Crosswatch: a Camera Phone System for Orienting Visually Impaired Pedestrians at Traffic Intersections

Volodymyr Ivanchenko, James Coughlan and Huiying Shen

The Smith-Kettlewell Eye Research Institute,
2318 Fillmore St.,
San Francisco CA 94115
USA
{vivanchenko, coughlan, hshen}@ski.org

Abstract. Urban intersections are the most dangerous parts of a blind or visually impaired person's travel. To address this problem, this paper describes the novel "Crosswatch" system, which uses computer vision to provide information about the location and orientation of crosswalks to a blind or visually impaired pedestrian holding a camera cell phone. A prototype of the system runs on an off-the-shelf Nokia camera phone in real time, which automatically takes a few images per second, uses the cell phone's built-in computer to analyze each image in a fraction of a second and sounds an audio tone when it detects a crosswalk. Tests with blind subjects demonstrate the feasibility of the system and its ability to provide useful crosswalk alignment information under real-world conditions.

1 Introduction

Urban intersections are the most dangerous parts of a blind or visually impaired person's travel. A pedestrian must locate a crosswalk before attempting to traverse an intersection, and proper alignment is necessary to enter the crosswalk in the right direction and avoid straying outside of it; these tasks are especially difficult for pedestrians with low or no vision. To address this problem, this paper describes the novel "Crosswatch" system, which uses computer vision technology to provide such information about crosswalks to a blind or visually impaired pedestrian holding a portable camera. A prototype of our system has been implemented on an off-the-shelf Nokia camera phone. The system runs in real time, automatically taking a few images per second, using the cell phone's built-in computer to analyze each image in a fraction of a second and sound an audio tone when it detects a crosswalk. By panning the cell phone left and right, the user can determine if any crosswalk is visible and if so, align him/herself to it.

We describe tests with blind subjects that demonstrate the success of the system in signaling the presence or absence of crosswalks as well as providing alignment information. This is the first example of a portable system we are aware of that provides orientation information at traffic intersections based solely on existing crosswalk patterns already used by sighted pedestrians, without the need for any added infrastructure (such as Audible Pedestrian Signals or other guidance systems requiring on-site installation).

2 Previous Work and Overview of Our Approach

Several types of technologies have been developed to assist blind and visually impaired individuals in crossing traffic intersections. Most prevalent among them are Audible Pedestrian Signals [1], which generate sounds that tell blind and visually impaired individuals when to cross intersections, and which also furnish acoustic cues that pedestrians can localize to help them properly orient themselves to the crosswalk. Talking Signs® [4] allow blind travelers to locate and identify landmarks, signs, and facilities of interest, at intersections and other locations, using signals from installed infrared transmitters that are converted to speech by a receiver carried by the traveler. However, the adoption of both Audible Pedestrian Signals and Talking Signs® is very sparse, and they are completely absent in most US cities. Finally, GPS (Global Positioning System) technology has the advantage of being widely available [6], but unfortunately on urban sidewalks near buildings the localization accuracy is only sufficient to resolve within a few street addresses – not nearly enough to provide useful guidance within a traffic intersection.

(a) **(b)**

Fig. 1. (a) Camera cell phone held by blind user. (b) Schematic diagram shows overhead view of zebra crosswalk and two users holding cell phone system: cell phone on left is aligned with crosswalk and makes an audio tone; cell phone on right is not aligned and makes no sound.

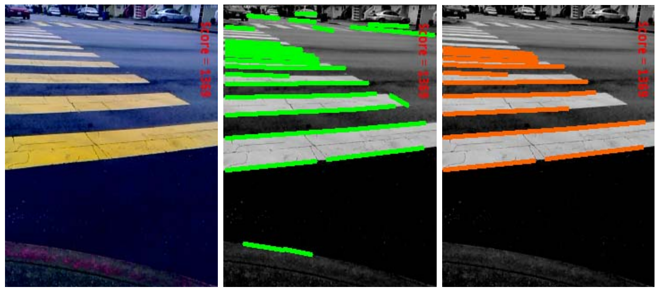
The alternative approach that we propose in this paper is to use a portable computer vision-based system to identify important features in an intersection. With this system (Fig. 1), the user takes an image of the intersection with a cell phone camera, which is analyzed by software run on the cell phone, and the output of the software is communicated to the user with audio tones. The great strength of the computer vision approach is that it offers virtually unlimited potential for acquiring information about one's surroundings, but does not require any infrastructure beyond what is already provided for sighted people – namely, painted street markings. A small amount of work has been done on developing computer vision algorithms specifically for analyzing traffic intersections [10,8,7,9], but such work has been conducted using near-ideal pictures of traffic intersections and has not been tested on more difficult images commonly encountered in which crosswalk visibility is impaired by occlusions, shadows and paint imperfections.

The use of a camera phone platform also has many important advantages. First, many pedestrians are already carrying cell phones, and our system won't burden them with having to carry an additional device. Second, the system is installed simply by downloading our software on an off-the-shelf cell phone, without requiring any special hardware or hardware modifications. Third, the cell phone is a mainstream consumer product which raises none of the cosmetic concerns that might arise with other assistive technology requiring custom hardware. Finally, we note that our system can easily be used in conjunction with standard orientation techniques used at traffic intersections. For instance, the user can hold the white cane in one hand to locate the curb and curb cuts near the crosswalk (note that in some cases the curb cut points directly to the crosswalk path while in other cases it points into the center of the intersection, which can be a source of confusion) while holding the camera phone system in the other hand.

3 Finding Crosswalks in Images

In this work we specialize to the zebra (striped) crosswalk (Fig.'s 1a, 2a), and leave the problem of finding the two-stripe crosswalk, another common type of crosswalk (that is much less visible than the zebra crosswalk), for later research. We draw on our past work [3] for finding zebra crosswalks under challenging real-world conditions such as cluttered traffic intersections with shadows, intensity saturation, slightly curved stripes (due to curvature in the road and/or camera distortion), crosswalk paint irregularities and occlusions that block the view of part of the crosswalk. Our approach (see [5] for details) first extracts straight-line segments from the image (Fig. 2b) and then applies a segmentation algorithm to group together those segments that conform to a valid crosswalk pattern. A learning procedure was used to train the algorithm from a database of typical images of traffic intersections, including those with and without crosswalks.

Once the crosswalk detection algorithm has identified which segments (if any) belong to the crosswalk (Fig. 2), the number and length of detected crosswalk lines are summarized in a single score that reflects how much of a crosswalk is visible in the image. If this score is above a particular threshold (500, an empirically chosen value), we decide that the image contains a significant portion of a crosswalk. (Note that although some of the crosswalk segments in Fig. 2 are missed by the algorithm, the score is still well above the detection threshold.) Several sample images are ranked according to this score, shown in Fig. 3.



(a) **(b)** **(c)**

Fig. 2. (a) Typical intersection image photographed by camera phone. (b) Straight-line segments (green) extracted from image. (c) Segments (orange) chosen by algorithm as belonging to crosswalk. The algorithm assigns a score of 1369 to the image, well above the minimum threshold (500) for detecting a crosswalk.

1369	1210	1079	1037	34	0
					

Fig. 3. Images ranked according to the score assigned by the crosswalk detection algorithm; scores (printed above each image) decrease from left to right. The first four images are above threshold (and are thus classified as containing crosswalks), while the last two are below threshold (and are thus classified as not containing crosswalks).

4 User Interface

The preceding crosswalk detection algorithm was developed and tested on a desktop computer and then ported to the Nokia N95 camera phone running Symbian C++. The Symbian cell phone operating system (OS) platform was chosen for two main reasons. First and foremost, Symbian is a popular operating system that is easily made blind or low vision-accessible because it "allows the installation of third-party screen-reading and screen-magnification software applications" [2]; as a result, Symbian phones (most of which are made by Nokia) are popular among visually impaired cell phone users. Second, Symbian C++ is also a well-supported dialect of C++ that provides performance superior to other languages supported on cell phones (such as Java).

The algorithm was implemented to run as quickly as possible on the cell phone, so that up to three images could be processed per second. Given the speed of the algorithm, we decided to have the system run in a video capture mode (in which several images are automatically captured every second). For each image, if the score produced by the detection algorithm (Fig.'s 2, 3) is above threshold, then a brief audio tone is sounded to indicate the presence of a crosswalk; otherwise no tone is sounded.

With this simple interface, the user slowly pans the camera left and right in front of him/her, listening to which camera directions are accompanied by audio tones. In this way, the user can determine the location of the left and right edges of the crosswalk, and thereby estimate where the center of the crosswalk is located.

5 Experimental Results and User Testing

We explained the purpose and operation of the Crosswatch system to two completely blind volunteer subjects, and advised them on how best to hold the camera phone. Particular emphasis was placed on the importance of moving the camera slowly to avoid motion blur and to ensuring the camera lens was not covered (e.g. by the user's fingers). After a brief training session for each of two experiments, we took the subjects to unfamiliar traffic intersections to test the system.

The first experiment was designed to test the usability of the system (e.g. can a blind subject point a camera phone accurately enough for the system to provide useful information?) and to test the ability of a subject using the system to determine whether or not there was a crosswalk visible at a given traffic intersection. 15 different traffic intersections were chosen in advance, such that there was a 50% chance that a zebra crosswalk was present at each intersection. Intersections without a zebra crosswalk either had no crosswalks or had one or more non-zebra (i.e. two-stripe) crosswalks. The experiment was designed so that there were no cues available to the blind subject (e.g. ambient environmental sounds or tactile cues from the white cane) about the presence or absence of a crosswalk except for those furnished by the cell phone system. The result of the experiment was that the subject answered correctly whether a zebra crosswalk was present or absent at all 15 intersections. An exact binomial test shows that the subject responded significantly above chance, with $p = 3.05 \cdot 10^{-5}$.

The purpose of the second experiment was to test the ability of another blind subject to align him/herself with a zebra crosswalk. We led the subject near a traffic intersection and told him/her to approach the crosswalk and to try to orient properly to it. When the subject felt correctly oriented, he/she pressed a button on the cell phone to take a picture of the crosswalk. In half the trials, the audio feedback from the Crosswatch system was turned off, and in the other half the feedback was turned on as usual. This procedure was repeated at all four corners of two separate crosswalks. Sample images captured during the experiment are shown in Fig. 4, showing that the subject was better able to align him/herself using the feedback from the system. Indeed, the only times that the subject took images in which little or no part of the crosswalk was visible was when feedback was not provided.

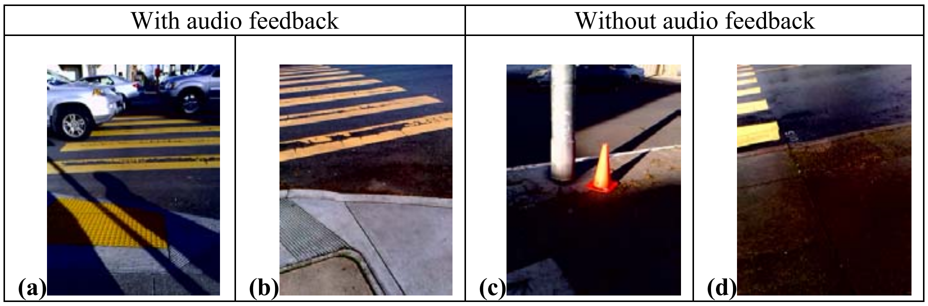


Fig. 4. Pictures taken by blind subject from Experiment 2. (a), (b) show images with audio feedback from the cell phone; (c), (d) show images without audio feedback.

6 Conclusion

Crosswatch is a novel camera phone-based system for helping blind and visually impaired pedestrians find crosswalks and align themselves properly to them before crossing. A prototype system has been implemented on the Nokia N95 camera phone, which searches for crosswalk stripes a few times per second and provides audio feedback whenever a crosswalk is detected. We have conducted preliminary experiments with blind volunteers to test the system, demonstrating its feasibility.

In the future we will improve the robustness of our system by adding an additional stage of processing to verify that the global geometry and appearance of a detected crosswalk pattern is valid, in order to reject false crosswalk detections (which are rare but obviously very important to avoid). We will also experiment with augmenting the user interface with detailed geometric information about the user's orientation relative to the crosswalk (e.g. the crosswalk direction is so many degrees to the right or left) to determine if such information is helpful, and if so then how the information should best be communicated. Ultimately we will expand the capabilities of the Crosswatch system to handle more types of crosswalk patterns, such as two-stripe crosswalks and variants of the zebra crosswalks we have explored in our experiments, and to provide more information about traffic intersections, such as the current state of traffic signal lights (e.g. "Walk" signals) and intersection layout.

Acknowledgments. The authors were supported by The National Institute on Disability and Rehabilitation Research (NIDRR) grant H133G030080, The National

Institute of Health grants 1 R21 EY015187-01A2 and 1 R01 EY018345-01, and the Claire Giannini Fund.

References

1. Barlow J. M., Bentzen, B. L., Tabor, L. Accessible pedestrian signals: Synthesis and guide to best practice. National Cooperative Highway Research Program (2003)
2. Burton, D. You Get to Choose: An Overview of Accessible Cell Phones. Access Issues, Vol. 6, no. 2. (2005) http://www.afb.org/afbpress/pub.asp?DocID=aw060206&select=1#1
3. Coughlan, J., Shen, H. A Fast Algorithm for Finding Crosswalks using Figure-Ground Segmentation. In Proc. 2nd Workshop on Applications of Computer Vision, a satellite workshop of European Conference on Computer Vision. Graz, Austria (2006)
4. Crandall, W., Bentzen, B., Myers, L., Brabyn, J. New orientation and accessibility option for persons with visual impairment: transportation applications for remote infrared audible signage. Clinical and Experimental Optometry, 84(3): 120-131 (2001)
5. Ivanchenko, V., Coughlan, J., Shen, H. Detecting and Locating Crosswalks using a Camera Phone. Submitted to Fourth Workshop on Embedded Computer Vision (ECVW), associated with Computer Vision and Pattern Recognition (2008)
6. Loomis, J, Klatzky, R, Golledge, R. Navigating without vision: basic and applied research. Optom Vis Sci. 78(5):282-9 (2001)
7. Se, S., Brady, M. Road Feature Detection and Estimation. Machine Vision and Applications Journal, Vol. 14, no. 3, pp. 157-165 (2003)
8. Shioyama, T., Wu, H., Nakamura, N., Kitawaki, S. Measurement of the pedestrian crossings and detection of traffic lights from image data. Meas. Sci. Technol. 13, pp. 1450-1457. 14 (2002)
9. Uddin, M.S., Shioyama, T. Bipolarity- and Projective Invariant-Based Zebra-Crossing Detection for the Visually Impaired. 1st IEEE Workshop on Computer Vision Applications for the Visually Impaired, a satellite workshop of Computer Vision and Pattern Recognition (2005)
10. Utcke, S. Grouping based on Projective Geometry Constraints and Uncertainty. Internation Conference on Computer Vision, Bombay, India. (1998)