Functional Assessment of a Camera Phone-Based Wayfinding System Operated by Blind Users

James Coughlan[*] and Roberto Manduchi[**]
[*]*Smith-Kettlewell Eye Research Institute, San Francisco*
[**]*Department of Computer Engineering, University of California, Santa Cruz*

Abstract

A major challenge faced by the blind and visually impaired population is that of wayfinding – the ability of a person to find his or her way to a given destination. We recently developed [4] a novel wayfinding aid based on a camera phone, which is held by the user to find and read aloud specially designed machine-readable signs in the environment (labeling locations such as offices and restrooms). These signs are simple color patterns (targets) that can be quickly and reliably identified using image processing algorithms running on the camera phone.

In order for this system to be truly usable, it is important to ensure that color targets in the environment can be quickly discovered by a user holding the camera phone. This depends on a number of factors, including the scanning strategy adopted by the user, the image acquisition and processing time, and the detection rate of the algorithm under different conditions of illumination and viewing geometry.

In this work, we develop simple models that allow us to tune a number of algorithmic parameters and to assess the performance of the system in terms of average time to discovery in typical conditions. Experiments with blind individuals are presented in order to validate our analytical models.

1. Introduction

There are nearly 1 million legally blind persons in the United States, and up to 10 million with significant visual impairments. A major challenge faced by this population is that of *wayfinding* – the ability of a person to find his or her way to a given destination. Well-established orientation and mobility techniques using a cane or guide dog are effective for following paths and avoiding obstacles, but are less helpful for finding specific locations or objects.

In recent work we have developed a new assistive technology system to aid in wayfinding based on a camera cell phone, which is held by the user to find and read aloud specially designed signs in the environment [4]. These signs consist of barcodes placed adjacent to special landmark symbols (see Fig. 1). The symbols are designed so as to be easily detected and located by a computer vision algorithm running on the cell phone; their function is to point to the barcode to make it easy to find without having to segment it from the entire image. Our proposed system, which we have already prototyped, has the advantage of using standard off-the-shelf cellphone technology – which is inexpensive, portable, multi-purpose and becoming nearly ubiquitous – and simple color signs that can be easily produced on a standard color printer. Another advantage of the cell phone is that it is a mainstream consumer product, creating none of the cosmetic concerns that might arise with other assistive technology requiring custom hardware.

2. Related work

A number of approaches have been explored to help blind travelers with orientation, navigation and wayfinding, most using modalities other than computer vision. Among the most promising include infrared signage that broadcasts information received by a hand-held receiver [6], GPSbased localization, RFID labeling, and indoor Wi-Fi based localization (based on signal strength) and database access. However, each of these approaches has significant limitations that limit their attractiveness as stand-alone solutions. Infrared signs require costly installation and maintenance; GPS has poor resolution in urban settings and is unavailable indoors; RFIDs can only be read at close range and would therefore be difficult to locate by blind travelers; and Wi-Fi localization requires extensive deployment to ensure complete coverage, as well as a time-consuming calibration process.

Research has been undertaken on computer vision algorithms to aid in wayfinding for such applications as navigation in traffic intersections [3,16] and sign reading [14]. The obvious advantage of computer vision is that it is designed to work with little or no infrastructure or modification to the environment. However, none of it is yet practical for commercial use because of issues such as insufficient reliability and prohibitive computational complexity (which is especially problematic when using the kind of portable hardware that these applications require).

Our approach, image-based labeling, is motivated by the need for computer vision algorithms that can run quickly and reliably on portable camera cell phones, requiring only minor modifications to the environment (i.e. posting special signs). Image-based labeling has been used extensively for product tagging (barcodes) and for robotic positioning and navigation (fiducials) [2, 15, 10, 1, 9]. It is important to recognize that a tag reading system must support two complementary functionalities: detection and data embedding. These two functionalities pose different challenges to the designer. Reliable detection requires unambiguous target appearance, whereas data embedding calls for robust spatial data encoding mechanisms. Distinctive visual features (shapes and textures or, as in this proposal, color combinations) can be used to maximize the likelihood of successful detection. Computational speed is a critical issue for our application. We argue that color targets have a clear advantage in this sense with respect to black and white textured patterns.

Variations on the theme of barcodes have become popular for spatial information encoding. Besides the typical applications of merchandise or postal parcel tagging, these systems have been demonstrated in conjunction with camera phones in a number of focused applications, such as linking a product or a flyer to a URL. Commercial systems of this type include the Semacode, QR code, Shotcode and Nextcode. An important limitation of these tags is that they need to be seen from a close distance in order to decode their dense spatial patterns. Our approach addresses both requirements mentioned above by combining a highly distinctive fiducial with a barcode.

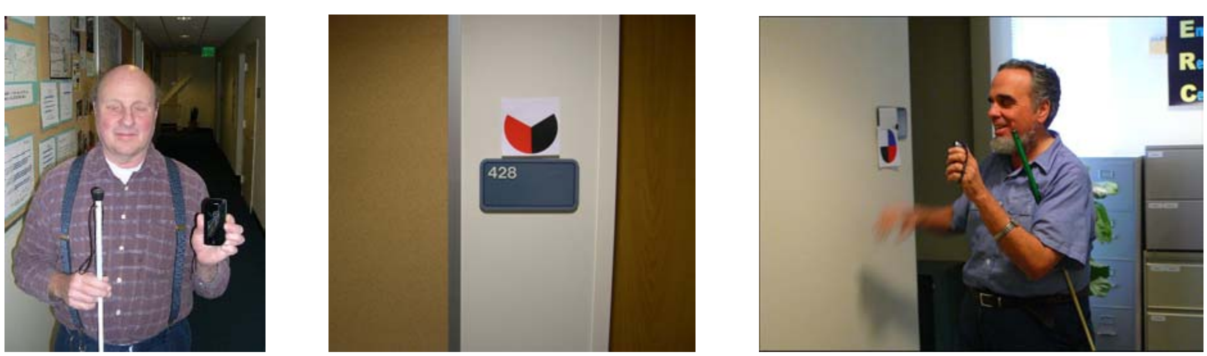
We originally introduced the concept of a color target for wayfinding, along with a fast barcode reader, in [4]. However, in [4] the target was designed based on purely heuristic criteria. In this paper we provide a sound approach to the joint design and testing of the color target and of the detection algorithm.

3. Color target design

Our ultimate goal is to enable a blind individual access to location-specific information. This information would typically be used for wayfinding purposes, such as locating a restroom, the elevator, the exit door, or someone's office. Information may be provided in graphical form, such as text (possibly in a standardized form) or other machine-readable format, such as 1-D or 2-D barcode. The image processing algorithms in the cell phone translate it into a form accessible by the blind user (e.g. via text-to-speech). In addition, our system provides the user with information about the direction towards the text or barcode, which is very useful for self-orientation.

Reading a text or barcode sign requires scanning the whole image in order to localize the sign. This can be computationally very demanding, thus slowing down the effective acquisition rate. For this purpose, we have proposed the use of a simple color pattern (*target*) placed near the barcode or text. This pattern is designed in such a way that detection can be done very quickly. If a color target is detected, then only the nearby image area undergoes further, computationally heavy processing for text or barcode reading.

Fast detection is obtained by a quick test that is repeated at all pixel locations. For each pixel x in the



(a) (b) (c)

Figure 1. Camera phone-based wayfinding system. (a): Blind user holding a Nokia 7610 camera phone. (b): Machine-readable sign consisting of 3-color target and barcode. The color target is designed to swiftly guide the system to the location of the barcode in images taken by the camera phone, which is then decoded and read aloud as synthetic speech. (c): A 4-color version of the target.

image, the color data in a set of *probing pixels* centered around *x* is analyzed (see Figure 2 for the geometry of probing pixels with a 3-color target). Detection is possible when the probing pixels are each in a distinct color area of the target. A cascade-type discrimination algorithm is used to compare the color channels in the probing pixels, resulting in a binary decision about whether *x* belongs to the color target or not. A subsequent fast clustering algorithm is used to rule out spurious false detections.

The choice of the colors in the target is very important in terms of detection performances. The color combination should be *distinctive*, to avoid the risk that image areas not corresponding to a target may mistakenly trigger a detection. At the same time, the chosen colors should provide *robustness* of detection in the face of varying light conditions. Finally, a limited number of colors in the pie-shaped target is preferable, in order to enable correct detection even when the cell phone is not kept in the correct orientation. For example, it is clear from Figure 2 that in the case of a 3-color target, the camera can be rotated by up to ±60° around its focal axis without impairing detection (because the probing pixels will still be located within the correct color areas). If 4 colors are used, each angular sector will be of 90°, and therefore the allowable rotation is only ±45°.

Note that distinctiveness and invariance are important characteristics of feature detection algorithms for numerous vision tasks such as object recognition [8,11] and tracking [13]. With respect to typical vision applications, however, we have one degree of freedom more, namely the choice of the target that we want to recognize.

In the following we present a method for the optimal choice of the colors in the target. For more details, the reader is referred to [5].

3.1. Color pattern selection

In previous work [4] we used a pattern with three colors (red, green, blue). This choice was suggested by the fact that these colors have well separated spectra. In fact, a principled color selection method should consider more than just spectral separation. In order for the color combination to be distinctive, it is important to take into account the statistics of typical background patterns in the environment where the color target will be placed. In addition, color variability due to different illumination and color reproduction should be accounted for.

As mentioned earlier, our detection algorithm is based on a cascade of tests on individual color channels of pairs of probing pixels. More specifically, let $c_m = \left(c_m^1, c_m^2, c_m^3\right)$ represent the color vector as measured by the probing pixel for the *m*-th patch. Then, a query involving the *m*-th and *n*-th color patches over the *k*-th color channels can be expressed as by $\left(c_m^k - c_n^k\right) \geq T_{m,n}^k$,where $T_{m,n}^k$ is a suitable threshold. The quadruplet $Q = \left(m, n, k, T_{m,n}^k\right)$ fully characterizes the query. The detection algorithm is thus defined by the sequence of *J* queries $(Q_1, Q_2, \ldots, Q_J)$. Only if a pixel satisfies the whole cascade of queries it is considered to be a candidate target location. The advantage of using a cascade structure is that, if the first queries are very selective, then only few pixels need to be tested in the subsequent queries.

In order to choose a set of (typically 3 or 4) colors from a pre-selected pool of possible candidates, we first take a number of pictures of the color patches in the pool under different illumination conditions (fluorescent and incandescent light, as well as natural light from clear and overcast sky). For each pair (m,n) of colors in the pool, and for each color channel *k*, we find the smallest threshold $T_{m,n}^k$ such that all pixels in the color patch pair pass the test *Q*. This guarantees that no false negatives (missed detection) will occur. However, the test *Q* may still pass a possibly large number of non-target instances (false positives). Our next step is to measure the false positives rate of each test *Q* over a set of representative "background" test images not containing any color targets. We then consider all possible cascades of *J* queries, drawn from the color pool, and select the cascade with associated minimum false positive rate. The resulting color set and associated *J* queries thus guarantees theoretical zero rate of missed detections and low rate of spurious detections. The queries in the set are ordered according to their associated false positive rate, so as to ensure that the first queries eliminate most points in the image.

We normally use J=4 queries in this work. This choice is justified by the experimental observation that the decline in false positives rate while increasing *J* from 4 to 5 is normally modest, hence larger values of *J* may not improve performances significantly. The pool of candidates is formed by the following colors: red, green, blue, black and white. The optimal triplet of color patches from this pool was found to be (white, red, black), with associated false positives rate of about10^{-3} over the selected "background" images. The optimal quadruplet was formed by (white, red, blue, black), with associated false positives rate of about $2 \cdot 10^{-4}$. The 3- and 4-color optimal targets are shown in Figure 1.

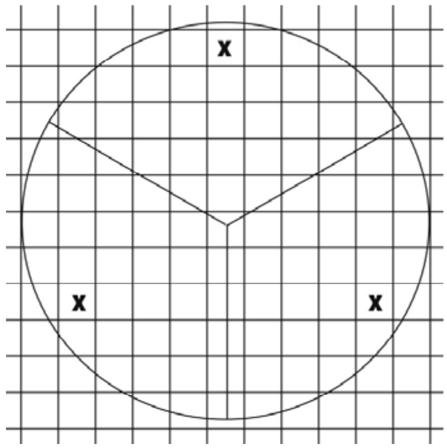


Figure 2. The location of the probing pixels for the 3-color scheme. In this case, the probing pixels separation is equal to 7 pixels.

4. Modeling the system performance

Two main factors contribute to the effectiveness of our wayfinding system: the maximum distance at which the target can be detected, and the speed at which the scene can be scanned while searching for a target. A large target may be easily recognizable from the distance, but of course the target size must be limited due to cosmetic and functional reasons. The maximum distance for detection can be increased by increasing the focal length of the camera. However, this would also imply a reduced field of view, resulting in longer time necessary to scan the scene.

In the following, we present some basic results relating the maximum distance $H_{\max}$ at which the target can be detected as a function of the system parameters, as well models for different possible scene scanning strategies, along with the expected scanning time before target detection.

4.1. Maximum distance for detection

This section summarizes the main results presented in [4] concerning the maximum distance at which a target of a given diameter can be detected when the camera is not moving. For simplicity's sake, we consider only a 3-color target here (the extension of these results to a 4-color target is immediate). We assume that the camera is held in the correct orientation, i.e. with the vertical axis parallel to the line separating the lower two regions of the target as shown in Figure 2 Three *probing pixels* are used to verify the presence of the target in a certain location of the image. The distance M (in pixels) between two such pixels is a very important parameter. A small value for M may impair detection due to edge effects (e.g., color bleeding). On the other hand, the probing pixel separation M directly affects the maximum distance $H_{\max}$ for detection. We showed in [4] that an acceptable form for the relationship between M and $H_{\max}$ is $H_{\max} = \dfrac{D}{IFOV \cdot (M+4)}$, where D is the diameter of the color target and $IFOV$ is the camera's instantaneous field of view (in our case, $IFOV$=1.5 mrads for 640×480 camera resolution). In all our experiments, we set the value of M to 7. This yields a theoretical value for $H_{\max}$ of about 7 meters for a target with diameter of D=12 cm. Note that when illumination is poor, the actual maximum distance for detection is typically lower than the theoretical one due to image noise.

4.2. Modeling the scanning strategy

There are at least two possible strategies for scanning a scene with a camera phone while searching for a color target. The first modality involves rotating the cell phone slowly around its vertical axis (*panning*). This would be suitable, for example, when the target can be expected to be on a wall facing the user. Another possibility is to walk parallel to the wall while holding the cell phone facing the wall (*translation*). A typical scenario for this modality would be a corridor in an office building, with a color target placed near every office door, signaling the location of a Braille pad or of other information readable by the cellphone (see Figure 3).

When images are taken while the camera is moving (as in both strategies described above), motion blur occurs, which may potentially impair the detection performance. We provided a model for motion blur effects in our system in [4]. More precisely, we gave a theoretical expression for the maximum angular speed that does not affect the detection rate, as well as empirical results for the maximum angular speed that ensures a detection rate of at least 50% in different light conditions. Based on the results in [4], we introduce in the following two simple models for describing the motion blur effect in the *panning* and *translation* strategies.

4.2.1. Scene scanning based on panning. Suppose a person is located at a certain distance from a wall where a color target is placed. Assume that the user has no idea of where the target may be (or whether there is a visible target in the scene at all), so she will scan the scene within a 180° span. Also assume that the person is holding the cell phone at the correct

(a) **(b)**

Figure 3. (a): Color targets placed in a corridor near existing Braille signs. (b) A blind user scanning the scene via translation (walking parallel to the wall).

height and elevation, so that, as long as the azimuth angle is correct, the target is within the field of view of the camera.

Motion blur may affect the detection rate p of the system. This depends on the rotational speed, the integration time (which in turn depends on the amount of light in the scene), and on the probing pixel separation M. For example, in [4] we empirically found that, with M=7 pixels, the maximum angular speed ω that ensures detection rate of p=0.5 for our system is ω=60°/s under average light and ω=30°/s under dim light for a color target of 12 cm of diameter at a distance of 2 meters.

Given the frame rate R (in our case, R≈2 frames/s) and the field of view FOV of the camera (in our case, FOV≈55°), we can estimate the average number of consecutive frames n in which a particular point in the scene is seen during scanning: $n=FOV/(\omega/R)$. For example, the color target would be visible in 3 to 4 consecutive frames with ω=30°/s, and in at most 2 consecutive frames with ω=60°/s. The probability that the target would be detected in one sweep is equal to $p_n=1-(1-p)^n$. The number of sweeps until the target is detected is a geometric random variable of parameter p_n. Hence, the expected number of sweeps until the target is detected is equal to $1/p_n$. For example, with ω=60°/s and p=0.5 (average light), an average number of at most 2 sweeps would be required for target detection, while with ω=30°/s and p=0.5 (dim light) the average number of sweeps before detection is $\leq 8/7$.

4.2.2. Scene scanning based on translation. Suppose the user is walking along a corridor while keeping the camera facing the side wall. In this case, it is desirable to ensure that the target is always detected. According to [4], this is achieved when the apparent image motion d (in pixels) during the exposure time T is less than $\lfloor M/2 \rfloor + 1$ pixels. The image motion within the exposure time is related to the actual velocity v of the user and the distance H to the wall as by $d=fvT/Hw$, where f is the camera's focal length and w is the width of a pixel. Note that we can approximate w/f with $IFOV$. For example, when walking at a distance of H=1 meter from the wall, assuming that T=1/60 s (this will depend on the ambient light), the highest allowable speed of the user when M=7 pixel is v=0.36 m/s. Note that if the user is relatively close to the wall, then the probing pixel separation may be increased, which allows for higher admissible translational speed.

5. Experiments with blind subjects

Previous experiments [4] have established that blind subjects are able to use the cell phone system to locate targets at a range of distances and that they are capable of orienting the camera properly and moving it smoothly and slowly enough to prevent interference from motion blur. To quantify these results, we conducted two additional experiments to measure the time it takes a blind person to perform simple wayfinding tasks using the cell phone system. The goal of each experiment was to locate and read aloud Braille signs bearing a single letter chosen at random (A through J) located in a corridor or small conference room (21 ft. by 16 ft.). For comparison we repeated the same experiments without the use of the cell phone system, so that the subjects relied exclusively on their sense of touch. Each experiment was performed by two blind volunteers who were informed of the purpose of the study. While the subjects were familiar with the layout of the corridor and conference room, randomization of the placement and content of the Braille signs ensured that differential effects due to familiarity and learning were minimized.

To guide the subjects towards color targets using the system, we used the same three-pitch audio feedback strategy used in [4]: low, medium or high tones signified the target appearing in the left, center or right part of the camera's field of view, respectively, and silence signified that no target was visible to the system. We used 640×480 camera resolution in the experiments, which allowed the system to process approximately two frames per second. The four-color target (white, blue, red and black) with 12 cm diameter was chosen instead of the three-color target to minimize false positives, and the experiments were sited in well-lit locations to avoid false negatives.

For each trial in the experiments, two different Braille signs were placed randomly at shoulder height in two different places. In the corridor experiment, the Braille signs were placed at seven possible locations, covering existing Braille signs (designating rooms and office numbers) running the length of the corridor. For the conference room experiment, two different Braille letter signs were placed randomly at shoulder height, with the two signs on different walls (also randomly chosen). The tasks were performed with and without the cell phone system on alternate trials, with a total of four trials per experiment for each subject. A brief training period was conducted before the experiments to familiarize the subjects with the tasks and possible strategies for completing them efficiently.

We recorded the total time for the subject to find and read aloud both Braille letter signs in each trial. The data for the corridor and conference room experiments are shown in Tables 1 and 2, respectively. Both subjects adopted similar search strategies for searching with the cell phone system: walking slowly down the corridor with the cell phone camera pointed at one wall (and pointing the camera at the other wall on the way back), and panning the camera in a circle in the middle of the conference room to find the directions to the color targets.

Table 1. Corridor experiment. Each cell contains time to find and read signs (in sec.) for two trials.

	Subject 1	Subject 2
Cellphone	164, 145	281, 84
No cellphone	60, 109	50, 38

Table 2. Conference room experiment. Each cell contains time to find and read signs (in sec.) for two trials.

	Subject 1	Subject 2
Cellphone	61, 65	68, 67
No cellphone	48, 37	47, 29

The results show that it was significantly faster for the subject to find Braille signs without the color target detector by rapidly sweeping one or both hands along the wall at shoulder height. This is not surprising because the experimental conditions were very favorable for finding the Braille signs by touch alone: the subjects were familiar with the layout of both the corridor and the conference room; they knew the height of the signs; and they knew they could safely touch all of the walls without any risk of encountering sharp or abrasive objects, which allowed them to move their hands quickly as they walked along walls. Indeed, when asked for feedback the subjects stated that they would have moved more slowly in an unfamiliar environment, and they speculated that some users would be reluctant to use this strategy (e.g. to avoid getting their hands dirty). In addition, the conference room was small enough (21 ft. by 16 ft.) that the subjects could quickly traverse the entire perimeter to find the signs by touch; this advantage would vanish in a sufficiently large room. Finally, the subjects have extensive training in orientation and mobility and were very experienced in searching for Braille, whereas they only had a brief period to practice with the cell phone system.

Even if the comparison between the two methods is unfair in some respects, it provides valuable insight into how blind subjects actually use the cell phone system, and suggests the most important areas for improvement. To operate the cell phone system, the subjects had to walk slowly and/or pan the camera slowly to avoid motion blur – so that they covered areas more slowly than when they walked to search for signs by touch. Moreover, they were instructed to hold the camera as level as possible (since the color target detection algorithm has limited rotation invariance), but the inability to maintain a level orientation – which is difficult to estimate for a blind or visually impaired person – was the most common cause of problems: if the camera was sufficiently off the horizontal, the subject could walk by a color target without detecting it. Without any evidence to signal a target, the subject would have to complete another entire circuit around the corridor in order to have another chance of finding it! This scenario explains an outlier in our data, the inordinately long time of 281 sec. that it took Subject 2 to complete the first cell phone trial (Table 1).

We plan to address the problems of motion blur and limited rotation invariance in the future. Barring future improvements in cell phone camera technology (i.e. faster exposure times), we could make the system less sensitive to motion blur – and thereby allow the user to walk or pan more quickly – by using larger color targets. This would permit the separation between the probe pixels to be increased, which would allow greater amounts of motion blur without sacrificing long-distance detection. A more aesthetically acceptable alternative may be to use the same size targets as before but to adopt a multi-scale detection approach: a greater probe pixel separation could be used in a first stage of detection for each image, and if nothing is detected in the image a second stage could be executed with a narrower separation (to detect targets at a greater distance).

One way to improve the system's range of rotation invariance is to use a target with three colors rather than four, since under ideal conditions the former is invariant to orientation deviations of ±60°, while the range for the latter is ±45°. However, we have found that using three colors creates more false positives than using four colors. Another possibility is to use the usual four-color target but to expand the color target search by including multiple orientations of the probe pixels (e.g. over the three orientations 0°, +20° and -20°). It is unclear whether this approach can be implemented without slowing down the detection process too much.

In the future we will conduct experiments to test the performance of our system under more complicated – but typical – real-world conditions, including unfamiliar buildings and locations, in corridors and rooms of different shapes and sizes (including rooms whose perimeter is long enough to be impractical to search by feel), with signs placed at varying heights.

Finally, we stress that the color target approach is intended as a component of an entire cell phone-based system for guiding users to barcodes and reading them aloud. This function will be irreplaceable for the many people with blindness or low vision who cannot read Braille and who would otherwise have no independent access to sign information.

6. Conclusion

We have previously demonstrated a camera cell phone-based wayfinding system that allows a visually impaired user to find and read barcode signs marked with color targets. In this work we develop simple models that allow us to tune a number of algorithmic and color target design parameters to optimize the color target detection and to assess the performance of the system in terms of average time to discovery in typical conditions. Experiments with blind individuals tasked with finding and reading aloud simple Braille signs are presented in order to validate our analytical models.

While the experiments confirm that our cell phone wayfinding system is usable by blind persons, comparisons with a conventional technique for finding Braille signs by touch alone highlight two main performance bottlenecks that we need to address in future work: motion blur, which limits the speed at which the user can translate and pan the camera, and limited rotation invariance, which creates false negatives when the user is unable to hold the camera sufficiently level. We discuss possible ways of overcoming these problems.

The experiments also underscore the need to conduct further tests under more complicated real-world conditions, in which factors such as familiarity with the environment and the likely locations of signs will be much less helpful to the subjects.

7. Acknowledgment

Both authors would like to acknowledge support from NIH grant 1 R21 EY017003-01A1, and JMC also acknowledges support from NIH grant 1R01EY013875 for preliminary work on computer vision-based wayfinding systems.

8. References

[1] Y. Cho and U. Neumann, "Multi-ring color fiducial systems for scalable fiducial tracking augmented reality", *Proc. IEEE VRAIS*, 1998.

[2] D. Claus and A.W. Fitzgibbon, "Reliable fiducial detection in natural scenes", *Proc. ECCV*, 2004.

[3] J. Coughlan and H. Shen. "A Fast Algorithm for Finding Crosswalks using Figure-Ground Segmentation." *2nd Workshop on Applications of Computer Vision* (in conjunction with *ECCV 2006*). Graz, Austria. May 2006.

[4] 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.

[5] J. Coughlan and R. Manduchi, "Color targets: Fiducials to help visually impaired people find their way by camera phone", submitted to *EURASIP Journal on Image and Video Processing – special issue on Image and Video Processing for Disability*, 2007.

[6] 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*, 84(3): 120-131, May 2001.

[7] M. Fiala, "ARTag, A fiducial marker system using digital techniques", *Proc. IEEE CVPR*, San Diego, June 2005.

[8] D.G. Lowe, “Distinctive image features from scale-invariant keypoints”, *International Journal of Computer Vision*, 60:2:91-110, 2004.

[9] L. Naimark and E. Foxlin, “Circular data matrix fiducial system and robust image processing for a wearable vision-inertial self-tracker”, *Proc. ISMAR*, 2002.

[10] D. Scharstein and A. Briggs, “Real-time recognition of self-similar landmarks”, *Image and Vision Computing* 19, 763-72, 2002.

[11] C. Schmid, R. Mohr, and C. Bauckhage, “Evaluation of interest point detectors, *International Journal of Computer Vision*, 37(2):15172, 2000.

[12] G. Sharma and J.J. Trussell, “Digital color imaging”, *IEEE Trans. on Image Processing*, 6(7):901-32, July 1997.

[13] J. Shi and C. Tomasi, “Good features to track”, *Proc. IEEE CVPR*, 1994.

[14] P. Silapachote, J. Weinman, A. Hanson, R. Weiss and M. A. Mattar, “Automatic sign detection and recognition in natural scenes”, *Proc. IEEE Workshop on Computer Vision Applications for the Visually Impaired (CVAVI ’05)*, San Diego, June 2005.

[15] A. State et al., “Superior augmented reality registration by integrating landmark tracking and magnetic tracking”, *Proc, SIGGRAPH*, 429-38, 1996.

[16] M.S. Uddin and T. Shioyama, “Bipolarity- and projective invariant-based zebra-crossing detection for the visually impaired”, *IEEE Workshop on Computer Vision Applications for the Visually Impaired (CVAVI ’05)*, San Diego, June 2005.

[17] P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features”, *Proc. IEEE CVPR*, Kauai, Hawaii, December 2001.