SIGACCESS NEWSLETTER

A regular publication of ACM SIGACCESS: Special Interest Group on Accessible Computing

Welcome to the October issue of the ACM SIGACCESS newsletter. This issue highlights three articles that tackle the challenges faced by visually impaired people through a variety of different approaches, ranging from educational games and artificial intelligence to smartglasses applications.

An Inclusive Educational Game Usable via Screen Reader on a Touch-Screen

Barbara Leporini and Eleonora Palmucci

Serious games are increasingly used for supporting education. Sadly, they are not accessible to visually-impaired people.

In the first article, Barbara Leoporini and Eleonora Palmucci present an accessible mobile educational game and investigate issues related to gesture input and enjoyment when accessing information via screen reader.

Teachable Machines for Accessibility

Hernisa Kacorri

How can accessibility research leverage advances in machine learning and artificial intelligence with limited data?

In the second article, Hernisa Kacorri explores a novel approach of empowering end-users to personalize technology by consciously providing training examples and actively interacting with machine learning algorithms to increase their accuracy.

Designing Smartglasses Applications for People with Low Vision

Shiri Azenkot and Yuhang Zhao

While our community has many active projects involving blind people, low vision is rarely addressed.

In the third article, Shiri Azenkot and Yuhang Zhao discuss their ongoing research on designing augmented reality applications for low vision users. Particularly, they provide an overview of three research projects that illustrate their research agenda.

Hugo Nicolau
Newsletter editor

Inside this Issue

About the Newsletter

SIGACCESS is a special interest group of ACM on Accessible Computing. The SIGACCESS Newsletter is a regular online publication of SIGACCESS that includes content of interest to the community. To join SIGACCESS, please visit our website **www.sigaccess.org**

SIGACCESS Officers

Shari Trewin
Chair

Matt Huenerfauth
Vice-Chair

Jinjuan Feng
Secretary Treasurer

Hugo Nicolau
Newsletter Editor

Tiago Guerreiro
Information Director

Want to see your article featured in the next issue?

Contributions and inquiries can be emailed to **editors_sigaccess@sigaccess.acm.org**

We encourage a wide variety of contributions, such as letters to the editor, technical papers, short reports, reviews of papers or products, abstracts, book reviews, conference reports and/or announcements, interesting web page URLs, local activity reports, and so forth. Actually, we solicit almost anything of interest to our readers and community.

You may publish your work here before submitting it elsewhere. We are a very informal forum for sharing ideas with others who have common interests. Anyone interested in editing a special issue on an appropriate topic should contact the editor. As a contributing author, you retain copyright to your article and ACM will refer requests for republication directly to you. By submitting your article, you hereby grant to ACM the following non-exclusive, perpetual, worldwide rights: to publish in print on condition of acceptance by the editor; to digitize and post your article in the electronic version of this publication; to include the article in the ACM Digital Library and in any Digital Library related services; to allow users to make a personal copy of the article for non-commercial, educational or research purposes.

An Inclusive Educational Game Usable via Screen Reader on a Touch-screen

Barbara Leporini and Eleonora Palmucci
ISTI - CNR
barbara.leporini@isti.cnr.it, palmuccieleonora@gmail.com

Abstract

Serious games are increasingly used for supporting education and many other activities via entertainment. Unfortunately, they are not accessible to visually-impaired people who have a very limited selection for games, especially in the mobile context. Usually they can only rely on specific games. Our aim is to investigate how to overcome this gap for visually-impaired people. For this, we are presenting a mobile educational game accessible also via screen reader on a touch-screen. Through the app we investigated: (1) a gesture-based interaction modality to perform exercises on a touch-screen when a screen reader is running; (2) an equal opportunity in enjoying perception also by those who cannot see the user interface. A pilot test confirmed a positive impact of the first prototype on the end-users.

Introduction

A combination of increasing interest in learning through games together with advances in information and communications technology (ICT), have led to the development of a number of digital (computer-based and mobile) educational games. Their use is likely to increase. There is also increasing recognition for the needs of disabled students and the importance of integrating them into mainstream education, as well as the importance of doing this appropriately with adequate resources of support. This makes it imperative to consider the requirements of disabled students (and staff) to ensure their full inclusion at a relatively early stage in the dissemination of educational computer-based games. This raises a number of issues, including the conditions to be met for educational digital games to be accessible to and usable by disabled students, so that the same games and environments can be used by disabled and non-disabled students.

Mobile and desktop applications and games are hence increasingly available on the market for several purposes, including those used in the educational and rehabilitation context. Although several apps are available for users interacting via assistive technologies, an adequate amount of accessible and usable games is not available for people with vision impairment. This occurs especially for blind children and students. Generally speaking, there are few educational games designed for visually-impaired people for both computer and mobile devices. There are many educational smartphone games for children, but few are accessible to blind children. BraillePlay, a suite of accessible games for smartphones that teach Braille character encodings to promote Braille literacy, is an example of specialized games for blind children [1].

In this paper a serious game as an educational tool accessible to visually-impaired users interacting via screen reader on a touch-screen device is presented. The main goal is to fill in the gap existing for visually-impaired people in the entertainment field, especially about the serious games in the mobile context. For this reason, the authors propose a mobile application to be used everywhere and at any time by everyone, including those who cannot see the user interface (UI) and must interact via screen reader.

The designed app is an educational game composed of various activities (e.g., matching questions, single or multiple choice questions, etc.) to be performed on a mobile device. Blind people interact via screen reader and specific gestures which may differ from those commonly used on touch-screens. For this reason, blind users often encounter several issues with the common interaction modalities used for exercises and questions (e.g. drag-and-drop). Thus, through the proposed prototype we intended to investigate: (1) a gestured-based interaction modality on a touch-screen to perform various types of tests which are commonly used for exercises and questions in the education context for practicing / assessing specific topics; (2) an auditory perception equivalent to a visual representation in order to make the game more inclusive for non-sighted and sighted people. In order to assure a good interaction, the interface should be well-designed in an accessible and usable way. In this perspective accessibility guidelines and WAI-Aria techniques have been considered when developing the mobile interface.

Games in the education of blind people

Numerous serious and entertaining games have been proposed in literature. Unfortunately, they are mainly visually-oriented [2] and do not offer alternative perception modalities. Therefore, blind people are excluded from the opportunity to choose any type of game available on the market. Several studies started to investigate how video games requiring different input and output modalities can be designed for people with different disabilities [3, 4]. Thus, various design suggestions and frameworks are proposed [5, 6]. Shoukry et al. [7] propose a framework for mobile educational games, but students with vision impairment are not included.

For visually-impaired people, predominantly auditory games (i.e. the UI is mainly based on the audio feedback) are developed, such as those proposed by [8] and [9], which are not commonly used by everyone. Consequently, the choices available for a blind person are dramatically limited. Song et al. [10] developed two audio-based learning games on TeacherMate, an inexpensive mobile device designed for people in developing countries. Although audio games certainly improve the abilities and special skills (e.g. orientation in the space, etc.), a blind person would prefer to be able to choose among various types of games as much as possible, like a sighted user can. This is a main reason why a universal design should be encouraged and further investigated [11]. Our approach is hence investigating how to design an inclusive game, accessible to screen reader users and touch-screen gestures as well. Our aim is to preserve the graphical interface while assuring accessibility and usability properties with a unique interface.

Our approach

Our approach is aimed at investigating the designing of an educational game suitable for sighted and non-sighted people by (1) overcoming some issues encountered in the interaction via gestures when a screen reader is running and (2) providing a user interface with a similar aural and visual perception. Two blind people (with a long experience of smartphones) were involved from the early phase of prototype design in order to discuss potential issues and multimodal aspects. During brainstorming activities, they described some problems they usually encounter based on their experience on a touch-screen when interacting with games, questions and multimedia contents. Specifically, the discussion was focused on the multimodal interaction related to the following aspects: (1) appropriate method to perform specific tasks (e.g. drag-and-drop) via gestures; (2) the design and perception of the user interface (e.g. audio, labels and messages).

The main issues pointed out by the blind users are especially about the exercises and practices, i.e. Tests and questions. Evaluation activities can be very interactive for the user. When the UI elements are not developed in a standard way, the assistive technology is not able to properly detect them. This can affect negatively the evaluation tests and practicing activities as well. They often are not suitably detected by the screen reader and sometimes item selection is difficult for a blind person, both via keyboard and gestures. This occurs especially when performing drag-and-drop (e.g., for matching) or certain actions such as the ordering or multiple selections.

The proposed prototype is a cross-platform app developed using the Cordova Framework in order to evaluate the accessibility of a User Interface designed for mobile devices. In [12, 13] WAI-Aria suite has been tested with Web-based interfaces. In this study the aim is exploiting the WAI-Aria suite support for interface accessibility on a mobile touch-screen. The prototype has been validated in terms of accessibility by the users involved during the design cycle.

An accessible mobile game prototype

Game description

The game has been graphically and structurally designed, as it is a "solar system" with eight planets. Each planet represents a "play" with a set of exercises / questions. Plays can be chosen from the home (first screen) by tapping on a planet (see Fig. 1). The player (i.e. who needs to practice / assess) must complete each play. According to the correctness and potential errors carried out while performing the game, a score is gained for each play. The plays represent the various types of questions or exercises which are usually used in a test environment. The player can carry out the plays in the preferred order. At the end of each play (and so at the end of the game) the player will have collected a final score which can be used as a level reached.

Based on the issues pointed by the users, to investigate potential interaction solutions we considered some main typologies of techniques used when preparing questionnaires, tests for practicing and evaluating specific educational topics **Single choice** (just one is right), **Multiple choice** (more than one may be right), **Matching choices** (match the choice on the left to the corresponding choice on the right), **True / False, Ordering, Gap-filling** (complete the sentence by filling in). In designing these typologies of exercises, we relied on simple gestures and VoiceOver-like interaction ways with menus (single tap to hear an option and double tap to select it) in line with earlier accessibility work. Color contrast, auditory and visual renderings have been used to assure different levels of interface accessibility.

The app has been designed to adapt the topics for any age. To do this the content is out of the implementation and stored in specific files. Thus several groups of questions and exercises can be prepared earlier by teachers and skilled people. By consequence the app can be suitable for many topics and different ages.

The main features considered for the proposed educational game can be summarized in: (1) Support for learning and practicing specific topics while enjoying; (2) Different typologies of questions and exercises designed for an interaction via gestures and on a touch-screen; (3) Attractive visual layout with different graphical themes (applied to the plays); (4) Audio support to provide an alternative and equivalent perception of the visual layout by a blind player; (5) Content customization in order to adapt the exercises and questions for different topics and ages.

Game architecture

Each play is characterized by three components: (a) template, (b) theme, and (c) content.

(a) Template. The template indicates what type the questions are, i.e., single-choice, multi-choice, matching, ordering, and so on. Different types of questions have been selected in order to investigate how to make them accessible via screen reader and gestures on a touch-screen. Each type of design structure identified to implement the question typologies is assigned to a template. So when using a certain template, the solution implemented is applied to design for that play. So we have templates like "single" or "multiple choice" to refer to the related typology of questions. As described in Play 1 and Play 2 (see next section), the two templates are very similar; there are just some differences in implementing the selection of one or more choices.

(b) Theme. The theme specifies the graphical and audio scheme to use for the UI. Graphical themes have been selected in order to reproduce certain appealing and suggestive scenarios. One of the main features offered by the game consists of reproducing a similar perception of the amusing aspect by a blind and a sighted user. To achieve this, we designed the visual themes so that it was possible applying given evocative sounds to reproduce a similar auditory effect. In the next section some examples are described.

(c) Content. The content is related to the questions to be proposed to the player. The questions and related answers are stored in external files which are loaded when selecting a specific play. This allows the adaptation of the contents with different topics and for various ages. The administrator (e.g. the teacher) can prepare them according to the topics, difficulty levels, and target. The game loads a group of questions from the files. This feature will be improved by defining a more structured database for the content s.

Play examples

As mentioned, each play is a set of questions of the same typology. So all questions and exercises related to a play have the same template and theme. Only the content changes.

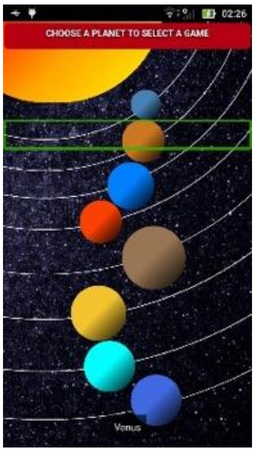

Figure 1: Home

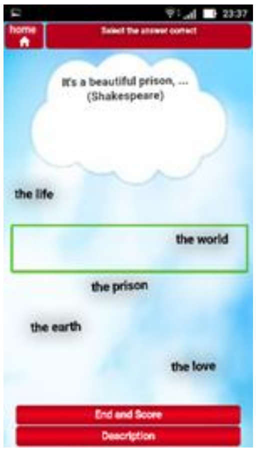

Figure 2: Gap-filling question

Play "Gap-filling".

The play "Gap-filling" consists in a set of questions to be completed (template). Each question is shown in a "cloud" (the theme) as in Figure 2. The sentence contains the ellipsis to indicate the missing term. In order to not make the editing too multifaceted, a list of potential answers is reported rather than asking the user to use the virtual keyboard. The terms are visually arranged in a scattered order, but the blind user can perceive them in a list when exploring the interface with the finger. The visual game perception has been equally audio rendered via sounds: "rain" as background sound, "thunder" for a wrong answer, and "bird singing" for a correct choice. Graphical and audio contents represent the suggestive scenario.

Play "Matching items".

These exercises consist in matching the elements belonging to two different sets (see Fig. 3). The elements are arranged in two lists: one on the left and one on the right. To match the corresponding elements, firstly the user selects by tapping one item on the left, and then on the corresponding one on the opposite side. By consequence, the two items are both shown on the left and marked as paired (no more selectable). When finished, on the right side there is no more elements. The user confirms via the specific buttons to complete the exercise.

The theme is a river. The elements to be paired are listed on the left and on the right of the river (Fig. 3). A rope is launched when the match is carried out (Fig. 4). In order to reproduce a similar effect for a blind player, the river sound and rope launch audio effect are used to reproduce a similar perception.

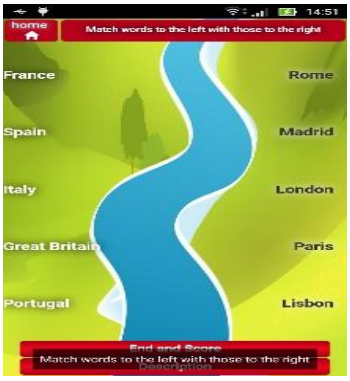

Figure 3: Matching question start

Figure 4: Matching question

EVALUATION AND DISCUSSION

For our first prototype we prepared a set of questions on topical interests we used to have a pilot evaluation with two blind adults (those involved in the design process) and two sighted adults (to verify the graphical and visual layout). Some tasks have been proposed to each user in order to collect feedback and potential issues about the interaction. We observed some problems in detecting the elements by the blind persons due to the object position in the interface and to some contents difficult to identify by gestures and screen reader interaction. The users' comments have been very positive although some issues have been encountered. The tasks were completed by all the users. The non-sighted users were able to interact with the interface via the screen reader and gestures. Minor inconveniences have been noticed with some gestures differing from those offered by VoiceOver. The users appreciated the sounds used to reproduce the visual scenes: they declared to be able to perceive the game scenarios thanks to the aural reproduction. Sighted users interacted with the app without significant issues and completed all the assigned tasks. Overall, both the sighted and visually-impaired users declared to enjoy with the game and to be interested in extending the possible plays with many other additional contents. The suggestions and the aspects observed during this pilot test will allow us to improve the first prototype and extend it with other features.

CONCLUSIONS

This work investigates how to design a potential mobile cross-platform app developed as an educational game for all, including screen reading users. It does not focus specifically on the educational purpose, but rather than on the methodology to design the user interface and the interaction via screen reader and gestures. A combination of audio and graphics allowed to obtain an equally content perception for both sighted and non-sighted people. This can represent a case study in which the graphical layout is not a limitation for assuring full accessibility. Furthermore, WAI-Aria techniques can support accessibility also for mobile interfaces when interacting via screen reader. Through the proposed UI design and interaction modality, questions and exercises can easily be performed also via gestures on a touch-screen by a blind user. Concluding, designing attractive apps for all is possible. Further investigation about interaction modalities for other typologies of questions and a more structured user testing will be conducted.

References

1. Milne, L. R., Bennett, C. L., Ladner, R. E., & Azenkot, S. (2014, October). BraillePlay: educational smartphone games for blind children. In Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility (pp. 137-144). ACM.
2. Annetta, L. A. (2008). Serious educational games. Theory into Practice, 83.
3. Yuan, B., Folmer, E., & Harris, F. C. (2011). Game accessibility: a survey. Universal Access in the Information Society, 10(1), 81-100.
4. Grammenos, D., Savidis, A., Stephanidis, C.: Designing universally accessible games. ACM Computers in Entertainment. 7, Article 8 (2009).
5. Torrente Vigil, F. J., Blanco Aguado, Á. D., Serrano Laguna, Á., Vallejo Pinto, J. A., Moreno Ger, P., & Fernández Manjón, B. (2014). Towards a low cost adaptation of educational games for people with disabilities. Computer Science and Information Systems, 11(1), 369-391.
6. Darin, T. G., Andrade, R., Merabet, L. B., & Sánchez, J. H. (2017, May). Investigating the Mode in Multimodal Video Games: Usability Issues for Learners who are Blind. In Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (pp. 2487-2495). ACM.
7. Shoukry, L., Sturm, C., & Galal-Edeen, G. H. (2015). Pre-MEGa: A Proposed Framework for the Design and Evaluation of Preschoolers' Mobile Educational Games. In Innovations and Advances in Computing, Informatics, Systems Sciences, Networking and Engineering (pp. 385-390). Springer International Publishing.
8. Drossos, K., Zormpas, N., Giannakopoulos, G., & Floros, A. (2015, July). Accessible games for blind children, empowered by binaural sound. In Proceedings of the 8th ACM International Conference on PErvasive Technologies Related to Assistive Environments (p. 5). ACM.Girard, C., Ecalle, J., & Magnan, A. (2013). Serious games as new educational tools: how effective are they? A meta-analysis of recent studies. Journal of Computer Assisted Learning, 29(3), 207-219.
9. Balan, O., Moldoveanu, A., Moldoveanu, F., & Dascalu, M. I. (2014). Audio games—a novel approach towards effective learning in the case of visually-impaired people. In Proceedings of seventh international conference of education, research and innovation, Seville.
10. Song, D., Karimi, A., & Kim, P. (2011, December). Toward designing mobile games for visually challenged children. In e-Education, Entertainment and e-Management (ICEEE), 2011 International Conference on (pp. 234-238). IEEE.
11. Iwarsson, S., & Ståhl, A. (2003). Accessibility, usability and universal design—positioning and definition of concepts describing person-environment relationships. Disability and rehabilitation, 25(2), 57-66.

12. Buzzi, M. and Leporini, B. (2009). Editing Wikipedia content by screen reader: easier interaction with the Accessible Rich Internet Applications suite. Disability and Rehabilitation: Assistive Technology, 4(4), 264-275.
13. Carvalho, L. P., Ferreira, L. P., & Freire, A. P. (2016). Accessibility evaluation of rich internet applications interface components for mobile screen readers. In Proceedings of the 31st Annual ACM Symposium on Applied Computing (pp. 181-186). ACM.
14. Madsen, H.S., Techniques in Testing. Oxford University Press
15. Kane, S. Wobbrock, J. O. and Ladner, R.E. 2011. Usable gestures for blind people: understanding preference and performance. In Proc. CHI '11. ACM, New York, NY, 413422.

About the Authors:

Barbara Leporini received at University of Pisa her Masters in Computer Science with full marks and honours in 1997 and in 2003 her PhD with a dissertation on accessibility and usability subject related to websites. Since the beginning of her PhD project, Barbara has done her research at the CNR of Pisa (ISTI) where she is now a researcher in the Human-Computer Interaction field, and in specific in accessibility and usability for people with special needs. Barbara investigates techniques and methods to make Web sites, electronic publishing, museum contents, search engines, e-learning, educational tools and identification systems for the electronic signature accessible and usable to disabled users. Generally speaking, she works on accessibility and usability of user interfaces for both desktop and mobile applications. Beyond research, Barbara has been teaching computer sciences classes. She has also been providing technical support for accessibility and usability at various levels. In particular, she has been participating on boards and groups working on different problems concerning visually-impaired people.

Eleonora Palmucci graduated in Computer Science in 2017 at the University of Pisa. Her thesis was on accessibility of educational tools on mobile devices, and especially via serious games. Currently she is a programmer.

Teachable Machines for Accessibility

Hernisa Kacorri
University of Maryland – College Park
hernisa@umd.edu

Abstract

How can accessibility research leverage advances in machine learning and artificial intelligence with limited data? In this article, we argue that teachable machines can empower accessibility research by enabling individuals with disabilities to personalize a data-driven assistive technology. By significantly constraining the conditions of the machine learning task to a specific user and their environment, these technologies can achieve higher robustness in real world scenarios. In contrast to automatic personalization, the end user is called to consciously provide training examples and actively interact with the machine learning algorithm to increase its accuracy. We demonstrate this concept with a concrete example: teachable object recognizers trained by and for blind users. Furthermore, we discuss open challenges in designing and building teachable machines with a focus on accessibility.

Introduction

Machine learning holds great promise for increasing independent living for people with disabilities. As accessibility researchers, we have embraced its possibilities. We have leveraged advances in natural language processing, speech, and computer vision to create novel assistive technologies, gain a deeper understanding of our users, and detect barriers in their environment. While not an extensive list, the SIGACCESS community has contributed to: predictive text-entry and brain-computer interfaces (*e.g.*, [4, 13]); sign language synthesis and recognition (*e.g.*, [9, 27]); automatic object recognition (*e.g.*, [30]); indoor localization (*e.g.*, [6, 21]); and physical world accessibility (*e.g.*, [7]).

At the heart of all machine learning applications is the need for data. This need becomes even greater as we transition to more complex deep learning architectures that offer state-of-the-art performance. In some cases, accessibility researchers can leverage existing large datasets obtained from a broader population. For example, in [30] photos of blind people are automatically labeled with models trained on millions of photos by sighted people. Similarly, many predictive models on text-entry for people with motor disabilities are trained on large text corpora by people with written language literacy [23]. However, for many assistive technologies, training data have to be specific to a population, *e.g.*, sign language videos from Deaf signers [8, 14, 15], mobility logs from people with visual impairment [3, 11], text corpora from people with dyslexia [18], and dysarthric speech recordings [19].

Research in data-driven assistive technologies is hindered by the scarcity of available data. Naturally, this is partly because focusing on a small portion of the population leads to smaller datasets. Yet there are other factors specific to the populations we study. People tend to vary in their individual preferences and environments, but people with disability lend another dimension with disparate characteristics, even within a given disability. This can lead to unique challenges for collecting big data sets. Moreover, data annotation often requires domain knowledge that few have. For example, creating sign language corpora requires sign language fluency, an additional challenge for obtaining annotations at scale, *e.g.* through crowdsourcing. While we observe a trend in our community for sharing resources, there are still many limitations for aggregating and combining datasets from different sources to facilitate large-scale and cross-context machine learning applications. This is partially due to privacy concerns and partially due to the fact that data are often related to very specific tasks, applications, or scenarios. For example, the outdoor mobility data for people with visual impairments in iMove [11] don't include user location or heading information and logs are tied to specific characteristics of the iMove interface and functionalities.

This leads to the question: *How can accessibility research leverage advances in machine learning and artificial intelligence with limited data?* This article explores **teachable machines**, a term borrowed from a recently released AI experiment [26], where anyone can train a neural network live in their browsers by providing a few examples through their camera. No prior machine learning or programming knowledge is required from the user and the training is done locally on the user's device. Under the hood, teachable machines usually deploy fully trained models on loosely related large-scale datasets and adapt them for new tasks using only a few samples.

In this article, we argue that **teachable machines can empower accessibility** research. They can enable individuals with disabilities to personalize a data-driven assistive technology. By significantly constraining the conditions of the machine learning task to a specific user and their environment, these technologies can achieve higher robustness in real world scenarios. Contrary to automatic personalization, the end user is called to consciously provide training examples and actively interact with the machine learning algorithm to increase its accuracy. We demonstrate this concept with a concrete example: teachable object recognizers for blind people. Furthermore, we discuss the open challenges in designing and building teachable machines with a focus on accessibility.

Related Work

While not a novel concept, the idea of having the end-user consciously provide examples for training has gain popularity with advances in neural networks. By leveraging prior work in transfer learning [16, 17] and k-shot learning [1, 25], we are now able to build robust systems that learn to represent a task in an abstract way and transfer this knowledge to new, previously unseen tasks given only one or few examples of the novel task. For example, users with no background in machine learning or programming can create new interactive applications such as gesturally controlled musical instruments by demonstrating a mapping between human action and responses [5, 29]. They can teach their phones to detect a specific object in their environment [28] or train smart home devices to be responsive to a specific user behavior [12].

Similar applications are emerging in accessibility research. Some of them have focused on user requirements and interface design, *e.g.* a sound detector app that is trained for and by deaf and hard-of-hearing users [2]. Others focus on the feasibility and accuracy of models trained on end-user data, *e.g.*, photos of everyday objects from video streams provided by blind users [22]. Our work on teachable object recognizers [10], presented in this article, falls under this second category. A unique challenge for deploying teachable machines in assistive technologies is that training requires similar skills to those the technology aims to fulfill. For example, to train a sound detector app, deaf users have to be able to provide accurate sound and label pairs. Similarly, to train an object recognizer, blind users need to provide accurate photo and label pairs. However, it is difficult for such users to assess the quality of their (annotated) training data. Thus, teachable machines in accessibility require innovative ways to provide feedback to end users about the quality of their training examples in different modalities.

Case Study: Teachable Object Recognizers

Blind people have been quick to adopt assistive technologies for object identification (*e.g.*, barcode readers and crowd-sourced image recognition). However, many factors, such as cost, accuracy, crowd availability, privacy, and photo-taking skills, are holding back their use for more independent living.

Teachable machines open new possibilities for this population by leveraging smartphone cameras and processors to overcome many of these challenges. By processing photos entirely on the user's mobile device, users can avoid sharing photos with servers or remote sighted assistants, thus reducing reliance on an Internet connection and ensuring privacy. For these reasons, we proposed a *personalized object recognition with teachable machines* [10]. Our work investigated the feasibility and challenges of this approach. Blind users trained a mobile application by providing a small sample of photos for objects of interest, along with personally meaningful names for the objects. Since recognition tasks were limited to a single user, we were able to constrain the number of object classes while reducing the variability between training and test images. This is true because both training and test photos would be taken by the same person under similar conditions and subjected similarly to any idiosyncratic (photo-taking) characteristics, as illustrated in Figure 1.

Blind people often use adhesive Braille labels, or ad hoc methods, to annotate their objects. However, the concept of training a recognizer though photo taking is largely unfamiliar territory for them. We solicited suggestions from participants on how such technology might be used, asked them to photograph a few snapshots of objects in their homes over a one-week period, and invited them to our lab for a round of *simulated* training and testing under controlled conditions and guidance. Our study uncovered user factors that degrade the performance and need to be addressed for any robust real world assistive application deploying teachable machines.

Figure 1: Examples of object instances that participants in our study chose for teaching their object recognizers.

Object Recognizer

Many recent advances in general image recognition tasks are due to the availability of object recognition models pre-trained on large datasets. One such state-of-the-art model is Google's Inception [24] trained on ImageNet [20], a benchmark dataset with millions of images. Inception and similar models allow researchers to build new recognizers by re-training or fine-tuning the existing model for new image classes with limited data while leveraging knowledge previously gained from very large datasets. Intuitively, as illustrated in Figure 2, low-level layers in the pre-trained neural network capture generic low-level representations for distinguishing objects, such as edges and colors. These representations can be re-purposed for different, but related, recognition tasks. To build a teachable object recognizer, we replace the top layer of pre-trained Inception with a new softmax layer. We re-trained the last layer using labels provided by our users. This newly trained user specific model forms the basis for the teachable object recognizer used to predict test images provided by the user.

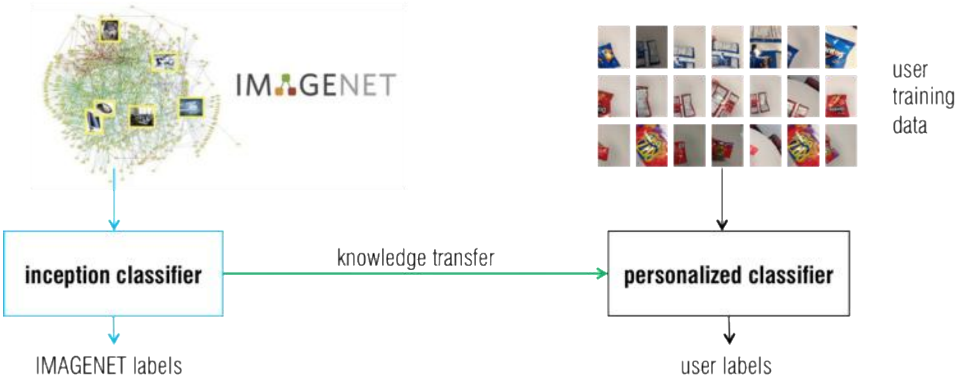

Figure 2: A visual representation of transfer learning in teachable object recognizers.

User Study

We explore the feasibility and challenges of teachable object recognizers for blind users with a two phase study. We recruited 8 blind participants (P1 – P8 with demographics shown in Table 1) from the local community who were familiar with smartphones. In Phase 1, we asked participants to photograph a few snapshots of objects in their homes over a one-week period. In Phase 2, we invited them to our lab for a round of *simulated* training and testing under controlled conditions and guidance (Phase 2).

Table 1: Participants' demographics and smartphone use period.

ID	*Gender*	*Age*	*Onset*	*Handedness*	*Smartphone*
P1	F	42	birth	right	2014
P2	M	40	birth (light)	right	2015
P3	F	68	birth	right	2008
P4	M	63	birth	left	2009
P5	F	46	birth	right	2016
P6	M	43	birth	right	2006
P7	F	61	birth (light)	right	2010
P8	F	58	10 months	right	2011

Phase 1. Participants were probed to imagine that they were teaching a phone application which uses the camera to recognize objects in their home. They were asked to take 5 photo examples for each object for teaching their personal machines. We received photos of up to 50 distinct objects per participant with about 1,500 total photos across participants. Objects spanned the following categories: food/drink, hygiene product, appliance, perfume, medication, and clothing. Figure 3 illustrates how one participant conceptualized a teachable object recognizer as a personal t-shirt recognition application.

Figure 3: Photos from a participant who conceptualized the application as a personal t-shirt recognizer. First row indicates the variation between the 5 photos per object and the rest of the rows the variation between different object for the classifier to distinguish.

By analyzing the labels and photos from all the participants in [10] we observed:

- ***Preference for personalized metadata*** such as brand, label, name, color, scent, flavor, location and date of purchase, washing instructions, and cost.
- ***Challenges in assigning labels*** to objects that were purchased in the past.
- ***Distinct and consistent training strategies*** across participants and similarly shaped objects with variation in viewpoints, distances from the object, rotation, and visible side.
- ***Exaggerated or non-discriminative viewpoints*** that could be an artifact of how training is perceived by the participants or due to the fact that training photos were obtained sequentially.
- ***Need for guidance and feedback*** to improve the quality of training images such as limiting background clutter, ensuring the object is in the camera scope, and obtaining more discriminative photos.

Phase 2. Blind participants P1-P8 and two sighted people S1-S2, who served as a baseline, were asked to train and test a personal object recognizer in a lab setting for 15 objects in the food/drink category (shown in Fig. 4). The collected data were used to assess the feasibility of this approach and explore effects of photo-quality variance, sample size, and object characteristics in a series of training experiments. We collected a total of 4,120 photos in training mode and 661 in test mode.

Figure 4: Our preliminary study shows high variation among how blind people take photos but idiosyncratic personal style per individual.

When looking at the participants' photos and video recordings of the sessions in [10], we observed:

- ***Presence of user's hand in training images*** to either hold an object or serve as a reference point to place the object in the photo. This behavior was mirrored in test mode across very few participants.
- ***Reinforced distinct and consistent teaching strategies***. With the option of more photos per object participants introduced many user-distinct variations such rotations and viewpoints.

- ***Variation in training unobserved in test*** where the number of 30 examples was perceived as high and often interpreted as an opportunity for variation covering edge cases.
- ***Different teaching strategies among sighted participants*** indicating that teaching strategies, and thus the performance of the classifier, may be affected by how people perceive machine learning concepts and not their photo-taking skills.

Experiment Results

The data obtained in Phase 2 were used to build object recognizers for each participant (P1-P8 and S1-S2) and report their performance in recognizing test images belonging to 15 object instances. Training images were always pulled from the set of photos collected in training mode. Test images included all photos taken in test mode, about 75 images per participant.

For blind participants P1-P8, the teachable object recognizers achieved accuracies ranging from 50% to 92%, as shown in Figure 5. Random prediction for a 15-way classification would yield about 7% accuracy. A recognizer taught by a sighted computer scientist (S1) with knowledge of the underlying algorithms achieved an average accuracy over 99%, and a recognizer taught by a sighted person unaware of the underlying methods (S2), achieved an average accuracy around 97%. The fact that some of our participants, P1 and P8, achieved performances comparable to those of a sighted participant S2 highlights the potential of teachable machines.

Error analysis of the results indicated that objects which are difficult to recognize are not uniformly distributed across participants. Similarly, reasons for lower accuracies vary across users due to idiosyncratic characteristics of the participant. For example, P2 provided examples photos lacking the object, while P8 provided photos with exaggerated viewpoints only while teaching the recognizer (but not while testing it).

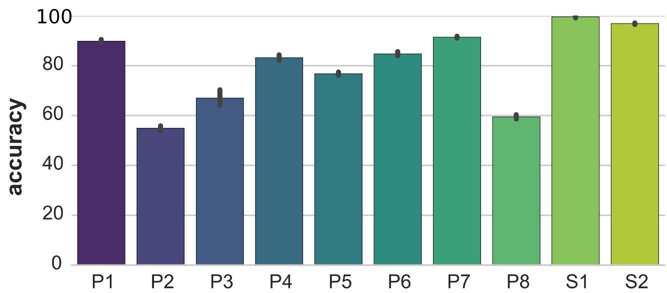

Figure 5: Performance of a teachable object recognizer across ten participants with 20 samples per object. Error bars were calculated over 10 random experiment runs.

We performed a series of experiments to explore effects of sample size, photo-quality variance, and object shape; and contrast models trained on photos by blind participants to those by sighted participants [10]. Our results indicate:

- No consistent effect of training examples order in the performance of the recognizer.
- Improved performance with additional training examples or data augmentation with distortion such as cropping and scaling. Marginal gains for limited variation or noise in training examples.
- With sufficient training data, blind users are better off training their own object recognizer rather than having a sighted person providing the training examples.

Open Challenges

In the teachable object recognizer study, participants indicated the importance of: 1) knowing whether the photos were good, 2) knowing the area of a package where the label or most discriminating information resides, 3) obtaining feedback from the application such as lighting conditions and number of photos taken, and 4) deciding on the distance between the object and camera lens. One of the participants stated "the most challenging and most fun is training the person [who will be teaching the machine]". We see the concerns raised by our participants as representative of challenges in the general case of teachable machines for accessibility.

With teachable machines, users are expected to be somewhat aware of the machine's learning process and foresee their future actions and settings to ensure that training and test examples and conditions are matched. This opens new research questions in how we design our application to communicate this underlying factor in the performance of teachable applications or make the application more robust to such factors. Teaching machines for accessibility pose a unique challenge since asking users to provide good training examples requires similar skills to those the technology aims to fulfill. For example, how can we develop applications that aid blind users to teach their applications to "see" on their behalf?

There are many parameters to be evaluated for teachable assistive applications such as incremental model learning and robustness to noisy conditions. More important is the issue of scalability over long periods of time. Although studies such as ours indicate success for limited number of classes and data, what happens as the number of objects increase over time and become more fine grained, increasing to hundreds or thousands of labels?

Acknowledgments

The work on personal objects recognizers, which served as a case study for this article, was conducted with Kris M. Kitani, Jeffrey P. Bigham, and Chieko Asakawa. The author would like to thank Shimizu Corporation and JST CREST for funding support for this study. The author is additionally grateful to the participants of the study for all their insight and contributions to the experiments.

References

1. Bauer, M., Rojas-Carulla, M., Świątkowski, J.B., Schölkopf, B. and Turner, R.E., 2017. Discriminative k-shot learning using probabilistic models. *arXiv preprint arXiv:1706.00326*.
2. Bragg, D., Huynh, N. and Ladner, R.E., 2016, October. A personalizable mobile sound detector app design for deaf and hard-of-hearing users. In *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility* (pp. 3-13). ACM.
3. Flores, G.H. and Manduchi, R., 2016, October. WeAllWalk: An Annotated Data Set of Inertial Sensor Time Series from Blind Walkers. In *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility* (pp. 141-150). ACM.
4. Fowler, A., Roark, B., Orhan, U., Erdogmus, D. and Fried-Oken, M., 2013. Improved inference and autotyping in EEG-based BCI typing systems. In *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility* (p. 15). ACM.
5. Giorgio Cam. https://experiments.withgoogle.com/ai/giorgio-cam/view/, 15 October 2017.
6. Gleason, C., Guo, A., Laput, G., Kitani, K. and Bigham, J.P., 2016. VizMap: Accessible Visual Information through Crowdsourced Map Reconstruction. In *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility* (pp. 273-274). ACM.

7. Hara, K. and Froehlich, J.E., 2015. Characterizing and visualizing physical world accessibility at scale using crowdsourcing, computer vision, and machine learning. *ACM SIGACCESS Accessibility and Computing*, (113), pp.13-21.
8. Huenerfauth, M. and Kacorri, H., 2014. Release of experimental stimuli and questions for evaluating facial expressions in animations of American Sign Language. In *Proceedings of the 6th Workshop on the Representation and Processing of Sign Languages: Beyond the Manual Channel (Vol 5)*.
9. Kacorri, H. and Huenerfauth, M., 2016. Continuous profile models in ASL syntactic facial expression synthesis. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics* (Volume 1: Long Papers) (Vol. 1, pp. 2084-2093).
10. Kacorri, H., Kitani, K.M., Bigham, J.P. and Asakawa, C., 2017, May. People with Visual Impairment Training Personal Object Recognizers: Feasibility and Challenges. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 5839-5849). ACM.
11. Kacorri, H., Mascetti, S., Gerino, A., Ahmetovic, D., Takagi, H. and Asakawa, C., 2016, October. Supporting orientation of people with visual impairment: Analysis of large scale usage data. In *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility* (pp. 151-159). ACM.
12. Karmann, B., Objectifier Spatial Programming. https://experiments.withgoogle.com/ai/objectifier-spatial-programming, 15 October 2017.
13. Koester, H.H. and Levine, S.P., 1994, October. Validation of a keystroke-level model for a text entry system used by people with disabilities. In *Proceedings of the first Annual ACM conference on Assistive Technologies* (pp. 115-122). ACM.
14. Lu, P. and Huenerfauth, M., 2014. Collecting and evaluating the CUNY ASL corpus for research on American Sign Language animation. *Computer Speech & Language*, *28*(3), pp.812-831.
15. Neidle, C., Liu, J., Liu, B., Peng, X., Vogler, C. and Metaxas, D., 2014. Computer-based tracking, analysis, and visualization of linguistically significant nonmanual events in American Sign Language (ASL). In *LREC Workshop on the Representation and Processing of Sign Languages: Beyond the Manual Channel (Vol. 5).*
16. Pan, S.J. and Yang, Q., 2010. A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10), pp.1345-1359.
17. Patricia, N. and Caputo, B., 2014. Learning to learn, from transfer learning to domain adaptation: A unifying perspective. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1442-1449).
18. Rello, L., Baeza-Yates, R. and Llisterri, J., 2017. A resource of errors written in Spanish by people with dyslexia and its linguistic, phonetic and visual analysis. *Language Resources and Evaluation*, 51(2), pp.379-408.
19. Rudzicz, F., Namasivayam, A.K. and Wolff, T., 2012. The TORGO database of acoustic and articulatory speech from speakers with dysarthria. *Language Resources and Evaluation*, 46(4), pp.523-541.
20. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M. and Berg, A.C., 2015. Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3), pp.211-252.
21. Sato, D., Oh, U., Naito, K., Takagi, H., Asakawa, C., and Kitani, K., 2017. NavCog3: An Evaluation of a Smartphone-Based Blind Indoor Navigation Assistant with Semantic Features in a Large-Scale Environment. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility* (pp. 270-279). ACM.

22. Sosa-García, J. and Odone, F., 2017. "Hands On" Visual Recognition for Visually Impaired Users. *ACM Transactions on Accessible Computing (TACCESS)*, *10*(3), p.8.
23. Speier, W., Arnold, C. and Pouratian, N., 2016. Integrating language models into classifiers for BCI communication: a review. *Journal of neural engineering*, 13(3), p.031002.
24. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. and Wojna, Z., 2016. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2818-2826).
25. Vinyals, O., Blundell, C., Lillicrap, T. and Wierstra, D., 2016. Matching networks for one shot learning. In *Advances in Neural Information Processing Systems* (pp. 3630-3638).
26. Teachable Machine. https://experiments.withgoogle.com/ai/teachable-machine, 15 October 2017.
27. Wang, H., Chai, X., Hong, X., Zhao, G. and Chen, X., 2016. Isolated Sign Language Recognition with Grassmann Covariance Matrices. *ACM Transactions on Accessible Computing (TACCESS)*, 8(4), p.14.
28. Warden P., iPhoneTracker. https://github.com/petewarden/iPhoneTracker, 15 October 2017.
29. Wekinator. https://experiments.withgoogle.com/ai/giorgio-cam/view/, 15 October 2017.
30. Zhong, Y., Garrigues, P.J. and Bigham, J.P., 2013, October. Real time object scanning using a mobile phone and cloud-based visual search engine. In *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility* (p. 20). ACM.

About the Author:

Hernisa Kacorri is an assistant professor in the iSchool at University of Maryland, College Park. Her research focuses on data-driven accessibility with an emphasis on rigorous, user-based experimental methodologies to assess impact. Hernisa received her Ph.D. in computer science in 2016 from City University of New York, and has conducted research at National and Kapodistrian University of Athens, IBM Research-Tokyo, Lawrence Berkeley National Lab, and Carnegie Mellon University. Hernisa is a recipient of a Mina Rees Dissertation Fellowship in the Sciences, an ACM ASSETS best paper finalist, and a CHI honorable mention award. She has been recognized by the Rising Stars in EECS program of CMU/MIT.

DESIGNING SMARTGLASSES APPLICATIONS FOR PEOPLE WITH LOW VISION

Shiri Azenkot and Yuhang Zhao
Jacobs Technion-Cornell Institute, Information Science, Cornell Tech, Cornell University
shiri.azenkot@cornell.edu; yz769@cornell.edu

Abstract

While our community has many active projects involving blind people, low vision is rarely addressed. People with low vision have functional vision, but their visual impairment adversely affects their daily life and it cannot be corrected with glasses or contact lenses. Over the last few years, we have been conducting research with this understudied demographic: understanding low vision people's needs and designing applications to address the challenges they face. In this article, we discuss our ongoing research in this area, focusing on designing augmented reality applications for low vision users. We begin this article by describing low vision and motivating our focus on augmented reality applications on smartglasses for low vision people. We then provide overviews of three research projects that exemplify our research agenda: a study where we observed low vision people conducting a navigation and shopping task, a study where we examined low vision people's perception of virtual text and shapes on smartglasses, and the design of a smartglasses application that facilitates a visual search task.

Low Vision

According to the Centers for Disease Control and Prevention (CDC), 3.3 million Americans who are 40 years old and older have low vision, and 19 million Americans have trouble seeing even when wearing glasses or contact lenses [9]. A person has *low vision* if he or she has difficulty seeing when performing daily activities, even with glasses or contact lenses [4,12]. Low vision is especially common among older adults, so the number of people with low vision is expected to rise significantly in the next 20 years with the general aging of the population. There are different kinds of low vision, including limited peripheral or central vision, blurry vision, light sensitivity, or blind spots [1]. While about 90% of people with a visual disability have functional vision, nearly all accessibility research has focused on nonvisual audio or tactile technology [3,16]. There is thus a major opportunity to advance low vision accessibility research and impact the lives of millions of people.

Low vision accessibility tools use simple image enhancement techniques and have not changed much in decades[1]. Low-tech low vision aids include handheld optical magnifiers and monoculars. Digital tools are available on desktop and mobile platforms, and most often allow users to magnify the contents of the screen, increase the contrast, or reverse the colors of the display. The simple "signal-to-signal" image processing techniques (image scaling, contrast enhancement, *etc*.) that are pervasive in low vision accessibility tools do not take the user's *context* into account: they are applied regardless of the contents of the screen, the user's task, or the user's visual condition.

[1] OrCam [18], a camera mounted on a user's glasses, is one of the very few recent innovations in low vision technology. If a user points to printed text, OrCam recognizes and speaks the text.

Augmented Reality on Smartglasses

Smartglasses [5] now present an exciting opportunity to revolutionize low vision accessibility: they have embedded sensors that can be leveraged to learn about the user's context and displays for presenting visual feedback that augments the users vision. These devices have the potential to: (1) solve important accessibility challenges and (2) become widely adopted by a wide range of low vision people. Since Google Glass was released in 2013, a variety of companies have already developed smartglasses for augmented and virtual reality. Unlike with smartphones or handheld magnifiers, smartglasses can provide faster hands-free access: the user does not need to pull out a phone and aim the camera to capture her surroundings. They are the next generation of corrective eyewear. As mainstream devices, smartglasses offer additional benefits. Prior research has shown that people with disabilities prefer to use mainstream devices (i.e., not specialized for users with disabilities) because they are more socially acceptable and non-stigmatizing, affordable, and easily available [7,8,10].

Research Questions

Our research on low vision people aims to address three high-level research questions:

1. What challenges do people with low vision face in their daily lives?
2. What are low vision people's perceptual abilities on smartglasses?
3. What designs for smartglasses applications address the challenges faced by people with low vision?

Below we describe examples of recent work that addresses each of our high-level research questions.

Understanding Low Vision People's Daily Challenges

When we first embarked on low vision research, we were surprised that no prior work had studied the challenges that low vision people experience when performing daily activities. Beyond their challenges, no research had explored what tools and accommodations low vision people used in their daily lives. Many low vision tools exist—from hand held magnifying glasses, to monoculars, to the Zoom magnification feature on iOS—but were people using these tools? If so, what were the patterns and challenges of their use?

To answer these questions, we conducted a study [11] where we observed low vision people conducting a navigation and shopping task, activities necessary for an active and independent life. Specifically, we tasked participants with finding a nearby pharmacy and buying a specific Tylenol product. We recruited 11 participants with low vision (5 males, 6 females) with a variety of visual conditions. Their age range was 20-68, with a mean of 41. Seven participants were born with low vision or became low vision as children, 3 participants became low vision as young adults, and one participant became low vision as an (slightly) older adult (age 55).

We used *contextual inquiry* to interview and observe participants in-situ. This method, of examining participants' behavior in a daily activity in real-time, allowed us to discover how people with low vision perform tasks, what kind of struggles they encountered, what aids they tended to use, and how these aids helped them.

During the task, participants used their vision extensively to gain information about their environment, but they experienced many challenges. The main aid participants used was their smartphone, which was mostly used for outdoor navigation (*e.g.*, using a map application).

Navigation using the smartphone was challenging, because the smartphone's accessibility tools did not provide adequate support. For example, participants continually readjusted the magnification level of the built-in magnifier to see text at different font sizes. Meanwhile, the gestures needed for these operations were difficult to perform. Navigating around the pharmacy was the most challenging part of the task. Participants did not use any tools for indoor navigation or searching for the product. When participants found the correct aisle, finding the correct product among rows of similar products was daunting. Most picked up each product individually, examined it closely (sometimes using a magnifying glass), until they found the desired product.

In sum, the study revealed many open problems for the accessibility community to consider. While many recent innovations have focused on blind people, our participants did not use any nonvisual accessibility tools. They used their vision extensively and desired better tools for their mainstream devices.

Understanding Low Vision People's Perceptual Abilities on Augmented Reality Glasses

While augmented reality smartglasses are advancing rapidly, these technologies are still limited in many ways. Consider the Epson Moverio or Microsoft HoloLens. Their displays are small compared to the user's field of view and the range of visual stimuli that can be projected is limited. Virtual elements are somewhat transparent so they have relatively low contrast. Given these limitations, we sought to determine whether and what low vision people can perceive on smartglasses. Is it feasible to design smartglasses applications with today's commercial platforms? What kinds of stimuli (text, shapes, colors) could low vision people see, if at all, on such platforms?

We conducted a study [14] with 20 participants with low vision (9 male, 11 female, mean age=45), evaluating their ability to see virtual elements on the Epson Moverio BT-200 glasses [2]. We used the Epson Moverio glasses because, by our assessments, they have advanced features for a consumer-grade device. Compared with Google Glass, for example, Epson Moverio glasses provide a larger projection area in the center of both the user's eyes, which makes them more versatile for people with different visual abilities.

The study consisted of one lab session where we assessed participants' visual perception of virtual objects on the smartglasses while they sat and walked in an office conference room. We displayed different sets of short phrases and shapes on a transparent background and varied their size, color, thickness, and, for text, font faces. We did this by varying one parameter as we held the rest constant. Participants identified randomly selected stimuli with the given parameters. After presenting the set of stimuli, we reviewed participants' performance and sought their feedback on their preferred parameters for short phrases and shapes.

Figure 1. Actual stimuli presented on smartglasses in our study, shown over dark (top) and light (bottom) backgrounds. The triangle and a short phrase are shown in the different colors used in the study. In addition to color, we also varied the stimulus size, thickness, and font face (for text).

On a high level, we found that people with low vision can benefit from commercial smartglasses. Participants identified the projected shapes and read short phrases while sitting and walking. While all participants could comfortably recognize shapes, only participants with mild to moderate low vision were able to comfortably read the short phrases.

We were pleasantly surprised with these findings, which indicated that there is much potential for designing smartglasses applications that use shapes and even text to augment people's vision. Text and shapes are basic visual output elements that can convey a wide variety of information in different contexts. For example, distant signs or nearby small text can be made accessible by displaying it on the smartglasses. Similarly, simple shapes like arrows can be used to provide guidance in navigation applications. Of course, our study only examined one smartglasses platform in a well-lit office environment with mostly white walls. However, our findings suggest that accessibility smartglasses applications may be feasible with today's technology.

Designing Novel Smartglasses Applications for Low Vision

As we found in our navigation and shopping study, certain daily tasks were not supported by existing low vision tools. Participants had difficulty with *visual search*, finding a known target among a set of distractors. Visual search tasks are common in our daily lives: people search for a friend in the crowd, the bathroom sign in an airport, or a certain word in a document. An example of a visual search task from our study was searching for the specified product on the store shelves. Low vision people could use a magnifier to see the details of each product, but no tool was available to help them quickly scan the shelves and recognize the desired product.

We designed *CueSee* [15], a video see-through smartglasses prototype to address product search for low vision people. CueSee located a specified product with computer vision and presented visual cues to direct the user's attention to the product. A cue that a sighted person can easily see (*e.g.*, a red dot) may be outside a low vision user's field of view or may not have sufficient contrast with the background for a low vision user. We thus designed the cues for different low vision conditions [1,17], based on cognitive psychology theories on attention [6,13]. The visual cues included: *Guideline*, *Spotlight*, *Flash*, *Movement*, and *Sunrays* (Figure 2). In addition to the cues,

Figure 2. **Visual cues in CueSee, a product search smartglasses system with five visual cues that guide the user's attention to the target product. CueSee was implemented on video see-through smartglasses and evaluated in a lab study with the mock grocery store shelf shown here** [15].

the system enhanced the target product with magnification and increased contrast to help users read the product information.

We explored the effectiveness of the cues in the product search context in a laboratory study with a mock grocery store shelf (Figure 2). We recruited 12 low vision participants with different visual conditions. Our results were promising: all participants found products significantly faster with CueSee than with their best-corrected vision. Moreover, participants did not make a single error (picking up an incorrect product) with CueSee. They had a variety of cue preferences: people with mild low vision preferred Guideline and Spotlight, while people with moderate to severe low vision preferred Flash and Sunrays. None of the participants liked Movement. While it attracted their attention, they felt this cue hindered their ability to see the product (which they wanted to be able to do).

The initial CueSee prototype was implemented on a *video see-through* display (an Oculus Rift with a web cam attached), unlike the perception study describe above which used an *optical see-through* display. Optical see-through displays tend to be lighter and more portable, designed for use in mobile scenarios. Moreover, they don't cover the user's eyes or cause disorientation. We plan to implement CueSee on optical see-through glasses, improve the cues with further studies, and add support for a variety of visual search tasks.

Conclusion

Studying low vision has been a fascinating and challenging journey. Unlike blindness, low vision manifests in different forms and degrees of severity. Visual perception is complex, and people's vision can change depending on the time of day, their task, and their environment. People with low vision often don't identify as "disabled" and don't necessarily associate with an advocacy or service organization. Sadly, many lose their sight as they age and view their condition as a medical problem that cannot be treated, rather than a disability they can overcome. Recruiting low vision users and understanding their experiences and visual abilities has thus been an interesting challenge. Similarly, we have also found it difficult at times to describe our work to researchers in the community. While low vision tools exist, there are many exciting open research problems that could significantly impact people's lives; simple tools like optical and digital magnifiers help many people, but they are not a one-stop solution for low vision. We hope you will join us in exploring new solutions for this large and growing population.

Acknowledgments

This research was partially funded by a gift from the Yang Tan Institute on Employment and Disability and by NSF Award No. 1657315.

References

1. AOA (American Optometric Association). Common Types of Low Vision.
2. Epson. Epson Moverio™ BT-200 Smart Glasses (Developer Version Only) - Product Information - Epson America, Inc.
3. JA Jacko and A Sears. 1998. Designing interfaces for an overlooked user group: considering the visual profiles of partially sighted users. *Proceedings of the third international ACM …*. Retrieved July 5, 2015 from http://dl.acm.org/citation.cfm?id=274512
4. Lighthouse International. All About Low Vision. Retrieved July 17, 2015 from http://www.lighthouse.org/about-low-vision-blindness/all-about-low-vision
5. S. Mann. 1997. Wearable computing: a first step toward personal imaging. *Computer* 30, 2, 25–32. http://doi.org/10.1109/2.566147

6. Peter McLeod, Jon Driver, and Jennie Crisp. 1988. Visual Search for a Conjunction of Movement and Form is parallel. *Nature* 336, 403–405.
7. TLB Pape, J Kim, and B Weiner. 2002. The shaping of individual meanings assigned to assistive technology: a review of personal factors. *Disability and rehabilitation*. Retrieved July 5, 2015 from http://www.tandfonline.com/doi/abs/10.1080/09638280110066235
8. Betsy Phillips and Hongxin Zhao. 1993. Predictors of Assistive Technology Abandonment. *Assistive Technology* 5, 1, 36–45. http://doi.org/10.1080/10400435.1993.10132205
9. JS Schiller and JW Lucas. 2012. Summary health statistics for US Adults: National health interview survey, 2010. *... National Health Survey*. Retrieved July 17, 2015 from http://europepmc.org/abstract/med/22834228
10. Kristen Shinohara and Jacob O. Wobbrock. 2011. In the shadow of misperception. *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11*, ACM Press, 705. http://doi.org/10.1145/1978942.1979044
11. Sarit Szpiro, Yuhang Zhao, and Shiri Azenkot. 2016. Finding a Store, Searching for a Product: A Study of Daily Challenges of Low Vision People. *The 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (Ubicomp'16).*, ACM New York, NY, USA.
12. VisionAware. Low Vision and Legal Blindness Terms and Descriptions. Retrieved September 22, 2014 from http://www.visionaware.org/info/your-eye-condition/eye-health/low-vision/low-vision-terms-and-descriptions/1235
13. Steven Yantis and John Jonides. 1990. Abrupt visual onsets and selective attention: Voluntary versus automatic allocation. *Journal of Experimental Psychology: Human Perception and Performance* 16, 1, 121–134.
14. Yuhang Zhao, Michele Hu, Shafeka Hashash, and Shiri Azenkot. 2017. Understanding Low Vision People's Visual Perception on Commercial Augmented Reality Glasses. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*, ACM Press, 4170–4181. http://doi.org/10.1145/3025453.3025949
15. Yuhang Zhao, Sarit Szpiro, Jonathan Knighten, and Shiri Azenkot. 2016. CueSee: Exploring Visual Cues for People with Low Vision to Facilitate a Visual Search Task. *The 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (Ubicomp'16)*, ACM New York, NY, USA.
16. What You Need to Know About Low Vision - American Foundation for the Blind. Retrieved July 19, 2015 from http://www.afb.org/info/low-vision/living-with-low-vision/23
17. What Causes Low Vision? Retrieved February 11, 2016 from http://www.aoa.org/patients-and-public/caring-for-your-vision/low-vision/what-causes-low-vision?sso=y
18. OrCam - See for Yourself. Retrieved June 18, 2015 from http://www.orcam.com/

About the Authors:

Shiri Azenkot is an assistant professor of information science at the Jacobs Technion-Cornell Institute at Cornell Tech, Cornell University, where she directs the Enhancing Ability Lab. Her research is in human-computer interaction, with a focus on accessibility and interaction on new platforms. Before arriving at Cornell, Shiri received a PhD in computer science at the University of Washington, where she was advised by Richard Ladner and Jacob Wobbrock.

Yuhang Zhao is a fourth-year Ph.D. student at Cornell University, majoring in Information Science. She is also an active member of the Enhancing Ability Lab in Cornell Tech. She conducts research in the area of human-computer interaction, especially accessibility, augmented reality, and virtual reality. Her current research focuses on innovating augmented reality systems to provide people with low vision equal access to information. Before joining Cornell, she received a bachelor's and a master's degree in Computer Science from Tsinghua University in China.