Blind Guidance Using Mobile Computer Vision: A Usability Study

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

Sri Kurniawan
Dept. of Computer Engineering
University of California, Santa Cruz
srikur@soe.ucsc.edu

Homayoun Bagherinia
Dept. of Electrical Engineering
University of California, Santa Cruz
hbagheri@soe.ucsc.edu

ABSTRACT
We present a study focusing on the usability of a wayfinding and localization system for persons with visual impairment. This system uses special color markers, placed at key locations in the environment, that can be detected by a regular camera phone. Three blind participants tested the system in various indoor locations and under different system settings. Quantitative performance results are reported.

Categories and Subject Descriptors
I.5.4 [Pattern Recognition]: Applications: Computer Vision.

General Terms
Algorithms, Performance, Experimentation, Human Factors.

Keywords
Wayfinding, Mobility, Orientation, Guidance, Recognition.

1. INTRODUCTION
Persons who are blind experience difficulties navigating efficiently and purposefully in an unfamiliar environment. On top of routine *mobility* tasks (avoiding obstacles), a blind person faces various challenges: assessing one's location in the environment (*positioning*), determining how to reach a destination (*wayfinding*), finding things such as a water fountain, the elevator switch, or an exit door (*localization*), and reading textual or graphical content such as exposed signs (*information access*). Coughlan and Manduchi [1][2] proposed the use of special markers to tag key locations in the environment. These markers are small and inexpensive, and are designed so as to be quickly detected by the camera in a regular cell phone. A small amount of information is encoded in the marker itself. A blind person can use his or her cell phone to find existing markers in the environment and decode their content. This requires that the user "scan" the environment by pointing the cell phones in different directions. Once the marker is in the camera's field of view, the user is warned, for example, by a beeping sound. This provides the user with an indication of the approximate bearing direction to the marker. In addition, the approximate distance to the marker can also be communicated to the user.

ASSETS'10, October 25–27, 2010, Orlando, Florida, USA.

This contribution is concerned with the usability of this marker-based system. Specifically, we study the mechanisms involved with *searching* for a marker and with *guidance* to a marker's location once the marker has been detected. We report qualitative and quantitative results from user studies with three blind subjects, testing the marker detection system in different outdoors environments. Preliminary tests [2] were mostly qualitative in nature, focused only on the *search* component of the process

2. SYSTEM DESCRIPTION
The markers used in our experiments are pie-shaped, with a diameter of about 16 cm. They are divided in 4 angular sectors, each colored with a different distinctive color. The detection algorithm is described in [1][2]. Using a Nokia N95 8GB cell phone acquiring images with resolution of 480 (vertical) by 640 (horizontal) pixels, detection is performed at about 7 frames per second (fps). In our indoor experiments, the algorithm gave virtually no false alarms. A marker viewed frontally is detected at a maximum distance of about 4.5 meters. Note that the cell phone needs to be held approximately vertically for the system to work correctly. In fact, the system produces correct detection if the camera is rotated around the principal axis (the line through the optical center orthogonal to the focal plane) by a roll angle of no more than approximately ±40°.

The user is informed of a detection by an acoustic signal (beeping) produced by the cell phone. If the marker is detected at a distance of more than 1.5 meters, then short beeps are repeated at a rate of approximately 2 beeps per second. At shorter distance, the beep repetition rate is of about 5 beeps per second. In order to ensure that the cell phone is held correctly (approximately vertical), a longer beep with lower pitch is produced whenever the user holds the cell phone at a roll angle of more than ±30° for more than half a second. The roll angle is measured using the onboard accelerometer.

The spatial combination of the four colors in the marker can be used to uniquely characterize each marker (24 different combinations are possible). Several markers (with different color combination) were placed in each testing site. At the beginning of each test, a sighted attendant punched in the ID (color combination) of the marker that was to be found in that test.

In our study, we experimented with a few variations of the original, *baseline* algorithm, in order to understand the effect of various system parameters on the effectiveness of the system as a wayfinding tool. *Smaller Marker size (SMS)* artificially reduces the maximum detectable distance to 2.5 meters (which is equivalent the maximum distance at which a marker with diameter of 9 cm can be detected). *Reduced Frame Rate (RFR)* artificially reduces the processing rate to 1 fps. *Reduced Field of*

View (RFOV) artificially reduces the camera's field of view (originally, 35° by 45°) by approximately one half. A theoretical model for the probability of detection of a marker while the user scans the environment with the camera at a certain angular rate as a function of the camera's frame rate and field of view was introduced in [1][2].

3. USER STUDIES

User studies were performed with three volunteers (two males and one female). All volunteers are blind, with only some residual light perception, and are independent ambulators. The experiments were performed in the Spring of 2010 in three locations within the E2 building in the UCSC campus: (1) A narrow corridor (NC) 14 meters long with 7 markers sticking out of the wall at 90 degrees, placed on both sides of the corridor at a height of about 1.5 m; (2) A wider corridor (WC, shown in Figure 1), 11 m in length, with 5 markers attached directly to the wall (flush with the wall) at a height of approximately 1.6 m; (3) An open space (OS), opening onto three corridors, and containing a large pillar, with 8 markers attached to the walls and to the pillar at heights variable between 1.15 m and 1.8 m.

Each participant was tested in all three locations, with the order of the locations balanced across participants. After proper training, the participant was asked to perform five practice runs using the *baseline* algorithm. At the beginning of each run, one marker was chosen at random, and its ID was punched into the phone by the sighted attendant. After a practice session, the phone was configured for one of the three settings (SMS, RFR, RFOV). The participant was explained the difference between the new setting and the baseline setting, and was taken to a marker in a specific location to familiarize themselves with the new setting. The participant was then taken to a starting position, from which five runs were conducted under the same modality as for the practice run, but with the new setting. This operation was repeated for all three settings, with the order of the settings balanced across participants and across locations. At each run, the participant was asked to find the marker, to walk towards it and to touch the marker with their hand. One of the investigators walked behind the participant and measured the *detection time* t_1 elapsed from the beginning on the run till the first detection (signaled by a beep) as well as the *homing-in time* t_2 elapsed from the first beep until the participant touched the marker. After the participants finished the tasks with all three settings, the whole process was started again in the next location until all three locations were tested. The whole experiment took about three hours for each participant. At the end of the experimental sessions, a debriefing interview was conducted.

3.1 Overall Quantitative Measurements

The main quantities of interest in our tests were the *effective speed* and the *homing-in time*. The effective speed (the average speed of the tester until he or she found a marker) describes how confidently and effectively a user of this technology walks while searching for a marker. The homing-in time t_2 (the average time it takes to reach a marker once it has been discovered, introduced earlier) describes how well the system supports micro-guidance to the precise location of the marker.

The median (computed over the five runs in a test with a given setting and a given location) effective speeds ranged from 6 cm/s to 30 cm/s in this narrow corridor. The median homing-in times ranged from 6 to 25 seconds. All participants reported lower homing-in times for the SMS case. This can be justified by the fact that under the SMS setting, detection occurs only at a closer distance, thus facilitating the homing-in task. Only Participant 1 was able to successfully complete all tests in the WC and in the OS location. Median homing-in times for this participant in the WC location were: 13 s (SMS), 17 s (RFR), and 14 s (RFOV), while in the OS location the median homing-in times were: 8 s (SMS), 12 s (RFR), 59 s (RFOV). This data suggests that the reduced FOV setting makes homing-in particularly challenging in the OS location.

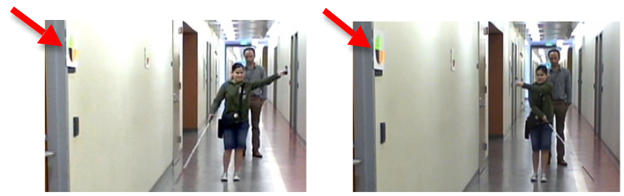

Figure 1. Participant P3 searching for targets attached to a corridor's walls in the WC location.

4. DISCUSSION AND CONCLUSIONS

Different participants had very different performances using the system. Whereas Participant 1 was successful in almost all tests, Participant 2 and 3 were successful only in the NC location. The marker placement also takes an important role. Marker sticking out of the wall as in the NC location seem to enable better detection and homing-in strategies. Among the different system settings, it seems that the reduced field of view setting is the most challenging one. The role of frame rate and of marker size did not result fully clear in the experiments.

Perhaps surprisingly, "homing-in" on a detected marker resulted in an often difficult and frustrating process. One might expect that, once the marker has been detected, the user simply needs to walk in the direction where the camera is pointing to, using proprioception to identify and follow this direction. This turns out to be not so simple in practice. Only one participant was able to successfully perform this lock-in task, and this only in the simplest environment (NC). Future work will investigate technological solutions to support homing-in once the marker has been detected, thus facilitating the whole guidance process.

5. ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grants No. IIS-0835645 and CNS-0709472. We thank the volunteers who participated in this study.

6. REFERENCES

[1] Coughlan, J., and Manduchi, R. 2007. Color targets: Fiducials to help visually impaired people find their way by camera phone. *EURASIP J. Im. Vis. Pr.*, 2007.

[2] Coughlan, J., and Manduchi, R. 2009. Functional assessment of a camera phone-based wayfinding system operated by blind and visually impaired users. *Int. J. Artif. Intell. Tools.,* 18, 3 (June 2009), 379-397