Cell Phones: Wayfinding Assistive Technology for the Visually Impaired

Robert P. Murray
Computer Engineering Department
University of California, Santa Cruz
rmurray@ucsc.edu

Roberto Manduchi
Computer Engineering Department
University of California, Santa Cruz
manduchi@soe.ucsc.edu

Abstract

A primary challenge the visually impaired and blind population faces today is that of wayfinding – the challenge of navigating effectively through the surrounding environment to a destination, obtaining sufficient yet minimal information about the surroundings. Various methods for aide in this task have been proposed, the most notable being the white cane. Of these proposals, one suggests the use of a cell phone camera and computing power in compliment with unique color targets as beacons in the environment to assist in wayfinding. Thus far, implementations of this in software on a cell phone have been made, making it possible to find beacons in the environment with the cell phone and receive feedback to guide the cell phone user. Since this development, it has been realized that feedback methods are a primary concern and several options have been considered and researched. Feedback methods that have been examined and investigated relevant to this work include region based tonal feedback, proximity based tonal feedback, and speech feedback, with considerations given to differences between continuous and non-continuous tonal feedbacks.

1. Introduction

Wayfinding is a problem often neglected by those with sight. One who sees does not consider the thought involved in keeping from colliding with obstacles such as people, railings, steps, doors, or trash cans. This is one facet of wayfinding, and an important one at that. But, now consider that one can see the obstacles, but cannot determine where the destination lies. This is but another aspect of wayfinding, one often more neglected by those of sight, and one which shall be discussed primarily in this work.

It is evident that the invention of the white cane brought great change to the habits of the visually impaired and blind. The white cane most unarguably provides a solution, if not optimal at least sufficient, for navigation in the first sense described above. With the same line of reasoning then, should it not be possible as well to create a device to aide in the latter form of wayfinding, and should it not be reasonable as well to do so? It is believed by the author and contributors to this work that it is most reasonable and possible to do so. How then is this done, it may be asked? The answer: it has been done to lesser degrees thus far as will be described, and the focus herein is to discuss in primary the methods used to enable this process of wayfinding, its development, new innovations in the project as of recent, and what shall be done to continue progress on this work. In particular, this work focuses on the development of new feedback methods and their application to existing work in the area of wayfinding technology.

1.1 Previous and Related Work

The BlindCam project, as it has been termed by previous contributors, was started in part by Professor Roberto Manduchi (listed above). The majority of this work expands on and extends the work of Roberto Manduchi [1] and related contributions to wayfinding by use of cell phones. Manduchi and other associates developed a method for color target identification using a cell phone camera and image processing algorithms which has been found to be an effective wayfinding method for the visually impaired. Furthermore, though less relevant for this work, Manduchi experimented with barcode reading upon successful target identification as a means of storing and relaying useful information to the user about the environment, working in many ways like an advanced form of brail placard. Other work in this field to create wayfinding devices includes development and experimentation with infrared signage broadcasts received by hand-held receivers [2], GPS-based localization, RFID labeling, and indoor Wi-Fi based localization (based on signal strength) and database access [3]. Each of these has its merits, though none have proven to be an effective, low cost, low maintenance method of implementing wayfinding beaconing technology for wide spread use. However, the method investigated by Manduchi proves to be a reasonable alternative. For this reason, further advances in the research performed in the BlindCam project seem the best approach at this time.

The BlindCam method of wayfinding has three primary components. First, color targets (Fig. 1) act as beacons in the environment. These may or may not be accompanied by barcodes for additional information, though for the purposes of the current work, barcodes are not necessary. These color targets are unique in that they provide a highly irregular color configuration for the normal environment of use. While it is possible that such a color configuration may be present in an environment in which this wayfinding technology will be utilized, it is far less likely than most other primary color configurations [1]. The second, a cell phone with a built-in camera, a common hand held device found accompanying a majority of the general population to this day, acts as the hardware for beacon identification. A camera on the cell phone in question captures the environment in an image, which is then sent to the cell phone's main memory for image processing in software. Once processing completes, the cell phone emits an appropriate audible response, either a tone in the case where a barcode is not present, or some speech in the case where one is. Except for the software that executes the processing algorithm and executes a request for feedback, this is all done in hardware and software via an operating system on the cell phone and is not controlled by the research developers. However, these other two crucial points in the procedure, the last component of the wayfinding method, are controlled and devised by the researchers, and are implemented in software by use of a user application running on the phone's OS.

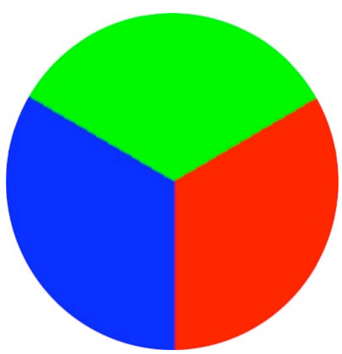

Fig. 1. Sample color target used by the BlindCam research as a beacon in the environment.

The image processing, explained in detail in Manduchi's paper [1], consists of scanning across the image in raster order, checking three points of interest, one for each color, and computing gradient thresholds by finding RGB differences between the three points. These gradient thresholds are used to create a cascade of filters which when applied to the image determines if a target appears in the image. Then, if a target is found, a series of tests to rule out false positives are executed and a final result is obtained. The second aspect of the software, the response to positive identification through feedback, is the primary interest for this work. In previous work, the BlindCam method determined region of the image that the target appeared, and responded with a tone of one out of three pitches, based on region. The region boundaries were determined as the boundaries of the thirds of the horizontal axis of the image (Fig. 2(a)). While this was sufficient for preliminary testing, it has become evident through use by human subjects with visual impairment that this method of feedback is sluggish, at times disorienting, and insufficient in relaying information about target location. In response, this work investigates other methods of feedback, both for use with and without barcodes.

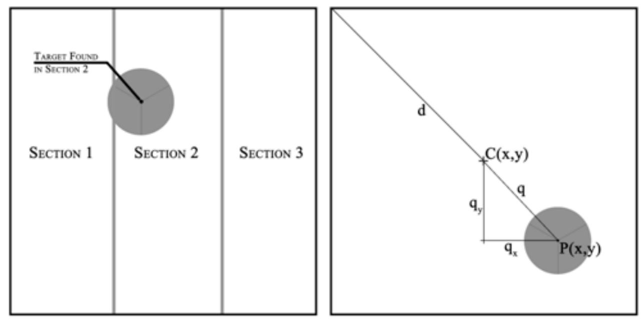

Fig. 2. (a) Left: Old feedback model in which image is divided into regions and feedback is based upon target's region. Used as a feedback method in previous work on BlindCam research. (b) Right: New feedback model dependent on Pythagorean distance of target vertex to center of image.

2. Feedback Methods

The primary response to the insufficiency of the previously utilized regioning method is to ask what is necessary of the feedback model implemented in this wayfinding procedure. What is ideal is vocal response, perhaps even two way conversational control of the device. Though vocal response is a possibility, conversation is highly unreasonable given current technology and the cell phone platform. As well, this is more readily implemented in combination with the barcodes previously discussed. Speech feedback will not be neglected or forgotten, but is not directly relevant here, and will be revisited later after other alternatives have been discussed. What is actually necessary is minimal information that indicates direction to and proximity of the center of the cell phone's camera lens to that of the target in the environment. Now, a simple switch in focus from the camera to the target will result in the location of the target in relation to the center of the image, or as previously stated, the center of the camera lens, being of relevance (Fig. 2(b)). This is the essence of the new feedback model proposed for BlindCam in this work.

Mathematically, this proximity method works as follows. Applying equations (1) and (2) to the vertices in question, the vertex of the target P, the vertex of the image center C, and the image origin (0,0), (usually the upper left of the image), gives the values d and q, two pixel distances relative to the image. The value d, the maximum distance the target vertex can be from the center and still be in the image (a case that never occurs[1]), and q, the actual distance of the target vertex from the image center, are used in equation (3) to arrive at the value r, the weighting factor of the pitch of the tone to be used as an audible feedback. Finally, the actual pitch is determined as f by equation (4) and a call to the OS is given with this pitch and the appropriate other parameters. Consequently, the pitch of the tone played is inversely proportional to the distance the target is from the center of the image at the time of capture.

$$d = [c_x^2 + c_y^2]^{1/2} \quad (1)$$

$$q = [|q_x - c_x|^2 + |q_y - c_y|^2]^{1/2} \quad (2)$$

$$r = 1 - q/d \quad (3)$$

$$f = 220 \text{ Hz} + r * 660 \text{ Hz} \quad (4)$$

It should be both noted and recognized at this time that this feedback model is not entirely meant to indicate what is in the environment, but is meant instead to be used as a tool in helping understand how the visually impaired navigate using such a technology. In the case of the barcode wayfinding model, having the target appear at the center of the image is almost wholly irrelevant, except in that the barcode is more likely to appear in the image in its entirety. For this reason, in that model this feedback is useless except in the instances where the barcode is only partially present in the image. Even in these instances, the combination of speech feedback from the barcode information and tonal feedback based on proximity could likely overwhelm the user of such a technology. As well, a primary goal of the BlindCam research is to provide assistive technology that gives sufficient yet minimal information about the environment to its user, meaning that the barcode based speech feedback system is far more important than proximity feedback in terms of an actual device and wayfinding method. Why then is this alternative important? It is fundamentally important to this research because understanding the ways in which these technologies are used is directly pertinent to the utility of a developed technology, as well as the way in which such technology should be designed and implemented.

This is clearly evident from the case of the regional tonal feedback model. In that model, the ways in which the visually impaired use the technology were not evaluated before implementation with the severity that was necessary. This was obvious when a test subject used the technology and continually lost the target either above or below the image. As well, the subject would overshoot when pivoting the device on its vertical axis due to the low resolution of the tonal spectrum being used. These two problems arose from a poorly designed interface for the user, and as a result, an attempt to understand how the user attempts to use the technology helps in creating a technology that can be used effectively and easily by the user. In this particular case, it was noted by the subject that the camera lens projects an imaginary vector in the subject's mind and that manipulation of this vector, turning the subject's hand in terms of pivots and pans, could be tracked mentally. Then, it was arrived at that if the audible feedback from the phone corresponded to this vector and its proximity to intersection with the target, the user could effortlessly determine which direction to move his or her hand. Hence, the model for proximity based tonal feedback was born, and pursuit of this model intends to arrive at a better understanding of how the visually impaired will use the cell phone wayfinding technology and will enhance the ease of use of versions which include barcode reading and speech feedback.

3. Implementation and Results

Using the Symbian Operating System for S60 based cell phones and the C++ programming language, existing BlindCam software was modified to incorporate the new feedback model developed above. This software runs on all Symbian OS 7.0s, S60 2.1 cell phones, but was used in particular on a Nokia 7610 (Fig. 3) for testing and to obtain results. The software runs in a loop until a request to terminate is given by user input. During this loop, an image is captured by the hardware, stored in memory, scanned for targets, and a response is generated. In the case of no target in the image, the response is nothing. In the case a target is found, a procedure is called to perform the pitch modulation algorithm discussed above, and then a request is given to the cell phone's operating system to play a tone with the given pitch and duration.

Fig. 3. Nokia 7610 cell phone used for testing.

[1] It is in fact not possible for the target vertex to be distance d from the center because at least one of the colors in the color target does not appear in the image, making it undetectable by the image-processing step.

A major problem faced in implementation not yet discussed is that of continuous tonal playback. During development, it was arrived at that the non-continuous feedback, a tone of a set duration followed by a pause followed by another tone and a pause, and o on, was not a useful feedback response. An attempt to convert to a continually playing tone that changed while playing in pitch or volume was made, but was unsuccessful given the hardware and software platforms. The delay between tones in the non-continuous version of the software was long enough that it caused the device to feel sluggish and inaccurate. By changing the feedback model to the proximity-based feedback, the delay was less a problem, but was still nonetheless evident and made use of the device cumbersome.

The change from the region based feedback model to the proximity-based feedback model was evident upon first use. The information received by the user based on a similar feedback method increased substantially by changing this one aspect of the device. The developers noticed first and foremost that direction of movement of the user's hand required to place the target at the center of the image is almost intuitive with this new feedback. If in fact a continual tone could be used in place of the non-continuous tone, the developer is certain that this aspect would be entirely intuitive after initial use and familiarity with the device. As of yet, no subjects with visual impairment have tested this revision of the devise, but it is planned that such subjects will begin testing in the near future to obtain more results and feedback on other improvements to the device.

4. Conclusion

The BlindCam research is far from complete, though with this recent work, some insight about the current work has been obtained and new avenues have been opened for future development in this research. The primary result obtained is that of the importance of understanding how individuals with visual impairment use assistive technology devices as extensions of themselves, not as precarious additions. Other secondary results include a new feedback model and implementation thereof with increased information retrieval by the user at little or no increase in effort of use, as well as ambition to find methods of implementing continuous tonal feedback models. Consequently, future work in the BlindCam project hopes to answer this implementation issue as well as explore possibilities in speech feedback and incorporation of these results into the final model and the barcode based feedback mechanism already in use.

In response to speech feedback, several options in text to speech (TTS) software as well as prerecorded speech playback are possible. TTS software has the two fundamental problems of large monetary cost for existing software use and unexpressive tone resultant from simplified speech engines, ones that are capable of running on cell phone platforms. TTS is a possibility that has been looked into, but the cost, even for academic research use, is large enough that use at this time is not practical. Prerecorded speech alternatively is a possible speech feedback mechanism, and in fact has small scales implementation in the BlindCam project already. The barcodes contain IDs that map into a small database of prerecorded speech segments. This database could be enlarged on the phone, or could be maintained on a massive scale on the Internet and the phone could perform Internet lookup and retrieval. This has a great potential and few limits in terms of scalability, but unfortunately has the drawback of high overhead to maintain the database and populate it initially. It is likely that the best speech mechanism for feedback is prerecorded speech segments and as such it will hopefully be investigated as a primary feedback mechanism for use in the future.

Integration of the current results into the previous work on barcode use in wayfinding is important and should not be neglected. Possible avenues of integration include a tonal variation in voice from a speech response based on locality of the target, tonal overlay in addition to speech feedback, such as a tone before the speech response or interval tones with speech upon successful reading of a barcode, or tonal response in absence of a correctly identified barcode to help the user reposition in order to find the target. All of these possibilities will be considered, and hopefully more will arise as work continues. Ultimately, the integration of the discoveries and insights brought on by this work into the previous work and overall ambition of the BlindCam research makes this work an important milestone in the long process of developing cell phone based assistive technology that will help the visually impaired and blind in their continual battle of wayfinding in the chaotic environment of every day life.

References

1. J. Coughlan, R. Manduchi, H. Shen, "Cell Phone-based Wayfinding for the Visually Impaired", 1st International Workshop on Mobile Vision, Graz, Austria, May 2006.
2. W. Crandall, B. Bentzen, L. Myers and J. Brabyn. New orientation and accessibility option for persons with visual impairment: transportation applications for remote infrared audible signage. Clinical and Experimental Optometry 2001 May, 84(3): 120-131.
3. A.M. Ladd, K.E. Bekris, A.P. Rudys, D.S. Wallach and L.E. Kavraki. 'On the feasibility of using wireless ethernet for indoor localization", IEEE Trans. On Robotics and Automation, 20(3):555-9, June 2004.