical Engineering Department, Politecnico di Milano, Milan, Italy
mario.covarrubias@polimi.it
[2] Universita' Telematica Pegaso, Napoli, Italy
[3] Department of Electronics and Telecommunications, Politecnico di Torino, Turin, Italy
[4] Design Department, Politecnico di Milano, Milan, Italy
[5] STIIMA, CNR, Biella, Italy

Abstract. Extended Reality (ER) stands out as one of the major technology trends currently, with expectations of even further growth in the near future. The primary objective of this paper is to develop an ER application for the Sharebot One and Sharebot Next Generation 3D Printers, specifically designed to assist students with special needs. The application follows an inclusive design perspective based on the International Classification of Functioning, Disability, and Health (ICF) by the World Health Organization (WHO), enabling students to work and interact with the machine safely.

To achieve this goal, three configurations have been developed and tested: a desktop version running on a PC, an Android tablet-based version, and a solution optimized for the Microsoft HoloLens device.

Keywords: Extended Reality · Training · Special Education · Neurodevelopmental Disorder · Down Syndrome · Autism · Inclusive Education · Special Educational Needs

1 Introduction

Nowadays, Education is experiencing a real challenge: on the one hand, an increasingly digital and digitized world and, on the other hand, the exponential growth of needs that points to the urgency of working for social equity as a real priority. In this race for modernization and digitization, a magnifying glass on humans' needs is fundamental. The technology could be a "bridge", a mediator between the complexity typical of our historical time and the opportunities offered by the new forms of doing, thinking, producing, and participating in society. Technological innovations, such as VR, AR, ER, and AI, represent a challenge and a great opportunity at the same time. Tools that simulate, extend, implement, or venture beyond reality, represent potential resources for equity and

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 377–384, 2024.
https://doi.org/10.1007/978-3-031-62846-7_45

inclusion when used with a multidisciplinary approach that sees technology in perfect synergy with pedagogical approaches to support special education. Technology has always represented, for many people with disabilities, a useful tool for enabling or compensating the individual in those areas of fragility determined by their diagnostic picture. In the beginning, the approach was considered purely "assistive," today thanks to new resources we can boast tools that not only compensate and enable but, at the same time are very attentive to user experiences. From special technology to technology for all: this has been a real innovation. People with disabilities can now experience technologies that are not only accessible and usable but also enjoyable, too. A new interest in ER is growing and its potential to improve special education based on interactive experiences by enhancing the real-life environment with virtual elements. This paper presents a best practice for using ER technology, specifically the 3D Printer Sharebot One, to support the learning processes of students with intellectual disabilities, pervasive developmental disorders, and cognitive fragility.

2 The ICF Model to Design the Learning Process

Activities were conducted from an initial observation of students based on a checklist on the ICF (International Classification of Functioning) by WHO model to identify students' strengths and weaknesses, barriers, and facilitators needed to support students during activities. It is necessary to explain functionality, capacity, and performance using standard vocabulary and tests. The International Classification of Functioning, Disability, and Health (ICF) was developed by WHO as an addition to the International Classification of Diseases (ICD) to carry out this mission. ICF was accepted as a standardized classification of health and health-related areas, including bodily structures and functions, activity, participation, and environmental factors, by the World Health Organization (WHO) in 2001 [1, 2]. Additionally, the United Nations' recommendations recognized ICF as a useful tool for protecting the rights of individuals with disabilities. ICF represents the biopsychosocial holistic knowledge of functioning, in contrast to the traditional biological approach. The biological approach contends that functional limitation is a person's static characteristic that can be changed by treating ailment or damage [3, 4]. The biopsychosocial model describes functional difficulty as a mismatch between a person's health status and the demands placed on him by his actual living circumstances [5, 6].

Setting the activities from an ICF perspective allowed us to start from a thorough knowledge of the students to plan activities that are manageable for them. Observation with checklist is essential. This specific observation, based on the ICF checklist, makes it possible to better define the pathway and to be able to personalize and differentiate it according to the student's abilities. In the ICF paradigm, it is important to differentiate performance from capacity: the performance records what the student has accomplished yet; the capacity makes it possible to identify what possible empowerment scenarios are viable and achievable for the student.

2.1 Users and Special Needs

This work is the result of a project that involves collaboration with an Italian Association, ASPOC (Association for Cognitive Enhancement), and, especially with a specific area of action: "ASPOCollege" which provides training and empowerment opportunities for young people with intellectual disabilities and pervasive developmental disorders aged 19 to 35. These students have completed their formal education processes and find themselves in that hard phase before job placement; although this phase is complex in general for all young people, turns out to be particularly difficult for students with disabilities who struggle to enter the job placement. Some Italian parents, in the province of Lecco (Lombardy), who have children with cognitive problems created the Aspoc Association. The group stands for an inclusive environment and promotes the creation of learning environments that will enhance inclusion not only in the classroom but also in society at large. The parents are invited to recognize the significant roles of educators, rehabilitation specialists, siblings, health professionals, and institutions as part of the mediated learning process. Aspoc, which stands for Association for the Development of Cognitive Potential in Italian, was established in 2004. The group maintains tight relationships with the facilities and services offered throughout the nation. According to ASPOC, the importance of context, the processes, and the quality of mediated learning processes play a fundamental role in the development of each. This belief is based on the confidence in current scientific theories on structural cognitive modifiability based on the evidence that people's neuronal structures are modifiable if exposed to sufficient stimuli. Making the cultural and social environment better prepared to help persons with disabilities and their parents is one of the first objectives for ASPOC families. The ASPOC project takes place in three settings where SENs are required to engage on a regular basis: Family, Shool and Health. In order to prepare SEN participants for job experiences, ASPOCollege seeks to improve cognitive enhancement and continue training residual skills in students who are 19 years old and older. Based on this experience, we are attempting to define best practices of rehabilitation paths suitable for these students using the Computing Module. We are doing this by identifying the best sets of rehabilitation sets to enable cognitive enhancement for these students, as well as by using the pedagogical approach.

2.2 Mixed Reality in Education

In the literature emerges that the usage of AR in general education opens new learning possibilities [1]. For instance, teaching astronomy by using AR tile, virtual 3D modeling package, web camera, projector, and a whiteboard to overtake traditional methods based on books. AR can be used to visualize abstract concepts like airflow, magnetic field, molecules, and so on but also can be used to create AR games that enhance students' engagement. An example is the one provided in the paper where students with an Android tablet are able to collect data on simulated radiation values on their campus which was hypothetically near a Nuclear Power Plant. Also, in [2] the beneficial use of such technology as a means of visualizing abstract concepts, phenomena that are too small, large, fast, or slow to be seen with human eyes, has been demonstrated. To summarize the key points obtained from the examples presented, AR could enable (i) learning content in 3D, (ii) collaborative learning; (iii) learners' senses of presence, immediacy, and immersion, (iv) visualizing the invisible, and (v) bridging formal and informal learning.

AR can also enhance students' knowledge with practical skills that are difficult to obtain with the traditional learning method. The main difference is that AR gives a new sense of learning to the students, enhancing their interest in the topic by immersing them directly inside it, and improving spatial and psychomotor cognitive skills. These outcomes have been obtained from Matt Bower et al. [3] during their research. They pointed out some of the pedagogic aspects that AR can support: (1) constructivist learning: AR makes students more interested in the topic and engaged on a deeper level of concepts, (2) situated learning: AR enables to bring the real world directly inside classrooms, (3) game-based learning: the use of a game-based interface facilitates students in learning some specific topics, like the examples provided in [1], (4) inquiry-based learning based on electronically gather data for future analysis. In [2,4] it is shown with a clear table all the advantages of AR in the educational setting: learning outcomes, pedagogical contributions, interaction, and others.

3 MR Application

3.1 Technical Specification

The concept of the application is that it should teach and guide students during their work using the Sharebot One 3D printer [5].

The application provides to the users all the steps that must be followed to setup and work on the machine from the turn-on to the final stop. These can be summarized as follows:

- Axes movements: with these group of animations the user will understand the axes motion as well as the software commands that are needed to control that movement.
- How it works: represent the procedure needed to turn on the 3D printer and to link it with the corresponding software.
- Procedure to print and replace PLA filament

3.2 Menu and Scenes

The overall application has been developed using the game engine software Unity, while the 3D models representing the twin of the Sharebot One and its tools have been designed using the CAD software Autodesk Inventor. This paper presents the application designed for three different target device: desktop PC, tablet

Android, and Microsoft HoloLens device. Even though the interaction mechanisms are completely different among these technologies, the overall structure of the application and the technical concepts described coincide.

The first scene highlights with a graphical interface the name of the project and its developing partners. The second one shows some technical data about the Sharebot One 3D printer. As can be seen in Fig. 1, these information include the machine's and table's dimensions, the axes strokes, the maximum speed, the resolution, etc. It allows students to easily move to the manufacturer website to obtain further details if necessary. Finally, on the right side, there are a picture of the machine and the list of its optional features. This scene is used to give a brief presentation about the 3D printer in terms of characteristics, dimensions, and performances.

Fig. 1. Scene with the Sharebot One's technical data

4 Desktop and Android Tablet Version

The first version of the application has been developed with a desktop windows PC as target device. The aim of this initial version of the application is to give the students the possibility to understand the Sharebot One with the assistance of an educator before directly see it and work with it in reality. This version can be controlled by the teacher and experienced by the entire class with the adoption of a video projector (Fig. 2 (a).

5 HoloLens R1 Version

The last version of the application was designed to be used with the Microsoft HoloLens R1 device. It is a MR device where the 3D virtual content overlap the

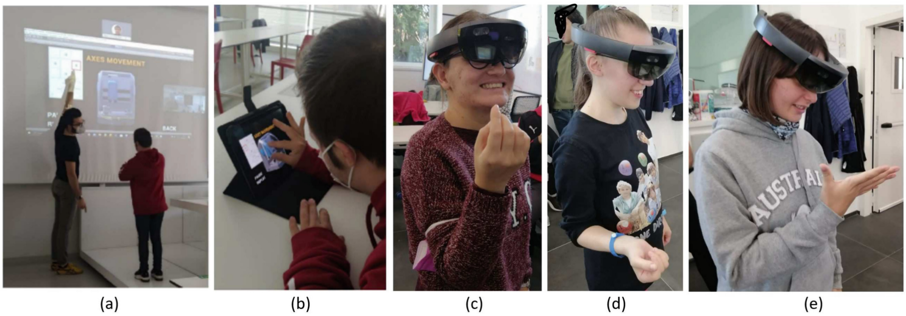


Fig. 2. First interaction with the desktop application thanks to a projector (a) and then the Aspoc student can practice alone with the tablet (b) and the Microsoft HoloLens R1 (c, d, e)

real environment allowing students to interact with both worlds. The interaction mechanism relies on a completely different approach compared to the previous. The user have to point using the head movement an element and select it by using specific gestures, as can be observed in some photographs taken during the evaluation phase with Aspoc students (Fig. 2 (c, d, e).

To allows students to familiarize with this method, a couple of scenes called ‘Before Starting’ and ‘360 Overview’ have been implemented.

6 Results

The application was tested and has been validated by using the System Usability Scale (SUS). 10 Aspoc Students were asked to score the set of questions from SUS that range from strongly agrees to strongly disagree. The SUS score is Calculated, and the Final average score comes out to be ***88.12***, which is an excellent rating for user validation. The average score is calculated by using interpreting scores through a google form survey. Here is the radar chart shown for the usability test as can be seen from Fig. 3.

7 ASPOC Logo 3D Design

Students have learned the basic comands of the Inventor Sorftware in order to create the ASPOC logo in 3D. Then they have created the G-code which is the input for the 3D printer. Figure 4 shows the Aspoc students with the real prototype of the ASPOC logo. In fact this product is used as a gift and can hold a metallic ring in which it’s possible to insert few keys.

For students with intellectual disabilities or pervasive disorders, learning through concrete experience represents an important facilitator. Theory meets practice and makes it tangible. This also trains students to internalize the processes of both Design and printing through an enjoyable experience that lowers the cognitive load and allows for a fruitful learning experience.

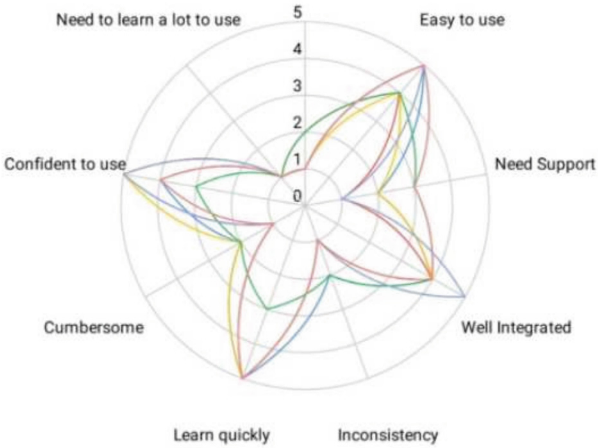


Fig. 3. Results.

Fig. 4. Final Step! Printig their Logo Association!

8 Conclusion

In this study, an Extended Reality applications is presented to help young students with special needs in the laboratory activity learning process, from the design phase through the actual printing of creations. The 3-D Printer application has been configured to meet the needs of the students. Students' engagement and learning can be enhanced by participating actively in the classroom, and this is especially true with laboratory activities involving specialized equipment. However, in order for students to learn and test the procedure, a qualified supervisor and a real machine are needed. This essay seeks to take a step forward by taking into account kids with disabilities who have neurodevelopmental problems. All of the students expressed curiosity and enthusiasm for the application. They love feeling "expert" and utilizing the same tools and machine they are interested in. The training procedure was important to enable students use the software to design the logo and also to print their products using the Sharebot 3D Printer. Overall, however, all of the tools were regarded as having sufficient usability, and it can be anticipated that they can be utilized for educating students with disabilities in a more systematic manner. It might be interesting to follow up to see if the skills students learned, both in using the design software and the printer, persist over time. Observations managed through ICF checklists show how students improved in performance by approaching attainable proficiency in different specific cases; this technology is recordable as a facilitator from an ICF perspective and as a tool that makes learning enjoyable. In the future, it could be interesting to make further evaluations to investigate the level of engagement and the emotional state are given by the interaction, too.

References

1. Wu, H.K., Lee, S.W.Y., Chang, H.Y., Liang, J.C.: Current status, opportunities and challenges of augmented reality in education. Comput. Educ. **62**, 41–49 (2013). https://doi.org/10.1016/j.compedu.2012.10.024
2. Nor Farhah Saidin, N.D.A.H., Yahaya., N.: A review of research on augmented reality in education: advantages and applications. Int. Educ. Stud. **8**(13) (2015). https://doi.org/10.1016/j.edurev.2016.11.002, https://www.sciencedirect.com/science/article/pii/S1747938X16300616
3. Bower, M., Howe, C., McCredie, N., Robinson, A., Grover, D.: Augmented reality in education - cases, places and potentials. Educ. Media Int. **51**(1), 1–15 (2014). https://doi.org/10.1080/09523987.2014.889400
4. Akçayir, M., Akçayir, G.: Advantages and challenges associated with augmented reality for education: a systematic review of the literature. Educ. Res. Rev. **20**, 1–11 (2017). https://doi.org/10.1016/j.edurev.2016.11.002
5. Sharebot: Sharebot NG, Next Generation. https://www.sharebot.it/stampanti-3dsharebot-next-generation/

The Effective Use of Generative AI for Personalized Learning

An Exploratory Study in the Norwegian Context

Simen Frogner Hellesnes(✉), Trym Simensen Nerem, Yavuz Inal, and Carlos Vicient Monllaó

Norwegian University of Science and Technology, Gjøvik, Norway
simenhell@gmail.com, {yavuz.inal,carlos.vicient}@ntnu.no

Abstract. This study explored the possibilities of using generative artificial intelligence (GAI) for text adaptation based on a created persona, i.e., a 12-year-old student with reading comprehension challenges caused by ADHD. The process entailed comparing three outputs generated in three different ways: "out-of-box" use of GAI, GAI-teacher collaboration, and teachers only. Each generated adapted text was evaluated by another group of teachers to understand their suitability for the persona. The results showed that the teacher-adapted text received the highest score, followed by the "out-of-the-box" adapted text and the GAI-teacher collaborative text. The resulting texts utilized different adaptation methods to fit the content towards the persona, engaging the reader, improving the structure, and adjusting the difficulty of the text. However, GAI could be a valid option for personalized learning through text adaptation, considering the close scores between the adapted texts. With the effective use of prompt design, we believe it could assist teachers in adapting content and save them valuable time.

Keywords: Artificial Intelligence · AI · Generative AI · Personalized Learning · ADHD

1 Introduction

When OpenAI released ChatGPT in November of 2022, it opened up new opportunities for advancements in many areas, such as education, programming, writing, and translating. Generative AI (GAI) is a type of artificial intelligence that gained worldwide recognition after the ChatGPT phenomenon. It generates various types of content, such as text, images, script, and audio data. One promising avenue of GAI is in the field of education. Possible use cases include ideation, improving writing skills, aiding creativity, critical thinking, and adaptive learning, e.g., text adaptation.

With the proliferation of widespread and available generative tools, the educational sector has gained access to one of the greatest advancements within personalized learning as they have the capabilities to generate required content based on given inputs. Using GAI, a teacher could easily adapt any text to each student, enabling them to learn from texts on their own level. Previous research has shown that ChatGPT can generate a proper

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 385–392, 2024.
https://doi.org/10.1007/978-3-031-62846-7_46

summary of a reading comprehension text (Vázquez-Cano et al., 2023). This optimization could point towards ChatGPT and other generative tools being useful for summarizing texts for a specific level to support a student-centered approach instead of “one-size-fits-all,” which assumes all students have the same learning paths and capabilities. This possible text adaptation could also improve students’ learning potential with reading disabilities or afflictions such as dyslexia (Michel-Villarreal et al., 2023).

AI can revolutionize pedagogical planning, paving the way for more engaging and individualized learning experiences (Bhutoria, 2022). It could be used in several ways to help with language studies and adaptation, as well as gain insights from large datasets to understand the students’ learning paths and adjust study materials that can follow their specific learning needs and interests (Bhutoria, 2022). This suggests that the potential of AI in education is vast, and it may offer the possibility of providing tailored education for students with various learning difficulties. As a follow-up example, GAI can be particularly impactful for students with attention deficit hyperactivity disorder (ADHD) by implementing some methods and strategies, such as text-to-speech or providing visual aids while reading (Feldt & Mekkelholt, 2020).

In Norway, adapting learning for each student is required by law, and in a Norwegian context, this means to “facilitate with varied assessment forms, learning resources, learning arenas, and learning activities such that everyone gets a satisfactory outcome of the education.” (Utdanningsdirektoratet, 2022b). Norwegian children benefit from individualization; however, teachers are in increasingly short supply as the last four years have seen fewer applicants for teacher education in universities (Samordnaopptak, 2023). Although personalized learning can benefit students, it requires additional effort and time for teachers. In a normal class, teachers need to differentiate for various challenges and levels, which could lead to the teacher being unable to adapt as much as they would like (Lilleås, 2022). Additionally, many Norwegian teachers experience increasing non-educative tasks hindering their pedagogical work (Vik & Korsmo, 2023). This adds to the problem of students not receiving the individualization they require because teachers lack sufficient time to adapt to the degree students need. We hypothesize that the challenge of time deficiency when adapting texts could be solved using GAI. However, no rigorous analysis has been conducted exploring the potential of GAI for personalized learning.

Taken together, the primary objective of this study was to evaluate the capability of GAI in adapting text for a specific target, 12-year-old children with special needs (i.e., reading comprehension challenges caused by ADHD). We assessed the performance of “out of the box” adapted text, GAI-teacher collaborative text, and those crafted by a committee of teachers. The “out of the box” approach refers to using general prompts such as “Adapt the text for a 12-year-old student with reading comprehension challenges caused by ADHD,” while the teachers provided specific and detailed instructions on how the text should be adapted. This comparison aimed to determine whether GAI could match or surpass human-generated adaptations and gain insights into how teachers perceive the efficacy of AI-generated text adaptations.

2 Methods

We conducted an exploratory study to assess the efficacy of using GAI for text adaptation, comparing it to texts adapted by teachers. The primary data collection method involved interviews with teachers across K-10 education levels. We created a persona representing an example student to define our target group. The participants generated an adapted text based on the needs of the persona, who was a student aged 12 years old with a reading difficulty found in a normal class that fit the school level of our teacher group. We focused on ADHD, as it is a learning difficulty faced by 3–5% of Norwegian students (Utdanningsdirektoratet, 2022a). By using this persona, teachers made their adaptations more specific, facilitating the evaluation and scoring of the quality of the outputs. As the baseline for the outputs, we chose a text from a national encyclopedia about sedimentary rocks. The technical level of this text was too high for most 12-year-olds, and it was filled with jargon, meaning that the text would have to be adapted to fit our persona. This ensured that there would be ample opportunity for both GAI and teachers to adapt the text, resulting in obvious differences from the original text.

The study is divided into three phases: The first phase involved a series of interviews where we explored the challenges teachers face through personalized learning and how they would tailor texts to students with ADHD. Through Phase 2, we generated the three adapted texts that were later used in Phase 3. These three outputs were: "out-of-the-box" adapted text (O1), GAI-teacher collaborative text (O2), and teacher-adapted text (O3). In the third phase, these three outputs were presented to a different group of teachers, unaffiliated with the first workshop, for comparison and scoring and for a short interview to gather their perspectives around the texts. All participants in this study, including those in the workshops and interviews, were teachers from the K-10 levels. They all had a minimum of four years of higher education and were currently in teaching positions at Norwegian public schools.

2.1 Data Collection Procedure

Phase 1: Exploratory Interviews on Personalized Learning
In the first phase, we aimed to explore the challenges teachers face through their work of personalized learning to different learning difficulties. We also wanted to learn how they adapt texts to students with ADHD. These interviews were conducted with 13 teachers distributed across 11 different schools and revolved around teachers' current practices and challenges in adapting the educational process.

Phase 2: Generation of Adapted Texts
"*Out-of-the-box*" *Adapted Text: (O1)*
Creating Output 1 was a simple process using the basic version of ChatGPT 4, where a simple prompt regarding text adaptation was given. The chosen text is in Norwegian to align with the rest of the study, so the prompt was also written in Norwegian. We only made one attempt and used a fresh chat, devoid of previous conversations, to ensure that the response would indicate its general knowledge. The prompt was made to be as simple as possible, only giving the task and the basic context. The resulting prompt

was “Adapt this text for a 12-year-old student with reading comprehension challenges caused by ADHD”, with the text in the next paragraph within the same message.

Workshop 1: GAI-Teacher Collaboration (O2)

In the first workshop, we aimed to generate a prompt that would effectively adapt the text to the persona through GAI (i.e., ChatGPT 4). An important quality of the final prompt was that it should function on various texts, and therefore, the teachers were provided two baseline texts on which they could test their prompt; this ensured that the prompt would properly adapt to the texts as intended. Five teachers participated in the workshop and used 45 min to complete the process. They used an iterative method, starting with a simple prompt, critiquing the reply, and discussing what they wanted to change before adapting it. They iterated the process thrice before they were satisfied with the resulting text.

Workshop 2: Teacher-Adapted Text (O3)

The purpose of the second workshop was to create a teacher-generated text. We gathered three teachers in this workshop and tasked them to manually rewrite the original text to fit our persona, focusing on aspects of reading comprehension, especially surrounding adaptation towards ADHD, based on their experience in personalized learning. The teachers spent 45 min evaluating which terms should be included; then, they carefully placed and explained the terms in a way that fit the persona. They also focused on keeping sentences short and concise, dividing the text into multiple paragraphs to help keep their student’s attention. Based on these three outputs, we continued to Phase 3 of the study, where another group of teachers evaluated the outputs.

Phase 3: Teacher Evaluation of Adapted Text

This phase aimed at comparing and evaluating the texts generated in Phase 2. We engaged 16 teachers from two primary schools and one lower secondary school for the interviews. All interviews were conducted in person. In the first part of the interview, the teachers were given the three adapted texts to read through in a random order. After reviewing all texts, they were asked to evaluate the quality of the texts in terms of how well each text was adapted for the persona on a scale of 1 to 7. They were also asked to evaluate what the adapted texts did correctly and highlight any problems encountered.

3 Results

3.1 Phase 1

Results from the first round of interviews (Phase 1) show that many teachers express a high workload or time constraints as primary challenges when adapting education for different reading levels. They mention the work of writing, finding, or adjusting suitable texts as a challenge; some also mention the number of students they need to adapt as a direct cause of their time constraints. The teachers make specific adaptations targeted at students with ADHD. These adaptations can be broadly categorized as handling of texts or organizational strategies. Regarding adapting texts, teachers emphasized the importance of avoiding overly lengthy texts, having a structured format, selecting engaging texts, and ensuring that the texts are tailored to the individuals’ skill level. On an organizational level, there were some differences between elementary and middle schools. In elementary schools, teachers focused on station-teaching for variety's sake and having short breaks to reset students' focus. In middle schools, adaptations were more specific: focusing on group work to increase engagement and understanding, ensuring the tasks were understood, systematic work, and following up closely with students.

3.2 Phase 2

We observed that all the resulting texts from Phase 2 were very different regarding structure, reading difficulty, term explanation, and content to a lesser degree. The first adapted text (O1) was generated with a simple prompt given to ChatGPT 4. This text used metaphors and language akin to a student's adventure, referencing building a sandcastle and playing with glue and glitter. Difficult terms were completely removed, and the text's difficulty level was greatly reduced. Structurally, it was divided into three paragraphs, compared to the two paragraphs from the original text, and did not include any headers.

O2 was generated by the prompt created in the second workshop through GAI-teacher collaboration: "Simplify this text by using simpler words and shorter sentences. Divide the text into several paragraphs with subheadings." (translated from Norwegian). The generated text relied almost solely on structure and short paragraphs. The text was divided into four sections with accompanying subheadings. The text's difficulty level was high as there were many advanced terms without proper explanation, such as non-clastic and metamorphic. The text kept more of the original content than the others.

O3, created by teachers, had a few difficult terms removed, and the remaining terms were well explained. The text's language was greatly simplified and structured into several free-standing sentences and short paragraphs with only one header. Additionally, some of the content from the original text was removed.

3.3 Phase 3

In this phase, we tasked 16 teachers with scoring the three output texts. The results showed that O3 (teacher-adapted text) received an average score of 4.8 (SD = 0.92), followed by O1, the "out-of-the-box" adapted text, (M = 4.7, SD = 1.40) and O2, the GAI-teacher collaborative text, (M = 4.1, SD = 1.45).

O3 was the most consistently scored text, and teachers appreciated its short sentences and good structure (n = 6, 37.5%), good explanations for terms (n = 6, 37.5%), and being more to the point than the other texts (n = 3, 19%), with one saying, "It is easy to understand this text, the explanations are good, and the connections between them are good." However, some teachers (n = 4, 25%) thought it was too stiff and boring, with one summing it up with: "*Might not be so fun to read.*"

O1 was praised for being good at relating to the persona and creating interest (n = 8, 50%), as well as for using vivid comparisons to explain the texts' themes (n = 6, 37.5%), with one saying: "*It has good descriptions, I envisioned building the sandcastle as I read the text.*" However, it was mentioned that the text was poorly structured (n = 8, 50%), would have benefited from having more relevant terms or academic content (n = 5, 31%), and was a bit childish (n = 2, 12.5%).

O2's most mentioned strength was its structure (n = 11, 69%), with one explaining that "*I would have written this text myself; I really like the structure and that there are* *many small paragraphs.*" Where O2 falls short is in its language and explanations (n = 13, 81%). A teacher claimed that students would give up when reading the text by saying, "*There were a lot of difficult words, a 12-year-old wouldn't understand this, and they would give up before they started reading.*"

4 Discussion and Conclusion

Adapting educational content for individual students is a time-consuming and complicated process. Through our interviews, we found high workload and time constraints to be the main challenges the teachers face when adapting content for different reading levels. This result aligns with previous research (e.g., Pedersen et al., 2022; Nerem et al., 2024), where teachers expressed being overwhelmed with work and insufficient time to provide their students with the best learning. Our findings also paint a picture of the complex work of personalizing learning for students, both when adapting text and when planning the organization of a class. Together these findings corroborate the claim that teachers need a solution to be able to support their work with personalized learning.

The teachers' opinions on how to adapt texts for students with ADHD were in line with suggestions from Phalke and Shrivastava (2022): improving the structure, engaging the reader, and adjusting the difficulty of the text. We observed that each of the generated texts in our study added extra focus on one or more of these: O1 focused on engaging the reader as well as adjusting the difficulty, O2 prioritized improving the structure, and O3 emphasized adjusting the difficulty of the text, especially through length of sentences and careful consideration of terms.

The teacher-adapted text (O3) received the highest score, followed closely by the "out-of-the-box" adapted text (O1). Both texts focused on simplifying the content to a large degree but in different ways. O3 included less difficult terms and had simple explanations, whereas O1 used metaphors and simple language. Additionally, O1 had more variance in the teachers' opinions, either evaluated as very good or bad, whereas O3 was consistently scored as adequate or better. The GAI-teacher collaborative text (O2) received the lowest score and had a similar variance in the score as O1. One significant difference compared to O1 was the prompt. In O1, the prompt instructed GAI to adapt the original text to a student with ADHD. O2's prompt did not mention ADHD but instead contained instructions to simplify the language and structure the text. This focus on structure could cause overfitting in the prompt, and the instructions for simplifying the text could be too ambiguous and lack context. Both of these are mentioned as potential pitfalls of prompt engineering by Giray (2023), resulting in GAI being too focused on structure in the adaptation. The GAI-teacher generated text (O2) was well-structured, as agreed by most teachers, but the information in the text was not as well-liked as O1 and O3, with several comments mentioning its complicated language and lackluster explanations.

We initially expected the GAI-teacher collaborative text to outperform "out-of-the-box" adapted text. However, proper prompt use has crucial effects on the quality of outputs. In our study, teachers' lack of experience using AI to generate adaptive content may have affected the results. An additional factor might have been insufficient time to test and develop the prompt. This could easily be mitigated by providing more in-depth education on the effective use of prompts for teachers, and a more detailed and well-made prompt could have shifted the scores in O2s favor.

Time is an important factor for overworked teachers. In the workshop, they spent 45 min to make the adapted text. The teachers in the second workshop spent the same amount of time making a generalized prompt and generating the text. The difference here lies in the next adaptations, where teachers would have to spend a significant amount of time to adapt it. GAI could reuse the same prompt and generate the text within a minute. The teacher-written text could be slightly better, but following the close scores found in this study, it is a tempting proposition to save time, highlighting a great advantage of GAI-assisted personalized learning.

Comparing the two GAI-generated texts, we observe that they each excel at one aspect of adaptation, language simplification, and text structuring. This suggests that a more optimized prompt could have succeeded at implementing both of these changes to a tailored text, resulting in an overall more suitable text for the target student. In combination with the "out-of-the-box" text already being capable of challenging teacher-written texts, this shows that GAI-generated texts have the potential to save teachers a significant amount of time by being an alternative to manually adapting texts, and the technology could feasibly be taken further by working on prompt design. The scope of this study was to explore whether GAI can automate the process of adapting texts by gauging teacher's opinions. Results are promising, but future research should investigate how improved prompt design could assist teachers in text adaptation for individual students and testing the effectiveness of AI-based text adaptation on students.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

Bhutoria, A.: Personalized education and artificial intelligence in the United States, China, and India: a systematic review using a human-in-the-loop model. Comput. Educ. Artif. Intell. **3**, 100068 (2022)

Feldt, Z., Mekkelholt, E.: Different methods and strategies to aid reading for pupils affected by ADHD. [Master's Thesis]. Malmö universitet (2020). https://urn.kb.se/resolve?urn=urn:nbn:se:mau:diva-34281

Giray, L.: Prompt engineering with ChatGPT: a guide for academic writers. Ann. Biomed. Eng. **51**(12), 2629–2633 (2023)

Lilleås, R.: Tilpasset opplæring for elever med ADHD i naturfag. [Master's thesis]. NTNU (2022). https://hdl.handle.net/11250/3018256

Michel-Villarreal, R., Vilalta-Perdomo, E., Salinas-Navarro, D.E., Thierry-Aguilera, R., Gerardou, F.S.: Challenges and opportunities of generative AI for higher education as explained by ChatGPT. Educ. Sci. **13**(9), 856 (2023)

Nerem, T.S., Hellesnes, S.F., Inal, Y., Vicient Monllaó, C.: The use of artificial intelligence for personalized learning: teacher perspective. In: The Seventeenth International Conference on Advances in Computer-Human Interactions, Barcelona, Spain, May 2024

Pedersen, C., Reiling, R.B., Vennerød-Diesen, F.F., Alne, R., Skålholdt, A.: Evaluering av norm for lærertetthet: Sluttrapport (2022)

Phalke, S.S., Shrivastava, A.: HI applications for ADHD children: a case for enhanced visual representations using novel and adapted guidelines. In: Mahmud, M., Mendoza-Barrera, C., Kaiser, M.S., Bandyopadhyay, A., Ray, K., Lugo, E. (eds.) Proceedings of Trends in Electronics and Health Informatics. TEHI 2022. LNNS, vol. 675, pp. 207–226. Springer, Singapore (2023). https://doi.org/10.1007/978-981-99-1916-1_16

Samordnaopptak. (2023). Flere velger økonomi og IKT - færre vil bli lærer. https://www.samordnaopptak.no/info/om/sokertall/pressemelding-240423-bm.pdf

Utdanningsdirektoratet. (2022a). ADHD. https://www.udir.no/laring-og-trivsel/spesialpedagogikk/spesialpedagogiske-fagomrader/adhd

Utdanningsdirektoratet. (2022b). Tilpasset opplæring. https://www.udir.no/laring-og-trivsel/tilpasset-opplaring/

Vázquez-Cano, E., Ramírez-Hurtado, J.M., Sáez-López, J.M., López-Meneses, E.: ChatGPT: the brightest student in the class. Think. Skills Creat. **49**, 101380 (2023)

Vik, M.G., Korsmo, E.K.: Arbeidstidsavtalen: Krever at lærere får være lærere (2023). https://www.utdanningsforbundet.no/nyheter/2023/arbeidstidsavtalen-krever-at-larere-far-vare-larere/

Blind and Low Vision: Orientation and Mobility

Accessible Mobility for Persons with Disabilities

Introduction to the Special Thematic Session

Alireza Darvishy(✉)

Zurich University of Applied Sciences, 8401 Winterthur, Switzerland
dvya@zhaw.ch

Abstract. This paper presents and summarizes nine Special Thematic Session (STS) papers focusing on navigational assistive technologies for persons with disabilities, particularly for blind and visually impaired (BVI) individuals. The selected papers cover a wide range of topics, including step length estimation, utilization of technology in orientation and mobility (O&M), thermal-tactile biofeedback, ICT trends in extraocular muscle prosthesis implants, sound-based navigation, indoor wayfinding applications, digital tools for navigation training, AI-based sensor orientation correction, and accessible wayfinding tools. Along with summarizing their approaches and key findings, this paper touches on the common methods, themes, and emerging trends across the papers.

Keywords: People with Disabilities · Visual Impairment · Navigation · Wayfinding · Mobility · Assistive Technologies

1 Introduction

According to statistics from the World Health Organization, approximately one-sixth of the global population is blind or visually impaired (BVI), and this number is on a significant upward trend [1]. Like anyone else, BVI individuals may require some assistance with everyday tasks such as exploring unfamiliar areas. Mobility is one important aspect of independent life. Yet for BVI individuals, travelling and navigating alone, whether indoors or outdoors, can present significant challenges. Traditional methods such as guide dogs and white canes offer only partial support for independent mobility. In recent decades, technological advancements have vastly expanded the range of solutions available [2]. Technologies like GPS, smartphone and Artificial Intelligence (AI) systems are revolutionizing wayfinding and safe navigation for BVI individuals, providing increasingly convenient and accurate solutions.

The papers presented in this session touch on a broad array of approaches. Two papers [4, 6] present overviews of existing digital solutions and emerging trends. Others present findings from user surveys, evaluations, and prototype testing. Built-in smartphone sensors are featured in two papers [3, 8]. The adaptability and accuracy offered by AI- and machine learning-based systems is also explored in 2 papers [3, 10]. Another shared theme is the integration of technology in orientation and mobility training, which

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 395–399, 2024.
https://doi.org/10.1007/978-3-031-62846-7_47

is emphasized in three papers [4, 9, 10]. Paper 9 addresses indoor wayfinding for a public museum setting, while [5] and [7] present solutions based on thermal-tactile biofeedback and sound to aid the user in navigation.

Notably, user-centric approaches feature prominently across the papers. The majority of authors explicitly emphasize the importance of considering user needs and preferences in the design and evaluation of orientation and mobility solutions.

2 Session Papers

The paper by Elyasi and Manduchi, entitled "Step Length Estimation for Blind Walkers" [3], focuses on step length estimation for blind walkers using inertial sensing, essential for indoor wayfinding where GPS isn't viable. Utilizing data from smartphone sensors, particularly accelerometers and gyroscopes, algorithms estimate location based on step count and length. However, maintaining a constant step length is challenging, especially for blind walkers. Previous work demonstrated machine learning algorithms' efficacy in step length estimation for sighted individuals. This study extends this by testing a similar algorithm on seven blind walkers. Results indicate that training the algorithm on data from blind individuals yields better results compared to sighted individuals. The findings suggest the feasibility of integrating such algorithms into inertial-based wayfinding systems for blind individuals, enhancing their independent navigation indoors.

In their paper "Use of Technology and Applications in Orientation and Mobility of Visually Impaired Persons in Bulgaria: A Contemporary Overview" [4], authors Tzvetkova-Arsova and Tomova provides an overview of the utilization of technology and applications in orientation and mobility (OM) for visually impaired individuals in Bulgaria. They discuss the historical development of OM programs to enhance independence for visually impaired individuals. Notable technologies include the WeWalk cane, the Lazarillo GPS application, and the Be My Eyes volunteer assistance app, as well as local public transportation apps. The research methodology involves distributing questionnaires to organizations and visually impaired individuals to assess the availability and usage of these technologies. Preliminary findings suggest a range of available technologies, with some being widely used, and highlight the importance of self-teaching and OM instructor support in learning to use these applications.

Frank and Kurschl [5] discuss the development and user survey of a prototype orientation aid for BVI individuals using thermal-tactile biofeedback at the lumbar region to prevent hazards. The prototype utilizes thermal modules to provide directional instructions and stop commands based on obstacle detection, aiming to enhance user protection and navigation. The study involved eight BVI participants, evaluating their perception thresholds and reactions to the thermal signals. Results indicate a reliable perception threshold for stop commands and directional signals, but synchronization issues between camera tracking and biofeedback were identified as major causes for collisions. Overall, participants showed positive surprise regarding thermal biofeedback, but concerns were raised about distractions and limited protection in certain scenarios, suggesting the prototype's potential but also highlighting areas for improvement and further research.

In "Enhancing Accessibility through ICT Trends for Extraocular Muscles Prosthesis Implant: A Comprehensive Literature Review," [6] authors Verma, Miesenberger and

Priglinger provide a comprehensive literature review on ICT trends in Extraocular Muscles (EOM) prosthesis implant surgeries, focusing on medical imaging, 3D modeling algorithms, and biomedical simulators. It discusses the challenges associated with EOM dysfunction, such as strabismus and nystagmus, and categorizes assistive technologies into invasive and non-invasive types, with EOM prosthesis implant falling under the invasive category. The study emphasizes the importance of precision in these surgeries and proposes the use of biomedical simulators for diagnosis and surgical planning. Furthermore, it highlights the need to integrate principles of Accessibility Technology to address broader accessibility concerns and advocates for the development of medical digital twins to enhance accuracy and predictive capabilities. Through a systematic literature review and analysis of research questions, the paper provides insights into imaging modalities, databases, algorithms, and mathematical models relevant to EOM issues, contributing to advancements in diagnosis, treatment, and assistive technologies for strabismus patients.

The paper by Erdenesambuu et al. [7] addresses the challenges faced by visually impaired individuals in navigating their environments independently due to the reliance on visual information in urban infrastructure. While Braille blocks have been traditionally used as a tactile aid, they offer limited information and may not be suitable in all situations. To overcome these limitations, the study proposes an accessible navigation system that utilizes sound effects and voice guidance to convey information effectively. By leveraging the superior auditory discrimination abilities of visually impaired individuals, the system aims to provide accurate and efficient route guidance. A walking simulator was developed to evaluate the proposed method, and experiments with visually impaired participants indicated that audio beacons based on sound effects may offer better accuracy in route understanding compared to traditional voice guidance. However, both methods were found to have similar usability within the study's scope. Future research aims to evaluate the proposed audio beacon under more realistic conditions to further enhance navigation assistance for visually impaired individuals.

Das and Hong's paper "FindMyWay: Toward Developing an Accessible Wayfinding Application for Visually Impaired in an Indoor Space" [8] introduces an iOS application designed to assist visually impaired individuals with indoor wayfinding using bluetooth low-energy beacons. Traditional wayfinding methods face challenges indoors due to the limitations of GPS signals, prompting the exploration of alternative technologies. FindMyWay employs a zone-based proximity approach and parallel processing through multithreading to improve localization accuracy and navigation performance. Testing scenarios, including exploration of points of interest and navigation across multiple floors, demonstrate the app's functionality and effectiveness. The study concludes that FindMyWay shows promise in addressing indoor wayfinding challenges for the visually impaired, with plans for further usability testing with visually impaired users.

Façanha et al. [9] introduce innovative digital tools aimed at enhancing OM training for blind individuals. The proposed system includes a web-based map editor and a mobile application that creates audio-haptic OM training environments. Through extensive evaluation involving human-computer interaction specialists, OM instructors, and PWB, the system's effectiveness in improving navigation skills is demonstrated. Feedback from users highlights the potential of these tools in promoting learner autonomy and independence by enabling customized training experiences tailored to individual needs. The study underscores the significance of further refinement to enhance user interface intuitiveness and audio-haptic feedback accuracy, emphasizing the importance of ongoing research and development in empowering PWB in navigation and spatial awareness.

The paper "An AI-based System for Predicting and Correcting the Sensor-orientation of an Electronic Travel Aid during use by a Visually Challenged Person," [10] by G. Singh et al., introduces an AI-based system designed to predict and correct sensor orientation in electronic travel aids used by visually impaired individuals. The system aims to reduce dependence on trainers and facilitate self-learning by continuously detecting and guiding users to correct sensor orientation through audio-vibratory feedback. By utilizing an embedded computing unit mounted on a cane equipped with an inertial measurement device, the system trains a neural network to predict orientation based on velocity and acceleration data. The neural network models, including LSTM and transformer architectures, are deployed on the device for real-time operation. Through self-supervised training data generation and on-device implementation, the system achieves accurate predictions and successful orientation correction. Evaluation results demonstrate high accuracy and real-time performance of the neural networks, indicating the system's potential to enhance mobility for visually impaired individuals.

Finally, Erdemli and Collins' paper "Accessible Wayfinding Tool" [11] discusses the co-design of an accessible wayfinding tool for public museums. It reviews existing wayfinding tools used by BVI folks, assessing their accessibility features. The study proposes a novel accessible wayfinding tool and outlines the methodology for evaluating its effectiveness through questionnaires, interviews, and collaborative discussions with BVI participants and specialists. Findings indicate interest among BVI individuals in using directional cues and audio descriptions to improve map accessibility, with plans to evaluate the tool with 10 BVI participants. The study underscores the importance of addressing issues such as screen reader support, visual hints, audio descriptions, and scaling methods to enhance overall accessibility for BVI individuals in digital wayfinding experiences.

3 Conclusion

There has been rapid development on the topic of autonomous mobility for BVI individuals and others with disabilities. Overall, future research in this area should focus on integrating emerging technologies, such as machine learning, AI, and augmented reality, to develop more effective and user-friendly solutions for enhancing the mobility and independence of BVI individuals in different environments, both indoors and outdoors. Additionally, collaboration between researchers, developers, OM specialists, and visually impaired individuals is crucial for ensuring that the developed solutions meet the real-world needs of the users.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. World Health Oraganization. Visual impairment and blindness. WHO Fact Sheet, 282. Geneva (Switzerland): WHO; 2010
2. Messaoudi, M.D., Menelas, B.J., Mcheick, H.: Review of navigation assistive tools and technologies for the visually impaired. Sensors (Basel) **22**(20), 7888 (2022). PMID: 36298237; PMCID: PMC9606951. https://doi.org/10.3390/s22207888
3. Elyasi, F., Manduchi, R.: Step length estimation for blind walkers. In: Miesenberger, K., Kobayashi, M., Penaz, P. (eds.), ICCHP 2024, Computers Helping People with Special Needs, 19th International Conference, LNCS, vol. 14750, pp. 400–407. Linz, Austria, Proceedings. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-62846-7_48
4. Tzvetkova-Arsova, M., Tomova, M.: Use of technology and applications in orientation and mobility of visually impaired persons in Bulgaria: a contemporary overview. In: Miesenberger, K., Kobayashi, M., Penaz, P. (eds.), ICCHP 2024, Computers Helping People with Special Needs, 19th International Conference, LNCS, vol. 14750, pp. 408–415. Linz, Austria, Proceedings. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-62846-7_49
5. Frank, V., Kurschl, W.: Orientation aid for blind and visually impaired people by using thermal-tactile biofeedback at lumbar region for hazard prevention: a user survey. In: Petz, A., Miesenberger, K. (eds.), ICCHP Open Access Compendium,Future Perspectives on Accessibility, AT and eInclusion. Linz, Austria, Proceedings. Johannes Kepler University Linz (2024)
6. Verma, A., Miesenberger, K., Priglinger, S.: Enhancing accessibility through ICT trends for extraocular muscles prosthesis implant: a comprehensive literature review. In: Miesenberger, K., Kobayashi, M., Penaz, P. (eds.), ICCHP 2024, Computers Helping People with Special Needs, 19th International Conference; Linz, Austria, Proceedings. Springer, Cham (2024)
7. Erdenesambuu1, D., Matsuo1, M., Miura, T., Sakajiri1, M., Onishi, J.: Advancing mobility for the visually impaired: a virtual sound-based navigation simulator interface. In: Miesenberger, K., Kobayashi, M., Penaz, P. (eds.), ICCHP 2024, Computers Helping People with Special Needs, 19th International Conference, LNCS, vol. 14750, pp. 416–427. Linz, Austria, Proceedings. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-62846-7_50
8. Das, U., Hong, B.: FindMyWay: toward developing an accessible wayfinding application for visually impaired in an indoor space. In: Miesenberger, K., Kobayashi, M., Penaz, P. (eds.), ICCHP 2024, Computers Helping People with Special Needs, 19th International Conference, LNCS, vol. 14750, pp. 428–435. Linz, Austria, Proceedings. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-62846-7_51
9. Façanha, A, Moraes Júnior, J. Carneiro Araújo, M., Ferreira, J., Viana, W., Sanchez, J.: Empowering Orientation and Mobility Instructors: Digital Tools for Enhancing Navigation Skills in people who are blind. In: Miesenberger, K., Kobayashi, M., Penaz, P. (eds.), ICCHP 2024, Computers Helping People with Special Needs, 19th International Conference, LNCS, vol. 14750, pp. 436–443. Linz, Austria, Proceedings. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-62846-7_52
10. Singh, G., Nadir, M., Thukral, R., Gambhir, V., Chanana, P., Paul, R.: An AI-based system for predicting and correcting the sensor-orientation of an electronic travel aid during use by a visually challenged person. In: Miesenberger, K., Kobayashi, M., Penaz, P. (eds.), ICCHP 2024, Computers Helping People with Special Needs, 19th International Conference; Linz, Austria, Proceedings. Springer (2024)
11. Erdemli, M., Collins, K.: Accessible wayfinding tool. In: Miesenberger, K., Kobayashi, M., Penaz, P. (eds.), ICCHP 2024, Computers Helping People with Special Needs, 19th International Conference; Linz, Austria, Proceedings. Springer (2024)

Step Length Estimation for Blind Walkers

Fatemeh Elyasi and Roberto Manduchi(✉)

University of California, Santa Cruz, USA
{felyasi,manduchi}@ucsc.edu

Abstract. Wayfinding systems using inertial data recorded from a smartphone carried by the walker have great potential for increasing mobility independence of blind pedestrians. Pedestrian dead-reckoning (PDR) algorithms for localization require estimation of the step length of the walker. Prior work has shown that step length can be reliably predicted by processing the inertial data recorded by the smartphone with a simple machine learning algorithm. However, this prior work only considered sighted walkers, whose gait may be different from that of blind walkers using a long cane or a dog guide. In this work, we show that a step length estimation network trained on data from sighted walkers performs poorly when tested on blind walkers, and that retraining with data from blind walkers can dramatically increase the accuracy of step length prediction.

Keywords: Wayfinding · Odometry · Navigation · Pedestrian Dead Reckoning

1 Introduction

Independent wayfinding can be challenging for people who are blind. While GPS localization can be very helpful in outdoor environments, GPS cannot be used inside buildings, and thus different mechanisms for localization and wayfinding need to be relied upon. In this work we focus on inertial sensing for localization. Inertial sensors (accelerometers and gyros) are contained in any standard smartphone. Data from these sensors can be used by pedestrian dead-reckoning (PDR) algorithms to estimate the user's location given a known starting point. One important advantage of inertial navigation is that it doesn't require an external infrastructure, as is the case, for example, for wayfinding systems based on Bluetooth Low Energy (BLE) beacons. Compared to vision-based systems, inertial sensing has the advantage that it does not require the user to hold the smartphone such that the camera gets a clear view of the environment. Users can conveniently keep their smartphone in their pocket, and be tracked by the system as they move around.

Standard PDR algorithms use inertial data to count the number of steps taken by the user, and to determine the walking direction. By multiplying the number of steps by each step's length, the distance traversed in a certain period of time can be determined. This approach, however, assumes that the length of each step be known in advance. This requires a prior calibration phase, during which

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 400–407, 2024.
https://doi.org/10.1007/978-3-031-62846-7_48

the "natural" step length of the user is computed. However, even after calibration, it is quite possible that the user may not maintain a constant step length. For example, walkers may take shorter steps when they are unsure of where they are going, or when negotiating an obstacle. In these cases, the odometry system would make localization errors and consequently provide wrong directions.

In prior work [4], Elyasi et al. presented a machine learning algorithm to estimate the length of each step taken by a walker, based on the same inertial data that is used by the smartphone-based PDR algorithm. This algorithm, which employs an LSTM recurrent network, was originally tested on sighted walkers, and was shown to produce rather accurate results. In this contribution, we present a study in which tested a similar algorithm on 7 blind walkers (5 using a long cane and 2 using a dog guide). Note that the gait of a blind walker using a cane is typically different from that of a sighted walker, and it is also different from that of walkers using a dog guide. It is thus important that the step length prediction system be tested with data from walkers from the same communities the wayfinding is designed for.

2 State of the Art

Inertial-based wayfinding systems for blind travelers were described in [1,12]. In particular, [1] proposed online estimation of step length in the context of Particle Filtering. Traditional methods for step length estimation were based, for example, on the relationship between the step length and the difference of max and min values of the vertical acceleration within the step [15]. Other methods used the magnitude of acceleration [7] or its local variance [9]. [3,11] used a combination of the user's step frequency and height. More recent approaches are based on machine learning. For example [6] used stacked autoencoders to learn valuable features from accelerometer and gyro data through stacked autoencoders, while [16] used deep belief networks. StepNet [8] uses a combination of high-level features with a convolutional neural network (CNN)-based. Wand et al. [14], used a combination of LSTM and autoencoder model. Elyasi et al. fed data from accelerometer and gyro to an LSTM followed by a fully convolutional layer.

3 Methodology

Our goal was to verify whether the algorithm of [4], applied to inertial data from a smartphone carried by blind users, could produce reliable step length measurements for our blind participants. This algorithm proceeds as follows. First, the recorded time series of inertial data is segmented into "steps", defined as the interval of time between two consecutive heel strikes from opposite feet. Heel strikes where computed using the algorithm of [10]. Then, from the data within a step period, a step length is inferred using an LSTM-based algorithm. This network was trained using ground-truth data recorded from foot-mounted sensors (Xsens DOT IMU packages, each tied to either shoe using an elastic band).

As well known [5], zero-velocity updates can be applied to data from foot-level sensors for precise dead-reckoning. We applied the ZUPT algorithm [5], along with HDR correction [2] to reduce gyro drift, to the data from the foot sensors to reconstruct the trajectory of the users' motion. From this, we computed the ground-truth length of each step, which was used in the loss function considered when training the step length measurement network from smartphone data. For more details, please refer to [4].

We recruited seven blind participants for this test (4 female, 3 male). Two participants (P1 and P7) used a dog guide, while the others used a long cane. The participants' ages raged from 53 to 76. Each participant walked with an iPhone 12, which was used to record inertial data. tucked in a pants pocket. In addition, they were equipped with the two foot sensors mentioned earlier. Each participant walked for 292 m through the corridors of a building. While walking, they kept the data-collecting iPhone in their back pocket. The histogram of ground-truth recorded step lengths is shown in Fig. 1, together with the histogram of step lengths recorded in the experiment with sighted participants described in [4].

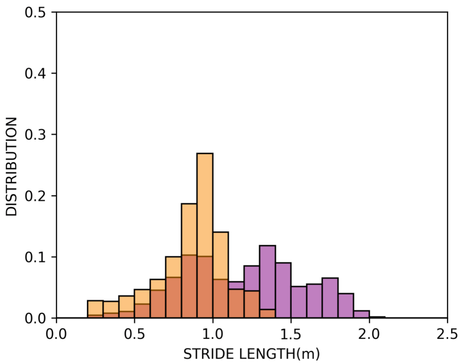


Fig. 1. Distribution of stride lengths for all blind participants in our study (orange bars), shown alongside the stride length distribution for the data set of sighted participants in [4] (purple bars). (Color figure online)

We considered four different training modalities for training the LSTM algorithm of [4]:

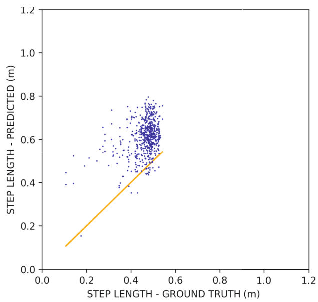
- *Train on Sighted (TS)*: In this modality, we used the original model of [4], which was trained on data from sighted walkers, and tested it on the data collected from all 7 blind participants.
- *Train in Community (TC)*: In this case, each walker using a long cane was tested with a system trained from all the 6 other walkers long cane users. Likewise, each walker with a dog guide was tested with a system trained on data from the other dog guide user.
- *Train on Blind (TB)*: Each blind participant was tested with a system trained on data from all other blind participants (regardless of whether they used a long cane or a dog guide).

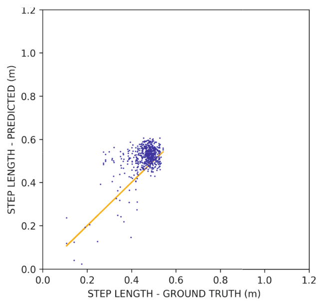
– *Train on All (TA)*: Each blind participant was tested with a system trained on all sighted participants, as well as all other blind participants.

Note that *TC*, *TB* and *TA* use the "leave-one-person-out modality". In all cases, data from a certain walker was never used for training the network tested on that user.

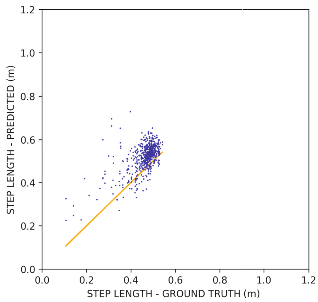
4 Results

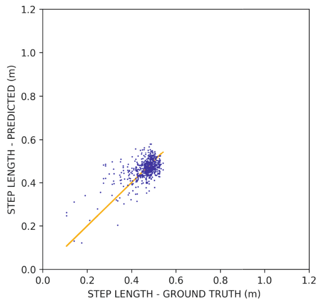
We processed the data recorded from the participants' iPhone to estimate each individual step length, then compared the results with the ground-truth step lengths from the foot-level sensors. Figures 2 and 3 show examples of estimated step lengths plotted against their ground-truth values for different training modalities. In these plots, the red line represents the locus of zero error estimations.





(a): $P1, TS : GD$ ($E_{sr} = 0.355$) (b): $P1, TC : GD$ ($E_{sr} = 0.15$)

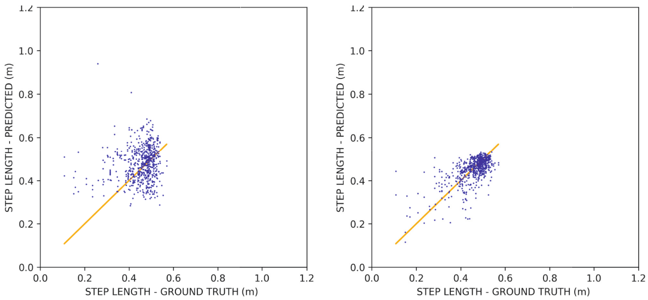




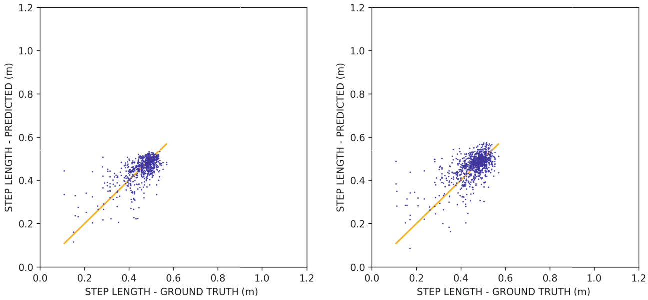
(c): $P1, TA : GD$ ($E_{sr} = 0.15$) (d): $P1, TB : GD$ ($E_{sr} = 0.09$)

Fig. 2. Examples of step length prediction for $P1$ (a dog guide user) for different Training/Test modalities, plotted against their ground truth values. (Color figure online)

To quantify the step length errors (difference between estimated and ground truth values), we e used the error metrics defined in [4], which are summarized in the following.



(a): $P3, TS : LC$ ($E_{sr} = 0.19$) (b): $P3, TC : LC$ ($E_{sr} = 0.10$)



(c): $P3, TA : LC$ ($E_{sr} = 0.11$) (d): $P3, TB : LC$ ($E_{sr} = 0.11$)

Fig. 3. Examples of step length prediction for $P3$ (a long cane user) for different Training/Test modalities, plotted against their ground truth values. (Color figure online)

– $E_d = \frac{|\sum_{i=1}^{N} l_i - \sum_{i=1}^{N} \hat{l}_i|}{\sum_{i=1}^{N} \hat{l}_i}$

– $E_s = \frac{1}{N} \sum_{i=1}^{N} |l_i - \hat{l}_i|$

– $E_{sr} = \frac{1}{N} \sum_{i=1}^{N} \frac{|l_i - \hat{l}_i|}{\hat{l}_i}$

– $R^2 = 1 - \frac{\text{RMSE}^2}{\sigma^2}$, where $\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(l_i - \hat{l}_i)^2}{N}}$ and σ^2 is the variance of the set of ground truth step lengths $\{\hat{l}_i\}$.

In these formulas, $\hat{l}_i$ represents the ground truth length of the i-th step in the test set for a certain participant, while l_i is the step length predicted by the LSTM system. E_d is the relative distance error, which is an appropriate measure for long paths where step-to-step error fluctuations tend to cancel out. E_s represents the average absolute error at each step. E_{sr} normalizes errors with the ground truth step length. R^2 (often called "coefficient of determination") is a number that is equal to 1 only in case of zero error, and is less than 1 otherwise.

A negative R^2 means that the prediction $\{l_i\}$ gives a worse RMSE than using a constant value equal to the average step length.

The average errors are reported in Table 1 for the participants using a long cane, and in Table 2 for the participants using a dog guide.

Table 1. Error metrics computed for all Training/Test modalities for long cane users.

	E_d	E_s (m)	E_{sr}	R^2
TS	0.14 ± 0.10	0.09 ± 0.03	0.29 ± 0.11	-0.95 ± 0.96
TC	0.02 ± 0.01	0.05 ± 0.01	0.14 ± 0.03	0.48 ± 0.08
TB	0.03 ± 0.02	0.05 ± 0.01	0.14 ± 0.03	0.43 ± 0.12
TA	0.02 ± 0.01	0.05 ± 0.01	0.15 ± 0.03	0.44 ± 0.03

Table 2. Error metrics computed for all Training/Test modalities for guide dog users.

	E_d	E_s (m)	E_{sr}	R^2
TS	0.19 ± 0.18	0.14 ± 0.01	0.353 ± 0.03	-3.79 ± 4.64
TC	0.18 ± 0.10	0.11 ± 0.07	0.23 ± 0.11	-0.75 ± 0.03
TB	0.07 ± 0.10	0.07 ± 0.04	0.15 ± 0.08	0.25 ± 0.07
TA	0.07 ± 0.08	0.07 ± 0.01	0.17 ± 0.02	-0.22 ± 0.94

5 Discussion

The main goal of this test was to evaluate whether a step length measurement system trained on sighted walkers would work well for blind walkers using a long cane or a dog guide, or retraining was in order. We hypothesized that the gait of blind walkers could be different from that of sighted walkers, motivating our research question. The results presented in Tables 1 and 2 show that the largest errors were observed when the network was trained with sighted walkers, while better values were obtained when data from blind walkers were included in the training set. Indeed, when the system was trained solely with data from sighted users, the value of R^2 was negative for both long cane and dog guide users, meaning that using a constant value equal to the average step length for each user would give a lower mean square error than using the prediction from the system. This can also be seen qualitatively in Fig. 2 (a) and Fig. 3 (a), which show a large spread of step length predictions for the same ground-truth value.

We note that for dog guide users, the *TC* modality resulted in a negative value for R^2. This is likely because in this case the system was only trained with data from only one other user (since there were only two dog guide users in our set of participants). The best results are obtained with the *TB* modality, where for each participant, data from all other blind participants was considered.

6 Conclusions

Step length estimation is a critical component of a PDR system used for localization and wayfinding in indoor environments. Prior work showed the feasibility of recurrent neural networks for step length estimation from inertial data. However, this algorithm was only tested with sighted walkers. Our study has shown that, with proper training, a similar architecture can be used successfully for blind walkers, and can thus be integrated in a complete inertial-based wayfinding system. In future work, we will integrate a properly trained step length measurement system in the indoor wayfinding and backtracking app described in [13].

Acknowledgment. Research reported in this article was supported by the National Eye Institute of the National Institutes of Health under award number R01 EY029260-01. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. Apostolopoulos, I., Fallah, N., Folmer, E., Bekris, K.E.: Integrated online localization and navigation for people with visual impairments using smart phones. ACM Trans. Interact. Intell. Syst. (TiiS) **3**(4), 1–28 (2014)
2. Borenstein, J., Ojeda, L., Kwanmuang, S.: Heuristic reduction of gyro drift for personnel tracking systems. J. Navigation **62**(1), 41–58 (2009)
3. Chen, R., Pei, L., Chen, Y.: A smart phone based PDR solution for indoor navigation. In: Proceedings of the 24th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS 2011), pp. 1404–1408 (2011)
4. Elyasi, F., Manduchi, R.: Step length is a more reliable measurement than walking speed for pedestrian dead-reckoning. In: 2023 13th International Conference on Indoor Positioning and Indoor Navigation (IPIN), pp. 1–6. IEEE (2023)
5. Foxlin, E.: Pedestrian tracking with shoe-mounted inertial sensors. IEEE Comput. Graph. Appl. **25**(6), 38–46 (2005)
6. Fuqiang, G., Khoshelham, K., Chunyang, Yu., Shang, J.: Accurate step length estimation for pedestrian dead reckoning localization using stacked autoencoders. IEEE Trans. Instrum. Meas. **68**(8), 2705–2713 (2019)
7. Kim, J.W., Jang, H.J., Hwang, D.H., Park, C.: A step, stride and heading determination for the pedestrian navigation system. J. Glob. Position. Syst. **01** (2004)
8. Klein, I., Asraf, O.: StepNet-deep learning approaches for step length estimation. IEEE Access **8**, 85706–85713 (2020)
9. Ladetto, Q.: On foot navigation: continuous step calibration using both complementary recursive prediction and adaptive Kalman filtering (2000)
10. Ren, P., Elyasi, F., Manduchi, R.: Smartphone-based inertial odometry for blind walkers. Sensors **21** (2021). 06/2021
11. Renaudin, V., Susi, M., Lachapelle, G.: Step length estimation using handheld inertial sensors. Sensors **12**(7), 8507–8525 (2012)
12. Riehle, T.H., Anderson, S.M., Lichter, P.A., Whalen, W.E., Giudice, N.A.: Indoor inertial waypoint navigation for the blind. In: 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 5187–5190. IEEE (2013)

13. Tsai, C.H., Elyasi, F., Ren, P., Manduchi, R.: All the way there and back: inertial-based, phone-in-pocket indoor wayfinding and backtracking apps for blind travelers. arXiv preprint arXiv:2401.08021 (2024)
14. Wang, Q., Ye, L., Luo, H., Men, A., Zhao, F., Huang, Y.: Pedestrian stride-length estimation based on LSTM and denoising autoencoders. Sensors **19**(4), 840 (2019)
15. Weinberg, H.: Using the ADXL 202 in pedometer and personal navigation applications (2002)
16. Yan, D., Shi, C., Li, T.: An improved PDR system with accurate heading and step length estimation using handheld smartphone. J. Navigation 1–19 (2021)

Use of Technology and Applications in Orientation and Mobility of Visually Impaired Persons in Bulgaria: A Contemporary Overview

Mira Tzvetkova-Arsova(✉) and Margarita Tomova

Department of Special Education, Faculty of Educational Studies and the Arts, Sofia University "St. Kliment Ohridski", Sofia, Bulgaria
{mcvetkova,mvenelinov}@uni-sofia.bg

Abstract. This paper examines for the first time in Bulgaria the use and popularity of some technological solutions and applications for orientation and mobility purposes among five special organizations for visually impaired in three different cities and 41 direct participants – students, clients or members of the same five organizations. The paper attempts to offer an overview of the participants' knowledge and direct use of technological devices and applications in Bulgaria.

Keywords: orientation and mobility · visually impaired · technologies

1 Introduction

In 1948 Lowenfeld [1] wrote that visual loss can be associated with three main limitations (published in selected papers in 1981):1. In the control of the environment and the self in relation to it; 2. In the ability to get about; 3. In the range and variety of concepts. To solve some of these limitations the special program of Orientation and Mobility (O&M) was developed. Orientation is defined by many researchers [2–4]. Pick (1987) states that orientation is "knowing where objects are in relation to each other and in relation to ourselves" (p. 80) [5]. Mobility is "the capacity or facility of movement, has two components. One is mental orientation and the other is physical locomotion…" (p.1) [6]. In time Orientation and Mobility became a key area of the education and the rehabilitation of the visually impaired individuals of all ages.

2 State of the Art

Along with the classical teaching of the techniques of the sighted guide, of the white cane and of moving alone, the Orientation and Mobility includes recently the implementation of new modern devices and applications [7]. Four O&M systems are described: human guide, canes, dog guides and electronic travel aids (ETA) (p. 241) [7]. The last are devices that "emit energy waves to detect the environment within a certain range or distance, process reflected information about the immediate environment in an intelligible and

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 408–415, 2024.
https://doi.org/10.1007/978-3-031-62846-7_49

useful manner (p. 250) [7]. The ETAs are divided into 2 main groups: primary (laser cane, Ultra cane etc.) and secondary (Sonicguide, Mini guide etc.) [7]. In addition, many applications were developed recently to be used on smartphones, iPads and other portable smart devices. A large theoretical overview of different technological solutions available for O&M purposes was made [8], dividing them into such appropriate for children and adults. In addition, research was done in regard with publications in the field of technology and O&M. A recent study [9] found a total of 2241 studies in this area in selected database. Our online search of internationally known devices and applications in the O&M, showed many popular solutions, e.g.:

- WeWalk cane is a smart cane that can detect obstacles and provide navigation by using ultrasound. A navigation feature is built, which by using clock directions, provides destination tracking through automatic voice about nearby places, transport stops and timetables. It works with application available for IOS and Android (https://wewalk.io/en/).
- Lazarillo is accessible GPS application for visually impaired people to explore the environment and to create routes with audio guidance. It provides real time messages in audio format and alerts while navigating (https://lazarillo.app). It works with application available for IOS and Android. Thanks to an official permission by the producing company, it was translated into Bulgarian language, optimized and introduced for free use in Bulgaria (https://play.google.com/store/apps/details?id=com.lazarillo).
- Be My Eyes connects visually impaired people with sighted volunteers who provide virtual assistance through a live video call. It is available in 180 languages. It works with application available for IOS and Android (https://www.bemyeyes.com/).
- Moovit is a popular mobile application designed to assist users with public transportation navigation. The application provides real-time public transport information, trip planning, and service alerts for buses, trains, subways etc. The Moovit app was one of Apple's Best Apps for 2017 (https://moovit.com/about-us/).
- BlindSquare is the world's most widely used accessible GPS-application developed for the blind, deafblind, and partially sighted. It is available in Bulgarian language. After determining the location, BlindSquare gathers information about the surroundings on Foursquare and OpenStreetMap (https://www.blindsquare.com/about/).
- DotWalker is a travel assistant application primarily designed for visually impaired users but it also enables eyes free control for those who need it. It provides navigation by handling a set of discrete points. Each point can be approached using several modes of control (selected point, nearest point, point on route) (https://play.google.com/store/apps/details?id=cz.lido.dotwalker&hl=en_US&pli=1). This application was additionally optimized to meet the needs of visually impaired users in Bulgaria. In 2017 the daily rehabilitation center 'Svetlina' in Sofia added files of the public stops in the capital city Sofia and of the railway stations in the whole country.
- In addition we included in the list a nationally developed application named Sofbus 24, for navigation within the capital Sofia. It is fully integrated and synchronized with the website for urban mobility of Sofia, but can be used on Androids only (https://play.google.com/store/apps/details?id=bg.znestorov.sofbus24.main&pli=1).

3 Methodology

The research was carried out in two phases:

1. First Questionnaire containing 4 questions was distributed to the CEOs of the five biggest special organizations for visually impaired in Bulgaria - 5 persons in total - in order to establish what technological devices and applications are known and used for Orientation and Mobility purposes in order these devices and applications to be later included in the second Questionnaire. The five organizations were: the two special schools for visually impaired in Sofia and in Varna, the Union of the Blind in Bulgaria; two rehabilitation centers for adults blind – one of daily type in the capital Sofia and the National rehabilitation center for Blind situated in the city of Plovdiv.
2. Second Questionnaire containing 11 questions was distributed to visually impaired persons from the same 5 organizations, who may be users of one or more of the technological devices and applications reported in the first questionnaire –students and adults. We were also somehow provoked by a study [10] which reported that technology use was among the least needed skills according to the participants in their research. The two questionnaires were developed with Google form and distributed to the CEOs of 5 organizations (Questionnaire 1) and to 41 visually impaired participants (Questionnaire 2) in an anonymous way, without requiring personal data, except of belonging to an organization. The research took place December 2023 – March 2024.

4 Analysis of the Results

4.1 Analysis of the Answers from Questionnaire 1

Questionnaire 1 was directed towards the CEOs of 5 special organizations where individuals with visual impairments were trained or received support. We collected information from 5 people, representing each of the institutions listed above.

Figure 1 shows that the respondents indicated that all applications and technologies were available except for BlindSquare. Five respondents (71.4%) mentioned the availability of DotWalker and Lazarillo, four (57.1%) noted the availability of WeWalk cane, Sofbus 24, and Moovit. Be My Eyes was marked as available by three respondents (42.9%), while two respondents (28.6%) included Google Maps. In addition, two individuals mentioned other apps such as Google Maps.

Question 3 in the first Questionnaire was open-ended and attempted to collect answers about other, not included in the list technologies and applications. The five CEOs did not provide any answers. Figure 2 shows the multiple answers about the availability and usage of the applications and new technologies. According to them leading in use were Lazarillo and WeWalk cane (57.1%). Three individuals (42,9%) indicated DotWalker, Be My Eyes, Sofbus24, and the usage of a white cane or a sighted guide, followed by Google Maps – 2 (28.6%). Only one person mentions Moovit (14.3%). Logically, BlindSquare was not mentioned as it was not available within the organizations.

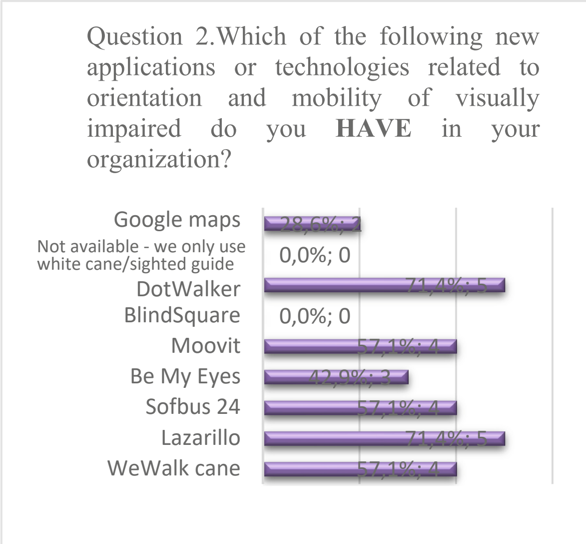


Fig. 1. Distribution based on the availability of new technologies and applications

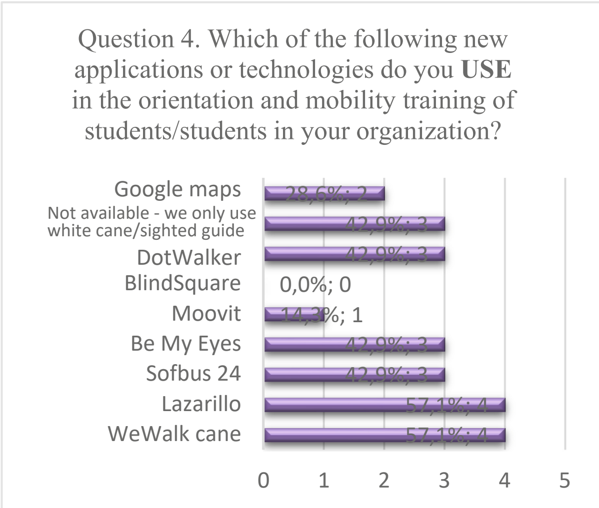


Fig. 2. Distribution based on the usage of new technologies and applications

4.2 Analysis of the Answers from Questionnaire 2

41 visually impaired individuals participated: from the Union of the Blind - 14 (34%), from Center for Social Rehabilitation and Integration “Svetlina” in Sofia - 11 (27%), the Special school for visually impaired “Louis Braille” in Sofia – 7 (17%), 5 from Special school for visually impaired “Prof. Dr. Ivan Shishmanov” in Varna– (12%) and 4 from the National Center for the Rehabilitation of the Blind in Plovdiv (10%).

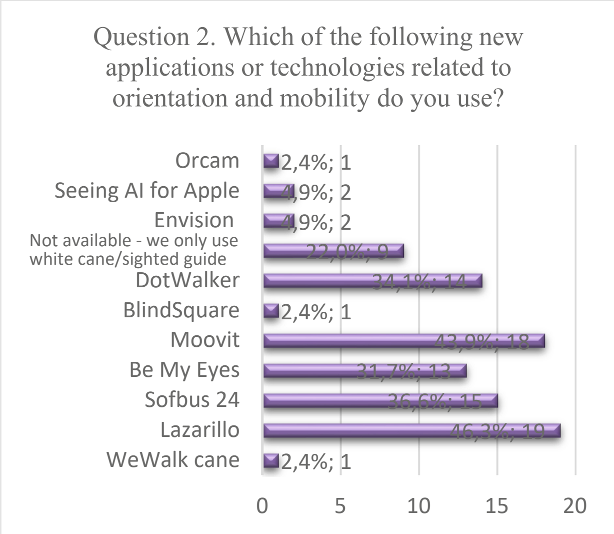


Fig. 3. Use of new technologies and applications from visually impaired persons

The results of the use of new technologies and applications among visually impaired users were particularly interesting and multiple answers were provided, evident in Fig. 3. Leading in usage were Lazarillo – by 19 participants (46.3%) and Moovit – by 18 (43.9%). Sofbus24 was used by 15 people (36.6%), while DotWalker was actively used by 14 individuals (34.1%), and Be My Eyes was noted by 13 (31.7%). The BlindSquare application was used by 1 individual (2,4%), as well as WeWalk cane (2.4%). Nine respondents (22%) indicated that they used mainly the techniques of the white cane and a sighted guide. Under "other applications" the Envision was pointed out by 2 persons (4.9%), Seeing eye for Apple by 2 (4.9%) and Orcam by 1 individual (2,4%).

Regarding the purposes for which technologies and apps were used, multiple answers were provided. Work duties – 26 persons (63.4%) and everyday activities – 24 individuals (58,52%) predominated. Fifteen respondents (36.6%) most commonly used them for shopping, and 14 (34,1%) applied them in the learning process. One user (2,4%) indicated public transport use, one (2,4%) indicated mobility purpose and one respondent (2,4%) noted he/she did not use any technologies or apps.

When asked about who taught them to use the mentioned technologies and applications, multiple answers were provided, shown in Fig. 4. 22 persons (53.7%) responded that they learned how them by themselves. For 10 individuals (24.4%), part of the training came by an O&M instructor, 13 (31.7%) received support by friends. Four users (9,8%) learned about the applications they used by an IT specialist in the center where they were trained. One individual (2.4%) stated he/she has not received any training, and 3participants (7,3%) were taught by a family member.

The answers to Question 5 showed that the majority of the visually impaired – 25 persons (61%) were definitely satisfied with the use of new technologies and applications. Eleven individuals (26.8%) were rather satisfied. Two (4,9%) were unable to determine,

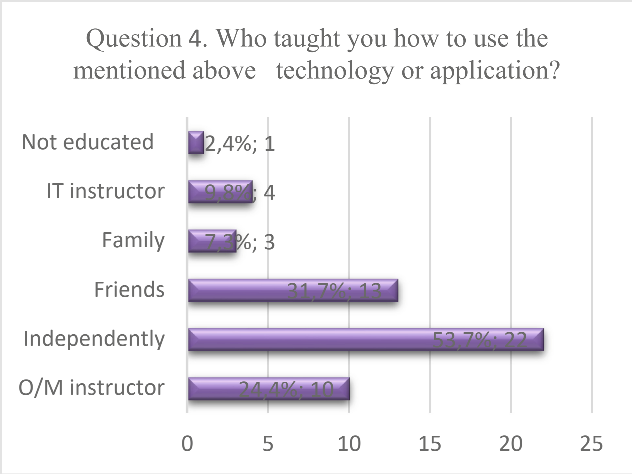


Fig. 4. Training on the use of new technologies and applications

another 2 (4.9%) noted they were rather dissatisfied, 1 person (2,4%) stated he/she did not use technologies.

Questions 6 and 7 were both open-ended and aimed to find out the positive characteristics (depicted in question 6) and the negative features (depicted in question 7). The positives prevail, and include the following: additional information about the environment, support for mobility with the white cane, control over the sequence of bus stops in transportation, information about location of nearby places, independence and confidence in O&M, integration into daily activities. The negative features were: signal delays, difficulties in working with the application and the need for support, inaccuracies in navigation, the need for a good internet connection, frequent updates of information are necessary.

To Question 8 about the practical application of the used technologies and apps in Bulgarian conditions, 31 participants (75,6%) replied positively with YES, 4 people (9,9%) thought that only some are applicable, 6 people (14,6%) could not tell. There were no negative answers giving NO as reply.

As shown in Fig. 5, most of the participants reported their orientation and mobility changed since the start of using technologies and apps. 36 people (87,8%) were fully convinced about it, only 2 (4,9%) reported no change and 3 (7,3%) did not use them.

Regarding the accessibility of the analyzed applications, 28 persons (68.3%) reported they were fully accessible, while 11 visually impaired (26.8%) indicated partial accessibility, and two participants (4,9%) reported no accessibility.

The last open-ended question, attempted to reveal reasons for lack of accessibility, if a negative answer was provided to the previous question. However, we found out a contradiction with the answers to question 2, in which 9 people indicated not using new technologies and apps for O&M purposes, and with the answers to question 3, where one respondent noted not using technologies or apps. Here 4 participants reported lack of

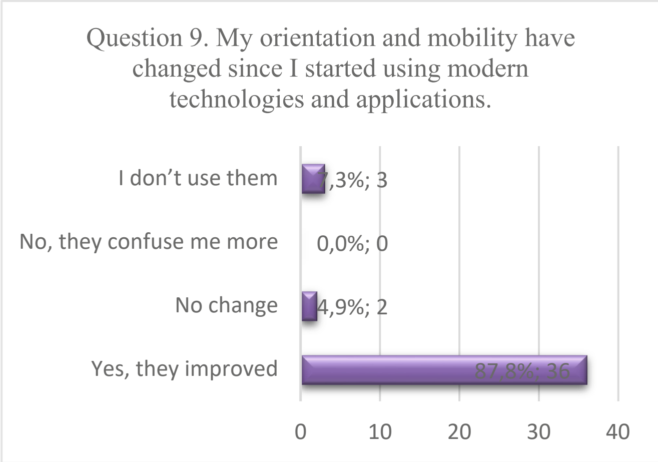


Fig. 5. Change in the quality of orientation and mobility with the use of applications

accessibility simply because they did not use technologies or apps. All other participants (37) reported technologies or apps were fully accessible or somehow accessible.

5 Conclusion

The results showed that many technologies and applications for O&M purposes are known, available and used among the visually impaired individuals in Bulgaria – foreign and national. A good sign is the development and adaptation of national technologies and apps, such as Sofbus24 and Lazarillo. Most of the visually impaired showed real practical experience by reporting about advantages and disadvantages as well as the use of technologies and applications in variety of cases – job purposes, education, shopping etc. Some of the participants knew, used and named additional apps, which were not originally included. A small number of visually impaired people relied on the classical white cane and sighted guide techniques only. However, part of the same participants later reported that the technologies and applications were accessible in Bulgaria, which can be interpreted as having some knowledge about them. Most of the respondents reported a positive change in their O&M skills after the start of using technologies and applications. There was also criticism, such as the Internet dependency, the need for updates with change in the environment, delays in signaling for obstacles etc. A recommendation can be made for more time in training in the use of technologies and apps for O&M purposes provided by schools, centers, instructors and specialists.

6 Limitations of the Research

The research included 5 special organization for visually impaired in Bulgaria from 3 different cities and 41 participants. Although these were the five biggest and mostly known organizations, there are more special organizations. The limited number of participants allowed to shape a picture about the surveyed topic, however more participants would allow a broader understanding of the use of new technologies and applications in O&M. Another limitation is the lack of participants from inclusive settings.

Acknowledgments. This study is financed by the European Union-NextGenerationEU, through the National Recovery and Resilience Plan of the Republic of Bulgaria, project SUMMIT BG-RRP-2.004–0008-C01.

References

1. Lowenfeld, B.: Berthold Lowenfeld on Blindness and Blind People: Selected Papers. AFB, New York (1981)
2. Blasch, B.B., Wiener, W.R., Welsh, R.I.: Foundations of Orientation and Mobility, 2nd edn. AFB Press, New York (1997)
3. VandenBos, G.R.: APA dictionary of Psychology. APA Press, Washington (2007)
4. Long, R.G., Giudice, N.A.: Establishing and maintaining orientation and mobility. In: Welsh, R.I., Blasch, B.B. (eds.). Foundations of Orientation and Mobility, 3rd ed., vol. I, pp. 45–62. AFB Press, New York (2010)
5. Pick, H.: Perception, locomotion & orientation. In: Welsh, R.I., Blasch, B.B. (eds.). Foundations of Orientation and Mobility, 3rd ed., vol. I, pp. 73–88. AFB Press, New York (1987)
6. Hill, E., Ponder, P.: Orientation and Mobility Techniques. A Guide for the Practitioner. AFB, New York (1976)
7. Smith, D., Penrod, W.M.: Adaptive technology for orientation and mobility. In: Wiener, W.R., Welsh, R.L., Blasch, B.B. (eds.) Foundations of Orientation and Mobility, 3rd ed., vol. I, pp. 241–276. AFB Press, New York (2010)
8. Cuturi, L.F., Aggius-Vella, E., Campus, C., Parmiggiani, A., Gori, M.: From science to technology: orientation and mobility in blind children and adults. Neurosci. Biobehav. Rev. **71**, 240–251 (2016)
9. Santos, A.D.P.D., Suzuki, A.H.G., Medola, F.O., Vaezipour, A.: A systematic review of wearable devices for orientation and mobility of adults with visual impairment and blindness. IEEE Access **9**, 162306–162324 (2021)
10. Bahtiyar, A., Can, B.: Orientation and mobility problems of adults with visual impairment and suggestions for solutions. J. Qual. Res. Educ. **28**, 300–330 (2021)

Advancing Mobility for the Visually Impaired: A Virtual Sound-Based Navigation Simulator Interface

Delgerbayar Erdenesambuu[1], Masaki Matsuo[1], Takahiro Miura[2], Msatsugu Sakajiri[1], and Junji Onishi[1(✉)]

[1] Tsukuba University of Technology, Kasuga 4-12-7, Tsukuba, Ibaraki, Japan
deegii@solab.work, ohnishi@cs.k.tsukuba-tech.ac.jp
[2] National Institute of Advanced Industrial Science and Technology, Kashiwanoha 6-2-3, Kashiwa, Chiba, Japan

Abstract. The study investigates a specialized voice navigation system designed for visually impaired users, enhancing voice guidance clarity and accuracy using virtual navigation simulator from autonomous driving systems. It examines the comprehension and proficiency of visually impaired individuals with voice-guided systems, involving a novel voice guide model that incorporates auditory beacons for more effective guidance. The evaluation involved seven visually impaired participants whose feedback was crucial in assessing the system's practicality and efficiency. Results indicated that all participants successfully reached their destinations using the model, showcasing a significant enhancement in navigation aid for visually impaired users. This research contributes to assistive technologies by offering insights into developing more effective and user-friendly navigation systems, emphasizing the importance of user-centric design to meet the specific needs of visually impaired individuals.

Keywords: Visual Impairments · Smartphone Navigation Apps · Enhanced Mobility · Audio Beacons · Orientation and Mobility

1 Introduction

Navigating through everyday environments poses a significant challenge for individuals with visual impairments owing to the inherent design of these spaces, which predominantly cater to those who rely on visual cues [1]. Common urban elements such as roadways, signage, and traffic signals predominantly use visual symbols, letters, and colors, rendering them inaccessible to the visually impaired population [2]. This reliance on visual navigation aids significantly hinders the autonomous mobility of visually impaired individuals and necessitates alternative solutions to foster their independence.

Among the various available aids, tactile paving, commonly known as tagtaile roads, plays a crucial role in assisting visually impaired individuals in navigating public spaces. These tactile surfaces provide essential navigational cues, such as indicating when to proceed straight, turn, and halt. However, the scope of information conveyed by braille

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 416–427, 2024.
https://doi.org/10.1007/978-3-031-62846-7_50

blocks is confined to immediate navigational decisions, leaving a gap in spatial awareness regarding broader environmental contexts, such as addresses and surrounding landmarks. The limitations of tactile paving are further exacerbated by crowded or adverse weather conditions, where maintaining a continuous path on these blocks is challenging or even impossible. Consequently, the mobility of visually impaired individuals is often restricted to familiar routes, or necessitates the assistance of sighted guides, which significantly affects their independence [2]. To address these challenges, this study focused on the development of an advanced navigation system tailored to the visually impaired.

The evolution of navigation aids has undergone a significant transition from the initial global positioning system-based devices introduced in the 1980 s, followed by the development of portable devices that could fit into a pocket in 2000, to the integration of sophisticated navigation capabilities into smartphones in the 2010 s. Despite these advancements, the current systems fail to provide the level of autonomy and ease of navigation desired by visually impaired users. In 2018, a notable breakthrough was achieved with the introduction of a smartphone application that leveraged the camera of a device to provide detailed auditory information regarding the surroundings of a user, thereby enhancing their understanding of the environment [3]. Consequently, users can obtain more information, such as accurate directions, detection of traffic lights, and detection of objects such as bicycles. Despite the potential of such technologies, a deluge of auditory information can overwhelm users, making it challenging to assimilate and act on data efficiently.

One way to address this problem is to utilize the sound discrimination and recognition abilities of the visually impaired [4], which are considered superior to those of the blind. For instance, visually impaired people with extensive hearing experience can detect an object that has been dropped by identifying the point at which it falls from the sound. Others can determine the relative positions of objects in the environment and themselves using environmental sounds as well as reflected and echoed sounds from walls and other objects [5–7]. As an application of this, they can also recognize information, such as the types of stores that are nearby and the stations they are at, based on the environmental sounds of stores and stations [8, 9]. Because various types of recognitions can be performed via the auditory sense, including environmental recognition using sound, designing an interface that exploits this ability is important.

We considered the use of sound effects to summarize information that is difficult to convey by voice alone. For instance, in a navigation system, the audio cue for "turn right" is not always consistent, such as "turn right" or "more to the right." This sometimes makes it difficult for the user to determine whether to "turn right," "go right," or "lean to the right" [10]. Therefore, even if the navigation system can generate an accurate route by mapping the current location of the user to the map data, guiding the user correctly on the route if they are not provided with accurate instructions is not easy. However, visually impaired people can distinguish sounds without problems, depending on the type of sound presented simultaneously [11]. Matsuo et al. developed the OTASCE Map, which helps visually impaired people understand routes by combining sound and sound effects with vibrations [12]. This system allows users to understand a route by tracing it with their fingers before visiting the site. However, constructed auditory presentations have not been used for real-time navigation.

This study proposes leveraging the heightened auditory discrimination and recognition capabilities often observed in visually impaired individuals to design a more intuitive and effective navigation interface. Specifically, we first proposed an instruction method based on sound effects and thereafter constructed a walking simulator for walking in a virtual space to evaluate this method. Next, we performed evaluations on seven visually impaired subjects by reproducing the proposed sound effects in a simulator and summarized the results. The reason for developing the simulator and conducting the evaluation in this study was to perform the test in a safe environment, where accidents with cars and people would not occur. Another reason is that COVID-19 does not require measures to prevent infections. The findings of this study have the potential to significantly impact the design of assistive technologies, contributing to the broader goal of creating inclusive and accessible urban environments.

Fig. 1. Overview of the voice navigation system architecture. This diagram illustrates the integration of auditory beacons with voice guidance technology, detailing the components involved in enhancing navigational aids for visually impaired users.

2 Walking Simulator Developed

2.1 System Overview

Figure 1 shows an example of a walking simulator screen. In this simulator, the user could walk while listening to sounds from a first-person perspective. The simulator was engineered using Python as the primary programming language with Pygame 2.1.2, which served as the foundational framework for the graphical and audio elements [13]. For auditory navigation cues and feedback, the simulator integrated accessible_output2 version 0.17 [14], ensuring compatibility with various screen-reading software to accommodate diverse user needs. Although the initial development and testing were performed on the Windows 11 platform, the cross-platform nature of Python suggests its potential applicability across different operating systems, including Linux.

The interface for the movement was a joystick (Microsoft Xbox wireless controller). It has been reported that visually impaired people have difficulty walking straight down a long street without tactile paving [15], and we assumed it might be possible to simulate this movement better. In the joystick operation, we assigned the direction of the downstroke to the walking direction in the simulator, and the magnitude of the downstroke angle to the walking speed. Auditory feedback, including the sound of footsteps and collisions with virtual obstacles, enhances the immersive experience. Footstep sounds were triggered when each virtual meter walked, whereas tactile vibration feedback was provided upon collision with the virtual walls. In addition, the simulator incorporates ambient sounds that mimic real-world environments, such as train stations and traffic signals, to simulate a realistic auditory landscape for users and their mental maps. These auditory cues were dynamically rendered to reflect spatial audio characteristics, such as volume and stereo panning, using the audio functions of Pygame's.

The map data in the simulator, an example of which is shown in Fig. 1, were created for evaluation purposes. We randomly generated multiple road lengths (30–130 m), converted them to comma-separated values (CSVs), and manually assigned sound effects and voices to them. The stores on either side of the road were described in the JSON format based on information collected from the Internet, including 24 convenience stores, eight electronics retailers, and 42 supermarkets. When the user operated the simulator, the following operations and events were logged every 60 fps:

- Player coordinates xy, orientation (degrees)
- Player's direction of movement per frame
- Angle of player's change of direction per frame
- Navigation progress (points on the route currently being guided)
- When the navigation phrase (voice for directions) is played.
- When the player hits the wall

2.2 Audio and Sound Effect Navigation Methods

In the walking simulator, we implemented a guidance system for the destination that simulated a navigation application on a smartphone or other device. The system was designed to simulate a situation in which a visually impaired person uses a navigation application on a smartphone, and various types of information are presented using voice and sound effects as the user walks through the simulator. Specifically, information on the distance to the next turn, direction of the turn, information necessary for crossing the road, such as traffic signals and pedestrian crossings, and information on landmarks along the route, such as convenience stores, electronics retail stores, and supermarkets, was presented using audio at the time the user approached each point.

Two types of audio beacons are used in the simulator. These audio beacons compare the heading of the route to the destination with the current heading of the user and provide audio and sound effects depending on the magnitude of the angle.

2.3 Audio Beacon Mechanisms

1. Speech Beacon: This beacon employs vocal cues to indicate the correct speech direction. If the heading of the user deviated by 15°or less from the intended direction, the vocal prompt confirmed the correct trajectory toward the next waypoint. The deviation exceeded 15°but remained below 90°, and the beacon advised a directional adjustment (left or right) through specific sound cues. For deviations greater than 90°, indicating a reverse direction, the beacon signals an "opposite" direction alert.

2. Sound Beacon: In contrast to a speech beacon, a sound beacon utilizes non-verbal sound effects to convey directionality: a positive tone for correct alignment, a buzzing sound in the opposite direction, and distinct pitch variations that are high for left turns and low for right turns. The volume of these sound effects decreases as the angular difference between the desired and current headings increases, aiding in directional discernment. These navigational aids simulate the experience of using a real-world navigation application, guiding users toward their destination with auditory cues and feedback.

3 Experimental Procedure

The purpose of the experiment was to evaluate the effectiveness of the voice navigation system, specifically focusing on the performance of speech and sound beacons in aiding visually impaired users to navigate urban environments accurately. This experiment aimed to determine whether these auditory cues could enhance spatial awareness and mobility, ultimately facilitating more independent and confident navigation for users with visual impairments. The study sought to gather feedback on the usability and user experience of the navigation aids to inform future improvements and ensure the system meets the needs of its intended audience effectively.

3.1 Experimental Participant

A total of seven participants were present, five males and two females aged between 19 and 29 years shown in Table 1. Five of them had low vision, and two were blind. All the participants used smartphones, tablets, and other devices daily, and four indicated that they used navigation and camera applications when they went out alone.

3.2 Evaluation Method

To assess the efficacy of our audio beacon guidance, we conducted an experimental study within a simulator, focusing on the accuracy of the navigational assistance provided by the audio beacons.
Experimental Setup:
First, we asked the participants to wear headphones and informed them that they would use a controller to move around in the simulator environment. Thereafter, using a practice route, we explained the use of voice guidance for route guidance and two types of virtual audio beacons, and familiarized the participants with the operation method.

The experiment was conducted once for each of the two enforcement actions, one using a speech beacon to reach the destination and the other using a sound beacon to reach the destination. The evaluation route was at 550 m, including traffic signals, stations, and convenience stores and included the following objects:

- Station sounds: one type (starting point)
- Intersection signal tone: three types
- Convenience stores: two types (names change with each experiment)
- Consumer electronics retailers: two types (names change with each experiment)

Table 1. Participant Demographics and Usage Details.

ID	1	2	3	4	5	6	7
Sex	male	male	female	male	female	male	male
Visual impairment	totally blind	totally blind	low vision	low vision	low vision	low vision	low vision
Mobility training experience	Learned at schools for the blind and universities	Acquired through classes at schools for the blind and universities, and training at rehabilitation centers	Learned at schools for the blind and universities	Learned at schools for the blind and universities	Learned at schools for the blind and universities	Learned at schools for the blind and universities	No mobility training (but capable of independent travel)
Frequency of going out alone	almost every day	Approximately 2–3 times a week	Approximately once a week	almost every day	Approximately once a week	Approximately 2–3 times a week	Approximately 2–3 times a week
Assistive devices/Services used during travel	White cane, station guidance, train boarding assistance	White cane, Guide Help	White cane	none	White cane	White cane	White cane, 15 × magnifier, light-emitting diode light, asks for guidance when lost

- Supermarket: one type (name changes with each experiment)
- Station sounds: one type (at goal)

The order in which the audio beacons were used was randomized for each participant. The simulator also logged the travel time, time when the guidance was presented, and reaction of the experimental participant at that time.

After completing each treatment, the participants were interviewed regarding the route they walked (rough directions: the order and direction of convenience stores, electronics stores, and supermarkets they passed; the number of traffic lights, stations, and bus stops; and the types of sound effects they heard in the simulator environment) to confirm whether they had learned the route correctly. In addition, the participants were asked to rate the intelligibility of the audio beacons by UMUX-Lite [16] using a 7-point Likert scale.

After the experiment, the participants were interviewed and asked to respond freely to their impressions and requests for improvement throughout the experiment.

4 Results and Discussion

4.1 Comparison Results Per Audio Beacon

This study focused on trends related to the evaluation results and comments from the participants in the experiment.

First, the evaluation results demonstrated that all experimental participants could reach their destination without being lost, regardless of the audio beacon used. However, the speed of response to the sound and speech beacons was different for each participant. For instance, one participant initially failed to recognize the left and right directions with the sound beacon and thereafter turned in the opposite direction and recognized the correct direction based on the loudness of the beacon. However, participants who were accustomed to using functions similar to the audio beacon used in this study in other systems, participants who could play musical instruments and had a relative sense of sound, and blind experimental participants quickly became accustomed to the sound beacon after only the practice time and did not turn in the opposite direction. Table 2 lists the distances of randomly generated road segments up to 1 km, along with the simulated navigation system guidance provided for each segment, demonstrating the system's response under varied navigational scenarios.

Table 3 lists the percentage of the experimental participants who learned the route correctly when using sound and speech beacons. In general, the percentage of correct routes learned was the same in all situations (10 routes) or the sound beacons were remembered correctly more often (sound beacons: eight routes, speech beacons: three routes).

Table 2. Navigation System Evaluation Results.

ID	Guidance
1	Go east. (chime sound plays)
2	convenience store on your right
3	Turn right
4	approximately 41 m, turn left
5	Turn left
6	approximately 106 m, turn right
7	electronics store on your left
8	Cross traffic lights
9	Turn right
10	approximately 126 m, turn left
11	Cross traffic light
12	electronics store on your right
13	the station on your left
14	Turn left
15	approximately 19 m, turn right
16	convenience store on your left
17	Turn right
18	approximately 168 m ahead, destination is on the left
19	Cross traffic lights
20	supermarket on your right
21	the destination is on your left

4.2 Comments from Experiment Participants

Firstly, there were reports from participants feeling uncomfortable due to the high frequency of the left and right sounds from the audio beacon.

For one participant, after changing to more comfortable sounds during the experiment, the response to the audio beacon became responsible, and the wrong turn errors decreased.

Furthermore, the initial audio beacon used had the frequencies set to C6 (1046.50 Hz) for the left and E6 (1318.51 Hz) for the right, based on musical notes.

This was done because the frequencies of piano notes are intuitively understandable, with the lower sounds assigned to the left side of the piano keyboard and the higher sounds to the right, making it easier for participants to orient themselves in that direction.

4.3 Comments from Experiment Participants

First, reports of participants feeling uncomfortable owing to the high frequency of the left and right sounds from the audio beacon were provided. For one participant, after changing to more comfortable sounds during the experiment, the response to the audio beacon became responsible and the wrong turn errors decreased. Furthermore, the initial audio beacon had frequencies set to C6 (1046.50 Hz) for the left and E6 (1318.51 Hz) for the right, based on musical notes. This was done because the frequencies of piano notes were intuitively understandable, with lower sounds assigned to the left side of the piano keyboard and higher sounds to the right, making it easier for the participants to orient themselves in that direction. However, some participants tended to orient themselves in a direction opposite to the assigned frequencies. In addition to assigning pitch, other methods, such as adjusting the panning (left-right balance), sense of spread, and stereo presentation, need to be considered in the future to become intuitive for each individual. Additionally, the comments indicated that familiar routes, for instance, made it easier to adapt to the system, suggesting that the preferred presentation method may vary depending on the experience of an individual with the route.

- Beacons that use sound instead of speech are more understandable. This was because of the use of soundscapes.
- Want the sound to be heard in surround
- Using stereo for left and right feels more intuitive
- Why the sound used for sound beacons was lower: It was my first use, and when I heard the buzzer sound, I could not associate the right or left direction with the sound. Once accustomed, listening through sound might be better than being told “right” or “left.” However, this requires getting accustomed to them.
- Might get used to audio beacons by using them in familiar routes where the path is known.
- Speech feels more intuitive.
- Being told the direction feels easier to follow.
- Sound beacons required understanding what the sound meant, which used mental resources.
- With pitch, slight changes can make it difficult to discern whether it is higher or lower.
- A clear difference is easier to understand.

- Changing the sound to something more pleasant to the human ear might be beneficial.
- Similar to speech beacons, clear high- and low-frequency sounds make them easier to understand.
- The ability to customize audio beacons may also be beneficial. Does not focus on navigation.
- What type of sound would be more understandable: pointing the pan toward the correct direction or adjusting the pan such that the direction of progress feels like it is coming from the front. No need to use pitch. The same sound in any direction. Volume might not need to be changed either.

Table 3. Navigation System Evaluation Results.

Guidance ID	sound beacon	speech beacon	Cases exhibiting a high average accuracy rate
1	71%	57%	sound beacon
2	71%	71%	equal
3	57%	43%	sound beacon
4	0%	0%	equal
5	29%	29%	equal
	0%	0%	equal
7	57%	71%	speech beacon
8	100%	100%	equal
9	29%	14%	sound beacon
10	0%	0%	equal
11	100%	100%	equal
12	71%	57%	sound beacon
13	71%	57%	sound beacon
14	14%	14%	equal
15	0%	0%	equal
16	57%	43%	sound beacon
17	29%	0%	sound beacon
18	0%	14%	speech beacon
19	86%	86%	equal
20	43%	29%	sound beacon
21	0%	14%	speech beacon
Ave	42%	38%	sound beacon

- Sometimes turning the opposite direction in sound beacons.
- Because piano keys get higher as you turn right, higher pitch when turning right might be better.

- Knowing about shops and signals in the vicinity before going out can make walking easier; therefore, it may be useful for simulations.
- Intend to create and practice routes from stations.
- Was not sure about the walking speed.
- If the maximum speed was slightly lower, it might not have been necessary to adjust the stick.
- Not good at remembering maps and routes; therefore, we attempted to walk at a speed close to the actual speed.
- Managed to distinguish between left and right sounds.
- Prefer the version with voice guidance.
- Thought additional announcements could be made such as “more to the left,” in addition to only volume control.
- Normally use Google Maps while checking the compass.
- Screens are hard to observe on sunny days.
- **Pan in speech beacons might also be an option.**
- **Having multiple functions that the user can customize would be good.**
- Pan might be hard to hear when walking outside.
- Walking with earphones can be scary.
- Changing the sound of the footsteps between the tactile paving and other areas may be an option.
- Instead of just “turn right in N meters,” would like to be informed about the intersection to turn at, such as “turn at the Nth intersection on the right/left.

5 Consideration

In a real environment, visually impaired people move around in their living environment by selecting a number of sounds. In this case, the reverberation sound used for echolocation and the sound of the white cane were not used in the evaluation experiment. In the future, we will construct an environment that simulates these sounds, which may enable us to evaluate navigation methods that are more suited to realistic situations.

In addition, using the directional guidance information that exists on the Web as the audio beacon proposed in this study, it may be easier for the visually impaired to navigate on their own based on navigation. However, when the voice guidance of the audio beacon is used in combination with tactile information such as sound and a white cane, which are used in a real environment, investigating the ease of understanding and the feeling of operation in detail is necessary. Therefore, evaluating this situation in a more realistic setting is necessary.

6 Summary and Future Issues

In this study, we constructed and evaluated a navigation method that can convey information accurately and efficiently through voice. Specifically, we proposed an audio beacon, which is a guiding method based on sound effects and voices and constructed a simulator for walking in a virtual space to evaluate the method. Next, we evaluated the proposed sound effects using a simulator with seven visually impaired people. The results suggest that audio beacons (sound beacons) based on sound effects may be more accurate than audio beacons alone for understanding routes. However, within the scope of this study, the usability of the two systems was similar. In the future, we intend to evaluate the proposed audio beacon under more realistic conditions such as many cars, people, and noisy environments.

Acknowledgments. This study was supported by KAKENHI (grant number 21K18483, 21H00885, 23K16919, 23K17582 and 23H00996).

References

1. Kanadalan, R.N., Namuduri, K.: A comprehensive survey of navigation systems for the visual impaired. Signal Processing. arXiv preprint arXiv:1906.05917 (2019)
2. Declerck, G., Lenay, C.: Living in space. A phenomenological account. In: Pissaloux, E., Velazquez, R. (eds.) Mobility of Visually Impaired People, p. 2018. Springer, Cham. (2018). https://doi.org/10.1007/978-3-319-54446-5
3. Use Live View on Google Maps. https://support.google.com/maps/answer/9332056?hl=en&co=GENIE.Platform%3DAndroid. Accessed 04 Nov 2022
4. Miura, T., Yabu, K.: Narrative review of assistive technologies and sensory substitution in people with visual and hearing impairment. Psychologia **65**(1), 70–99 (2023)
5. Miura, T., Muraoka, T., Ifukube, T.: Comparison of obstacle sense ability between the blind and the sighted: a basic psychophysical study for designs of acoustic assistive devices. Acoust. Sci. Tech. **31**(2), 137–147 (2010)
6. Chebat, D.R., Harrar, V., Kupers, R., Maidenbaum, S., Amedi, A., Ptito, M.: Sensory substitution and the neural correlates of navigation in blindness. In: Pissaloux, E., Velazquez, R. (eds.) Mobility of Visually Impaired People, pp. 167–200. Springer International Publishing, Cham (2018). https://doi.org/10.1007/978-3-319-54446-5_6
7. Rojas, J.A.M., Hermosilla, J.A., Montero, R.S., Espi, P.L.L.: Physical analysis of several organic signals for human echolocation: oral vacuum pulses. Acta Acust. Acust. **95**(2), 325–330 (2009)
8. Nakamura-Funaba, H., Ueda, M., Iwamiya, S.: Questionnaire survey of the use of sound to support the mobility of the visually impaired. J. Acoust. Soc. Jpn. **62**(12), 839–847 (2006)
9. Miura, T., Ebihara, Y., Sakajiri, M., Ifukube, T.: Utilization of auditory perceptions of sounds and silent objects for orientation and mobility by visually-impaired people. In: 2011 IEEE International Conference on Systems, Man, and Cybernetics, pp. 1080–1087 (2011)
10. Pissaloux, E., Velázquez, R.: On spatial cognition and mobility strategies. In: Pissaloux, E., Velazquez, R. (eds.) Mobility of Visually Impaired People, p. 2018. Springer, Cham. (2018). https://doi.org/10.1007/978-3-319-54446-5_5
11. Harrar, V., Aubin, S., Chebat, D.R., Kupers, R., Ptito, M.: The multisensory blind brain. In: Pissaloux, E., Velázquez, R. (eds.) Mobility of Visually Impaired People, pp. 111–136. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-54446-5_4
12. Matsuo, M., Miura, T., Ichikari, R., Kato, K., Kurata, T.: OTASCE map: a mobile map tool with customizable audio-tactile cues for the visually impaired. J. Technol. Persons Disabil. **8**, 82–103 (2020)
13. Pygame. https://www.pygame.org/news. Accessed 04 Apr 2024
14. accessibleapps/accessible_output2. https://github.com/accessibleapps/accessible_output2. Accessed 04 Apr 2024

15. Kallie, C.S., Schrater, P.R., Legge, G.E.: Variability in stepping direction explains the veering behavior of blind walkers. J. Exp. Psychol. Hum. Percept. Perform. **33**(1), 183–200 (2007)
16. Lewis, J.R., Utesch, B.S., Maher, D.E.: UMUX-LITE: when there's no time for the SUS. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 2099–2102. ACM, New York, NY (2013)

FindMyWay: Toward Developing an Accessible Wayfinding Application for Visually Impaired in an Indoor Space

Uddipan Das(✉) and Brian Hong

The College of New Jersey, Ewing, NJ 08628, USA
{dasu,hongb2}@tcnj.edu

Abstract. Due to the limitations of satellite-based navigation systems, such as GPS, indoor wayfinding has been very challenging for people with disabilities. Unfamiliarity and infrastructure complexity in an indoor space could put an additional burden on people in their wayfinding, who have various types of disabilities such as vision, mobility, hearing, etc. An accessible wayfinding application could be of great help to them in navigating an indoor space. To fulfill the wayfinding needs of visually impaired persons in an indoor space, an iOS application named "FindMy-Way" has been developed in this work based on Bluetooth Low Energy (BLE) beacons. FindMyWay provides both exploration and navigation features for the visually impaired for indoor wayfinding. This work incorporated a parallel processing approach into the app while implementing both exploration and navigation functionalities. The exploration feature provides customized information about a point of interest (PoI) in an indoor space based on user preference. On the other hand, the navigation feature provides guidance to reach a destination in a multi-floor environment. The preliminary results indicate the effective performance of the app while navigating an indoor space.

Keywords: Assistive Technology (AT) · Indoor Wayfinding · Visually Impaired · Accessibility · People with Disabilities

1 Introduction

Wayfinding for individuals with vision impairments has long been a difficult task, particularly in indoor environments, such as office buildings, shopping malls, airports, and hospitals, due to the limitations of GPS signals. To address this issue, researchers and developers have been exploring alternative navigation technologies specifically designed for indoor environments. These technologies utilize a variety of approaches, including Wi-Fi, Bluetooth, radio frequency identification (RFID), and beacons, to provide accurate positioning and navigation information within buildings. By leveraging alternative navigation systems and incorporating visual and audio cues, it is possible to create more accessible and inclusive

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 428–435, 2024.
https://doi.org/10.1007/978-3-031-62846-7_51

spaces for individuals with visual impairments. All available assistive technologies have different advantages and disadvantages when used in a wayfinding application. The benefits and limitations mostly depend on the deployment density, deployment and maintenance cost, and processing and storage capability of the assistive devices having these technologies.

In this work, an iOS application, named FindMyWay, has been developed for the wayfinding of visually impaired persons in indoor spaces using Bluetooth Low Energy (BLE) beacons. The user interface of the app has been designed to address the accessibility needs of visually impaired users. An indoor space has been represented in the app as a graph-based network, G (V, E) of point-of-interests (PoIs). A set of BLE beacons deployed at a testing site are associated with PoIs. To provide wayfinding assistance to the users, zone-based proximity of the PoI, categorized information about any PoI based on user preference, and a parallel processing/multi-threading implementation approach have been incorporated into the underlying algorithms of FindMyWay.

2 State of the Art

Various assistive technologies have been used in the past in different location-based systems to provide wayfinding in an indoor environment. Although tag reader technology, such as RFID (e.g., [5]), was widely used in many early systems, this technology comes with some inflexibility in modifying information on tags which could be vital in providing any location-based services. There has been much work done on computer vision-based localization in the literature though there are very few works reported aimed at BVI persons, e.g., [2]. Moreover, a computer vision-based wayfinding application requires heavy computational resources and battery power from the device. However, bluetooth technology-based BLE beacons became significantly popular in recent years in developing wayfinding applications as BLE beacons reduced many issues related to previously mentioned technologies.

Some recent efforts have been reported in systems, such as NavCog3 [7], GuideBeacon [1], StaNavi [4], ASSIST [6], PerCept [3], etc., that used BLE beacons for providing wayfinding assistance to people who are blind or visually impaired (BVI). These BLE beacon-based systems used various approaches to the localization of a user in an indoor space. All these systems reported an increased ability to navigate indoor spaces by BVI persons; however, there is still a lot of scope to improve localization accuracy in providing wayfinding assistance to BVI persons and to increase the scalability of the app.

The FindMyWay app in this work used zone-based proximity along with a multi-threading approach to localize a user in providing wayfinding assistance to the visually impaired in an indoor space. Moreover, a multi-threading approach has also been incorporated into the navigation algorithm of the FindMyWay app to increase the scalability of the app.

3 Application

The accessibility needs of visually impaired persons have been considered while designing the user interface (UI) of the FindMyWay app on iOS. Only black, blue, and white colors have been used as a color combination for the UI of the app. The positions of buttons in the app are strategically placed at the corners of the UI of the app. However, a big button has been used only at the bottom of the UI on the welcome page of the app to provide basic instructions/protocols for using the app. On all pages of the UI, except the welcome page, there are two buttons placed at the bottom corners of the UI named "BACK" and "HOME" to provide convenience to users to easily navigate the UI of the app.

A big interactive speaker icon has been placed at the center of the UI across all pages of the app for the convenience of users so that they can easily tap on the icon for voice interaction with the app since the voice/audio interaction feature is the primary way of interacting with this app. In addition to audio instructions, FindMyWay provides textual instructions as well so that users having low vision can still use the textual information on the app. To provide textual information on the app, borderless textboxes have been used throughout different pages of the UI. There might be a possibility of missing any given instructions while using the app in a real-world setting. To facilitate getting any missed instructions, the app provides the option to hear previous instructions while using the app. This option helps a user to hear the instruction one more time if it is missed due to noise or any other reason in the environment.

4 Methodology

4.1 Zone-Based Proximity

There are various ways to localize a person in an indoor space using BLE beacons. Most of them rely on estimating distance based on the received bluetooth signals. In many cases, distance calculation is error-prone due to the variation of bluetooth signals. Instead, we used zone-based proximity to find out the proximity of a beacon through our app. The zone-based concept could help to cover multiple PoIs per beacon based on the range of proximity. FindMyWay app has two zones per beacon, such as the "inner zone" and the "outer zone". The outer zone has been used strategically for specific beacons in the testing site to cover multiple PoIs for a single beacon. This zone design scenario could differ based on different indoor spaces. The inner zone of a beacon is designed to describe an associated PoI (Fig. 1 - Scenario A), whereas the outer zone could be used to describe some other adjacent PoI (Fig. 1 - Scenario B). This zone design scenario essentially helped us in reducing the number of beacons in an indoor space.

4.2 Parallel Processing

Two use cases have been identified in this work to perform a parallel processing/multithreading approach to improve the performance of the FindMyWay app and the navigational experience of the user. The parallel processing/multithreading approach can help an application to distribute the underlying

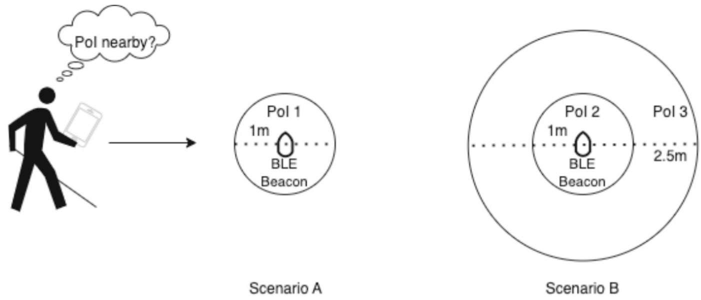


Fig. 1. Zone-based proxmity

tasks among available resources which can eventually increase the scalability of the application and facilitate avoiding potential app crashes.

Avoiding Conflicts Between Multiple Zones of a Particular Beacon. Through empirical testing, it was observed in some cases that there were localization conflicts between multiple zones of the beacon to localize a user based on the proximity ranges of a PoI. To avoid this kind of conflict between multiple zones of a particular beacon, a multi-threaded approach was incorporated into the FindMyWay app. Since two zones have been designed to define the range of proximity for beacons/PoIs, parallel processing has been applied to avoid zone conflicts between these two zones of a particular beacon. As such, the proximity observation for beacons in the FindMyWay app was performed using two different threads; one is for observing the inner zone and the other one is for observing the outer zone of the beacon.

Improving Navigation Experience. A parallel processing approach was applied to find the shortest path from one floor to another floor in a multi-floor building to improve the navigation experience. One of the main challenges of navigation in indoor spaces across multiple floors is to identify connecting nodes/points of interest across multiple floors in a graph-based network representation of a building. To find the shortest path based on the predefined weight on the path in an indoor space, the navigation algorithm may need to work concurrently to improve the navigation time to find the shortest path in a route that spans multiple floors. In this type of scenario, a multi-threading approach has been incorporated into the underlying navigation algorithm to find the shortest path toward a destination across multiple floors. This multi-threading approach certainly increases the scalability of the application to be used in a multi-story building. As such, six threads were designed to be used in the multi-floor navigation of the testing site by the FindMyWay app. These six threads were run concurrently to facilitate a faster runtime of finding the shortest path between source and destination across two floors in the testing site.

4.3 Categorized Description of a PoI

One of the crucial design decisions in developing a wayfinding app is how to deliver detailed descriptions of a PoI in an indoor space to the visually impaired. Application designers need to be very careful about minimizing cognitive load while delivering information through the app. It is expected that the app will provide detailed information about any PoI only based on user preference. As such, the detailed description of PoIs is organized into various categories in a pre-constructed database of the FindMyWay app. The detailed categorized description of a PoI would only be delivered to a user based on their willingness to hear additional information about any PoI. This approach can keep a good balance between reducing the cognitive load on the user and providing detailed descriptions about any PoI in a given indoor space.

4.4 Rerouting

It could be a common issue for visually impaired users to deviate from the guided path provided by the app while navigating in an indoor space. So, it would be very useful if the navigation algorithm for a wayfinding app would incorporate the rerouting feature during navigation to guide the user from the deviated path to the intended path toward a destination. Therefore, FindMyWay integrates the rerouting feature into the navigation algorithm to provide navigational instructions while away from the previously given path toward a destination. If the user is away from the given route, the FindMyWay app essentially localizes the user within the proximity of the nearest PoI within the floor graph and then steers the route from the new localized PoI to the previously selected destination.

5 Results

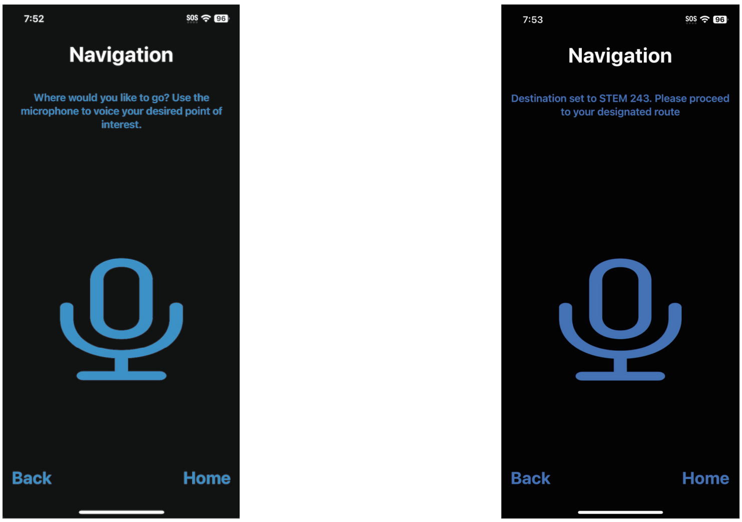
The functionalities of FindMyWay have been preliminary tested by three sighted individuals to check how the functionalities of the app perform. Although this preliminary testing is to conduct the functionality testing of the app, we have a plan to conduct the usability testing of the app by visually impaired persons upon receiving IRB approval. To facilitate functionality testing we have identified various testing scenarios at a testing site. The testing site has two typical floors in an academic setting. Since the app has two different units such as exploration and navigation, a set of testing scenarios was identified and tested for both units in multiple iterations.

5.1 Testing of the Exploration Unit

First, a PoI, such as an elevator or a restroom, was tested that involved a single level of user interaction through voice command to provide a general description of the PoI. Then, a PoI, such as an office room, lab, or other academic room, etc., was tested that involved multiple levels of user interaction through voice command to provide a general description and other categorized information about the PoI. When a user was within the proximity of a particular PoI, the user was localized and user interaction through voice command for all levels worked as intended. The elevator and restrooms were covered by the inner zone of respective beacons. As such, when a user was within 1m of the beacon the user was localized near to the corresponding PoI. Similarly, some of the office rooms, labs, and other academic rooms were covered by the inner zone of respective beacons. As such, a user was localized near to those PoIs when the user was within 1 m of the corresponding beacon. However, some office rooms were covered by the outer zone of the beacons that contained other PoI in their inner zone. For instance, an office room, STEM 119 was covered by the outer zone of a beacon placed near Stair B on the first floor of the testing site. So, when the user was within the proximity of 2.5 m of the beacon, the corresponding PoI was STEM 119. However, when the user was within the proximity of 1 m of that beacon, the corresponding PoI was Stair B. The multiple levels of user interaction were designed to provide different types of information based on user preference and to avoid providing a forceful bombardment of information. All different levels of interaction were tested and the app worked as intended for a given PoI.

5.2 Testing of the Navigation Unit

Since this preliminary testing was the functionality testing of the app by sighted individuals, our focus was to observe whether the navigation unit of the app could provide the correct information on the direction, localization within the



(a) Beginning of the navigation unit (b) Setting a destination

Fig. 2. Navigation using FindMyWay app

proximity of subsequent PoIs along the route, the approximate number of steps (or approximate distance) to reach the next PoI from the current PoI, and any turns required to reach the next PoI along the route, etc. First of all, navigation from a source to a destination PoI on the first floor was tested. We tested two destinations (STEM 101 and STEM 112) separately from a source PoI (TSC park entrance). The navigation path from the TSC park entrance to STEM 101 involved one turn and one intermediate PoI along the route. Moreover, navigation from the TSC park entrance to STEM 112 involved two turns and seven intermediate PoIs along the route. It was observed that the instruction to take a turn on these navigation paths was correctly provided by the app on time. Moreover, the intermediate PoIs information along the path and the distance between the PoIs, in terms of the number of steps, were correctly provided by the app. The beacons associated with the first path didn't involve any PoI covered by their outer zone. So, a user was localized in this path near a PoI when the user was within the proximity of 1m of the beacon. However, the second path involved one PoI covered by the outer zone of a corresponding beacon. In this case, a user was localized near that particular PoI when the user was within the proximity of 2.5 m of the corresponding beacon. In both cases, users were able to reach the destination successfully without any errors such as loss of signal from beacons, missed PoI along the route, or incorrect direction of the route.

Secondly, the navigation paths were tested on the second floor where both the source and destination PoIs were located. The first testing path (from the second-floor elevator to STEM 243) was straightforward as the path had only two intermediate PoIs and no turns were involved. However, the second path (from Stair C to STEM 203 opposite corner) involved two turns and two intermediate PoIs in between. Since a particular area on this navigation path involved several PoIs located within a short distance, it was observed that having multiple zones for every PoI could end up with localization conflicts. As such, to avoid any conflict in this particular area, we kept one PoI within the outer zone of another PoI while designing the app. This design decision helped to avoid localization conflicts for PoIs located in such a small area and eventually the user was able to reach the destination successfully following the navigational guidance. This is one of the areas on the testing site where localization accuracy was improved by introducing zone-based proximity and multi-zones for beacons. So, the user was localized near the intermediate PoIs in this path as per the proximity range, such as 1m for the inner zone and 2.5 m for the outer zone, set for the corresponding beacons.

Third, a multi-floor navigation path (from the TSC park entrance to STEM 243) was tested from the first floor to the second floor (Fig. 2). This path was tested with a preference for "Stairs" initially and then with "Elevator". This path involved multiple intermediate PoIs and turns along the route regardless of transition preference. In both cases, the navigation instructions and turn-by-turn directions were correctly provided by the app. In the case of multi-floor navigation, the transition node/floor connecting PoI is very crucial in graph-based navigation. In this path, a user was localized near the transition node(s) on both floors correctly, and the multi-threading approach in the navigation algorithm helped in finding the shortest route quickly without any issues.

6 Conclusions and Future Work

The FindMyWay application presented in this paper implements a zone-based proximity approach to localize a person within the proximity of a PoI in an indoor environment. Although the zone-based proximity approach worked reasonably well during preliminary testing, the inclusion of the pedestrian dead reckoning (PDR) system as an additional method in the application would provide improved localization accuracy. The inclusion of a parallel processing approach in the app through multi-threading significantly improved the performance of the app. This approach certainly increases the scalability of the app where a wayfinding application is required to design for a significantly large and high-rise building. The preference-based categorizing information delivery approach in this app for a particular PoI through a voice command increases convenience for users. In the future, we plan to conduct comprehensive quantitative usability testing of the app using various evaluation metrics by recruiting visually impaired users upon IRB approval. Future works also include finding customized proximity ranges for individual beacons based on the location of PoIs and developing the app on Android for a comparative study.

References

1. Cheraghi, S.A., Namboodiri, V., Walker, L.: GuideBeacon: beacon-based indoor wayfinding for the blind, visually impaired, and disoriented. In: IEEE International Conference on Pervasive Computing and Communications (2017)
2. Fusco, G., Cheraghi, S.A., Neat, L., Coughlan, J.: An indoor navigation app using computer vision and sign recognition. In: 17th International Conference on Computers Helping People with Special Needs (2020)
3. Ganz, A., et al.: Percept navigation for visually impaired in large transportation hubs. J. Technol. Persons Disabilities **6**(1), 336–353 (2018)
4. Kim, J.E., Bessho, M., Kobayashi, S., Koshizuka, N., Sakamura, K.: Navigating visually impaired travelers in a large train station using smartphone and bluetooth low energy. In: Proceedings of the 31st Annual ACM Symposium on Applied Computing, pp. 604–611. SAC 2016 (2016)
5. Mooi, W.S., Eng, T.C., et al.: Efficient RFID tag placement framework for in building navigation system for the blind. In: 8th Asia-Pacific Symposium on Information and Telecommunication Technologies, pp. 1–6. IEEE (2010)
6. Nair, V., Budhai, M., Olmschenk, G., Seiple, W.H., Zhu, Z.: ASSIST: personalized indoor navigation via multimodal sensors and high-level semantic information. In: The European Conference on Computer Vision (ECCV) Workshops (2018)
7. Sato, D., et al.: Navcog3 in the wild: large-scale blind indoor navigation assistant with semantic features. ACM Trans. Accessible Computi. **12**(3), 1–30 (2019)

Empowering Orientation and Mobility Instructors: Digital Tools for Enhancing Navigation Skills in People Who Are Blind

Agebson R. Façanha[1], José Martônio L. de M. Júnior[2], Maria da C. Carneiro Araújo[2], Joel dos Santos[3], Jaime Sánchez[4], and Windson Viana[2](✉)

[1] Federal Institute of Education, Science and Technology of Ceará, Fortaleza, Brazil
agebson@ifce.edu.br

[2] Federal University of Ceará, Fortaleza, Brazil
{martoniojunior,mariana.carneiro,windson}@virtual.ufc.br

[3] Federal Center for Technological Education Celso Suckow da Fonseca (Cefet/RJ), Rio de Janeiro, Brazil
jsantos@eic.cefet-rj.br

[4] Universidad de Chile, Santiago, Chile
jsanchez@dcc.uchile.cl

Abstract. Spatial navigation could present challenges for People Who are Blind (PWB), impeding their ability to effectively determine locations, navigate, and interact with their environment. This study proposes an innovative orientation and mobility (OM) virtual environment. Our research encompasses the development of a map editor and an audio-haptic OM training environment for mobile devices, empowering OM instructors to create customizable maps for training PWB, thus enhancing learner autonomy and independence. We evaluated it through an extensive assessment involving 25 human-computer interaction (HCI) specialists, 24 OM instructors, and 10 PWB. The evaluation aimed to gauge the system's effectiveness in improving the navigation and wayfinding skills of PWB. Findings suggest that our approach significantly aids PWB in creating mental maps, facilitating navigation and spatial awareness. Feedback from HCI experts, OM instructors, and participants who were blind was instrumental in identifying critical areas for further refinement, particularly in enhancing the intuitiveness of the user interface and the accuracy of audio-haptic feedback. By enabling OM instructors to tailor learning materials to the specific needs of their students, this tool has the potential to make a meaningful impact on the autonomy and mobility of PWB. Further research and development are warranted to refine the system and explore its full potential in real-world applications.

Keywords: Orientation and Mobility · Virtual Reality · People Who Are Blind

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 436–443, 2024.
https://doi.org/10.1007/978-3-031-62846-7_52

1 Introduction

Carlos' imagination has no limits, even having lost much of his vision at the age of 15. His bright and curious eyes do not miss any of the countless discoveries of everyday life! At his school, Carlos is the only student with visual disabilities among 200 pupils. There, he began to get the first notions of Orientation and Mobility[1] (OM). With them, Carlos could follow other colleagues and have more autonomy. From this experience and given a challenge proposed by Carlos' OM instructor, we started our research presented in this paper. We aim to provide people who are blind (PWB), like Carlos, with mobile virtual environments to learn and practice OM safely and playfully. With our approach, PWB can playfully explore places that are not well-known. Without being physically there, he will learn spatial information to navigate these places (*e.g.*, a new school floor).

PWB (if there are no other disabilities) have full capacity to develop motor and mental abilities. For that, they need to have opportunities for learning and having significant educational experiences. Frequently, the OM process is confused with learning to use the cane. OM involves many other strategies and resources (*e.g.*, self-protection techniques, canes, and guide dogs) [5,9]. The goal is to allow PWB to know, to relate, and to move independently and naturally in the most diverse structures and spaces [5,6,11]. For PWB, hearing is one of the most important senses because it allows the establishment of spatial relationships and the development of cognitive abilities [11]. Based on this evidence, researchers started to investigate 3D audio for enabling the spatial construction of mental maps, suggesting that spatial images do not rely solely on visual information [6].

Following this technique, we present an approach for creating OM practices tailored for PWB. It encompasses two digital tools: a web-based platform (**E3 Web Editor**) that enables instructors to design OM practice maps, and a mobile application (**ENA Player**) that renders a virtual environment derived from these maps. The proposal aims to provide digital support to OM instructors in making their practices by allowing map customization according to the students' profiles and skills. HCI experts, OM instructors, and PWB evaluated this proposal [10]. In this paper, in particular, we report the evaluation conducted by an OM instructor and her ten blind students. The ten PWB used the mobile application for a month with ten maps created by their OM instructor.

2 Background

We carried out a Systematic Literature Review (SLR) that addresses the OM theme of PWB [6]. We analyzed 987 papers and identified 32 OM virtual environments that provide practice with indoor maps. In those papers, OM researchers have shown that using auditory cues, sonar techniques, and 3D audio combined

[1] Area of special education aimed at the rehabilitation of people with visual disabilities, whether due to congenital or acquired problems.

with haptic feedback devices (*e.g.*, joysticks, electronic canes) improves user interaction, helps promote technology acceptance, and positively affects their cognitive impact in the OM area. In this section, we describe some of these studies. We focus on solutions that assist the acquisition of spatial knowledge before the user navigates the real situation.

2.1 OM Virtual Environments for Mobile Devices

The Hungry Cat game [4], for example, is a serious game for Android smartphones, which uses tactile and audio feedback. Users get spatial information about the virtual environment by controlling a cat avatar in search of food. The mobile app allows players to move around 2D house rooms. For that, players use touchscreen controls or gestures with the device itself. Virtual objects and obstacles are identified when colliding with the player's avatar.

In the VirtualWalk application [7], the environment simulates the user' walking at different speeds using 3D audio effects (*e.g.*, sounds of human footsteps). A narrator transmits a sequence of instructions and the location of the points of interest. That work aims to facilitate the execution of navigation tasks (going from one point to another in the virtual environment). The mobile application uses sensors (*e.g.*, gyroscope, accelerometer) on the smartphone to detect input gestures, which the app uses to control the avatar's walking and rotation.

Guerron et al. [8] proposed a similar mobile application. Their proposal combines voice, beeps, and vibration feedback to help users when exploring a virtual scenario. The smartphone vibration informs the user that his avatar is walking around the virtual world. The research goal is to enable users to create a mental model of the virtual space. After, the users are expected to move in the real world successfully. Results showed the voice interface afforded better effectiveness, efficiency, and cognitive mapping when compared to the other interface modalities (with statistically significant results).

2.2 Virtual Reality and OM Virtual Environments

A recent trend in the HCI community is the evaluation of the new generation of Virtual Reality (VR) devices. VR Head-Mounted Displays (HMDs) (*e.g.*, HTC Vive, Oculus Quest) are receiving researchers' attention. Similar behavior occurred in the Assistive Technology community [2,12,13,17]. More specifically, we already identified three studies integrating HMD into OM virtual environments [1,15,16].

For example, Baker et al. [1] developed a first-person game targeting PWB. During the game, the player rides a horse avoiding obstacles, and has to hit a target or engage in combat. The player needs an HTC HMD, two controllers, and a VR horse simulator. Similarly, Siu et al. [15] propose a treasure hunt game that uses an HTC Vive and a haptic controller (Canetroller) that enhances the sensations of a real cane. The "augmented" cane vibrates when the user collides it with virtual objects. It allows PWB to navigate in a mixed-reality environment while maintaining the tactile and sound sensations of the cane physical usage.

X-Road [16], on the other hand, is a visual and audio street simulator. It is a mobile application with a VR system designed for OM classes. They use a smartphone and a 3D-printed support to simulate a street environment. The user hears ambient noises and the sounds of cars approaching. PWB will have to analyze the scenario, understand which direction to take, and choose the safe time to cross the street to achieve the goal of the game.

Our mobile solution is inspired by the above studies and combines audio (audio cues, 3D audio, Text To Speech - TTS) and haptic feedback (vibration) to assist users in OM practices. Similar to [8], the ENA Player allows the user to customize the TTS speed and voice. As [16], we use the smartphone as the central execution platform for our gamified experience. However, ENA Player differs from both by adding a Bluetooth joystick to ease the control and reduce the interaction barriers in the use of touchscreen-based interfaces. Our mobile application is less immersive than those allowing users to move in a natural environment using HMDs (*e.g.*, X-Road). However, our goal is a cheaper, safer solution for the first indoor OM training tasks.

Furthermore, the above tools were not designed aiming at customization. When they have those features, the tools offer either rudimentary customization options or highly complex ones that require particular knowledge from OM instructors for interface manipulation. This complicates engagement for use and hinders their popularization for creating audio-haptic maps, which is expected to be possible in our proposal using the E3 Web Editor.

3 Methodology

The main objective of our research was to enable OM instructors to customize virtual environments for their learners. We propose two tools: a Web Editor for map construction and an Android mobile app for practicing activities. The methodology used during the research was User-Centered Design (UCD), in which usability goals, workflow, and environment were studied and tested with the help of OM instructors, HCI experts, and PWB [10,14].

The Web Editor, implemented in HTML + JavaScript, allows the creation of maps for indoor environments, such as buildings, malls, airports, and houses. Using a 2D drawing approach, the editor is characterized as an application for authoring audio-haptic maps, in which the OM instructor designs a map to be used in their class, organizing spaces, inserting elements (*e.g.*, doors, glass windows, chairs, tables), and defining the objectives that students will achieve (*e.g.*, finding a specific object, going to a place and returning).

When the mobile user's avatar walks on the floor, a footstep sound is played (*e.g.*, the sound of footsteps on a wooden floor). The walls have two sounds: one for collision and one for description, the latter being played after three collisions. The other map elements, such as furniture and appliances, have three types of audio: a collision sound, a clue audio that describes the object, and an audioquake that will be activated if the object is a goal in the gamified activity.

Figure 1 shows an example of creating a map in which the instructor selected objects from the left column to position them and build the activity. Besides creating their map, instructors can customize parameters such as size, type, rotation, and location of the objects. A video[2] is available and shows a time-lapse of this process. Further information about the design process of the editor is presented in [14].

Fig. 1. Proposed Digital Tools.

The mobile application has a complementary function of generating the virtual environment corresponding to the activity designed by the OM teacher for use by the student. Characterized as a first-person serious game, the app was developed using the Unity engine along with Google's VR development SDK and the UAP plugin for accessibility functions. The app requires headphones and a Bluetooth device. Optionally, it integrates with smartphone-based HMDs. In it, the user must move in a virtual environment to complete objectives and then return to the starting point, interpreting relevant information and acting according to the demands and aids. To use this game, the user must rely on their auditory and haptic perception, knowledge, and OM skills. Such activities aim to practice: Lateralization - perception of the direction of the sound source; Echolocation - perception of the proximity of the object/obstacle; Perimeter tracking - exploration following the environment's walls; and Origin/destination movement - finding a sound-emitting object (*i.e.*, audioquake). Figure 1 and Fig. 2 show examples of activities running on the app. A video[3] is available and shows an example of a map rendering in which the user must locate objectives by sound and return to the starting position.

4 Evaluation

During our research, we conducted various tests and evaluations with HCI experts, OM instructors, and PWBs, some of which are described in [10,14]. These initial evaluations focused on the usability and cognitive effort of using

[2] https://youtu.be/NjpizlleKfA.
[3] https://youtu.be/F7N1PnQyU4c.

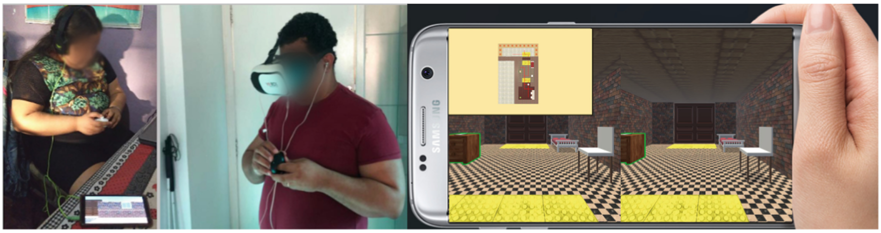


Fig. 2. On the left, 2 PWB using the ENA Player mobile and HMD modes, respectively. On the right, rendering the map in HMD mode.

the Web editor and mobile app. In this paper, we present a deeper case study with an OM instructor and ten of her students. The evaluation was approved by our university Ethics Committee (CAEE: 1652 5319 3 0000 5054).

The OM teacher has a degree in physical education and a master's in education, with over 20 years of experience teaching OM for PWB. The ten PWBs (4 females and six males) were of legal age and used screen readers and mobile devices. They were between 20 and 36 years old; eight had been totally blind for at least 18 years, and 2 had low vision (congenital). The level of OM skills varied, with eight regularly using a cane and two still in the rehabilitation process.

The students' evaluation occurred in two sessions over a one-month interval. The first one focused on familiarizing with the app, commands, and available maps and measuring usability (SUS - System Usability Scale) [3]. The teacher prepared maps according to the specific profile of each user. The mobile app was installed on the user's device, and they were supposed to use it at least once a week for a month. The second session focused on measuring the app's impacts over the month. Before the first session, the teacher conducted an OM pretest. After the second one, the teacher applied the same OM verbal test (posttest).

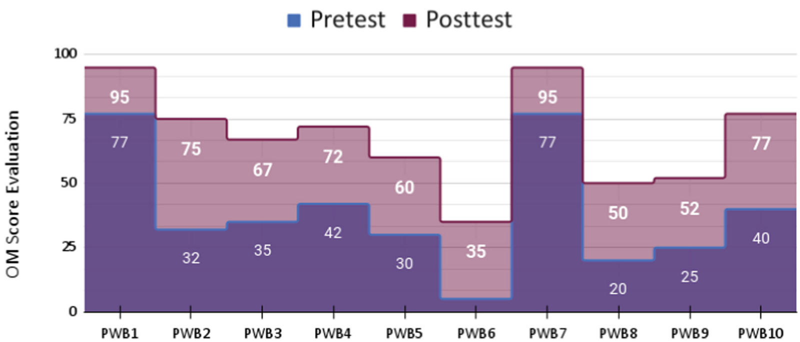


Fig. 3. Comparison of pre and posttest.

The results show that the app received an excellent usability rating [3], measured by a mean score of 81.25 points according to the SUS procedures (one user rated in the range of 60–75, 3 users between 75–80, 5 between 80-90, and 1 above 90). To analyze the impacts of using the app, the focus was on two main axes: *(i)* orientation skills and mental mapping and *(ii)* orientation and mobility techniques. Figure 3 shows the results of pretests and post-tests for all PWBs. All 10 participants increased their scores in the post-test. However, four of them still did not reach 60% of the maximum possible score. We applied the Wilcoxon Test to the results obtained from the tests. Those results indicated a significant difference between the scores ($p - value < 0.01$).

5 Final Considerations

This work presented an audio-haptic map editor designed to allow OM teachers to create and customize maps useful for their students. The studies conducted indicate a latent need to build various scenarios and situations in digital environments, enabling PWBs to practice OM techniques in innovative, safe, and affordable-cost ways.

Our results identified that practice in the virtual environment promotes OM improving skills, precisely, spatio-temporal orientation, orientation according to landmarks, and route planning. The OM teacher, who monitored the use of the ENA player, claimed that there was a noticeable improvement in OM skills, especially, because she was able to design a map for each student.

However, it should be noted that the assessment of cognitive-related aspects is still preliminary. Given the limitations imposed on participant recruitment, even with ten users for the tests conducted, we believe it is still necessary to carry out additional experiments with a larger and more diverse number of PWBs.

Ackowledgments. This research was partially funded by CNPQ, under grant number 409184/2021-7.

References

1. Baker, R.M., Ramos, K., Turner, J.R.: Game design for visually-impaired individuals: creativity and innovation theories and sensory substitution devices influence on virtual and physical navigation skills. Irish J. Technol. Enhanced Learn. **4**(1), 36–47 (2018)
2. Boyd, L.E., et al.: VRsocial: toward immersive therapeutic VR systems for children with autism. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, pp. 1–12 (2018)
3. Brooke, J.: SUS: a retrospective. J. Usability Stud. **8**(2), 29–40 (2013). http://dl.acm.org/citation.cfm?id=2817912.2817913
4. Chai, C., Lau, B.T., Pan, Z.: Hungry cat-a serious game for conveying spatial information to the visually impaired. Multimodal Technol. Interact. **3**(1), 12 (2019)
5. Dove, G., et al.: Digital technologies in orientation and mobility instruction for people who are blind or have low vision. Proc. ACM Hum.-Comput. Interact. **6**(CSCW2), 1–25 (2022)

6. Façanha, A.R., Darin, T., Viana, W., Sánchez, J.: O&m indoor virtual environments for people who are blind: a systematic literature review. ACM Trans. Access. Comput. (TACCESS) **13**(2), 1–42 (2020)
7. Guerreiro, J., Sato, D., Ahmetovic, D., Ohn-Bar, E., Kitani, K.M., Asakawa, C.: Virtual navigation for blind people: transferring route knowledge to the real-world. Int. J. Hum Comput Stud. **135**, 102369 (2020)
8. Guerrón, N.E., Cobo, A., Olmedo, J.J.S., Martín, C.: Sensitive interfaces for blind people in virtual visits inside unknown spaces. Int. J. Hum Comput Stud. **133**, 13–25 (2020)
9. India, G., Jain, M., Karya, P., Diwakar, N., Swaminathan, M.: Vstroll: an audio-based virtual exploration to encourage walking among people with vision impairments. In: Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility, pp. 1–13 (2021)
10. Júnior, J.M.L.D.M., Araújo, M.D.C.C., Façanha, A.R., Viana, W., Sánchez, J.: Uma ferramenta para a customização de ambientes virtuais para práticas de orientação e mobilidade. In: Anais Estendidos do XXVIII Simpósio Brasileiro de Sistemas Multimídia e Web, pp. 103–106. SBC (2022)
11. Lahav, O., Schloerb, D.W., Srinivasan, M.A.: Rehabilitation program integrating virtual environment to improve orientation and mobility skills for people who are blind. Comput. Educ. **80**, 1–14 (2015)
12. Miura, T., Ando, G., Onishi, J., Matsuo, M., Sakajiri, M., Ono, T.: Virtual museum for people with low vision: comparison of the experience on flat and head-mounted displays. In: International Conference on Computers Helping People with Special Needs. pp. 246–249. Springer (2018)
13. Mott, M., et al.: Accessible by design: an opportunity for virtual reality. In: 2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), pp. 451–454. IEEE (2019)
14. Rocha Façanha, A., Viana, W., Sánchez, J.: Editor of O & M virtual environments for the training of people with visual impairment. In: Antona, M., Stephanidis, C. (eds.) HCII 2019. LNCS, vol. 11572, pp. 617–627. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-23560-4_45
15. Siu, A.F., Sinclair, M., Kovacs, R., Ofek, E., Holz, C., Cutrell, E.: Virtual reality without vision: a haptic and auditory white cane to navigate complex virtual worlds. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, pp. 1–13 (2020)
16. Thévin, L., Briant, C., Brock, A.M.: X-road: virtual reality glasses for orientation and mobility training of people with visual impairments. ACM Trans. Access. Comput. (TACCESS) **13**(2), 1–47 (2020)
17. Weir, K., Loizides, F., Nahar, V., Aggoun, A.: Using virtual reality to enable individuals with severe visual disabilities to read books. In: Lamas, D., Loizides, F., Nacke, L., Petrie, H., Winckler, M., Zaphiris, P. (eds.) INTERACT 2019. LNCS, vol. 11749, pp. 680–684. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-29390-1_62

An Embedded AI System for Predicting and Correcting the Sensor-Orientation of an Electronic Travel Aid During Use by a Visually Impaired Person

Gagandeep Singh[2], Mohammad Nadir[3], Rachit Thukral[1], Varun Gambhir[1], Piyush Chanana[1](✉), and Rohan Paul[1,2,3]

[1] School of Information Technology, Delhi, India
{anz228562,piyush.chanana}@sit.iitd.ac.in,
vg.varungambhir@gmail.com, rohan@cse.iitd.ac.in

[2] Yardi School of Artificial Intelligence, Delhi, India
aiy227511@iitd.ac.in

[3] Deparatment of Computer Science Engineering, Indian Institute of Technology Delhi, Delhi, India
ird600069@iitd.ac.in

Abstract. People with visual impairment frequently rely on Electronic Travel Aids for navigation, with sensor-equipped guide canes being a prevalent choice. These ETAs alert users of nearby obstacles using directive sensing technology, effectively covering the navigation corridor of the person motion, thus facilitating obstacle avoidance. However, the efficacy of these devices depends upon the user's ability to maintain them in the correct orientation, a requirement that may not always be met due to various factors, including lack of training and excess cognitive load on the user in mobility, resulting in injuries. This paper proposes the development of an advanced AI-based embedded system, designed to be integrated into a traditional white guide cane. This proposed system predicts the instantaneous orientation of the ETA from raw proprioceptive measurement of ETA's instantaneous velocity and acceleration. If the estimated angle lies beyond a nominal range, audio or vibration feedback is provided proportional to the error.

Keyword: Embedded System · Artificial Intelligence · Assistive Technologies

1 Introduction

Electronic Travel Aids (ETAs) are instrumental in enhancing the mobility of persons with visual impairment (PVI), offering the potential for greater independence and safety. Many of these aids are integrated into canes, providing the ability to detect obstacles from the ground to head height using one or more sensors. These sensors alert users of obstacles in their path by emitting audio-vibratory feedback, significantly reducing the risk of collisions and potential injuries. Despite the apparent benefits of ETAs, their effective

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 444–453, 2024.
https://doi.org/10.1007/978-3-031-62846-7_53

utilization depends on maintaining the proper alignment of sensors for obstacle detection as the user navigates their environment. Trainers play a crucial role in monitoring the PVI's use of the ETA, providing immediate feedback and guidance when sensors are tilted sideways or held in an incorrect orientation. Based on our discussions and feedback from over 100 ETA users in India from 2015–22, it was observed that the majority do not find the devices effective until they have received comprehensive mobility training and engaged in regular practice sessions. However, in areas where there is a scarcity of experienced mobility trainers, access to training for correct use and holding of ETA is unmet, resulting in incomplete and incorrect training by help of other untrained family members or by self, leading to a continued reliance on others for navigation, negating the potential benefits of ETAs for PVIs (Fig. 1).

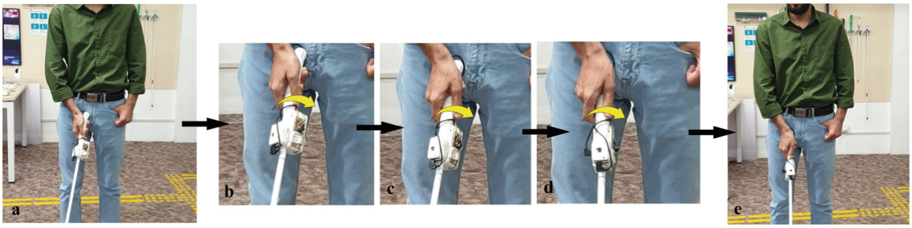


Fig. 1. Orientation correction by user. (a) is the Visually Impaired person holding the ETA incorrectly. (b), (c), (d) shows the ETA being transitioned to the correct orientation aided by our system's auditory and vibratory feedback. (e) shows the Visually impaired person correctly corrected the orientation of the ETA.

This work presents design, implementation and evaluation of a cane mountable AI based intelligent embedded system that learns the ETA holding style of the user, continuously detects the ETA orientation and guides the user to self-correct it through intuitive audio-vibratory feedback. Thereby, minimizing the dependence on the trainer and facilitating self-learning.

2 Related Works

In recent years, there have been several efforts to augment the white cane with sensing technology to improve its detection range. These advancements provide a pre-warning to users about obstacles in their navigation corridor. These efforts are based on ultrasonic ranging [1–3], lasers [4], wearable cameras [5, 6] and sensors combined with a camera [7] to not only avoid obstacles but also identify them, enhancing the overall navigation experience for PVI. Also, there are devices that recognize user gestures [8]. Furthermore, existing research has addressed rule-based solutions to orientation correction by adjusting sensor angles to maintain a constant angle from the ground [9].

However, these solutions often fall short in accounting for incorrect device use, all the available solutions assume the PVI is holding the ETA in correct fashion. Hence, an unparalleled necessity arises for a comprehensive orientation detection and correction system that allows individuals to self-correct the orientation.

3 Technical Approach

We have developed a system which comprises of a battery powered Embedded processing device interfaced to an Inertial Measurement Unit (IMU) sensor, mounted on a cane. This system uses proprioceptive sensing to predict the orientation of the ETA as the PVI is holding it. We train an auto-regressive neural network that predicts the instantaneous orientation of the cane (continuous values along the vertical and rotational axes). The predictions are then discretized and (i) converted to modulated audio buzzer sounds and vibratory output to the user and (ii) visualized and recorded through a graphical display and logs on the computer for the mobility trainer to analyze (Fig. 2).

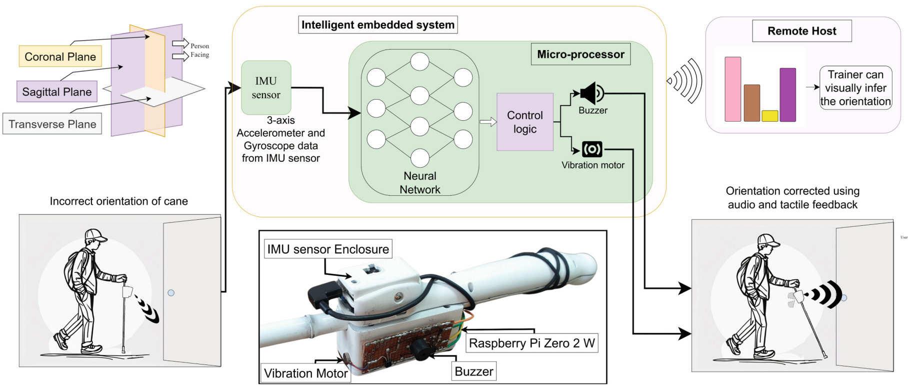


Fig. 2. Overall System Architecture. The system consists of an Embedded processing device with integrated wireless transmission unit and an IMU sensor, A neural network implemented on the device which takes raw sensor data and predicts the orientation of the device, these predictions are modulated for controlling auditory-sensory feedback and wirelessly transmitted to remote server for visual inference.

3.1 Neural Network and Training

Our aim is to design a neural network that takes sensor measurements as input and predicts the instantaneous angle in both sagittal and coronal plane. Since, the choice of the architecture was based on factors like performance, computational requirements, parameter size, power consumption and inference time when deployed on the system. Thus, we tested two different neural architectures: LSTM and Transformer based network. The Noisy raw data from the sensor was given to the neural network to predict the angle of the ETA. The angle predicted by the network is of 4 independent classes, each class $\in [0, 1]$. The predicted classes are Slant, Right, Left and Vertical. We define $\hat{y} = \text{Prediction}(x_{0:t})$, where "$x_{0:t}$" is the raw sensor signal of sequence length 300 and dimensionality of 6. Whereas the "prediction" is the neural network prediction (Fig. 3).

The Neural Networks were optimized using TFlite [11]. TFlite optimizes the models using operations like quantization and pruning. Additionally, it uses FlatBuffers for the optimization of the network. With FlatBuffers, the data can be accessed directly from

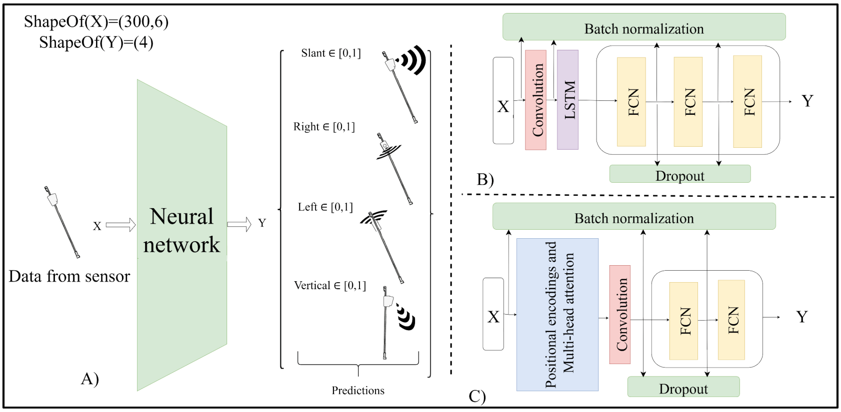


Fig. 3. Neural Network Architecture. A) The neural network takes input(X) from sensor on ETA and predicts(Y) orientation, B) Network architecture of LSTM based neural network contains 1391 learnable parameters, C) Network architecture of transformer based neural network contains 3412 learnable parameters.

the serialized buffer without parsing/unpacking. Allowing TFlite to load and run models faster and with less memory consumption (Table 1).

Table 1. Neural Network Architecture and parameters of LSTM and Transformer models.

LSTM			Transformer		
Layer name	Output shape	Parameters	Layer name	Output shape	Parameters
Batch normalization	300,6,1	4	**Batch normalization**	300,6	24
Convolution	9,4,1	151	–	–	–
Batch normalization	9, 4, 1	4	**Multi-head Attention**	300,6	1424
LSTM	8	320	**Convolution**	9,1,4	1204
FCN	16	144	–	–	–
Batch normalization	16	64	**Batch normalization**	36	144
Dropout	16	0	**Dropout**	36	0
FCN	32	544	**FCN**	16	592
Batch normalization	32	128	**Batch normalization**	16	64
Dropout	32	0	**Dropout**	16	0
FCN	4	132	**FCN**	4	68

3.2 Data Collection and Training Data Generation

To create our training set, 3-axis acceleration and gyroscope data was recorded from an Inertial Measurement Unit (IMU) sensor at a sampling rate of 6660 Hz. Labelling raw IMU data was challenging due to high data rate, noise (due to ETA holding/tapping by the person) and high cost of expensive motion capture systems to record ground truth. Hence, to enable large scale data generation without requiring specialized external hardware, we used a combination of classical algorithms for pre-processing and scoring of the data to generate labels. In the data labeling process, labels were assigned continuous values to accurately indicate the angle of the ETA within the plane. We have labelled the data such that all the classes independently provide the angle of the ETA.

Data Collection. We developed a prototype system with a battery and a micro-controller with which we interfaced a microSD module such that we can collect the raw data from the sensor and store it in the storage device. Using this system, we collected the raw sensor data in 3 passes, (i) Slowly moving the cane from almost vertical position (around 80 degrees) to almost horizontal position (around 10 degrees) in about 2 min. This data was collected to generate labels for the slant class, we will refer to this data as "raw sensor slant data". (ii) Slowly moved the cane from the position of extreme right to extreme left in around the same time, this data was collected to label the right and left orientation, we will refer to this data as "raw sensor right-left data". (iii) For the collection of the vertical class in which the cane was held in vertical position for about 30 s and we will refer to this data as "raw sensor vertical data". After the collection of raw data, we then recorded a reference signal which was a few seconds of cane held in correct orientation.

Training Data Generation. The data generation or labeling process uses a similarity measure technique, we used Dynamic Time Warping (DTW) [10] algorithm to find the distance/similarity of the unlabeled data and a single correct labelled data sample. The DTW is a useful technique for comparing time domain signals. For the similarity calculation of the data with a reference signal, we used the 3-accelerometer axis of both the unlabeled and reference signal, we omitted the gyroscope axis data from the signal to avoid the dependence of velocity of motion on the labelling of data.

In the process of data generation, (i) We filtered the raw sensor data using a series of time domain filers. (ii) We segmented the filtered unlabeled data from the filters into sequence lengths of 300. We used this unlabeled data and calculated the distance of this sequence and the reference signal to compute a distance between the unlabeled data and the reference signal using DTW algorithm, (iii) We normalized the DTW distance/similarity scores to make labels for the slant class. Similarly, we did the same for the right left orientation data using a different reference signal to get the labels. We define the pseudo labels generated by the label generating algorithm as $\tilde{y} = Estimate(x_{0:t})$, where "*Estimate*" is the Estimated labels generated by the supervised label generation process. We train the Neural Network using backpropagation of the loss. Mathematical equation of Loss $= \mathcal{L}(\tilde{y}, \hat{y})$ is as follows (Fig. 4).

$$\mathcal{L}(\tilde{y}, \hat{y}) = \alpha\mathcal{L}_{slant}(\tilde{y}, \hat{y}) + \beta\mathcal{L}_{right}(\tilde{y}, \hat{y}) + \gamma\mathcal{L}_{left}(\tilde{y}, \hat{y}) + \delta\mathcal{L}_{vertical}(\tilde{y}, \hat{y}) \quad (1)$$

$$\mathcal{L}_{class}(\tilde{y}, \hat{y}) = \frac{1}{n}\sum\nolimits_{i=1}^{n}(\tilde{y}_{class} - \hat{y}_{class})^2 \quad (2)$$

$$\text{The co-efficient, } \alpha = \beta = \gamma = \delta = 1. \tag{3}$$

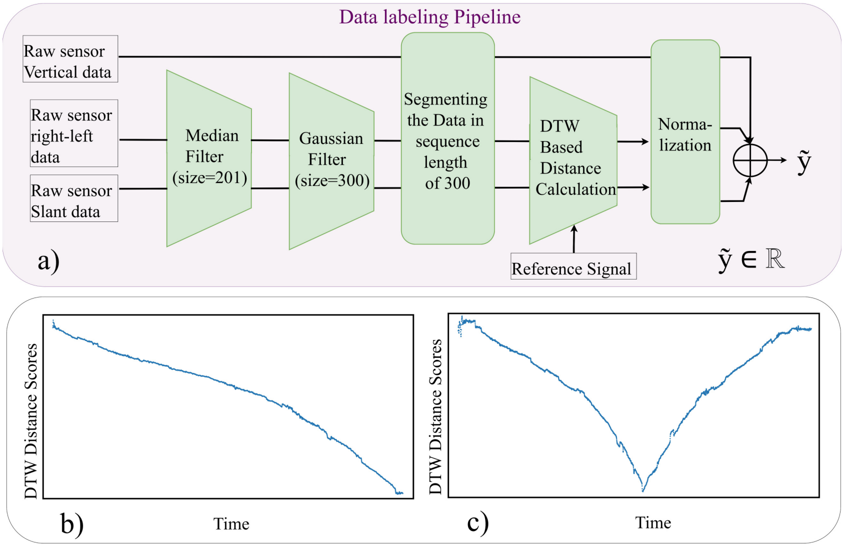


Fig. 4. Label Generating Process. (a) The collected raw sensor data for *slant,* and *right-left* is passed through a median and gaussian filter. After filtering the data is segmented to a sequence length of 300. The segmented data is passed to a Dynamic Time Warping based labelling algorithm for labelling. The resulting data labels are combined to provide the labels $\tilde{y}$. (b) Normalized DTW distance scores generated using the above pipeline for the slant data, as the ETA is moved from vertical position to horizontal position (near ref. Signal) the distance score is decreasing. (c) DTW distance score for right-left data, as the ETA moved from extreme right position to correct position (near ref. Signal) then continued to move towards extreme left position the distance scores first decreased then increased.

3.3 Hardware Implementation

Utilizing the TensorFlow Lite (TFlite) API, we convert the neural models to low-footprint embedded processor compatible TFlite format and deployed them on the device for real time operation. The system calculated the neural prediction and generated the Pulse Width Modulation (PWM) signal using the same core of the CPU. To circumvent the delay due to PWM signal's time-period, a multi-core processing approach was employed. Specifically, one microprocessor core was allocated for neural network computations, while another core was dedicated to PWM generation.

Considering the constraints of limited RAM on the microprocessor, we implemented memory swapping with flash storage, resulting in the successful real-time operation of the system. For user feedback mechanisms, auditory and sensory cues were employed, utilizing a buzzer and vibration motor. These components were controlled by a variable duty cycle signal originating from the microprocessor. The system was intricately designed to alert right and left orientations using vibrator feedback, while slant and vertical orientations were alerted using buzzer beeps as feedback. All the system operations were powered by a 4.44-W hr. Battery. Moreover, the predictions generated by the neural network were wirelessly transmitted to a remote server connected to the system, facilitating visual inference.

3.4 Graphical Interface and Logging

To effectively interpret the system's predictions, it was imperative to develop a Graphical User Interface (GUI) capable of visualizing the results in real-time. To achieve this, a web server implementation was employed, owing to its rapid data transmission capabilities, and reduced computational demands on the server side. The neural network's predictions are stored within the system, from which a host device retrieves the data and displays the graph, the data is also logged for an orientation and mobility (O&M) trainer to analyze (Fig. 5).

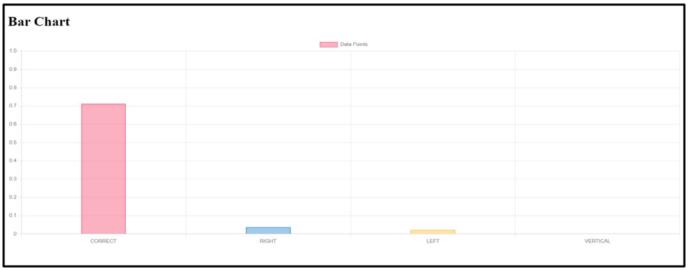


Fig. 5. Visual Inference. The Graphical User Interface on remote host shows the orientation of the ETA. The latency of the interface is minimized to visualize the ETA orientation in real-time.

4 Evaluation and Results

The evaluation covers three primary areas: predictive accuracy of the LSTM and Transformer models, user adaptability in orientation correction, and the system's operational efficiency on low-power edge devices.

4.1 Model Accuracy

Our system underwent evaluation on both neural networks. The labelled dataset was partitioned into training, testing, and validation sets, consisting of 4013, 1255, and 1004 samples, respectively. The performance of each network was assessed on the test dataset, revealing an accuracy of 96.73% for the LSTM and 97.72% for the Transformer network. Although the labelled dataset consists of pseudo labels generated by DTW technique, in our findings they align closely with what would be expected from actual human-annotated data. Both networks achieve high accuracy, they still struggle in scenarios where the right, and left orientations are very close to correct orientation (Fig. 6).

Predicted Class \ Actual Class	Slant	Right	Left	Vertical
Slant	429	13	0	0
Right	17	263	0	0
Left	8	0	245	0
Vertical	3	0	0	277

(a)

Predicted Class \ Actual Class	Slant	Right	Left	Vertical
Slant	442	0	0	0
Right	15	265	0	0
Left	13	0	240	0
Vertical	0	0	0	280

(a)

Fig. 6. Confusion matrix. The above image shows the confusion matrix of the (a) LSTM and (b) Transformer neural network using the pseudo labels generated by label generating algorithm.

4.2 Self-Correction by User

In the control trial, an adult visually impaired male participant was provided with the cane in arbitrary orientations and tasked with adjusting it to the correct orientation using feedback from the buzzer and vibration. Across 10 iterations with a PVI, the following observations were made: The user successfully corrected the cane eight times, with an average correction time of less than 5 s. The user also added that "*this device can be very useful for mobility training as they don't have to correct the first-time user again and again*". Notably, the user demonstrated an improvement in correction speed as the trial progressed. Recognizing the user-specific variability in slant preferences, we identified the need for fine-tuning the feedback mechanism to adapt to individual holding patterns. This adaptation can be achieved through modulation equation adjustments, eliminating the need for repetitive AI model training. The videos of user trial and GUI output can be seen here (Fig. 7).

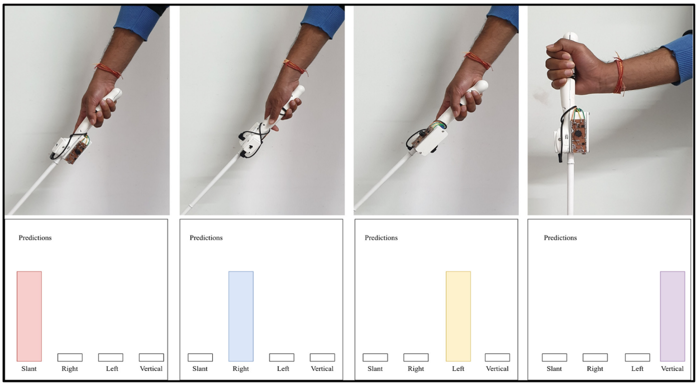


Fig. 7. ETA orientation and prediction. Views of different orientations of holding the cane with their corresponding orientation angle predictions on GUI. The continuous predictions are discretized and converted to buzzer sound and vibration feedback to the visually challenged user as well as logged/shown to the mobility trainer.

4.3 Real-Time Evaluation

The two neural architectures had a small memory footprint with LSTM having 107 KB and the transformer having 42 KB of space in the computer. Successful inference runs were achieved on the Raspberry Pi Zero 2W in 18 ms and 48 ms for the LSTM and transformer networks, respectively. In user trials, the Transformer network outperformed the LSTM-based counterpart in prediction of correct state (Fig. 8).

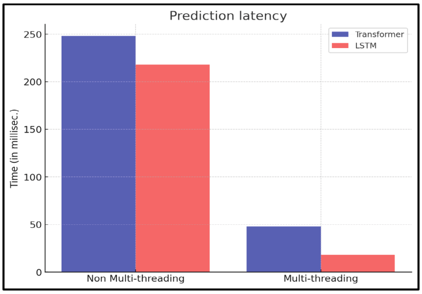


Fig. 8. Neural Network comparison. The prediction latency of the neural networks before and after multi-threading of neural network calculations and GUI calculations.

Additionally, a noteworthy observation revealed that the transformer architecture consumed approximately 1.5 times more power than its LSTM counterpart during inference. This substantial power consumption resulted in elevated temperatures within the microprocessor, highlighting a critical consideration for energy efficiency in system deployment. Furthermore, the graphical user interface (GUI) offers a real-time visualization of cane orientation for a trainer or family member to analyze, updating at a refresh rate of 20 ms, enhancing user interaction.

5 Conclusions

In our research, we developed a system that detects and corrects cane orientation in real-time, offering auditory and vibratory feedback for proper alignment. Evaluation shows the system's effectiveness in slow movement scenarios, yet its accuracy in orientation prediction declines when rapid motion is experienced. Additionally, it cannot accurately predict orientations beyond 90 ° in the coronal plane nor provide the cane's angle in the transverse plane. Future efforts will aim to improve the model's robustness, especially in fast-motion situations, and to extend predictions to include transverse plane angles during movement.

Acknowledgement. Authors acknowledge research funding support from DST, GoI under TiDE program and from ICMR, GoI under National Center for Assistive Health Technologies (NCAHT) project at IIT Delhi. We thank volunteers from Saksham Trust who contributed to user evaluation. We thank Prof. P. V. M. Rao for providing inspiration and mentoring for undertaking this project. We acknowledge Ms. Sunita Negi for administrative support and the IIT Delhi HPC team for the high-performance computing infrastructure used for this work.

References

1. Chanana, P., Paul, R., Balakrishnan, M., Rao, P.: Assistive technology solutions for aiding travel of pedestrians with visual impairment. J. Rehabil. Assist. Technol. Eng. **4** (2017). https://doi.org/10.1177/2055668317725993
2. Hoyle, B., Waters, D.: Mobility AT: the Batcane (UltraCane). In: Hersh, M., Johnson, M. (eds.) Assistive Technology for Visually Impaired and Blind People. Springer, London (2008). https://doi.org/10.1007/978-1-84628-867-8_6
3. Gearhart, C., Herold, A., Self, B., Birdsong, C., Slivovsky, L.: Use of ultrasonic sensors in the development of an electronic travel aid. In: 2009 IEEE Sensors Applications Symposium, New Orleans, LA, USA, pp. 275–280 (2009). https://doi.org/10.1109/SAS.2009.4801815
4. Farcy, R., et al.: Electronic travel aids and electronic orientation aids for blind people: Technical, rehabilitation and everyday life points of view. In: Conference & Workshop on Assistive Technologies for People with Vision & Hearing Impairments Technology for Inclusion, vol. 12 (2006)
5. Kedia, R., et al.: MAVI: mobility assistant for visually impaired with optional use of local and cloud resources. In: 2019 32nd International Conference on VLSI Design and 2019 18th International Conference on Embedded Systems (VLSID), IEEE (2019)
6. Everding, L., Walger, L., Ghaderi, V.S., Conradt, J.: A mobility device for the blind with improved vertical resolution using dynamic vision sensors. In: 2016 IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom), Munich, Germany, pp. 1–5 (2016). https://doi.org/10.1109/HealthCom.2016.7749459
7. Bhatlawande, S., et al.: AI based handheld electronic travel aid for visually impaired people. In: 2022 IEEE 7th International conference for Convergence in Technology (I2CT), Mumbai, India, pp. 1–5 (2022). https://doi.org/10.1109/I2CT54291.2022.9823962
8. Patil, S.G., et al.: GesturePod: Enabling On-device Gesture-based Interaction for White Cane Users. In Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology (UIST 2019). Association for Computing Machinery, New York, NY, USA, pp. 403–415 (2019). https://doi.org/10.1145/3332165.3347881
9. Suresh, M., Pariti, J., Oh, T.: Dynamic sensor orientation unit for the intelligent mobility cane. In: Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS 2019). Association for Computing Machinery, New York, NY, USA, pp. 548–550 (2019). https://doi.org/10.1145/3308561.3354627
10. Sakoe, H., Chiba, S.: Dynamic programming algorithm optimization for spoken word recognition. IEEE Trans. Acoust. Speech Signal Process. **26**(1), 43–49 (1978). https://doi.org/10.1109/TASSP.1978.1163055
11. TensorFlow Authors. *TensorFlow Lite*. GitHub 2024. https://github.com/tensorflow/tensorflow/tree/master/tensorflow/lite

Designing an Inclusive Audiovisual Tool for Museum Exploration

Mahmut Erdemli(✉)

Carleton University, Ottawa, Canada
mahmut.erdemli@carleton.ca

Abstract. The aim of this study is to collaboratively design an audiovisual tool for familiarizing oneself with the interior and exterior of museums. It explores the experiences, preferences, challenges, navigation methods, and use of wayfinding tools in museum settings among individuals with vision and hearing impairments. This research aims to explore auditory and visual elements of maps and street previews through a survey, interviews, and design discussions involving five participants who are blind, four with low vision, and one with deafness, in collaboration with accessibility specialists, to create guidelines for designing inclusive interaction strategies that enhance wayfinding and spatial awareness for both mobile and large screens.

Keywords: Accessibility · Wayfinding · Visual impairment

1 Introduction

When visiting indoor locations, people with vision impairments (VI) encounter challenges related to environmental factors, including noise level, lighting conditions, and layout complexity [9]. Digital maps present limited accessibility for people with blindness (BL) to learn about structural, non-geometric, and accessible details such as spatial layout, signs, and tactile guiding information [1,4]. In the recent years, digital maps were often used for planning, wayfinding, and social participation for individuals with blindness and low vision (BLV) [3]. Navigation studies demonstrated that augmenting route instructions with additional cues decreases cognitive load when following route instructions [5]. A significant challenge in route instructions is the reliance on verbal guidance alone, which proves less effective than using tactile maps supplemented by auditory cues such as spatial relations and soundscape information, crucial for improving navigation efficiency among individuals with VI [8]. Recent research evaluated the accessibility of web browsing systems using spatial audio cues, testing horizontal panning effects, elevated element indicators with higher pitch, and directional beeps for page scrolling [2]. Spatial exploration techniques, enhanced with augmented audio cues, improve understanding of image arrangement compared to relying solely on audio descriptions [7]. In addition to guided tours, pre-recorded audio

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 454–461, 2024.
https://doi.org/10.1007/978-3-031-62846-7_54

descriptions of art pieces are now common in museums and galleries [6]. The use of sound is a logical solution to visual-based approaches when dealing with people with VI. Here, we seek to add to existing research by evaluating tools that have been used by people with BLV for accessible wayfinding to design and propose a novel accessible wayfinding tool for public museums.

2 State of Art

Table 1 outlines the accessibility features of wayfinding tools used by BLV, who often rely on diverse tools for navigating to public locations. We assessed the accessibility of features, marking accessible features with an 'x' and indicating inaccessible ones with a dash. All digital maps offer menu access through screen reader support and but lack audio messages. Digital maps lacked screen reader interaction features such as pinning points and provided visual hints only when using Bing maps. Google Maps offers audio hints for PoIs, while Microsoft Bing had audio hints for icons, arrows, contours, and PoIs. All navigation assistants provide menu access and route instructions, with Waze having location information such as opening hours. Navigation assistants like SoundScape and RightHear provide audio descriptions and street previews, with SoundScape using cardinal directions and RightHear offering clock position and customizable relative directions. Camera apps are used for scanning surroundings with a camera to hear descriptions. Camera apps revealed variations in features such as screen reader support and functionalities related to text, objects, documents, QR codes, scene review, persons, color, light, and additional features like Envision Glasses and in-person support Camera apps such as Seeing AI and Envision support screen readers, with Seeing AI and Envision also providing color recognition.

Table 1. Accessibility Review of Maps, Navigation Assistants, and Camera Apps

Digital Maps				
Criteria	**Google**	**OpenStreet**	**Apple**	**Microsoft Bing**
Hints	PoIs	–	–	Icons, Contours, PoIs
Coordinates	x	x	–	–
Landmarks	Mouse	Nearby	Search this area	Around & ahead of me
Navigation Assistants				
Criteria	**GoodMaps**	**SoundScape**	**RightHear**	**Lazarillo**
Directional GPS	–	Cardinal	Clock Position	Relative (Customizable)
Audio Descriptions	–	x	Audio Recording	x
Street Preview	–	x	x	Tracking
Camera Apps				
Criteria	**Seeing AI**	**Magnifier**	**Envision**	**Be My Eyes**
Screen Reader Support	–	x	x	x
Color	x	–	x	In Person

Previous studies have examined how spatial information is effectively conveyed through various representations of visual, tactile, and audio maps [5]. Studies have shown that auditory and visual exploration tends to be more selective, on the other hand tactile exploration requires methodical approaches. Accessible visual maps are advised to avoid semantic overload [8], while tactile map have clear elements ensuring clear elements and well-spaced labels to improve readability and comprehension [11]. Table 2 groups spatial information into graphic, map, tactile, and audio types, including size, area, dimension, descriptions, showing how different senses help understand space.

Table 2. Spatial information conveyed through different senses

Graphic variables	Map variables	Tactile variables	Audio variables
Size	Area	Dimension	Descriptions
Value	Icons, text	Label, braille	Audio cues
Texture	Geometry	Texture	Environmental sound
Color	Terrain	Elevation	Descriptions
Orientation	Pinned indicators	Landmarks	Directional cues
Shape	Environment	Facets	Alert

3 Methodology

3.1 Apparatus

In the proposed museum audiovisual tool for exterior and interior of the museum two concepts were presented: large screen prototype for exterior and mobile prototype for interior. In the large screen prototype, participants experienced a demonstration of audiovisual interaction patterns for maps and detailed auditory route descriptions for street preview levels, facilitating wayfinding to a public museum. Facilitator demonstrated zooming and camera interaction techniques for exploring points of interest initiating from a bird's-eye view, complemented by additional audio cues. The mobile prototype featured a portrait layout design. The middle screen displayed the interior view of the museum, the top screen provided a minimap of the museum, and the bottom screen was designated for closed captions. Participants listened to audio cues and audio descriptions to familiarize themselves with the museum's interior wayfinding, while the facilitator adjusted settings based on participants' preferences (See Fig. 1).

Graphical Elements. For the indoor audiovisual tool, a virtual view and a minimap was redesigned with Adobe Illustrator from the current exhibits on the website of the museum. Locational data was exported using geolocation databases such as Google Maps API, and Google Earth was used for creating a route for exterior wayfinding experience of the museum. Adobe After Effects and

Fig. 1. Interior of museum, minimap and interior view with closed captions

Premiere was used to generate audio beacons triggered by the proximity of the camera. As the camera approached, accompanying text displayed alongside the audio cues faded in, enhancing the participant's awareness of nearby elements, and facilitating navigation within the virtual environment. Graphical elements of the audio cues were icons and text elements placed of the map. Cues were recognizable based on their saliency and relevance within the context. The graphical elements consist of accessible information in the environment. For this inclusive design we created audiovisual cues for doors (e.g., entrances, exits, elevators), paths (e.g., crosswalks, stairs, ramps), landmarks (e.g., train station, bus stop, river), zones (e.g., terrain, trees), with directional cues (e.g., routes, arrows) (See Fig. 2).

Audio Elements. Audio elements were designed to evaluate if they convey information about the environment, objects, interaction, and obstacles. Audio elements include audio descriptions or audio clips that were compared by participants with VI. Researchers have examined this technique and discovered that integrating verbal explanations with audio elements proves serve as an efficient method for wayfinding [10]. Audio cues played crucial roles in enhancing participants' familiarization with both the interior and exterior of the museum, providing orientation, guiding them to key areas, and immersing them in their surroundings.

Interior museum prototype was collaboratively designed with audiovisual cues to enhance spatial awareness. Ambient sounds were used to increase immersion and provide locative information, while audio descriptions offered

Fig. 2. Exterior of museum, audiovisual cues with closed captions

route descriptions and obstacle warnings. Audio descriptions provided landmark descriptions to aid in map navigation and route exploration. Directional cues supported spatial orientation. Participants listened to audio descriptions of routes and surrounding accessibility features, with audio beacons serving as cues for navigating from person-to-object. Object-to-object cues were conveyed through audio beacons, in accordance with directional proximity. For the exterior museum prototype, various audio cues were presented to participants, including audio descriptions of point of interests, sonification pitches indicating map zooming, directional cues indicating north or south, alerts for barriers, audio descriptions of nearby points of interest, directional audio beacons to adjacent accessibility features, and background ambient sound while panning, zooming, and transitioning between areas.

3.2 Procedure

Our data collection procedure involved conducting a survey, in-person and online design meetings and collaborative discussions. Participants first completed a twenty-minute survey with single-select and multiple-select multiple-choice and open-ended questions about their museum experience including challenges and barriers in wayfinding. For the second phase of the survey a video for the conceptual design was presented and participants answered questions about the auditory cues and visual cues and their preference for auditory interaction methods when used with wayfinding tools. The audiovisual cues were customized based on participant feedback during interviews. By customizing and positioning audiovisual elements within the prototype, our aim was to enhance navigation for participants with VI by enriching auditory information on visual wayfinding tool. Our goal was to provide participants with VI with comprehensive auditory information for effective wayfinding and spatial orientation. Additionally, auditory route descriptions with closed captions are intended to improve wayfinding for participant with hearing impairment, enhancing the immersive and inclusive experience of audiovisual tool.

3.3 Participants

The study included ten participants, averaging forty-five years old. Five reported blindness (BL) and four low vision (LV), and one deafness(DF). Five out of ten participants with BLV having vision difficulties before the age of eight, with one having hearing difficulties since childhood, with an average onset age of around eighteen. Eight participants were from North America, one from Europe, and one from Asia. Educationally, three have High School degree, two Bachelor's, four Master's, and one associate degree. Most visited museums yearly, one monthly, and one weekly. Many used digital maps and navigation assistants daily, while those with vision impairments integrated camera apps into their routines; the participant with DF did not use them at all.

4 Results

4.1 Survey

Participants have different habits when traveling to museums alone, one traveling weekly, one once a month, and the rest once a year. Out of the total ten participants, nine have visited the art museum, seven have visited the history museum, and six have visited the science museum. All participants were interested in indoor museums, five outdoor exhibitions and three online museums. Participants have diverse preferences when visiting museums, eight out of ten participants prefer guided tours, lectures, or seeking assistance, with the exception of one individual with DF and one with BL. Among the participants seeking assistance, only two used indoor mobile apps or audio guides, and one participant with LV received verbal directions from family members. During indoor navigation, only one participant with BL used camera apps to follow directional signages at museums or to read museum maps. Regarding museum collections, eight prefer science and technology displays, three prefer historical exhibits, five prefer art galleries, and seven are interested in special exhibitions. Five enjoy interactive activities. Participants were asked about museum navigation barriers. Six chose physical obstacles such as stairs or ramps, seven considered crowded areas as barriers, eight mentioned lack of accessible signages, five selected difficulty discerning between sections and exhibits. Regarding accessible cues for museum navigation, six out of nine BLV used structural landmarks such as handrails and walls. Two participants used custom signage or landmarks within the museum, while two relied on sidewalk patterns or tactile guiding systems. Seven prioritized barriers inside a museum, with six considering doors and stairs essential, and four deeming art pieces necessary. When visiting, five typically search for information about specific artworks or artifacts, five are interested in exhibit history or culture, and seven seek interactive activities. Locating amenities such as restrooms or cafes was also highlighted by five respondents.

4.2 Interviews

During the interviews participants shared their views on the effectiveness of auditory cues and audio descriptions for museum navigation. Majority of the participants found audio descriptions for describing the map pinpoints helpful, while others had varied opinions. Most participants found the sonification sounds indicating map zooming effective. Audio descriptions on maps were generally heard as effective, though some participants found them more of a hindrance. Participants preferred measurements as steps for short distances, meters/feet for rooms, avoided counting stairs individually, was familiar with cardinal directions for object-to-object orientation, and used clock directions for self-orientation. Directional cues like sounds indicating north or south were considered helpful, and alerts for barriers were deemed very important. Participants preferred clear distinctions between bigger and smaller areas on maps and enjoyed hearing environmental sounds and verbal descriptions while exploring with audio cues. Sonification of landmarks and background sounds were helpful for navigating the museum exterior, and ambient background sound during map panning was important for immersion. Overall, audio beacons guiding to adjacent features, verbal descriptions of nearby points of interest, and background sound changing by location were found most helpful for navigating museum maps. Participants mentioned the importance of audio cues in navigating public spaces like museums. They highlighted how audio cues can enhance accessibility for individuals with vision impairments, providing vital information about their surroundings. Participants emphasized the need for clear and consistent audio cues, such as announcements and verbal directions, to aid in wayfinding. Additionally, they recognized the role of technology in delivering audio information effectively, suggesting the integration of audio guides and smartphone applications. The interviews highlighted how audio cues promote inclusivity and independence for individuals with visual impairments when familiarizing themselves with locations.

5 Conclusion

Our review of wayfinding tools has highlighted some essential issues, such as lack of visual hints, audiovisual audio beacons, detailed audio descriptions, and scaling methods for audiovisual maps. Our design objectives focused on designing an interaction flow specifically designed to get familiar with interior and exterior wayfinding for a museum. The assessment process involved testing audio beacons and sound effects to provide directional cues, aiming to enhance spatial awareness with audio descriptions to enhance overall an accessible digital wayfinding experience for VI. Sequential methods enhanced learning, offering structured guidance, while auditory cues improved spatial awareness, allowing participants to experience accessible elements in an immersive wayfinding experience. Exploring new technologies, such as machine learning for object recognition, could facilitate the integration of audio cues for providing automated geographical information, aligning with accessibility guidelines.

References

1. Brock, A.M., Truillet, P., Oriola, B., Picard, D., Jouffrais, C.: Interactivity improves usability of geographic maps for visually impaired people. Hum.-Comput. Interact. **30**(2), 156–194 (2015). https://doi.org/10.1080/07370024.2014.924412
2. Chheda-Kothary, A., Rios, D.A., Smith, K.S., Reyna, A., Zhang, C., Smith, B.A.: Understanding blind and low vision users' attitudes towards spatial interactions in desktop screen readers. In: Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility, pp. 1–5. ASSETS 2023, Association for Computing Machinery, New York, NY, USA, October 2023. https://doi.org/10.1145/3597638.3614490
3. Engel, C., et al.: Travelling more independently: a requirements analysis for accessible journeys to unknown buildings for people with visual impairments. In: ASSETS 2020 - 22nd International ACM SIGACCESS Conference on Computers and Accessibility (2020). https://doi.org/10.1145/3373625.3417022
4. Erdemli, M., Collins, K.C.: Uses of wayfinding tools by people who are blind and low vision. In: Assistive Technology: Shaping a Sustainable and Inclusive World, pp. 389–396. IOS Press (2023). https://doi.org/10.3233/SHTI230649
5. Koustriava, E., Papadopoulos, K., Koukourikos, P., Barouti, M.: The impact of orientation and mobility aids on wayfinding of individuals with blindness: verbal description vs. audio-tactile map. In: Antona, M., Stephanidis, C. (eds.) UAHCI 2016. LNCS, vol. 9739, pp. 577–585. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-40238-3_55
6. Li, F.M., et al.: understanding visual arts experiences of blind people. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, CHI 2023, pp. 1–21. Association for Computing Machinery, New York, NY, USA, April 2023. https://doi.org/10.1145/3544548.3580941
7. Nair, V., Zhu, H., Smith, B.: ImageAssist: tools for enhancing touchscreen-based image exploration systems for blind and low vision users. In: Conference on Human Factors in Computing Systems - Proceedings (2023). https://doi.org/10.1145/3544548.3581302
8. Papadopoulos, K., Koustriava, E., Koukourikos, P.: Orientation and mobility aids for individuals with blindness: verbal description vs. audio-tactile map. Assistive Technol. **30**(4), 191–200 (2018). https://doi.org/10.1080/10400435.2017.1307879
9. Quinones, P.A., Greene, T., Yang, R., Newman, M.: Supporting visually impaired navigation: a needs-finding study. In: CHI 2011 Extended Abstracts on Human Factors in Computing Systems, CHI EA 2011, pp. 1645–1650. Association for Computing Machinery, New York, NY, USA, May 2011. https://doi.org/10.1145/1979742.1979822
10. Srikrishnan, A., Joshi, A.: Comparison of verbalized navigation by visually impaired users with navigation based on instructions from Google maps. In: Proceedings of the 10th Indian Conference on Human-Computer Interaction, IndiaHCI 2019, pp. 1–6. Association for Computing Machinery (2019). https://doi.org/10.1145/3364183.3364193
11. Štampach, R., Mulíčková, E.: Automated generation of tactile maps. J. Maps **12**(sup1), 532–540 (2016). https://doi.org/10.1080/17445647.2016.1196622

Blindness, Low Vision: New Approaches to Perception and ICT Mediation

Blindness, Low Vision: New Approac Hes to Perception and ICT Mediation Introduction to the Special Thematic Session

Katerine Romeo[1](✉), Marion Chottin[2], and Edwige Pissaloux[1]

[1] LITIS, University of Rouen Normandy, Mont-Saint-Aignan, France
Katerine.Romeo@univ-rouen.fr
[2] IHRIM, CNRS-ENS Lyon, Lyon, France

Abstract. Blind and low vision people tell us that they often face unexpected accessibility issues. Videos, images, some digital documents, graphics, artworks such as paintings or sculptures are accessible mainly through vision. They need other alternative solutions, involving other senses, in order to experience daily life challenges and enjoy a (multimodal) aesthetic experience induced by artistic or sonic interactions. We introduce "Blindness Gain" where blind and partially blind people benefit from access to a multisensory way of being which celebrates inventiveness, imagination and creativity. The perception is presented through a practical-sensitive approach as the interweaving of the five senses that communicate with each other, one in the other. Finally, the articles in this session are introduced as ICT mediation tools with different perceptions to advance in the usability of the proposed systems.

Keywords: Blindness · Low Vision · Perception · Information and Communication Technology

1 Introduction

In our daily life, blind and partially blind (BPB) people face unexpected accessibility issues. Videos, images on digital documents, graphics, artworks such as paintings or sculptures are accessible mainly through vision ("visual art"). But they benefit from alternative solutions, like audio descriptions, involving other senses, in order to experience daily life challenges and enjoy a (multimodal) aesthetic experience induced by artistic or sonic interactions. While BPB people are traditionally considered as the primary users of these alternatives solutions, some of them can also benefit to young children, particularly those acquiring language, elderly people who are losing visual acuity, people with limited attention spans or other cognitive disabilities, people who are temporarily disabled through illness or injury, migrants and tourists for whom a local language is not well known, economically and socially disadvantaged groups who are unfamiliar with or alienated by traditional ways of experiencing and engaging with visual challenges, and people who do not process information in a primarily visual way.

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 465–472, 2024.
https://doi.org/10.1007/978-3-031-62846-7_55

In general, any public will benefit from multimodal initiatives such as audio-tactile accessibility to the "visual life"; this includes both blind and non-blind visitors. This requires a necessary move beyond ocular-centric assumptions about perception towards a truly inclusive enriched experience for all. ("post-ocular centric perception"). Trials of tactile presentations of objects exist in research laboratories and are evaluated by BPB people.

Issues such as new multi-modal stimulation devices, models of perception emerging from stimulations, the simplified presentation of objects whilst still achieving the "conservation of meaning", multimodal data fusion, etc. are frequently investigated around the world.

However, BPB persons are not only interested in multimodal presentations of objects or documents, but also in the physical accessibility for everyday challenges as for example the expiration date detection and recognition of products while shopping and self-guidance through the buildings. Specific wearable technologies and specific adaptable guides for indoor (and outdoor) navigation are necessary.

Therefore, the objective of this session is to provide an introduction of Information and Communication Technologies (ICT) based solutions for accessible documents, objects, music and artworks, virtual work space and possible progress beyond. In Sect. 2 we introduce perception as a multisensory fusion and as a practical-sensitive approach. In Sect. 3 we present the concept of Blindness Gain and the inclusive co-creation of audio descriptions. Finally, in Sect. 4 the articles in this session are presented through an approach of ICT mediation. The Sect. 5 closes this introduction, and proposes some items for a possible « road map» for *multimodal inclusive* approach.

2 Perception

In the most popular meaning, the world « perception» is assimilated to « sensory perception», i.e. the "immediate" perception that our senses deliver to us, as direct information (multimodal approach). The term "sensation" is also used in a broader sense also covering emotions»; this broader definition is essential in the case of "impaired" senses.

Perceptual phenomena have at least the two following characteristics in the case of "not-impaired" sensors: the permanence of operation [12], and the permanent « on/input» status. Our sensors transport to the brain the acquired data (raw or partially (pre-)processed (as in the housefly eye), [13, 14]). In some specific situations (such as medical or voluntary cognitive anesthesia) our brain cannot ignore the perceived data; it fuses/integrates them (consciously or unconsciously) in a percept (a perceived object) (*a multisensory fusion*).

Today's world is built by sighted people for … sighted people, despite of the fact that many information on the observed objects are not visual. If you see an image of an orange, you can recognize it as such, but you cannot directly appreciate its weight, its ruggedness, neither its taste nor smell. Consequently, the world presentations based on visual data only is a *barrier* to the real-world perception, interpretation and ad hoc interaction with it.

Finally, the world perception should be studied with *a multimodal inclusive approach*, so any person, with/without "impairment", can understand it, share its perception with others and tune it for their own (interactive or not) purposes., and better memorize it [11].

Marion Ink [10] proposes to set aside the sensory understanding of perception in favor of the practical-sensitive approach to perception. She insists on the idea of conceiving perception as the interweaving of the five senses that communicate with each other, one in the other. Three concepts emerge in her paper:

1) A particular situation can be interpreted with some central or diffuse indices in a more complete way. Diffuse indices are usually "taken for granted" but our perception records them despite us. Everybody has the experience of these indices but BPB people use them to solve their everyday challenges. Our blind colleague, Frederic, told us one day that when he listened to the rain, he could feel all the objects in his environment by the diffuse sound every object produced with the raindrops.
2) The perception of a situation necessarily includes the perception of individuals who perceive the situation themselves. In this respect, it is collective and affective from the outset. The observation is indirect, any change will have an effect on other elements. Blind and low vision people search for diffuse indices to solve problematic situations and often use these skills individually or collectively.
3) Individuals learn sensitive-practical skills from their interaction partners. They also transmit these skills with time to people close to them. They sometimes even develop a perceiving body together by synchronizing and completing their practical-sensitive skills. In this respect, the term "practical" makes it possible to account for the collective, social, processual and affective character that constitutes the perception of each individual.

3 Blindness Gain and Audio Description

The term "blindness gain" was introduced by Hannah Thompson [8, 15] as a parallel to the notion of "deaf gain" coined by Bauman and Murray [16], Rosemarie Garland-Thomson's concept of "disability gain" [17] and Georgina Kleege's reflections on "gaining blindness" rather than 'losing sight' [18]. Blindness gain considers that blindness in not a problem to be solved, but a benefit of a new perception of the world. Blind and partially blind people may discover in a multisensory way the world and that boosts their inventiveness, imagination and creativity. "Non-visual living is an art". The blindness gain benefits to ANY person, and leads to new ways to share the same space (cultural and work spaces, schools).

Thompson and Chottin [9] add that for BPB people that way of considering life brings productiveness and appreciation of "visual art" through audio descriptions.

3.1 Pragmatic Approach to Perception: Inclusive and Co-created Audio Descriptions, a Case Study

Hereafter, an experiment to create inclusive audio-description of some artworks is presented. This experiment has been led jointly by the University of Rouen, France, Royal Holloway University of London and by the National Research Centre (CNRS, France).

Using the concept of blindness gain, which is part of critical disability studies and more precisely critical blindness studies, and the idea of a practical-sensitive approach to perception named "Co-Created Inclusive Audio Descriptions" have been produced. Several types of works (paintings, embroidery, Bayeux Tapestry [19], The Apocalypse Tapestry [20]) belonging to so-called "visual" art have been used to realize the relevant ICAD [23].

A partially blind person offers an initial description of the work, which is preserved as it is. The partial blindness indeed offers an original point of view on artwork, which deserves to be known and disseminated. Then, a dialogue is established between blind, sighted and partially blind people, and a second description is created, guided by the blind people's questions.

Our methodological approach is in line with that of Sophie Calle, who, as early as 1991, with her exhibition entitled "Blind Colour" presented in MET, NY [21], asked partially blind people to describe monochrome paintings, and, in 2023, on the occasion of the exhibition "À toi de faire, ma mignonne" [22], offered the public multi-vocal descriptions of Picasso's paintings hidden from view (See Fig. 1). For Calle, blindness is not a loss, deficiency or "disability", but has a whole range of aesthetic potentialities (blindness gain).

From there, it is possible to consider that all the access devices made available to blind and partially blind people do not constitute tools, created by volunteer benefactors, intended to compensate for an "impairment", but co-created instruments that expand the always collective, social, processual and affective dimension of the act of perceiving.

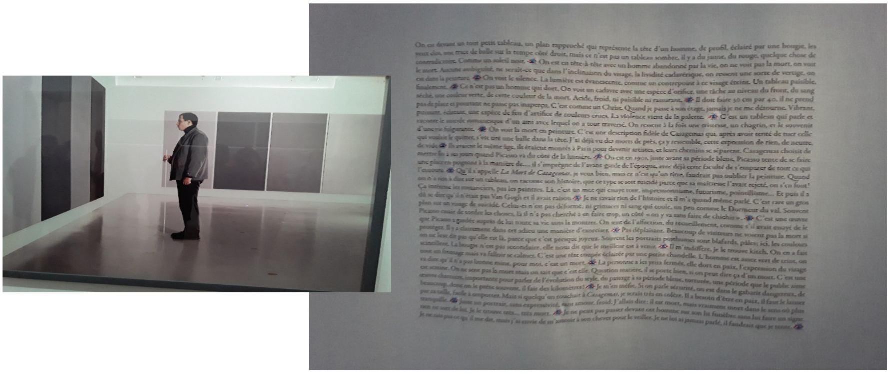


Fig. 1. Sophie Calle's exhibition "À toi de faire, ma mignonne", offered the public multi-vocal descriptions of Picasso's paintings hidden from view.

The implementation of the inclusivity principle in the design of new ICT tools is addressed in Sect. 4 which introduces this special session papers.

4 Information and Communication Technology Mediation

Seven papers constitute the Special Thematic Session. They can be grouped in three categories:

- Four communication present the design of inclusive multimodal devices for VIP accessibility to different data;
- Two communications consider the inclusivity in the workspace;
- One communication proposes an extension of the accessibility implementation with recent technologies.

4.1 Design of Inclusive Multimodal Devices

Redon et al. [5] study different technologies, tactile graphic techniques and supports to allow haptic exploration of bi-dimensional artworks. To evaluate their impact on the museum visitors' experience, the authors seek to establish the benefits and drawbacks and look for further improvements. Several tactile supports have been used in this experiment: a tactile book with embossed paper, a swell paper representation, a 3D printed puzzle, and a haptic device, named F2T (Force Feedback Tablet). BPB and sighted participants were tasked to explore these supports accompanied by audio descriptions.

Luo et al. [4] explore human visual attention when observing artworks to create audio descriptions that will guide tactile exploration via a force feedback tablet F2T. To find the semantically important elements (as the centered focus – to be completed by diffuse information), authors tested with an Eye-tracker people's behavior when observing scenes with and without audio description. The collected data and the small dataset constituted from images of the Bayeux Tapestry is used to train a deep learning model. This model aims to predict saliency on other images.

These two communications contribute to the design of a new multimodal (tactile/haptic, audio and vision) tablet (named Force-Feedback Tablet or F2T). The F2T is based on multimodal data which assist fusion done by human brain. This inclusive tablet (today of the TRL 2–3) aims to be helpful to any museum visitor (BPB and sighted people) as it provides tactile/haptic presentations and/or audio description of a painting. Both representations of an artwork will induce a new aesthetic experience and better understanding of the artwork content of the considered artwork, and will make more memorable museum visits ([19, 21]). Such system awakes our imagination and the creativity.

Vetter et al. [6] present the research results of an in-depth exploration of inclusive and interactive digital music practice with regard to accessibility barriers in the field of computer music software and hardware. Based on the collaboration of BPB and sighted musicians and researchers, new concepts were developed for tangible computer interaction, focusing on interfaces that are both accessible and enable professional artistic expression. The system inspires creativity and may also be used by sighted for their immediate music generation. It is "inclusive" and based on multimodal (music, haptic touch) perception and fusion done by human brain.

Treml et al. [2] analyze the potential of a virtual reality hand tracking system from Ultraleap to be used for braille finger tracking. The aim is to see if the system is accurate enough to reliably detect the braille cell that is currently the focus of the reader's tactile perception. The system introduces the concept of Active Tactile Control (ACT) to improve the accessibility of an information displayed on a braille device.

4.2 Accessibility Implementation with Recent Technologies

Remote collaboration especially in mixed-ability teams makes different individuals face different challenges. Kaschnig et al. [1] analyze remote collaborative interaction between sighted and non-sighted users who worked together on a problem-solving task. They consider the perception and closeness of collaboration, assistance and communication among team members, disruptive elements, workspace awareness, territoriality, or triggers for problematic situations, and indicates both, challenges but also potentials of collaboration within mixed-ability teams.

Ina [7] explores the challenges faced by "visually impaired" individuals using screen readers in the workplace and proposes the development of assistive tools to address difficulties related to screen reader compatibility, image and character recognition, inaccessibility of PDFs, icons without alternative text, and inadequate visual layouts. The paper identifies also new challenges for future screen readers.

4.3 Inclusivity in the Workspace

Takeuchi and Suzuki [3] propose a system for recognizing expiration dates of perishable food products. To solve the problems of skewed, misaligned, or partially missing date images, authors use Spatial Transformer Network (STN). The experimental results showed that the system was able to achieve a high recognition rate of 99.42% for the dataset without dates images with spaces and 98.44% for the dataset with dates images without spaces. Authors pledges in the favour of a clear integration of the object recognition concept when implementing accessibility principle.

5 Items for "Road Maps"

The design of new ICT devices having a significant social benefit expected by the targeted population (BPB in our case) and a benefit of social progress as well as personal development is very complex. Addressing all implementation principles listed in the paper requires multidisciplinary knowledge (philosophy and social science included), excellent high-level technical know-how, and deep understanding of "impairment" and its potential temporal evolution.

To our best knowledge, there is no clear road map for such ICT system developments, despite of the European Commission « The Strategy for the Rights of Persons with Disabilities 2021–2030» [24]. Some points, which may be parts of a such road map, are proposed.

- Establishment of philosophic bases of perception and percept emergence.
- Investigation of physiological and neurocognitive bases for sensory data acquisition and stimulations communications throughout the brain.
- Co-creation of inclusive tools with sustainable design.
- Careful usage of Artificial Intelligence (as usually AI designed solutions are more quantitative than qualitative).
- Open-design platforms (which may be improved in collaborative way) for such tools design and simulation, before their physical realization.

- Ergonomic (displayed data refreshment and precision, relevant presentation, supports for haptic exploration (automated or not), wearability, etc.)
- Evaluation with new, precise methodology and tools allowing to quantify in details the obtained performances of the prototypes of TRL 6–8.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Kashnig, F., Neumayr, T., Augstein, M.: Remote collaboration between sighted and non-sighted people in mixed-ability teams. In: Miesenberger,K.; Kobayashi, M.; Penaz, P. (eds.): ICCHP 2024, Computers Helping People with Special Needs, 19th International Conference; Linz, Austria, Proceedings. Springer, Cham (2024)
2. Treml, M., Zagler, W.L., Busboom, M.: Analyzing the accuracy of braille finger tracking with ultraleap's virtual reality hand tracking system. In: Miesenberger, K., Kobayashi, M., Penaz, P. (eds.): ICCHP 2024, Computers Helping People with Special Needs, 19th International Conference; Linz, Austria, Proceedings. Springer, Cham (2024)
3. Takeuchi, Y., Suzuki, K.: Expiration date recognition system using spatial transformer network for visually impaired. In: Miesenberger, K., Kobayashi, M., Penaz, P. (eds.): ICCHP 2024, Computers Helping People with Special Needs, 19th International Conference; Linz, Austria, Proceedings. Springer, Cham (2024)
4. Luo, W., Djoussouf, L., Lecomte, C., Romeo, K.: Detecting areas of interest for blind people. In: Miesenberger, K., Kobayashi, M., Penaz, P. (eds.): ICCHP 2024, Computers Helping People with Special Needs, 19th International Conference; Linz, Austria, Proceedings. Springer, Cham (2024)
5. Redon, M., Djoussouf, L., Romeo, K., Quéau, Y., Elmoataz, A.: A comparison of audio-tactile exploration methods to discover the apocalypse tapestry. In: Miesenberger, K., Kobayashi, M., Penaz, P. (eds.): ICCHP 2024, Computers Helping People with Special Needs, 19th International Conference; Linz, Austria, Proceedings. Springer, Cham (2024)
6. Vetter, J., Kaltenbrunner, M., Schmid, E.: Sonic interactions – towards accessible digital music-making. In: Miesenberger, K. (eds.): ICCHP Open Access Compendium "Future Perspectives on Accessibility, AT and eInclusion"; Linz, Austria, Proceedings. Johannes Kepler University Linz (2024)
7. Ina, Y.: Survey for development of an assistive tool for screen reader users to utilize icons without alternative text. In: Miesenberger, K. (eds.): ICCHP Open Access Compendium "Future Perspectives on Accessibility, AT and eInclusion"; Linz, Austria, Proceedings. Johannes Kepler University Linz (2024)
8. Thompson, H., Blind Spot, Blog: https://hannah-thompson.blogspot.com/2018/10/blindness-gain-and-art-of-non-visual.html, Accessed 07 Apr 2024
9. Chottin, M., Thompson, H.: Blindness gain as worldmaking: audio description as a new partage du sensible. L'Esprit Créateur **61**(4), 32–44 (2021)
10. Ink, M.: Pour une approche protico-sensible de la perception. Ce que les handicaps visuels et auditifs nous apprennent des sens. In: Battesti, V., Candau, J. (eds.) : Apprendre les sens, apprendre par les sens. Anthropologie des perceptions sensorielles. PETRA (2023)
11. Eardley, A.F., Mineiro, C., Neves, J., Ride, P.: Redefining access: embracing multimodality, memorability and shared experience in Museums Curator: Museum J. **59**(3), 263–286 (2016)
12. D'Humières, R.: Quand la Cognition rencontre la Psychanalyse, Edilivre, p. 20 (2013)

13. Franceschini, N.: Early processing of colour and motion in a mosaic visual system. Neurosci. Res. **2**(Suppl.), S17–S49 (1985)
14. Franceschini, N.: Small brains, smart machines: from fly vision to robot vision and back again. In: Proceedings of the IEEE, vol. 102, no. 5 (2014)
15. Chottin, M., Thompson, H.: Blindness gain or the unexpected advantages of audio description, disability and disciplines. In: International Conference on Educational, Cultural, and Disability Studies, Centre for Culture and Disability Studies, Liverpool Hope University (2019)
16. Bauman, H-D.L., Murray, J.J. (eds.): Deaf gain: Raising the stakes for human diversity. University of Minnesota Press, Minneapolis (2014)
17. Fox, A.M., Krings, M., Vierke, U.: Disability gain' and the limits of representing alternative beauty. In: Liebelt, C., Böllinger, S., Vierke, U. (eds.) Beauty and the Norm. Palgrave Studies in Globalization and Embodiment; Palgrave Macmillan, Cham (2019). https://doi.org/10.1007/978-3-319-91174-8_5
18. Kleege, G.: Blind rage: an open letter to Helen Keller. Sign Lang. Stud. **7**(2), 186–194 (2007)
19. https://www.bayeuxmuseum.com/en/ Accessed 14 Apr 2024
20. The tapestry of the Apocalypse: the essential in 2 minutes! (chateau-angers.fr), Accessed 14 April 2024
21. https://www.metmuseum.org/art/collection/search/284117, Accessed 14 Apr 2024
22. https://www.museepicassoparis.fr/fr/toi-de-faire-ma-mignonne, Accessed on 14 Apr 2024
23. Romeo, K., Thompson, H., Chottin, M.: Inclusive multimodal discovery of cultural heritage: listen and touch. In: Miesenberger, K., Kouroupetroglou, G., Mavrou, K., Manduchi, R., Covarrubias Rodriguez, M., Pen´az, P. (eds.), Computers Helping People with Special Needs. ICCHP-AAATE 2022. LNCS, vol. 13341, pp. 278–285 (2022)
24. European Commission, The strategy for the rights of persons with disabilities 2021–2030
25. https://eur-lex.europa.eu/EN/legal-content/summary/strategy-for-the-rights-of-persons-with-disabilities.html, Accessed 10 Mar 2024

Detecting Areas of Interest for Blind People: Deep Learning Saliency Methods for Artworks

Wenqi Luo(✉), Lilia Djoussouf, Christèle Lecomte, and Katerine Romeo

Rouen Normandy University, Normandy Univ, LITIS UR 4108, Rouen, France
wenqi.luo@univ-rouen.fr

Abstract. The purpose of this study is to explore human visual attention when observing artworks to create audio descriptions that will guide tactile exploration via a force feedback tablet F2T. To find the semantically important elements, we tested with an Eye-tracker people's behaviour when observing scenes with and without audio description. The collected data and the small dataset constituted from images of the Bayeux Tapestry will be used to train a deep learning model. This model aims to predict saliency on other images. We use three models to predict images of Bayeux Tapestry: Resnet50, TransalNet, SAM-LSTM-Resnet. We can see that after the training phase on our dataset, the predictions of chosen model (SAM-LSTM-Resnet) are closer to the ground truth and have better correlation, which is a significant improvement over the same model learnt with only the original dataset.

Keywords: Accessibility · Artworks · Saliency detection · Blind prople · Deep learning

1 Introduction

In museums, blind and partially blind people are restricted in their movement and find it difficult to appreciate artworks in an autonomous way. To facilitate access to cultural places for people with different visual abilities, an academic prototype was designed to assist museum visitors. This haptic device, F2T (Force Feedback Tablet) [1] was developed with the aim of allowing everyone to perceive an image of a two-dimensional work of art (see Fig. 1) with the guidance of an audio description. However, to avoid too much information to transmit via F2T, we need to analyze important elements of the image. The research question will be: what kind of elements human eye is naturally attracted to while observing artworks?

To find the semantically important elements, we tested with an Eye-tracker people's behaviour when observing scenes with and without audio description. The small dataset constituted from images of the Bayeux Tapestry will be used to train a deep learning model. This model aims to predict saliency on other images.

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 473–480, 2024.
https://doi.org/10.1007/978-3-031-62846-7_56

Fig. 1. F2T Force Feedback Tablet used to transmit a haptic impression of a virtual image. The joystick moves in 2D under the impulse of two motors on the horizontal and vertical axes.

2 State of the Art

Before starting the eye movement experiment, we should first understand what visual attention is. Visual attention consists of two different types of attention: overt visual attention and covert visual attention.

As in the study presented by Olivier Le Meur [2], overt visual attention can be observed. When people consciously want to focus on a certain area or object, they make noticeable head or eye movements to focus their gaze on that area. Covert visual attention uses external perception to perceive the environment without moving the head or eyes. This type of attention is not easy to detect. So, in our experiments, we use overt visual attention, because it can be directly detected by Eye-trackers through eye movements. However, our attention abilities are affected by many influences, especially prior knowledge (top-down attention) and the influence of stimuli (bottom-up attention), as presented in Orquin's study [3]. Some researchers work on attention methods and are interested in deep learning methods.

Special attention is given to predicting saliency maps related to human eye gaze, Cornia et al. [4] used the saliency-based attention model LSTM. Lou et al. [5] used Transformer to improve the performance of salient regions detection. Simonyan et al. [6] used Resnet to create saliency maps. As part of this paper, we start by evaluating different deep learning models on a database consisting of Bayeux Tapestry images [8]. This is a historical embroidery depicting the story of William the Conqueror's conquest of England. It is 70 m long and 50 cm high.

3 Methodology

3.1 Eye-Tracker

48 participants ($n_m = 33$ males, $n_f = 15$ females), all volunteers (between 21 and 64 years old), took part in our Eye-tracking tests. Since the Eye-tracker cannot detect correctly the gaze of partially sighted participants, we chose to have only sighted participants. The test of one partially sighted participant produced data that was not usable. The other participants were all sighted.

The participants were invited to sit in front of a computer and observe images of the Bayeux Tapestry. Before the experiment began, participants were informed of the purpose of the experiment. They were asked to fill in their personal data and sign an informed consent form for the use of their data. In the experiment, the Tobii Pro Fusion brand Eye-tracker device [9] was placed below the screen (see Fig. 2). It operates at a sampling frequency of 250 Hz. The distance between the device and participant's eyes is 60 cm in average.

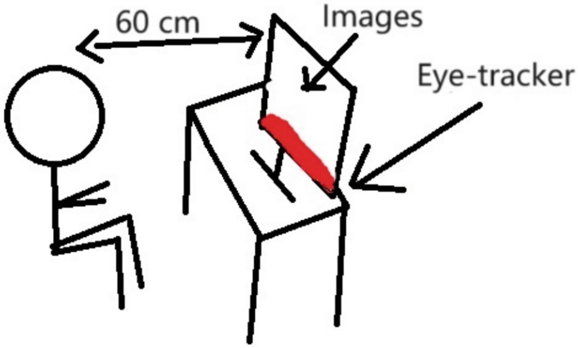


Fig. 2. A participant observes the images on the screen.

In the experiment, first, each participant calibrated the eye tracker, 9 points were continuously displayed at the different positions on the screen in a certain order. Participants were asked to observe them as instructed in the calibration procedure. Then the images of the Bayeux Tapestry were displayed continuously on the screen for 6 s. Each participant will observe the images of several scenes. The images of one scene included a brief audio description and the other scenes were without audio description. Each participant will observe approximately 85 images. To avoid eye strain, participants had the option to take a 2-minute break after observing approximately 40 images. Between two images, a gray background image appears with a cross in the center of the screen. This gray background will be displayed for 1 s to concentrate the gaze in the centre of the screen. The display of the images starts with the full image and then images subdivided into multiple thumbnails for one scene (see Fig. 3). A total of 314 images from 28 scenes of Bayeux Tapestry were tested. Finally, participants were asked to answer a questionnaire with a Likert scale and some open-ended questions.

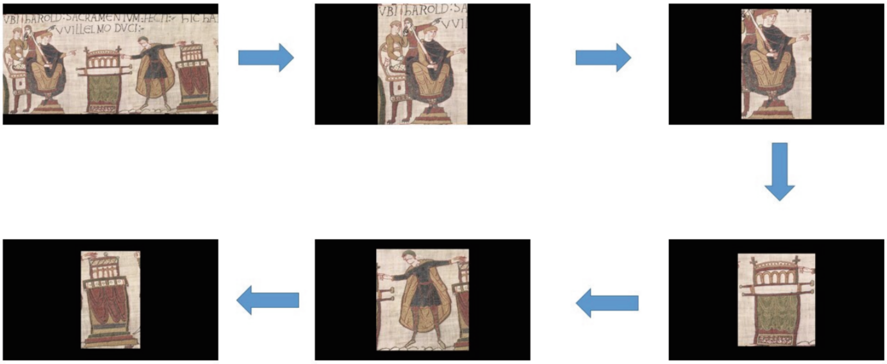


Fig. 3. Example of images displayed.

3.2 Deep Learning

In the literature, some researchers have worked on attention methods with Deep learning methods. The learning models used by the authors of these articles are supervised. In supervised learning, the model attempts to discover structure or regularities from the data with labelling of the input data. These models cover a wide range of architectures in the field of supervised learning, including Convolutional neural network, Transformers, and Long Short-Term Memory (LSTM) Networks. When training these deep learning models, researchers typically use a variety of loss functions to measure the difference between the model predictions and the true labels and tune the model parameters by optimising the algorithm to minimise the loss. When adjusting the parameters inversely based on the loss values obtained, the model encounters problems such as gradient explosion and gradient vanishing, which are solved by residual connection, different activation functions. For images from Bayeux Tapestry, we tested three different models using Salicon and ImageNet database: ResNet50 [7] (a model with a simple architecture), SAM-LSTM-RESNET [4] and TranSalNet [5]. ResNet50 is using a 50-layer Convolutional Neural Network with some residual connection to avoid exploding gradient or vanishing gradient. SAM-LSTM-RESNET is using an LSTM. TranSalNet is using a Transformer model, Encoder with 3-layer Convolutional neural network, Decoder with 7-layer Convolutional neural network. We use these three models to predict our images. We chose the model whose results are the closest to the ground truth data from the Eye-tracker, and we used this model to train a new dataset consisting of eye-tracking test results. The new dataset has the same format as Salicon.

4 Results

4.1 Eye-Tracker

Two types of visualization were generated from the collected Eye-tracker data: heat maps and gaze plots. We obtained approximately 33,000 fixation points and 3,000 saccades points per participant. We obtained heat maps per image for 7 people in average (see Fig. 4 column 2). We observed that, most participants focused on the faces of the human figures as well as the text.

4.2 Deep Learning

Comparison. We use three models to predict images of Bayeux Tapestry. As presented in the picture 4, we can observe that the ResNet50 (see Fig. 4 column 3) and TranSalNet (see Fig. 4 column 5) model results are diffused and that the eye-tracker results are more specific to the region of interest. So, we chose SAM-LSTM-Resnet model (see Fig. 4 column 4) for further training. As our dataset only has 314 images, we chose the model whose results are the closest to the ground truth for training.

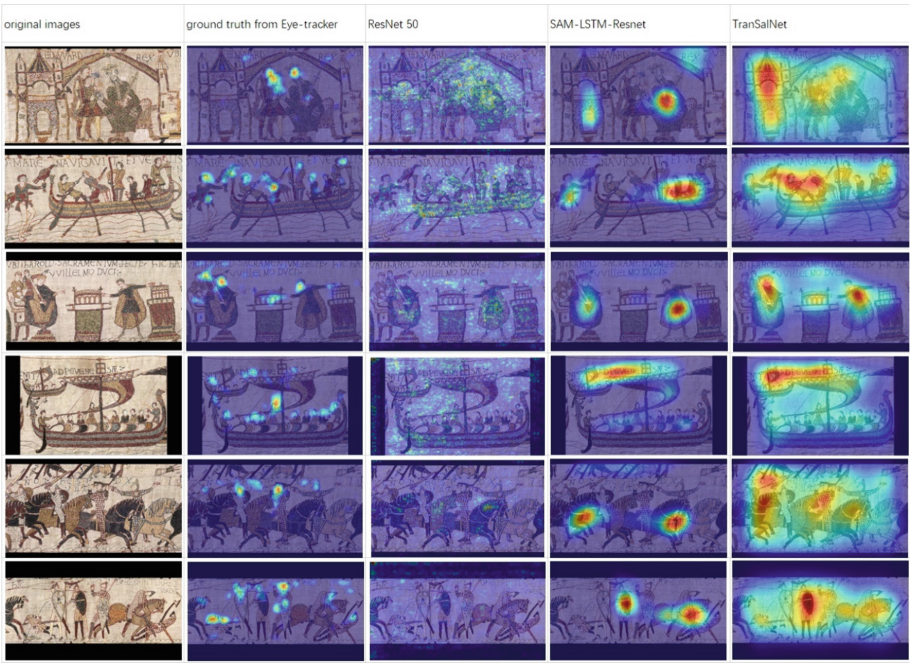


Fig. 4. Comparison of ResNet50, SAM-LSTM-RESNET, TranSalNet models.

Training Results. After the training phase with our dataset, we can compare the learned predictions with the Eye-tracker results. We can observe that the main salient points given in the Eye-tracker results are successfully predicted. In the example of Fig. 5(see Fig. 5 column 4), the faces are usually well detected. However, other elements (e.g. boats) are not considered salient regions.

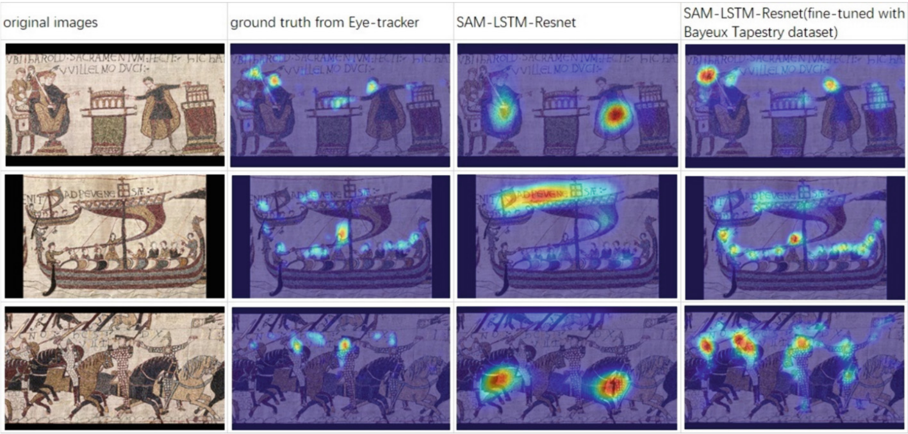


Fig. 5. Prediction by SAM-LSTM-RESNET.

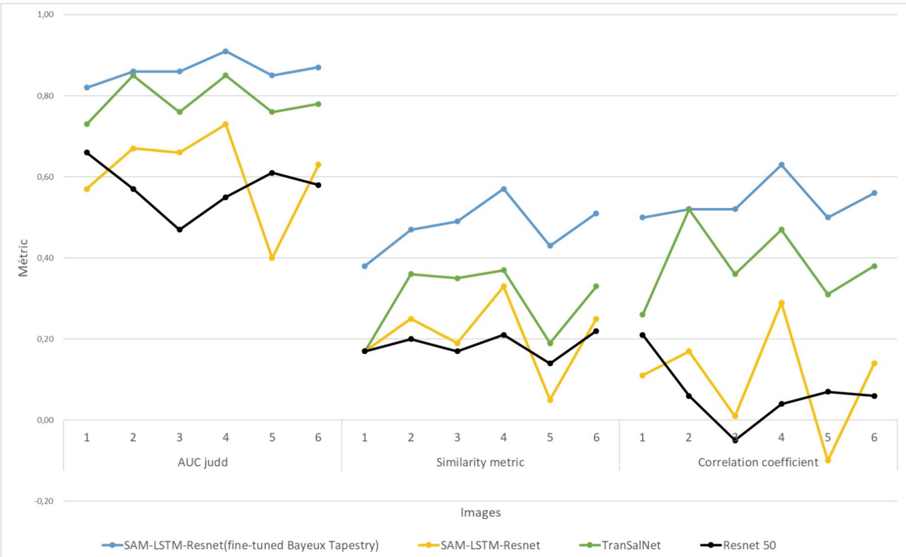


Fig. 6. Metric.

We calculated multiple metrics to compare the results before and after deep learning on our dataset, and we chose three typical metrics to draw as line graphs (see Fig. 6). AUC judd calculates the area under the curve of the ROC curve, Similarity metric calculates the similarity between the predicted image and the ground truth, Correlation coefficient calculates the correlation between the predicted image and ground truth. The closer the value of each of these three metrics is to 1, the closer the predicted image is to the ground truth and the more relevant it is. We can see that after the learning phase on our dataset, the predictions of SAM-LSTM-Resnet (blue line) are closer to the ground truth and have better correlation, which is a significant improvement over the same model learnt with only Salicon dataset.

5 Conclusion

In this article, we conducted eye-tracking experiments on 314 images of the Bayeux Tapestry under two experimental conditions (observation of images with and without short audio descriptions). With a small database, we were able to train a learning model to predict areas of salience and thus generalise the location of the observer's attention. Characters (faces and hands) and text were observed to be the most salient elements in the images. The next stage of the project will involve adapting databases dedicated to the observation of works of art. An exhaustive comparison of learning models will have to be carried out in order to obtain more accurate predictions from images of works of art that include features (architecture, characters, animals, plants, etc.) from the Bayeux Tapestry. Then, the Eye-tracker will be used to guide the haptic exploration of the elements to be discovered in the scene observed. In the future, our work will involve these salience maps to create representations of works of art that can be used by the F2T (see Fig. 1) to help blind people perceive these works of art more effectively, in particular for adapting audio descriptions with haptic exploration.

Acknowledgments. We would like to thank all the participants in our study. This research work was sponsored by NormaSTIC, an organisation of Normandy University. The database was produced as part of the ANR IMG ANR-20-CE38-0007 project.

References

1. Gay, S.L., Pissaloux, E., Romeo, K., Truong, N.-T.: F2T: a novel force-feedback haptic architecture delivering 2d data to visually impaired people. IEEE Access **9**, 94901–94911 (2021). https://doi.org/10.1109/ACCESS.2021.3091441
2. Le Meur, O., Le Pen, T., Cozot, R.: Can we accurately predict where we look at paintings? Plos One **15**(10), e0239980 (2020). https://doi.org/10.1371/journal.pone.0239980
3. Orquin, J.L., Loose, S.M.: Attention and choice: a review on eye movements in decision making. Acta Psychologica **144**(1), 190–206 (2013). ISSN 0001-6918. https://doi.org/10.1016/j.actpsy.2013.06.003

4. Cornia, M., Baraldi, L., Serra, G., Cucchiara, R.: Predicting human eye fixations via an LSTM-based saliency attentive model. IEEE Trans. Image Process. **27**(10), 5142–5154 (2018). https://doi.org/10.1109/TIP.2018.2851672
5. Lou, J., Lin, H., Marshall, D., Saupe, D., Liu, H. TranSalNet: towards perceptually relevant visual saliency prediction. Neurocomputing **494**, 455–446 (2022). https://doi.org/10.1016/j.neucom.2022.04.080
6. Simonyan, K., Vedaldi, A., Zisserman, A.: Deep inside convolutional networks: visualising image classification models and saliency maps. In arXiv preprint (2013). https://doi.org/10.48550/arXiv.1312.6034
7. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, pp. 770–778 (2016). https://doi.org/10.1109/CVPR.2016.90
8. Bayeux tapestry homepage. https://www.bayeuxmuseum.com/la-tapisserie-de-bayeux/, Accessed 17 Jan 2024
9. Tobii pro fusion Homepage. https://www.tobii.com/products/eye-trackers/screen-based/tobii-pro-fusion, Accessed 11 Jan 2024

A Comparison of Audio-Tactile Exploration Methods to Discover the Tapestry of the Apocalypse

Marjorie Redon[1(✉)], Lilia Djoussouf[2], Katerine Romeo[2], Yvain Quéau[1], and Abderrahim El Moataz[1]

[1] Normandie Univ, UNICAEN, ENSICAEN, CNRS, GREYC, Caen, France
marjorie.redon@unicaen.fr

[2] Rouen Normandy Univ, Normandy Univ, LITIS UR 4108, Rouen, France
lilia.djoussouf@univ-rouen.fr

Abstract. Artwork accessibility is a major issue for blind and partially blind museum visitors. To that extent, tactile graphic techniques and supports have been developed to allow haptic exploration of bi-dimensional artworks which is not always available in museums. This study aims to compare different technologies with a focus on the Tapestry of the Apocalypse, located in the Castle of Angers, France. To evaluate their impact on the museum visitors' experience, we seek to establish the benefits and drawbacks and look for further improvements. Several tactile supports have been used in this experiment: a tactile book with embossed paper, a swell paper representation, a 3D printed puzzle, and a haptic device, named F2T (Force-Feedback Tablet). Blind, partially blind, and sighted participants were tasked to explore these supports accompanied by audio descriptions. Participants were presented with several Likert scale and open-ended questions to rate their general feeling. Overall, all the conditions received good ratings from the participants. However, more nuanced comments were gathered on each condition. New ways to improve the technologies have been proposed through users' feedback.

Keywords: artwork accessibility · multisensory exploration

1 Introduction

The Tapestry of the Apocalypse is a unique artwork of the fourteenth century. Recognized as the world's largest collection of medieval tapestries, this document is a 104 m long and 450 cm height medieval set of tapestries. Throughout 84 scenes, mythology and middle age reality are mixed. To facilitate the accessibility of this artwork to all, the Castle of Angers offers special guided visits to blind and partially blind (BPB) people. However, those visits are limited and must be scheduled with a minimum of five persons. Independence of navigation and art appreciation have been demonstrated to improve satisfaction and increase the motivation to discover art [2].

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 481–488, 2024.
https://doi.org/10.1007/978-3-031-62846-7_57

In this paper, we will focus on art appreciation. We attempt to find different audio-tactile systems adapted to the BPB audience. A tactile book of the castle and its tapestry is already available in the castle shop [3]. Yet, this tactile board is not in open access inside the castle. Furthermore, to our knowledge, no study has been published on tests made with BPB people on this tactile book. Moreover, this book has not been designed to be used with complete autonomy, with some advice being given inside for helpers.

To reduce the cognitive overload generated by the large-scale integration of tactile information through touch we propose alternative methods of tactile exploration. These new audio-tactile exploration methods are composed of swell paper representations, 3D prints, and a force-feedback haptic device. Our three new experimental systems aim to be inclusive and in the end, to be used with complete autonomy. By comparing all of those methods we attempt to find the advantages and drawbacks of each method with the aim of improving them in order to get the most adapted system.

In the first part of this paper, we describe all of the systems and the experienced method. Then we compare the participants' feelings and comments for each system. A conclusion on those tactile systems follows, along with a discussion on the possible ways to improve them.

2 State of the Art

In the literature, the representation of 2D artworks like tapestries or paintings is mostly transcribed to 2.5D presentation [14]. Edman [5] listed a multitude of materials to support tactile graphics (e.g. swell paper, vacuum-forming). These supports are still used by experts specializing in the transcription of images for BPB people through simplification. To create these representations, guidelines have been established to achieve better haptic perception by the Braille Associations [1]. Those techniques and guidelines are used to create the swell paper representations presented in this paper.

However, it has been shown that the tactile sense is not enough to promote independent exploration, especially for 2D artworks. It is recommended to complement the experience by presenting auditory information, such as audio-description [4,13]. To make visual presentation accessible to BPB people, most researchers will look into cross-modal concepts through visual-to-tactile or visual-to-auditory cues. As such, all of our tactile experiences will be supported by audio descriptions (AD). Though there is no consensus known on how to design artwork representations for BPB people, most researchers in this domain will look into an iterative approach involving the targeted population. Kobayashi *et al.* [8] proposed an iterative approach involving BPB people to create a representation with progressive details of the "Storm below Mount Fuji" by Hokusai on swell paper. Raynal *et al.* [11] proposed the use of tangible graphics and an interactive interface in an educational setting in the form of a board game for the BPB. They showed that this type of device improves collaboration between people with different levels of visual acuity, thus promoting an inclusive design.

Our work is in the same vein and intends to use this experience to improve our tactile models based on the feedback we receive. However, it is difficult to evaluate if the created devices will be effectively used in the museum context. Our goal is to not isolate BPB museum visitors from other visitors. This is why it is important to involve people with different visual abilities to approach an inclusive design [9], which we will apply for our experiment. According to our knowledge, no research has been conducted to compare multiple audio-tactile devices in an inclusive setting. To that extent, this article has the objective of looking into the relevance of art representation methods and their impact on the museum experience for an inclusive audience. The next section will focus on describing the methodology for the techniques used.

3 Methodology

3.1 Experimental Conditions

For the experiment conducted here, several types of presentations, illustrated in Fig. 1, have been chosen to be compared, with two presentations based on a tactile book provided by the staff of the Castle of Angers, and each accompanied by an AD to assist the exploration. For each condition, free exploration was not temporally restricted. An estimated time of 15 min was intended by the supervisors, however, participants took more time to explore. Thus not all participants were able to test every experimental condition in one day.

Condition A. The tactile book uses an embossing technique via Pachica paper to represent elements from scenes of the artwork. An AD extracted from the CD, delivered with the tactile book, is provided to accompany the exploration. However, it has been specified that a helper is required to use this tactile book, thus not allowing independent exploration from blind users.

Condition B. The same tactile book and AD are used accompanied by a swell paper representation of one or several elements of the scene designed by experts from “Centre Normandie Lorraine”. The particularity of the swell paper representation is the simplification of the element of a scene through contour presentation from the tactile book. It introduces “blank space” textures for better tactile perception and high contrast colors.

Condition C. A 3D printed puzzle was designed to make detachable pieces from scenes using the pipeline of Redon *et al.* [12]. Each scene element has been segmented using the Segment Anything Model (SAM, [7]), while their height value was manually attributed. An adapted AD from the tactile book’s CD has been used to accompany users’ exploration. The ADs are divided into several sections. In this way, free exploration time is possible for each described object.

Condition D. The F2T (Force Feedback Tablet) is a 2-DOF haptic device that allows haptic exploration via a flat joystick [6]. The presentation of a scene is made by deploying three exploration modes composed of: a global and a detailed

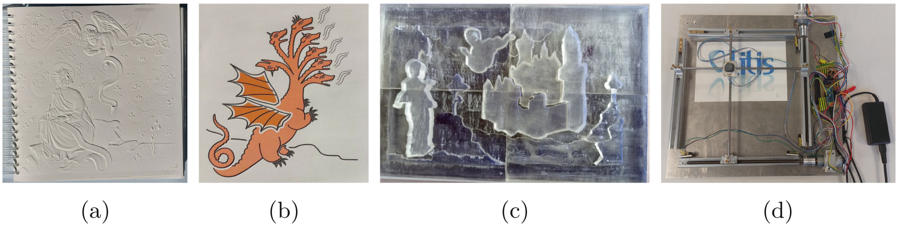


Fig. 1. Tactile supports. (1a) vacuum-forming support. (1b) Swell paper. (1c) 3-D puzzle. (1d) The F2T prototype.

guidance; and an exploration without guidance [10]. This method uses annotated paths and virtual images with haptic effects and audio feedback areas.

Conditions A, B, and C were conducted in the gallery of the Tapestry of the Apocalypse, while condition D was tested in one of the buildings of Castle of Angers, to limit noise in the gallery. To avoid learning bias, each condition represented distinct scenes of the tapestry with a minimum of shared information.

3.2 Participants' Profile

A total of 13 participants aged between 41 and 77 years old have been recruited, with nine BPB people. Among them, seven had already used ADs and six had already visited the Tapestry of the Apocalypse. Before undertaking the investigations, ethical clearance was obtained from the ethical committee. The participants number varies from one condition to another: nine for condition A; eight for condition B; ten for condition D; and eleven for condition C.

4 Results

Following each experience, a questionnaire was completed by the participants and transcribed by helpers. It is composed of a Likert scale, with ratings ranging from 1 ("not at all") to 5 ("a lot"), and open-ended questions. The results have been analyzed by using the median, the lowest, and the highest values. Conditions B, C, and D are still in an experimental phase, while condition A is tested to have a base for comparison.

4.1 Likert-Scale Results

We observe in Table 1 that all tactile supports were highly appreciated, but not at the same level. If we focus our attention on the highest and lowest value condition B is the most satisfying system, with less variability in answers. For the detail level in the ADs, the tactile systems, and the experiment's length no preference can be identified, see Tables 2 and 3. However, if we look at the comments real differences appeared. This will be discussed further below.

Table 1. General ratings about experiences. Five answers were possible and transcribed as numbers, 1 meaning "not at all" and 5 "a lot". Here are represented the median, the lowest, and the highest marks.

	Did you enjoy the experience?	Did you find the experience interesting?	Was the experiment understandable?	Has the experience improved your feelings about the Tapestry of the Apocalypse?	Are you satisfied with the amount of information provided by the represented scene?
Condition A	5 [2; 5]	4 [2; 5]	4 [2; 5]	5 [2; 5]	4 [2; 5]
Condition B	5 [4; 5]	5 [4; 5]	5 [4; 5]	5 [1; 5]	4 [3; 5]
Condition C	4 [2; 5]	4 [2; 5]	5 [1; 5]	4 [1; 5]	4 [4; 5]
Condition D	4 [1; 5]	4 [1; 5]	4 [1; 5]	4 [1; 5]	4 [1; 5]

Table 2. Experiences details rating, 1 meaning "not enough" and 5 "too much".

<table>
<tr><th></th><th>What do you think of the detail level in the audio description?</th><th colspan="2">What do you think of the tactile level ?</th></tr>
<tr><td>Condition A</td><td>3 [3; 4]</td><td colspan="2">3 [3; 5]</td></tr>
<tr><td>Condition B</td><td>3 [1; 3]</td><td>Tactile book: 3 [3; 4]</td><td>Swell paper: 3 [1; 3]</td></tr>
<tr><td>Condition C</td><td>3 [2; 5]</td><td colspan="2">3 [1; 5]</td></tr>
<tr><td>Condition D</td><td colspan="3">3 [1; 5]</td></tr>
</table>

Table 3. Experiences length rating, 1 meaning "too short" and 5 "too long".

	What do you think of the experiment's length?
Condition A	3 [1; 3]
Condition B	3 [1; 3]
Condition C	3 [2; 4]
Condition D	3 [3; 4]

4.2 Participants' Comments

During the questioning phase, participants' suggestions were noted by the supervisor. This allows more nuanced appreciation. Globally, participants found that all the systems and supports lacked of guidance.

Some participants found condition A instructive and liked its esthetics: participant 10 said "I liked it, it was a good discovery and instructive". Some comments highlighted the sensation of frustration, sometimes due to the complexity of the tactile support or due to the presence of element descriptions that are not tacitly represented. Participant 2 commented "It is stressful and frustrating not being able to tactilely identify what's on the audio". Another issue was the height of the reliefs which did not allow enough sensation, and the speed of the

AD, which did not leave time to find elements according to the participants: P3 said "Audio flow should be slowed down and relief heights increased". To that extent, some participants wished to control ADs, P2 said: "I would prefer to manage the audio flow myself and have more guidance".

For condition B the speed was also a problem, which makes sense because we used the same type of audio support: P1 said "The AD is too fast, when I touch I need to really explore". The other point was about the level of detail when the tactile book was found overloaded with details, and the swell paper representation was under-loaded, representing only certain elements of the scene. Moreover, some elements considered important were missing. As we prepared the swell paper as a complement for some details: P2 said "Two pages were necessary to explain the scene, here, it's reductive. Missing elements are essential to give the meaning or content of the image". The principle of two supports with different levels of difficulty was highly appreciated: P6 said "I like the different treatment of the two supports".

For condition C, ADs were partitioned. Some participants appreciated it: P2 said "It's good that the ADs of each item are detached". It was also a new approach so it was disturbing for people, making it too complicated to grasp: P6 said "The change of dimension between objects is disturbing"; P9 found that it was not easy to understand with touch and the guidance with audio was not clear enough. Moreover, it was complicated for BPB to reassemble the puzzle once it was unraveled: P9 said "It's complicated to replace the removable elements"; P2 explained "When it's all off, you don't know which piece you're holding". However, participants found the idea innovative and with great potential: P1 said "The idea is brilliant"; P10 commented "I've never seen this type of device"; P4 opined "It was a good idea, but you would have to be used to the object to be able to put it back on the support". During the experience, participants offered us new perspectives. For example, P9 suggested adding a marker to start with and to move around. P12 suggested other perspectives: adding a magnetic side like on children's puzzles, cutting into 6 side-by-side frames to help find the objects, and creating a new system consisting of a flat surface with shapes that rise with the AD. P2 suggested adding a plate with boxes and numbers for easier storage and retrieval of objects.

Regarding condition D, participants generally appreciated how the content of the tapestry was presented to them through complementary exploration modes (global and detailed guidance). Some of them were surprised, by the fine details of the contours provided by the detailed guidance mode. However, as it is a novel technology, people needed a learning phase to enjoy this new tactile support as commented by P6: "It would be great to have a learning phase". They also suggested the inclusion of sound signals to accompany guided exploration: P2 said "There is no signal to distinguish between elements (for the detailed guidance mode)" and P5 noted "Silent areas should be explained". Other participants noted that certain haptic effects hinder further exploration. The motor noise disturbed some participants, even while wearing headphones, and some effects were found to be hard to understand.

5 Conclusion and Perspectives

In this article, we compared several tactile supports intending to make the Tapestry of the Apocalypse accessible to BPB persons. We first described these supports, then we showed that all of them have been appreciated. However, some improvements must be made whether in terms of speech speed or complexity of information, tactile, and audio. The addition of swell paper representation support is intended to represent the elements of a scene in a simplified way. It complements the tactile book use, which alone is difficult to grasp. It contains interesting details that are pleasant to touch, but difficult to interpret by touch only.

The 3D print is a new interactive way to discover art, allowing interesting interactions. However, it was disturbed by its presentation method and object shapes. It needed to be simplified and clarified in the tactile information and the orientation. For future 3D puzzles, we will add some physical landmarks to guide during the exploration. Moreover, a magnetic part will be put on the support and removable objects to make the puzzle reassembling effortless. Another major modification will be to create a storage plate to easily retrieve objects.

The F2T is a modular support that allows one to explore a scene. Global guidance mode was found to be a good way to locate elements from a scene, and overall all exploration methods complement each other. Nevertheless, in addition to a learning phase that needs to be reviewed to make it easier to use. Each presented exploration mode will be improved, notably by looking into how multiple haptic effects can be integrated for better perception of a scene. Improvements will enable better interpretation during the exploration without guidance. These modes would be more understandable by adding sound signals to indicate changes between elements. Regarding system design, a new prototype is under construction to lessen motors' noise.

For future works, it could be interesting to combine these different methods according to the participants' feedback. The fine relief of the book could be combined with the interactivity of the 3D print. The F2T global guidance mode could reinforce the 3D print exploration. We will continue to improve our prototypes by further involving participants in the design and evaluation.

Acknowledgments. Access to the Tapestry of the Apocalypse was made possible by Castle of Angers and the city of Angers, under the supervision of Catherine Leroi and Damien Perdriau. We would also like to show our gratitude to the voluntary supervisors involved. This study was funded by ANR IMG Project (ANR-20-CE38-0007) and a grant from Dassault Systems.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Braille Authority of North America, Guidelines and Standards for Tactile Graphics (2010), https://www.brailleauthority.org/sites/default/files/tg/web-manual/index.html, (Accessed 25 Mar 2024)
2. Asakawa, S., et al.: An independent and interactive museum experience for blind people. In: 16th International Web for All Conference, pp. 1–9 (2019)
3. Catherine, L., Hoëlle, C.: La tenture de l'Apocalypse d'Angers - CMN. Editions du patrimoine - Braille publishing (2015)
4. Djoussouf, L., Romeo, K., Chottin, M., Thompson, H., Eardley, A.F.: Inclusion for cultural education in museums, audio and touch interaction. In: Archambault, D., Kouroupetroglou, G. (eds.) Studies in Health Technology and Informatics. IOS Press (2023)
5. Edman, P.: Production methods and materials. In: Tactile graphics. American Foundation for the Blind (1992)
6. Gay, S.L., Pissaloux, E., Romeo, K., Truong, N.T.: F2t: a novel force-feedback haptic architecture delivering 2d data to visually impaired people. IEEE Access, 94901–94911 (2021)
7. Kirillov, A., et al.: Segment anything (2023)
8. Kobayashi, Y., Miyagi, M., Ishikawa, E.: Introducing art to people with visual impairment through tactile pictures: a survey of art museums in japan. In: ICCHP-AAATE 2022 Open Access Compendium Assistive Technology (2022)
9. Lloyd-Esenkaya, T., Lloyd-Esenkaya, V., O'Neill, E., Proulx, M.J.: Multisensory inclusive design with sensory substitution. Cognitive Res. Principles Implicat. (1), 37 (2020)
10. Pissaloux, E., et al.: A new approach to assist virtual image accessibility for visually impaired people. In: Antona, M., Stephanidis, C. (eds.) Universal Access in Human-Computer Interaction. User and Context Diversity, pp. 251–261. Springer International Publishing (2022). https://doi.org/10.1007/978-3-031-05039-8_18
11. Raynal, M., Ducasse, J., Macé, M.J.M., Oriola, B., Jouffrais, C.: The FlexiBoard: tangible and tactile graphics for people with vision impairments. Multimodal Technol. Interact. (3), 17 (2024)
12. Redon, M., Pizenberg, M., Quéau, Y., Elmoataz, A.: 3d surface approximation of the entire bayeux tapestry for improved pedagogical access. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops, pp. 1593–1602 (2023)
13. Romeo, K., Thompson, H., Chottin, M.: Inclusive multimodal discovery of cultural heritage: listen and touch. In: Computers Helping People with Special Needs : 18th ICCHP-AAATE 2022, pp. 278–285 (2022)
14. Vaz, R., Freitas, D., Coelho, A.: Blind and visually impaired visitors' experiences in museums: increasing accessibility through assistive technologies. Inter. J. Inclusive Museum (2) (2020)

Sonic Interactions - Towards Accessible Digital Music-Making

Jens Vetter[1(✉)], Martin Kaltenbrunner[1], and Erich Schmid[2]

[1] Tangible Music Lab, University of Arts Linz, Linz, Austria
{jens.vetter,martin.kaltenbrunner}@kunstuni-linz.at
[2] Institute for the Blind, Vienna, Austria
erich.schmid@aon.at

Abstract. Can tangible computer music interfaces inspire creativity and enhance artistic stage performances for visually impaired and blind musicians? What are the benefits of engaging with tangible interfaces in the context of computer music and accessibility? This paper presents the research project *Tangible Signals* and its results. It summarizes the research process including the development of an accessible music software, the design of three tangible user interfaces and their use in workshops, expert feedback sessions and artistic live performances by visually impaired and blind individuals. The resulting material is grouped into four key topics and will be presented in detail. This includes a discussion of inclusion and computer music, sonic interaction with tangible interfaces, observations of artistic practice and suggestions for future use cases.

Keywords: Human-Computer Interaction · Tangible Interaction Design · Assistive Technology · Musical Interfaces · Sonic Interaction

1 Introduction

Making music is a joy, an artistic way of expressing one's own feelings, engaging with the world and connecting with others. The computer as one of the most versatile and suitable tools for digital music-making is used in many music studios, in schools and education and has long since conquered the stage as a digital musical instrument.

Nevertheless, using the computer intuitively, expressively and artistically poses accessibility barriers for visually impaired and blind musicians (VIB), whether it is for digital content creation [21], computer programming [9], as a means for digital audio creation or as musical instrument [11]. Assistive technologies such as Braille displays, screen readers or text-to-speech software do enable VIB individuals to use computers and mobile devices. But popular music software still has limited support for these assistive technologies, e.g. Digital Audio Workstations (DAW) like *Reaper* still require additional scripts to be accessible for VIB producers. Extensions such as VST plugins often have no screen reader support at all and can only be controlled with the mouse [11]. Yet it is the

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 489–496, 2024.
https://doi.org/10.1007/978-3-031-62846-7_58

interactivity and the ability to communicate that makes the computer a musical instrument, just like the mechanical dynamic coupling of traditional acoustic instruments with the musician, as described by O'Modhrain and Gillespie [10].

Over the last decades new technologies expanded the interaction with the computer and introduced modalities like haptic feedback [10], actuated surfaces [14] or tangible interaction [6]. In order to overcome barriers and to improve access to computers as musical instruments, these interaction modalities promise to offer improved accessibility, new musical expressions and enhanced digital music work-flows for VIB individuals and musicians.

2 Related Work

Several attempts have already been made towards the artistic use of computers for persons with disabilities by exploring new interaction concepts and developing external hardware interfaces. Amongst the various projects that have been realized is Zach Lieberman's *Eyewriter*, a device that enabled a former graffiti artist diagnosed with amyotrophic lateral sclerosis (ALS) to paint again. Similarly, the *EyeHarp* project enables persons with ALS to learn, perform and compose music using only their gaze as control mechanism [16]. Others explored the possibilities of tangible interaction to empower spatial thinking [2] or using tangible music blocks for teaching programming for children [13].

In the field of computer music and audio production by VIB individuals several attempts have been made to address accessibility barriers by exploring novel interfaces, haptic feedback and tangible interaction. Sile O'Modhrain and Brent Gillespie presented *The Moose*, an interface for VIB individuals to improve the use of computers as digital sound studios by providing a mouse, that responds with haptic feedback when moving over graphical content, allowing for more intuitive computer use [5].

Atau Tanaka and Adam Parkinson developed the *Haptic Wave*, a device, that enables the translation of digital audio information into the physical domain and in return allows the manipulation of the source material [15]. Aaron Karp and Bryan Pardo focused on the interactive tactile display of software equalizers with their project *HaptEQ* [8], similarly Jakub Pesek et al. developed the *TouchEQ*, an eyes-free audio equalizer for a surface haptic interface [12].

A growing number of accessible musical interfaces are developed under the umbrella term Accessible Digital Musical Instrument (ADMI), that Emma Frid defines as *"accessible musical control interfaces used in electronic music, inclusive music practice and music therapy settings"* [4]. Frid's systematic review of ADMIs unfortunately shows that only 3.6% of 83 analysed ADMIs aim towards VIB individuals [4]. One example is the *LoopBlocks* by Förster and Komesker, which is a tangible musical sequencer, that uses small wooden blocks to create a musical sequence by providing a physical grid of multiple rows and columns [3], similarly to the earlier project *BeatBearing* by Peter Bennett and Sile O'Modhrain [1] or the playful sequencer *Orbita* by Playtronica. These accessible tangible sequencers allow both experienced VIB musicians and beginners to play and experiment with musical compositions, helping to stimulate musical creativity and artistic expression. At the same time, this shows that the number of accessible interactive interfaces that enable and support exploratory musical expression for VIB musicians is still relatively limited, and further research is needed.

3 Research Questions and Methods

To investigate and explore accessible computer music, musical expression and interactive physical interfaces for VIB individuals, this interdisciplinary research project *Tangible Signals* was launched at the department Tangible Music Lab of the University of Arts Linz, under the supervision of the second author and in collaboration with VIB musicians and professionals at the Institute for the Blind in Vienna (BBI), in particular with the third author as musician and blind expert, whereas the first author contributes to the project with his experience both in the design of new musical instruments [7,17] and professional stage experience as musician and performer [20].

With regard to digital inclusion, performative music-making and with a focus on tangible interaction, the question was explored as to what are the benefits of interactive physical representation of sounds and sonic features for VIB musicians in the context of computer music. The project also investigated how tangible interaction and interactive physical representations of sound could enrich artistic stage performance and musical expression.

In relation to the artistic topics of interactive music, composition and stage performance and based on the expertise and artistic practice within the University of Art Linz, artistic research methodologies were applied to answer the aforementioned questions, in particular practice-based research methods. These included collaborations and workshops with VIB pupils and professionals at the BBI Vienna, the development of an accessible web-based music software and the iterative design of three tangible user interfaces, as well as user inquiries and expert feedback. Additionally the developed tools were then used for artistic practice in form of stage-based performances and live concerts. Finally, the collected material was reviewed and evaluated in the form of an interpretative analysis.

4 Practical Work

In order to answer the research questions, the initial research intention was the iterative development and design of a tangible interface to represent and manipulate envelopes, which are an essential element of computer music and describe changes over time for sonic elements like amplitude, frequency or pitch.

4.1 Software and Hardware

To approach the development of this tangible interface a series of workshops was arranged, involving six visually impaired and blind pupils aged 14 to 16 with diverse musical backgrounds and the third author who is also their blind teacher and a musician at the BBI Vienna. The workshops aimed to introduce and practice concepts of sound and computer music in order to inform the design and development of the tangible interface.

However, during the workshops it became apparent that there was a lack of accessible music software suitable for providing VIB beginners with convenient access to computer music [18]. This discovery undermined the research process and inhibited the answering of the research questions, as the development of a tangible interface requires the availability of an appropriate and interactive music software for the digital representation. As a consequence and in order to enable access to computer music for the participants, the simplified web-based music software *Welle* was developed in collaboration with the pupils and the teacher [18]. *Welle* made it possible to explore accessible digital sound generation, to perform rhythmic patterns and to compose, record and save music in a web browser. The software was further used and developed during workshops at the BBI Vienna and at the Austrian Computer Camp (OCC), involving a total of 17 visually impaired and blind pupils aged 10 to 17 and three blind teachers and supervisors.

During this software development process all the participants involved were able to gain experience through the artistic practice with *Welle*. Based on their observation, inspired by his own artistic experience and in order to expand the range of tangible interfaces and thus support the research project, the author identified two more sonic features to be developed as part of the research project. This resulted in a total of three proposals for tangible interfaces, which were then initiated and developed in an iterative design process involving regular user inquiries, VIB expert feedback and artistic experimentation [19]. The resulting interfaces are the *Tangible Pins*, *Tangible Wheel* and *Tangible String* (see Fig. 1). These interfaces each represent individual sonic features of digital computer music and were attached to the *Welle* software via MIDI messages and provided access to pattern sequencing, continuous amplitude control and envelope settings. Their motorized design allowed direct manipulation and interactive parameter setting [19].

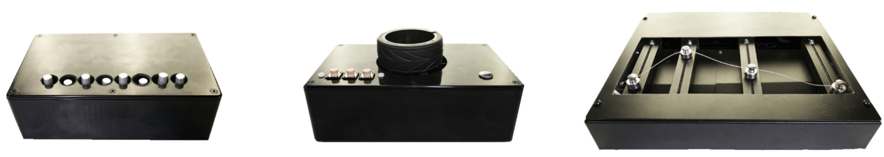


Fig. 1. Images of three tangible interfaces: Tangible Pins, Tangible Wheel and Tangible Envelope

4.2 Usage and Exploration

The interfaces and their use in a musical setting was then explored by four VIB experts and three VIB pupils mainly during the Austrian Computer Camp at the BBI Vienna in 2023, who subsequently reflected on their impressions in the form of unstructured interviews. In order to gain insights into the use of the tangible interfaces in a professional musical context, a musical composition and stage performance with the devices was developed by the third author and performed at the same event in front of an audience of around 40 people. For the performance, instead of using *Welle*, a custom-made music software was provided, that suited the artistic and compositional needs of the performer. The music piece was artistically centered around the third authors echolocation practice, including sound recordings and prepared sample loops. The composition was arranged for the stage by making use of the three tangible interfaces in conjunction with improvisation techniques, spatial distribution in a multi-channel speaker setup and generative echo effects.

5 Expert Feedback and Observations

During the workshops and presentations, the participants were asked to share their impressions regarding the computer as a musical instrument, the tangible interfaces and their respective interaction modalities for making music, furthermore they were asked about the benefits of the interactive interfaces for representing sonic features and finally possible future use cases. Out of a total of 20 VIB pupils and five VIB professionals that participated in the entire project, four VIB professionals and three VIB pupils gave feedback in the form of unstructured interviews.

The expert feedback can be structured into four topics, namely *music and inclusion*, *sonic interaction*, *observations of artistic practice* and *suggestions for future use cases*. In summary, the participants confirmed the exclusion of VIB musicians from mainstream music platforms, including many external MIDI controllers, due to the lack of support for screen readers and the focus on visual representation. With regard to the research project and the developed hardware, they emphasize that the presented tangible interfaces are helpful for the music creation especially due to their motorization and thus their use as input and output devices. Both the *Tangible Pins* as an interface for discrete value input and output and the *Tangible Wheel* with its ability to change values continuously were found to be useful for musical interaction and performance, especially on the stage. The *Tangible String* interface was identified as the most novel and exciting interface to explore sounds and musical effects, based on its interaction with four control handles connected by a string, of which the two middle handles can be moved freely in a two-dimensional arrangement. The entire setup as such, consisting of the three tangible interfaces and the computer with the corresponding software, was perceived as versatile constellation of motorized, tactile and interactive means to practice computer music intuitively and expressively, even if it required concentration to coordinate this amount of devices. Compared to working with music software using a computer keyboard, "touching" the sequences and envelopes leads to a more vivid musical experience, which "feels like joy" compared to adjusting discrete values in the software via the keyboard.

Suggestions that were made for future applications include the use of the *Tangible Pins* interface as a standalone trigger sequencer for modular electronic synthesizers. Accordingly, the *Tangible Wheel* could be used for scrubbing through recordings in audio editors or DAWs, setting markers and editing the waveform. One participant mentioned that it could also be used to emulate a theremin. Another suggestion was to use the *Tangible Envelope* to set a waveform for wavetable synthesis or to use it as a parametric equalizer with additional buttons for zooming and scrolling. Overall, the wish was expressed to combine all three devices into one device with the display of the individual parameters in Braille, which could then serve as a multi-functional composition and performance device in an inclusive music studio.

6 Results

Despite limitations of the research project, such as a small number of professional VIB participants and limited access to the BBI Vienna for workshops due to the Covid-19 pandemic, conclusions and insights can be drawn from the practical work with regard to the research questions.

The research shows that actuated tangible interfaces can provide benefits for VIB individuals in computer music, especially the interactive physical representation of musical data and sonic features, which offers new ways of integrating the computer as a musical instrument into composition and performance practice. It demonstrates that the continuous physical control of digital data enables direct manipulation of musical parameters, which is common for sighted computer users, e.g. by the use of a computer mouse as a pointer within the graphical user interface (GUI), but remains challenging for VIB individuals using mainly Braille displays, screen-readers and keyboard-based interaction. The research also shows that interaction with multiple tangible interfaces simultaneously can increase the bandwidth of interactions with the computer by enabling access to multiple sonic features at the same time, which was found to be valuable and inspiring. It shows that tangible interaction can enhance the immediacy and connection of VIB musicians towards the computer as a musical instrument and that it encourages musical exploration when playing with the physical representations of sonic elements, leading to a more playful experience and unexpected yet inspiring artistic results.

7 Conclusion

This paper summarizes the research project *Tangible Signals* and its focus on tangible interaction in the context of musical expression and accessible computer music. It describes the practical work including the development of an accessible music software and three hardware devices, as well as their use in workshops, expert feedback sessions and artistic stage-based performances. It shows that tangible interaction can offer new opportunities regarding musical expression for VIB musicians and that the interactive physical access to digital sound parameters encourages experimentation and artistic exploration. However, access to computers for musical purposes for VIB individuals remains limited, mainly due to the deficits of music software manufacturers in terms of screen reader support and accessibility. External hardware devices can only compensate for this to a limited extent, and screen readers and the Braille display remain the most important tools. Nevertheless, or rather because of this, research on assistive technologies and interaction modalities in the context of computer music is essential to improve the accessibility of the computer as a musical instrument and as a sonic environment for an inspiring, inclusive and artistically stimulating music practice.

Acknowledgments. Many thanks to the Institute for the Blind Vienna (BBI) and the pupils who participated in the workshops and presentations. Also sincere thanks to the participants of the user inquiries and feedback rounds, especially Ben Hofer, Martin Mayrhofer, Mario Lang and Angela Videbis. This research is funded by a DOC Fellowship of the Austrian Academy of Sciences at the Tangible Music Lab at the Institute of Media Studies, University of Arts Linz, Austria.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Bennett, P., O'Modhrain, S.: The BeatBearing: a tangible rhythm sequencer. In: Proceedings of NordiCHI 2008 (2008)
2. Brock, A.: Tangible Interaction for Visually Impaired People: why and how. In: World Haptics Conference - Workshop on Haptic Interfaces for Accessibility. Fuerstenfeldbruck, Germany (2017)
3. Förster, A., Komesker, M.: LoopBlocks: design and preliminary evaluation of an accessible tangible musical step sequencer. In: Proceedings of the International Conference on New Interfaces for Musical Expression (2021). https://doi.org/10.21428/92FBEB44.F45E1CAF
4. Frid, E.: Accessible digital musical instruments–a review of musical interfaces in inclusive music practice. Multimodal Technol. Interact. **3**(3) (2019). https://doi.org/10.3390/mti3030057
5. Gillespie, R.B., O'Modhrain, S.: The moose: a haptic user interface for blind persons with application to the digital sound studio. In: Third WWW6 Conference (1997)
6. Jordá, S., Geiger, G., Alonso, M., Kaltenbrunner, M.: The reacTable : exploring the synergy between live music performance and tabletop tangible interfaces. In: TEI 2007 Proceedings of the 1st International Conference on Tangible and Embedded Interaction, pp. 139–146 (2007). https://doi.org/10.1145/1226969.1226998

7. Kaltenbrunner, M., Vetter, J.: Tquencer: a tangible musical sequencer using overlays. In: TEI 2018, pp. 429–434 (2018). https://doi.org/10.1145/3173225.3173233
8. Karp, A., Pardo, B.: HaptEQ: a collaborative tool for visually impaired audio producers. In: AM 2017: Proceedings of the 12th International Audio Mostly Conference on Augmented and Participatory Sound and Music Experiences. ACM Press, New York, USA (2017). https://doi.org/10.1145/3123514.3123531
9. Mountapmbeme, A., Okafor, O., Ludi, S.: Addressing accessibility barriers in programming for people with visual impairments: a literature review. ACM Trans. Accessible Comput. **15**(1) (2022). https://doi.org/10.1145/3507469
10. O'Modhrain, S., Gillespie, R.B.: Once more, with feeling: revisiting the role of touch in performer-instrument interaction. In: Papetti, S., Saitis, C. (eds.) Musical Haptics. SSTHS, pp. 11–27. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-58316-7_2
11. Payne, W.C., Xu, A.Y., Ahmed, F., Ye, L., Hurst, A.: How blind and visually impaired composers, producers, and songwriters leverage and adapt music technology. In: ASSETS 2020 - 22nd International ACM SIGACCESS Conference on Computers and Accessibility (2020). https://doi.org/10.1145/3373625.3417002
12. Pesek, J., De Man, B.: touchEQ: an eyes-free audio equalizer for a surface haptic interface. In: 150th Convention of the Audio Engineering Society (2021)
13. Sabuncuoglu, A.: Tangible music programming blocks for visually impaired children. In: TEI 2020 - Proceedings of the 14th International Conference on Tangible, Embedded, and Embodied Interaction (2020). https://doi.org/10.1145/3374920.3374939
14. Siu, A.F., Gonzalez, E.J., Yuan, S., Ginsberg, J.B., Follmer, S.: Shapeshift: 2D spatial manipulation and self-actuation of tabletop shape displays for tangible and haptic interaction. In: Conference on Human Factors in Computing Systems - Proceedings (2018). https://doi.org/10.1145/3173574.3173865
15. Tanaka, A., Parkinson, A.: Haptic wave: a cross-modal interface for visually impaired audio producers. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI 2016 (2016). https://doi.org/10.1145/2858036.2858304
16. Vamvakousis, Z., Ramirez, R.: The EyeHarp: a gaze-controlled digital musical instrument. Front. Psychol. **7** (2016). https://doi.org/10.3389/FPSYG.2016.00906/BIBTEX
17. Vetter, J.: Netz 2.0: towards site-specific performative topologies. In: xCoAx 2019 - Proceedings of the Seventh Conference on Computation, Communication, Aesthetics & X (2019)
18. Vetter, J.: WELLE - a web-based music environment for the blind. In: NIME - Proceedings of the International Conference on New Interfaces for Musical Expression (2020). https://doi.org/10.5281/zenodo.4813208
19. Vetter, J.: Tangible signals - prototyping interactive physical sound displays. In: TEI 2021 - Proceedings of the 15th International Conference on Tangible, Embedded, and Embodied Interaction (2021). https://doi.org/10.1145/3430524.3442450
20. Vetter, J., Leimcke, S.: Homo restis - constructive control through modular string topologies. In: Proceedings of the International Conference on New Interfaces for Musical Expression (2017). https://doi.org/10.5281/ZENODO.1176179
21. Zhang, L., Sun, S., Findlater, L.: Understanding digital content creation needs of blind and low vision people. The 25th International ACM SIGACCESS Conference on Computers and Accessibility (2023). https://doi.org/10.1145/3597638.3608387

Analyzing the Accuracy of Braille Finger Tracking with Ultraleap's Virtual Reality Hand Tracking System

Michael Treml[1,2](✉), Wolfgang L. Zagler[2], and Mike Busboom[2]

[1] Institute of Visual Computing and Human-Centered Technology, TU Wien, Vienna, Austria
michael.treml@tetragon.at
[2] TETRAGON Braille Systems GmbH, Vienna, Austria

Abstract. Tracking the fingers of users of refreshable braille displays can be useful for studying braille reading habits and for user-centric features, such as in braille learning tools. Ultraleap offers a low-cost system for hand tracking in virtual reality use cases, which could be utilized for braille finger tracking if its accuracy is sufficient. In an experimental setup using a 20-cell refreshable braille display, the x-coordinates of two test persons' fingertips were repeatedly captured on four distinct braille cells. The largest range of values measured for one cell was 0.01 Ultraleap units (~10 mm), and two instances of overlapping values between neighboring cells were observed in the study, indicating that the accuracy is insufficient to capture a single braille cell in the user's center of attention. However, the system could still be utilized for use cases like reading out the word that is currently touched as the system was off by only one braille cell in the worst case.

Keywords: Braille · Finger Tracking · Accuracy · Ultraleap

1 Introduction

Technical solutions for tracking the fingers of braille readers have been in use for over 100 years [1, p. 28], but most of them require specific laboratory settings that do not resemble real-life braille reading experiences. In this paper, the authors analyze the potential of a virtual-reality hand tracking system from Ultraleap [2] that could allow for braille finger tracking with only a minimum of such inconvenience.

In previous work (still unpublished), the authors found that the system achieves the highest reliability of finger detection on a 20-cell refreshable braille display when the display surface contrasts with the fingers' color (white) in infrared light (the system utilizes a stereo infrared camera), when the camera is positioned above the display, and when the tracking mode in the included software is set to "head mounted". A setup like this is flexible, unobtrusive, cost-effective and easy to replicate by using an off-the-shelf neckholder mount for smartphones to position the camera (see Fig. 1).

The same setup is utilized for the work presented in this paper. While the previous work focused on whether fingers in such a context can be detected at all, this paper

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 497–504, 2024.
https://doi.org/10.1007/978-3-031-62846-7_59

answers the question of whether the system provides sufficient accuracy to reliably detect the braille cell currently in the center of the reader's tactile perception.

The long-term goal of this study is to inform the design of braille learning software.

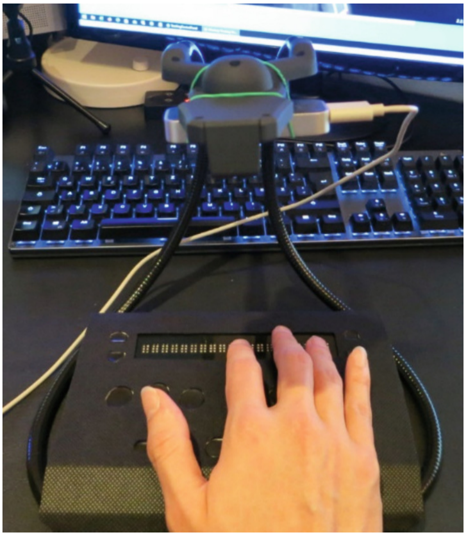


Fig. 1. Ultraleap's stereo infrared camera is positioned above a refreshable braille display using a smartphone neckholder mount.

2 State of the Art

Accuracy is crucial for braille finger tracking due to the small dimensions of braille. For instance, the German standard DIN 32976 [3] recommends a dot diameter of approximately 1.5 mm, a distance from one dot's center to its neighboring dot's center within the same cell of 2.7 mm, and a distance from the first dot's center in a cell to the first dot's center in its right neighbor of 6.6 mm. Similar specifications (~1.45 mm diameter, ~2.34 mm distance within the same cell and ~6.22 mm distance from cell to cell) are given by the Library of Congress in the USA [4].

Finger tracking with low accuracy is not necessarily useless, but it limits the use cases for which it can be utilized. A common application is the study of braille reading patterns, for example researching which fingers are used for reading, how (if at all) both hands are used, and whether movements are linear or scrubbing [5, 6]. Refreshable braille displays can also implement user-centric features such as audio output of a character or word that is currently touched [7] and it has been suggested that such features could be used for braille learning [8]. Not all of these use cases require the same level of tracking accuracy.

The first known device to capture the finger movements of braille readers was, therefore, a simple pen that was attached to a person's reading finger [1, p. 28]. Almost

100 years later, researchers investigated a similar approach using the pen of a digital graphics tablet [8, 9]. Other researchers attached light-emitting diodes (LEDs) to the fingers of readers to provide distinct signals for automatic processing of camera footage [10, 11]. One approach aimed at being cost-effective and easy to use involved using Nintendo's Wii Remote game controller as a camera to track light emitting diodes on index fingers [11].

While all of these solutions require readers to attach something to their fingers, there have also been approaches without this inconvenience. On the downside, these require very specific laboratory settings. Several researchers have used camera footage from underneath a transparent surface with transparent braille sheets [6, 12, 13]. Another concept made use of "transparent braille paper" on the touchscreen of an Android tablet [14]. While such approaches are effective for braille on paper or similar media, they cannot be used to track fingers on refreshable braille displays.

The current solution that comes closest to a natural setting is Active Tactile Control (ATC), a technology that utilizes the piezoelectric elements that set braille dots in refreshable braille displays to also act as touch sensors [8, 15]. However, this technology is only available in few refreshable braille displays because it is patented [15]. Additionally, an instruction manual also mentions a possibly unreliable detection rate for readers who only use a light touch [7], and the concept does not allow for finger detection on blank spaces where no braille dots are set.

Ultraleap's [2] virtual-reality hand tracking system is based on an infrared stereo camera and could allow for a natural setting without requiring a minimum force of touch. The system's accuracy has already been scientifically tested in various settings other than braille reading, yielding very different results depending on the setting [16–19]. When tracking a pen on an industrial robot instead of a human finger, a position accuracy of 0.2 mm was achieved for static setups [16], but in a more realistic setting where actual human fingers were tracked when pointing at targets, only an accuracy of 17.3 mm was achieved [17]. However, it was observed that accuracy decreases as the distance between the hand and sensor increases [18, 19]. This higher accuracy in close proximity could be advantageous for braille finger tracking setups, where mounting the camera directly to the refreshable braille display also results in a more robust and less obtrusive setup.

3 Methodology

3.1 Setup for Experiments

The sensor used in the experiment was Ultraleap's Leap Motion Controller, with following specifications according to its data sheet [20]: 2 near-infrared cameras (850 $\pm$ 25 nm spectral range) with a resolution of 640 pixels $\times$ 140 pixels each, spaced 40 mm apart; 140° $\times$ 120° typical field of view; depth up to 60 cm preferred, up to 80 cm maximum; typically operating at 120 Hz; 3 LEDs on each side and between the cameras, baffled to prevent overlap; connection via USB 2.0; dimensions: 80 mm $\times$ 30 mm $\times$ 11.3 mm; weight: 32 g. The sensor was operated on Microsoft Windows 10 with Ultraleap's tracking software in version 5.5.3, and it was positioned 14 cm above a braille display by using a neckholder mount (originally intended as a smartphone mount to be worn around a person's neck, see Fig. 1). The braille device used in the setup was Innovision's Braille Me, a refreshable braille display with 20 six-dot cells and a six-key Perkins-style keyboard.

3.2 Accuracy Measuring

To measure the accuracy of finger detection, a simple data logging software was created using Ultraleap's plugin for the Unity game engine [21]. This software displays the current coordinates of the detected index fingers' fingertips on the computer screen, and when one of the keys from 1 to 4 on the keyboard is pressed, these coordinates are written to a log file along with the number of the key pressed. The keys represent four distinct positions on the refreshable braille display:

- Key 1 has to be pressed every time the reading finger (index finger) of the tracked hand is supposed to be on the very first braille cell of the line.
- Key 2 corresponds to the second braille cell.
- Key 3 corresponds to the second to last braille cell.
- Key 4 corresponds to the last braille cell.

Capturing the first and the last character was chosen as part of the experiment to ensure coverage of the full length of the display (max. distance). Positions 2 and 3 were included to test two sets of adjacent cells (1–2, 3–4) (min. cell distance).

The refreshable braille display was set up to show a line of full cells (all dots set) to ensure clear tactile detectability of each individual cell. In this setup, the reading finger of a test person (sighted, with some beginner's experience in tactile braille reading) was moved sequentially to the following positions: 1, 4, 1, 4, 2, 3, 2, 3, 1, 2, 3, 4, 2, 1, 4, 3. At each position, the coordinates were logged. The specific sequence was designed to ensure that each position was accessed exactly four times, and so that each possible transition between two position appears at least once. The first four positions consist of two repeated movements (1 to 4 and 2 to 3) to make potentially large discrepancies easily visible right from the start of the experiment. The sequence was tested in three variations: left-handed, right-handed and with both hands. As the previous hand detection experiments had shown minor weaknesses to detect a hand consistently, it was decided to proceed with processing collected data if only one out of the 16 positions could not be captured, but to repeat the iteration if more values are missing.

Although all three coordinates were logged, only the x-axis values were further analyzed, as this dimension determines the position of each braille cell in relation to the others. The mean x positions for each cell and the minimum and maximum measured for each position were analyzed to identify any overlaps between neighboring cells.

To ensure that the results do not apply only to one single person's specific hands and movements in a particular laboratory setting, the experiment was repeated with a second person (a blind braille expert) in a different location. As the initial iteration with the first person used a stationary computer, the software had to be transferred to a laptop first. To ensure that there were no potential differences resulting from the hardware change, the accuracy test was repeated once more with the original test person in the original location for the left and the right hand before conducting the test at the new location. Unlike the first test person, the second person was not instructed to use his left, his right, or both hands, but to move just as he normally does when reading braille in order to obtain results that are not biased by forced conditions that may not occur in real-life situations.

Both test participants were members of the research team. Therefore, no ethics committee approval or informed consent was necessary. The purpose of this study was to analyze the technical performance of the hand tracking system when looking at a braille display. User involvement was technically necessary to provide actual hands to track, but as the system is already established and tested for other hand tracking use cases, the involvement of more and external participants was not considered crucial for this particular study.

4 Results

Table 1. X axis metrics for each position in Ultraleap units (overlapping values for neighboring cells in bold).

Position		Test Person #1 (Beginner)					Test Person #2 (Expert)
		PC			Laptop		Laptop
		Left Hand	Right Hand	Both Hand	Left Hand	Right Hand	Right Hand
1	Mean	−0,061	−0,049	−0,053	−0,045	−0,074	−0,074
	Min	−0,064	−0,052	−0,058	−0,047	−0,076	−0,076
	Max	−0,058	**−0,045**	−0,048	−0,044	−0,072	−0,072
	Range	0,006	0,007	0,010	0,003	0,004	0,004
2	Mean	−0,052	−0,043	−0,045	−0,039	−0,067	−0,067
	Min	−0,055	**−0,045**	−0,046	−0,040	−0,069	−0,069
	Max	−0,050	−0,040	−0,045	−0,037	−0,066	−0,066
	Range	0,005	0,005	0,001	0,003	0,003	0,003
3	Mean	0,060	0,058	0,058	0,069	0,041	0,041
	Min	0,058	0,055	0,054	0,066	0,040	0,040
	Max	0,062	0,060	0,060	**0,071**	0,042	0,042
	Range	0,004	0,005	0,006	0,005	0,002	0,002
4	Mean	0,067	0,066	0,065	0,075	0,051	0,051
	Min	0,065	0,066	0,061	**0,070**	0,047	0,047
	Max	0,068	0,066	0,069	0,078	0,056	0,056
	Range	0,003	0,000	0,008	0,008	0,009	0,009

Table 1 displays characteristics of the x-axis values measured for each of the 4 positions on the refreshable braille display. The table includes mean, minimum, maximum, and distance between minimum and maximum (= range). The largest range measured was 0.01 units given by the Ultraleap plugin. Two instances were found where the values of a cell overlapped with those of its neighboring cell. In the first case, the maximum value measured for the first cell was the same as the minimum value measured for the cell on its right. In the second case, the maximum value of the left cell was 0.001 Ultraleap units higher than the minimum value of the right cell.

Table 2. Mean distances between positions in Ultraleap units.

Position Span	Test Person #1 (Beginner)					Test Person #2 (Expert)
	PC			Laptop		Laptop
	Left Hand	Right Hand	Both Hand	Left Hand	Right Hand	Right Hand
1–2	0,009	0,006	0,008	0,007	0,005	0,006
3–4	0,007	0,008	0,007	0,005	0,005	0,010
1–4	0,128	0,115	0,118	0,120	0,116	0,124

Table 2 displays the distances between adjacent cells as well as between the first and the last cell, calculated by determining the difference between their mean position values. The mean distance between neighboring cells ranged from 0.005 to 0.01 Ultraleap units, while the distance from the first to the last cell ranged from 0.115 to 0.128 Ultraleap units. The actual physical distance between the first and the last cell measured on the refreshable display was 12.35 cm, indicating that 0.01 Ultraleap units correspond to approximately 1 cm in real life.

Observing the x values of a static fingertip on the monitor output indicates a static noise of no more than 0.001 Ultraleap units.

5 Discussion

The accuracy of the system is not sufficient to reliably detect the single braille cell that is currently supposed to be in the center of the user's tactile perception. This is demonstrated by the largest range of values measured for one cell being 0.01 Ultraleap units (~10 mm) and the smallest mean distance from one cell to its neighbor being 0.005 Ultraleap units (~5 mm). Additionally, two concrete examples of overlaps occurring in this study further support this conclusion. As the system already provides very steady values with a fluctuation of only 0.001 Ultraleap units (~1 mm), smoothing the values over time would not significantly increase accuracy.

However, the accuracy in this braille reading setting clearly outperformed the accuracy of 17.3 mm achieved in a pointing-at-a-screen task [17]. This is likely due to the camera being close to the fingers. The measured values suggest that the system will usually not be off by more than a single braille cell.

It will depend on the use case whether this accuracy is acceptable. While it is clearly not precise enough to research micro-movements within a single braille cell, it is at least sufficient to identify the word currently being read by the user, which can be a useful feature for braille learning software. It should also be kept in mind that fingers are physically wider than braille cells, allowing them to touch two neighboring cells at once. Therefore, identifying a wider area instead of just a single cell may often provide a more accurate depiction of reality.

Since the camera has a very wide field of view (140° × 120°) that reaches far beyond the edges of the 20-cell braille display used in this study, it could also cover larger braille displays from the same vertical viewing distance. However, a separate study would be required to analyze if fingers are still sufficiently tracked with the bigger horizontal distance to cover in such settings. As the long-term goal of this study was to find a cost-effective tracking solution for braille beginners, larger displays were not within the research scope since they are typically considerably more expensive.

6 Conclusion

The accuracy of the measurements (~10 mm in the worst case) is not sufficient to reliably detect a single braille cell that is currently in the user's focus of touch perception. However, it can be expected that the system will only deviate by a single cell at worst, which is adequate for use cases like detecting the currently touched word.

The study was focused on the general suitability for the specific use case of tracking fingers on a refreshable braille display. User involvement was limited to two test persons since more elaborate studies for accuracy in other use cases had already been conducted. Including more test persons in future braille display studies may provide more precise data, but more precise numbers would not alter the main conclusion of this study.

All of these results were achieved using an Ultraleap Leap Motion Controller positioned 14 cm above a 20-cell refreshable braille display. The impact of other distances and braille display sizes was not analyzed in this study, but existing literature suggests that accuracy would decrease as distance increases.

System speed was not analyzed in this study because braille beginners, who are the main target group for this work, are typically slow readers. However, speed may be important in other contexts, especially when working with braille experts, and may then need to be researched in a separate study.

Disclosure of Interests. TETRAGON Braille Systems GmbH is a TU Wien spin-off company and was founded by the authors with the aim of developing innovative braille technology. However, the academic work presented in this paper is not directly tied to any of the work done at TETRAGON.

References

1. Bürklen, K.: Touch-reading of the Blind (in German). Verlag von Johann Ambrosius Barth, Leipzig (1917). ISBN: 978-1173243111

2. Ultraleap. https://www.ultraleap.com/. Accessed 03 Jan 2024
3. German Institute for Standardization. DIN 32976: Braille - Requirements and Dimensions. Berlin, Germany (2007)
4. Library of Congress: Specification 800:2014. Braille Book and Pamphlets. National Library Service for the Blind and Physically Handicapped (2014)
5. Foulke, E.: Investigative approaches to the study of braille reading. J. Visual Impairment Blindness **73**(8), 298–308 (1979)
6. Davidson, P.W., Valenti, D., Reuter, S., Wiles-Kettenmann, M., Haber, R.N., Appelle, S.: Some functions of reading by Hand. In: Annual Meeting of the American Psychological Association. Public Health Service (DHEW), Rockville (1980)
7. Handy Tech Elektronik GmbH: Instruction Manual Active Braille (in German). Horb, Germany (2011)
8. Kipke, S.: Sensitive Braille Displays with ATC Technology (Active Tactile Control) as a Tool for Learning Braille. In: Miesenberger, K., Klaus, J., Zagler, W., Karshmer, A. (eds.) ICCHP 2008. LNCS, vol. 5105, pp. 843–850. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-70540-6_125
9. Hughes, B.: Movement kinematics of the Braille-reading finger. J. Visual Impairm. Blindness **105**(6), 370–381 (2011). https://doi.org/10.1177/0145482X1110500608
10. Hughes, B., Van Gemmert, A.W.A., Stelmach, G.E.: Linguistic and perceptual-motor contributions to the kinematic properties of the braille reading finger. Hum. Mov. Sci. **30**(4), 711–730 (2011). https://doi.org/10.1016/j.humov.2010.05.005
11. Noblet, A., Ridelaire, H., Sylin, G.: Equipment for the study in braille of operating processes reading. Behav. Res. Methods Instrum. Comput. **17**, 107–113 (1985). https://doi.org/10.3758/BF03200904
12. Aranyanak, I., Reilly, R.G.: A system for tracking braille readers using a Wii remote and a refreshable braille display. Behav. Res. Methods **45**, 216–228 (2012). https://doi.org/10.3758/s13428-012-0235-8
13. Millar, S: Reading by Touch. Routledge, London (1997). ISBN: 978-0415068376
14. Breidegard, B.: Computer-based automatic finger- and speech-tracking system. Behav. Res. Methods **39**, 824–834 (2007). https://doi.org/10.3758/BF03192975
15. Aranyanak, I.: A finger-tracking system for studying braille reading on a tablet. In: 2018 3rd International Conference on Computer and Communication Systems (ICCCS), Nagoya, pp. 109–112 (2018). https://doi.org/10.1109/CCOMS.2018.8463350
16. Kipke, S.: Patent DE 10 2004 046 526 B4 (2007)
17. Weichert, F., Bachmann, D., Rudak, B., Fisseler, D.: Analysis of the accuracy and robustness of the leap motion controller. Sensors **13**(5), 6380–6393 (2013). https://doi.org/10.3390/s130506380
18. Tung, J.Y., Lulic, T., Gonzalez, D.A., Tran, J., Dickerson, C.R., Roy, E.A.: Evaluation of a portable markerless finger position capture device: accuracy of the Leap Motion controller in healthy adults. Physiol. Meas. **36**, 1025–1035 (2015). https://doi.org/10.1088/0967-3334/36/5/1025
19. Guna, J., Jakus, G., Pogačnik, M., Tomažič, S., Sodnik, J.: An analysis of the precision and reliability of the leap motion sensor and its suitability for static and dynamic tracking. Sensors **14**(2), 3702–3720 (2014). https://doi.org/10.3390/s140203702
20. Valentini, P.P., Pezzuti, E.: Accuracy in fingertip tracking using leap motion controller for interactive virtual applications. Int. J. Interact. Design Manuf. (IJIDeM) **11**, 641–650 (2014). https://doi.org/10.1007/s12008-016-0339-y
21. Ultrahaptics Ltd. UH-003206-TC Issue 6 Leap Motion Controller Data Sheet. (no date available)
22. Unity. https://unity.com/. Accessed 03 Jan 2024

Remote Collaboration Between Sighted and Non-Sighted People in Mixed-Ability Teams

Francesca Kaschnig[2](✉), Thomas Neumayr[1], and Mirjam Augstein[1]

[1] Research Group Personalized Environments and Collaborative Systems, University of Applied Sciences Upper Austria, Softwarepark 11, 4232 Hagenberg, Austria
{thomas.neumayr,mirjam.augstein}@fh-hagenberg.at

[2] Degree Program Human-Centered Computing, University of Applied Sciences Upper Austria, Softwarepark 11, 4232 Hagenberg, Austria
francesca.kaschnig@outlook.com
https://peec.fh-hagenberg.at

Abstract. Remote work settings became increasingly important in recent years, which often inherently require collaboration at a distance. Hereby, especially in mixed-ability teams, different individuals might face different challenges. In this paper, we thus aim to analyze remote collaborative interaction between sighted and non-sighted users based on a qualitative study with twelve participants (six non-sighted) in three mixed-ability teams, who worked together on a problem solving task. Our analysis considers a variety of aspects related to the perception and closeness of collaboration, assistance and communication among team members, workspace awareness, territoriality, and triggers for problematic situations, and indicates both, challenges but also potentials of collaboration within mixed-ability teams.

Keywords: remote collaboration · mixed ability teams · mixed visual ability teams · synchronous remote collaboration

1 Introduction

According to the World Health Organization (WHO), at least 2.2 billion people have a near or distance vision impairment [25] and the World Blind Union (WBU) estimates that there are 253 million people who are blind or partially sighted [24]. Many of these individuals rely on assistive tools (e.g., screen readers, Braille displays) which enable them to access technology and content [5,23]. At the same time, recent studies show that remote work, which especially gained importance during the COVID-19 pandemic, remains a work setting with rapidly increasing popularity (e.g., [2,15]). Usually, working remotely inherently includes collaboration at distance, which is not only generally challenging, e.g., related to boundary management or lack of social interaction [1], but even more so for mixed-ability teams. Potluri et al. e.g., mention a considerable coordination overhead for blind or visually impaired team members, especially in tightly coupled

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 505–516, 2024.
https://doi.org/10.1007/978-3-031-62846-7_60

synchronous collaboration [16]. They further point out that for blind or visually impaired participants, existing tools and work practices might "imped[e] their abilities to collaborate effectively, compromising their agency, and limiting their contribution" [16]. In this paper, we thus further investigate potentials and challenges in synchronous remote collaboration for mixed-ability teams, based on a qualitative user study with twelve participants in three groups, each involving sighted and non-sighted participants, who collaborated on a problem solving task. Our results, besides general findings regarding the interaction process itself, reveal positive as well as negative aspects, related to, e.g., assistance among team members, but also barriers in the form of unexpected accessibility issues.

2 Related Work

In this section we describe related research, focusing on *remote collaboration* for *blind people*, and on *collaboration* between *blind* and *sighted people* in general.

2.1 Remote Collaboration and Blind People

Recent years have seen a boost of remote collaboration, leading to a drastic increase in global usage of video conferencing tools such as Microsoft Teams or Zoom. E.g., Microsoft reported a drastic increase of meeting minutes spent on Teams in the first hot phase of the COVID-19 pandemic[1]. While the following months and years brought a cool-down related to the pandemic, the trend to videoconferencing remained [21]. The transition to online work and collaboration might lead to a greater level of inclusion for some individuals, e.g., with mobility impairment [7], but impose additional challenges for others, including visually impaired and blind people. E.g., Doush et al. report limited keyboard accessibility and rare provision of alternative audio information [7] and further refer to Hersh et al.'s 2020 accessibility evaluation of video conferencing tools, which found none of the evaluated tools to be fully accessible [9]. In 2021, Leporini et al. provided another analysis of accessibility in video conferencing tools for blind users [13], focusing on Zoom, MS Teams and Google Meets. Their evaluation comprised a survey with blind users and an inspection by accessibility experts. It considered different roles (host and participant) and typical tasks (e.g., joining a meeting, hand raising, turning the microphone on and off, or accessing shared content). In summary, the authors found that while the tools' main functions were accessible via keyboard with screen reader (with Zoom being best accessible in summary), access was not always easy (e.g., due to a lack of shortcuts or "other problems resulting from poor design for screen reader feedback" [13]). The latter point related to feedback is relevant not only for the users' interaction with the system but also for awareness of the status of other users, according to Buzzi et al. who discuss access to a collaborative editor (Google Docs) via screen

[1] https://www.microsoft.com/en-us/microsoft-365/blog/2020/04/09/remote-work-trend-report-meetings/, last access: 2024-01-23.

readers and (back in 2010) came to the conclusion that several main functions were "practically inacessible" via keyboard [4]. Further, Lee et al. more recently discussed an approach to accessible awareness in collaborative document editing, which is often part of remote collaboration and usually challenging for blind people although most commercial tools have screen reader support [12]. Part of these challenges is related to a lack of awareness (as, e.g., provided by visual markup for sighted users). Lee et al. in their research suggest six design considerations including provision of "easy navigation", contextual information, on-demand and automatic updates, and support of common collaboration activities (e.g., information about collaborators or text changes) [12]. These are relevant within the scope of collaborative text editors but also beyond, e.g., during navigation in shared document repositories as in our study (see Sect. 3).

2.2 Collaboration Between Blind and Sighted People

In addition to challenges that generally arise for blind people working together with others remotely, there might be further issues in collaboration among people with different perceptual abilities [19]. Sánchez et al. argue that it is necessary to establish a "common ground" for people perceiving a collaboration environment in different ways [19]. While this might be based on the WISIWYS paradigm in synchronous collaboration of sighted people, an alternative is required in groups with blind members. Tanhua-Piiroinen et al. state that collaboration "depends on an unrestrained dialogue and on the fact that people working together can maintain a common ground of the context of joint action including the features of the workspace and the continuous status of the work process" [22], and argue that a shared understanding cannot be easily maintained if the collaboration environment does not allow mutual access to information or perception of others' activities [22]. Sánchez et al. thus introduce an environment offering different, e.g., sound-based, interfaces [18,19]. Kuber et al. argue that blind and sighted users "develop different mental representations of their respective interfaces" [11] which might cause problems in communication and collaboration. They suggest a non-visual approach based on haptic technology for blind members of mixed-ability teams collaborating on web navigation tasks, and found in a study that blind and sighted users could successfully communicate relative positions and directions, which was not always achieved with screen readers. Another haptic approach to interaction of blind and sighted people is presented in [17].

Examples for collaboration practices between sighted and blind people in a non-virtual context are discussed by Branham & Kane [3], who explain collaboration at the workplace, or Yuan et al. [26], who describe collaborative shopping practices. Yuan et al. identify three sources of "common ground contributing to successful collaboration" that might also apply for virtual collaboration contexts: knowledge about how to assist a visually impaired person, knowledge about common experience and interpersonal relationship history, and domain knowledge [26]. Branham & Kane found that often times blind participants hesitated to ask their sighted colleagues for help with accessibility problems, and that blind participants often lacked awareness of resources in the work environment [3].

3 Study

To gain in-depth information about the nature of remote collaboration between sighted individuals (SIs) and non-sighted individuals (NIs), we conducted a qualitative user study. Three mixed-ability teams (cf. Sect. 3.1) engaged in two-hour problem solving sessions. The setup included video-conferencing, cloud storage, browser, document editing tools, and assistive technology (for NIs). Before the study, task and setup (cf. Sect. 3.2) were tested for accessibility with the help of an expert of the Austrian Association for the Visually Impaired and Blind[2]. Additionally, the task was pretested (with several SIs unrelated to the study) and the setup was tested with all participants in a preparation meeting before the study. This latter preparation phase led to a change in the task setup concerning the chosen collaboration toolset (cf. Section 3.2).

3.1 Participants

Twelve participants (9m and 3f, aged between 23 and 64) including six NIs were recruited via associations of visually impaired people and our university, see Table 1. While no special knowledge or skills were required for participation, we presumed the availability of an own computer, an internet connection, and, for NIs, the use of assistive technology based on screen readers: one NI used NVDA[3], the others JAWS[4] (a part additionally used a Braille display). One of the NI has a residue of sight, but is considered blind by legal means. The participants were randomly assigned to three teams, composed by two NIs and two SIs, respectively.

3.2 Task and Setup

We created a custom problem solving task inspired by [8,10] in which a fictional character addressed the team in a letter, asking for help in destroying Dracula. This character disappeared and left behind **accessible** and **inaccessible** information in the form of videos, images, texts, geolocations, links to websites, and PDFs that helped the team develop a strategy. The information was grouped and made accessible only via passwords, which had to be found by solving riddles (e.g., converting a Morse code audio into natural text using a table). In a strategy paper to be elaborated jointly, the teams had to determine Dracula's strengths and weaknesses, the perfect time and location to defeat him, mechanisms to protect oneself, and a detailed defeat plan. The setup included OneDrive as a cloud storage, Zoom for video conferencing, and Browser, Word (after accessibility problems with Google Docs had been identified during preparation), or Excel, for document viewing and editing (web and desktop versions).

---

[2] https://www.blindenverband.at/en, last access 2024-04-02.

[3] https://www.nvaccess.org/, last access 2024-04-02.

[4] https://www.freedomscientific.com/products/software/jaws/, last access 2024-04-02.

Table 1. Study participants, their backgrounds and team assignments.

Team	ID	Visual Ability	Screen Reder Usage	Remote Collaboration Tool Usage	Home Office Working Habits
Team 1	P02_SI	Sighted	-	Very often	Exclusively
	P05_SI	Sighted	-	Very often	Regularly
	P09_NI	Non-sighted	JAWS	Now and then	Very often
	P11_NI	Non-sighted	NVDA	Often	Never
Team 2	P01_SI	Sighted	-	Very often	Very often
	P03_SI	Sighted	-	Very often	Exclusively
	P08_NI	Non-sighted	JAWS	Often	Exclusively
	P10_NI	Non-sighted	JAWS	Now and then	Never
Team 3	P04_SI	Sighted	-	Very often	Often
	P06_SI	Sighted	-	Very often	Exclusively
	P07_NI	Non-sighted	JAWS	Never	Never
	P12_NI	Non-sighted	JAWS	Often	Exclusively

3.3 Data Sources and Analysis

During the study, an **observer** was present to take notes describing the respective team's work. After collaboration, a 45-minute, semi-structured **group interview** was conducted via Zoom, where participants shared their own impressions and experience on collaboration. These interviews were recorded to facilitate analysis (through transcription and grouping of statements). Additional data was collected via a **questionnaire**. Different online questionnaire tools were tested for accessibility before SoSci Survey[5] was identified as a suitable tool. Further, we analyzed participants' screen **recordings** (including system audio and microphone input) for generally interesting observations rounding out the questionnaire and interview responses to the division of labor, or the perception of collaboration. Additionally, recordings were coded considering usage of territories and assistance provided among participants.

4 Results

In this section, we report on the results of our original user study.

4.1 Perception of Collaboration

Participants generally stated that the collaboration was enjoyable, although especially the NIs pointed out that the amount of information provided was too extensive. The NIs of Team 1 stated that this contributed to a difficulty in collaborating, as they had to invest a lot of time in reading the documents. In

[5] https://www.soscisurvey.de/, last access 2024-04-02.

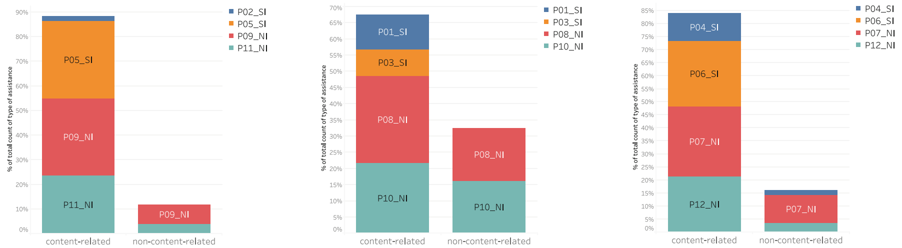
addition, one of them claimed that the task would not have been solvable without the help of SIs (cf. Sect. 4.3), as information was simply not accessible, which was partly intended by the design of the study. Some NIs felt overwhelmed and under pressure concerning their reading speed, although the SIs did not notice any difference in pace. In Team 3, one SI e.g., mentioned that they did not have the impression that the non-sighted colleagues needed more time to read through documents. All SIs of the study stated that they found the collaboration very enriching. None of them had experienced working with NIs before, and many were unaware of the barriers they have to face or the way NIs work.

4.2 Division of Labor

When approaching their tasks, all teams followed a similar path of first individual overviews, before discussing and deciding how to proceed. For most people, it became clear by the third riddle that task division was necessary to reach the goal together. They determined who was going to work on which task and was responsible for which information resource by how accessible it seemed. An illustrative example was looking up coordinates in the provided tool, where one NI pointed out that map services should be taken over by an SI, as they are not easy to use for NIs. Another example was translating Morse code into natural text, where two NIs repeated heard Morse code for the SIs so that they could look up the correct letter in an inaccessible translation table. Expected reading speed was another criterion for dividing the tasks. Some NIs reported a feeling of being rather slow in reading the provided information. This may also be due to the amount of information that was only accessible for NIs via audio. Interestingly, this was not perceived in this way by SIs, and yet it partly influenced the allocation of tasks. In one team (Team3), on the other hand, division of labor was not always related to accessibility or anticipated reading speed. In fact, even as it became clear that a table was poorly accessible, a NI was observed to navigate this table until they got an approximate overview. While all teams completed the task, elaborating a strategy within the given time, only Team 3 solved all the riddles.

4.3 Assistance Among Participants

From our recorded material, we coded instances of content-related and -unrelated assistance whenever a participant requested help from others (see Fig. 1). In total, 52 assistance requests were coded for Team 1, 37 for Team 2, and 56 for Team 3. The results of the questionnaire additionally showed that the participants who asked for or provided help felt that this was part of teamwork. Help was asked for and openly and willingly accepted. Further, we found awareness (see below) to play an important role in the context of assistance: it is necessary to know what the person to be assisted is doing to provide adequate support.



(a) Assistance in Team 1. (b) Assistance in Team 2. (c) Assistance in Team 3.

Fig. 1. Assistance requests by the individual members of our mixed-ability teams.

4.4 Awareness

During the group interviews, we asked whether it was always clear what the respective other team members were doing. All participants agreed that it was *not* clear in the beginning but became clearer as time progressed. This increase in workspace awareness was mainly attributed to the always-on Zoom audio connection. Especially Team 3 pointed out that workspace awareness was achieved by the fact that individual team members e.g., read small text sections aloud, thought aloud while working on them, or acoustically took responsibility for a certain task or resource. Interestingly, screen reader output which was picked up by the microphone was identified to take the function of an indirect awareness mechanism. At least one screen reader per team could be constantly or partly perceived by all team members and therefore provided awareness information.

Explicit awareness mechanisms in our tool set were partly asymmetrical. Word Web at least provided such mechanisms for SIs (e.g., remote cursors), and NIs were informed via screen reader when someone was currently revising a paragraph or when they "stumbled" over another user with the screen reader. Within OneDrive (directories, images, PDFs and videos), no explicit awareness mechanisms could be observed. Although a version history indicated which documents had been edited by whom, no real-time information, e.g., showed colleagues' location in the shared directory. This could also be a reason why participants were not always aware of their colleagues' activities and further indicates a lack of synchronous awareness mechanisms for both, NIs and SIs.

Related to the assistance requests discussed in Sect. 4.3, our data shows that the majority noticed when someone needed help. For example, P08_NI of Team 2 frequently assisted their non-sighted colleague in accessibility issues and was aware of their colleague needing help without them explicitly asking for it.

4.5 Territoriality

Territoriality according to Scott et al. [20] was previously shown to be helpful for the analysis of collaborative interaction (e.g., learning about the visibility of and shared access to certain workspace zones [14]). In our study, we coded

(a) Territoriality in Team 1. (b) Territoriality in Team 2. (c) Territoriality in Team 3.

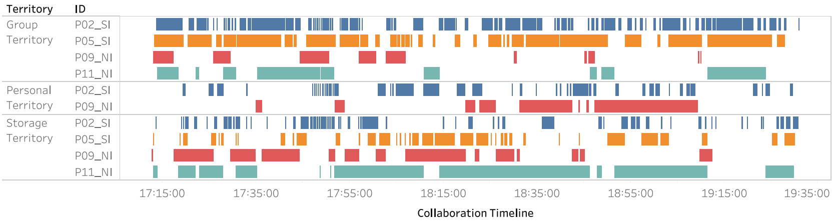
Fig. 2. Territoriality within our teams. The first column shows how much time was spent in the GTs, the second in the PTs, and the third in the STs.

participants' interaction location according to the three territory types defined in [20]: **Personal Territory (PT)**, **Group Territory (GT)**, and **Storage Territory (ST)**. In cases where several territories were visible on screen (mostly by SIs), we coded only the one that was currently interacted in. Downloaded files, participants' notes and pages opened in the course of internet research were considered **PT**. Documents that could be edited by several people at the same time, e.g., Word Web and Excel files, were considered **GT**. Finally, the OneDrive directory and information provided in form of images, links and videos were considered **ST**, since participants could not modify them. Regarding the permanence in territories (cf. Fig. 2), there was a trend in Teams 1 and 2 that SIs seemed to spend most of their time in GTs while NIs relied more on PTs and STs. This was however not true for Team 3, where it was particularly striking that all the territories seemed to be used to the same extent by all its members.

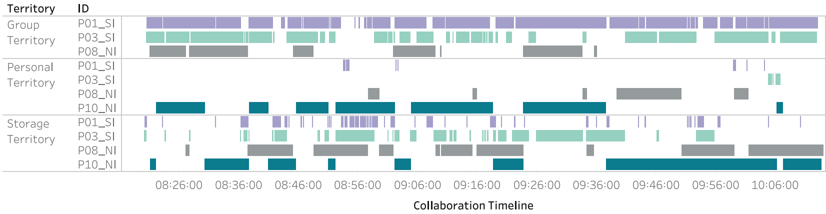
Our analysis showed that territories had different functions for NIs and SIs, and that territories existed simultaneously and were not concretely separable. It additionally indicated that a PT can have a stress-relieving effect. A NI, for example, claimed they felt overwhelmed and could cope with stress by switching to their own PT. Further, we found that responsibility for the content of the ST, in this certain setting, was not influenced by proximity ([20]) but rather by accessibility. During the analysis, we noticed that transitions between different territories can occur quickly (see Fig. 3). While most NIs tended to switch less and had longer stays in the respective territories, there was an exception in Team 3 (cf. Fig. 3c) where all members switched quickly between territories.

4.6 Expected and Unexpected Accessibility Issues

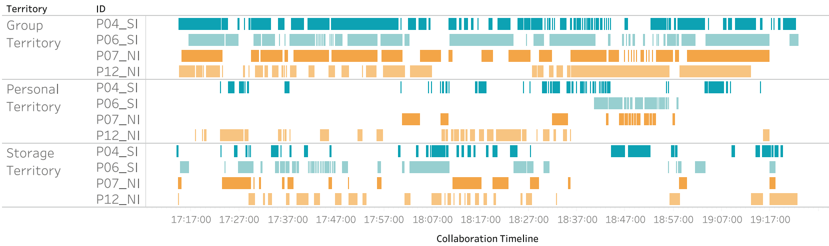
Intended by study design, some visual data was inaccessible to NIs (due to lack of alternative text). Partly, this material was jointly interpreted in close teamwork, partly, SIs took over the interpretation, and there was one instance where a NI used an online tool to make sense of a Morse code (instead of the provided table).



(a) Changes in territoriality in Team 1.



(b) Changes in territoriality in Team 2.



(c) Changes in territoriality in Team 3.

Fig. 3. Changes in use of territories throughout the collaboration per user.

Additionally, we observed numerous unexpected accessibility issues. A table designed to be accessible, proved to be tedious to navigate with screen readers, and when navigating in OneDrive with a screen reader, some shortcuts did not work, making it necessary to go over each and every item to arrive at the target. Further, there were ambiguities on which operation was carried out with files (e.g., select, download, or open a file) and it was sometimes not apparent if a new item would open in the same or a new tab resulting in accidental closing of tabs. While a possible explanation for these accessibility issues would be lack of experience on the users' side, most of our participants had worked with the provided tools before (a majority had used OneDrive, only Word Web was rarely used). The problems can thus not be exclusively attributed to lack of experience. Interestingly, some technical issues were only observed for part of the NIs, hinting at a high dependence on the employed ecology of tools and assistive technology.

5 Discussion and Conclusion

In this paper, we described a qualitative study exploring remote collaboration between SIs and NIs in mixed-ability teams. With this research, we intended to gain a deeper understanding of the nature of collaboration in this setting, but also to identify potentials as well as challenges and fault lines. The results of the study actually indicate both, positive and negative aspects in collaboration.

On one hand, NIs had to face numerous unexpected accessibility problems (in addition to intentionally included non-accessible material), lack of awareness and navigation difficulties. Additionally, it showed to be hard for NIs to gain and retain an overview of the task-related resources. This might be due to the fact that information and navigation was reduced to the auditory channel, and in some cases unfamiliar key combinations had to be used for navigation, increasing learning efforts for NIs. On the other hand, our observations indicate that a team may have a supportive role in collaboration, as e.g., non-accessible visual clues could be described for the NIs by SIs, and NIs could help each other in circumventing accessibility issues. This could suggest that non-sighted people are well able to support each other in collaboration, because they know best where difficulties exist and how they can be overcome.

Regarding the general collaboration process in mixed-ability teams, our findings indicate that screen readers are only partially perceived as disruptive (extending and emphasizing results of previous research [3]). The NIs in our study, e.g., confirmed that in remote settings, listening to different auditory inputs simultaneously is demanding, but also, that remote settings may allow listening to a screen reader loudly, without worrying about disturbing (co-located) colleagues (cf. [3]). In future research, other forms of assistive technology could thus be investigated in more detail within this setting. E.g., using Braille displays might impact the collaboration experience differently, since no multiple auditory inputs need to be processed simultaneously, but instead different senses (touch and hearing) are engaged. Further, our findings suggest that direct (verbal) communication played an important role in the collaboration process, as it turned out to be crucial for assistance and awareness. Alternative communication channels (e.g., chat or comment functions) were not used by any of the three teams, although they were generally available; it became apparent already during the preparation phase that, e.g., the comment function in Microsoft Word Web is difficult to access by NIs (also see [6]). Thus, such alternative communication channels could be dealt with in a more targeted way in future research.

We summarize that in remote collaboration of teams with blind and sighted members, NIs face several challenges, which may not only prevent light-hearted interaction but also influence mood and overall perception of collaboration. Still, we have seen many instances where appropriate assistive technology and an inclusive mindset helped shape an environment of equal participation and inclusivity that can be a great chance for future remote collaboration in similar settings.

Acknowledgements. This paper is based on F. Kaschnig's Master Thesis. T. Neumayr's and M. Augstein's work on the paper has been conducted as part of the *Hybrid Collaboration Spaces (HYCOS)* project funded by the Austrian Science Fund FWF [P 34928].

References

1. Augstein, M., Neumayr, T., Schönböck, J., Kovacs, C.: Remote persons are closer than they appear: home, team and a lockdown. In: Proceedings of CHI'23. ACM, Hamburg, Germany (2023). https://doi.org/10.1145/3544548.3580989
2. Bamieh, O., Ziegler, L.: Are remote work options the new standard? Evidence from vacancy postings during the Covid-19 crisis. Labour Econom. **76** (2022)
3. Branham, S., Kane, S.: The invisible work of accessibility: how blind employees manage accessibility in mixed-ability workplaces. In: Proceedings of ASSETS '15, pp. 163–171. ACM, Lisbon, Portugal (2015)
4. Buzzi, M.C., Buzzi, M., Leporini, B., Mori, G., Penichet, V.M.R.: Accessing Google Docs via Screen Reader. In: Miesenberger, K., Klaus, J., Zagler, W., Karshmer, A. (eds.) Computers Helping People with Special Needs, pp. 92–99. Springer Berlin Heidelberg, Berlin, Heidelberg (2010). https://doi.org/10.1007/978-3-642-14097-6_17
5. Curiel, J.: It accessibility for the visually impaired. https://www.gbnews.ch/accessibility-for-all/
6. Das, M., Gergle, D., Piper, A.M.: "it doesn't win you friends": understanding accessibility in collaborative writing for people with vision impairments. Proc. ACM Hum.-Comput. Interact. (2019-11)
7. Doush, I.A., Al-Jarrah, A., Alajarmeh, N., Alnfiai, M.: Learning features and accessibility limitations of video conferencing applications: are people with visual impairment left behind. Univ. Access Inform. Society **22**, 1353–1368 (2023). https://doi.org/10.1007/s10209-022-00917-4
8. Grinstein, G., O'Connell, T., Laskowski, S., Plaisant, C., Scholtz, J., Whiting, M.: Contest - A Tale of Alderwood, pp. 215–216
9. Hersh, M., Leporini, B., Buzzi, M.: Accessibility evaluation of video conferencing tools to support disabled people in distance teaching, meetings and other activities. In: Petz, A., Miesenberger, K. (eds.) ICCHP Opern Access Compendium Future Perspectives of AT, eAccessibility amd eInclusion (2020)
10. Johnson, D., Johnson, F.: Joining Together Group Theory and Group Skills. Pearson Education Limited, 11 edn
11. Kuber, R., Yu, W., McAllister, G.: A non-visual approach to improving collaboration between blind and sighted internet users. In: Stephanidis, C. (ed.) UAHCI 2007. LNCS, vol. 4556, pp. 913–922. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-73283-9_99
12. Lee, C.Y.P., Zhang, Z., Herskovitz, J., Seo, J., Guo, A.: Collabally: accessible collaboration awareness in document editing. In: CHI '22: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems. ACM, New Orleans, LA, USA (2022). https://doi.org/10.1145/3491102.3517635
13. Leporini, B., Buzzi, M., Hersh, M.: Distance meetings during the Covid-19 pandemic: are video conferencing tools accessible for blind people? In: W4A '21: Proceedings of the 18th International Web for All Conference. ACM, Ljubljana, Slovenia (2021). https://doi.org/10.1145/3430263.3452433

14. Neumayr, T., Augstein, M., Kubicek, B.: Territoriality in hybrid collaboration. Proc. ACM Human-Comput. Interact. **6**(CSCW2) (2022). https://doi.org/10.1145/3555224
15. Pew Research Center: Covid-19 pandemic continues to reshape work in america, https://www.pewresearch.org/social-trends/2022/02/16/covid-19-pandemic-continues-to-reshape-work-in-america/
16. Potluri, V., Pandey, M., Begel, A., Barnett, M., Reitherman, S.: Codewalk: facilitating shared awareness in mixed-ability collaborative software development. In: ASSETS '22: Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility. ACM, Athens, Greece (2022). https://doi.org/10.1145/3517428.3544812
17. Sallnas, E.L., Moll, J., Severinson-Eklundh, K.: Group work about geometrical concepts among blind and sighted pupils using haptic interfaces. In: Second Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (WHC'07), pp. 330–335. IEEE, Tsukuba, Japan (2007). https://doi.org/10.1109/WHC.2007.61
18. Sánchez, J.: Audiobattleship: blind learners cognition through sound. Int. J. Disability Human Develop. **4**(4), 303–309 (2005). https://doi.org/10.1515/IJDHD.2005.4.4.303
19. Sánchez, J., Baloian, N., Hassler, T.: Blind to sighted children interaction through collaborative environments. In: de Vreede, G.-J., Guerrero, L.A., Marín Raventós, G. (eds.) CRIWG 2004. LNCS, vol. 3198, pp. 192–205. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-30112-7_16
20. Scott, S.D., Carpendale, M.S.T., Inkpen, K.: Territoriality in collaborative tabletop workspaces. In: Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work, pp. 294–303. CSCW '04, Association for Computing Machinery, New York, NY, USA (2004). https://doi.org/10.1145/1031607.1031655
21. Statista: Zoom video communications meetings worldwide from 2nd quarter 2020 to 3rd quarter 2021. https://www.statista.com/statistics/1322970/zoom-meetings-global/. Accessed Dec 4 2023
22. Tanhua-Piiroinen, E., Pasto, V., Raisamo, R., Sallnäs, E.-L.: Supporting collaboration between visually impaired and sighted children in a multimodal learning environment. In: Pirhonen, A., Brewster, S. (eds.) HAID 2008. LNCS, vol. 5270, pp. 11–20. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-87883-4_2
23. Winberg, F., Bowers, J.: Assembling the senses: towards the design of cooperative interfaces for visually impaired users. In: Proceedings of the CSCW'04, pp. 332–341. CSCW '04, Association for Computing Machinery, New York, NY, U. https://doi.org/10.1145/1031607.1031662
24. World Blind Union: World blind union. https://worldblindunion.org/. Accessed Dec 04 2023
25. World Health Organization: Blindness and visual impairment. https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment. Accessed Dec 04 2023
26. Yuan, C.W., Hanrahan, B.V., Lee, S., Rosson, M.B., Carroll, J.M.: I didn't know that you knew i knew: collaborative shopping practices between people with visual impairment and people with vision. Proc. ACM Human-Comput. Interact. **1**(CSCW) (2017). https://doi.org/10.1145/3134753

Expiration Date Recognition System Using Spatial Transformer Network for Visually Impaired

Yoshinori Takeuchi(✉) and Kengo Suzuki

Daido University, 10-3 Takiharu-cho Minami-ku, Nagoya, Japan
ytake@daido-it.ac.jp

Abstract. In this paper, we propose a system for recognizing an expiration date of a beverge package. A visually impaired user takes a photo of the product. Then, the system automatically recognizes the expiration date of the product and tells it to the user. The dates in the image were sometimes skewed, misaligned, or partially missing. To solve these problems, we use Spatial Transformer Network (STN) to recognize skewed, misaligned, or partially missing date images. STN explicitly allows the spatial manipulation of data within the neural network. We propose a dedicated CNN with STN expiration date recognition. The input to this network is an image of the expiration date. The network has six STNs to detect each digit of the date. Each STN outputs the rectified image of the digits. The image of the digits is then recognized by the traditional CNN. We trained the expiration date recognition network. The network converged quickly, and the training was stopped in 10 epochs. We tested this network on 1088 date images that were not used during training. The experimental results showed that the system was able to achieve a high recognition rate of 99.42% for the dataset without dates with spaces and 98.44% for the dataset with dates with spaces.

Keywords: Character Recognition · Expiration Date · Spatial Transformer Network

1 Introduction

People who are visually impaired can experience difficulties in many aspects of daily life. Among them is shopping, which requires visual information. According to the results of an interview with 14 visually impaired individuals about shopping [1], they can go to the front of the product shelves by themselves, because they remember the arrangement of the shelves in the supermarket. However, they are unable to obtain the name, price, or expiration date of perishable foods. They express a desire to enjoy shopping independently without the assistance of others. In this paper, we develop a system that provides expiration date information to the visually impaired.

Seker *et al.* created and published a dataset of 1,767 images and 12,000 date samples including synthesized images to train 13 forms of expiration date

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 517–524, 2024.
https://doi.org/10.1007/978-3-031-62846-7_61

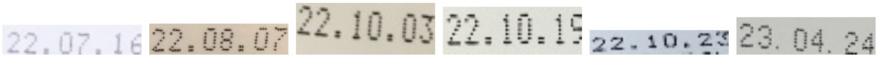


Fig. 1. example of the expiration date image written as dot matrix characters. Some of them are skewed, misaligned, or partially missing

display formats [2]. In an experiment using the dataset, the recognition rate was 97.74%. Zheng *et al.* proposed a character detection algorithm using DBNet (a segmentation-based text detection method) as the base network, combined with the Convolutional Block Attention Module (CBAM) to improve its feature extraction of characters in complex contexts. They collected images of expiry dates for 162 food items, a total of 2,161 images and tuned the model to be better applied to expiry date recognition. They also adopt affine transformation to address the skewed text lines that appear during printing characters. In their experiment, the character detection accuracy reaches 97.9% on Jetson Nano implementation.

Expiration dates are often written as dot-matrix characters. Such characters are represented by a set of dots, as shown in Fig. 1. To recognize them, Ashino and Takeuchi proposed an automatic recognition system using combination of deep neural networks [4]. These works used object detection techniques to detect the date region from an input image. When we examined the images of the detected dates, we found that the dates were skewed, misaligned, or partially missing. These images make date recognition difficult. To solve these problems, a spatially invariant character recognition method is required. Jaderberg *et al.* proposed the Spatial Transformer Network (STN), which explicitly allows the spatial manipulation of data within the neural network [5].

In this paper, we propose a system for recognizing an expiration date of perishable food products. We use STN to recognize skewed, misaligned, or partially missing date images. We propose a dedicated expiration date recognition neural network with STN.

2 Expiration Date Recognition System

2.1 Overview

We propose a system to recognize the expiration date of the products and output it as audio. We assume that the expiration date is written in YY.MM.DD or YY MM DD format. Figure 2 shows the flow of the proposed system. First, the system detects the expiration date region from the input image. If the region is detected, then the system recognizes the expiration date from the image. Finally, the system tells the expiration date to the user.

2.2 Detection of Expiration Date Region

The expiration date region is detected by a conventional object detection technique. We use Faster R-CNN ResNet50 V1 640×640 [6] as the date detection

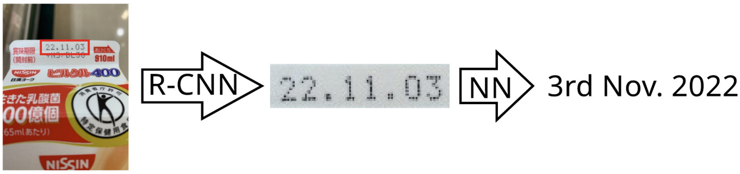


Fig. 2. Flow of proposed system

Table 1. Detailed implementation of the localisation network

layer	no. of filters/units	kernel size	activation
Conv2D	8	7×7	relu
MaxPooling		2×2	
Conv2D	10	5×5	relu
MaxPooling		2×2	
Dense	32		relu
Dense	6		linear

R-CNN. The input image size is resized to 300×400 pixels. We manually labeled the product images containing the expiration date to enclose its region. Then we trained the R-CNN by using the labeled images. The trained R-CNN is able to detect the expiration date region in the image.

2.3 Recognition of Expiration Date

The system recognizes the expiration date from the detected expiration date image. We propose a dedicated CNN with STN to recognize the expiration date, as shown in Fig. 3. First, the image is input to six STNs. Each STN crops the image of the region corresponding to each digit of the expiration date. Then, the cropped images are recognized by traditional CNNs.

The input to this network is a 128×32 pixel image of the expiration date. The network has six STNs to detect each digit of the date. The STN consists of localisation network and grid sampling. The localisation network estimates the parameter of the spatial transformation from the input image. Grid sampling samples the pixel from a transformed input image using the estimated parameters. In this way, the STN crops the image region of the corresponding digits from the input image. The localisation network can correct image distortions to make the image suitable for later image recognition.

Table 1 shows the detailed implementation of the localisation network. Note that the last layer is initializeed all zeros and adds different bias to the output. The outputs of six units are used as the parameters of the spatial transformation. We use affine transformation as the spatial transformation.

All six STNs are identical except for the output bias. It is set to crop the corresponding position of the expiration date image, i.e. only the horizontal transformation differs. Each STN outputs the rectified image of the digits with a size of 16×32 pixels. This digit image is then recognized by the traditional CNN.

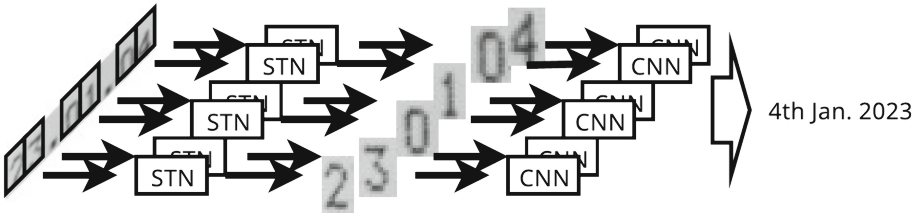


Fig. 3. Proposed expiration date recognition neural network

Table 2. Detailed implementation of the recognition CNN

layer	no. of filters/units	kernel size	activation
Conv2D	32	3×3	relu
Conv2D	32	3×3	relu
BatchNormaiization			
MaxPooling		2×2	
DropOut			
Conv2D	32	3×3	relu
Conv2D	32	3×3	relu
BatchNormaiization			
MaxPooling		2×2	
DropOut			
Flatten			
Dense	255		relu
Dense	10		softmax

The advantages of this method are as follows:

1. The STN can rectify the distortion of the image.
2. The output of the STN is a cropped image. Therefore, a conventional CNN can be used for digit recognition, and the training time can be reduced by pre-training the CNNs.
3. This method can output the whole date, not only each digit of the date as in [4].

Table 2 shows the detailed implementation of the recognition CNN. The output of this network is the probability of the corresponding number. Combine the outputs of the six CNNs to obtain the final expiration date.

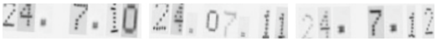


Fig. 4. example of synthesized training images. The images are translated and rotated randomly.

3 Experimental Results

3.1 Training

Detection Neural Network. In the detection part, we trained the R-CNN using images of the beverage package including the expiration date. We manually labeled the expiration date region for a total of 1,301 product images. We used 1,170 images for training and 131 images for validation. The training batch size was one and the number of steps was 25,000.

Recognition Neural Network. In the recognition part, we cropped out 100 images for each number from 0 to 9. These images were used to pre-train the CNNs in our method. We used 80 images for training, 10 for validation and 10 for testing. The batch size was 100 and we trained 10,000 epochs. The loss function was categorical cross-entropy and the optimizer was Adadelta.

We synthesized the date images to increase the number of training images and to improve the recognition of future dates. We also added random translation and rotation to simulate the skewed, misaligned, or partially missing images. We synthesized 90,000 images for training and 10,000 images for validation. Examples of the synthesized images are shown in Fig. 4. These images were used to train our proposed date recognition network. The batch size was 100 and the training time was 100 epochs. We also used early stopping which stops the training if there is no improvement in the validation loss during 5 epochs. The loss function was the sum of six categorical cross-entropy and the optimizer was Adam.

3.2 Results and Discussion

Detection Neural Network. We tested the R-CNN detection by using 147 images of the beverage package, which is not used in training. 140 out of 147 images were correctly detected, giving a detection rate of 95.2%.

Recognition Neural Network. We pre-trained the CNN in our network and achieved a test result of 98%.

We trained the expiration date recognition network using this pre-trained CNN. The network converged quickly, and the training was stopped in 10 epochs. Figure 5 shows the loss and accuracy of the network. We tested this network on 1034 date images that were not used during training and obtained a test result of 99.42%. All the distorted images in Fig. 1 were correctly recognized. We used the distorted training images. Therefore, this method can archive the high

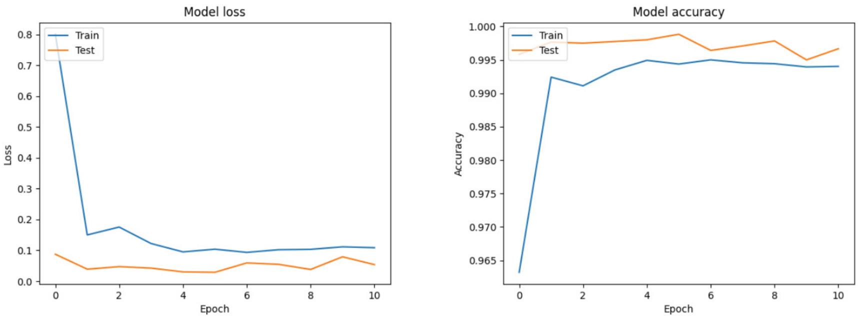
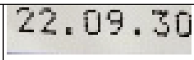
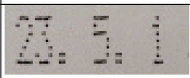
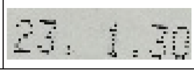


Fig. 5. Training loss and accuracy for date recognition network

Table 3. All the misrecognized images

	221102		221911		220939
	221622		230510		230132

recognition rate for real date images. All the misrecognized images are shown in Table 3. These images can be correctly recognized by adding more training digits to the synthesized training images.

3.3 Adaptation of Other Date Format

In the previous experiment, we excluded date images without spaces, as shown in Fig. 6. Dates in this format account for 5% of all dates. We tested on 54 images with this format, only 10 images were recognized correctly. The overall recognition rate dropped to 95.40%. To overcome this problem, we synthesized the date images without spaces and used them for training. After the training, we tested this network on 1088 date images. The result was 98.44%.

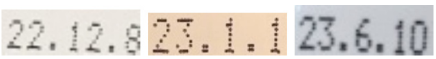


Fig. 6. Date image without spaces

3.4 Comparison with Other Methods

We have compared the performance of the expiration date recognition results from recent studies (Table 4). Note that it is not possible to make a fair comparison due to the different data used. All the results are obtained from the original papers.

Table 4. Quantitative comparison of the expiration date recognition methods.

method	detection		recognition (%)	
	no. of test images	accuracy (%)	no. of test images	accuracy (%)
Ashino *et al.* [4]	30	90	30	97
Florea *et al.* [7]	—	—	60	72.7
Gong *et al.* [8]	727	98.20	727	95.44
Seker *et al.* [2]	665	98.31	665	97.74
Zheng *et al.* [3]	—	90.2	—	98.1
Ours w/o space	147	95.2	1,034	99.42
Ours w space	147	95.2	1,088	98.44

Ashino and Takeuchi [4] collected 138 images of a beverage package and use them for training and test. They used two DNNs to improve the recognition performance. Florea and Rebedea [7] collected approximately 660 photos, taken using a smartphone, which contain expiration date from a diverse range of food products. They used convolutional recurrent neural network (RCNN) to detect and recognize the expiry date in the image. Gong *et al.* [8] used 2,424 images of a variety of food package images from different superstore sites. They used 70% of the images for training and 30% for test. They used FCN and, which were fine-tuned from dealing with texts to detect/recognize use-by date. Seker *et al.* [2] collected 1767 real-world expiration date image by capturing food, beverage, and medicine products, and used them training and test. Their method could handle 13 expiration date formats. Zheng *et al.* [3] collected images of expiry dates for 162 food items, a total of 2,161 images. They did not describe how many images they used for the test.

Our method uses images of a beverage package, such as milk, as a data set. It achieved a high recognition rate. In general, the recognition rate is high only for a specific expiration date format and low in different formats.

4 Conclusion

In this paper, we proposed a system for recognizing the expiration date of perishable food products. The system uses Spatial Transformer Network to correct the distortion of the date image. In the detection part, the experimental result showed that the system was able to detect 95.2% of the expiration date region in the image. In the recognition part, the experimental results showed that the system was able to achieve high recognition rate of 99.42% for the dataset excluding dates without spaces and 98.44% for the dataset including dates without spaces.

As a future work, it is necessary to extend the system to output the expiration date as audio. It is also necessary for visually impaired people to use the system and provide feedback on it. We plan to implement the proposed system in a smartphone application.

References

1. Tanaka, N., Doi, Y., Matsumoto, T., Takeuchi, Y., Kudo, H., Ohnishi, N.: A system helping the blind to get merchandise information. In: Miesenberger, K., Karshmer, A., Penaz, P., Zagler, W. (eds.) ICCHP 2012. LNCS, vol. 7383, pp. 596–598. Springer, Cham (2012). https://doi.org/10.1007/978-3-642-31534-3_87
2. Seker, A.C., Ahn, C.A.: A generalized framework for recognition of expiration on product packages using fully convolutional networks. Expert Syst. Appli. **203**, 117310 (2022) https://doi.org/10.1016/j.eswa.2022.117310
3. Zheng, J., Li, J., Ding, Z., Kong, L., Chen, Q.: Recognition of expiry data on food packages based on improved DBNet. Connection Sci. **35**(1), 1–16 (2023) https://doi.org/10.1080/09540091.2023.2202363
4. Ashino, M., Takeuchi, Y.: Expiry-date recognition system using combination of deep neural networks for visually impaired. In: Miesenberger, K., Manduchi, R., Covarrubias Rodriguez, M., Peňáz, P. (eds.) ICCHP 2020. LNCS, vol. 12376, pp. 510–516. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-58796-3_58
5. Jaderberg, M., Simonyan, K., Zisserman, A., kavukcuoglu, K.: Spatial transformer networks. In: Cortes, C., Lawrence, N.D., Lee, D.D., Sugiyama, M., and Garnett, R. (eds.) Advances in Neural Information Processing Systems 28, pp. 2017–2025. Curran Associates, Inc. (2015)
6. Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks (2015). http://arxiv.org/abs/1506.01497
7. Florea, V., Rebedea, T.: Expiry date recognition using deep neural networks. Inter. J. User-Syst. Interact. **13**(1), 1–17 (2020). https://doi.org/10.37789/ijusi.2020.13.1.1
8. Gong, L., et al.: A novel unified deep neural networks methodology for use by date recognition in retail food package image. Signal, Image Video Process. **15**, 449–457 (2021) https://doi.org/10.1007/s11760-020-01764-7

Correction to: Computers Helping People with Special Needs

Klaus Miesenberger, Petr Peňáz, and Makoto Kobayashi

Correction to:
K. Miesenberger et al. (Eds.): ***Computers Helping People with Special Needs*****, LNCS 14750,**
https://doi.org/10.1007/978-3-031-62846-7

The original version of this book had a typing error in the name of the third editor. "Makato Kobayashi" should have been "Makoto Kobayashi". This has now been corrected.

The updated version of this book can be found at
https://doi.org/10.1007/978-3-031-62846-7

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, p. C1, 2024.
https://doi.org/10.1007/978-3-031-62846-7_62

Author Index

K. Miesenberger et al. (Eds.): ICCHP 2024, LNCS 14750, pp. 525–529, 2024.
https://doi.org/10.1007/978-3-031-62846-7_0

S

T

V

W