Digital Sign System for Indoor Wayfinding for the Visually Impaired

B.S. Tjan[1], P.J. Beckmann[2], R. Roy[3], N. Giudice[4], and G.E. Legge[3]

[1]3620 South McClintock, SGM 1016, University of Southern California, Los Angeles, CA 90089

[2]Beckmann Research, LLC, 1882 South Ln, Mendota Hgts, MN 55118

[3]75 East River Road, University of Minnesota, Minneapolis, MN 55455

[4]University of California, Santa Barbara CA, 93106

Abstract

Mobility challenges and independent travel are major concerns for blind and visually impaired pedestrians [1][2]. Navigation and wayfinding in unfamiliar indoor environments are particularly challenging because blind pedestrians do not have ready access to building maps, signs and other orienting devices. The development of assistive technologies to aid wayfinding is hampered by the lack of a reliable and cost-efficient method for providing location information in an indoor environment. Here we describe the design and implementation of a digital sign system based on low-cost passive retro-reflective tags printed with specially designed patterns that can be readily detected and identified by a hand-held camera and machine-vision system. Performance of the prototype showed the tag detection/recognition system could cope with the real-world environment of a typical building.

I. INTRODUCTION

Independent mobility is an important prerequisite for full participation in modern society. Reduced mobility and associated social isolation are among the most debilitating consequences of vision loss. Currently, there are about 3.3 million Americans over the age of 40 with impaired vision. Of these, about 937,000 are legally blind and another 2.4 million have milder low vision [3]. Among the legally blind, approximately 200,000 are totally blind, that is, have no useful pattern vision. Because of demographic trends, particularly the aging of the American population, the same study estimates that by 2020, there will be an increase in these numbers by about 70%, bringing us close to 5.7 million, with additional people under 40 years of age not included in these estimates.

The mobility problem for visually impaired pedestrians can be divided into two components: obstacle avoidance and wayfinding. Obstacle avoidance refers to the local problem of taking the next step safely without bumping into things, or tripping. Wayfinding refers to the global problem of planning and following routes from place to place while maintaining a real-time understanding of current position and heading. Many people who are visually disabled deal effectively with obstacle avoidance using a white cane, guide dog, or their residual vision. Compared with obstacle avoidance, much less is known about wayfinding with vision impairment, and there is no technology equivalent to the success of the white cane or guide dog.

Our research focuses on indoor wayfinding for three major reasons:

1) The advent of the GPS system has driven tremendous innovation in the development of navigation instrumentation for outdoor environments. Much less is known about methods for tracking position and orientation indoors.
2) GPS has already been exploited for speech-based navigation for visually-impaired wayfinding outdoors (e.g., Loomis et al. [4][5]), and accessible software for the same purposes is commercially available through SenderoGroup LLC [6].
3) Inability to easily access signage continues to be a major impediment to indoor wayfinding by people with vision disabilities.

There is interest in several technologies for indoor wayfinding for visually impaired people. A partial list of these technologies include: Braille signs, Talking Signs, Talking Lights, RFID tags, dead reckoning (gyroscopic systems, computer-readable pedometers, etc), and systems using Wi-Fi signals. While several of these technologies may share in solving the problem of visually-impaired indoor wayfinding, they all have a major limitation, namely, they are restricted to providing fixed messages about the immediate local environment. Individually, each of these technologies also has its own limitations:

- Braille signs: Often found by the elevator buttons, ATM machines, and on regular signage in public buildings. Braille signs have perhaps the largest installation base. However, many visually impaired people do not read Braille, but even those who do read Braille have trouble locating Braille signs in large spatial layouts.
- Talking Signs [7]: An infrared transmitter encodes a fixed verbal message that is decoded by a hand-held receiver and converted to audio [8]. Talking signs have been deployed in a number of large public spaces, such as the BART system in San Francisco, but are quite costly on a per unit basis.
- Talking Lights [9]: Fluorescent lights are temporally modulated to encode a message. The user’s receiver converts the fixed verbal message to audio. As with talking signs, this technology requires substantial investment in infrastructure.
- RFID Tags: In typical usage, a powered stationary reader decodes “passive” tags on inventory, ID cards etc. in its vicinity using radio frequency. In principle, inexpensive passive tags could be used to label salient points in the environment, and a blind pedestrian could carry an RFID reader to capture information from the tags. Although RFID technology is rapidly developing for commercial

applications, an appropriate configuration for wayfinding application is not yet available due to a lack of portable readers capable of decoding tags at a range of meters (rather than centimeters – this is because a relatively strong RF signal from the reader is needed to power tags) and with directional (as opposed to omni-directional) sensitivity profile.

- Dead Reckoning: Gyroscopic systems, computer-readable pedometers, etc could be used to track a pedestrian through a building. Such devices require periodic recalibration because of accumulating errors.
- Wi-Fi Signal Strength: Technology that measures profiles of signal strength from known 802.11 wireless computer access points can be used to locate a pedestrian within a building [10]. Accuracy of this method is limited by the complex relationship between distance and signal strength in an indoor environment. The signal strength to distance map can also be significantly altered by the installation of objects or furniture with large conductive structures.

A more flexible system would couple an inexpensive method for determining a pedestrian's location and heading with readily accessible information about the building environment, capable of guiding pedestrians along routes, supporting free exploration, and describing points of interest to the pedestrian. This paper focuses on the design and implementation of a location determination system, referred to as the Digital Sign System (DSS). The DSS uses machine vision to detect and identify specially designed tags within a distance of about 3 meters. Each tag contains a numeric code that is associated with location specific data in a spatial database. The DSS serves as a sensor component for a more interactive system – the 'Indoor Guidance System' (IGS). The IGS provides the spatial database containing information about building layout and physical features, and provides an interactive speech-based interface for facilitating wayfinding. This paper briefly discusses the structure of the IGS as well.

II. System Overview and Design Considerations

A. System Overview

The Digital Sign System consists of three components: passive retro-reflective tags, a hand-held sensor module dubbed the "Magic Flashlight", and the machine-vision software that identifies the tags. DSS is a part of a portable/wearable indoor wayfinding system under development, which also includes a building database and a user interface with speech output, jointly referred to as the Indoor Guidance System. Given the complex characteristics of indoor environments, the functionalities of the IGS are indispensable, and set this system apart from other methods of location determination, which lack extended information about the surrounding environment.

The tags are slightly larger than a credit card, designed to be posted as a part of or next to the various indoor signs for room numbers, exits, and public facilities (e.g. elevators, restrooms, public telephones, water fountains, etc.). They can also be used to label light switches, fire alarms, or other items of importance. It is even possible to reserve a range of tag codes for common warning messages such as "floor wet", or "caution: construction" for use in all environments.

A visually-impaired user finds the tags by using a hand-held device called the Magic Flashlight. The device contains an array of infrared (IR) emitters, IR detectors, and a camera. The Magic Flashlight is in the "search" mode when it is switched on. A tone varying in intensity and pitch guides the user to point the Magic Flashlight toward a potential tag. The tone's intensity and pitch reaches a maximum when the potential tag is at the center of the camera's field of view. The user then presses a button to identify the tag.

Each tag codes a 16-bit number, the meaning of which is stored in a database specific to a building. We envision that the manager of a DSS-enabled building would maintain this building database and would distribute it to a DSS user via the Internet or wirelessly each time the user enters the building.

When a tag has been identified, the IGS performs a database lookup and outputs a verbal message based on the user's current preference. The user can vary the level of environmental details provided, ranging from simply reading out what the tag designates (e.g. "Room 612, Dr. Richard Lowe, MD.") to providing a description of the spatial layout at the location of the tag (e.g. "You are on the north hallway of the sixth floor. Rooms 601 thru 616 are along this hallway. You are facing Room 612.").

B. Design Objectives

A key feature of the DSS system is that it requires active involvement of a user. This is similar to other navigation aids such as a white cane or a guide dog. Our objective is to assist the user and not to think for the user. A fully automated navigation aid, even if it were technologically feasible, would require a user to relinquish control and to place a high level of trust in the device, which is exactly the opposite of a typical user's intent to be mobile and independent.

In addition to locating the user's position in a building, it is important to establish the user's current orientation (heading). [cf. 11]. Sweeping the Magic Flashlight to locate a tag enforces such orienting behavior. As a result, a user always knows the direction of the tag that has just been identified.

The requirement for user interaction must be balanced against the consideration of a typical user's physical and cognitive load during navigation. A typical user may need to operate other navigation aids (e.g. a cane) and/or carry other objects (e.g. a purse or briefcase). Therefore, our device needs to be simple to use and of a size easy to be stowed away when not needed. A device that requires continuous user interaction is not desirable, nor is it desirable for a device to produce continuous or complicated auditory output. It can be hazardous for the user if the device becomes a distraction or masks important environmental sounds. In Section III, we will address the issues and techniques for providing concise and intuitive verbal navigational information conducive to the formation of a cognitive map of indoor environments.

Another design consideration is cost. This includes cost to the user and cost to the owner or manager of a building. Cost to a user is best considered in terms of amortized cost over the lifetime of the device. A reasonable lifetime for a device like ours is about five years, after which keeping the computer's operating system and database engine up-to-date without upgrading the hardware will be a challenge. A reasonable per annum cost for the proposed device is probably around $500. Subtracting $100 a year for cost of servicing the device, a lifetime of five years will put the acceptable price tag for the device at around $2000.

With regard to the general acceptance of the system, the cost to a building's facility management is more important than the cost to an individual user. This is the reason why we decided on a system that uses passive tags. The cost for the retro-reflective material and for printing a pattern on the material is very low, less than a few pennies per tag when mass-produced. These tags need no power source and require minimal maintenance. A change of tenant does not require replacing any tag. All that needs to be modified is an entry in the IGS building database, which can be part of the process for updating the building's physical directory. The biggest one-time cost to a building owner is in creating the building database, which represents the spatial layout of the building's interior and associates the tags with their physical locations in the building. Because the interior spatial layout of a building does not change frequently, modification of the building database at the spatial layout level is fortunately rather rare. Nevertheless, the design of the building database can significantly influence the cost for creating a new database for a given building. This database and overall design of the IGS, which also includes a navigation interface for an end-user, is discussed in the next section. A detailed discussion of the DSS follows in the subsequent sections.

III. THE INDOOR GUIDANCE SYSTEM

The Indoor Guidance System works as a platform for integrating multiple wayfinding technologies and providing a uniform user interface to access the information. The IGS is built as a plugin system where each type of user interface or sensing device acts as a single component. The IGS also provides a building database that can be used to link information gathered from the sensors and generate relevant navigation information for the user. In the particular setup mentioned here, the DSS acts as an input plugin to the IGS, supplying a 16-bit integer for a location within a building. This integer is then looked up in the building database and the actual location and surrounding features are retrieved. The IGS also runs a 'List-based Navigation' interface to present the navigational information. The interface is briefly described at the end of this section.

The IGS is composed of three functionally different layers. At the core is a relational database management system (RDBMS) which holds structured digital maps of multiple buildings. Although the RDBMS can be accessed directly, to maintain uniformity and ease of access for all IGS components, there exists a separate request broker – the Integrator layer. The integrator layer provides a well-defined application programming interface (API) tailored towards 2 objectives – 1) to provide an extensive set of functions for retrieving detailed layout data, and 2) to synchronize the state of all input and output components using a message passing system. The topmost layer of the IGS is composed of multiple input/output components. These components are programmed as plugins and can be dynamically loaded or unloaded to support different sensor hardware and user interface modalities. The DSS connects to the IGS in the form of an input plugin. The diagram in Figure 1 presents an overview of the architecture.

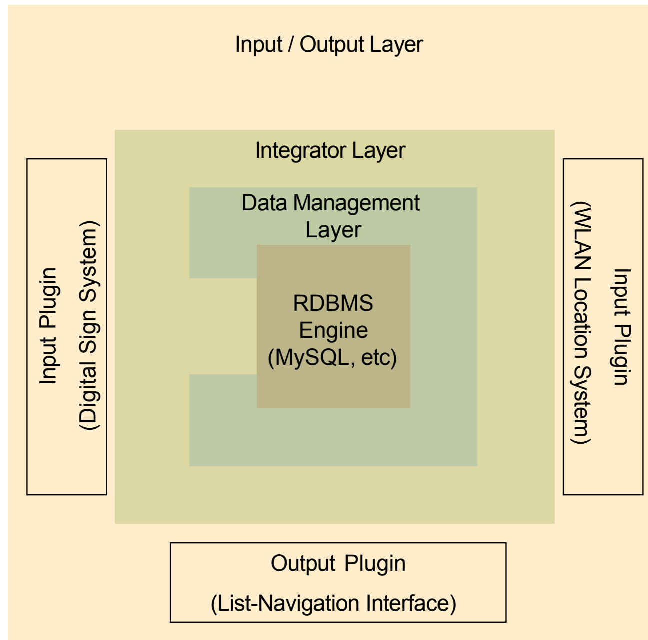

Figure 1. IGS Architecture.

A. The Building Database

The building database stores information about the building layout and physical features within each floor. Multiple buildings can be stored in the database and linked to terrestrial location information to enable seamless integration with outdoor GPS systems.

The database contains two distinct types of polygons defining geometric shapes on one floor. The 'base' layer consists of non-overlapping convex polygons associated with physical building features. An 'overlay' polygon superimposed on the base polygon and logically associated with the base polygon gives meaning to the area. The overlay polygons have detailed semantic information regarding the particular geometric shape. For example, one type of overlay polygon can represent rooms where another can represent hallways. A room can in turn be 'connected' to a hallway through a door overlay polygon which has logical associations with both the room and hallway polygons. This structure enables extraction of both logical associations and physical shape of building features.

B. User Interface – List based Navigation

The user interface for the IGS is targeted towards visually impaired people. Hence, the primary mode of information presentation is synthetic speech. The biggest challenge here is structuring information about building layouts using consistent and unambiguous terminology. As opposed to roads and highways, hallways and lobbies do not have names. The spatial density of information is also much higher than outdoors. A verbal protocol for describing geometric properties of indoor layouts (corridor structure), called the List-Navigation interface, has been developed and investigated in several experiments with blindfolded-sighted and blind participants as described in [12]. The results of these studies demonstrated that when well-crafted verbal descriptions are employed, people are readily able to learn and navigate large-scale unfamiliar environments.

In List-based Navigation, the environment is depicted as a set of 'feature points' that can be navigated to. At each feature point, the user is presented with a list of nearby features. To mentally navigate to a nearby feature, the user can scroll within the list and hit a key to "move" to the selected feature. The user can get more detailed information about the selected feature by hitting a different key. Users are also given the choice of getting egocentric and allocentric descriptions of features around them at any point. An egocentric description provides the distance and direction to a feature from the user's current location and chosen heading. Allocentric descriptions present information with respect to a set of absolute reference directions such as North, South, East and West. For example, an egocentric description could be "In north lobby facing south. Door to Room N119 is 31 feet at approximately 10 o'clock. Entry to east west hallway 1 is 28 feet at approximately 9 o'clock. Start of north south hallway 1 is 45 feet at approximately 11 o'clock". The allocentric description at the same layout could be "In north lobby. East wall: Door to Room N119, entry to east west hallway 1. North wall: North Entrance. South wall: start of north south hallway 1".

When connected to the IGS, a tag ID obtained from the DSS is looked up in the building database and an overlay polygon is located within the floor. The association of this overlay polygon with a base polygon is then used to determine the location of nearby features and thereafter, a description of the surroundings is provided to a user via the List-based interface. The interface also allows virtual navigation of the whole layout through a point-to-point movement process. In the near future, we intend to add on-the-fly routing capabilities that can generate best routes to the desired targets based on the measured current location.

IV. DSS TAG

A DSS tag consists of a specially designed pattern printed on a retro-reflective sheet. Retro-reflective material has a unique property that it reflects light back along that same direction as the incident ray, regardless of the incident angle. In other words, the material reflects light back to its source, and in the case of the retro-reflective material we used (made by 3M), the reflected light is within a 0.25 deg cone centered on the incident ray. This retro-reflection takes place for any incident rays within 60 degs of a tag's surface normal. We used a ring of IR LEDs mounted around the camera lens as the illumination source. With IR illumination turned on, a DSS tag appears as a bright object in the scene. This property of the retro-reflective material is used to facilitate tag segmentation and identification.

Retro-reflective material comes in different colors to match the interior design of a building. Because DSS operates in the IR range, it is also possible to print the code pattern in a color the same as the retro-reflective backing, resulting in a uniform tag without any visible pattern.

Figure 2 shows a DSS tag. A DSS tag is to be posted in a "portrait" orientation such that oblique viewing does not affect the critical spacing of the code elements. Each tag contains four vertically oriented tracks of code elements. The two tracks on the right are rotated copies of the two tracks on the left. This arrangement leads to a tag that is rotationally symmetric and eliminates any error of posting the tag upside-down. It also provides the redundancy for error checking and correction. A phase code is used to encode a 16-bit number with two code tracks. Each binary digit is represented be a light bar and a dark bar. On the two left tracks, a light bar on top of a dark bar represents 0, while a dark bar on top of a light bar represents 1. The left-most track is the most-significant byte, and the second left track is the least-significant byte. With this phase coding scheme, any combination of 1's and 0's produces only four code elements: a bar that is one of two sizes (thin or thick) and one of two intensities (light or dark). This property greatly improves the robustness of the tag identification algorithm. The coding scheme also has the property that the average intensity of a tag (the total light-area to dark-area ratio) is a constant regardless of the numeric ID it encodes, which is advantageous for tag detection and segmentation. Tag segmentation is further assisted by having a dark-and-light segmentation border surrounding the code region, setting it apart from the background.

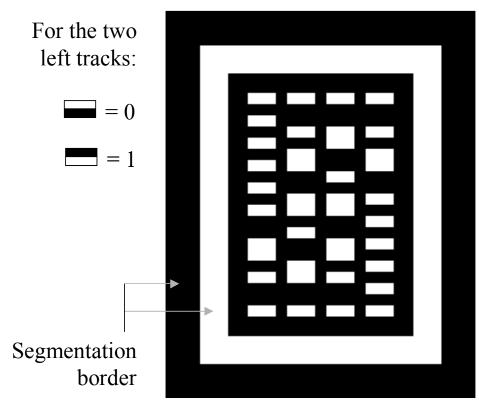

Figure 2. A DSS tag coding for 1234. See text for the coding scheme.

V. THE MAGIC FLASHLIGHT

The Magic Flashlight (MF) is an illumination and sensor module held by a user to search for and identify DSS tags within a range of approximately 3 meters, which is the typical distance between doors in a building. It consists of IR illuminators, IR phototransistors, and a black-and-white camera. The MF has two modes of operation: search and identification. Most of the time, the flashlight is in the search mode. Analog detection circuitry controls the frequency and amplitude of a tone, providing real-time auditory feedback to guide a user to aim the MF toward a potential tag. When the tone reaches maximum amplitude and frequency, a DSS tag is likely close to the center of the camera's field of view. A user then presses a button to acquire the tag, and a pair of images will be taken and processed by a machine-vision algorithm to retrieve the 16-bit tag ID. Figure 3 shows a picture of the current prototype.

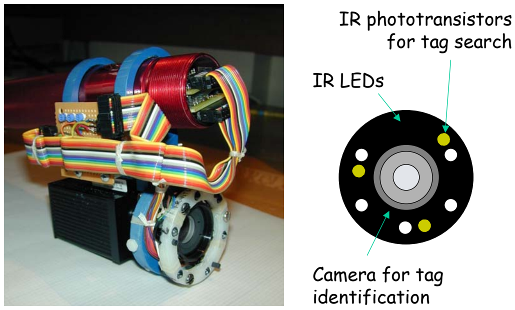

Figure 3. A prototype of the “Magic Flashlight” and an illustration of its sensor ring. The circuit boards implement the analog circuitry for the search mode and synchronization of LEDs with image acquisition.

A. The Search Mode

The search mode is supported by three IR LEDs and three associated IR phototransistors. Using “lock-in amplifier” techniques, the LEDs are strobed on and off at a consistent rate, roughly 470Hz. This frequency was chosen to avoid frequencies commonly found in built environments (e.g., 60 Hz) and their harmonics. A synchronous detection and amplification circuit within the flashlight is sensitive only to light returning from the environment at the same rate used to strobe the LEDs. This technique prevents the MF from incorrectly identifying an environmental source of IR as a DSS tag. The retro-reflective characteristic of the tag allows the sensitivity of the detector to be set low enough to avoid detecting diffuse surfaces incorrectly as DSS tags since the return from the tag is so much stronger than that from a diffuse surface. Figure 4 shows a block diagram of this search mode system.

Bruggeman et al. [13] found that when blindfolded subjects searched for a target using a flashlight with four beam widths, subtending 0.25 to 73 degs of visual angle, a 35 degree cone was optimal for both finding the target on a wall during search and for keeping the target localized within the beam as they approached and touched the target. This empirical finding was used to specify the field of view of the IR detectors for the search mode to be 35 degs. Due to several engineering constraints, the current beam width of the IR LEDs and the acceptance angle of the phototransistors support a 15 degs wide cone of sensitivity, still within the empirical range where good performances were obtained.

The illumination level and amplifier gains are set for a detectable tone shift at a maximum distance of 2.5 meters. This is adequate for the situation in which DSS tags are mounted on or near to existing room signage mounted on a wall near the room entrance. If additional range is desired, for example, the placement of a tag at the end of a hallway, the size of the DSS tag may be scaled up to provide a return signal strong enough to be detected at that distance.

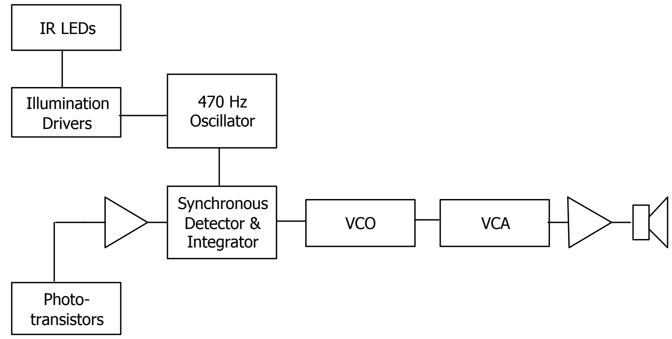

Figure 4. Block diagram for the analog circuitry that implements the search mode.

B. Tag Identification

The Magic Flashlight carries a black-and-white camera with sufficient sensitivity to near-IR (most CCD cameras have reasonably high near-IR sensitivity once a IR-blocking filter at the lens is removed). An illumination strobe consisting of three IR LEDs is mounted on a ring around the camera's lens. (This ring, incidentally, also carries the LEDs and phototransistors for the search mode.) The three IR illuminators for the camera have their optical axes parallel to the optical axis of the camera so that light reflected from the retro-reflective tag will be seen by the camera.

The tag identification process consists of a segmentation step followed by a decoding step. Tag segmentation begins with a pair of gray-level images (640x480) taken in successive frames (frame rate = 30 fps). The IR illuminators are turned on during the first frame, and off during the second. The output of tag segmentation is a list of the regions in the image that contain the tag proper. The following steps are taken to ensure robustness:

1) Co-register the IR-off image (Figure 5b) with the IR-on image (Figure 5a) by computing the cross-correlation in Fourier domain of the Sobel edge maps of the two images. Mis-registrations due to hand-jitter within 30ms are mostly translations.
2) Subtract the co-registered IR-off image from IR-on image and remove pixels with intensity lower than the 15-th percentile from the subtracted image (Figure 5c)
3) Use median filter to remove noise from the subtracted image. Divide the subtracted image into 6x6 blocks. Correct for illumination inhomogeneity by fitting a plane thru the pixels with the maximum intensity within each block. Apply the correction to the subtracted image (Figure 5d).
4) Threshold the normalized image at 30% of the maximum intensity. Remove isolated pixels and H-connected pixels from the thresholded image. Use asymmetric closing (dilate, erode more, and dilate less) to allow the light pixels of a tag to grow into contiguous regions. Fill in any hole. The result is a set of unconnected regions. (Figure 5e)
5) Designate a region from Step 4 as a “tag stencil” if its size is greater than 1/60 of the total image size

and with a figure complexity ($perimeter^2$/area) less than 25. This step rejects any small and highly elongated regions. (Figure 5f)

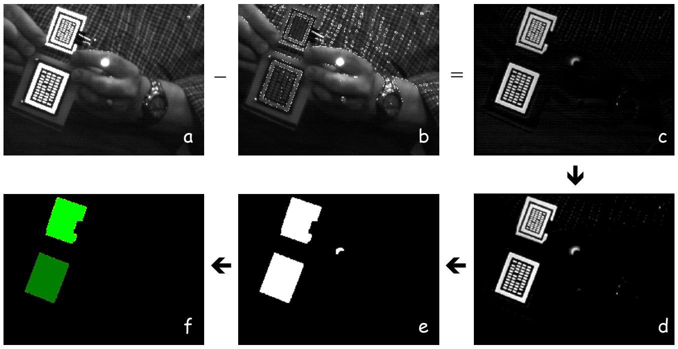

Figure 5. Tag segmentation: (a) Image acquired with IR LEDs turned on. (b) Image acquired with IR LEDs off. (c) Difference between IR-on image and the co-registered IR-off image. (d) The difference image with illumination in-homogeneity corrected. (e) Unconnected regions found by asymmetric closing and hole-removing. (f) Tag stencils – output of the segmentation process, which marks the potential tag regions.

The decoding step begins with the list of tag stencils from the segmentation step and applies them to the unprocessed IR-on image. The stencils are ordered by their size. The tag with the largest area, and presumably closest to the user, is read first. The outputs of this stage are a tag code, a quality assignment of the code (0.0-1.0, 1.0 being the best), and an estimate of the distance and orientation of the tag. Tag decoding proceeds as follows:

1) Extract the image region from the raw IR-on image defined by a tag stencil.

2) Apply Harris corner detector [14] to find the four inner corners of the tag's code region (Figure 6a,d).

3) Morph the code region into a standard template using projective mapping [15] (Figure 6b,e). The parameters of the projective mapping, along with the fixed dimension of a DSS tag, are used to estimate the orientation and distance of the tag.

4) Extract the gray-level images of the four code tracks according to the template. Two additional code tracks are synthesized by averaging the intensity values between the left track pair and the rotated right pair, which should contain the same code. Within each track (synthetic or actual), sum the image intensity horizontally. (Figure 6c,f)

5) Threshold the horizontally summed track intensities by their mean. Construct a "pre-code" by classifying each code element as either light or dark and either narrow or wide. Determine the quality of the readout by comparing the pre-codes between the three pairs of code tracks. Determine, based on quality, which pair of code track should be selected to give the output. Provide the tag ID output by converting the "pre-code" from the selected pair of code tracks to binary according to the coding scheme of DSS tags. Also report readout quality and the tag's distance orientation and distance.

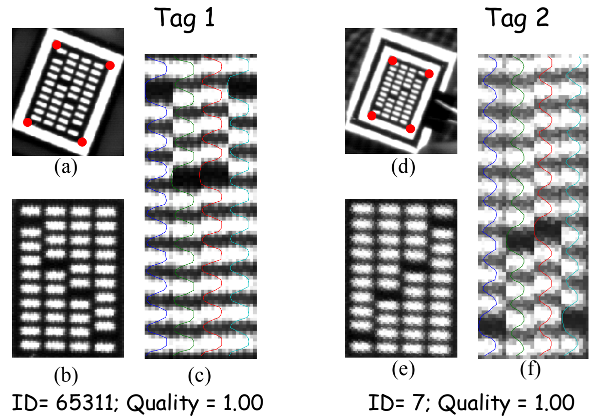

Figure 6. Tag decoding. Tag 1 & 2 correspond to the two tag regions from Figure 3. (a, d) Interior corners of a code region within a tag stencil. (b, d) Normalized code region. (c, f) Extracted code tracks and their horizontally summed intensities.

VI. Discussion

The tag identification subsystem, which runs the tag segmentation and decoding algorithms, has been successfully tested in a wide range of conditions. These included scenes with multiple tags, tags of different sizes, viewed at oblique angles with uneven illumination, acquired with large hand-jitter, etc (Figure 7). At an earlier development phase (before we committed to the current tag design), successful testing of tag segmentation was achieved in challenging indoor environments (Figure 8).

A condition that is difficult for the tag identification subsystem to cope with occurs when a tag is brightly illuminated by direct sunlight when posted opposite a window. In this case, the intensity from the IR illuminators is too low to be discriminated from the bright background. However, we have not encountered a condition where the same failure happened with artificial lights.

The code for the tag identification subsystem is written in Matlab and has not been optimized for real-time operation. On a 650MHz Pentium III laptop, it takes 20 seconds to acquire and process an image pair containing two tags, with 16s for segmentation and 4s for decoding. A factor of 10 to 20 speedup is possible by a combination of using lower-resolution images, optimized and natively complied code, and faster hardware.

The analog circuitry that implements the search mode of the Magic Flashlight operates in real-time. However, its robustness has yet to be fully evaluated. Our initial findings suggested that because the input to the system is a single value of the reflected intensity, it appears difficult to adjust the circuit to exclude other retro-reflective structures in the environment (e.g. inner corners between walls), even if they are very different in shape from a DSS tag. This shortcoming is being addressed by incorporating automatic gain control

(AGC) circuitry in the sensor amplifier stages of the Magic Flashlight. We are simultaneously experimenting with a digital solution that uses a low-resolution video stream acquired from the camera on the Magic Flashlight to perform this search function in real-time. Future evaluations regarding the search mode subsystem will investigate the effects of the tone's baseline and modulation in frequency and amplitude on a user's search performance.

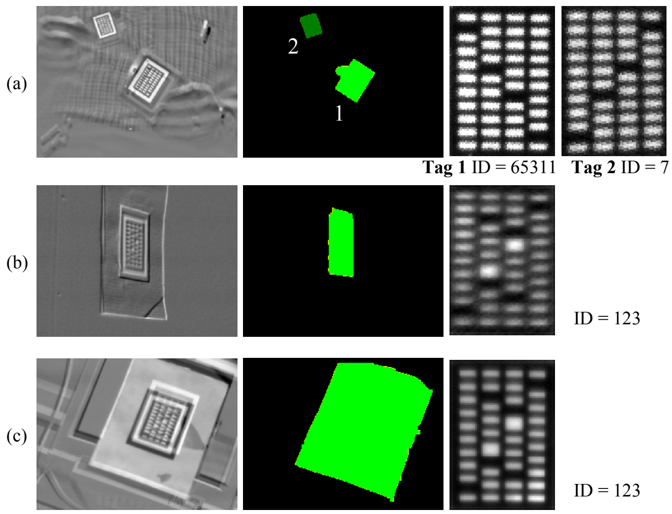

Figure 7. Tag identification using tags of multiple sizes (a), from an oblique viewpoint (b), and with uneven illumination and large hand jitter (c). The left column shows the difference between IR-on and IR-off images without co-registration to reveal the extent of hand jitter. The middle column shows the result of tag segmentation. The right column shows the normal code region extracted from the images. In all three examples, the tags were correctly identified with quality of 1.0.

VII. SUMMARY

We have built a prototype of a hand-held system, called the Digital Sign System (or DSS), that can read specially designed signs printed on cheap retro-reflective material. We have demonstrated the robustness of the tag segmentation and identification algorithm running on a low-power laptop. This system provides location and orientation information to a visually impaired user in an indoor environment. A database containing the layout information of a building and a cognitively efficient method for communicating with the user are provided by the Indoor Guidance System (IGS). Since no semantic information is stored with the sign itself, the potential domain of use of the IGS/DSS system is quite vast. The level of interactivity provided by the system is also quite unique for indoor applications, rivaled only by outdoor GPS based navigation devices.

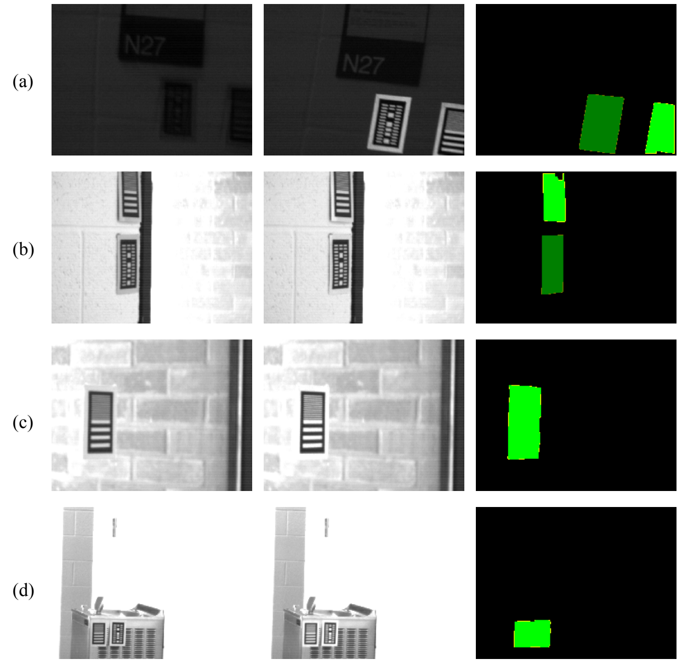

Figure 8. Potentially challenging indoor environment for DSS. From left to right: raw images with IR off, raw images with IR on, results of tag segmentation (tag stencils). (a) A typical installation in an interior hallway along with other signage. (b) Installation in hallway with large windows and natural light. (c) A tag posted against a window. (d) A pair of tags posted next to each other on the side of a water fountain. Tag segmentation was accurate except for the water-fountain condition, when the system mistakenly merged the two tags into one.

VIII. ACKNOWLEDGMENTS

Supported by NIH Grant EY02857 and NIDRR grant H133A011903.

IX. REFERENCES

[1] Golledge, R. G. (1993). Geography and the disabled: A survey with special reference to vision impaired and blind populations. *Transactions of the Institute of British Geographers*, 18, 63-85.

[2] Marston, J. R., & Golledge, R. G. (2003). The hidden demand for participation in activities and travel by persons who are visually impaired. *Journal of Visual Impairment & Blindness*, 97(8).

[3] The Eye Diseases Prevalence Research Group. (2004). Causes and prevalence of visual impairment among adults in the United States. *Archives of Ophthalmology*, 122, 477-485.

[4] Loomis, J. M., Golledge, R. G., Klatzky, R. L., Speigle, J. M., & Tietz, J. (1994) Personal guidance system for the visually impaired. *Proceedings of the First Annual ACM/SIGGAPH Conference on Assistive Technologies*, Marina Del Ray, CA, October 31-November 1, 1994, pp. 85-91. New York: ACM.

[5] Loomis, J. M., Golledge, R. G., & Klatzky, R. L. (1998). Navigation system for the blind: Auditory display modes and guidance. *Presence: Teleoperators and Virtual Environments*, 7, 193-203.

[6] Davis, CA www.SenderoGroup.com

[7] Talking Signs, Inc., California

[8] Crandall W, Brabyn J, Bentzen BL, Myers L. (1999) Remote infrared signage evaluation for transit stations and intersections. *J Rehabil Res Dev.*36(4):341-55.

[9] Talking Lights LLC, www.talking-lights.com

[10] Commercial software EPE 2.1 is available from Ekahau, www.ekahau.com

[11] Rieser, J.J. (1989) Access to knowledge of spatial structure at novel points of observation. *Journal of Experimental Psychology* 15:1157-1165.

[12] Giudice, N., Navigating Novel Environments: A Comparison of Verbal and Visual Learning. Unpublished doctoral dissertation, University of Minnesota, Minneapolis, MN

[13] Bruggeman, H., Giudice, N., Stankiewicz, B.J. and Legge, G.E. (2000) Distal target localization by the blind. *Annual Meeting of the Association for Research in Vision and Ophthalmology*.

[14] Harris, C.G. and Stephens, M.J. (1988) A combined corner and edge detector, *Proceedings Fourth Alvey Vision Conference*, Manchester. pp 147-151, 1988.

[15] Heckbert, P. (1989) *Fundamentals of Texture Mapping and Image Warping* Master's thesis, UCB/CSD 89/516, CS Division, U.C. Berkeley.