2016 IEEE 8th International Conference on Intelligent Systems

Computer vision guidance system for indoor navigation of visually impaired people

Kabalan Chaccour
TICKET Lab, Antonine University
Hadat-Baabda, Lebanon
P.O. Box 40016
Email: kabalan.chaccour@ua.edu.lb

Georges Badr
TICKET Lab, Antonine University
Hadat-Baabda, Lebanon
P.O. Box 40016
Email: georges.badr@ua.edu.lb

***Abstract*—Visually Impaired (VI) and blind people suffer from reduced mobility, as they cannot detect the terrain and their environment. They always need assistance and walking support systems in their daily life. Solutions have been proposed many decades ago and are rapidly improving nowadays due to the technology evolution and integration. A large number of assistance aids have been deployed in real life situations whereas other concepts remained as research ideas. This paper describes a new approach of an ambient navigation system that would help the visually impaired or blind person to move freely indoor (house, office, etc.) without the assistance of anyone. The system is composed of IP cameras attached to the ceiling of each room and the smart phone of the subject is used as human machine interface (HMI). Frames are sent to a computer that analyzes the environment, detects and recognizes objects. A computer vision guidance algorithm is designed to help the user reach his destination (or his personal item) with obstacle detection. The system is commanded by voice messages via a simple mobile application. Feedbacks (alerts, route) are voice messages returns by the application to the user. This system provides a reliable solution to assist those users in their indoor navigation providing them a correct route with obstacle avoidance.**

***Index Terms*—Visually impaired, computer vision, indoor navigation, mobile application, image processing.**

I. Introduction

According to the World Health Organization (WHO), 285 million people are estimated to be visually impaired worldwide, 39 million are blind and 246 have low vision. Whether it happened accidently or by a disease, this physical impairment has major impacts on daily life activities. As a matter of fact, motion is noticeably refrained. Moreover, visually impaired people may lose orientation and have a higher risk of falling. But, people need to move, whether at home, at work or leisure. The quality of life highly depends on the ability to move, to calculate itineraries, to take decisions and avoid obstacles [3]. For those people, losing sight is like losing independence. Most of visually impaired subjects depend on other humans or pets for movement and environmental perception. They also have a long stick, called the white cane, introduced after the World War I [8], or they follow the walls of houses or shops, memorizing all locations and details. Even in their own home, blind people must exhibit efforts and cognitive burden to navigate from one place to another, avoiding obstacles even in the same room particularly when the subject does not know his spatial location relatively to the room. The cane could help him avoid the obstacle, but how could he know his direction especially when he is away from a wall? These questions any many others need answers from scientists and researchers who strive to provide technical solutions that can answer partially or wholly the problem of mobility of visually impaired and blind subjects. Henceforth, an assistance system is therefore required so that the person could be aware of his position and in which direction he should commence his relocation in order to reach his final destination.

In this context, the current research proposes a novel solution for indoor navigation of visually impaired and blind people. To our knowledge, the proposed approach is a new idea that has not been yet exploited. Our system architecture is simple and easy to operate. It doesn't require extra skills or additional hardware. It is simply composed of a voice commanded mobile application that interacts with a remote processing system where photos from IP cameras located in each room are captured and analyzed using computer vision algorithms. The system can identify the orientation of the blind person through a marker with an imprinted pattern on his head. The person can command the system which will guide him by voice messages to his final destination while avoiding obstacles and walls.

The rest of the paper is organized as follows. Section II describes the related work of most existing solutions. It gives a general overview on successful products and research techniques that have already been deployed on the market. Section III illustrates our proposed system approach to the problem of reduced mobility in visually impaired subjects. The section exposes the system architecture and its main components. Section IV carries the implementation of each of the architecture components. It focuses on the development of the mobile application and the remote processing system. The computer vision guidance algorithm is also detailed in this section. Experiments is carried out in section V where some drawbacks of the system are shown. We conclude our paper in section VI.

978-1-5090-1354-8/16/$31.00 ©2016 IEEE

II. Related Work

Designing assists and aids for visual impaired persons is a challenging and ambitious research area. Engineers and designers provide systems in order to allow blind people navigate freely and independently. Before technology, blind people relied on the help of guide dogs that require extensive training and high cost. Moreover, the white cane could not help them navigate neither indoor nor outdoor because of the lack of details provided. The problem of indoor navigation remains largely unsolved. Critical aspects should be always taken into consideration: sensing the environment and informing the blind user [2]. Detecting the environment can be done with several sensors (ultrasonic, infrared, pressure, etc.), or cameras. Audio messages or tactile senses are used to inform the user.

Echolocation [16] is a system that was designed in the early 1990s and uses two ultrasonic sensors attached to eyeglasses. Data are sent to a microprocessor and converted to audio messages via headphones to the subject. It used different ultrasonic frequencies to discriminate different objects in front of the person. The simplicity and the portability of this device make the product very advantageous; however it presents a lack in guiding the user due to many blind zones. In 1992, Navbelt [15] was developed as a guiding system. It used 8 ultrasonic sensors with an obstacle avoidance mobile robot, a computer and earphones to generate a map of angles and the distance to each object. This system was able to guide the user to his destination using sound beeps. This system required extensive training for the user in order to understand the audio feedback (beeps). Guide-cane [17] is the updated version of NavBelt. It uses a cane with wheels, some ultrasonic sensors and a computer. The system notifies the user of encountered obstacles by audio feedback. It also offers an alternate direction algorithm to bypass the obstacle. Such system has a limited scanning area and heavy to carry. CyARM [5] alerts the user of the presence of obstacles by calculating the distance with ultrasonic sensors and provides a tension signals via a wire attached to the user. High tension indicates close distance, while a lower tension indicates longer distance. A similar sensing device was presented by UCSC project [18]. The computer of the system deals with 1-D of data retrieved by ultrasonic sensors. The data is analyzed by the computer to detect environmental features that are critical for mobility, such as curbs, steps, and drop-offs by means of an extended Kalman filter tracker. Many commercial products were developed for this goal. They consist mainly of a cane with embedded range scanners (laser or sonar). Obstacle detection is translated into vibrations or sound feedback; however these systems are very expensive and not effectively reliable.

Other types of walking aids use cameras to scan the environment. Cameras are widely used in environment sensing because they can provide rich information concerning the environment where they are deployed. Thus, cameras have gained close attention in designing electronic mobility aid systems. For example, NAVI [14] (Navigation Assistance for Visually Impaired) uses a camera to classify image taken into background or objects in order to prevent the user from hitting obstacles. The system in [13] uses a color histogram to detect obstacles and alert the user via his smart phone. Other systems use stereo cameras in order to sense the environment to compute depth from the pair of images; however the processes are more complicated. This technology was used in TVS (Tactile Vision System) [6] and Tyflos [1] systems that convert depth into vibrations on a device attached to the user's abdomen. In VOICE [11], the system retrieves environment information by a digital camera attached to eyeglasses. A portable computer performs an image to sound mapping and sends the audio feedback directly to user's ear without any filter. The University of Stuttgart has developed a wearable assistance system for blind user for indoor orientation [4]. The system consisted of a portable computer, a cane to which is attached a module containing two cameras, a keyboard, a digital compass, a 3D inclinometer and a speaker. The sensors scan the environment and the computer sends an audio feedback to the speaker. The SmartVision prototype [7] provides global navigation for guiding the user to some destination. The navigation is done indoor or outdoor, with avoidance of static and moving obstacles. The user still uses his cane and wears a stereo camera at chest level, a portable computer and earphones.

Other navigation systems such as ENVS (Electron-Neural Vision System) [10] are used for indoor and outdoor. The system has two stereo cameras, a digital compass, a portable computer and a GPS. The feedback is an electro-tactile simulation sent to a specific glove. The tactile feedback is also used in University of Guelph [12]. The prototype detects the presence of obstacles and sends the direction as a vibration to a different finger of the tactile glove. Tactile handler system [9] is a device dedicated to help visually impaired people navigate in both familiar and non-familiar environments. The prototype is a glove with embedded microcon- troller, tactile array (on each finger) and 4 sonar sensors. The intensity of the vibration represents the distance to the obstacle. Training the user is a necessary part and requires the user to use only one hand while the second one is kept for scanning.

In this paper, an indoor guidance system is proposed. The contribution of this paper can be divided into different axes. First, we are presenting a non-wearable assistance aid. The blind person needs to carry his mobile phone only. Second, the cameras are attached to the ceiling of the rooms so that the entire surface is covered. The user wears only a marker with an imprinted pattern in order to be detected. His position and orientation are therefore known for processing. The system is voice commanded and the guidance is done using a GPS like navigation which doesn't require extra skills for operation.

III. Proposed System Approach

The main goal of our proposed system approach is to provide a new concept in safe indoor navigation with minimum system complexity. The system must be reliable, easy to operate and implemented with minimum cost. The challenge of our model is to deliver high precision indoor navigation with obstacle avoidance capability.

The proposed system approach has a simple architecture. The two principal actors in this architecture are the visually impaired person and the remote processing system. The concept of our indoor navigation system is distributed between the mobile application and the remote processing computer as shown in the figure below. The subject needs to carry his smart phone during his displacement without any other additional hardware. The mobile application will react to the user voice commands by sending text messages to the remote processing system for analysis. These messages are transmitted through WIFI or Bluetooth communication to the remote processing system. The latter is based on image analytic algorithms applied on photos taken from different cameras installed in the environment. These photos are scaled and processed to compute the orientation of the visually impaired and guide him to his destination safely. As a matter of fact, the user commands his mobile of the destination he wishes to reach and the system responds with audible navigational directions on the mobile application.

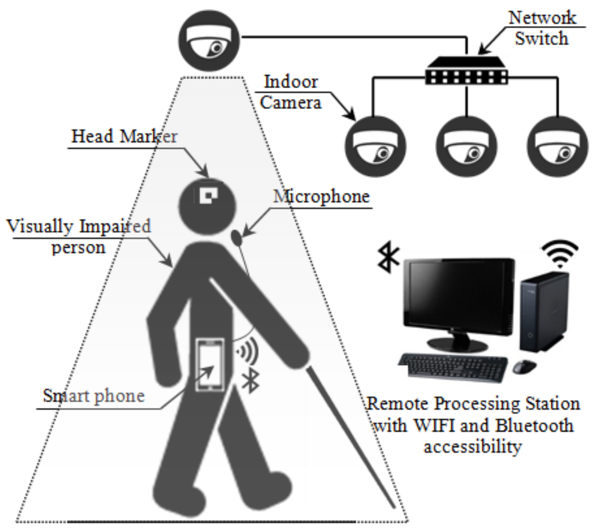

Fig. 1. Proposed System Approach Architecture

A. Mobile Application

The mobile application is the interactive interface with the user. It enables the blind person to invoke voice commands and provides him with audible notification feedback answers. The core of the mobile application consists of two basic functions: the ”speech to text” and the ”text to speech” functions. The first function takes the voice commands from the user and converts them to text. The text is afterwards transmitted to the processing system through an indoor WIFI connection. Simultaneously, the image processing system responds with text containing navigational information. Finally, the application converts this information into audible messages in order for the person to make his move. In addition to these basic functions, the application has a WIFI connection capability for data transmission. This connection is provided through an indoor wireless network between the smart phone and the image processing system. Maintaining the connection between the blind person and the system is very critical. Any connection loss may induce an emergency situation such as a fall incident. To avoid connectivity issues, the application has a redundant point to point connection. This connection is ensured using the Bluetooth communication interface of the smart phone and the remote processing computer. The application automatically switches to Bluetooth when the WIFI network is absent.

The mobile application is downloaded on the user’s smart phone. Thus, the person must hold his phone all the time to be able to navigate. It can be fixed on his belt to free his hands for other tasks. In addition, the phone has to be connected to the local WLAN of the house for operation. Input voice commands can be apprehended from the user through a wired microphone. Audible output messages are transmitted loudly through the speakers of the phone. We avoided using the earplugs to let the person stay aware of other surrounding sounds. The figure below illustrates the functional architecture of the mobile application.

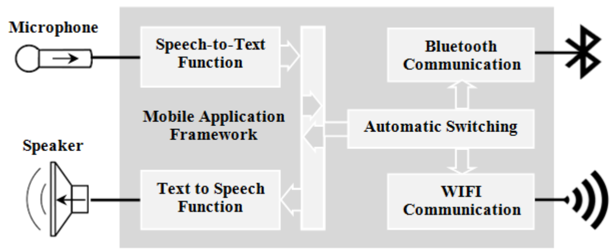

Fig. 2. Mobile application functional architecture.

The mobile application is developed using Android Studio environment. The application is afterwards downloaded on an Android device such as a Samsung smart phone for testing.

B. Remote Precessing System

The processing system is the core of our proposed approach. The system has two input data sources and one output command. The first input source is the IP camera network while the second is the WLAN interface that captures text-converted voice commands. The system processes these two inputs and generates directional text messages for indoor navigation using a computer vision guidance algorithm. Text-converted voice commands and navigational data are interchanged on the same WIFI interface. Table 1 describes the voice commands used by the user to navigate or to locate himself and find his items. The general system behavior uses built-in image processing functions to estimate the position of the person with respect to the room and therefore to the entire house.

The image processing functions for location identification and navigation are developed on Visual Studio environment.

TABLE I
USER VOICE COMMANDS DESCRIPTION

#	Voice Command	Command Description
1	*"Destination Name"*	Name of the location intended (Kitchen, Bedroom, etc.)
2	*"Item name"*	Name of the personal item to be located (Book, Wallet, etc.)
3	*"My location"*	Name of the current location
4	*"Power"*	Percentage of the battery level

The real time image processing techniques use built-in functions found in OpenCV library and wrapped in EmguCV. The latter enables OpenCV library functions to be accessed using different programming languages.

C. Computer Vision Guidance Algorithm

Before the computer vision guidance algorithm is developed, a learning phase must be done to the system. Three objects are currently proposed in our approach. They are declared and identified in a relational database created for this purpose. These objects are defined as follows:

- ***Destination***: Destination objects are locations where the user intends to go. They are the names of the rooms in a living home or an office environment. Locations are identified by their Cartesian coordinates in the database.
- ***Obstacle***: Obstacle objects are indoor assets of a living home or an office. For a VI subject, all assets are considered as obstacle objects and therefore the system must identify these assets to alert the subject. Obstacle objects are identified using their shape or predefined images captured and saved in the database.
- ***Item***: Item objects are personal assets that belong to the subject himself. They are identified using imprinted markers, colors, or shapes.

The properties of these objects can be easily parameterized by the image processing functions used in the computer vision guidance algorithm. As a matter of fact, before processing the text-converted voice commands, the algorithm identifies the object of the command in the database and reacts with the processing accordingly. For his needs, the visually impaired user commands his application with destination and item objects only since obstacle detection is a common-sense feature in any guidance algorithm. In this case two complementary algorithms are developed. The first algorithm provides navigation when destination objects are desired whereas the second algorithm provides location of the desired personal item. In the second case the user is free to reach the item or not. If he decides to reach the item, he must therefore command the system with the location of the item.

The algorithm of the remote processing system is triggered when motion is detected in the house. It captures the image and searches for the imprinted marker at the head of the person. The marker has a particular shape that fits the properties of the image processing functions used in estimating the location and the direction in the house. This step requires a learning phase to the system where each camera is identified and located using its Cartesian coordinates with respect to the entire floor coordinates. When the user location is identified, the system waits for a text-converted voice command. When a command is received, the system analyzes it and responds accordingly, otherwise, the system tracks the user inside the same room providing him only with obstacle detection functionalities. The system responds with directional information for indoor navigation similar to an outdoor GPS. This computation is also preceded by a learning phase where all possible room coordinates in the house are known. Angular directions are provided through the visual analysis of the marker and responses are transmitted to the application for audible conversion. Similarly when the subject asks for a missing personal item, he commands his application. The algorithm begins the search from the last known location where the item was identified. The system replies with the location of the item when it is found. The flow chart of these processing algorithms are shown in figure 3.

IV. EXPERIMENTS

To prove our proposition, experiments were conducted on real subjects. Eight blind folded persons (5 males and 3 females) volunteered to validate the reliability of the system. Subjects wore an imprinted marker located on their head to provide the system with directional information. Experiments were held in three rooms of a living home (kitchen, living room, bedroom) where a network of 3 IP cameras were installed. Each camera can cover a global view of the whole room. Subjects were initially at different locations holding their mobile phone. One by one, they started commanding their application by saying the name of the destination. For example, if the user is in the kitchen and wants to go to the bedroom, he says *"bedroom"* on the mobile. After receiving the command, the remote processing system analyzes the start and the end points of the track providing the user with the number of steps required to reach his destination.

Several scenarios have been tested. From the living room to the bedroom then to the kitchen, from the kitchen to the bedroom and the living room, and from the bedroom to the kitchen and the living room. In all cases, the system has provided the correct path to the destination. On their way to their destination, different obstacles were detected and the system was able to alert the user from hitting them and changing the direction accordingly. Similar results were obtained when obstacles were added dynamically to the path provided that the shape has been already acquired and identified by the system. The positioning of the obstacle must highly correlate with the predefined image for a successful detection. The system needs also to separate between an obstacle and house asset where in most cases they are the same. The experiments were conducted in an empty space where obstacle are only considered.

The search for objects is also another feature of our system. The scenario was conducted where the user is searching for

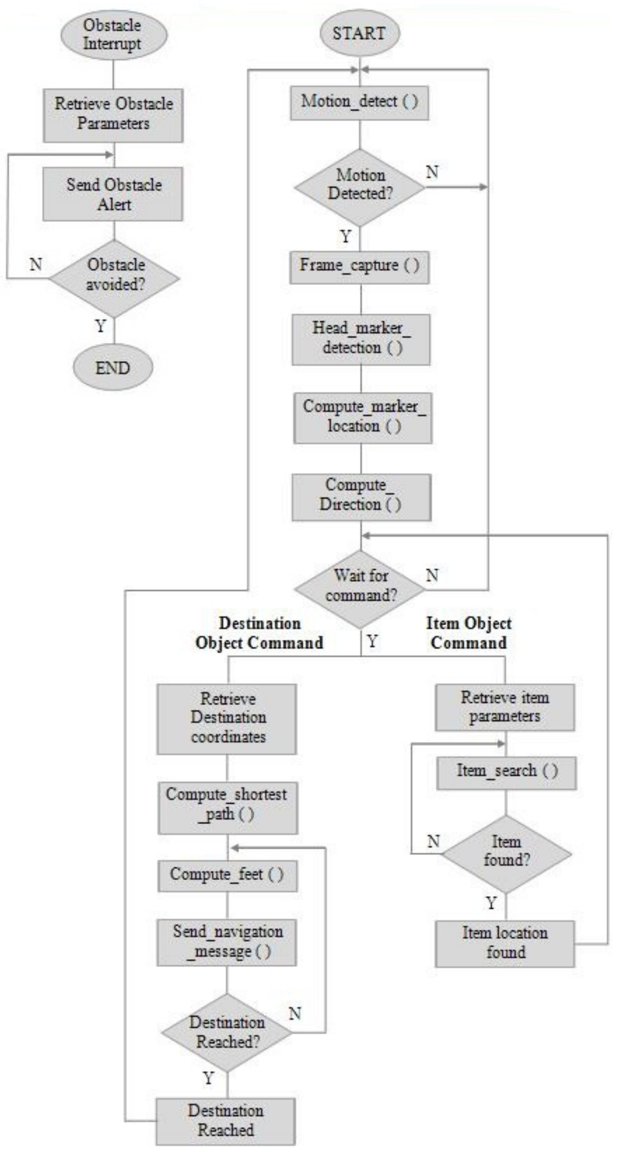

Fig. 3. Flow chart of the computer vision guidance algorithm

a specified personal item. We have provided the system with different items needed by the user (i.e. preferred braille book, preferred chair, wallet, etc.). These items were put randomly in the testing rooms. When the user requests an object, the system begins the search from the last found position. If the object is not found, all cameras are launched to scan the other rooms to look for the missing item. Different places and positions were also tested to validate to reliability of the system. Items were found no matter where they are located provided that the rooms are lightened. The system in its present state can find objects using image characteristics such as the shape or the color. More complex objects such as a preferred chair were unable to be detected by the system. The system must be customized to stop the visually impaired from walking while he is searching for the requested item. Whether the item is found or not, an audible notification must be sounded. The system does not allow multiple commands at the same time. In other words, until now only one person can benefit from this system otherwise the system freezes. One possible solution could be associating the messages to the head pattern inside a queuing stack for sequential processing. This will be considered in our future work also.

The systems requires little training from the user. The latter considered as visually impaired has some messages to memorize. These messages must include the names of the rooms inside the house which the person is already familiar with. Moreover, the system can interact with the person using comprehensible audible English messages. More can be added depending on the user preference.

V. Conclusion

This paper presents a novel design for an indoor navigation system for visually impaired and blind people. The proposed approach has a simple architecture that allows the subject to be fully independent in his home or work. The system provides navigation assistance, obstacle avoidance and object recognition functionalities in indoor premises. Unlike other systems in the related work, the subject needs only to hold his smart phone during his displacement and doesn't require any particular skills to be operated. The complexity of the system resides in computer vision processing algorithm seamlessly to the user. Primary experiments show promising results as the user can feel safe using the system. He can freely navigate inside his house or office without additional assistance. The user is also alerted when an obstacle is encountered and he can locate easily his missing objects. The functionalities of the watchful eye of the system are made available to the user by a simple voice command on his mobile phone.

We are still far from a complete system. Future development is planned for the system expansion. On the application level, we are working to automatically run the application when motion is sensed from the subject (chair rise, walking, etc.). It can also revert to its sleep mode in static conditions to minimize the battery consumption. Battery level can also be communicated loudly through voice messages to avoid critical situations. The application may also offer gait analysis for elderly and visually impaired subjects and may prevent the subject from a potential fall. On the remote processing system level, models are being developed to enhance the learning of the environment and the response time to commands. Moreover, the computer vision algorithm will be redefined to afford the shortest path to the desired destination. These issues will be addressed in future research activities.

References

[1] D. Dakopoulos, S. K. Boddhu, and N. G. Bourbakis. A 2d vibration array as an assistive device for visually impaired. pages 930–937. IEEE Computer Society, 2007.

[2] D. Dakopoulos and N. G. Bourbakis. Wearable obstacle avoidance electronic travel aids for blind: a survey. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 40(1):25–35, 2010.

[3] 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*, pages 63–85, 1993.
[4] A. Hub, J. Diepstraten, and T. Ertl. Design and development of an indoor navigation and object identification system for the blind. pages 147–152. ACM, 2004.
[5] K. Ito, M. Okamoto, J. Akita, T. Ono, I. Gyobu, T. Takagi, T. Hoshi, and Y. Mishima. Cyarm: an alternative aid device for blind persons. In *Extended Abstracts Proceedings of the 2005 Conference on Human Factors in Computing Systems, CHI 2005, Portland, Oregon, USA, April 2-7, 2005*, pages 1483–1488, 2005.
[6] L. A. Johnson and C. M. Higgins. A navigation aid for the blind using tactile-visual sensory substitution. *Conf Proc IEEE Eng Med Biol Soc*, 1:6289–6292, 2006.
[7] J. o. Jose, M. Farrajota, J. M. F. Rodrigues, and J. M. H. Du Buf. The smartvision local navigation aid for blind and visually impaired persons, 2011.
[8] P. Kelley. Historical development of orientation and mobility as a profession. 1999.
[9] K. J. D. L. M. Bouzit, A. Chaibi and C. Mavroidis. Tactile feedback navigation handle for the visually impaired, 2004.
[10] S. Meers and K. Ward. The smartvision local navigation aid for blind and visually impaired persons, 2005.
[11] P. B. Meijer. An experimental system for auditory image representations. *IEEE Trans Biomed Eng*, 39(2):112–121, 1992 Feb.
[12] N. Molton, S. Se, J. M. Brady, D. Lee, and P. Probert. A stereo vision-based aid for the visually impaired. In *Image and Vision Computing*, pages 251–263, 1998.
[13] E. Peng, P. Peursum, L. Li, and S. Venkatesh. A smartphone-based obstacle sensor for the visually impaired. In *UIC*, volume 6406, pages 590–604. Springer, 2010.
[14] G. Sainarayanan, R. Nagarajan, and S. Yaacob. Fuzzy image processing scheme for autonomous navigation of human blind. *Appl. Soft Comput.*, 7(1):257–264, 2007.
[15] S. Shoval, J. Borenstein, and Y. Koren. Mobile robot obstacle avoidance in a computerized travel aid for the blind. In *IEEE International Conference on Robotics and Automation*, pages 8–13, 1994.
[16] T. S. T. Ifukube and C. Peng. A blind mobility aid modeled after echolocation of bats. *IEEE Transactions on Biomedical Engineering*, 38(5):461–465, 1991.
[17] I. Ulrich and J. Borenstein. The guidecane - applying mobile robot technologies to assist the visually impaired. *IEEE Transactions on Systems, Man, and Cybernetics, -Part A: Systems and Humans*, 31:131–136, 2001.
[18] D. Yuan and R. Manduchi. A tool for range sensing and environment discovery for the blind. *2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 3:39, 2004.