Published in final edited form as:

Int Symp Mix Augment Real. 2017 October ; 2017: 288–292. doi:10.1109/ISMAR-Adjunct.2017.89.

AR4VI: AR as an Accessibility Tool for People with Visual Impairments

James M. Coughlan^{*} and Joshua Miele[†]

The Smith-Kettlewell Eye Research Institute, 2318 Fillmore St., San Francisco, CA 94115 USA

Abstract

Although AR technology has been largely dominated by visual media, a number of AR tools using both visual and auditory feedback have been developed specifically to assist people with low vision or blindness – an application domain that we term Augmented Reality for Visual Impairment (AR4VI). We describe two AR4VI tools developed at Smith-Kettlewell, as well as a number of pre-existing examples. We emphasize that AR4VI is a powerful tool with the potential to remove or significantly reduce a range of accessibility barriers. Rather than being restricted to use by people with visual impairments, AR4VI is a compelling universal design approach offering benefits for mainstream applications as well.

Keywords

Haptics; tactile graphics; accessibility; visual impairment; low vision; blindness

Index Terms

H.5.1 [Multimedia Information Systems]; Artificial; augmented and virtual realities—; H.5.2 [User Interfaces]; Interaction Style—

1 Introduction

The World Health Organization (WHO) estimates that roughly 285 million people are visually impaired worldwide, of whom 39 million are blind [1]. In the US alone, these numbers correspond to 10 million people who are visually impaired, of whom 1.3 million are blind [2]. These numbers are escalating rapidly as the population ages, and imply a growing barrier to a wide range of daily activities that require access to visual information. Such activities include navigation and travel and interaction with a wide variety of media and devices that contain important visual information, such as printed documents and graphics, maps (flat or relief), 3D models used in STEM education, and appliances with digital displays (e.g., microwave ovens and thermostats). Despite the existence of an increasing number of access technologies, many daily activities are partially or completely inaccessible to people with visual disabilities [3].

^{*} coughlan@ski.org. [†] jam@ski.org.

While most AR applications target users with normal vision, AR holds the promise of making a wide range of information available to people with visual impairments by adding accessible annotations or enhancements to physical objects or the environment. Accessible AR enhancement can be provided in a number of ways, including speech or audio cues, haptic or tactile feedback, or even image enhancement for people with some usable vision.

Incorporation of gesture recognition into the AR approach enables an entirely new *audio-haptic interaction* paradigm. A gestural interface allows the user to explore nearby objects or the environment in a natural way by touching or pointing, or otherwise indicating a location or direction of interest. The reality augmentation comes in the form of additional information or annotation (whether audio, tactile, or visual enhancement) related to the location or direction indicated by the gesture. We coin the term “AR4VI” to describe this type of assistive AR application for people with visual impairments.

In this paper we provide an overview of existing AR4VI applications for people with visual impairments. We then describe two specific applications we have conceived, drawing on our extensive experience developing accessibility technology (and on the personal experience of one author, who is blind): overTHERE, for providing directional information to nearby points of interest in the neighborhood, and CamIO, for providing information about locations of interest on objects such as 3D models, appliances and relief maps. Finally we conclude with a discussion of the implications of AR4VI for universal design.

2 Related Work

We divide AR4VI applications into two broad categories: global applications, which augment the physical world in the user’s vicinity, and local applications, which augment physical objects that the user can touch and explore haptically. We begin by describing image enhancement applications, which cut across both categories because they can be applied both globally and locally.

2.1 Image enhancement applications

A variety of image enhancement AR4VI applications have been researched and developed for people with low vision. Research on image enhancement has explored several types of image transformations, including contrast enhancement [4], color transformations [5] and the enhancement of intensity edges [6]. Smartphone/tablet apps such as Brighter and Bigger¹ and Super Vision² perform these types of image enhancements, in addition to providing features such as magnification and image stabilization.

Wearable (electronic glasses) devices such as eSight³ and OxSight⁴ use a form factor that delivers image enhancement directly to the eyes rather than requiring the user to manipulate and view an external screen. OxSight goes beyond basic image enhancement to provide higher-level functionality [7], such as rendering the scene in cartoon-like layers to highlight

¹<http://asada.tukusi.ne.jp/brighterandbigger/e/>

²<http://www.schepens.harvard.edu/superVision>

³<https://www.esighteyewear.com/>

⁴<http://smartspecs.co/about-us/>

important foreground features (e.g., a person nearby) relative to background features. OrCam⁵ allows the user to point their finger to a location of interest, such as a text region on a document, and have the text in that region read aloud.

The next two subsections focus on the global and local applications of AR4VI that are targeted primarily to blind users, and which typically deliver feedback in audio or other non-visual form.

2.2 Global applications

Global applications of AR4VI provide directional information about nearby points of interest (POIs) in the environment to facilitate wayfinding and navigation. Many such navigation systems are based on GPS to obtain localization, such as Loomis's Personal Guidance System [8] and Nearby Explorer⁶ from American Printing House for the Blind, and are now standard features on smartphones [9] that include turn-by-turn directions.

Alternative localization systems have also been developed to overcome the limitations of GPS, which is only available outdoors and has a spatial resolution of approximately 10 meters in urban environments [10]. One such alternative is Remote Audible Infrared Signage (RIAS), an example of which is Talking Signs [11] – a system of infrared transmitters and receivers that Smith-Kettlewell originally developed to make signage accessible for blind pedestrians. Other popular localization approaches rely on Wi-Fi triangulation [12], Bluetooth beacons [13] [14] or computer vision, such as the SLAM (simultaneous location and mapping)-based Tango Tablet [15].

Each localization approach has unique advantages and disadvantages in terms of localization accuracy and range, cost and the type of physical or digital infrastructure that is required, if any. At present no single alternative to GPS has emerged as a dominant winner, but the goal of making navigation tools for travelers of all abilities function in GPS-denied indoor environments is an area of active research, with entire conferences dedicated to this subject [16], and increasing presence in the marketplace [17] [18].

2.3 Local applications

The other main category of AR4VI is for local applications, which provide audio-haptic access to tactile documents and maps and 3D models. While the simplest and most common way to enhance the accessibility of an object or document is to apply braille labels to its surface, this approach has significant drawbacks, including the limited space available for them and their inaccessibility to those who don't read braille. These drawbacks motivate the need for more powerful AR4VI approaches.

An increasingly popular alternative to braille labels is the use of audio labels, which enable audio-haptic exploration of the object, whereby the user explores the surface naturally with the fingers and queries locations of interest; note that blind people typically prefer to explore objects with multiple fingers or hands instead of just one [19], which is compatible with the

⁵<http://www.orcam.com/>

⁶<http://tech.aph.org/ne/>

use of audio labels. Audio labels can be associated with specific locations on tactile models and graphics in a variety of ways, including by placing touch-sensitive sensors at locations of interest, overlaying a tactile graphic on a touch-sensitive tablet, or using a camera-enabled “smart pen” that works with special paper overlaid on the surface of the object [20]. While effective, these methods require each object to be customized with special hardware and materials, which is costly and greatly limits their adoption.

Computer vision is an AR4VI approach to creating audio labels, in which a camera tracks the user’s hands as they explore an object and triggers audio feedback when a finger points to a “hotspot” of interest. It has the key advantage of offering audio-haptic accessibility for existing objects with minimal or no customization. One of the first such computer vision approaches was KnowWare [21], which focused on the domain of geographical maps. Other related projects include Access Lens [22] for facilitating OCR, the Tactile Graphics Helper [23] to provide access to educational tactile graphics, augmented reality tactile maps [24] and VizLens [25] for making appliances accessible; an iOS app called Touch & Hear Assistive Teaching System (“THAT’S”)⁷ provides access to tactile graphics. We describe our “CamIO” project in Sec. 4, which uses a conventional camera to make 2D and 3D objects and graphics accessible.

3 overTHERE

overTHERE⁸ is a free iPhone app [26] (see Fig. 1) that gives blind pedestrians quick and easy access to accurate location information about businesses and other points of interest in the world [27]. The user simply points their phone in a direction to hear what’s “over there,” and audio cues make it easy to quickly and accurately pinpoint exact locations of businesses and addresses. These cues are generated by estimating the user’s location and pointing (heading) direction using the smartphone GPS, magnetometer and other sensors, and thereby determining the directions and distances to nearby points of interest (POIs) listed on the Google Places API.

The overTHERE app grew out of Smith-Kettlewell’s Virtual Talking Signs project [28]. The app uses an interface based on Talking Signs, which in its original form allowed the location of a transmitter to be determined by moving the receiver back and forth across a transmitter beam until the strongest signal is heard. overTHERE uses a similar interface, offering many of the advantages of Talking Signs, but now the transmitters are Google map pins and the receiver is the iPhone. It also allows the user to add custom virtual signs to any location of interest.

overTHERE is a free, powerful demonstration of this compelling virtual Talking Signs interface for use in providing accessible environmental information. overTHERE does not attempt to compete with the scope and power of other accessible GPS apps with their rich feature sets and route navigation. Rather, overTHERE is a simple app that performs a single job – telling the user quickly and intuitively what is “over there.”

⁷<http://thats.wiki/en/home/>

⁸<http://www.ski.org/project/overthere>

The app has been downloaded thousands of times since its release in 2016. While the interface has the potential to support indoor as well as outdoor navigation, limitations in the accuracy of GPS, magnetometer and other sensors make indoor use impractical at this time. However, indoor wayfinding systems with greatly improved localization and heading accuracy, such as Tango, could be used with the overTHERE approach in the future.

4 CamIO

CamIO⁹ (short for “Camera Input-Output”) is a computer vision-based system to make physical objects (such as documents, maps, devices and 3D models) accessible to blind and visually impaired persons, by providing real-time audio feedback in response to the location on an object that the user is touching [29]. The project was conceived by one of the co-authors (who is blind) and has been iteratively improved based on ongoing user feedback with a user-centered research and development approach. It builds on our earlier work [30], but now uses a conventional camera (like Magic Touch [31], which specializes to 3D-printed objects) instead of relying on a depth camera.

The CamIO system (Fig. 2) consists of two main components: (a) the “CamIO Explorer” system for end users who are blind or visually impaired, and (b) the “CamIO Creator” annotation tool that allows someone (typically with normal vision) to define the *hotspots* (locations of interest on the object) and the text labels associated with them. Both Explorer and Creator will be released as open source software.

CamIO Explorer uses computer vision to track the location of the user’s *pointer finger* (usually the index finger), and whenever the fingertip nears a hotspot, the text label associated with the hotspot is announced using text-to-speech. Note that the user can explore the object naturally with one or both hands, but receives audio feedback on the location indicated by the pointing finger. Computer vision processing is simplified using a fiducial marker board and a multi-color fingertip marker. The current prototype version of Explorer runs in Python on a Windows laptop, using the laptop’s built-in camera. We also have a preliminary version of Explorer running in Android, and plan to add the iOS platform in the future.

CamIO Creator allows a sighted assistant to define digital annotations, which specify the hotspot locations and associated text labels for an object of interest. A sighted annotator indicates hotspot locations by clicking on the appropriate location in two different images of the object taken from different vantage points and typing in the corresponding text labels. This process allows the annotations to be authored remotely, i.e., even when the annotator lacks physical access to the object. Creator also runs in Python on a Windows laptop, and we plan to port this to a browser-based platform, which will simplify the annotation process by eliminating the need to download and install the tool. It also opens the possibility of crowdsourcing the annotation process, so that multiple people can contribute annotations for a single object.

⁹<http://www.ski.org/project/camio>

We have conducted multiple usability studies on CamIO, including a focus group with five blind adults and a formative study with a sighted person using CamIO Creator. Feedback from the studies was very positive, and has already given us the opportunity to identify and correct bugs and add new UI features. Overall, the user studies confirmed our belief that the CamIO project is likely to be of great interest to many people who are blind and that it has real-world potential. The studies suggest many avenues for future development, including UI improvements and new functionalities such as making authoring functions accessible.

We are also exploring possible applications of the CamIO approach for users who have low vision, who may prefer a visual interface that displays a close-up view of the current hotspot region, instead of (or in addition to) an audio interface. Note that this visual interface could act as a magnifier that is steered by the fingertip; it could also include the option of digitally rendering this view *without* the fingers obscuring it – effectively rendering the controlling finger invisible, providing an elegant interface with no interference from the finger.

5 Discussion

AR4VI is an example of an application domain of AR that expands the accessibility of the technology beyond the usual target of people with normal vision, and thereby fulfills many of the principles of universal design [32]. Universal design is a design philosophy that aims to create spaces, tools, and systems that minimize inherent barriers for people with disabilities, and thereby improve usability for a broader range of users with and without disabilities. Universal Design for Learning (UDL) [33] applies this philosophy to the design of educational technologies and techniques, based on the principle that multiple means of presentation, expression, and engagement support learning by students with different learning styles, cultural backgrounds, and physical/cognitive abilities.

AR4VI tools that support visual impairment use primarily text-based information to augment the local or global environment. This inherently supports the multiple modes required by UDL by flexibly allowing the transformation and re-representation of the annotations into a variety of presentation modes (e.g., language translation, text-to-speech, magnification/enlargement/visual enhancements, refreshable braille, etc.). This makes AR4VI highly consonant with both universal design principles and UDL.

Indeed, we argue that AR4VI is not only consonant with universal design, but its power and cost-effectiveness actively facilitate it. A good illustration of this principle is the motivation for the overTHERE and CamIO projects, both of which may be viewed as the application of AR to make previous-generation technologies more cost-effective, convenient and practical. overTHERE is the modern successor to the original Talking Signs system, which required the installation of one infrared beacon for each POI (requiring ongoing maintenance to replace batteries and aging components) and a dedicated hardware receiver for each user. Similarly, CamIO replaces customization of objects with braille labels, built-in touch sensors and other physical modifications with digital annotations.

In both cases the expense of the physical infrastructure required for the previous-generation technology hindered its adoption; by virtualizing the infrastructure, and separating the

physical object from its annotations, the AR4VI approach can provide more information, with much less expense, and deliver it in a richer variety of presentation modes to accommodate a diversity of user needs and preferences. The power of the AR4VI approach will only grow as the underlying technology matures, such as the availability of more accurate indoor/outdoor localization for overTHERE enabled by systems like Tango, and the advent of markerless AR (e.g., supported by Apple's ARKit¹⁰) for CamIO.

6 Conclusion and Future Work

Gaining quality access to information is one of the most significant issues facing blind and visually-impaired people, with access to spatial information being particularly challenging. AR4VI is a general approach to tightly coupling labeling and annotation information with spatial location, improving access to information for people with visual disabilities in both local and global contexts (e.g., CamIO's accessible annotation of features on a map, model, or document, or overTHERE'S interactive augmentation of the nearby environment). These technologies and others are part of a growing trend in accessibility that, rather than requiring specialized materials or modifications to the built environment, takes advantage of accelerating improvements in processing power, connectivity, and cloud-based resources, making accessible spatial information available more quickly, in more contexts, and at lower cost than ever before.

With the increasing availability and reliability of such free geographical resources as Open Street Map and Google Places, as well as the broad adoption of sensor-rich and highly connected smartphones, our ability to overlay accessible and relevant information on the environment is reaching more blind and visually impaired users than ever before. The gestural model of AR4VI demonstrated by overTHERE is already being adopted by other accessible GPS apps (e.g., Nearby Explorer), and Microsoft's Cities Unlocked Project¹¹. As mainstream AR technologies continue to improve in quality and affordability, we anticipate continued expansion of AR4VI in the accessible wayfinding world.

Similarly, the trends and tools of AR4VI that support access to maps, models, objects, and other materials points to significant expansion of this approach in the coming years. The quality of webcams and graphics processing power are improving simultaneously, paving the way for faster and more accurate hand tracking for interactive object-based AR4VI. At the same time, there is a significant growth in the availability of human- and computer-based resources for generating descriptive annotations. For example, remote sighted assistance services such as Be My Eyes¹² and Aira¹³ could be used to remotely annotate maps and models to enable haptic AR4VI interaction. There are also a growing number of free and low-cost APIs offering such services as optical character recognition (OCR), object recognition, and image description. Such resources could be incorporated into a framework capable of automatically annotating certain categories of documents and objects, further expanding the availability and immediacy of AR4VI interaction.

¹⁰<https://developer.apple.com/arkit/>

¹¹<http://www.guidedogs.org.uk/cities-unlocked>

¹²<http://bemyeyes.com/>

¹³<https://aira.io/>

It is common for accessibility technologies to find their way into mainstream usage. Examples of this kind of universal-design cross-fertilization include the mainstream adoption of text-to-speech (speech synthesis), speech-to-text (dictation), and OCR. With the proliferation of AR technologies and techniques, it is inevitable that some of the approaches to non-visual presentation of spatial information pioneered in AR4VI will find applications in the mainstream. While it may not be possible to anticipate exactly where these contributions may be made, one imagines that auditory directional information in the environment may be useful in the context of driving and operation of heavy machinery, providing auditory feedback to avoid overloading visual displays; supporting travel and tourism by providing audio directional information for guidance and to draw attention to points of interest in the environment; or in retail to guide or draw consumers' attention to merchandise out of immediate view. Object-based AR4VI approaches such as demonstrated by CamIO may assist drivers and pilots in finding needed controls without using vision, and will almost certainly impact mainstream educational UDL technologies by providing multiple modes of interaction with educational models and manipulables.

Ultimately, AR4VI is a rapidly expanding field with strong potential not only in accessibility for the blind and visually impaired, but also for mainstream commercial applications and a wide array of interdisciplinary HCI-related research.

Acknowledgments

This work was supported by NIH grant 5R01EY025332 and NIDILRR grant 90RE5024-01-00.

References

1. Pascolini D, Mariotti SP. Global estimates of visual impairment: 2010. *Br. J. Ophthalmol.* Dec.2011
2. [Accessed: 02-Jul-2017] Fact Sheet Blindness and Low Vision | National Federation of the Blind. [Online]. Available: <https://nfb.org/fact-sheet-blindness-and-low-vision>
3. Geruschat, D., Dagnelie, D. *Assistive Technology for Blindness and Low Vision*. CRC Press; 2012. *Low Vision: Types of Vision Loss and Common Effects on Activities of Daily Life*.
4. Peli E, Woods RL. Image enhancement for impaired vision: the challenge of evaluation. *Int. J. Artif. Intell. Tools.* Jun.2009 18(03):415–438. [PubMed: 20161188]
5. Jefferson, L., Harvey, R. *Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility*. New York, NY, USA: 2006. *Accommodating Color Blind Computer Users*; p. 40-47.
6. Satgunam P, et al. Effects of contour enhancement on low-vision preference and visual search. *Optom. Vis. Sci. Off. Publ. Am. Acad. Optom.* Sep.2012 89(9):E1364–E1373.
7. [Accessed: 30-Jun-2017] OxSight uses augmented reality to aid the visually impaired | TechCrunch. [Online]. Available: <https://techcrunch.com/2017/02/16/oxsight-uses-augmented-reality-to-aide-the-visually-impaired/>
8. Loomis, J., Golledge, R., Klatzky, R. *Personal guidance system for the visually impaired using GPS, GIS, and VR technologies*; *Proc. Conf. on Virtual Reality and Persons with Disabilities*; 1993. p. 17-18.
9. [Accessed: 02-Jul-2017] Built-in Smartphone Mapping Apps from Google and Apple - American Foundation for the Blind. [Online]. Available: <http://www.afb.org/info/living-with-vision-loss/using-technology/smartphone-gps-navigation-for-people-with-visual-impairments/built-in-smartphone-mapping-apps-from-google-and-apple/1235>
10. Brabyn J, Alden A, Haegerstrom-Portnoy G, Schneck M. GPS performance for blind navigation in urban pedestrian settings. *Proc Vis.* 2002; 2002

11. Crandall W, Brabyn J, Bentzen BL, Myers L. Remote infrared signage evaluation for transit stations and intersections. *J Rehabil. Res. Dev.* 1999; 36(4):341–355. [PubMed: 10678457]
12. Gallagher, T., Wise, E., Li, B., Dempster, AG., Rizos, C., Ramsey-Stewart, E. Indoor positioning system based on sensor fusion for the Blind and Visually Impaired; 2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN); 2012. p. 1-9.
13. Ahmetovic, D., Gleason, C., Ruan, C., Kitani, K., Takagi, H., Asakawa, C. NavCog: a navigational cognitive assistant for the blind; Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services; 2016. p. 90-99.
14. Ruffa AJ, Stevens A, Woodward N, Zonfrelli T. Assessing iBeacons as an Assistive Tool for Blind People in Denmark. *Worcester Polytech. Inst. Interact. Qualif. Proj. E-Proj.-050115-131140*. 2015
15. [Accessed: 10-Jul-2017] Navigation for the Visually Impaired Using a Google Tango RGB-D Tablet – Dan Andersen. [Online]. Available: <http://www.dan.andersen.name/navigation-for-the-visually-impaired-using-a-google-tango-rgb-d-tablet/>
16. [Accessed: 03-Jun-2017] International Conference on Indoor Positioning and Indoor Navigation. [Online]. Available: <http://ipin-conference.org/>
17. [Accessed: 03-Jun-2017] indoo.rs - Professional indoor positioning service. [Online]. Available: <https://indoo.rs/>
18. [Accessed: 03-Jun-2017] Indoor Maps – About – Google Maps. [Online]. Available: <https://www.google.com/maps/about/partners/indoormaps/>
19. Morash VS, Pensky AEC, Miele JA. Effects of using multiple hands and fingers on haptic performance. *Perception*. 2013; 42(7):759–777. [PubMed: 24344552]
20. Miele, J. Talking Tactile Apps for the Pulse Pen: STEM Binder. presented at the 25th Annual International Technology & Persons with Disabilities Conference (CSUN); San Diego, CA. Mar-2010;
21. Krueger MW, Gilden D. KnowWare [TM]: Virtual Reality Maps for Blind People. *Stud. Health Technol. Inform.* 1999:191–197.
22. Kane, SK., Frey, B., Wobbrock, JO. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York, NY, USA: 2013. Access Lens: A Gesture-based Screen Reader for Real-world Documents; p. 347-350.
23. Fusco, G., Morash, VS. Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility. New York, NY, USA: 2015. The Tactile Graphics Helper: Providing Audio Clarification for Tactile Graphics Using Machine Vision; p. 97-106.
24. Ichikari R, Yanagimachi T, Kurata T. Augmented Reality Tactile Map with Hand Gesture Recognition. *Computers Helping People with Special Needs*. 2016:123–130.
25. Guo, A., et al. Proceedings of the 29th Annual Symposium on User Interface Software and Technology. New York, NY, USA: 2016. VizLens: A Robust and Interactive Screen Reader for Interfaces in the Real World; p. 651-664.
26. [Accessed: 30-Jun-2017] overTHERE on the App Store. App Store. [Online]. Available: <https://itunes.apple.com/us/app/overthere/id1126056833?mt=8>
27. [Accessed: 30-Jun-2017] CSUN Conference 2017 Session Details. [Online]. Available: <http://www.csun.edu/cod/conference/2017/sessions/index.php/public/presentations/view/73>
28. Miele J, Lawrence MM, Crandall W. SK Smartphone Based Virtual Audible Signage. *J Technol. Pers. Disabil.* 2013:201.
29. [Accessed: 30-Jun-2017] CSUN Conference 2017 Session Details. [Online]. Available: <http://www.csun.edu/cod/conference/2017/sessions/index.php/public/presentations/view/333>
30. Shen, H., Edwards, O., Miele, J., Coughlan, JM. Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility. New York, NY, USA: 2013. CamIO: A 3D Computer Vision System Enabling Audio/Haptic Interaction with Physical Objects by Blind Users; p. 41:1-41:2.
31. Shi, L., McLachlan, R., Zhao, Y., Azenkot, S. Magic Touch: Interacting with 3D Printed Graphics; Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility; 2016. p. 329-330.

32. Story, MF., Mueller, JL., Mace, RL. The Universal Design File: Designing for People of All Ages and Abilities. Revised Edition. Center for Universal Design, NC State University, Box 8613; Raleigh, NC: 1998. 27695-8613 (\$24)
33. Rose D. Universal Design for Learning. J Spec. Educ. Technol. Sep.2000 15(4):47–51.

Figure 1.
(a) overTHERE app icon. (b) App screenshot showing list of nearby signs.

Figure 2.
CamIO system shows laptop, object (relief map), fiducial marker board and user pointing to a hotspot.