

Mobile Accessibility: Camera Focalization for Blind and Low-Vision Users On the Go

Chandrika Jayant
Computer Science and Engineering
University of Washington
Seattle, Washington
cjayant@cs.washington.edu

ABSTRACT

There is a significant need for access to low-cost assistive mobile services through standard devices. Many special purpose devices that provide services are specialized for few functions, are usually not networked, and are expensive and difficult to sustain. Mobile devices and their sensors, along with audio or tactile output, show immense promise for improved daily accessibility to the world for blind people on the go. Particularly, the camera has the potential to expand environmental knowledge using context and computer vision. This research will investigate general camera interaction techniques, mobile device and camera accessibility for blind users, and computer vision techniques for recognition given mobile phone limitations. These studies along with formative studies and focus groups with users will inform the design, implementation, and evaluation of future lab and field studies for semi-autonomous focalization with the camera for blind users.

Categories and Subject Descriptors

K.4.2 [Social Issues]: Assistive technologies for persons with disabilities; H.5.2 [Information Interfaces and Presentation]: User Interfaces; I.2.10 [Vision and Scene Understanding]: Video Analysis; I.4.8 [Scene Analysis]: Object Recognition

General Terms

Design, Experimentation, Human Factors

Keywords

Mobile, camera, blind, low-vision, web services, focalization.

1. INTRODUCTION

Mobile devices present people with disabilities with new opportunities to act independently in the world. The World Health Organization estimates that there are more than 37 million blind and 124 million low-vision people in the world [2006]. For most of these people, common everyday activities, such as transportation, shopping, eating out, and simply taking a walk, can be major challenges. Most of these blind and low-vision people have not realized the full potential of the mobile revolution due to lack of sustainability and adoption of specialized devices which aid in some of these activities, partly due to the expense and proprietary nature of the companies which provide them, and the lack of accessible mainstream devices. The proposed platform leverages the sensors that modern cell phones have to keep devices cheap, and uses remote web services to process requests. Both human-powered and fully-automated web services are used to balance cost, accuracy, and timeliness of the services available. Being able to effectively use the camera to send essential data to these services is a huge challenge for blind people, and motivates this particular research.

1.1 Background and Motivation

There is scant computer science research on blind and low-vision users' interaction with mobile devices and general computing practices, not to mention interaction with cameras. There is work on the design of auditory user interfaces for blind users but not in the context of activities on the go. Watanabe et al. [7] surveyed the present state of mobile phone usage by blind and low-vision people in Japan, finding that users felt that current applications with speech output were insufficient for supporting their needs. Through interviews and diary studies with 20 visually and motor impaired users, Kane, Jayant, et. al. [5] ran a formative study that examined how people with visual and motor impairments select, adapt, and use mobile devices on the go.

Very few mainstream phones are completely accessible off the shelf. The iPhone and Google Android-enabled phones now provide free screen readers. While this is promising, this only makes accessible certain functions of the phone. Many specialized portable devices fill a specific need of blind and low-vision users and are prohibitively expensive. The kNFBReaderMobile provides optical character recognition (OCR) for documents, but is very expensive and is particular to just one phone. Other examples are portable navigation devices, like Access Ingenuity's GPS Trekker, barcode readers such as the ID Mate OMNI talking scanner, and talking book readers like the Victor Stream.

There is barely any literature on blind people's interactions with the camera. There have been a few exciting camera prototypes developed recently (e.g. TouchSight and Haidary's Camera for the Blind) using audio notes and tactile screens, but these have not investigated the actual interaction while taking a photo. A currency reader has been made for a mainstream Windows Mobile phone which works very well as currency is very distinct in its pattern [6]. Motivating scenarios that should propel more research in this area include recognizing faces, creating art, taking well-framed photos for family members, reading signs on office doors, reading street signs, perusing a menu with a non-standard font, locating a bus stop number, matching a pair of socks, and differentiating the women's from the men's bathroom. These recognition techniques themselves are not trivial, but for them to work at all necessitates blind people taking the "right" pictures to send to services to reach their goals.

2. PROPOSED SOLUTION

MobileAccessibility will provide a mixed-initiative user interface for blind and low-vision people which will be integrated with remote automated and human-powered services (Figure 1). The user-centered design of the MobileAccessibility platform will involve blind and low-vision people throughout its development as both evaluators and researchers. One of the main HCI challenges is the development of intelligent, multi-modal user interfaces to the cell phone that choose how and when to interrupt and present information to users. The research for my thesis will start to tackle this challenge, in the realm of *camera focalization*.

I will work on gathering the appropriate camera data and feedback for the mixed-initiative mobile user interface. With contextual information from the phone and the user, this can be a semiautonomous procedure, as in the Blind Driver Challenge [4]. The goal is to provide environmental information without object modeling or tagging. I will design, implement, and evaluate interactive procedures for blind users to point the camera in the right direction, which I call *focalization*. Two main areas that need to be researched to reach this goal are interaction and vision.

Camera Interaction: Some issues to be investigated are the combined use of tonal, verbal, and tactile output modalities and what input to get from the users about their environment (e.g. when should this be task oriented or location oriented?) How does panning the scene with the camera with real-time feedback fare versus taking single photographs and getting more discrete feedback? How do we know we've taken the picture we needed?

How do blind users perform with feedback about absolute distances to move the camera? What about relative distances?

Computer Vision: Though mobile phones boast cameras with higher resolution these days (5 megapixels and up), there are still many hardware limitations with them compared to expensive high-end cameras. Issues such as lower resolution, motion blur, lighting, and perspective must be investigated. Also, OCR is currently only reliable with standard text documents. Finding text in a scene and then using OCR is an open problem which people have researched for awhile but have not found a sufficient solution for, so this must be carefully taken into account.

Eventually, we will conduct lab and field studies to verify usability for the user interfaces, preceded by research into experience sampling for blind and low-vision users..

3. RESEARCH GOALS AND STATUS

I have conducted a preliminary formative study about mobile technology adoption and accessibility with Kane et. al. [5]. We have already begun work on creating prototype applications on the Google G1 phone using the Android platform. So far these prototypes include an accessible barcode reader, a color recognizer, and a bus navigation system. We will also use existing OCR applications such as the kNFBreaderMobile and TextScout. These prototypes will span different sensors and different remote services, both automated and human. After building initial prototypes, we will conduct focus groups and iterate on designs. Next we will conduct lab studies studying the interaction of our work with the provider web interfaces studied by other members of our team to test out latency, cost, and quality of results. Examples of services we will test out include OCR (using human services and also automated services with human verification), and specific object recognition of already defined objects.

Richard Ladner, my advisor, has extensive research experience on access technology for blind people since the 1980s. Our team has a close relationship with the National Federation of the Blind, Jernigan Institute, Google, Seattle Lighthouse for the Blind, and the Accessibility Group at IBM Tokyo Research Lab. Jeffrey Bigham, a professor at U. Rochester who has worked on WebInsight and WebAnywhere [2,3], will be a consultant and researcher on the MobileAccessibility platform.

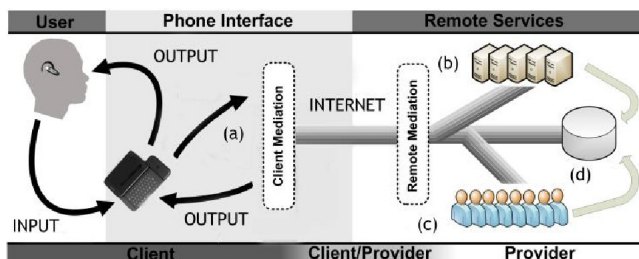


Figure 1. MobileAccessibility platform showing the user, the phone interface, and the remote services. (a) Sensor (camera) and user input from the phone via user interface (b) automated web services (c) human web services (d) Accessibility Commons (shared data repository).

4. ENVISIONED CONTRIBUTIONS TO THE ACCESSIBILITY FIELD

An accessible cell phone that can accomplish multiple tasks has the potential to provide blind and low-vision people with more independence than they have currently. Furthermore, the MobileAccessibility solution has the potential to be inexpensive and more sustainable than current accessibility solutions. Being able to harness the power of the camera and leverage the potential power of human and automated web services are two major components for making this framework succeed. Blind students will be recruited for the project giving them a chance to participate in work directly affecting them. Lessons learned from these studies can extend to users with different disabilities or situational impairments, and push towards more universal design.

5. REFERENCES

- [1] American Foundation for the Blind: Cell Phone Accessibility Overview. <http://afb.org>
- [2] Bigham, J, Kaminsky, R., Ladner, R., Danielsson, O., and Hempton, G. Webinsight: Making Web Images Accessible. ASSETS '06. Portland, OR.
- [3] Bigham, J., Prince, C., Ladner, R. Webanywhere: A Screenreader On-the-go. W4A '08.
- [4] Hong, D., Kimmel, S. et al. Development of a Semi-Autonomous Vehicle Operable by the Visually-Impaired. In Multisensor Fusion and Integration for Intelligent Systems. Spring Berlin Heidelberg, 2009.
- [5] Kane, S., Jayant, C., Wobbrock, J., and Ladner, R. 2009. Freedom to roam: A study of mobile device adoption and accessibility for people with visual and motor disabilities. ASSETS '09. Pittsburgh, PA.
- [6] Liu, Xu. A camera phone based currency reader for the visually impaired. ASSETS '08. Halifax, Nova Scotia.
- [7] Watanabe, T., Miyagi, M., Minatani, K., and Nagaoka, H. A survey on the use of mobile phones by visually impaired persons in Japan. Proc. ICCHP '08. Berlin: Springer-Verlag, 1081-1084.