

International Journal on Artificial Intelligence Tools
© World Scientific Publishing Company

FUNCTIONAL ASSESSMENT OF A CAMERA PHONE-BASED WAYFINDING SYSTEM OPERATED BY BLIND AND VISUALLY IMPAIRED USERS

JAMES COUGHLAN

*The Smith-Kettlewell Eye Research Institute
San Francisco, California 94115, U.S.A.
coughlan@ski.org*

ROBERTO MANDUCHI

*Department of Computer Engineering, University of California, Santa Cruz
Santa Cruz, California, 95064
manduchi@soe.ucsc.edu*

Received (Day Month Year)
Revised (Day Month Year)
Accepted (Day Month Year)

We describe a wayfinding system for blind and visually impaired persons that uses a camera phone to determine the user's location with respect to color markers, posted at locations of interest (such as offices), which are automatically detected by the phone. The color marker signs are specially designed to be detected in real time in cluttered environments using computer vision software running on the phone; a novel segmentation algorithm quickly locates the borders of the color marker in each image, which allows the system to calculate how far the marker is from the phone. We present a model of how the user's scanning strategy (i.e. how he/she pans the phone left and right to find color markers) affects the system's ability to detect color markers given the limitations imposed by motion blur, which is always a possibility whenever a camera is in motion. Finally, we describe experiments with our system tested by blind and visually impaired volunteers, demonstrating their ability to reliably use the system to find locations designated by color markers in a variety of indoor and outdoor environments, and elucidating which search strategies were most effective for users.

Keywords: Wayfinding; Assistive technology; Fiducial design.

1. Introduction

The ability to move independently in a new environment is an essential component of any person's active life. Visiting a shopping mall, finding a room in a hotel, negotiating a terminal transfer in an airport, are all activities that require orientation and wayfinding skills. Unfortunately, many individuals are impeded from such basic undertakings due to physical, cognitive, or visual impairments. In particular, these tasks are dauntingly challenging for those who cannot see, and thus cannot make use of the visual information that sighted individuals rely on.

It is estimated that more than 250,000 Americans are totally blind or have only some light perception³⁰. Among the challenges faced by these individuals, independent mobility is a very relevant one, as it affects employment and education as well as activities such as shopping, site visit and travel. Although some blind individuals are quite skilled at negotiating the difficulties involved with using public transportation and self-orientation in difficult environments, about 30 percent of persons with blindness make no independent trips outside the home⁸. Indeed, as stated in¹⁶: “The inability to travel independently and to interact with the wider world is one of the most significant handicaps facing the vision-impaired.”

Orientation (or *wayfinding*), which is the object of this paper, can be defined as the capacity to know and track one’s position with respect to the environment, and to find a route to destination. This is traditionally distinct from *mobility*, which is the ability to move “safely, gracefully and comfortably”⁵, and refers to tasks such as the detection of obstacles or impediments along a chosen path, normally accomplished using a white cane or relying on a guide dog. Whereas sighted persons use visual landmarks and signs in order to orient themselves, a blind person moving in an unfamiliar environment faces a number of hurdles²⁸: accessing spatial information from a distance; obtaining directional cues to distant locations; keeping track of one’s orientation and location; and obtaining positive identification once a location is reached. All of these problems would be alleviated if a sighted companion (a “guide”) were available to answer questions such as “Where am I?”, “Which direction am I facing?”, and “How do I get there?”.

Our work aims to produce a technological solution to wayfinding that provides a similar type of support. The user’s own camera cell phone acts as a guide, extracting spatial information from special signage in the environment. The cell phone identifies signs placed at key locations, decodes any ancillary information, and determines the correct route to destination. Guidance, orientation and location information are provided to the user via synthesized speech, sound, or vibration.

This article describes a key component of this system: the use of specialized color markers that can be detected easily and robustly by a camera cell phone. The design of these color markers was described elsewhere⁹; in this paper, we concentrate on algorithms for accurately detecting a marker and determining its distance. Additionally, we present some initial user studies with blind subjects that show the potential of this modality for wayfinding.

This article is organized as follows. After discussing a variety of wayfinding methods in Sec. 2, we describe the design of our color markers and algorithms for robust detection in Sec. 3. Theoretical models for the detection performance under different scanning strategies are presented in Sec. 4. Indoor and outdoor experiments with blind subjects are reported in Sec. 5. Sec. 6 has the conclusions.

2. Related Work

There are two main ways in which a blind person can navigate with confidence in a possibly complex environment and find his or her way to destination: piloting and path integration²⁵. Piloting means using sensorial information to estimate one's position at any given time, whereas path integration is equivalent to the "dead reckoning" technique of incremental position estimation, used for example by pilots and mariners. Although some blind individuals excel at path integration, and can easily re-trace a path in a large environment, this is not the case for most blind (as well as sighted) persons.

Path integration using inertial sensors² or visual sensors³² has been used extensively in robotics, and a few attempts at using this technology for orientation have been reported^{23,17}. However, the bulk of research in this field has focused on piloting, with very promising results and a number of commercial products already available. For outdoor travelers, GPS, coupled with a suitable Geographical Information System (GIS), represents an invaluable technology. At least three companies (HumanWare, Sendero Group and Freedom Scientific) offer GPS-based navigational systems which have met with a good degree of success³¹. GPS, however, is viable only outdoors. In addition, none of these systems can help the user in tasks such as "Find the entrance door of this building", due to the low spatial resolution of GPS reading and to the lack of such details in available GIS databases.

A different approach, and one that doesn't require a geographical database or map, is based on specific landmarks or beacons placed at key locations, which can be detected by an appropriate sensor carried by the user. Landmarks can be active (light, radio or sound beacons) or passive (reflecting light or radio signals). A person detecting a landmark may be given information about the pointing direction and distance to it. Thus, rather than absolute position as with GPS, the user is made aware of his or her own relative position and attitude with respect to the landmark only. This may be sufficient for a number of navigational tasks, such as finding specific nearby locations. For guidance to destinations that are beyond the "receptive field" of a landmark, a route can be built as a sequence of waypoints. As long as the user is able to reach (or "home in on") each waypoint in the sequence, even complex routes can be easily followed. This is not dissimilar from the way sighted people navigate in complex environments such as airports, by following a sequence of directional signs to destination.

In the following, we draw a taxonomy of technologies that enable wayfinding by piloting and that have been proposed for use by blind individuals.

Light beacons. The best-known beaconing system for the blind is Talking Signs, now a commercial product based on technology developed at the Smith-Kettlewell Eye Research Institute (SKERI), which has been already deployed in several cities. Talking Signs^{4,13} uses a directional beacon of infrared (IR) light, modulated by a speech signal. This can be received and decoded at a distance of several meters by a

specialized hand-held device carried by the user. The speech signal is demodulated from the signal received and presented via a speaker. Since the field of view of the receiver is limited, the user can detect the pointing direction to the beacon by scanning the area until a signal is received.

The technology developed by Talking Lights, LLC, is based on a similar principle, but rather than a specialized IR beacon, it uses ordinary fluorescent light fixtures that are normally already present in the building. Although simpler to install (only the light ballasts need to be changed), it lacks the directionality of reception provided by Talking Signs.

RF beacons. Radio frequency (RF) beacons have been used extensively for location-based services. The simplest way to provide location information via radio signal is by limiting the transmission range of the beacon, and communicating the beacon's location to any user who receives the beacon's signal. Of course, in order to provide full connectivity, a possibly large number of beacons would have to be deployed. Location accuracy can be increased using multilateration, based on the received signal strength indication and on the known location of the transmitters. A number of commercial services (e.g., Skyhook and Navizon) provide location by multilateration based on signals from 802.11 (Wi-Fi) access points and cellular towers. This represents an interesting alternative to GPS (especially where the GPS signal is not accessible), although resolution remains on the order of several meters. A higher spatial resolution is available for indoor networks of Wi-Fi beacons^{34,22} or ultra-wide band (UWB) transmitters³⁵, although accuracy is a function of environmental conditions, including the density of people in the area, which may affect the radio signal.

RFID. The use of RFID technology, originally developed as a substitute for bar codes, has been recently proposed in the context of wayfinding for the blind²¹. RFID embedded in a carpet have also been used for guidance and pathfinding^{6,15}. Passive RFID, which work on the principle of modulated backscattering, are small, inexpensive, and easy to deploy, and may contain several hundreds bits of information. The main limitation of these systems is their reading range and lack of directionality. For example, the range of systems operating in the HF band (13.56 MHz) is typically less than one meter. Commercially available UHF (400-1000 MHz) tags allow for operating distances in excess of 10 meters provided that the power of the "interrogator" is at least 4 Watts, but the range reduces to less than 4 meters at 500 mW. Of course, longer range is available with active tags, but the cost of installation (with the required power connection) is much higher.

Retroreflective digital signs. This type of technology, developed by Gordon Legge's group at the University of Minnesota³⁷, is closely related to our proposed work. Special markers, printed on a retroreflective substrate, are used instead of conventional signs. These markers contain a certain amount of encoded information,

represented with spatial patterns. In order to detect and read out such markers, one uses a special device, containing a camera and a IR illuminator, dubbed “Magic Flashlight”. Light reflected by the target is read by the onboard camera, and the spatial pattern is decoded by a computer attached to the camera.

Visual recognition of urban features. Algorithms for automatic recognition and localization of buildings and other urban features based on visual analysis of mobile imagery are the object of investigation by researchers in academia^{41,39,14} and industry alike (e.g. Microsoft’s Photo2Search¹⁹ and work in Augmented Reality at Nokia⁷). Such systems typically perform a certain amount of local image processing (geometric rectification, feature extraction), followed by a matching phase that may be carried out at a remote server, or on images or feature vectors that are downloaded onto the user’s mobile computer or cell phone.

User interfaces. Communicating directional information effectively and without annoying or tiring the user is critical for the usability of the system. The fact that blind persons often rely on aural cues for orientation precludes the use of regular headphones, but ear-tube earphones²⁹ and bonephones³⁸ are promising alternatives. The most common methods for information display include speech (synthesized or recorded, as in the case of Talking Signs); simple audio (e.g., a “Geiger counter” interface with beeping rate that is proportional to the distance to a landmark¹⁸); spatialized sound, generated so as it appears to come from the direction of the landmark^{29,20}; “haptic point interface”²⁹, a modality by which the user can establish the direction to a landmark by rotating a hand-held device until the sound produced has maximum volume (as with the Talking Signs receiver); and tactual displays such as “tappers”³³.

User studies. Experimental user studies involving an adequate number of visually impaired subjects are critical for the assessment of the different systems and usage modalities. In a published study²⁸, 30 visually impaired subjects tested the Talking Signs system in a urban train terminal, and the time to reach particular locations was recorded. The same system was tested at traffic intersections with 20 visually impaired subjects¹³ (a similar experiment was run with 27 subjects in Japan³⁶). Specific data measuring the safety and precision when crossing the intersection with and without the device was recorded by the experimenters.

Indoor navigation tests were conducted at the University of Minnesota³ using their “Magic Flashlight”³⁷. A mixed set of subjects, including blind, low-vision, and blindfolded sighted volunteers, performed tasks that involved navigating through a building and finding specific locations. The time to destination was measured and compared with direct physical localization (i.e., finding the location by relying on Braille pads on the wall).

The UCSB group developed the Personal Guidance System testbed, comprised of a GPS, compass, and different types of user interface. This system was tested

extensively in outdoor environments^{24,26,29}. Blind subjects (varying in number between 5 and 15) were asked to follow a route defined by a sequence of waypoints, and quantitative functional measurements were taken in terms of the time to destination or travel distance. These experiments showed that efficiency in the route guidance tasks is directly correlated with the choice of interface modality. The analysis of different user interfaces was also the focus of outdoor tests conducted at the Atlanta VA Center³³ with 15 subjects.

3. Color Markers

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 (marker) placed near the barcode or text. This pattern is designed in such a way that detection can be done very quickly. If a color marker 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 image, the color data in a set of probing pixels centered around x is analyzed (see Fig. 1 for the geometry of probing pixels with a 4-color marker). Detection is possible when the probing pixels are each in a distinct color area of the marker. 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 marker or not. A subsequent fast clustering algorithm is used to rule out spurious false detections.

The choice of the colors in the marker 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 marker 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 marker 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 Fig. 1 that in the case of a 4-color marker, the camera can be rotated by up to $\pm 45^\circ$ around its focal axis without impairing detection (because the probing pixels will still be located within the correct color areas).

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

3.1. Marker Design

The marker selected for this system is pie-shaped, with 4 differently colored sectors. This design allows for a very simple detection strategy based on recognition of the color pattern, as briefly explained in the following. The camera in the cell phones continuously acquires images (frames) of the scene. Each frame is analyzed on a pixel-by-pixel basis. For each pixel, a “probe”, that is, a set of nearby pixels in a specific geometric pattern, is considered, as shown in Fig. 1. The color values of the pixels in the probe are analyzed to evaluate the hypothesis that the probe may reveal the marker. If a certain number of contiguous probes pass this color analysis test, then the presence of a marker in the image is declared.

Since this algorithm is relatively invariant to the size of the marker in the image, the marker is detected over a fairly wide range of distances, as shown in Fig. 1. Other factors, however, may impair detection quality. For example, the user may involuntarily hold the cell phone at an angle, so that the marker’s image is not upright. For a large enough rotation angle, the order of the pixels in the probe may not match the sequence of colors in the marker, causing the algorithm to fail. This effect can be kept under control by increasing the number of tests on the probe pixels. A more insidious problem is “motion blur”, which is caused by the motion of the camera during exposure. As discussed in¹², motion blur may indeed reduce the maximum distance for detection. Illumination conditions may also be a concern. Our system uses a regular camera without an illuminator (which would use too much power and therefore severely limit the battery life).

As a consequence, the color of the marker’s image is not a constant, as it depends on the type of illuminant (e.g., fluorescent or incandescence lamp or natural light). Finally, there is always the risk that a color pattern similar to our marker may exist in the environment (e.g., reproduced in a wallpaper), which may deceive the system and generate confusing false alarms. Ideally, the color pattern in the marker should be customized to the environment, in order to avoid such occurrences. We developed a procedure that, starting from a number of images taken from the surroundings where the markers are to be posted, determines the most “distinctive” color pattern for the marker, as the one that minimizes the likelihood of false alarms⁹.

3.2. Detection Algorithm

The marker detection algorithm needs to run in real time on the cell phone’s computer, at a rate of several frames per second. If there is a delay between image

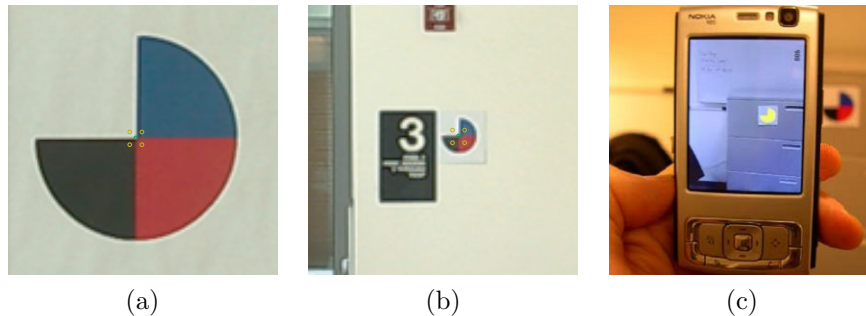
8 *James Coughlan, Roberto Manduchi*

Fig. 1. (a),(b): Our color marker seen at a distance of 1 meter and of 6 meters respectively. The figures shows a 121x121 pixel subimage cropped from the original 640x480 pixel image. Superimposed are the locations of the 4 pixels of a “probe” placed at the center of the marker’s image. Note the probe’s size is the same for the two images, and that correct detection is attained in both cases. (c): The color marker region is segmented (shown in yellow) after detection.

acquisition and output, then interaction with the marked environment (e.g., determining the correct direction to the marker) may become difficult. Image processing is a demanding task, and only in recent years it became practical to program portable, hand-held devices to acquire and analyze images in real time¹¹. For example, the popular Nokia N95 has an ARM-11 processor running at 330 MHz, which is programmable in several languages under the Symbian operating system.

Nevertheless, even such a top-of-the-line model is much less powerful than a regular desktop or laptop computers, and therefore the computational complexity of the algorithms to be implemented needs to be carefully evaluated. In our case, testing the colors of the “probes” centered at each pixel represents the critical bottleneck of our processing pipeline. We have designed an algorithm that minimizes the amount of operations per pixel by performing a cascade of very simple intermediate color tests. The wide majority of non-marker pixels (about 90 percent in our experiments) are promptly discarded after the first test. The remaining pixels undergo up to three additional tests, and the pixels that survive are candidate marker locations. Each intermediate test simply compares one color component of one pixel in the probe with the same color component of another pixel in the probe. We have found that this intermediate test sequence is remarkably resilient to variations in the illumination conditions.

Another functionality provided by our software is the measurement of the distance to the marker. This is obtained by segmenting the color marker region (as shown in Fig. 1 (c)), from which the distance can be estimated knowing the marker’s size and the camera’s parameters (focal length, pixel size). The border of the color marker region is computed using a fast region growing segmentation algorithm that expands a rectangle until it reaches the outer white border, as explained in Sec. 3.3. On a Nokia N95 phone, marker detection and distance estimation is performed at a rate of 3 frames per second (5 frames per second if the distance estimation is turned off).

The user interface for our current implementation is very simple. At any time, the application is in one of two possible modes. In the first mode, marker detection is communicated to the user via a “beep” sound with three possible pitches, depending on whether the marker is located in the central portion of the image or to the left or to the right of it. The purpose is to allow the user to figure out whether he or she is pointing the camera directly to the marker or to the side. In the second mode, the cell phone reads aloud the distance to the marker upon detection using a pre-recorded voice.

3.3. Marker Segmentation

We devised our own novel segmentation algorithm to determine the borders of any color marker that is detected. The segmentation is useful for two purposes: first, the apparent size of the marker in the image can be used to estimate the distance of the marker from the camera; and second, the shape of any marker candidate can be compared to that of the true color marker shape to verify that the candidate is indeed a color marker and not a false positive (this second function is an area of current research).

There is an enormous amount of research in computer vision on segmentation⁴⁰, but very few segmentation algorithms are fast enough for real-time use, especially given the limited processing power of the mobile phone CPU. Our approach was inspired by Seeded Region Growing¹, which incrementally grows a segmented region from one or more seed pixels or regions (which are usually manually selected) in such a way that pixels in the growing region remain homogeneous in terms of intensity, color, etc. In our application, seed pixels from each of the four patches of the marker (the white, blue, red and black regions) are available for each detected marker as a by-product of the marker detection process, which identifies pixels near the center of a marker candidate, and thus also furnishes the associated probe pixels that lie in the four regions of the marker.

Given the seed pixels from the white, blue, red and black regions, we compute the average RGB values for each color region. These values are treated as color *centroids* in a vector quantization scheme that allows us to classify any pixel in the color marker region into the most similar of the four colors. Specifically, given any pixel, the L_1 distance (in RGB space) is computed to the four color centroids, and the closest color centroid is chosen as the classification of that pixel. Given the color input image $I^{(k)}(i, j)$, where $k = 1, 2, 3$ specifies the color channel (red, green or blue) and (i, j) are the row and column pixel coordinates, we will denote the color classification of each pixel as $C(i, j)$, where $C(i, j) = 0, 1, 2, 3$ corresponds to the colors white, blue, red and black, respectively. Fig. 2(b) shows a map of $C(i, j)$ given the original image in Fig. 2(a).

Empirically we find that most of the pixels in a color marker region are correctly classified by this method. The most common misclassification error is confusing colors such as red with blue, but it is important to note that a white pixel is almost

never misclassified as a non-white one (or vice versa). Thus, since the color marker consists of non-white colors bordered by white, we can segment it by incrementally growing a region of non-white pixels beginning with the center of the color marker candidate (this is the center of mass of the cluster of pixels identified by the marker detection process).

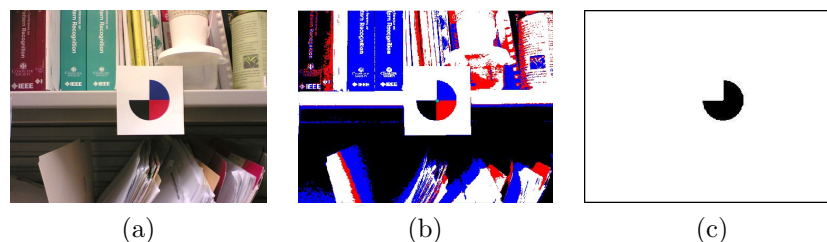


Fig. 2. Segmentation algorithm. (a) Original image. (b) Color classification map $C(i, j)$ shows entire image quantized into the four colors (white, blue, red and black); green dot shows detected center of color marker. (c) Binary segmentation map $S(i, j)$ (image is framed by black rectangle).

Given the color classification map $C(i, j)$, the following region growing segmentation procedure is performed to segment the color marker. First, a binary segmentation map $S(i, j)$, which will specify the color marker region when the algorithm is complete (see Fig. 2(c)), is initialized everywhere to 0. The marker region is initialized to the center pixel of the color marker candidate (i.e. $S(i, j) = 1$ at the center pixel), and working radially outwards, each pixel (i, j) at the “frontier” (border) of the growing region is added to the marker region whenever the pixel value is non-white, i.e. whenever $C(i, j) \neq 0$. The frontier is a rectangular strip of pixels that is specified by four “radii” relative to the center pixel, r_1, r_2, r_3 and r_4 ; the frontier grows whenever one of the radii is incremented by one pixel. Once a frontier is found to contain all white pixels, then the algorithm terminates, since the color marker is surrounded by a white region.

One implementation detail to note is that $C(i, j)$ need only be calculated incrementally as the frontier grows, rather than calculating it all at once over the entire image. Moreover, since the algorithm stops as soon as the white region surrounding the color marker is detected, only a small portion of the image need be evaluated to calculate $C(i, j)$ and $S(i, j)$. Also note that while the Seeded Region Growing algorithm needs to maintain a sorted list of pixels that is updated as pixels are assigned to segmentation regions, our algorithm needs no such procedure because it can exploit the known color and geometry of the color marker region to grow in a “greedier” – and faster – way.

4. Modeling the Scanning Strategy

In previous work⁹ we studied the maximum distance at which a marker can be detected under different illumination conditions. Here, we discuss the expected per-

formance of the system as a function of the user's scanning strategy.

There are at least two possible strategies for scanning a scene with a camera phone while searching for a color marker. The first modality involves rotating the cell phone slowly around its vertical axis (panning). This would be suitable, for example, when the marker 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 marker placed near every office door, signaling the location of a Braille pad or of other information readable by the cellphone.

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 in⁹. 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 in⁹ we introduce in the following two simple models for describing the motion blur effect in the panning and translation strategies¹⁰.

4.1. Scene Scanning Based on Panning

Suppose a person is located at a certain distance from a wall where a color marker is placed. Assume that the user has no idea of where the marker may be (or whether there is a visible marker in the scene at all), so she will scan the scene within a 180 degrees span. Also assume that the person is holding the cell phone at the correct height and elevation, so that, as long as the azimuth angle is correct, the marker 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, in⁹ we empirically found that, with $M = 7$ pixels, the maximum angular speed that ensures detection rate of $p = 0.5$ for our system is $\omega = 60$ degrees/second under average light and $\omega = 30$ degrees/second under dim light for a color marker of 12 cm of diameter at a distance of 2 meters.

Given the frame rate R (in our case, $R \approx 2$ frames/second) and the field of view (FOV) of the camera (in our case, $\text{FOV} \approx 55$ degrees), we can estimate the average number of consecutive frames n in which a particular point in the scene is seen during scanning: $n = \text{FOV}/(\omega/R)$. For example, the color marker would be visible in 3 to 4 consecutive frames with $\omega = 30$ degrees/second, and in at most 2 consecutive frames with $\omega = 60$ degrees/second. The probability that the marker would be detected in one sweep is equal to $p_n = 1 - (1 - p)^n$. The number of sweeps until the marker is detected is a geometric random variable of parameter p_n . Hence, the expected number of sweeps until the marker is detected is equal to $1/p_n$. For example, with $\omega = 60$ degrees/second and $p = 0.5$ (average light), an

average number of at most 2 sweeps would be required for marker detection, while with $\omega = 30$ degrees/second and $p = 0.5$ (dim light) the average number of sweeps before detection is less than or equal to $8/7$.

4.2. 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 marker is always detected. According to⁹ 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 $T = 1/60$ seconds (this will depend on the ambient light), the highest allowable speed of the user when $M = 7$ pixel is $v = 0.36$ meters/second. 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

We conducted a number of experiments using color markers in indoor and outdoor environments with the help of four visually impaired individuals²⁷. Two subjects were fully blind, while the other two had a minimum amount of remaining vision which allowed them to see the marker at very close distances (less than 30 cm). Hence, all subjects could be considered functionally blind for the purposes of our experiments. Although quantitative measurements were taken during the tests (such as the success rate or the time to perform selected tasks), these trials were meant mostly to provide a qualitative initial assessment of the system. More precisely, our broader goals in these experiments were (a) to validate the effectiveness of color markers for labeling specific locations, and (b) to investigate different search strategies for marker detection. Four types of experiments were considered (three indoors and one outdoors) as described below.

5.1. Marker Perpendicular to the Wall

These tests was conducted in an office corridor approximately 35 meters long. Color markers (12 cm in diameters) were printed on white paper, and then attached on both sides of square cardboards (18 cm in side). Using Velcro stickers, marker boards could easily be attached to and detached from the mail slots that line the corridors walls beneath each office door (see Fig. 3 (a)). Three markers were attached at equidistant locations on one wall, and two on the opposite wall. However, at most one such marker was placed upside up, with all remaining markers kept upside down. (Note that the current detection algorithm only detects a marker when it is oriented



Fig. 3. Representative scenes for the three indoor tests. (a): Marker perpendicular to the corridor wall. (b): Marker flush with the corridor wall. (c): Marker in cluttered conference room.

correctly.) The main reason for choosing this strategy rather than just placing at most one marker at a time was that a blind subject may potentially be able to feel the presence of a marker by touching it as he or she passes by. Another reason was that one of the testers had a minimal amount of vision left, which sometimes enabled him to detect the presence of the marker at approximately one meter of distance. However, the subject had basically no color perception, and was certainly unable to establish the marker's orientation.

These markers were detectable at a distance of 5-6 meters in this environment. Since they are printed on both faces of the board, markers are visible from any point of the corridor, except when the subject is exactly at the level of the marker (seeing it from the side). A sequence of 15 runs was devised, such that no marker was present for 6 randomly chosen runs, while one marker was placed at a random location on either side of the wall in the remaining runs. At each run, the subject started from one end of the corridor, and was instructed to walk towards the other end while exploring both walls for the presence of a marker. If a marker was found, the subject was instructed to reach for the handle of the door closest to the marker. The subject was not allowed to change direction (walk back) but could use as much time as he or she wanted for the task. The approximate height of the marker location was fixed and known to the subject. The time elapsed until a door handle was reached or the subject arrived at the other end of the corridor was measured. An initial training phase was conducted for each subject. In particular, subjects were instructed about the correct way to hold the phone and operate the toggle button for distance measurement, and were reminded not to cover the camera with their fingers.

One blind individual (subject A) and one legally blind individual (subject B) took part in testing sessions in different days. Subject A normally uses a guide dog for mobility, although he elected not to use the dog during the test, as he was confident that there would be no obstacles on the way. Subject B used his white cane during the tests. Each session consisted of 15 runs, in 6 of which no marker was posted, while in the remaining ones one marker was posted in random locations

on either wall in the corridor. At each run the subject walked through the corridor from end to end looking for a marker.

No false detection, leading to reaching the wrong door handle, was recorded during the tests. Indeed, save for a few sporadic false alarms (generating an isolated beep), the corridor did not present any patterns that would confuse the system. For the runs where the marker was present, subject A found the marker in all but one run, while subject B found the marker in all runs. Remarkably, both subjects were able to walk at approximately their normal speed while searching for markers. Both of them developed a search technique based on rotating the phone at $\pm 45^\circ$ around the vertical axis with a rhythm synchronized with their gait (one full scan every three steps for subject A and every two steps for subject B). Subject A regularly checked for the distance to the marker as soon as the marker was detected and found this operation to be very important. Subject B stopped using the distance measurement function after a while.

5.2. *Marker Flush with the Wall*

The second test sequence was run on a different corridor, with the same markers attached directly to the wall (see Fig. 3 (b)) and with a similar test sequence. It was found that it is impractical for the user to search for markers on both walls in this configuration. Since the camera can only detect the marker within a certain angle from the normal, the subject would need to rotate the camera over a 180° span around the vertical axis, which is a somewhat awkward operation. Instead, it was decided to let the subject scan only one wall at the time (the subject was instructed about which wall he may expect to find the marker if present). This marker layout required a different strategy, which was quickly learned by both subjects. Rather than scanning the scene by rotating the cell phone, the optimal strategy turned out to be walking closer to the opposite wall, keeping the phone at about 45° with respect to the wall normal. The marker was detectable at a smaller maximum distance than in the previous case, with a higher risk of missing it if the subject walked too fast. Still, no missed detection was recorded out of 10 runs for subject A and 15 runs for subject B. Both subjects were able to walk at their normal speed after a while, with subject B maneuvering a white cane.

5.3. *Marker in Cluttered Conference Room*

The purpose of this experiment was to test the search strategies of a blind subject locating a color marker in a cluttered conference room. In each of 10 trials, a 24 cm. diameter color marker was placed in a random location in a conference room (approximately 7 m by 12 m). A third tester (subject C, who is blind) was told that the marker would appear at approximately shoulder height either flush on a wall or other surface (e.g. podium, shown in Fig. 3 (c)) or mounted on a portable easel (a stand for displaying a large pad of paper) somewhere in the room. At the start of each trial, the experimenters brought the subject to another random location in the

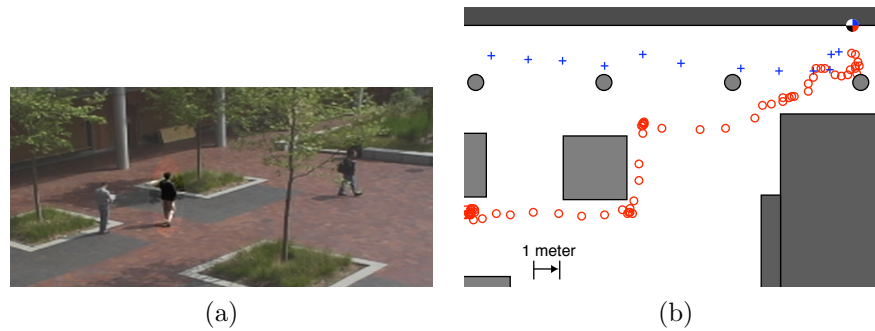


Fig. 4. (a) The E2 courtyard where the outdoor tests took place. Highlighted is subject D during one of the sessions. (b) Blueprint of a portion of the E2 courtyard. Non-accessible grey areas indicate pillars, walls, or grass patches. Two partial trajectories leading to the color marker are shown via blue crosses (subject B) and red circles (subject D). Each mark corresponds to the location of the subject measured every 0.5 seconds.

room, chosen so that the color marker was visible to the camera phone system from this location (i.e. the subject would have to pan the cell phone but wouldn't have to move to a different location to detect it); in most, but not all, trials there was also at least one intervening obstacle between the subject and the marker, such as a table, chair or podium. The experimenters timed how long it took the subject to find and touch the color marker in each trial. After a few practice trials, subject C adopted a fairly consistent search strategy. Holding the cell phone in one hand and the white cane in the other, he began each test trial by panning to locate the color marker, and after detecting it he pressed the button to obtain range information. He then used his white cane to find a clear path to approach the marker by a few meters, and walked rapidly in that direction (often while holding the camera phone down at his side in a position that would not track the marker). He would then locate the marker in his new location before approaching still closer. Sometimes he turned on the range function when he was within a few meters of the marker, especially when he felt he was close to the marker but detected an obstacle with the white cane that prevented him from directly approaching it. While he was successful in finding the marker in all ten trials, his style of alternating between tracking/ranging the marker and walking towards it meant that he lost track of the marker from time to time, which delayed him and forced him to backtrack a few steps to locate it again.

5.4. Outdoor Tests

In order to investigate the utility of this system in more challenging scenarios, we designed a number of outdoor tests, performed in the courtyard of the E2 building that hosts the Computer Engineering Department at UCSC. One or more markers of 24 cm of diameter were placed in different locations on the external walls (typically near a door) and at different heights. The subjects were informed about which one of the two buildings facing the courtyard hosted the sign. The sign was never visible

from the starting location; the subjects had to look for the marker while negotiating any obstacles on the way (such as patches of grass, pillars, small walls, or people walking by). Fig. 4 (a) shows a picture of the environment, while Fig. 4 (b) shows the blueprint of a portion of the courtyard. Three subjects (A, B and D) participated to the experiment at different times. Beside measuring the times to destination, all trials were filmed by a video camera placed on a passageway above the courtyard.

One of the main findings of this experiment was that different subjects may adopt different search strategies with dramatically different results. Subjects A and B walked at a fast pace keeping close to the wall where the marker was expected to be found, and methodically scanned the wall from beginning to the end. This allowed them to find the marker very quickly at each test. Subject D, instead, moved to a location farther away from the wall, spent some time scanning the scene with the phone, then moved to another nearby location and started the process again. Although he was always able to find the marker, it took him substantially more time than for the other subjects. Two partial trajectories for subjects B and D are shown in Fig. 4 (b), where each mark corresponds to the location of the subject measured every 0.5 seconds (as recorded by the video camera). These trajectories highlight the difference in speed and efficiency between the two search strategies.

6. Conclusions

This paper describes a wayfinding system for blind and visually impaired persons using a camera phone to automatically detect color markers posted at locations of interest. The color markers have been designed for maximum detectability in typical building environments, and a novel segmentation algorithm locates the borders of a color marker in an image in real time, allowing us to calculate the distance of the marker from the phone. Experimental results demonstrate that the system is usable by blind and visually impaired persons, who use varied scanning strategies to find locations designated by color markers in different indoor and outdoor environments.

Our future work will focus on adding functionality to the system to provide appropriate directional information for guidance to specific destinations, rather than simply reporting the presence of each marker that is detected. This added functionality will employ a simple map of a building that associates each location of interest with a unique color marker, allowing the system to determine the user's location and provide the corresponding directional information (e.g. "turn left") to any desired destination.

Acknowledgments

Both authors would like to acknowledge support from NIH grant 1 R21 EY017003-01A1. James Coughlan acknowledges support from the National Institute of Health grant 1R01EY013875 for preliminary work on computer vision-based wayfinding systems. Part of Roberto Manduchi's work was supported by the National Science Foundation under Grant No. CNS-0709472.

References

1. R. Adams and L. Bischof. Seeded region growing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(6):641–647, 1994.
2. B. Barshan and H. Durrant-Whyte. An inertial navigation system for a mobile robot. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '93)*, 3:2243–2248 vol.3, 1993.
3. P. J. Beckmann, G. Legge, A. Ogale, B. Tjan, and L. Kramer. Behavioral evaluation of the Digital Sign System (DSS). In *Fall Vision Meeting: Workshop Session on Computer Vision Applications for the Visually Impaired*, Berkeley, CA, 2007.
4. B. Bentzen, W. Crandall, and L. Myers. Wayfinding system for transportation services: Remote infrared audible signage for transit stations, surface transit, and intersections. *Transportation Research Record*, pages 19–26, 1999.
5. B. Blasch, W. Wiener, and R. Welsh. *Foundations of Orientation and Mobility*. AFB Press, 1997. Second Edition.
6. T.-H. Chang, C.-J. Ho, D. C. Hsu, Y.-H. Lee, M.-S. Tsai, M.-C. Wang, and J. Y.-j. Hsu. iCane - A partner for the visually impaired. In *Proceedings of the International Symposium on Ubiquitous Intelligence and Smart Worlds (UISW 2005), Embedded and Ubiquitous Computing Workshops (EUS)*, pages 393–402, 2005.
7. W.-C. Chen, Y. Xiong, J. Gao, N. Gelfand, and R. Grzeszczuk. Efficient extraction of robust image features on mobile devices. In *Proceedings of the 6th IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR 2007)*, 2007.
8. D. Clark-Carter, A. Heyes, and C. Howarth. The efficiency and walking speed of visually impaired people. *Ergonomics*, 29(6):779–89, 1986.
9. 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*, 2007(2):10–10, 2007.
10. J. Coughlan and R. Manduchi. Functional assessment of a camera phone-based wayfinding system operated by blind users. In *Proceedings of the IEEE-BAIS Symposium on Research on Assistive Technology (RAT'07)*, Dayton, OH, 2007.
11. J. Coughlan, R. Manduchi, and H. Shen. Cell phone-based wayfinding for the visually impaired. In *1st International Workshop on Mobile Vision*, Graz, Austria, 2006.
12. J. Coughlan, R. Manduchi, and H. Shen. Computer vision-based terrain sensors for blind wheelchair users. In *10th International Conference on Computers Helping People with Special Needs (ICCHP 06)*, Linz, Austria, July 2006.
13. W. Crandall, B. L. Bentzen, and L. Meyers. Talking signs®: Remote infrared auditory signage for transit, intersections and atms. In *Proceedings of the California State University Northridge Conference on Technology and Disability*, Los Angeles, CA, 1998.
14. R. Duncan and R. Cipolla. An image-based system for urban navigation. In *Proceedings of the British Machine Vision Conference (BMVC '04)*, 2004.
15. N. Fukasawa, H. Matsubara, S. Myojo, and R. Tsuchiya. Guiding passengers in railway stations by ubiquitous computing technologies. In *Proceedings of the IASTED Conference on Human-Computer Interaction*, 2005.
16. R. G. Golledge. Geography and the disabled: A survey with special reference to vision impaired and blind populations. *Transactions of the Institute of British Geographers*, 18(1):63–85, 1993.
17. J. Hesch and S. Roumeliotis. An indoor localization aid for the visually impaired. *2007 IEEE International Conference on Robotics and Automation (ICRA 2007)*, pages 3545–3551, 2007.
18. S. Holland and D. Morse. Audio GPS: Spatial audio in a minimal attention interface.

18 James Coughlan, Roberto Manduchi

- In *Proceedings of Human Computer Interaction with Mobile Devices*, pages 28–33, 2001.
19. M. Jia, X. Fan, X. Xie, M. Li, and W.-Y. Ma. Photo-to-Search: Using camera phones to inquire of the surrounding world. *7th International Conference on Mobile Data Management (MDM 2006)*, pages 46–46, 2006.
 20. R. L. Klatzky, J. R. Marston, N. A. Giudice, R. G. Golledge, and J. M. Loomis. Cognitive load of navigating without vision when guided by virtual sound versus spatial language. *Journal of Experimental Psychology: Applied*, 12(4):223–232, 2006.
 21. V. Kulyukin, C. Gharpure, J. Nicholson, and S. Pavithran. Rfid in robot-assisted indoor navigation for the visually impaired. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2004)*, 2:1979–1984 vol.2, 2004.
 22. A. Ladd, K. Bekris, A. Rudys, D. Wallach, and L. Kavraki. On the feasibility of using wireless ethernet for indoor localization. *IEEE Transactions on Robotics and Automation*, 20(3):555–559, June 2004.
 23. Q. Ladetto and B. Merminod. An Alternative Approach to Vision Techniques - Pedestrian Navigation System based on Digital Magnetic Compass and Gyroscope Integration. In *6th World Multiconference on Systemics, Cybernetics and Information, Orlando, USA, 2002*.
 24. J. M. Loomis, R. G. Golledge, and R. L. Klatzky. Navigation system for the blind: Auditory display modes and guidance. *Presence: Teleoperators and Virtual Environments*, 7:193–203, 1998.
 25. J. M. Loomis, R. G. Golledge, R. L. Klatzky, and J. R. Marston. Assisting wayfinding in visually impaired travelers. In A. G., editor, *Applied Spatial Cognition: From Research to Cognitive Technology*, pages 179–202. Lawrence Erlbaum Assoc., Mahwah, NJ, 2007.
 26. J. M. Loomis, J. R. Marston, R. G. Golledge, and R. L. Klatzky. Personal guidance system for people with visual impairment: a comparison of spatial displays for route guidance. *Journal of Visual Impaired and Blindness*, 99(4):219–32, 2005.
 27. R. Manduchi, J. Coughlan, and V. Ivanchenko. Search strategies of visually impaired persons using a camera phone wayfinding system. In *11th International Conference on Computers Helping People with Special Needs (ICCHP '08)*, Linz, Austria, 2008.
 28. J. Marston. *Towards an accessible city: Empirical measurement and modeling of access to urban opportunities for those with vision impairments, using remote infrared audible signage*. PhD thesis, University of California, Santa Barbara, 2002.
 29. J. R. Marston, J. M. Loomis, R. L. Klatzky, R. G. Golledge, and E. L. Smith. Evaluation of spatial displays for navigation without sight. *ACM Trans. Appl. Percept.*, 3(2):110–124, 2006.
 30. NEI. Statistics on blindness in the model reporting area, 1969-70. Technical Report NIH 73-427, National Eye Institute, 1970.
 31. NFB Access Technology Staff. GPS technology for the blind, a product evaluation. *Braille Monitor*, 49:101–8, 2006.
 32. D. Nister, O. Naroditsky, and J. Bergen. Visual odometry. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '04)*, 01:652–659, 2004.
 33. D. A. Ross and B. B. Blasch. Wearable interfaces for orientation and wayfinding. In *Proceedings of the fourth international ACM conference on Assistive technologies (Assets '00)*, pages 193–200, New York, NY, USA, 2000. ACM.
 34. S. Saha, K. Chaudhuri, D. Sanghi, and P. Bhagwat. Location determination of a mobile device using IEEE 802.11b access point signals. In *Proceedings of the IEEE Conference on Wireless Communications and Networking (WCNC 2003)*, volume 3,

- 2003.
35. P. Steggles and S. Gschwind. The Ubisense smart space platform. In *Advances in Pervasive- Computing, Adjunct Proceedings of the Third International Conference on Pervasive Computing*, volume 191, 2005.
 36. T. Tajima, T. Aotani, K. Kurauchi, and H. Ohkubo. Evaluation of pedestrian information and communication systems-a for visually impaired persons. In *Proc. 17th CSUN conference*, 2002.
 37. B. S. Tjan, P. J. Beckmann, R. Roy, N. Giudice, and G. E. Legge. Digital sign system for indoor wayfinding for the visually impaired. In *Proceedings of the IEEE Workshop on Computer Vision for the Visually Impaired*, page 30, Washington, DC, USA, 2005.
 38. B. N. Walker and J. Lindsay. Navigation performance in a virtual environment with bonephones. In *Proceedings of the International Conference on Auditory Display (ICAD2005)*, pages 260–3, Limerick, Ireland, 2005.
 39. T. Yeh, K. Tollmar, and T. Darrell. Searching the Web with mobile images for location recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '04)*, volume 02, pages 76–81, Los Alamitos, CA, USA, 2004. IEEE Computer Society.
 40. H. Zhang, J. E. Fritts, and S. A. Goldman. Image segmentation evaluation: A survey of unsupervised methods. *Computer Vision and Image Understanding*, 110(2):260–280, 2008.
 41. W. Zhang and J. Kosecka. Image based localization in urban environments. In *Proceedings of the Third International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT'06)*, pages 33–40, Washington, DC, USA, 2006. IEEE Computer Society.