

Designing and Evaluating a Customizable Head-mounted Vision Enhancement System for People with Low Vision

YUHANG ZHAO, Information Science, Cornell Tech, Cornell University, USA

SARIT SZPIRO, The Special Education Department, University of Haifa, Israel

LEI SHI, Information Science, Cornell Tech, Cornell University, USA

SHIRI AZENKOT, Jacobs Technion-Cornell Institute, Cornell Tech, Cornell University, USA

Recent advances in head-mounted displays (HMDs) present an opportunity to design vision enhancement systems for people with *low vision*, whose vision cannot be corrected with glasses or contact lenses. We aim to understand whether and how HMDs can aid low vision people in their daily lives. We designed *ForeSee*, an HMD prototype that enhances people's view of the world with image processing techniques such as magnification and edge enhancement. We evaluated these vision enhancements with 20 low vision participants who performed four viewing tasks: image recognition and reading tasks from near- and far-distance. We found that participants needed to combine and adjust the enhancements to comfortably complete the viewing tasks. We then designed two input modes to enable fast and easy customization: speech commands and smartwatch-based gestures. While speech commands are commonly used for eyes-free input, our novel set of onscreen gestures on a smartwatch can be used in scenarios where speech is not appropriate or desired. We evaluated both input modes with 11 low vision participants and found that both modes effectively enabled low vision users to customize their visual experience on the HMD. We distill design insights for HMD applications for low vision and spur new research directions.

CCS Concepts: • **Human-centered computing** → **Accessibility technologies**;

Additional Key Words and Phrases: Accessibility, head-mounted displays (HMD), interaction techniques, low vision

ACM Reference format:

Yuhang Zhao, Sarit Szpiro, Lei Shi, and Shiri Azenkot. 2019. Designing and Evaluating a Customizable Head-mounted Vision Enhancement System for People with Low Vision. *ACM Trans. Access. Comput.* 12, 4, Article 15 (December 2019), 46 pages.

<https://doi.org/10.1145/3361866>

1 INTRODUCTION

Low vision is a visual impairment that falls short of blindness but cannot be corrected with glasses or contact lenses. According to the Center for Disease Control and Prevention (CDC), there are

This work was supported in part by the National Science Foundation under Grant Award No. 1657315 and by a gift from Verizon Media.

Authors' addresses: Y. Zhao, L. Shi, and S. Azenkot, Information Science, Cornell Tech, 2 W Loop Rd, New York, NY 10044; emails: {yz769, ls776, shiri.azenkot}@cornell.edu; S. Szpiro, Faculty of Education, University of Haifa, 199 Aba Khoushy Ave. Mount Carmel, Haifa, Israel 3498838; email: sarit.szpiro@gmail.com.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2019 Association for Computing Machinery.

1936-7228/2019/12-ART15 \$15.00

<https://doi.org/10.1145/3361866>

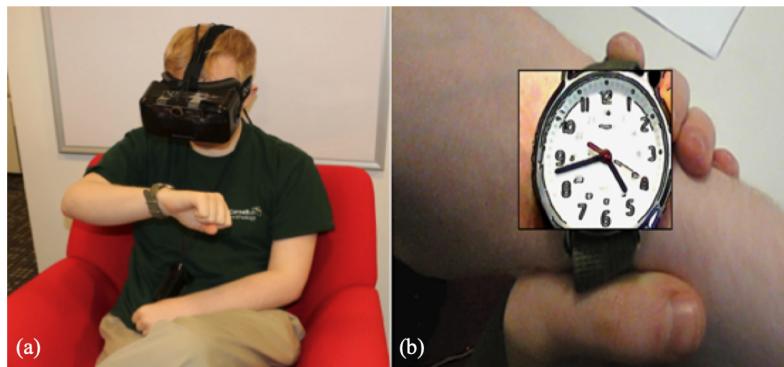


Fig. 1. ForeSee: (a) A low vision person uses ForeSee to check the time. (b) The user sees the enhanced watch surface with a combination of the Magnification, Contrast, and Edge Enhancements in Window Display Mode.

roughly 19 million people in the United States who have difficulty seeing when performing daily activities, even with corrective contact lenses or glasses [1]. Meanwhile, the National Federation of the Blind reports that there are only about 1.3 million legally blind people in the U.S. [66]. This means that there are millions of people who have functional vision and could leverage their vision for performing daily activities more efficiently and comfortably [92]. There has been a wealth of research on people with vision disabilities in the technology and disability community, yet rarely does this research focus on vision enhancement innovations and address low vision people's specific needs. Research on vision disabilities has largely focused on audio or vibrotactile interaction [22, 48, 115], overlooking the fact that the vast majority of people with vision disabilities have functional vision.

We seek to seize an opportunity to enhance the vision of people with low vision using head-mounted display systems, an emerging mainstream technology. In the past few years, a variety of companies have developed head-mounted display (HMD) systems for augmented reality [25, 34, 65] and virtual reality [71, 86] applications. Some HMDs provide an always-available display [34, 93]. These systems often include embedded cameras and other sensors so they can capture information from the users and their surroundings while presenting altered or enhanced visual information on their displays. HMD systems have received attention from the press [114], consumers, industry, and researchers and will likely continue to develop and become widely adopted. As with mass-market technology, they will be relatively affordable and available to consumers, unlike specialized assistive technology [49, 82, 84]. HMDs could therefore become effective, versatile, and practical vision enhancement platforms, which low vision people would likely accept as accessibility tools in their daily lives.

To explore the potential of HMD systems for people with low vision, we explore two high-level research questions in this article focusing on output and input:

- RQ1. Output: How do different visual enhancements on an HMD affect low vision people's visual experiences?
- RQ2. Input: What interaction techniques enable people with low vision to interact with an HMD system effectively?

To answer these questions, we developed ForeSee (Figure 1), a prototype video see-through augmented reality platform that has five visual enhancement methods and two display modes based on conventional low vision tools [119]. ForeSee's five enhancement methods include Magnification,

Contrast, Edge Enhancement, Black/White Reversal, and Text Extraction; and its two display modes include Full Display Mode and Window Display Mode, which apply the enhancements to the full field of view or only to a region within the view, respectively. The user can customize her experience with ForeSee according to their vision and a given environment or task by selecting combinations of enhancement methods in either display mode and adjusting their parameters.

To answer our first research question, we conducted an extensive user study with 20 participants with a variety of visual abilities to evaluate the different enhancement methods and display modes in ForeSee (some results were presented in Reference [119]). Participants used ForeSee to complete four visual tasks: (1) recognizing celebrities in handheld photos, (2) reading a page of printed text, (3) describing traffic signs, and (4) describing signs of numbers and writings that were hung at eye-level at a distance of about three meters. Participants tried all enhancement methods and display modes in each task and described their visual experience, in terms of what they saw and how they felt. We discuss the benefits and weaknesses of each enhancement method and display mode, as well as how participants with different visual conditions customized their visual experiences.

To answer our second research question, we designed two sets of interaction techniques to allow users to interact with ForeSee. In particular, these interaction techniques enabled users to customize the enhancement methods. We first designed a set of speech commands, a commonly used eyes-free input method that requires only a microphone. We also sought an alternative that can be used in a more private and discreet way. As such, we designed gesture-based techniques on a smartwatch. Smartwatches are mainstream devices that are relatively inexpensive and easy to obtain, so that they can serve as “always available” [67] input devices if connected to an HMD (via BlueTooth, for example).

We followed a user-centered design process [20] to design the smartwatch gestures. To our knowledge, no prior work has explored touchscreen-based interactions on a smartwatch for low vision people to interact with HMDs. We iteratively refined the design of the gestures with 8 low vision users. We evaluated both the speech and the smartwatch input modalities in a letter-reading task and found that participants accurately read small letters (1–3 inches in height) on a large display at a 10-foot distance with a mean time of 12.23 seconds using the smartwatch interactions, and 9.06 seconds using speech input. Our study demonstrated that touchscreen-based smartwatch interaction is a promising input method, which enables low vision people to interact with HMDs quickly and easily in daily tasks with a short learning curve.

In summary, this article has two contributions:

- (1) An extensive evaluation of ForeSee, which sheds light on the effectiveness of different enhancement methods and display modes in an HMD setting, and their impact on people with different visual abilities performing different viewing tasks for near and distant targets.
- (2) The design and evaluation of smartwatch-based interaction techniques that allow low vision users to adjust the enhancement methods in ForeSee. More generally, they are the first set of HMD interaction techniques that are private, accessible, and do not require specialized hardware.

We note that this article is an expanded version of a conference paper published at ASSETS 2015 [119]. In the original paper, we presented ForeSee and described our study of low vision people’s experience with different enhancement methods (RQ1). However, we only presented two of the four tasks we present here. In the current article, we present additional findings from the two reading tasks and findings from another two tasks that involved image recognition (near- and far-distance). *Our work on interaction techniques (RQ2) is new and published here for the first time.*

2 BACKGROUND AND RELATED WORK

2.1 What is Low Vision?

Low vision has different definitions. Some organizations use visual acuity, a measure of the sharpness of vision, to define low vision. For example, the World Health Organization (WHO) defines low vision as a visual impairment where visual acuity is less than 20/60 in the better eye with best correction [112]. In layman's terms, this means that if a person with typical vision can read a sign from 60 feet away, a person with low vision needs to be 20 feet away to read that sign. The term "best correction" simply refers to one's visual ability with glasses or contact lenses. The International Classification of Disease, Clinical Modification (ICD-10-CM) coding system uses the same visual acuity threshold for low vision [45]. Medicare carriers also regard this as the eligibility criterion for the coverage of low vision services [68]. Meanwhile, different visual acuity thresholds are used by other organizations to define low vision, such as 20/40 based on the ability to obtain an unrestricted driver's license [105], and 20/70, which is used by Medicare to establish medical necessity for low vision rehabilitation [88].

While many low vision definitions are based on visual acuity, Massof and Lidoff believe that low vision must be defined in terms of function but not quantitative test results, incorporating aspects of visual function that cannot be captured by the measures of visual acuity such as contrast sensitivity, glare, and visual field loss [60]. They defined low vision as "uncorrectable vision loss that interferes with daily activities." Similarly, in Foundations of Low Vision, Corn and Erin defined a person with low vision as "a person who has measurable vision but has difficulty accomplishing or cannot accomplish visual tasks, even with prescribed corrective lenses, but who can enhance his or her ability to accomplish these tasks with the use of compensatory visual strategies, low vision devices, and environmental modifications" [21]. This definition better reflects the impact of visual impairments on people's daily lives and emphasizes the way in which low vision people use their vision in visual tasks. In this article, we followed this functional definition of low vision from Corn and Erin, because our research focuses on qualitatively understanding low vision people's visual experiences with and without vision enhancement technology.

2.2 Conventional Low Vision Enhancement Methods and Devices

People with low vision use a variety of optical and digital tools to access information in the physical world. For seeing nearby objects or text, typical low vision aids include handheld optical magnifiers, handheld or head-mounted telemicroscopes, and reading glasses (ultra-high prescription glasses) [59, 90]. For seeing objects at a distance, people can use handheld monocular telescopes or Bioptic telescopes that are mounted on a pair of glasses [43]. People with low vision also use portable (or handheld) or stationary Closed Circuit Televisions (CCTVs) [56, 81, 113] that magnify a video feed captured by a camera onto a digital display. Like most assistive technology, CCTVs are expensive, ranging from about \$600 for a handheld device to several thousands of dollars for a stationary device.

In addition to low vision aids, people with low vision often adjust their surroundings to improve access to information. Optometrists and low organizations [109] recommend adjusting the light indoors (some magnifiers include built-in lights as well), and move closer to objects to improve visibility [21]. For printed materials, they recommend using bold large high-contrasting fonts and dark pens for writing [3, 21]. Some people may also use nonvisual materials, such as audiobooks or text-to-speech, especially for extended reading tasks, which may be more efficient and avoid eye fatigue [90].

For accessing digital information, software applications create similar effects to the tools described above. Many mainstream platforms such as Mac OS and Windows include built-in

magnifier, text-to-speech applications, and adjustable contrast and color settings [30]. There are also standalone software applications such as Magic [58] and ZoomText [94] that provide the same basic functionality as built-in applications but with additional customization options. The iOS and Android mobile platforms also offer similar built-in accessibility tools [33, 41]. To our knowledge, no recent published research quantified the prevalence or anecdotal use patterns of these technologies.

2.3 Low Vision Accessibility Research

Even though the vast majority of people with visual impairments have functional vision, very little research has (1) studied the challenges low vision people face and (2) produced innovative vision enhancement tools. Jacko and Sears [46] advocated in 1998 for the importance of research on designing systems for low vision people, emphasizing the lack of innovation in the area. However, today's common low vision tools such as CCTVs, optical and screen (software) magnifiers remain largely the same as they were 17 years ago. In this section, we highlight accessibility research that specifically targeted users with functional vision, as opposed to the vast array of research that explores eyes-free technology.

A few studies have shed some light on the challenges that low vision people experience in daily activities. Cimarolli et al. [13] conducted in-person interviews with 364 older adults with significant vision loss due to age-related macular degeneration to discover what challenges they face in reading and outdoor mobility in functional, social, and psychological domains. They interviewed each participant three times during the 2-year period and found that the vision-related challenges may change overtime: functional challenges increased, social challenges remained relatively stable, while psychological challenges decreased. Cimarolli et al. further investigated the social challenges of working age people [12, 14–17] and found that overprotecting and underestimating people's abilities caused stress and increased their disability. Recently, Szpiro et al. [99] conducted contextual inquiry with 11 low vision participants in wayfinding and grocery-shopping tasks to understand their challenges and strategies in this process. They found that low vision people struggled but still preferred using their vision during navigation. Meanwhile, shopping was overwhelming with many challenges for low vision participants, since there was little to no technology they can use for indoor navigation and product search. Zhao et al. [118] observed 14 low vision participants navigating environments with surface level changes (e.g., stairs, street curbs). They revealed the distinct needs of low vision people from blind people in such tasks, emphasizing the lack of navigation tools that facilitate safe navigation for people with low vision.

Some researchers focused on the challenges that low vision people face when using technologies. In 2009, Kane et al. [49] conducted a diary study with people with vision and motor impairments, four of whom had low vision, to discover patterns and challenges of technology used on-the-go. They found that low vision participants faced challenges using mobile devices resulting from different environmental conditions, health and ability changes, and device maintenance, and revealed the strategies adopted such as holding the devices in unconventional ways and using other devices such as handheld magnifiers to access their mobile devices in everyday tasks. Theofanos et al. [103] conducted a user study with participants with different vision problems who use software to magnify computer text, and described several challenges low vision people face when navigating websites: most critically they found that participants missed crucial information and context when navigating in magnified displays. Moreover, Szpiro et al. [100] conducted a user study with 11 low vision people, observing how they used computing devices such as desktop computers and smartphones. They found that although low vision participants used various accessibility tools, they had to constantly adjust the settings to read different visual contents

comfortably. The current accessibility tools did not provide enough control and the interaction was difficult and inefficient.

Several research efforts have attempted to improve the accessibility of computing tasks such as reading documents or browsing the web. Bernard et al. presented Navisio [7], a software tool that magnified selected regions in a PDF document (as opposed to the entire document) to facilitate navigation and increase reading comfort; a user study confirmed that reading comfort and speed were higher with Navisio than with a CCTV. Bigham presented an approach to improve website accessibility [8] that magnified elements in a webpage in an optimal way, without causing negative side effects like scrolling and element overlap. Bigham's approach was not evaluated in a user study, and the mean magnification level achieved was only 1.6 times the original size. Researchers have also addressed the challenge of using a mouse, since a cursor can be difficult for low vision people to see. Kline and Glinert [51] proposed a visualization of cartoon eyes that indicated the direction and distance to the mouse cursor, and Fraser and Gutwin [31] presented a framework of attributes such as sound or tactile feedback to improve mouse cursor accessibility. In addition to the traditional user interfaces on desktop or laptop displays, Zhao et al. [116] addressed the accessibility of virtual reality for people with low vision. They designed SeeingVR, a set of 14 low vision tools that augment VR scenes with both visual and auditory feedback.

While there has not been research, to our knowledge, to improve the accessibility of mobile device interfaces, researchers have used mobile devices to improve access to a classroom environment. Kushalnagar et al. [53] and Hayden et al. [36–38] developed systems that provide low vision students magnified views of lecture visuals on their personal smartphones or tablets. Surprisingly, we have not found any research that attempts to improve access to the physical environment in other settings using mainstream mobile devices. In the following section, we discuss research that specifically involves HMDs, mostly with custom hardware, to improve access to the user's physical environment.

2.4 Head-mounted Systems for People with Visual Disabilities

The development of cameras, head-mounted displays, and image processing technology created opportunities for low vision devices that enhance people's vision directly.

Some researchers have focused on correcting specific low vision conditions. Optometrists have contributed several head-mounted systems and algorithms to enhance the vision of people with different eye conditions [54, 57, 76–79, 87, 107]. For example, Peli [76] presented a head-mounted system with a one-dimensional analog video processing algorithm that presented more detailed real-time gray-scale video. Using a one-dimensional algorithm reduced the cost, weight, and power consumption of the device compared with a two-dimensional algorithm. Peli and Wood evaluated this system [79] with 19 participants with central field loss who were tasked with recognizing celebrity faces. Forty-two percent of the participants showed significant improvement with this system. Their lab also developed a method for optical see-through displays that aimed to expand people's field of view by overlaying the contours of a minified view of a wide field over the center of the user's original vision [107]. They evaluated their method with (only) two participants with Retinitis Pigmentosa who completed visual field tests with this enhancement method on head-mounted displays. They found that participants' field of view was effectively expanded. To our knowledge, this system was not evaluated with real-world tasks.

Instead of developing custom hardware platforms, researchers have recently explored commercial smartglasses as low vision tools. For example, Zhao et al. [117] conducted an exploratory study with 20 low vision participants to understand the feasibility of commercially available smartglasses (Epson's Moverio BT-200) for low vision users in both stationary and mobile situations. The results showed that this mainstream smartglass platform could be a powerful low vision tool, because low

vision people identified shapes and short phrases in certain sizes and colors on the smartglasses. Stearns et al. designed different AR magnification methods on the Microsoft HoloLens [95, 96]. They also used a handheld smartphone and a finger-worn camera as input devices to capture text in a reading scenario. Researchers also created enhancement tools for Google Glass [34]. For example, Tanuwidjaja et al. developed Chroma [101], a system that projected a filtered image on the Google Glass display that highlighted or shifted certain colors to make them discernable to a user with color blindness. They evaluated this system with 23 colorblind participants who performed color identification tasks in different contexts. In over half the tasks, Chroma improved participants' performance. Hwang and Peli implemented two edge-enhancement algorithms based on positive and negative Laplacian filters on Google Glass to increase contrast for people with age-related macular degeneration (AMD) [44]. They conducted a preliminary study with three sighted participants by measuring their contrast sensitivity, finding that their contrast sensitivity improved to a limited degree when using the system.

While some have focused on compensating for certain eye conditions, other researchers aimed to address challenges for low vision people in specific tasks, such as visual search [120], mobility [27, 39, 50, 106], and reading [42, 64]. For example, Zhao et al. designed CueSee [120], an HMD system that facilitates the process of searching for a product in a grocery store aisle. CueSee used computer vision to automatically recognize the target product and rendered visual cues to direct the user's attention to the target. Everingham et al. [27] developed a neural-network classification algorithm for a head-mounted device that segmented scenes rendered in front of users' eyes and recolored objects to make obstacles more visible. Hicks et al. [39] and Rheede et al. [106] built a real-time head-mounted LED display with a depth camera to aid navigation by detecting the distance to nearby objects and changing the brightness of objects to indicate their distances. In addition to visual search and mobility, some research systems and products have been developed to assist low vision people with reading. For example, Huang et al. [42] designed a sign-reading system on the HoloLens, which identified real-world text and visually enhanced the text by rewriting it in a larger and higher contrast font to enable low vision users to read it. Merino-Gracia et al. [64] proposed a real-time text detection algorithm for a head-mounted device with audio feedback that aimed to help blind people read. However, only detection performance on a particular image dataset was measured, and no user study was conducted to evaluate its effect on blind people. OrCam [72] is a similar product on the market, currently selling for \$3,500. It is a portable device with a camera but no display that is mounted on the frame of the user's eyeglasses. OrCam recognizes text and products and provides audio feedback with a bone-conducting earpiece. To our knowledge, no studies have evaluated the effectiveness of OrCam. Instead of using computer vision to recognize the environment, Aira [4] is a pair of glasses with a camera, which captures the environment and sends the video feed to a remote human agent who describes the visual information for the user.

Even though different head-mounted systems and enhancement methods were developed, they were mostly designed for specific low vision conditions or scenarios. Little research has designed systems that are customizable and can be widely used cross different visual conditions in various tasks. Massof et al. designed the first Low Vision Enhancement System (LVES), an HMD with three video cameras, on which image enhancement algorithms (e.g., contrast enhancement via reemphasis spatial filtering, pixel remapping for central scotomas) could be implemented to improve low vision people's visual abilities [61, 62]. Jordy [108] was a similar HMD that cost \$3,620. However, these early HMD devices have faded from the marketplace because of the limitations of both the processing systems and the hardware [23]. Recently, a commercial product called eSight [26] came to market. The eSight is a custom hardware head-mounted video see-through display that costs \$15,000, supporting magnification, contrast enhancement, and color adjustment of the real-time video-feed. This device is connected to an additional controller via wires so that a user can adjust

the enhancements with hardware buttons and switches. Zolyomi et al. [122] interviewed 13 eSight users to understand how the eSight mediated participants' daily social experiences. They discussed participants' social and emotional experiences with eSight and its effect on their social engagements and daily activities. However, no study evaluated its functionality, such as the effectiveness of different enhancements for different low vision users, or how effective this tool was in facilitating different daily tasks.

In this article, we describe a study of different vision enhancements on an HMD with 20 low vision people and the design and evaluation of two sets of interaction techniques to customize the vision enhancements. To our knowledge, no prior work has addressed the interaction techniques on HMDs for people with low vision.

Our study on vision enhancements on HMDs also differed from most prior research. Previous work on head-mounted vision enhancement systems (and vision enhancement systems in general) mostly focused on quantitative evaluations where measures like speed and accuracy were measured on specific tasks (e.g., References [27, 57, 104, 110]), with no consideration of the user's preference and experience. However, studies on assistive technology abandonment have repeatedly shown that people are more likely to abandon assistive technology that is prescribed without considering their subjective experiences and "feelings" [24, 40, 82, 111]. Meanwhile, assistive technology is more likely to be used continually if the user feels that it is socially acceptable, minimizes stigma [74, 91], and supports long-term life goals [12, 17]. Different from prior work, our evaluation on vision enhancement methods focused on users' qualitative experiences. With qualitative interviews and observations, we explored low vision people's experiences with different enhancement methods and display modes in an HMD context, and how they customize the enhancements according to their visual conditions and different viewing tasks. Based on our findings, we discuss the challenges, research opportunities, and design recommendations for head-mounted vision enhancement systems for low vision accessibility.

3 FORESEE

ForeSee, first presented in our previous publication [119], is a head-mounted device that enhances people's vision through a video see-through display. It has five vision enhancement methods and two display modes, allowing people with low vision to adjust their visual experience according to their visual condition and viewing task. In this section, we describe the design of ForeSee.

3.1 Design Guidelines

We developed ForeSee according to the following guidelines:

DG1. Enhance users' vision directly: We wanted to develop a system that enables people with low vision to use their residual vision for performing daily tasks. Meanwhile, we sought to seize an opportunity to use computer vision technology, which is able to perform various real-time transformations to video streams, to enhance people's vision directly in all their daily living activities, without being highly task specific like prior work (e.g., Reference [35]).

DG2. Use standard enhancement methods: As a starting point, we wanted to use standard enhancement methods that were widely used and accepted by low vision professionals. Users could adjust to the new platform more easily when using familiar enhancement methods.

DG3. Enable customization: We wanted our system to be customizable to suit users' needs and preferences. This is especially critical for a technology that targets low vision people, since there are many different kinds and degrees of vision impairments [2]. Moreover, people's visual abilities frequently change overtime or in different settings [66] (e.g., a sunny sidewalk, a dimly lit restaurant, and an office cubicle). Thus, customization is not only important to accommodate different users but also to maximize the technology's effectiveness for any particular user.

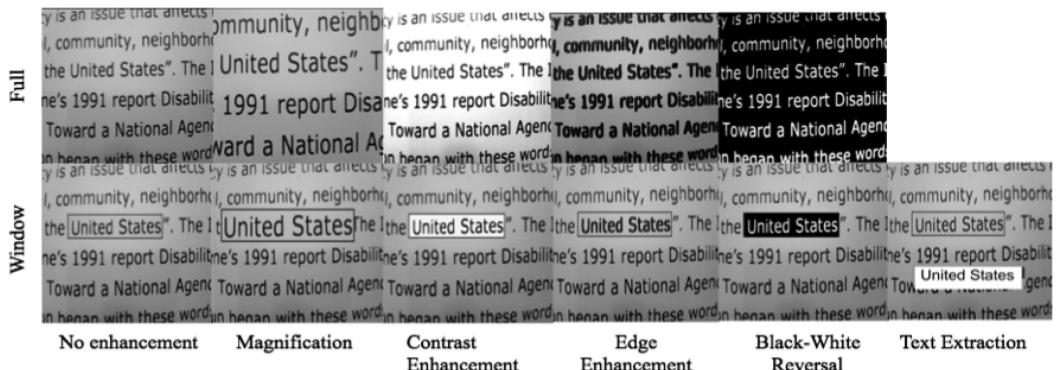


Fig. 2. The five enhancement methods in ForeSee: Magnification, Contrast, Edge Enhancement, and Black/White Reversal; and the two display modes: Full Display Mode and Window Display Mode in a near-distance text reading task. While the Window Display Mode includes only two words in this example, users can adjust the size and the position of the window to enhance any part of the display.

DG4. Enable fast and easy interaction: As mentioned in DG3, many people's visual abilities vary in different environments so they may need to adjust the system often. Moreover, people may need different enhancement settings for different tasks. For example, when looking for a friend in a crowd, a user may use a lower level of magnification than when reading a newspaper, since she does not want to lose an overview of the environment. As such it is not only imperative for the system to be customizable, but the user must be able to efficiently interact with the system in different situations to customize it.

3.2 System Description

ForeSee is a video see-through [6] augmented reality head-mounted device that is designed to work with commodity hardware. It consists of a display, a camera, and a processor. ForeSee captures a video feed of the user's surroundings from the user's point of view, sends the feed to the processor, and then displays the processed feed on the display. There is only one camera, so the user sees the world in two dimensions (in the future, we may experiment with using two cameras to achieve a stereo 3D effect). Unlike Google Glass, the display covers the user's entire field of view. This allows us to enhance the user's field of view directly, following DG1. The processor enhances the image in various ways, as described below.

ForeSee enhances the user's vision by applying some combination of *enhancement methods* in different *display modes*. An enhancement method represents a transformation that is applied to the captured video feed. A display mode represents a way of incorporating the transformed content into the original captured video. ForeSee includes five enhancement methods and two display modes. Following DG2, we developed the enhancement methods and the display modes by examining commonly-used low vision aids and techniques for improving the visibility of printed and digital materials for people with low vision, as discussed in Section 2.

We created two display modes (Figure 2): Full Display Mode and Window Display Mode.

Full Display Mode. In this mode, we simply apply enhancement methods to the user's entire field of view. On a video see-through platform, the enhancements are rendered on the entire frame of the video feed.

Window Display Mode. When the user selects the Window Display Mode, a rectangular black-border window overlay appears on the original video feed, and all enhancement methods are only applied to this rectangular area. The user can change the width, height, and position of the

window to enhance a specific area of her visual field. We modeled the Window Display Mode after a handheld magnifier. It also resembles the “lens” mode in many screen magnifiers [58, 94] where a rectangular area around the cursor is magnified.

We created five enhancement methods (Figure 2):

Magnification. Since magnification is the most common vision enhancement method on both optical and digital devices, we created a Magnification Enhancement Method where the user can adjust the level of magnification.

Contrast. In this method, we increase the luminance contrast of the video feed but maintain the color hues. We used a standard contrast enhancement algorithm that can be found in [98]. This enhancement is widely used in low vision aids such as MAGic [58] and CCTVs [11].

Edge Enhancement. This method darkens and thickens the edges (or contours) in the video feed. We used Canny’s edge detection algorithm [10] to detect the edges in an image and then darkened the pixels on and surrounding the edges in the original image. This method is adopted by several previous head-mounted vision enhancement research [44, 54, 107].

Black/White Reversal. In this enhancement method, we used Otsu’s Thresholding Method [73] to convert the video to binary (only black and white) with the threshold of 100, and then reversed the colors. People often use this type of effect (white text on a black background) when reading with CCTVs [11].

Text Extraction. Text extraction, which involves optical character recognition (OCR), is used by several low vision accessibility tools such as OrCam [72] and screen readers (e.g., JAWS [89]). It is always combined with text-to-speech technology, providing audio feedback to visually impaired users. To explore the effectiveness of this method in an HMD setting, we use visual feedback instead of audio. In ForeSee, this method is only available in the Window Display Mode. It detects [69] and recognizes [102] the text in the window and displays a digital version in a standard font with high contrast in a separate region on the display. Currently, we only allow users to select 1–3 words at a time to make sure that the size of the digital text is big enough for most participants.

We address DG4 in Sections 5 and 6, where we discuss the design and evaluation of input methods to customize the enhancement methods.

4 EVALUATING VISION ENHANCEMENT METHODS

We conducted an extensive user study to address our first research question, which we posed in the Introduction:

RQ1. How do different visual enhancements on an HMD affect low vision people’s visual experiences?

We operationalized this question by breaking it down into the following two questions:

- (1) How effective was each enhancement method and display mode in ForeSee for people with different visual abilities in different viewing tasks?
- (2) How did people with low vision customize their visual experiences (select and combine vision enhancement methods and display modes) in different viewing tasks when using ForeSee?

4.1 Method

4.1.1 Participants. We recruited 20 people with low vision to participate in our study (13 females, 7 males) with a mean age of 46 (range: 21–68). In our recruiting materials, we called for volunteers who had “low vision” (see definition in Section 2.1). When volunteers contacted us, we conducted a brief screening interview over the phone to ensure they were indeed low vision. If

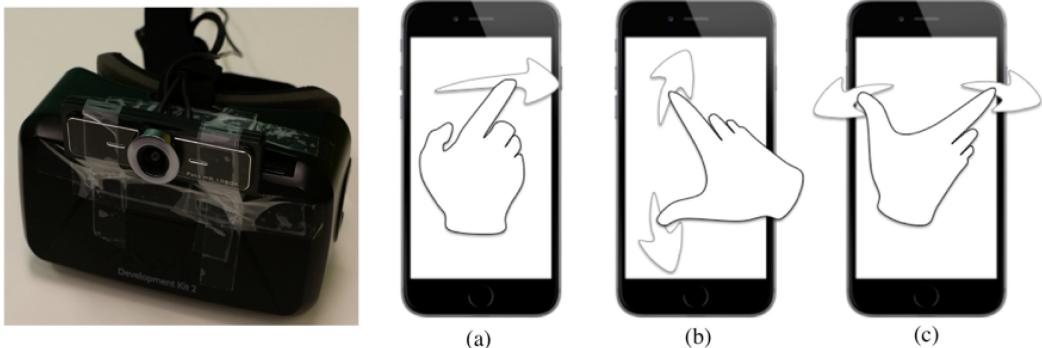


Fig. 3. Left: The ForeSee HMD prototype. Right: The gestures to control the window in the Window Display Mode: (a) dragging gesture to adjust the position of the window, (b) pinching in/out gesture in the vertical direction to adjust the height of the window, (c) pinching in/out gesture in the horizontal direction to adjust the width of the window.

a volunteer said they used tools that enhanced their vision such as handheld magnifiers, screen magnifiers, or a CCTV, then we had enough evidence that they had low vision and invited them to participate. Volunteers who only used screen readers were not considered suitable, since they likely had little to no functional vision.¹ We considered people who are legally blind² as eligible participants as long as they can use their vision to complete some visual tasks in daily activities (e.g., using a magnifier to read some text). Our participants had a variety of vision conditions, shown in Table 1. Participants were compensated \$20 per hour and received up to \$60 in reimbursement for travel expenses.

We were unable to conduct the whole study with one participant, P3, who arrived late and did not understand the study instructions. He completed only the near-distance task for image recognition. We only report his data of this one task in the Findings Section (4.2).

4.1.2 Apparatus. We implemented ForeSee on an Oculus Rift that we combined with a webcam (Figure 3 Left). We used a WideCam F100, a 1080HD webcam with a 120 Degree ultra wide angle lens to capture the surrounding environment, and an Oculus Rift DK2 as the display. The webcam and Oculus Rift were both connected to a laptop that was used as the processor of the vision enhancement system. The webcam was attached to the front of the Oculus Rift between the wearer's eyes. The two screens of Oculus rendered the same images of the enhanced environment (biocular vision). Although the webcam has a high resolution, we reduced the image resolution to 480×640 to make the system real time.

To enable the participants to customize the enhancements on ForeSee in the user study, we developed a smartphone application as a controller, with which users could control the size and the position of the window in the Window Display Mode using eyes-free gestures. Users can regard the smartphone touchscreen as a touchpad, moving the position and adjusting the size of the window using touch gestures (Figure 3, right). To move the Window, the user can conduct a dragging gesture as if she was dragging the Window; the Window will move towards the same direction on the HMD display with the same distance. The dragging gesture can start at any point on the smartphone screen, so that the user does not need to look at the phone. The user can also use a

¹Functional vision is used to describe how well the person functions in vision-related activities. It refers to how the person functions, which is distinct from visual functions (how the eye functions) [18, 19].

²A person is “legally blind” if either (1) his or her best-corrected visual acuity in their better eye is 20/200 or worse, or (2) his or her visual field is 20 degrees or narrower. Some people with low vision are legally blind.

Table 1. Demographic Information of the 20 Participants Who Evaluated the Vision Enhancements

ID	Age/ Sex	Diagnosis	Visual Field	Visual Acuity	Color Vision	Tool Used
P1	23/M	Nystagmus	Full	20/120	Good	Bifocal, telescope, software magnifier, zoom-in on computer
P2	29/F	Optic Atrophy	Full	20/400	Could distinguish primary colors	Cane, black/white reversal on computer
P3	62/M	Retinopathy of prematurity	Very limited, a little vision on the right side in the right eye	Left: no vision; Right: <20/200	Nearly cannot distinguish colors	Jaws, Handheld/ stationary CCTVs, reading glasses, monocular, ZoomText and Magic
P4	45/F	Pituitary gland optic nerve tumor	Limited	Unknown, but good	Good	ZoomText, CCTV, magnifier for brief reading tasks
P5	53/M	Stargardt's Disease	Full	Left: 20/800; Right: 20/600	Have color vision but not good at subtle colors	CCTV, handheld magnifiers, handheld telescopes
P6	21/F	Leber's Congenital Amaurosis	Very limited	20/600	Have color vision but not good at subtle colors	Braille notetaker, Jaws, iPhone with VoiceOver, zoom-in to see images
P7	34/F	Retinitis Pigmentosa	Limited; <30 degrees	Unknown, but good	Good	ZoomText, cane at night
P8	34/F	Pathological Myopia; Detached retina in left eye; Cataract	Full	20/200	Have color vision but not good at subtle colors	Screen magnifiers for phone and tablet, speech input and Siri, ZoomText, color reversal
P9	56/M	Bilateral optic atrophy; Nystagmus	Full	20/220	Have color vision but not good at subtle colors	Screen magnifier, black/white reverse
P10	57/M	Usher's syndrome, Retinitis Pigmentosa	Very limited, <5 degrees	Left: <20/170; Right: blind	Need high contrast to see colors	CCTV, color reversal
P11	34/F	Retinitis Pigmentosa	Limited	20/60	Color blind	Magnifier, flash light at night
P12	40/M	Detached retina	Full	Left: 20/60; Right: blind	Have color vision but not good at subtle colors	Reading glasses, cane, iPhone with VoiceOver, Laptop with Zoom, CCTV
P13	65/M	Central areolar choroidal sclerosis	Unknown	Unknown	Need high contrast to see colors	CCTV, ZoomText
P14	58/F	Retinopathy of prematurity	Full	Left: 20/400; Right: 20/300	Need high contrast to see colors	Handheld monocular, CCTV, handheld magnifier, ZoomText
P15	51/F	Cytomegalovirus Retinitis	Very limited	Left: unknown; Right: blind	Color blind	Jaws, ZoomText
P16	64/F	Retrolental Fibroplasia	Very limited	Left: blind; Right: unknown	Have color vision but not good at subtle colors	Cane, screen readers, used to use magnification
P17	55/F	Steven Johnson Syndrome	Very limited	Left: 20/150; Right: 20/200	Good	Magnifier, CCTV
P18	23/M	Stargardt's Disease	Full	Limited, but unknown	More issues with contrast	Screen magnifier, color reversal
P19	36/M	Reversal class of Retinitis Pigmentosa	Full	20/300	Cannot distinguish colors of low contrast	Screen magnifiers, screen readers, color reversal
P20	68/F	Retinal myopathy, glaucoma	Left: limited; Right: almost none	Left: limited; Right: only has light perception	Cannot distinguish colors of low contrast	ZoomText, CCTV, color reversal, handheld magnifiers, iPhone with Zoom and Siri

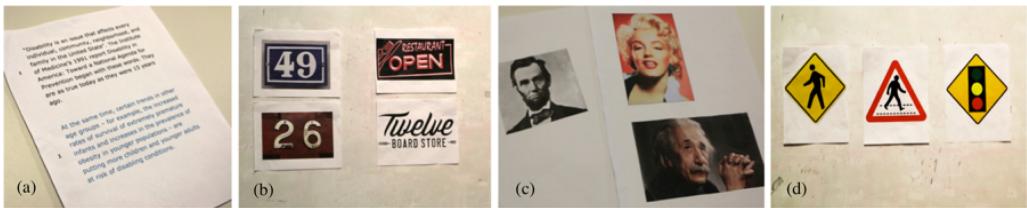


Fig. 4. Experiment Materials: (a) a handheld printed page, (b) four printed signs of numbers or words hung on the wall, (c) three handheld celebrity photos, and (d) three traffic signs hung on the wall.

pinching in or pinching out gesture to increase or reduce the size of the window. We also used Wizard of Oz (WOZ) [85] to allow users to choose and adjust different enhancement methods with natural speech commands. The researcher, acting as the wizard, responded to the commands and adjusted parameters of the system. For example, the user can say, “move the window down,” to adjust the window position in the Window Display Mode, or “make it larger,” to increase the magnification level when using Magnification.

4.1.3 Procedure and Materials. The study consisted of one session that lasted two to four hours (about 2.5 hours on average). Participants sat on a chair inside a well-lit office throughout the study. We offered them water and coffee and allowed them to take breaks in between tasks for as long as they needed to reduce the impact of fatigue on their performance. We began each session by conducting a short interview, where we asked participants for their demographic information, including age, gender, and visual condition. We then introduced ForeSee, explaining its functionality and the goals of the study to get the participants familiar with the system as well as the study procedure.

We then asked participants to perform four viewing tasks, including (1) near-distance reading: reading a handheld printed page, (2) far-distance reading: reading signs with words and numbers, (3) near-distance image perception: recognizing celebrity photos, and (4) far-distance image perception: recognizing traffic signs. We chose these tasks because they represent important daily living challenges for low vision people [13]. We described the two reading tasks in our conference paper [119], but here we additionally present the results from the image perception tasks.

The experiment materials we used included: a printed page with two paragraphs of black or blue letters in Verdana, size 20 (Figure 4(a)), four signs of letters or numbers printed in a decorative font with different colors hung three meters away right in front of the participants (Figure 4(b)), three photos with one black and white photo of Abraham Lincoln and two color photos of Marilyn Monroe and Albert Einstein (Figure 4(c)), as well as three printed traffic signs hung three meters away in front of the participants (Figure 4(d)).

For each task, we asked participants about their visual experience under different conditions. Specifically, we asked them to describe what they saw, whether the current condition was effective, and their comfort level (i.e., whether they experienced any eyestrain). Participants were also asked to describe the image content or read the text. In total, we evaluated 12 conditions for each task:

- (1) Completing the task without using ForeSee.
- (2) Using ForeSee but without any enhancement methods.
- (3) Using ForeSee in the Full Display Mode, with all enhancement methods, except for Text Extraction, which can only be used in the Window Display Mode. Conditions 3a-3d correspond to using magnification, contrast, edge enhancement, and black/white reversal in the Full Display Mode, respectively.

- (4) Using ForeSee in the Window Display Mode, with each of the five enhancement methods. Conditions 4a–4e correspond to using magnification, contrast edge enhancement, black/white reversal, and text extraction in the Window Display Mode, respectively.
- (5) Using ForeSee with whatever combination of enhancement methods and display mode worked best for the participant in the current task. The user was able to customize the system to optimize their visual experience.

To avoid order effects, we counterbalanced the order of some conditions and randomized the order of others as follows. For the tasks, we counterbalanced the order based on distance (near vs. far distance). And for each distance level, we randomized the order of reading and image perception tasks. For the enhancement conditions, we first counterbalanced conditions 1 and 2. We then presented the enhancement methods in a random order (conditions 3a–3d and 4a–4e), since there were too many enhancement methods to properly counterbalance them. We always presented condition 5 at the end, because participants had to be familiar with all the enhancement methods and both display modes to choose their preferred combination.

We connected ForeSee to the researcher’s laptop, mirroring the display of the Oculus on the laptop. Since the participants may have difficulty locating the target objects, especially for the far-distance tasks, the researcher provided guidance to help the participants look in the right direction, ensuring the targets were showed on the display.

4.1.4 Analysis. We video-recorded the study and took notes during each session. After the study, we transcribed the videos using a professional service and analyzed the transcripts following the general coding method described by Burnard [9], organizing them by tasks and enhancement conditions. Two researchers discussed the themes and categories of the data together, while one of them coded the interview data by labeling important words and sentences with corresponding themes.

4.2 Findings

In this section, we report the benefits and weaknesses of different vision enhancement methods and display modes, and how participants with different visual abilities selected and combined different enhancement methods with ForeSee to create the best visual experience in different viewing tasks.

4.2.1 User Fit. ForeSee is designed to enhance users’ vision directly, so it requires some amount of functional vision. During the study, we found that some low vision people did not have enough functional vision, which prevented them from benefiting from ForeSee as much as other users could. Two participants (P6 and P16) could not complete the evaluation in some viewing tasks. Both of them only finished the near-distance image task and could not perceive much information from the materials in the other three tasks.

We found that both P6 and P16 had difficulty locating the target objects in their field of view. In the near-distance reading task, when we asked P16 about what she saw on the reading material, she said, “It’s just white paper. No, I can’t tell you what’s on this paper.” At the beginning of the far-distance task, P6 also said, “Honestly, what I see right now is just the white.”

Moreover, both P6 and P16 could not distinguish among the different enhancement methods. For example, when we asked P16 to describe the differences between the magnified and the original text in the near-reading task, her feedback didn’t relate to the magnification. P16 described her visual experience: “This glasses have some light shining through them.”

4.2.2 Video See-Through View (No Enhancement). We compared participants’ original vision (condition 1 in the study) with their vision through the digital display without any enhancement (condition 2). Considering the limited focal range of the webcam and the low resolution of the

Oculus display, we expected the unenhanced view to be blurry and distorted. However, some participants preferred the video see-through display to their original vision.

In the near-distance tasks, most participants (16 in the image task and 9 in the reading task) indicated that using ForeSee without enhancement improved their visual experience. When perceiving images, the participants (e.g., P2 and P7) felt the images became sharper and could see more colors and details. “I can’t tell the color of this image [Einstein’s photo], but with these [glasses], I can tell there is more, I can see more skin tone” (P2), and “It’s sharper actually with the glasses. I can actually see the eyes of Einstein there. It’s much sharper, and I can see the hair on the top of his head” (P7). When reading text, some participants (e.g., P13, P15) who could not read the materials with their original vision were able to read with ForeSee. “I can’t read standard printed material without magnification. With [ForeSee], I can see an individual word” (P15). Moreover, some participants (e.g., P12, P17) felt ForeSee increased the distance at which they could comfortably hold the page. “The difference is that I really have to look very close [without ForeSee], so if I had [ForeSee] I wouldn’t have to look this close” (P17).

In the far-distance tasks, most participants preferred their original vision because of the limited resolution of the prototype. However, there were still some participants (five in the image task and three in the reading task) who believed that the visual experience with ForeSee was better, because they felt the view in ForeSee was brighter. As P7 mentioned, “It’s sharper in the sense of light.” P9 also believed that the images from ForeSee were more focused than his original vision, “[My original vision] is just less focused and blurrier.”

Based on the participants’ feedback, we speculate three possible reasons for why people prefer seeing contents on digital displays. First, the brightness of the digital displays provided higher luminance contrast. Second, the Oculus Rift itself has lenses between the digital display and the users’ eyes, which may have helped the participants to focus, especially for those who did not have lenses in their eyes (e.g., P17). Third, the headset also blocked some elements in the real environment (e.g., natural light, stimuli in the periphery), which may reduce distractions and help people focus on the displays.

4.2.3 Perceived Effectiveness of Vision Enhancements. We analyze participants’ perceived effectiveness of different enhancement methods in each display mode in the four viewing tasks. For each enhancement, we organize the findings by display modes or tasks (reading vs. image), according to which factor has a greater impact on participants’ experience.

(1) Magnification

Full Display Mode. Most participants felt magnification was effective: 17 participants in the image task (including P3 who only finished near-distance image task), and 15 participants in the reading task. Many remarked that magnification enabled them to see more colors and details. “I certainly see more detail than before. Now the color [of the top left sign] is green or blue, like a light green or blue” (P19). Even though the magnified video has a lower resolution than the original video, participants felt they perceived more information and were better able to distinguish details and read. “Now that it’s bigger, it’s blurry but I can see more” (P2).

While providing more detailed information, magnification had some disadvantages. First, it reduced users’ visual field and made it difficult for them to locate targets. As P15 indicated, “If it’s too big, you feel it’s losing the orientation, then you have to work harder to look around.” Second, it reduced the speed at which participants perceived the content because of the loss of the context, especially for reading long text:

“So much of reading is recognizing general forms and not necessarily seeing every letter, so [magnification] is not the most [effective enhancement]. Instead of actually

looking at every single letter or every single syllable. I'm doing a lot from context."
(P12)

The loss in visual field negatively impacted people with tunnel vision even more. P10 insisted that magnification did not improve her vision in any tasks and P4 did not prefer magnification in reading tasks: "Everything that I look at cuts in half, so although they are becoming clearer, they're cut in half for me" (P4).

Window Display Mode. In the near-distance tasks, participants did not like as much magnification in the Window mode as they did in the Full mode: Only nine participants preferred magnification for viewing images and eight for reading. One reason is that the size of the window was limited, which reduced the users' field of view more. For the near-distance image, some participants emphasized the difficulty of understanding the whole picture (e.g., P5, P9). "I think [perceiving] images is much more right brain and you need things not to be chopped up so much. You need the whole 'gestalt' of the thing, you know? The whole [image]" (P5). And for the reading materials, the window made it hard to track the words. "Because I'm looking inside the box... You could miss a sentence, you could skip over a line" (P10).

In the far-distance tasks targets appeared smaller, which reduced target "cut-offs" and "losing track." This led to 10 participants finding magnification useful in the image recognition task and 11 participants finding magnification useful in the reading task.

Participants tended to modify the size of the window as they adjusted the magnification level, increasing the window to keep the target in the magnified view. Some suggested that the window should automatically increase as the magnification level increases to keep the target in the window. P9 suggested: "The magnification is helping but you need to magnify the border [of the window] with it, because once you magnify and if you don't [enlarge the window], what you see inside the border decreases."

(2) Contrast Enhancement

Full Display Mode. In the near-distance tasks, most participants (15 in the image task and 16 in the reading task) felt that contrast enhancement was beneficial. They noted that the contrast enhancement emphasized the color and increased the sharpness: "It's brighter. It's sharper. That makes it easier to recognize what it is right away" (P17). Some participants also felt less eyestrain, because the targets became clearer. "It is easier to see, because I find it less of an eye strain... I can see the people clearer" (P14). However, in the far distance conditions, only six participants in the image task and seven in the reading task preferred contrast enhancement. As the targets are far away and some parts are too small or thin for the webcam, the contrast enhancement makes the contours more invisible, especially the lines between two similar colors: "The contrast has helped with the colors, but it has hurt the shape. I can tell that there's a sign there more but I can't quite tell what kind of sign it is" (P1).

Some participants (e.g., P7, P14) mentioned the problem of distortion aroused by contrast enhancement. Even though the colors are enhanced, the image is not as natural as original. As P7 commented, "It just doesn't feel like the colors are real." Contrast enhancement also had a negative effect on some participants' visual comfort, especially those who were sensitive to light (e.g., P2, P4, P11). They felt that the brightness hurt their eyes and distracted them from focusing on the target. Some participants even had more eyestrain with a higher contrast. "I couldn't read like that, because the glare is too much for my eyes" (P4).

Window Display Mode. Contrast enhancement was effective for more participants in the Window mode than in the Full mode. Participants no longer reported discomfort. The limited size of the window reduced the glare of the contrast. P2 indicated, "It is not that bright and would not hurt much, because it's not covering the whole paper."

Moreover, we found that the window actually increased the effect of the contrast enhancement. P14 explained, "I realized something: When you add the borders [of the Window Display Mode], it increases the light in the area. The reading area seems brighter when you add the borders." The high contrast in the window, which was described as a flashlight by P13, highlighted the target and helped users locate and concentrate on the content they wanted to see: "It does highlight it. If you wanted me to focus on the 'Z', it would help find it... so it's easier to focus" (P13).

(3) Edge Enhancement

Image Tasks. In the near-distance image (face recognition, Figure 4(c)) task, most of the participants (14 out of 20), especially those who had more functional vision (e.g., P1, P13), felt edge enhancement was distracting, because it added "unfamiliar" information to the human faces. The edges darkened the scene and occluded the original colors and details of the images. As P14 commented, "I'm not seeing the actual skin. I see the skin but it's almost like the details have been removed. Almost like now it's a mask." For some participants, it even changed the image style and the way they perceive the images: "It sort of turns things into a cartoon" (P1), and "It looks like abstract art" (P4). However, some participants (4 out of 20), especially those who had more difficulties with distinguishing colors (e.g., P3, P19), believed that edge enhancement could help with the face recognition, because it increased the contrast. "I think it gives it more contrast. The features stand out a little bit more" (P3). P15 also indicated that the edges added 3D effect to the images, "The colors look more realistic. The details I saw before look more 3D, more of a 3D effect."

Meanwhile, more participants (7 out of 20) thought edge enhancement improved the visual effect in the far-distance image task (traffic signs, Figure 4(d)) than in the near-distance task. They commented that the edges emphasized the outline and defined the shape, "The shape is clearer. Actually, now I'm seeing something going like around it, now I see a diamond shape. Then to the right of that looks to me like a triangle shape" (P13). Some participants (e.g., P12, P20) also felt that edge enhancement helped to discover objects in their view and locate them quickly. "With the edge definition, the objects are more obviously there, because they're dark" (P12). Some of them mentioned that the edges could help with concentration: "I think it helps me concentrate on the image more, the black contour helps me to focus my vision more on the image" (P14).

Reading Tasks. Participants had strong but differing opinions about edge enhancement in the reading tasks, especially the near-distance reading. When reading the handheld printed page (Figure 4(a)), all participants except for P6 and P16 agreed that the font appeared bolder. Some participants indicated they could benefit from the bold letters. As P11 explained, "It's bold. Let's say with low vision, you always have a problem with differentiating 'e' and 'a'. This is very [clear]. It gives you a clear definition to the letter. It defines the letter." However, the others (seven participants) did not like this enhancement method while reading. Since the contours of each letter was thickened, it created a crowding effect [80], in which the letters got closer than a critical spacing, making the reading content difficult to be identified. "They're a little bit bolder and a little bit closer together. It's a little bit harder to distinguish them" (P12).

(4) Black/White Reversal

Image Tasks. Since black/white reversal was modeled after a common effect people use when reading with CCTV's, it largely distorted the color and detail information of the objects in one's environment. Unsurprisingly, most participants (18 out of 20) did not like this enhancement method in the image tasks. "You lose all details. You lose color cues, especially in a picture like Marilyn's where color really is what defines the different areas of the photo" (P18). Some participants also described the visual effect as an x-ray image (e.g., P11, P18) or a Rorschach test (e.g., P14), "[The image is] like an x-ray image. It's a little bit creepy" (P11).

However, one interesting result was that, for participants with ultra-low vision³ (e.g., P6, P16), black/white reversal was the only method that provided more information. Both P6 and P16 could barely see anything through ForeSee until we triggered black/white reversal. With this enhancement, P6 could distinguish the frame of the image and also discern some basic contours in the near-distance image task. P16 could even recognize there's a human face in the image: "Okay, there's an image here. Wow, let's see what's on this paper... Is this a face? It's a face!" (P16).

According to this finding, we believed that black/white reversal created extreme high contrast that helped people with ultra-low vision get more information from the images. It may also simplify the images and eliminate the distraction from complicated structures and various colors. However, for participants with better vision, this method oversimplifies the images, leading to loss of critical information.

Reading Tasks. Ten participants preferred black/white reversal to the original black-on-white text. They said there was too much glare on the original white background, which they found distracting. As P19 mentioned, "My eyes are so light sensitive, they pick up on the brightness of the light against the dark background more easily than trying to pick up the dark letters on a bright white background."

However, seven participants thought black-on-white was better than the reverse. They felt that the white letters were too bright, which hurt their eyes and caused fatigue and eyestrain. "This would be more tiring over a longer period of time, because I would be getting too much glare from the letters" (P14).

(5) Text Extraction

In the far-distance reading task, 11 participants benefited from the text extraction, because it was much clearer than the original one. "I feel like if it was used for like a street sign, then that would be great. I really like the text [extraction]. I pretty much can't see the numbers without the text detection" (P2). In the near-distance reading task, fewer participants (six out of 20) liked text extraction, because they were concerned about its efficiency and practicality. Text extraction only displays a few words at a time, making it difficult to read long paragraphs. As P18 indicated, "It's great if you're looking at distance. Usually a sign or something that is not going to be more than a few words, but to read like [a long paragraph] would take you a very long time. I don't think it's practical, maybe somehow superimposing it over [the original text]." Some of them (e.g., P5, P12) suggested using an eye-tracker to locate the words that the user is looking at and recognizing that text automatically, so that they would not need to manually select the words that they want to read.

Redirecting their gaze between the original text in the window and the extracted digital text at the bottom of the display was challenging for some participants (e.g., P7, P12). As P7 said, "I won't see the two [phrases] at the same time. I have to look at the [original phrase] and then I have to look down. I have to change the position of my gaze." Some participants (P7, P11) suggested that the extracted text could replace or appear next to the original text to avoid the gaze switching.

Participants proposed improvements to the style and location of the digital text presented in text extraction. Some participants felt the current location of the text (at the bottom of the display) was best, because the text did not cover the center of their view (e.g., P2, P11), and looking down felt natural for reading (P8). While eight participants (including P6 and P16 who have ultra-low vision), especially those with tunnel vision, had difficulty seeing the extracted text at the bottom of the display. The style of the digital text, including font size, color, and weight needed to be adjustable as well. P13 said text extraction did not help, because the text was too small for him to see, and P19 suggested reversing the color of the text to white-on-black and using a bold font.

³Ultra-low vision: visual acuity is less than or equals to 20/500 in the better seeing eye [47].

(6) Window Display Mode

We summarize general user reflections and use patterns with the Window Display Mode. Overall, 14 participants benefited from the Window in some viewing tasks.

Window placement. As we expected, most participants (14 out of 20) put the selection area at the location where their sight was best (usually in the center; for P4, it is the left side, because she only had vision on the left side of her right eye) and adjusted its size to fit the target object. Interestingly, P17 had a different approach and positioned it in a weaker region in her field of view. She put the window in the center of her visual field to compensate for her central vision loss. Another three participants (P1, P8, and P12) adjusted the position of the window according to the distance of the viewing target. P11 described his preferences: “For [near-distance] reading purposes, I think [the window] would be closer to the bottom. When you are looking at something far away, you are more likely to have your head upright. [The window] should be a little bit higher than the middle line. I think that is just the natural set up of this.” Participants used this positioning pattern to simulate a bifocal lens that they were familiar with, “Because that’s essentially what a bifocal lens is doing. That’s just what my glasses are for” (P1).

Benefits. Participants reported several benefits from using the Window as described below.

Concentration: Some (e.g., P14, P17) said the window helped them block the distracting surroundings and concentrate on the target. As P4 explained, “There’s less visual confusion because of the fact that I’m focusing on [particular] objects. When there are too many objects to look at, I’m trying to look at all of them at the same time, trying to distinguish the colors and the features. It’s just more like distracting. Here, it’s easier to see, because I’m focusing on one object at a time.” P4 even refused to conduct the reading tasks without the Window Display Mode.

Multitasking: Three participants (P8, P18, and P19) used the Window to support multitasking. They conducted different viewing tasks between looking inside and outside the window. For example, they put the targets that need enhancement inside the window and kept the target that didn’t need to be enhanced outside the window. This was because it was difficult for participants to switch from one task to another when using enhancements like magnification and black/white reversal in the Full Display Mode. For example, P18 felt disoriented when using black/white reversal: “As long as you don’t take your eyes off the page [with text] you’re good. The second you look up, you’re very disoriented.” While in the Window Display Mode, some participants felt the window enabled them to conduct both the task that needs enhancement inside of the window and the task needs no enhancement outside of the window at the same time. P7 explained:

“I’m trying to think of a situation where I was reading something at my desk and if I wanted to be able to keep track of the things going on around me, then I would use the enhancement box set to a thin line so I could get the text.” (P7)

Seeing Context: Finally, some participants (e.g., P1, P18) mentioned that the Window Display Mode allowed people to see more context (i.e., information about the original image and the background of the targets).

“It gives me a point of comparison, because I can be not looking at [the enhancement inside the window]. I need help to zoom in, but [with the Window Display Mode] I see contextually. As opposed to when it’s full blown, maybe the issue I had before was I don’t know where I am in a world that large. I think this is just [more] contextual.” (P18)

Weaknesses. Participants also mentioned some weaknesses of the Window.

Inconsistency: Participants with better peripheral vision (e.g., P1, P5, P18) mentioned the inconsistency between inside the window where information was enhanced and outside the window

where information was not enhanced. They felt this inconsistency distracted them. For example, P18 commented on the magnification in Window Display Mode: “There are too many text sizes going on at once. I’m seeing the text changing size as I’m reading it peripherally and it’s a little distracting” (P18).

Difficulty Locating the Window: Some participants (e.g., P2, P13) found it difficult to find the window in their visual field. “It was probably worse because you have to find the specific spot” (P2). Most participants preferred to set the window in a fixed position and then move their head to position the target in the window. Some of these participants (e.g., P17, P20) felt this required too much effort, especially for the crowded text in the near-distance reading task, “It’s hard to put the specific line [of text] into this [window]. It’s frustrating” (P17).

4.2.4 Combining Enhancement Methods. Most participants improved their visual experience by combining enhancement methods. We present the methods and display modes participants chose to achieve their best visual experience in Figure 5. We found that preferences among participants differed a lot. Meanwhile, each participant’ preferences also differed across tasks.

Most participants (more than 70% in each task) used magnification. For some participants, other enhancement methods were effective only with magnification, especially in the far-distance tasks. P12, for example, didn’t feel that edge enhancement alone was beneficial in the far-distance task. However, when combining it with magnification, he said: “Edge enhancement probably helped. Plus you’ve got magnification going too, right? Now I could see the contour of the inside part. The edging really helped me distinguish the number.”

Even though combining enhancements improved their vision, some participants (e.g., P2, P11) felt this was a little complicated for them. There were too many enhancement methods and display modes, so that it was difficult for them to determine what combination was optimal for them. They commented that the effort they spent to configure the system should match the improvement in their experience. “I feel like the effort has to match the efficiency. So, if the effort is too much, I’m not going to do it. But if the effort is minimal and get good benefit, I will use it” (P2). This emphasized the importance of designing interaction techniques that are easy and efficient for low vision people to use when customizing their visual experience. P11 also suggested providing a small set of options that worked best for each particular user:

“I think different people with low vision could have different options, you know, like contrast works for me, zoom works for other people. I think the options are important, but not having too many options either. You know, just having couple of different ways to go, just couple, like the best couple, not like having ten different choices. Maybe three choices that work best for me. Too many options will be too hard to use.” (P11)

4.2.5 Control with Touchscreen Gestures. Although this study focused on evaluating the vision enhancements on the HMD, we also observed participants’ use of the touchscreen-based gestures on the smartphone controller (Figure 3, right) to adjust the size and position of the window in the Window Display Mode. We found that all the participants learnt the gestures quickly and could control the window easily with the smartphone application. All of them felt it was easy and feasible to use touch gestures to control ForeSee. This motivated us to design a complete set of touchscreen-based interaction techniques to enable low vision users to customize their visual experience by selecting different enhancements and adjusting the parameters (e.g., magnification level). We describe our design and evaluation of interaction techniques for ForeSee in the following sections.

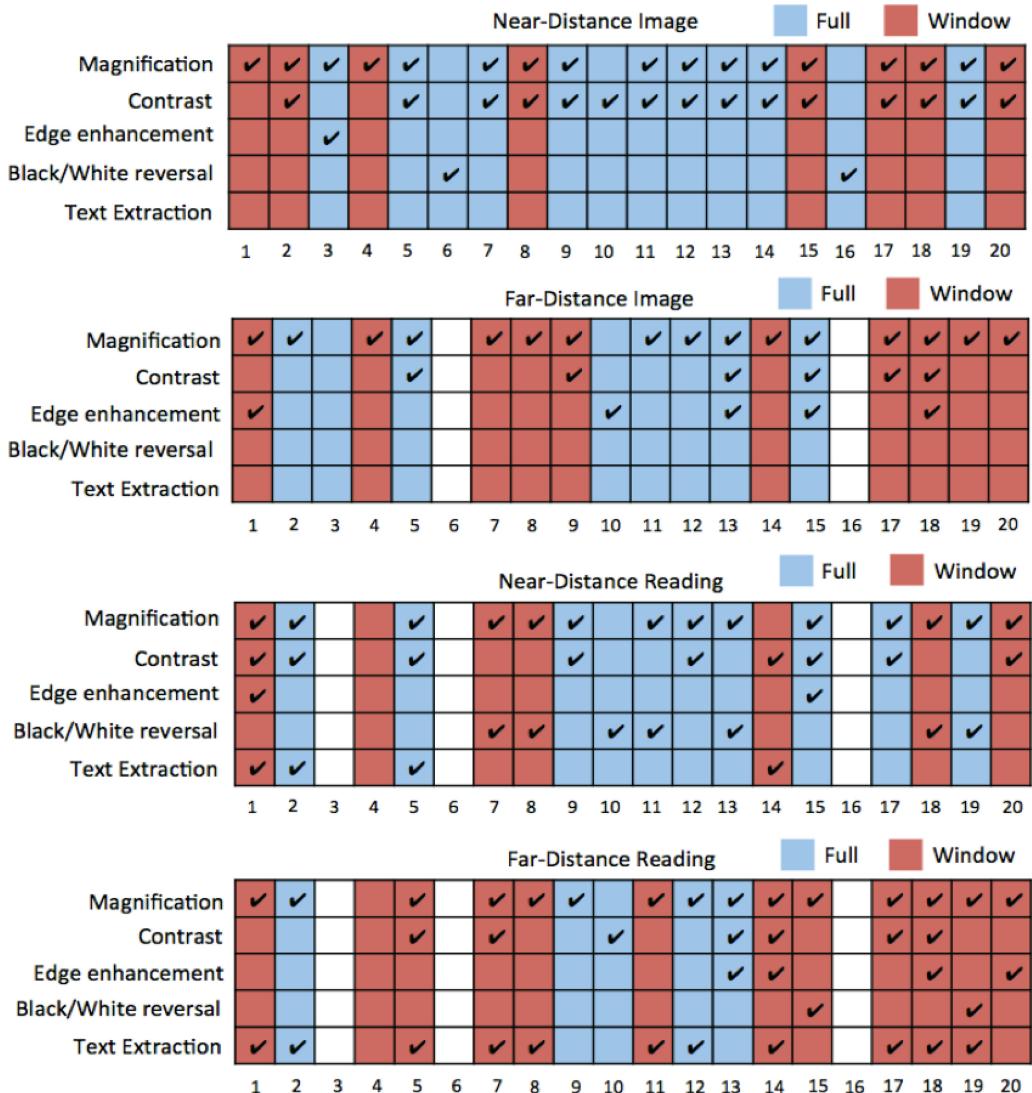


Fig. 5. Participant preferences for combined enhancements in the near- and far-distance tasks. The x-axis lists the 20 participants, while y-axis lists the five enhancement methods. Blue columns represent the Full Display Mode, and orange columns represent the Window Display Mode. If a participant chose to use an enhancement in the combination condition, then we added a check to the square corresponding to that enhancement. For some participants who had ultra-low vision and could not provide their preferences for all the tasks, we keep their columns blank. P4 preferred the Window Display Mode without any enhancement in the near- and far-distance reading tasks, so there are no check boxes in the column.

5 DESIGNING INTERACTION TECHNIQUES

After studying the effect of different vision enhancements, the output of ForeSee, we now turn to understanding the input and interaction. We address the second research question posed in the Introduction:

RQ2. What interaction techniques enable people with low vision to interact with an HMD system effectively?

As our first study confirmed, an effective vision enhancement system must have usable interaction techniques that allow the user to control and customize the system. During our study of vision enhancements, we allowed participants to control ForeSee with speech commands (implemented by WOZ) and gestures on a smartphone touchscreen (Figure 3, right). These interaction techniques were merely a way to allow participants to customize the system to their liking—we did not aim to evaluate the effectiveness of these interaction techniques. However, they did provide insights into preliminary input options for an HMD system for low vision people.

Speech is a standard interaction method and, as such, a natural candidate modality for interaction. Many smart devices support speech assistants and dictation, allowing users to use speech to interact with the systems, for example, Siri in Apple devices, Cortina in Window devices, and Alexa in Amazon Echo. Since speech interaction is eyes-free, it is accessible to people with low vision. Prior research found that speech can be an effective mode of input [5] and interaction [83] for people with visual impairments. Our study of vision enhancements also showed that all participants could easily use speech (implemented by WOZ) to customize ForeSee. Thus, we decided to use speech as the basic input method of ForeSee. However, speech is not suitable for all situations, for example, a quiet public environment, or giving private commands in public. Using speech control in unsuitable situations can be especially problematic for people with disabilities, since it can expose or emphasize their disabilities.

When speech is not an appropriate or desirable option, it is critically important to have an alternative. Touchpad gestures proved to be a viable option in our first study, in which participants could easily use touch gestures on a smartphone screen to control the size and position of the window in the Window Display Mode. However, interactions on a touchpad or a smartphone would require the user to retrieve such a device out of a pocket or purse, posing an unnecessary barrier to convenient and fast interaction. We decided instead to explore gestures on a smartwatch, since it is a wearable “always available” [67] touchscreen that a user can access easily in many situations. It can be wirelessly connected to an HMD and used as an input sensor. As with HMDs, smartwatches are growing in popularity and are becoming mainstream, relatively low cost devices that can serve as socially-acceptable accessibility tools.

In the following sections, we describe the design and evaluation of the interaction techniques we designed in ForeSee, which include a set of speech commands and a set of smartwatch gestures.

5.1 Interaction Techniques on a Smartwatch

5.1.1 *User-Centered Design of Interaction Techniques on a Smartwatch.* To design these interaction techniques for low vision users, we followed a user-centered design method [20], where we conducted informal formative studies with eight low vision participants (two female, six male, age ranges from 23 to 64 with the mean of 36). During the formative studies, we asked the eight participants what input gestures and output feedback they preferred, demonstrated different prototypes to them, and improved the prototypes based on their suggestions. We describe our iteration process and the final design of our interaction techniques below.

We designed three interaction techniques on a smartwatch to support the following operations in ForeSee: *select an enhancement*, *adjust the enhancement level*, and *toggle an enhancement on and off*. With these interaction techniques, a user can have full control of their visual experience by combining different enhancements and adjusting the level of each enhancement method.

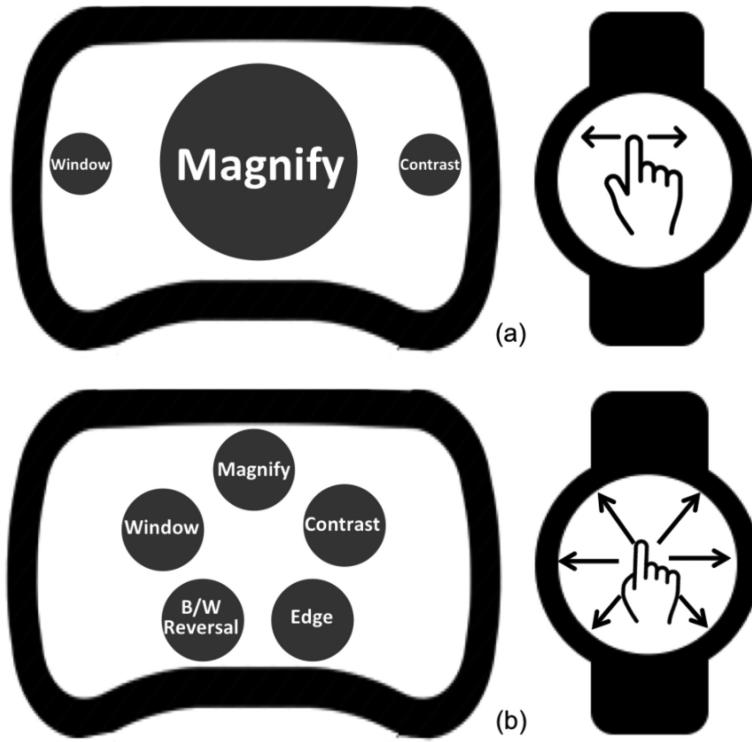


Fig. 6. Two menu types: (a) a linear menu that a user can switch menu items by swiping to left or right on the smartwatch surface; (b) a pie menu where all menu items are shown in a circle at a time. A user can select one item by swiping to the corresponding direction on the smartwatch.

(1) Selecting an Enhancement

Since we have five enhancement methods and two display modes (Figure 2) for a user to select, we designed a menu that includes all these options. A user can navigate the menu and select a specific enhancement or display mode for further control.

Menu Hierarchy. Since there are two visual effect categories (enhancement and display mode), we originally designed a two-level menu. The users needed to first select which category they wanted to navigate (enhancement or display mode), and then select a specific item in the corresponding category. However, when showing this design to our participants, they all commented that the two-level menu was too complicated and that it took too long to navigate the menu. Thus, we designed a one-level menu by regarding the Full Display Mode as the default: the menu items included the five enhancements and the Window Display Mode. If a user only selects the enhancements, then they are presented in the Full Display Mode. If she selects the Window Display Mode, then all the enhancements she has selected will be rendered in the Window.

Menu Type. We prepared two types of menus for participants to compare: a linear menu and a pie menu. The first was a linear menu, in which only one menu item was presented on the HMD at a time and a user must swipe left or right on the smartwatch to move to the previous or next item. When moving forward from the last item, the user returned to the first item, as shown in Figure 6(a). The second menu design was a pie menu, in which all menu items were presented in a circle on the HMD. A user must select one item by swiping along the direction from the smartwatch center to the location of the item on the display, as shown in Figure 6(b). The swipe gesture can

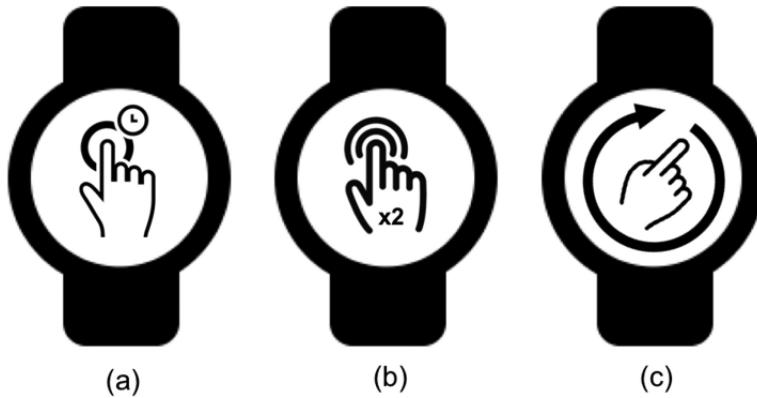


Fig. 7. Touch gestures to control ForeSee: (a) a long press gesture to trigger the menu; (b) a double tap to select a menu item; (c) a clockwise or counterclockwise slide to adjust the enhancement level.

start at any point on the smartwatch screen. We select these two types of menus, because they represent the two extreme cases: the smallest (one for linear menu) and the largest (all items for the pie menu) number of menu items that could be shown at a time.

Most participants preferred the linear menu, since the menu items in the pie menu were too small to see (or read easily). However, two participants who had relatively good vision liked the pie menu, since they could see all menu items at one time. The pie menu enabled them to select an enhancement quickly. The pie menu was challenging for people who had peripheral vision loss. Since their visual field was limited, they could not see all menu items at one time and had to scan around to find and read all items, which increased the time to traverse all items. Considering the fact that too few menu items shown on the screen at a time would increase the time of navigating the whole menu, while too many menu items at a time would make each item too small to see for the low vision users, we tried to get the preferable number to optimize the menu design. We asked participants how many menu items they wanted to see on the screen at a time. Most participants (six out of eight) preferred to have only one item presented on the screen at a time, explaining that it was easier for them to focus on just one item. They also wanted the menu item text to be as large as possible with high contrast, for example, using white text on a black background. Thus, according to participants' preferences, we used the linear menu (presenting one menu item at a time) as our final design (Figure 6(a)).

Smartwatch interaction. Using the linear menu, we asked participants what gestures they preferred to trigger the menu and select a menu item. All participants preferred a double tap gesture (Figure 7(b)) to select a menu item, explaining that this was a standard method of selection used on smartphone screen readers. Although they had functional vision, all participants had at least some experience with screen readers on smartphones. To trigger the menu, most participants chose a one-finger long press (Figure 7(a)). Although some participants suggested two-finger gestures, such as a two-finger double tap, or a two-finger long press, we found that one finger gestures were more effective because of the limited size of the smartwatch screen.

(2) Adjusting the Enhancement Level

When an enhancement or display mode is selected, the user enters the control mode of this specific enhancement or display mode. For each enhancement method (except for Black/White Reversal), we allowed users to adjust the enhancement level (e.g., magnification level, edge thickness, window size for the Window Display Mode) from level 0 to 10. For the Window Display

Mode, the user can also adjust the window size from level 0 to 10. Adjusting a parameter is a slider-type interaction that needs to be performed with the limited screen of the smartwatch. As such, we leveraged the edge of the smartwatch surface and allowed users to slide around the watch screen in a clockwise (or counterclockwise) direction to increase (or decrease) the enhancement level, as shown in Figure 7(c). The user can start this gesture at any point along the edge of the screen. The system computes the enhancement level based on the rotation angle around the watch surface center between the start and the current touch point. The enhancement level increases (or decreases) one level when the user slides clockwise (or counterclockwise) around the watch 36 degrees. When we demonstrated this interaction technique, all participants felt it was both effective and easy to use.

In terms of output, some participants suggested showing the value of the current enhancement level on the ForeSee display during the adjustment, since it was hard to know what level they were at, especially whether the enhancement has been adjusted to the highest or lowest level, only by observing the enhancement effect. Thus, we provided this visual feedback at the top of the HMD (Figures 8(b)–8(d)).

(3) Toggling an Enhancement On and Off

In addition to adjusting the enhancement level, we designed a way for users to turn on and off an enhancement. Recall that when a user selects an enhancement from the menu with a double tap gesture, she enters the enhancement control mode. At this time, she can double tap again on the screen to toggle the enhancement on and off. For an enhancement method that only needs to be triggered (i.e., black/white reversal, text extraction), a user can use a double tap to trigger and cancel the enhancement. For an enhancement method that has multiple enhancement levels (e.g., magnification), a user can turn off the enhancement directly without adjusting the enhancement level to zero. When she turns on this enhancement with a double tap, it will reload its previous state (the level it was previously set to). All participants were satisfied with this design.

(4) Interaction Flow

Based on the three interaction techniques above, a user can easily select, combine, and adjust different enhancements to customize their visual experience. We describe the interaction process below (Figure 8).

When the system is launched for the first time, the default setting is the Full Display Mode without any enhancement. To add an enhancement, a user can long press on the smartwatch surface to trigger the linear menu that includes all enhancement methods and the Window Display Mode, and then swipe to the left or right to switch the menu item (Figure 8(a)). When the menu item is switched, we generate a movement animation to simulate the effect of the current menu item moving away and the next one moving to the screen center, which can better attract the user's attention [63]. After the desired enhancement or display mode is shown on the display, the user can double tap to select it and enter the control mode of this enhancement.

If an enhancement is selected for the first time, then its default enhancement level is 0 (Figure 8(b)). If the user selects Black/White Reversal, then the enhancement is triggered directly. When an enhancement is selected, the user can move her finger clockwise or counterclockwise around the watch face to adjust the enhancement level. The system presents the corresponding enhancement level on the top of the user's view on the HMD (Figure 8(c)). The user can also double tap to turn off or back on the current enhancement (Figure 8(d)).

To add another enhancement, a user can simply long press to trigger the menu again and repeat the above interaction process. To help the user easily continue the navigation, the system saves the menu's state. When the user triggers the menu again, the menu loads the menu item that she selected last.



Fig. 8. Interaction Flow of ForeSee (the images on the right column are the screen shot images a user can see through ForeSee in each interaction step, and the smartwatch figures on the left column represent the gestures to trigger a corresponding step): (a) Long press to trigger the menu and swipe to the left or right to switch to a specific enhancement method; (b) Double tap to select an enhancement; (c) Clockwise or counterclockwise slide to adjust the enhancement level; (d) Double tap to turn off/on the enhancement.

If the user wants to use the Window Display Mode, then she can select it in the menu with the same interaction and all enhancements she has selected will be rendered only in the Window. To simplify the interaction, our current interaction design does not involve adjusting the window position, but always keeps the window at the center of the HMD display in a square shape. The user can slide clockwise (or counterclockwise) to increase (or decrease) the size of the window. We illustrate the interaction flow of using the Window Display Mode in the Appendix (Figure A1).

5.1.2 Summative Evaluation of the Interaction Techniques on a Smartwatch. After completing the user-centered design process, we conducted a summative study with 11 low vision participants to evaluate the usability of the touchscreen-based interaction techniques on a smartwatch.

Method

Participants. We recruited 11 people with low vision to participate in our study (5 females, 6 males) with a mean age of 55.6 (range: 33–74). None of these participants had participated in the design iterations, while two (I4, I9) participated in the first study where we evaluated the vision enhancements on ForeSee. Our recruitment criteria were the same with the first study. We required that participants have “low vision,” and use vision enhancement tools such as magnifiers or a CCTV on a regular basis. Volunteers who only use screen readers were not considered suitable, since this served as an indication that they had too little vision to be considered “low vision.” Our participants had a variety of vision conditions, as shown in Table 2. They were compensated \$15 per hour and received up to \$30 in reimbursement for travel expenses.

Apparatus. We implemented a new ForeSee prototype on a Google Nexus 5 smartphone in an OCT17 VR headset, so that the system would not have to be connected to a laptop as the original prototype did. We used the phone’s back camera to capture the surrounding environment and rendered the different enhancement methods on the phone screen. We split the phone’s screen into two parts as the displays for the user’s left and right eyes, and rendered the same images of the enhanced environment, creating binocular vision. We did not implement the Text Extraction method in this study, since it is only useful for text reading in the Window Display Mode.

We implemented the smartwatch-based interaction techniques on a Motorola Moto 360 smartwatch. The smartwatch communicated with the smartphone through Bluetooth.

Procedure. Our study consisted of one lab session that lasted about 45 minutes. We started the study with a brief interview asking the participants about their demographic information. We then introduced ForeSee to the participants, including all the enhancement methods and display modes, and the interaction techniques on the smartwatch. We asked them to freely adjust ForeSee to look at the surrounding environment until they felt familiar with the system. We continued the study when the participants confirmed that they memorized all touch gestures.

When evaluating the effectiveness and usability of the smartwatch interaction techniques, we focused on two tasks: select an enhancement and adjust the enhancement level. We did not evaluate the interaction technique of toggling on and off an enhancement, because it is too simple (only a double tap gesture).

(1) *Selecting an Enhancement.* Participants first conducted the task of selecting an enhancement for 10 trials. This task includes three steps: long press to trigger the menu, swipe to reach the target enhancement, and double tap to select it. For each trial, we assigned the participant a target menu item and said “start.” The participant then completed the task as soon as possible. We recorded the time for each trial.

We counterbalanced the target enhancements in the 10 trials: Since the system saves the last menu item that a user has selected, the menu might start from a different menu item in each trial. Thus, we could not simply counterbalance the target items by assigning each menu item twice. Instead, we counterbalanced the swipe behavior (including the minimum number of swipes and

Table 2. Demographic Information of the 11 Participants Who Evaluated the Interaction Techniques

ID	Age/ Sex	Diagnosis	Legally Blind	Visual Field	Visual Acuity	Color Vision	Tool Used
I1	62/M	Aniridia; Glaucoma	Yes	Not full; a little loss on the left	Not sure	Good; not light sensitive	CCTV, color reversal on computer, ZoomText, VoiceOver on iPhone
I2	74/F	Pseudoxanthoma elasticum (PXE)	Yes	Full	Left: 20/400; Right: 20/300	Problem with subtle colors, not light sensitive	Magnifying glasses, ZoomText, handheld CCTV
I3	74/M	End-stage Glaucoma; Scotoma;	No	Limited, one eye does not have central vision	20/60 after correction; 20/800 before	Problem with subtle colors; a little light sensitive	Reading glasses, KNFB reader, magnification on computer and iPad, VoiceOver on iPhone
I4	60/F	Retinopathy of Prematurity	Yes	Full	Left: 20/400; Right: 20/300	Good	CCTV, magnifier glasses, monocular, ZoomText, handheld magnifier
I5	51/F	Stargardt's Macular Dystrophy	Yes	Has peripheral vision, but no central vision	Unknown, but low	Problem with subtle colors; light sensitive	Handheld magnifier, Ruby digital magnifier, magnification app with audio on iPhone
I6	33/M	Retinopathy of Prematurity, Nystagmus	Yes	Full	20/200	Good color vision; light sensitive	Dictation and magnification on iPhone, Grabber (OCR app), AmpliVision (magnification)
I7	53/M	Flecked retina syndrome; Central scotoma	Yes	Limited central visual field; good peripheral vision	Left: 20/200; Right: 20/400	Problem with low contrast; light sensitive	Glasses for contrast, Zoomtext, CCTV, magnification on android phone
I8	56/F	Stevens-Johnson Syndrome	Yes	Very limited	Left: 20/150; Right: only light perception	Good color vision; light sensitive	Handheld magnifier, ZoomText, CCTV, Magic on laptop
I9	38/M	Retinitis pigmentosa	Yes	Full	20/400 after correction	Problem with subtle colors; light sensitive	Color reversal and magnification on computer, magnification and dictation on iPhone
I10	44/M	Retinitis pigmentosa	Yes	No peripheral vision	20/200	Color blindness; not sensitive to light	ZoomText, Amigo illuminated handheld magnifier (10×)
I11	67/F	Glaucoma	Yes	Limited on the right eye	Left: Unknown but good; Right: 20/2500	Good color vision; light sensitive	Optelec handheld magnifier (2.5×/6D), magnification on computer

the swipe direction) of each trial, since it represents the distance between different menu items. With five items in the menu (four enhancement methods and the Window Display Mode), there were five different swiping behaviors to reach an enhancement: no swipe, swipe once to the left, swipe once to the right, swipe twice to the left, and swipe twice to the right. In the 10-trial user study, participants should conduct each swipe behavior twice for counterbalancing purpose. For each participant, we randomize the order of these swipe behaviors and generated a corresponding enhancement sequence to assign to her in each trial. Note that the menu always started from Magnification in the initial trial for each participant.

(2) *Adjusting the Enhancement Level.* Participants then were tasked with adjusting the enhancement level for 10 trials. We used the magnification enhancement mode across all trials and participants. We first set the system to the control mode of Magnification and adjusted the magnification level to 0. For each trial, we assigned participants a target level. When we said “start,” participants completed the adjustment as quickly as possible. We recorded the time for each trial.

We counterbalanced the target enhancement level using the similar method as the enhancement-selection task: Since the participants conducted the adjustment based on the enhancement level of the last trial, we counterbalanced the difference between the target level of a current and previous trial. All participants should complete 10 different adjustment conditions in

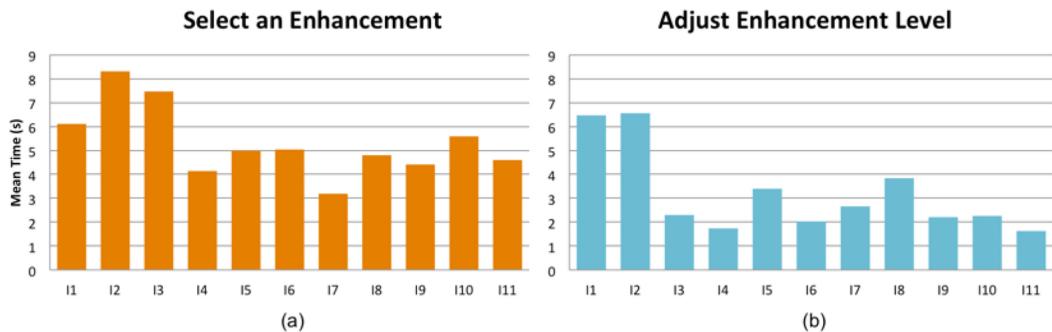


Fig. 9. (a) The mean time each participant spent when selecting an enhancement. (b) The mean time each participant spent when adjusting enhancement level.

the 10 trials: increasing 1 to 5 levels, and decreasing 1 to 5 levels. For each participant, we randomized the order of the conditions and generated a corresponding enhancement level sequence to assign to each trial. We ensured all the generated enhancement levels were between 0 and 10. If they were not, then we generated a new sequence. The enhancement level always started from Level 0 in the initial trial for each participant.

After participants completed the two tasks, we conducted a short interview, discussing the benefits and weaknesses of the interaction techniques. We asked participants to assess the effectiveness and usability of each interaction technique on a 7-point Likert Scale, where 1 represented not effective or usable at all and 7 represented extremely effective or usable.

Analysis. We video recorded the study and took notes during the study. We recorded the time that participants spent in completing each interaction task from the video. We watched the videos and transcribed meaningful participant quotes and summarized participants' feedback on the design of the interaction techniques.

Results

When selecting an enhancement method, participants spent a mean time of 5.33 s to complete the task (ranging from 3.19 s to 8.31 s, $SD = 1.48$ s). We present the mean time that each participant spent in the 10 trials in Figure 9(a). For the task of adjusting the enhancement level, the mean time participants spent was 3.19 s (ranging from 1.63 s to 6.56 s, $SD = 1.77$ s). Figure 9(b) shows the mean time each participant spent when adjusting the enhancement level in the 10 trials. During the study, we found that I1 and I2 spent relatively longer time than most other participants in both interaction tasks, and I3 spent a long time when selecting an enhancement. We speculate the reason: I1 had the lowest visual ability among all the participants and struggled reading the visual information (e.g., the menu item, and the enhancement level) on the HMD. I2 and I3 were the two oldest participants in our study (both 74 years old) and I2 did not have any experience of using a touchscreen, so they were slower in performing touchscreen gestures on the smartwatch: I3 sometimes had difficulty with the swipe gestures, while I2 generally moved slowly.

Participants rated the effectiveness and usability of the interaction techniques highly, as shown in Table 3. For the technique of selecting an enhancement, participants' mean score was 6.27 ($SD = 1.19$) for Effectiveness, and 6.91 ($SD = 0.30$) for Usability. For the technique of adjusting the enhancement level, participants' mean score was 6.45 ($SD = 1.04$) for Effectiveness and 6.64 ($SD = 0.67$) for Usability.

In general, most participants liked our design and felt the interaction techniques on the smartwatch were effective and easy to use. As I10 described, "I think it's very simple. I thought it would be more cumbersome, but the gestures are adequate. It's not overly clunky, it's not difficult to

Table 3. The Histograms and Means of the Effectiveness and Usability Scores of the Two Interaction Techniques: Selecting an Enhancement and Adjusting Enhancement Level

1	2	3	4	5	6	6	Mean
Effectiveness score of Selecting an Enhancement							6.27
		1			4	6	
Usability score of Selecting an Enhancement							6.91
					1	10	
Effectiveness score of Adjusting the Enhancement Level							6.45
			1	1	1	8	
Usability score of Adjusting the Enhancement Level							6.64
			1	2	8		

manage, and it seems like even if I'm in motion I would still be comfortable with the watch." Even I2 who had never used any touchscreen devices learnt all our interaction techniques within 10 minutes. "It took some time to learn it, but it was not difficult. I wanted to use [ForeSee] for shopping to look at the price tag ... and looking at the street sign" (I2).

However, I2 gave a low score (3) for the effectiveness when selecting an enhancement, because she felt the text on the menu item should be larger. I3 felt it was challenging to adjust the enhancement level and gave relatively low scores (4 for Effectiveness and 5 for Usability), because he was not sure how much he should move his finger around the watch to get to a specific enhancement level. He sometimes moved his finger over the target value and had to move back, which extended his time. As I3 indicated, "It took me longer to learn, because I don't know how far I should go [on the watch surface] to go from [level] 7 to [level] 3."

Participants suggested ways to improve the interaction techniques. Although they were able to read the current menu item, some participants (e.g., I2, I7, I10) wanted the text to be larger. They suggested using the first letter or a standard icon to represent an enhancement instead of using a word. I7 said, "I like [the menu]. [But] maybe instead of a word, using a big W, a big M, E, C, BW, kind of the way Google does their icons in the email. Another thing you might be able to do is using a graphic for the menu items. Something instantly recognizable is good."

When adjusting the enhancement level, participants accidentally triggered the menu, because they rested their finger on the smartwatch surface (conducting a long press gesture) when deciding how much they should slide. I6 suggested using a "tap and tap hold" to trigger the menu, since it is a standard gesture for smartphone screen reader to represent a long press and it is not easy to be accidentally triggered.

5.2 Speech Commands

We also designed a set of speech commands, providing eyes- and hands-free audio input to control ForeSee, shown in Table 4. Besides speech commands that trigger the same functions as the touch gestures on the smartwatch (Commands 1–7), we also provide two short cut commands that allow users to select an enhancement and set the enhancement level to a specific value directly without navigating the menu, see Commands 8 and 9 in Table 4. For the current version of ForeSee, the user has to use the exact speech commands. If they say the wrong words, then the system will verbally report "wrong command."

6 EVALUATING INTERACTION TECHNIQUES FOR FORESEE

We evaluated the effectiveness of ForeSee's two sets of interaction techniques separately (i.e., smartwatch interaction techniques and speech control) in a viewing task: reading small letters

Table 4. Speech Commands to Control ForeSee

ID	Speech Commands	Function
1	Call Menu	Trigger the menu
2	Left	Switch to the previous menu item
3	Right	Switch to the next menu item
4	Select	Select the current menu item
5	Increase Magnification, Contrast, Edge, Window	Increase the level of the corresponding enhancement by one
6	Decrease Magnification, Contrast, Edge, Window	Decrease the level of the corresponding enhancement by one
7	Turn On, Off Magnification, Contrast, Edge, Black/White Reversal, Window	Turn on/off the corresponding enhancement or display mode
8	Use Magnification, Contrast, Edge, Black/White Reversal, Window	Get into the control mode of the corresponding enhancement or display mode directly
9	Set Magnification, Contrast, Edge, Window to 0–10	Adjust the corresponding enhancement level to the specific number

on a faraway display. We chose this task because seeing small text at a far distance is a challenging but important task in people's daily lives, for example, reading a street sign on the street. The two types of interaction methods can be used together, but we evaluated them separately to understand the benefits and weaknesses of each on their own.

6.1 Method

6.1.1 Participants. The same 11 low vision participants who took part in the evaluation of the smartwatch-based interaction technique (Section 5.1.2) participated in this study. Table 2 shows their demographic information. They were compensated \$15 per hour and received up to \$30 in reimbursement for travel expenses.

6.1.2 Procedure. The study consisted of one lab session that lasted about 45 minutes. We introduced ForeSee to the participants, including all the enhancement methods and display modes, the interaction techniques on the smartwatch, and the speech commands. We asked them to adjust ForeSee to look at the surrounding environment until they felt comfortable with the system and the two sets of interaction techniques. We continued the study when the participants confirmed that they memorized all touch gestures and speech commands.

We then asked participants to wear ForeSee and read small letters on a screen at a 10-foot distance. We simplify this reading task by using letters instead of text to reduce the time participants spent in reading itself, since we wanted to evaluate the interaction with the system and not their reading speed.

Participants sat 10 feet away from a large digital screen and read letters with ForeSee in two conditions: (1) using touch gestures on the smartwatch; and (2) using speech commands to interact with the system. Each participant conducted 15 trials of reading tasks in each condition. For each trial, we first removed all the enhancements in ForeSee and presented a letter on the large screen. We said "start," indicating that the participants should begin adjusting the enhancements until they could recognize the letter and read it aloud. If they misrecognized the letter, then we informed them immediately and allowed them to continue adjusting the system until they recognized the letter correctly. We asked the participants to complete the task as quickly as possible. We recorded the time and the enhancements each participant selected for each trial.

We varied the size of the letters in three levels: 1, 2, and 3 inches in height. We randomized the letter and the letter size for each trial but made sure each participant conducted five trials of

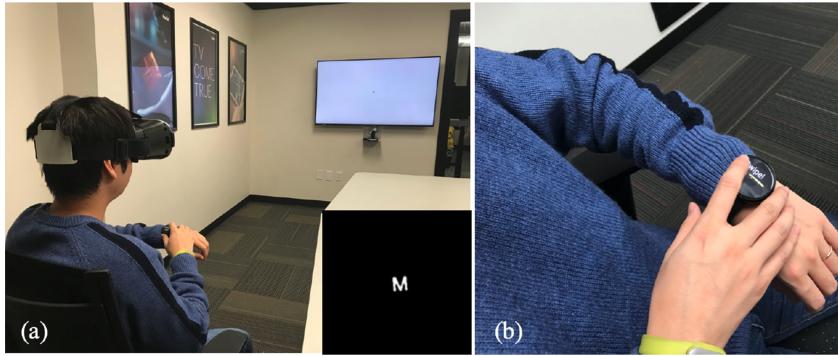


Fig. 10. (a) A user who adjusted ForeSee to read a letter on a large screen. The inset on the bottom right shows the visual effect on ForeSee's head-mounted display after the user's adjustment (Magnification+Black/White Reversal). (b) A user is conducting a swipe gesture on the smartwatch surface.

reading tasks with each letter size in each condition. We also counterbalanced the order of the two interaction conditions.

We concluded the session with a short interview. Participants discussed their experience of using the smartwatch and speech commands to control ForeSee when reading the letters, including the benefits, the weaknesses, and their suggestions for improvement. We also asked them to assess the effectiveness and usability of both interaction conditions on a 7-point Likert Scale, where a higher score represents a better experience.

6.1.3 Apparatus. The study was conducted in a well-lit conference room. The large screen we used to present the letters was a 60.7×39.2 -inch Panasonic Full HD display (TH-80LFC70). We used black Arial (a sans serif font) letters on a white background during the study. Figure 10 shows the setup of the experiment room and how a participant used the smartwatch to control ForeSee and read a letter on the large screen.

Although we implemented speech recognition on ForeSee, we used Wizard of Oz to simulate the speech control in the user study to minimize the confounding effect of speech recognition errors. The researcher, acting as the wizard, responded to the participants' spoken commands and adjusted the parameters of the system remotely with another Google Nexus 5 smartphone via WiFi. If the participants spoke an incorrect command, then the researcher said "wrong command."

6.1.4 Analysis. We recorded the time that participants spent in completing each task from the video. We analyzed the data with ANOVA. Our study had one within-subjects factor, Condition (smartwatch interaction, speech), and one measure, Time. We defined a Trial (1–15) as one reading task. To validate counterbalancing, we added another between-subjects factor, Order (two levels: smartwatch-speech, speech-smartwatch), into our model. An ANOVA showed no significant effect of Order on reading time ($F_{(1,9)} = 1.978, p = 0.193$, for $\alpha = 0.05$).

We video recorded the study and took notes during each session. We watched the videos and transcribed meaningful participant quotes and summarized participants' feedback on the design of both sets of interaction techniques.

6.2 Results

6.2.1 Effectiveness. In the letter-reading task, no participants could read the letter with their best correction (using their own standard glasses or contact lenses). However, with ForeSee, all of them were able to read the letters correctly by adjusting the enhancements. We recorded the

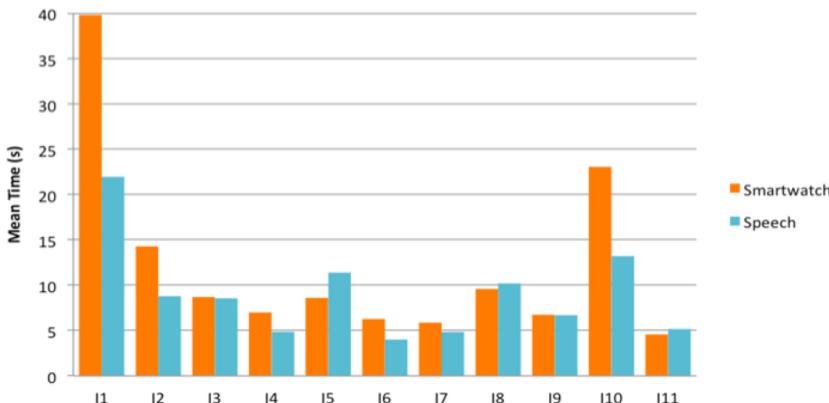


Fig. 11. The mean time each participant spent adjusting ForeSee to read letters on the large screen using the smartwatch interaction techniques (orange) and the speech commands (green).

Table 5. The Histograms and Mean Scores of the Effectiveness and Usability of the Two Sets of Interaction Techniques: The Smartwatch Interaction Techniques and the Speech Commands

1	2	3	4	5	6	7	Mean
Effectiveness score of the smartwatch interaction techniques							6.09
	1			1	3	6	
Usability score of the smartwatch interaction techniques							6.73
				1	1	9	
Effectiveness score of the speech commands							6.91
					1	10	
Usability score of the speech commands							6.18
				3	3	5	

time participants spent when adjusting ForeSee with the two different interaction techniques in the letter-reading task. When using the touchscreen interaction techniques on the smartwatch, participants spent a mean time of 12.25s ($SD = 10.53$ s) adjusting the system to read letters on the screen that is 10 feet away. When using the speech commands, participants spent a mean time of 9.06s ($SD = 5.22$ s) adjusting the system to read the letters. Figure 11 shows the mean time each participant spent when adjusting ForeSee with the smartwatch and the speech commands, respectively. With an ANOVA, we found no significant effect of interaction conditions ($F_{(1,9)} = 3.095$, $p = 0.109$) on Time.

Most participants (7 out of 11) adjusted the system and read the letters in less than 10 seconds. I1 and I10 spent more time than the other participants in both interaction conditions, since they combined two enhancement methods, Magnification and Black/White Reversal, to read the letters (see Section 6.2.3). I1 spent a particularly long time, because his vision was very low and struggled seeing the menu items as he was adjusting the enhancement. This was similar to the result from our summative study of the smartwatch interaction techniques (Section 5.1.2), when I1 spent a relatively long time performing the interaction tasks with the smartwatch (Figure 9).

In general, participants enjoyed using ForeSee to read the letters. “It’s fantastic. It’s very simple to use” (I7). Most participants gave high scores to both interaction conditions, as shown in Table 5. When using the smartwatch, participants gave a mean score of 6.09 ($SD = 1.51$) for Effectiveness of the interactions, and gave a mean score of 6.73 ($SD = 0.65$) for Usability. Compared with the

smartwatch, participants gave a higher score for Effectiveness when using the speech commands (mean = 6.91, SD = 0.30), but gave a lower score for Usability of the interaction techniques (mean = 6.18, SD = 0.87). We explain the reasons by comparing the two interaction conditions in the following section. However, we found no significant effect of interaction conditions on the subjective scores for Effectiveness ($V = 0, p = 0.098$) and Usability ($V = 23, p = 0.141$) with Wilcoxon signed-rank tests.

6.2.2 Subjective Responses to the Two Interaction Conditions. We compared the two interaction conditions: touch gestures on the smartwatch and the speech commands. We found that most participants (8 out of 11) spent more time performing the touch gestures on the smartwatch than the speech commands (Figure 11). This is not surprising, because the interactions on the smartwatch were more complex than the speech commands: While a user could use only one speech command to add an enhancement (Command 8 in Table 3) or adjust the level of an enhancement (Command 9), she had to trigger and navigate the menu to access a specific enhancement when using the smartwatch. As I9 indicated, “The speech I think is easier to use, because you can set the parameter directly. You don’t have to go through the menu.” This also explained why participants gave a higher score to Effectiveness for the speech commands (Table 5).

Although the touch gestures on the smartwatch did not require vision, using a separate device in addition to the HMD seemed to distract the participants. It was difficult for participants to locate and track the letters on the screen, especially under a high magnification level. Many participants (e.g., I2, I8, I10) felt that the letters “moved” much more when using the smartwatch than when using the speech commands, even though they used the same HMD to look at the same set of letters. We observed that although the participants did not look at the watch surface during the control, their head and body moved with their hand gestures, which caused the magnified letter to appear to move. Most participants did not notice their head movement, saying they kept their head still. This may be because when they started adjusting the enhancements, their attention switched from the HMD to their hands, and they unconsciously moved their head to the object of their attention.

Interestingly, I2 noticed her body movement and said she did this to better conduct the touch gestures:

“The speech is much easier. [When using the smartwatch], there’s some kind of movement on the screen of that letter. I had a much harder time finding the letter...I’m moving my body, to increase [the magnification level] I’m moving to the right, and to decrease it I’m moving to the left. When I’m doing a gesture, my whole body is moving with it. [The touch gesture] is a whole different experience to me.” (I2)

This was also the main reason that I2 gave a very low Effectiveness score (2) for the touch gestures on the smartwatch (Table 5).

However, three participants (I5, I8, and I11) spent more time using the speech commands than using the smartwatch. They felt the commands were not natural enough, because they had to memorize specific language. Since the current design only accepted specific words to trigger the commands, participants sometimes spoke incorrect commands and could not recall the correct ones immediately, which resulted in longer times to complete the reading task. For example, some participants used “reduce magnification” (the correct command is “decrease magnification”), but got the feedback of “wrong command.” Even participants who were faster when using the speech commands felt that it was challenging to memorize all the commands. I6 said, “I hate fixed commands, because we don’t talk like that... Every time when I wanted to say a command, I have to think about it, because I have to say it in a certain way.” This explained why the speech commands

got lower Usability scores than the touch gestures. Moreover, I7 preferred the smartwatch control, because it protected his privacy better than the speech commands.

All participants liked the fact that they had two interaction options, so that they could choose a suitable one for a given situation. “If I was outside, the only con about speech is the noise. It may be too loud. [The system] may not hear me unless. Also, the [smartwatch] controls are good for a quiet place where you don’t want to be singled out. But when you are outside in a public and open space, and maybe you have some bags and walk down the street, you want to look at what the street sign is, you can just say ‘set magnification to blah blah’ or ‘increase contrast,’ that is totally cool! [Combining the two interaction modalities] makes the system comprehensive and powerful” (I6).

6.2.3 Enhancement Selections When Reading the Letters. All participants used Magnification, indicating that Magnification was the most effective enhancement for this reading task. I1 and I10 also used Black/White Reversal at the same time. Although all participants used Magnification, we found that participants used different magnification levels when reading the letters. Many of them also varied the magnification levels to see letters at different sizes. This echoed our prior finding that customization was important for low vision tools.

When observing participants as they optimized their visual experience, we found that some tried to find a specific pattern that worked best for them, and then used that pattern in the following trials. For example, I3 and I9 used Magnification, Black/White Reversal, and Contrast Enhancement for the first few trials, and found that Magnification was the most effective enhancement. So they continued to use only Magnification in the rest of the study. As another example, when I1 and I7 used the smartwatch, they tried different Magnification levels in the first few trials. When they found a specific level that worked well for all different letter sizes (I1: Black/White Reversal + Magnification with level 7, I7: Magnification with level 10), they used this setting in the rest of the study. Even participants who did not show clear pattern seeking behavior tended to use the same enhancement combinations throughout the study (i.e., I1 and I10 used both Black/White Reversal and Magnification, while others only used Magnification).

Participants emphasized the importance of saving a preferred combination of enhancements and using a shortcut to trigger it. As I9 mentioned, “It would be nice if they stayed, like you set personal preferences and then that’s just the way it always was...you are using the device as an enhancement, so whatever you just setting up the enhancements is the way the best for you to typically see things. And every time you turn it on, it should just set to those preferences.”

6.3 Discussion: Designing and Evaluating Interaction Techniques

Our study demonstrated that both the smartwatch gestures and the speech commands were effective interaction techniques, which enabled low vision users to customize their visual experience and complete a far-distance letter reading task within a reasonable amount of time. All participants liked the design of our interaction techniques and felt they were easy to learn and use. Moreover, participants enjoyed having two interaction options, thus they could choose suitable interactions for different scenarios. For example, using the speech commands on the go, while using the touch gestures in a quiet environment (e.g., a classroom).

The most challenging aspect of using the smartwatch was participants’ tendency to unconsciously move their head and body when interacting with it. This made it difficult to find and track the letters on the HMD. This may be because participants were not used to conducting gestures on a smartwatch, especially when not looking down at its screen. It is possible that their head and body movements would lessen after some practice. Future work should consider a long-term user study, observing the participants’ behaviors in their daily activities after they are fully

familiar with the interaction techniques. A different hardware design can also mitigate head and body movements. Instead of a smartwatch, we can add a touchpad to the HMD itself, as in Google Glass, allowing users to interact with the HMD without switching their attention to another device. However, using the smartwatch does eliminate the need to rely on specific HMD designs, allowing us to leverage mainstream devices.

In terms of limitations, our study involved only one task: reading a single letter that appeared in high-contrast on a digital display. Although this is a challenging and important task, many other viewing tasks need to be considered to evaluate the effectiveness of our interaction techniques, for example, reading a long paragraph, or recognizing faces. As ForeSee has provided a platform to enable low vision users to select and adjust different enhancements, many other user studies could be conducted to investigate low vision people's visual experiences in different visual tasks based on this platform. Moreover, since ForeSee provides such a new experience, the scores participants raised might be influenced by a novelty effect. Future research should consider a long-term evaluation to reduce this impact.

7 DISCUSSION

Our research has demonstrated the potential of HMDs to benefit people with low vision. In the first study, we showed that typical low vision enhancements such as magnification and contrast enhancements are powerful in a head-mounted setting. Importantly, people adjusted the parameters of the enhancements and combined them for the task and their particular visual ability. In the second study, we showed that our designs of speech commands and smartwatch-based interaction techniques enabled low vision people to effectively adjust and combine enhancements. These interaction techniques can be used to adjust an HMD in a variety of scenarios according to the user's preference, allowing people to customize their visual experience for different daily tasks. Of course, one prominent limitation at this time is the sheer size of the HMD itself. In its current form, people with low vision are unlikely to use such a device, especially when on-the-go. We hope that as HMD platforms continue to evolve, our work will spur innovation on how these devices can best be used to provide access for people with low vision.

We discuss design considerations for future HMD applications for low vision people that emerged from our research. First, customization (or "adaptability"⁴ [29, 97]) was extremely useful for low vision users to accommodate their visual experiences to their visual abilities and the current viewing tasks. Participants showed different preferences for enhancements in our study. In fact, they wanted even more ways to customize the system. When using Text Extraction, for example, they wanted to be able to modify the properties of the extracted digital text. Following a medical model of disability, prior work on low vision has tended to focus on specific eye diseases, such as Glaucoma or Retinitis Pigmentosa (e.g., Reference [107]). However, we found that there was as much diversity of preferences among people with similar vision conditions as there was among people with different vision conditions. For example, P6, P9, and P10 all had Retinitis Pigmentosa with a limited field of view, but their preferences for vision enhancements were largely different in both the reading and the image perception tasks (Figure 5). The medical model of low vision rehabilitation typically uses clinical measures and etiologies to guide design. However, our research demonstrates the importance of empowering users by elevating their subjective experiences. Giudice [32] has also argued that the medical practices are not necessarily the best technique for the design of assistive technologies, emphasizing the importance of users' subjective

⁴Adaptability means that the system provides customization mechanisms but relies on the user to use those mechanisms. Meanwhile, adaptivity means that the system has the capacity to dynamically adjust its interface in a way that is intended to support the user [29].

requirements and experiences. Performing daily tasks such as reading a novel or looking at a loved one's face are more than just a function of light entering a sensory organ. They are the convolution of physical, social, and emotional processes based on our physiological makeup as well as our attitudes and desires. We are only beginning to understand how vision can be enhanced to improve people's daily experiences, but, as a starting point, our research demonstrates the importance of providing a flexible system that allows users to adjust their visual experience according to their ability, preference, and circumstance.

Although customization was beneficial to low vision users, adjusting the system could be difficult and time-consuming. Participants wanted the HMD systems to be adaptive by saving and reloading their preferences based on use scenarios. Although there are no clear guidelines for which enhancements are best, the system could automatically recommend the most suitable enhancement combination according to the environment (e.g., amount of light), the user's visual conditions (e.g., the user has limited peripheral vision), and the task (e.g., the user is reading materials up-close) to make interaction more efficient. Designers could consider using machine learning and additional sensors to learn the user's preferences in different environmental conditions to create an adaptive system that learns from the user's past inputs. Users should still be able to adjust and reinforce these automatic settings.

Our research showed that traditional enhancements (those used in ForeSee) are not a one-stop fix for low vision. As described in the design of ForeSee (Section 3), we studied enhancement methods that were commonly used for low vision accessibility, abundant in low and high-tech tools. In our study, however, we found that some enhancement methods did not work well for some participants. For example, contrast enhancement is supported in almost all video and screen magnifiers, but we found that it did not benefit people who were sensitive to light, because it generated too much light. Moreover, although participants generally found magnification to be the most effective enhancement, many participants (e.g., I2, I8, I10) struggled to locate the target under high magnification levels, making the reading task slow and strenuous. This is an example of one visual challenge (visual search) that was not addressed by traditional enhancements. Researchers must continue to study visual perception and experience and innovate new methods to complement (if not replace) these traditional enhancements in different tasks. For example, when the magnification level was high, visual cues [120] could be used to orient the users' attention and help them locate the target text or object. Beyond visual augmentation, designers could also consider multimodal feedback, leveraging audio and tactile feedback to compensate for the lack of visual information.

We now turn to broader research directions that are spurred by our work. Our research indicates that there is much potential for video see-through smartglasses. Video see-through (VST) glasses are smartglass platforms that capture the user's view with a camera, modify it, and present it on a digital display that is mounted on the user's face. By contrast, optical see-through (OST) glasses project light over the user's vision and allow users to see the real world with their natural vision (e.g., Google Glass, Epson Moverio, HoloLens). As AR and smartglass technologies rapidly advance, VST glasses have drawn much less attention than OST glasses, since they are bulky, cover the users' eyes, and cause disorientation or "VR sickness" [121]. However, for people with low vision, VST glasses could be powerful assistive tools. They allow a complete manipulation of the user's view of the world, providing a unique opportunity to improve low vision people's vision directly using vision enhancements and augmentations. Interestingly, none of our low vision participants indicated feelings of VR sickness during all the studies in this article. This is perhaps explained by the fact that low vision people tend to be less sensitive to the quality of the visual displays (e.g., resolution, frame rate) and many of them had a limited field of view. Simulating these conditions (e.g., increasing the frame rate of the rendering, reducing the field of view) has

been shown to reduce VR sickness in typically sighted people [28, 52, 55, 75]. VST glasses for low vision are already commercially available. For example, eSight [26] and NuEyes [70] are relatively lightweight VST glasses that look more like sunglasses than a phone mounted on one's face. Both these VST products and our promising findings showed that, with the rapid advance of computer technology, VST glasses could also be the future of smartglasses, not only for people with low vision but also for sighted users. When designing applications for smartglasses, designers should also take into account VST glasses as a potential platform, designing features that can leverage the advantages of VST platforms rather than only focusing on OST glasses.

With HMD devices getting more and more popular, people need suitable and flexible interaction techniques to control the HMDs. We are the first to explore non-speech interaction techniques on HMDs for low vision (and one of the first to explore interaction techniques for low vision on any platform). Although existing smartglass products for low vision include controllers to allow users to adjust their visual experience (e.g., eSight is connected to a controller with buttons, and NuEyes has a wireless controller with buttons), they used specialized control devices. Moreover, these controllers simply used specific physical buttons to control specific enhancements, so that the number of buttons would increase linearly with the increase of supported enhancements or system complexity. It is necessary to post design rules for smartglass interaction techniques for people with low vision, and design rich interaction techniques to enable low vision users to interact with different smartglass applications rather than just selecting and adjusting basic vision enhancements. Our research was the first exploration in this direction, yielding design insights for non-speech interaction techniques on HMDs for low vision. For example, touch gestures on touchscreen-based devices (e.g., smartphones, smartwatches) were effective to control smartglasses instead of a specialized device. A linear menu with one menu item on the glass display at a time is preferable, since it is easier for low vision users to locate and read the menu item. Moreover, continuous sliding on the touchscreen can be used for a slider-type interaction, however, it is important to design noticeable feedback to inform users of the current status of the slider. While our study investigated some basic interaction techniques (i.e., selection, menu navigation, slider control), more research is needed to refine these interactions and also to extend to more interaction techniques such as target acquisition, and text entry.

Last, we summarize the insights and lessons we learnt from our studies with people with low vision. In our study, we used a qualitative method to evaluate the enhancement methods on an HMD. Most previous work on HMD systems used quantitative studies to evaluate the system by asking participants to conduct a specific task (e.g., a visual search task [27, 57]) and measuring the speed and accuracy of the system performance. None of them focused on users' visual experience and preferences when using a head-mounted vision enhancement system. Beyond quantitatively evaluating the interaction techniques, we adopted a different approach—a qualitative study to evaluate the enhancement methods by analyzing users' feedback and preferences for each enhancement in different viewing tasks. Users' subjective feedback revealed more than expected. Because a qualitative study is open-ended, participants discussed their experiences with the different enhancements during and beyond the study; they mentioned similar tools they had used in the past, described specific situations where certain configurations of ForeSee may be helpful, and suggested improvements to the current implementation. Our findings yielded a deeper and more complete understanding of different conventional enhancement methods for different tasks and visual conditions, providing insights that helped us understand low vision, and form guidelines for designing future HMD applications.

We also faced several challenges when conducting the user studies that were specific to low vision participants. Visual perception is complex, and it was challenging to understand the users' visual experience. Although we asked participants' visual conditions at the beginning of the study,

they sometimes did not know how to quantify their visual ability (e.g., acuity, visual field), and their visual experience did not always align with their diagnosis. Moreover, during the study, some participants also had difficulty describing what they saw and how the enhancement methods affected their visual experience. For example, when we showed P20 different enhancements in the study of evaluating the enhancement methods, she usually gave vague comments such as “this is good” or “I can see better,” but had a hard time describing why she could see better and what the difference was between enhancement methods. We had to explain our questions several times and verify our understanding of the participants’ description to ensure that we understood their experience.

While understanding visual perception was difficult, we found that asking participants to customize their experience by combining enhancements helped us understand their visual experience and the effect of the enhancements. For example, when experiencing different enhancements, P2 gave similar feedback for the enhancements in full display mode and window display mode, but she finally chose full display mode for most viewing tasks and only window display mode in the near-distance reading. This implied that she actually had a different visual experience when using these two display modes, so we asked her why she preferred the different display modes for each task. Customizing the system allowed users to clearly express a preference and allowed us to better understand their experiences.

Our studies had some limitations, which provide directions for future work. First, some of the vision enhancements can be further improved to better evaluate their effect on users’ experiences. For example, when enhancing text with edge enhancement, we will consider adjusting the distance between letters to reduce the crowding effect (i.e., the letters get closer than a critical spacing, making the reading content difficult to be identified [80]). Second, while our current study focuses on people’s qualitative experiences, quantitative measures would provide a more comprehensive understanding of the effectiveness of different vision enhancements on the HMD, such as measuring users’ reading speed when using ForeSee. More data regarding users’ behaviors (e.g., head coordination, gaze direction) could also be collected to understand how people use ForeSee in terms of both vision enhancements and interaction techniques.

8 CONCLUSION

In this article, we explored both the output and input of a customizable vision enhancement HMD system for people with low vision. We first described an extensive user study with ForeSee to explore low vision people’s visual experiences when using different vision enhancements in different viewing tasks. We then designed speech commands and smartwatch-based interaction techniques to enable users to flexibly customize their visual experience. Our study demonstrated the potential of HMDs to benefit people with low vision. Participants were able to read letters and recognize images by adjusting and combining different vision enhancements in an HMD setting. The interaction techniques we designed effectively enabled low vision people to interact with the HMD system and customize their visual experience. We discussed design considerations for HMD applications and spurred new research directions. As augmented and virtual reality technologies continue to advance, we believe HMDs with well-designed interaction techniques will become powerful vision enhancement tools for people with low and even typical vision.

A APPENDIX

A.1 Interaction Flow for the Window Display Mode

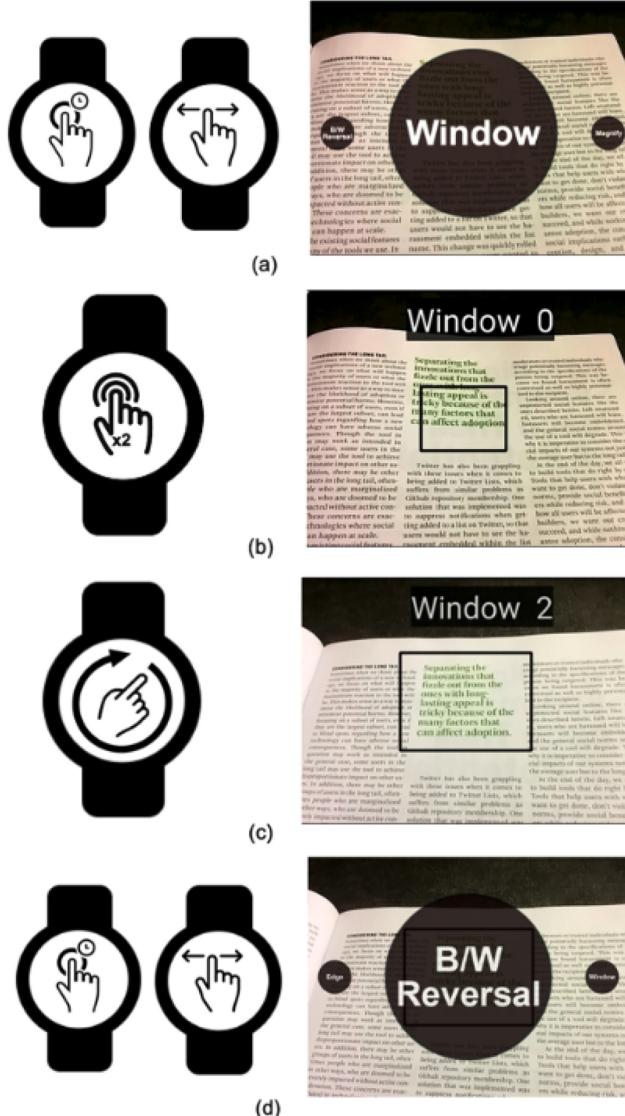


Figure A1. Interaction Flow of ForeSee when using the Window Display Mode (the images on the right column are the screen shot images a user can see through ForeSee in each interaction step, and the smartwatch figures on the left column represent the gestures to trigger a corresponding step). This figure illustrates the steps of turning on the Window Display Mode and adding Black/White Reversal and Magnification in the Window. (a) Long press to trigger the menu and swipe to the left or right to switch to the “Window” menu item; (b) Double tap to turn on the Window Display Mode; (c) Clockwise slide to increase the window size. The user can move her head to make sure the text she wants to read is in the Window; (d) Long press to trigger the menu and swipe to the left or right to switch to the “B/W Reversal” menu item; (e) Double tap to trigger the Black/White Reversal enhancement in the Window; (f) Long press to trigger the menu and swipe to the left or right to switch to the “Magnification” menu item; (g) Double tap to trigger the Magnification enhancement in the Window; (h) Clockwise slide to increase the Magnification level in the Window.



Figure A1. (Continued).

ACKNOWLEDGMENTS

We thank all participants for their time and helpful feedback.

REFERENCES

- [1] P. F. Adams and P. M. Barnes. 2006. Summary health statistics for the US population: National Health Interview Survey, 2004. *Nat. Center Health Stat.* 10, 229 (2006), 1–104.
- [2] American Optometric Association (AOA). 2015. What Causes Low Vision? Retrieved from <http://www.aoa.org/patients-and-public/caring-for-your-vision/low-vision/what-causes-low-vision?ss0=y>.
- [3] Aries Ardit. 2004. Adjustable typography: An approach to enhancing low vision text accessibility. *Ergonomics* 47, 5 (2004), 469–482.
- [4] Aria. 2018. Aria. Retrieved from <https://aria.io/>.
- [5] Shiri Azenkot and Nicole B. Lee. 2013. Exploring the use of speech input by blind people on mobile devices. In *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, 11.
- [6] Ronald T. Azuma. 1997. A survey of augmented reality. *Presence: Teleop. Virtual Environ.* 6, 4 (1997), 355–385.
- [7] Jean-Baptiste Bernard, Emilien Tlapale, Geraldine Faure, Eric Castet, and Pierre Kornprobst. 2008. Navisio: Towards an integrated reading aid system for low vision patients. In *Workshop on Computer Vision Applications for the Visually Impaired*. <https://hal.inria.fr/hal-02321739/>.
- [8] Jeffrey P. Bigham. 2014. Making the web easier to see with opportunistic accessibility improvement. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*. ACM, 117–122.
- [9] Philip Burnard. 1991. A method of analysing interview transcripts in qualitative research. *Nurse Edu. Today* 11, 6 (1991), 461–466.
- [10] John Canny. 1986. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.* 6 (1986), 679–698.
- [11] Low Vision Center. 2018. LVC—Closed Circuit Television. Retrieved from <http://www.lowvisioninfo.org/hightech.htm>.
- [12] Verena R. Cimarolli and Kathrin Boerner. 2005. Social support and well-being in adults who are visually impaired. *J. Visual Impair. Blind.* 99, 9 (2005), 521–534.
- [13] Verena R. Cimarolli, Kathrin Boerner, Mark Brennan-Ing, Joann P. Reinhardt, and Amy Horowitz. 2012. Challenges faced by older adults with vision loss: A qualitative study with implications for rehabilitation. *Clin. Rehab.* 26, 8 (2012), 748–757.
- [14] Verena R. Cimarolli, Kathrin Boerner, and Shu-wen Wang. 2006. Life goals in vision rehabilitation: Are they addressed and how? *J. Visual Impair. Blind.* 100, 6 (2006), 343–352.
- [15] Verena R. Cimarolli, Joann P. Reinhardt, and Amy Horowitz. 2006. Perceived overprotection: Support gone bad? *J. Gerontol. Series B: Psychol. Sci. Soc. Sci.* 61, 1 (2006), S18–S23.
- [16] Verena R. Cimarolli, Carol J. Sussman-Skalka, and Caryn R. Goodman. 2004. “Program for partners”: Support groups for partners of adults with visual impairments. *J. Visual Impair. Blind.* 98, 2 (2004), 90–98.
- [17] Verena R. Cimarolli and Shu-wen Wang. 2006. Differences in social support among employed and unemployed adults who are visually impaired. *J. Visual Impair. Blind.* 100, 9 (2006), 545–556.
- [18] August Colenbrander. 2005. Visual functions and functional vision. In *International Congress Series*, Vol. 1282. Elsevier, 482–486.
- [19] August Colenbrander. 2010. Assessment of functional vision and its rehabilitation. *Acta Ophthalmologica* 88, 2 (2010), 163–173.
- [20] Mike Cooley. 2000. Human-centered design. *Information Design*, Robert E. Jacobson (Ed.), 59–81. <https://books.google.com/books?hl=en&lr=&id=vnax4nN4Ws4C&oi=fnd&pg=PA59&dq=Human-centered+design+mike+cooley&ots=PaK4-7osBa&sig=QPKr1tSKhxPMuUTjhmvHVa3vRU#v=onepage&q=Human-centered%20design%20mike%20cooley&f=false>.
- [21] Anne Lesley Corn and Jane N. Erin. 2010. *Foundations of Low Vision: Clinical and Functional Perspectives*. American Foundation for the Blind.
- [22] Dimitrios Dakopoulos and Nikolaos G. Bourbakis. 2009. Wearable obstacle avoidance electronic travel aids for blind: A survey. *IEEE Trans. Syst. Man Cybernet. Part C (Appl. Rev.)* 40, 1 (2009), 25–35.
- [23] Ashley D. Deemer, Christopher K. Bradley, Nicole C. Ross, Danielle M. Natale, Rath Itthipanichpong, Frank S. Werblin, and Robert W. Massof. 2018. Low vision enhancement with head-mounted video display systems: Are we there yet? *Optom. Vision Sci.* 95, 9 (2018), 694–703.
- [24] Bradley E. Dougherty, K. Bradley Kehler, Richard Jamara, Nicole Patterson, Denise Valenti, and Fuensanta A. Vera-Diaz. 2011. Abandonment of low vision devices in an outpatient population. *Optom. Vision Sci.* 88, 11 (2011), 1283.
- [25] Epson. 2015. Epson Moverio BT-200 Next Generation Smart Glasses—Epson America, Inc. Retrieved from http://www.epson.com/cgi-bin/Store/jsp/Landing/moverio-bt-200-smart-glasses.do?ref=van_moverio_2014.

- [26] eSight. 2018. eSight Electronic Glasses. Retrieved from <https://www.esighteyewear.com/homex>.
- [27] M. R. Everingham, B. T. Thomas, T. Troscianko et al. 1999. Head-mounted mobility aid for low vision using scene classification techniques. *Int. J. Virtual Real.* 3, 4 (1999), 3.
- [28] Ajoy S. Fernandes and Steven K. Feiner. 2016. Combating VR sickness through subtle dynamic field-of-view modification. In *Proceedings of the IEEE Symposium on 3D User Interfaces (3DUI'16)*. IEEE, 201–210.
- [29] Leah Findlater and Joanna McGrenere. 2004. A comparison of static, adaptive, and adaptable menus. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 89–96.
- [30] American Foundation for the Blind (AFB). 2018. Screen Readers and Screen Magnifiers: An Introduction to Computer Accessibility Software. Retrieved from <http://suo.im/5rv11h>.
- [31] Julie Fraser and Carl Gutwin. 2000. A framework of assistive pointers for low vision users. In *Proceedings of the 4th International ACM Conference on Assistive Technologies*. ACM, 9–16.
- [32] Nicholas A. Giudice. 2018. Navigating without vision: Principles of blind spatial cognition. In *Handbook of Behavioral and Cognitive Geography*. Edward Elgar Publishing.
- [33] Google. [n.d.]. Android accessibility overview—Android Accessibility Help. Retrieved from <https://support.google.com/accessibility/android/answer/6006564?hl=en>.
- [34] Google. [n.d.]. Google Glass Help. Retrieved from <https://support.google.com/glass#topic=4363345>.
- [35] Robert Harper, Louise Culham, and Christine Dickinson. 1999. Head mounted video magnification devices for low vision rehabilitation: a comparison with existing technology. *Brit. J. Ophthalmol.* 83, 4 (1999), 495–500.
- [36] David Hayden, Dirk Colbry, John A. Black Jr., and Sethuraman Panchanathan. 2008. Note-taker: Enabling students who are legally blind to take notes in class. In *Proceedings of the 10th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, 81–88.
- [37] David S. Hayden, Michael Astrauskas, Qian Yan, Liqing Zhou, and John A. Black Jr. 2011. Note-taker 3.0: An assistive technology enabling students who are legally blind to take notes in class. In *Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, 269–270.
- [38] David S. Hayden, Liqing Zhou, Michael J. Astrauskas, and John A. Black Jr. 2010. Note-taker 2.0: The next step toward enabling students who are legally blind to take notes in class. In *Proceedings of the 12th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, 131–138.
- [39] Stephen L. Hicks, Iain Wilson, Louwai Muhammed, John Worsfold, Susan M. Downes, and Christopher Kennard. 2013. A depth-based head-mounted visual display to aid navigation in partially sighted individuals. *PLoS One* 8, 7 (2013), e67695.
- [40] Clare Hocking. 1999. Function or feelings: Factors in abandonment of assistive devices. *Technol. Disabil.* 11, 1, 2 (1999), 3–11.
- [41] Bill Holton. 2014. A review of iOS access for all: Your comprehensive guide to accessibility for iPad, iPhone, and iPod touch, by Shelly Brisbin. *AccessWorld Mag.* 15, 7 (2014).
- [42] Jonathan Huang, Max Kinateder, Matt J. Dunn, Wojciech Jarosz, Xing-Dong Yang, and Emily A. Cooper. 2019. An augmented reality sign-reading assistant for users with reduced vision. *PLoS One* 14, 1 (2019), e0210630.
- [43] Chuck Huss and Anne Corn. 2004. Low vision driving with bi optics: An overview. *J. Visual Impair. Blind.* 98, 10 (2004), 641–653.
- [44] Alex D. Hwang and Eli Peli. 2014. An augmented-reality edge enhancement application for Google glass. *Optom. Vision Sci.* 91, 8 (2014), 1021.
- [45] ICD.Codes. [n.d.]. ICD-10-CM Code H54.10. Retrieved from <https://icd.codes/icd10cm/H5410>.
- [46] Julie A. Jacko and Andrew Sears. 1998. Designing interfaces for an overlooked user group: Considering the visual profiles of partially sighted users. In *Proceedings of the 3rd International ACM Conference on Assistive Technologies*. ACM, 75–77.
- [47] Pamela E. Jeter, Collin Rozanski, Robert Massof, Olukemi Adeyemo, and Gislin Dagnelie. 2017. Development of the ultra-low vision visual functioning questionnaire (ULV-VFQ). *Translat. Vision Scie. Technol.* 6, 3 (2017), 11–11.
- [48] Shaun K. Kane, Jeffrey P. Bigham, and Jacob O. Wobbrock. 2008. Slide rule: Making mobile touchscreens accessible to blind people using multi-touch interaction techniques. In *Proceedings of the 10th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, 73–80.
- [49] Shaun K. Kane, Chandrika Jayant, Jacob O. Wobbrock, and Richard E. Ladner. 2009. Freedom to roam: A study of mobile device adoption and accessibility for people with visual and motor disabilities. In *Proceedings of the 11th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, 115–122.
- [50] Max Kinateder, Justin Gualtieri, Matt J. Dunn, Wojciech Jarosz, Xing-Dong Yang, and Emily A. Cooper. 2018. Using an augmented reality device as a distance-based vision aid—promise and Limitations. *Optom. Vision Sci.* 95, 9 (2018), 727.
- [51] Richard L. Kline and Ephraim P. Glinert. 1995. Improving GUI accessibility for people with low vision. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM Press/Addison-Wesley Publishing, 114–121.

- [52] Eugenia M. Kolasinski. 1995. *Simulator Sickness in Virtual Environments*. Technical Report. Army Research Institute for the Behavioral and Social Sciences, Alexandria, VA.
- [53] Raja S. Kushalnagar, Stephanie A. Ludie, and Poorna Kushalnagar. 2011. Multi-view platform: An accessible live classroom viewing approach for low vision students. In *The Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, 267–268.
- [54] MiYoung Kwon, Chaithanya Ramachandra, PremNandhini Satgunam, Bartlett W. Mel, Eli Peli, and Bosco S. Tjan. 2012. Contour enhancement benefits older adults with simulated central field loss. *Optom. Vision Sci.* 89, 9 (2012), 1374.
- [55] J. J.-W. Lin, Henry Been-Lirn Duh, Donald E. Parker, Habib Abi-Rached, and Thomas A. Furness. 2002. Effects of field of view on presence, enjoyment, memory, and simulator sickness in a virtual environment. In *Proceedings of the IEEE Conference on Virtual Reality*. IEEE, 164–171.
- [56] Jan E. Lovie-Kitchin and George C. Woo. 1988. Effect of magnification and field of view on reading speed using a CCTV. *Ophthalm. Physiol. Optics* 8, 2 (1988), 139–145.
- [57] Gang Luo and Eli Peli. 2006. Use of an augmented-vision device for visual search by patients with tunnel vision. *Invest. Ophthalmol. Visual Sci.* 47, 9 (2006), 4152–4159.
- [58] MAGic. [n.d.]. The best screen magnifier—MAGIC. Retrieved from <http://www.freedomscientific.com/Products/LowVision/MAGIC> last accessed 23 August 2019.
- [59] Tom H. Margrain. 2000. Helping blind and partially sighted people to read: The effectiveness of low vision aids. *Brit. J. Ophthalmol.* 84, 8 (2000), 919–921.
- [60] Robert W. Massof and Lorraine Lidoff. 2001. *Issues in Low Vision Rehabilitation: Service Delivery, Policy, and Funding*. American Foundation for the Blind.
- [61] Robert W. Massof and Douglas L. Rickman. 1992. Obstacles encountered in the development of the low vision enhancement system. *Optom. Vision Sci.* 69, 1 (1992), 32–41.
- [62] Robert W. Massof, Douglas L. Rickman, Peter A. Lalle et al. 1994. Low vision enhancement system. *Johns Hopkins APL Tech. Digest* 15, 2 (1994), 120–125.
- [63] Peter McLeod, Jon Driver, and Jennie Crisp. 1988. Visual search for a conjunction of movement and form is parallel. *Nature* 332, 6160 (1988), 154.
- [64] Carlos Merino-Gracia, Karel Lenc, and Majid Mirmehdi. 2011. A head-mounted device for recognizing text in natural scenes. In *Proceedings of the International Workshop on Camera-based Document Analysis and Recognition*. Springer, 29–41.
- [65] Microsoft. [n.d.]. Microsoft HoloLens | Official Site. Retrieved from <https://www.microsoft.com/microsoft-hololens/en-us>.
- [66] J. Elton Moore, B. J. LeJeune, and Low Vision Devices. 2008. Low vision. *International Encyclopedia of Rehabilitation*. Retrieved from <http://cirrie.buffalo.edu/encyclopedia/article.php>.
- [67] Dan Morris, T. Scott Saponas, Desney Tan et al. 2011. Emerging input technologies for always-available mobile interaction. *Found. Trends Hum.-Comput. Interact.* 4, 4 (2011), 245–316.
- [68] Alan R. Morse, Robert W. Massof, Roy G. Cole, Lylas G. Mogk, Annemarie M. O’Hearn, Yu-Pin Hsu, Eleanor E. Faye, Stanley F. Wainapel, and Mary Lou Jackson. 2010. Medicare coverage for vision assistive equipment. *Arch. Ophthalmol.* 128, 10 (2010), 1350–1357.
- [69] Lukáš Neumann and Jiří Matas. 2012. Real-time scene text localization and recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 3538–3545.
- [70] NuEyes. [n.d.]. NuEyes—Electronic Glasses for Low Vision. Retrieved from <https://nueyes.com/>.
- [71] Oculus. [n.d.]. Oculus Rift Development Kit 2 Order Page | Oculus Rift—Virtual Reality Headset for 3D Gaming. Retrieved from <https://www1.oculus.com/order/>.
- [72] OrCam. 2015. OrCam—See for Yourself. Retrieved from <http://www.orcam.com/>.
- [73] Nobuyuki Otsu. 1979. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybernet.* 9, 1 (1979), 62–66.
- [74] Phil Parette and Marcia Scherer. 2004. Assistive technology use and stigma. *Edu. Train. Dev. Disabil.* 39, 3 (Sept. 2004), 217–226.
- [75] Randy Pausch, Thomas Crea, and Matthew Conway. 1992. A literature survey for virtual environments: Military flight simulator visual systems and simulator sickness. *Presence: Teleop. Virtual Environ.* 1, 3 (1992), 344–363.
- [76] Eli Peli. 1999. Simple 1-D image enhancement for head-mounted low vision aid. *Visual Impair. Res.* 1, 1 (1999), 3–10.
- [77] Eli Peli, Gang Luo, Alex Bowers, and Noa Rensing. 2007. Applications of augmented-vision head-mounted systems in vision rehabilitation. *J. Soc. Info. Display* 15, 12 (2007), 1037–1045.
- [78] Eli Peli, Gang Luo, Alex Bowers, and Noa Rensing. 2009. Development and evaluation of vision multiplexing devices for vision impairments. *Int. J. Artific. Intelligence Tools* 18, 03 (2009), 365–378.

- [79] Eli Peli and Russell L. Woods. 2009. Image enhancement for impaired vision: The challenge of evaluation. *Int. J. Artific. Intell. Tools* 18, 03 (2009), 415–438.
- [80] Denis G. Pelli. 2008. Crowding: A cortical constraint on object recognition. *Curr. Opin. Neurobiol.* 18, 4 (2008), 445–451.
- [81] Rachael C. Peterson, James S. Wolffsohn, Martin Rubinstein, and John Lowe. 2003. Benefits of electronic vision enhancement systems (EVES) for the visually impaired. *Amer. J. Ophthalmol.* 136, 6 (2003), 1129–1135.
- [82] Betsy Phillips and Hongxin Zhao. 1993. Predictors of assistive technology abandonment. *Assist. Technol.* 5, 1 (1993), 36–45.
- [83] Alisha Pradhan, Kanika Mehta, and Leah Findlater. 2018. Accessibility came by accident: Use of voice-controlled intelligent personal assistants by people with disabilities. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. ACM, 459.
- [84] Marti L. Riemer-Reiss and Robbyn R. Wacker. 2000. Factors associated with assistive technology discontinuance among individuals with disabilities. *J. Rehabil.* 66, 3 (2000).
- [85] Daniel Salber and Joëlle Coutaz. 1993. Applying the wizard of oz technique to the study of multimodal systems. In *Proceedings of the International Conference on Human-Computer Interaction*. Springer, 219–230.
- [86] Samsung. [n.d.]. Samsung Gear. Retrieved from <http://www.samsung.com/global/microsite/gearvr/>.
- [87] Prem Nandhini Satgunam, Russell L. Woods, Gang Luo, P. Matthew Bronstad, Zachary Reynolds, Chaithanya Ramachandra, Bartlett W. Mel, and Eli Peli. 2012. Effects of contour enhancement on low-vision preference and visual search. *Optom. Vision Sci.* 89, 9 (2012), E1364.
- [88] Mitchell Scheiman, Maxine Scheiman, and Steven Whittaker. 2007. *Low Vision Rehabilitation: A Practical Guide for Occupational Therapists*. Slack Incorporated.
- [89] Freedom Scientific. [n.d.]. JAWS (Job Access with Speech) for Windows (Professional). Retrieved from <https://www.freedomscientific.com/products/software/jaws/>.
- [90] Liz Segre. 2017. Low Vision Aids for Reading. Retrieved from <https://www.allaboutvision.com/lowvision/reading>.
- [91] Kristen Shinohara and Jacob O. Wobbrock. 2011. In the shadow of misperception: Assistive technology use and social interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 705–714.
- [92] Liz Simon. 2008. Low vision and rehabilitation for older people: Integrating services into the health care system. *Commun. Eye Health* 21, 66 (2008), 28.
- [93] Sony. [n.d.]. Develop Apps for SmartEyeglass. Retrieved from <https://developer.sony.com/develop/wearables/smarteyeglass-sdk/>.
- [94] Ai Squared. [n.d.]. Ai Squared—We've got accessibility covered. Retrieved from <http://www.aisquared.com/zoomtext>.
- [95] Lee Stearns, Victor DeSouza, Jessica Yin, Leah Findlater, and Jon E. Froehlich. 2017. Augmented reality magnification for low vision users with the microsoft hololens and a finger-worn camera. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, 361–362.
- [96] Lee Stearns, Leah Findlater, and Jon E. Froehlich. 2018. Design of an Augmented Reality Magnification Aid for Low Vision Users. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, 28–39.
- [97] Constantine Stephanidis, Demosthenes Akoumianakis, and Anthony Savidis. 1995. Design representations and development support for user interface adaptation. In *Proceedings of the 1st ERCIM Workshop on User Interfaces for All (UI4All'95)*. ICS-FORTH.
- [98] Richard Szeliski. 2010. *Computer Vision: Algorithms and Applications*. Springer Science & Business Media.
- [99] Sarit Szpiro, Yuhang Zhao, and Shiri Azenkot. 2016. Finding a store, searching for a product: A study of daily challenges of low vision people. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 61–72.
- [100] Sarit Felicia Anais Szpiro, Shafeqa Hashash, Yuhang Zhao, and Shiri Azenkot. 2016. How people with low vision access computing devices: Understanding challenges and opportunities. In *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, 171–180.
- [101] Enrico Tanuwidjaja, Derek Huynh, Kirsten Koa, Calvin Nguyen, Churen Shao, Patrick Torbett, Colleen Emmenegger, and Nadir Weibel. 2014. Chroma: A wearable augmented-reality solution for color blindness. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 799–810.
- [102] tesseract ocr. 2018. tesseract-ocr-GitHub. Retrieved from <https://github.com/tesseract-ocr/>.
- [103] Mary Frances Theofanos and Janice Ginny Redish. 2005. Helping low-vision and other users with web sites that meet their needs: Is one site for all feasible? *Techn. Commun.* 52, 1 (2005), 9–20.
- [104] Soeren Thierfelder, Barbara Lege, and Frank Ulrich. 1998. First clinical experience on the Low-Vision-Enhancement System (LVESR)—A new type of optoelectronic rehabilitation device. *Der Ophthalmol.* 95, 11 (1998), 781–783.

- [105] James M. Tielsch, Alfred Sommer, Kathe Witt, Joanne Katz, and Richard M. Royall. 1990. Blindness and visual impairment in an American urban population: The Baltimore eye survey. *Arch. Ophthalmol.* 108, 2 (1990), 286–290.
- [106] Joram J. van Rheede, Iain R. Wilson, Rose I. Qian, Susan M. Downes, Christopher Kennard, and Stephen L. Hicks. 2015. Improving mobility performance in low vision with a distance-based representation of the visual scene. *Investigat. Ophthalmol. Visual Sci.* 56, 8 (2015), 4802–4809.
- [107] Fernando Vargas-Martin, Eli Peli, et al. 2002. Augmented-view for restricted visual field: Multiple device implementations. *Optom. Vision Sci.* 79, 11 (2002), 715–723.
- [108] Enhanced Vision. 2015. Jordy. Retrieved from <https://www.enhancedvision.com/low-vision-product-line/jordy.html>.
- [109] VisionAware. [n.d.]. Organizing & Modifying Your Home. Retrieved from <https://www.visionaware.org/info/everyday-living/home-modification-12>.
- [110] P. Weckerle, S. Trauzettel-Klosinski, G. Kamin, and E. Zrenner. 2000. Task performance with the low vision enhancement system (LVES). *Visual Impair. Res.* 2, 3 (2000), 155–162.
- [111] R. Wessels, B. Dijcks, M. Soede, G. J. Gelderblom, and L. De Witte. 2003. Non-use of provided assistive technology devices, a literature overview. *Technol. Disabil.* 15, 4 (2003), 231–238.
- [112] World Health Organization (WHO). 2018. Blindness and Vision Impairment. Retrieved from <https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment>.
- [113] James S. Wolffsohn and Rachael C. Peterson. 2003. A review of current knowledge on electronic vision enhancement systems for the visually impaired. *Ophthalm. Physiol. Optics* 23, 1 (2003), 35–42.
- [114] Molly Wood. [n.d.]. Video Feature: Signs That Virtual Reality Is on the Verge of Taking Off. Retrieved from http://www.nytimes.com/2015/01/29/technology/personaltech/video-feature-signs-that-virtual-reality-is-on-the-verge-of-taking-off.html?_r=1.
- [115] Wai Yu, Katri Kangas, and Stephen Brewster. 2003. Web-based haptic applications for blind people to create virtual graphs. In *Proceedings of the 11th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, (HAPTICS'03)*. IEEE, 318–325.
- [116] Yuhang Zhao, Edward Cutrell, Christian Holz, Meredith Ringel Morris, Eyal Ofek, and Andrew D. Wilson. 2019. SeeingVR: A set of tools to make virtual reality more accessible to people with low vision. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. ACM, 111.
- [117] Yuhang Zhao, Michele Hu, Shafeeka Hashash, and Shiri Azenkot. 2017. Understanding low vision people's visual perception on commercial augmented reality glasses. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. ACM, 4170–4181.
- [118] Yuhang Zhao, Elizabeth Kupferstein, Doron Tal, and Shiri Azenkot. 2018. It looks beautiful but scary: How low vision people navigate stairs and other surface level changes. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, 307–320.
- [119] Yuhang Zhao, Sarit Szpiro, and Shiri Azenkot. 2015. Foresee: A customizable head-mounted vision enhancement system for people with low vision. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, 239–249.
- [120] 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. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 73–84.
- [121] Feng Zhou, Henry Been-Lirn Duh, and Mark Billinghurst. 2008. Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR. In *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*. IEEE Computer Society, 193–202.
- [122] Annuska Zolyomi, Anushree Shukla, and Jaime Snyder. 2017. Technology-mediated sight: A case study of early adopters of a low vision assistive technology. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, 220–229.

Received May 2019; revised August 2019; accepted September 2019