Accessible Spaces: Navigating through a Marked Environment with a Camera Phone

Kee-Yip Chan
Dept. of Information Syst. Mgmt.
University of California
Santa Cruz, CA 95064
kyc@soe.ucsc.edu

Roberto Manduchi
Dept. of Computer Engineering
University of California
Santa Cruz, CA 95064
manduchi@soe.ucsc.edu

James Coughlan
Smith-Kettlewell
Eye Research Institute
San Francisco, CA 94115
coughlan@ski.org

ABSTRACT

We demonstrate a system designed to assist a visually impaired individual while moving in an unfamiliar environment. Small and economical color markers are placed in key locations, possibly in the vicinity of other signs (bar codes or text). The user can detect these markers by means of a cell phone equipped with a camera. Our demonstration highlights a number of novel features, including: improved acoustic interfaces; estimation of the distance to the marker, which is communicated to the user via text-to-speech (TTS); increased robustness via rotation invariance, which makes the system easier to use for users with reduced dexterity.

Categories and Subject Descriptors

I.5.4 [**Pattern Recognition**]: Applications: *Computer Vision*.

General Terms

Algorithms, Performance, Experimentation, Human Factors.

Keywords

Wayfinding, Orientation. Blindness.

1. INTRODUCTION

Wayfinding, the ability to orient oneself and to find the path to destination in a possibly unknown environment, is a necessary skill for independent traveling. Individuals who, due to visual impairment, cannot access critical visual information (signage or landmarks), face a daily challenge to their ability to move about alone and, ultimately, to their independence.

Various types of solutions have been proposed and implemented, often aimed at specific environments. Braille signs, as mandated by the ADA for public accommodation, are effective only when the user knows where they are located and can reach them with his or her hand (and is obviously useful only for those who know how to read Braille). Detectable warnings such as raised truncated domes and grooves on sidewalks (normally painted in bright yellow) may help detecting the end of a sidewalk or a ramp and directing a person toward a nearby destination. Environment marking by means of RFID tags has been proposed recently [2][7], however, RFID tags can only be read at limited distances.

ASSETS'07, October 15-17, 2007, Tempe, Arizona, USA.
ACM 978-1-59593-573-1/07/0010.

Acoustic signals, such as those implemented in Accessible Pedestrian Signs (APS) at traffic intersections, have been deployed in many cities throughout the world, but concerns have been raised about their effectiveness and ensuing acoustic pollution. Light beaconing using infrared transmitters (Talking Signs®) or fluorescent light fixtures (Talking Lights®) may provide location-aware information (in the form of analog or digital signals) to the user carrying a special receiver/decoder. In particular, Talking Signs® allows one to estimate the pointing direction to the beacon, since the receiver only produces an acoustic output when it is pointed towards the beacon. Location-aware on-demand information can also be provided via RF, for example using WiFi or Bluetooth. In particular, the wide availability of Bluetooth-enabled cell phones makes this an attractive solution [1]. Image understanding algorithms, implemented on one's cell phone (e.g. the "Cross Watch" project [6]), or perhaps on a remote server, hold great promise. Even state of the art automatic image recognition systems, however, are not reliable enough for practical use as yet.

We have introduced [5] a new approach based on the deployment of small, unobtrusive, and economical environment markings, which can be detected at a distance of several meters using a regular camera cell phone. The markers are designed in such a way that they are quickly and robustly recognized even considering the limited computational power of the cell phone. Pointing direction and even approximate distance to the marker can be inferred directly. No additional specialized equipment is needed, thus avoiding the stigma often associated with assistive technology. Although the markers do not carry information *per se*, they can be used as pointers to other signals, for example, a bar code or some text placed nearby. Whereas detecting a bar code or a text signal in an unknown position in an image is a challenging task, our markers allow the system to immediately focus on the region(s) of interest in the image. Then, the text or bar code can be analyzed using all resources available, or, if necessary, the user may move closer to the marker in order to get a higher resolution image of it. Our proposed marker is pie-shaped, containing 3 or 4 areas with distinctive colors (see Fig. 1). Detection is performed by running a cascade of tests at each pixel location. Each test involves comparing two nearby pixels over one color component. Image points that are unlikely to correspond to a marker are normally discarded during the initial tests. Those points that pass the whole sequence are then tested further using a spatial coherence criterion. Only clusters of points larger than a predefined size are labeled as marker locations. For more details about the algorithm, please see [5].

Figure 1. Left: Two versions of our proposed color marker. Right: A blind user operating the cell-phone to detect a marker.

Previous work focused on optimal marker design [4], on fast algorithms for robust marker detection [5], and on performance analysis with blind users [3]. This demonstration, which will be performed on a regular cell phone (Nokia N70), will highlight a number of new system features, as described in the next section.

2. NEW FEATURES

2.1 Improved User Interface

Our previous implementation communicated the presence of the marker in the image to the user by a sequence of impulsive sound bits. We have designed an improved user interface, based on continuous or intermittent sound of variable pitch and loudness as well as on text-to-speech (TTS). A combination of pitch and loudness is used to identify the location of the detected marker in the camera's field of view, thus allowing a visually impaired user to quickly identify the pointing direction to the marker. TTS (from Loquendo, S.p.A.) is used for higher-level information, such as text decoded from a bar code, or the distance to the marker when requested.

2.2 Estimating Distance to Marker

If the size of the marker is standardized, and the optical parameters of the camera are known, it is possible to estimate the approximate distance to the marker by measuring its apparent size in the image. The amount of foreshortening as well as the marker's apparent aspect ratio depend on its distance as well as on the viewing direction (which can be rather skewed when the marker is seen from an angle, for example when it is located on the side wall of a narrow corridor). We have implemented a simple segmentation algorithm that, starting from the detected center of the marker, quickly finds its contour, to which an ellipse is fitted. From the ellipse parameters, the distance to the marker is estimated. The accuracy of estimation depends both on the quality of the image (which may be blurred due to motion) and on the distance to the marker (due to the finite resolution of the pixel grid). Information about the distance is provided to the user when he or she requests it by pressing a key. Depending on the expected accuracy, different textual formats are used, such as: "Distance is larger/smaller than x meters", "Distance is between x_1 and x_2 meters", or "Distance is approximately x meters".

2.3 Improved Orientation Invariance

Our previous search algorithm tested each pixel in the image for the presence of a marker by centering a 3- or 4-point pattern (depending on the marker type) at the pixel, and by running a sequence of comparisons involving the color content of the pixels in the pattern. While this approach assumes that the user holds the camera in an upright position, the structure of the marker allows for a certain degree of rotation invariance (nominally, up to 60° for 3-color markers and 45° for 4-color markers). Our experience shows that a wider range of invariance is needed in practice, especially when the user is unable to precisely control the camera orientation. We have implemented an improved search algorithm that looks for markers at different orientations. The additional computational cost is marginal, however, the increased rotation invariance slightly increases the number of false positives, which then need to be removed by the final clustering stage.

3. ACKNOWLEDGMENTS

This work was completed as part of UCSC's SURF-IT summer undergraduate research program, an NSF CISE REU Site supported under Grant No. CCF-0552688. J. Coughlan and R. Manduchi are supported in part by the NIH under Grant 1-R21-EY017003-01A1.

4. REFERENCES

[1] S. Bohonos, A. Lee, A. Malik, C. Thai, and R. Manduchi. Universal Real-Time Navigational Assistance (URNA): An Urban Bluetooth Beacon for the Blind. *1st International Workshop on Systems and Networking Support for Healthcare and Assisted Living Environments (HealthNet 2007)*, June 2007.

[2] Tsung-Hsiang Chang, Chien-Ju Ho, David Chawei Hsu, Yuan-Hsiang Lee, Min-Shieh Tsai, Mu-Chun Wang, and Jane Yung-jen Hsu. iCane - A Partner for the Visually Impaired. *2nd International Symposium on Ubiquitous Intelligence and Smart Worlds (UISW 2005), Embedded and Ubiquitous Computing Workshops (EUS)*, 2005: 393-402.

[3] J. Coughlan and R. Manduchi. Functional Assessment of a Camera Phone-Based Wayfinding System Operated by Blind Users. *IEEE-BAIS Symposium on Research on Assistive Technology*, Dayton, OH, April 2007.

[4] J. Coughlan and R. Manduchi. Color Targets: Fiducials to Help Visually Impaired People Find Their Way by Camera Phone. *EURASIP Journal on Image and Video Processing*, in press.

[5] J. Coughlan, R. Manduchi, and H. Shen. Cell Phone-based Wayfinding for the Visually Impaired. *1st International Workshop on Mobile Vision,* Graz, Austria, May 2006.

[6] J. Coughlan and H. Shen. A Fast Algorithm for Finding Crosswalks using Figure-Ground Segmentation. *2nd Workshop on Applications of Computer Vision*, Graz, Austria, May 2006.

[7] N. Fukasawa, H. Matsubara, S. Myojo, and R. Tsuchiya. Guiding Passengers in Railway Stations by Ubiquitous Computing Technologies. *IASTED Human- Computer Interaction*, Phoenix, AZ, 2005.