Designing a Virtual Environment to Evaluate Multimodal Sensors for Assisting the Visually Impaired

Wai L. Khoo, Eric L. Seidel, and Zhigang Zhu

Department of Computer Science, CUNY City College
Convent Avenue and 138^{th} Street, New York, NY 10031
{khoo, zhu}@cs.ccny.cuny.edu, eseidel01@ccny.cuny.edu

Abstract. We describe how to design a virtual environment using Microsoft Robotics Developer Studio in order to evaluate multimodal sensors for assisting visually impaired people in daily tasks such as navigation and orientation. The work focuses on the design of the interfaces of sensors and stimulators in the virtual environment for future subject experimentation. We discuss what type of sensors we have simulated and define some non-classical interfaces to interact with the environment and get feedback from it. We also present preliminary results for feasibility by showing experimental results on volunteer test subjects, concluding with a discussion of potential future directions.

1 Introduction

Based on the 2002 world population survey, there are more than 161 million visually impaired people in the world today, of which 37 million are blind [2], [6], [7]. Research into alternative perception will have direct impact on these people with regards to navigation and orientation. We define alternative perception as using machines or devices to sense the environment and present the user with meaningful information about his or her surroundings, allowing the user to navigate the area. The machine then adapts based on the decisions made, so that it can intelligently present meaningful information (i.e. based on user's preference).

To realize alternative perception, we must determine what kinds of sensors (or combination of sensors) are better suited as "input" devices. In addition, we must also address the inherited limitations of these sensors and what compromises are needed, e.g., infrared has limited sensing distance. An efficient and robust system must be able to present meaningful information to the user without overloading their senses, which is the downfall of current electronic travel aid (ETA) technologies [1], [2]. The question boils down to which human senses can best be exploited for alternative perception without overloading the user, which is the main scope of this project. In the process of doing this project, some comparisons will be drawn on a sensor's pros and cons with other types.

The remainder of the paper is organized as follows. In Sect. 2 we present an overview of our approach and comparisons to related works. In Sect. 3 we describe the design of the experiment and environment. In Sect. 4 we show some preliminary results based on volunteer subjects. Finally, in Sect. 5 we discuss future work and extensions.

2 Overview of Our Approach

How is our research different from the state of the art ETAs? We need to define the fine line between human and computers. In other words, how much influence should we place on the computer for decision-making? If we rely heavily on the computer, then a minor error in the system or decision-making process will result in a potentially catastrophic error. Conversely, if we rely heavily on the human, then the enormous amount of raw data will overwhelm the user and potentially affect his or her decision adversely, making them ignore the technology all together [1], [6].

As such, we want to study this fine line by testing out various sensors and various non-classical interfaces ("display"). One metric that we can use to compare various approaches of different "display" is to measure a subject's brain [13] and motor activities [14]. Beauchamp, et al measured brain responses to vibrotactile somatosensory, auditory, and visual stimuli using magnetic resonance imaging [13], while Prilutsky, et al discussed how to quantify motor cortex function and the movement kinematics of the corresponding limb [14]. In order to measure human performance in navigation, an accurate tracking system is needed and ground truth of the environment is needed. This will be very hard for a real environment. Torres-Gil, et al [12] have developed a virtual reality simulator that will track the user's head orientation and position in a designated room and generate a virtual view of what the user is seeing. However, instead of presenting the view to the user graphically, an auditory representation of the scene is transduced to the user. Their results are mostly empirical. To have a better understanding of how virtual reality can help us evaluate multimodal sensors for the visually impaired, we are going to look at different brain scans and action measurements, and see which method the users show affinity for or respond well to, thus allowing us to quantify the results. However, brain activity can only be accurately measured when the subject is stationary, which is another reason why we decided to use virtual reality. To do this, we can have the user sit in front of a computer and perform a navigation task in the virtual environment while we obtain brain scans as well as action recordings of the user. The user reach his or her specified destination by relying on various stimulators on his or her body. Further details on simulated sensors and stimulators will be discussed in Sect. 3. The subject is, of course, blindfolded (or is actually blind) so that he or she has to rely on the devices. Using virtual reality not only allows us to determine which "display" is suitable, but also allows us to determine which combination of sensors (homogeneous and/or heterogeneous) is optimal.

We use Microsoft Robotics Developer Studio [15] to construct the 3D virtual environment with an avatar sitting on a wheelchair to approximate a real setting. The user will use an XBox controller to steer the wheelchair. He or she will be sitting on a chair in front of a computer with electrodes on his or her head and various sensors will be strategically placed on the avatar's body, the "display" device will be placed on the user's corresponding body parts. The user (who is either blindfolded or visually impaired) will have to navigate the avatar in an obstacle course and the virtual sensor readings will be translated to the real "display" devices. The user will have to make navigation decision based on the feedback received. Figure 1 illustrates the basic idea and setup of our approach.

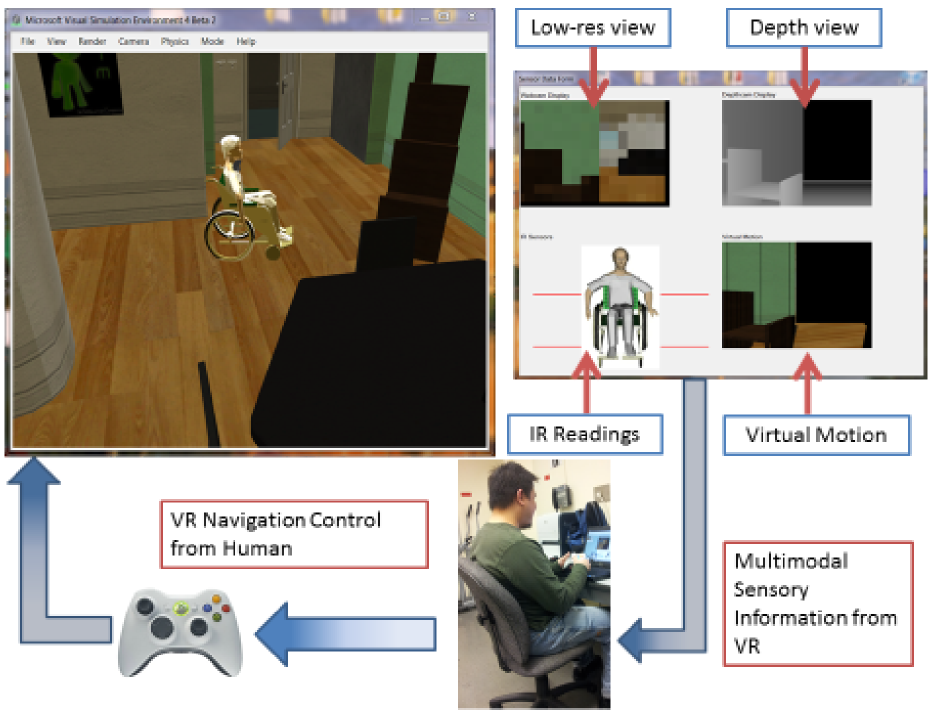


Fig. 1. Sensing and navigating a virtual environment

3 Sensors and Stimulators: Experiment Design

In the setup of Fig. 1, the overlooking view on the left provides the tester (and other sighted people) progress of the testee (subject). The sensor data window on the right shows some of the sensors we are simulating and will be programmatically fed into the corresponding stimulators on the subject. Currently the setup includes simulating low resolution image, a depth view, a simulated motion map, and infrared (IR) sensors.

Fig. 2. Brainport tongue stimulation

The low resolution image will be fed into a tongue stimulating array such as Brainport's vision technology [11] (2). The Brainport is a limited clinical trial device; the white rectangular plastic house the 20 × 20 stimulating array, the signal is obtained from a camera mounted in the center of the glasses and processed on-board in the black handheld device. The depth view is obtained from a simulated Microsoft Kinect. The simulated motion map is derived from the depth view using known intrinsic and extrinsic parameters of Microsoft Kinect, i.e. calculating the disparity value of each pixel. Then two views are generated by shifting all of the pixel locations to the left (and right) by its disparity divided by 2. These two views are displayed consecutively such that you can see objects nearer to you shifting more than objects furthest from you. This map is also an alternative data that we can feed into the tongue stimulating array in an attempt to capture object's presence by virtually "moving" it. The IR sensors (on arms and legs) will be sensing the virtual environment and trigger the corresponding vibrators based on the scene [18]. This method of "display" is called vibrotactile [5]. Figure 3 shows a prototype designed by our lab [18], where the frequency of vibration corresponds to the measured distance (i.e. obstacle that is very closer to the user will have stronger/faster vibration).

3.1 Sensors

We also plan to simulate other sensors including stereo cameras, laser range sensors and ultrasonic range (sonar) sensors. We will simulate more sensors as

Fig. 3. Vibrotactile Prototype

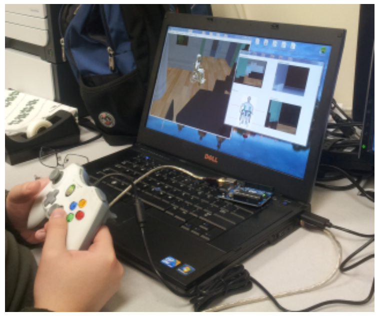
we come across them and if it is necessary to the study of this project. The stereo cameras will be used as a comparison to the Microsoft Kinect, studying its tradeoff in depth of field and computation complexity with regards to aiding the visually impaired. In two pieces of accompanying work, our lab has developed a segmentatation-based stereo vision algorithm for obaining high-level 3D description that can be provided to users [17], and people and obstacle detection algorithms using the Kinect [19]. Similarly, the laser range and sonar sensors will be used as a comparison to IR sensor, studying its tradeoff in range and field of view.

3.2 Stimulators

In addition to vibrators and tongue stimulating array as stimulators (or "display"), we can use braille to indicate range or intensity [16], auditory representation [12], converting 3D space into vibration array [9], and haptic feedback [16]. Braille is a traditional method for visually impaired people to communicate. However, it may be too slow (user "reading" speed may vary as well) to convey all of the spatial information needed for navigation and orientation. Auditory representation is similar in principle to echolocation, as used by bats. We can convert distance information into stereophonics which can be used to localize an objectc's location [9]. However, it may overwhelm the user since he/she will have to constantly "listen" to the scene. This may pose a danger to their safety especially in an urban area.

4 Experimental Results

Here we show the feasibility of the design and prototype mentioned in Sect. 3. As shown in Fig. 4(a), the user were able to navigate the virtual environment with an easy-to-use XBox controller. The controller is an intuitive tool to use for navigation compared to using a mouse and keyboard. Figure 4(a) also show what a typical experiment setup will look like. The user will sit in front of a computer



(a) Example of a physical setup

(b) IR data extracted to Arduino

Fig. 4. Setup

with the controller in hand while wearing a set of stimulators, in this case a set of vibrators. The monitor will mainly be used by the test administrator(s) to monitor the progress of the subject.

Although the vibrotactile device (Fig. 3) is still in the prototyping phase, we were able to demonstrate that IR readings in the virtual reality can be interface out to Arduino, which can be used to control the vibrators on the user's body (Fig. 4(b)).

5 Future Work

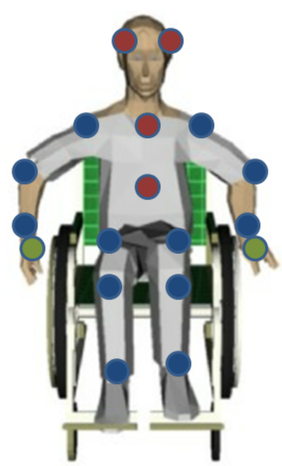


Fig. 5. Range field visualization: IR rangers (blue), sonar rangers (green) and laser rangers (red)

In order to show the capacity of our virtual environment sensor simulation, our next step is to generate a whole body simulation of range sensors. As an illustration, Fig. 5 visualizes the sensor setup of what we have in mind - including 12 IR rangers, 4 sonar sensors, and 4 laser rangers. The IR ranges are from 10 cm to 80 cm, mounted on arms and legs, which can be used for measuring proximity of doors, walls and close-by obstacles and will be transduced to vibrotactile stimulators with increasing levels of vibrations based on the measured distance. The sonar ranges up to 12 meters, while mounted on the user's wrists, two facing front and two facing back, could be used to detect farther obstacles with wider field of view. The laser range sensors are more accurate in both distance and angle, ranging up to 80 meters, which are mounted on user's head (2) and chest (2) for far environment obstacle detection.

To begin our experiments, we will be recruiting some human subjects and collaborating with our colleagues in the psychology department who will collect and analyze brain scans while we configure the sensors-stimulators setup. We will run several experiments with various sensor combination and placements, and various groups of subjects (sighted but blindfolded, low-vision, and totally blind). Finally, with the collected brain scan results and analysis, we will study the optimal combination and placements of sensors, and user's learning curve.

The goal of this project is to provide a research platform for the development of assistive devices for the visually impaired. Specifically, this project aims to determine which interface method (i.e. what type of information and how to present to the user) is the most efficient, reliable, and robust based on the various brain and action measurements that we will collect and study. This is different from mainstream assistive devices which focus on the technological and mechatronics aspects of the design. The key idea here is to test how visual and non-visual stimulation can enhance vision through understanding the underlying neural mechanisms, enhancing effects of the specific non-visual stimulation on vision.

Acknowledgement

This work is supported by U.S. National Science Foundation Emerging Frontiers in Research and Innovation Program under Award No. EFRI-1137172 and City College 2011 Grant for Interdisciplinary Scientific Research Collaborations. We thank Mr. Nathan Slobody and Ms. Ania Wisniewska for their assistance in connecting the XBox controller and Arduino to the Microsoft Robotics Developer Studio.

References

1. Electronic Travel Aids:New Directions for Research, Working Group on Mobility Aids for the Visually Impaired and Blind, Committee on Vision, 1986
2. D. Dakopoulos and N. G. Bourbakis, "Wearable Obstacle Avoidance Electronic Travel Aids for Blind: A Survey," in IEEE Trans. On Systems, Man, and Cybernetics, Vol. 40, No. 1, January 2010.

3. P. Meijer, "An Experimental System for Auditory Image Representations," in IEEE Trans. On Biomedical Engineering, Vol. 39. No. 2, February 1992
4. J.L. Gonzalez-Mora, et al. "Development of a new space perception system for blind people, based on the creation of a virtual acoustic space." Available online: http://www.iac.es/proyect/eavi
5. L. A. Johnson and C. M. Higgins, "A navigation aid for the blind using tactile-visual sensory substitution," in Proc 28th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., New York, 2006, pp. 6298-6292.
6. J. Liu and X. Sun, "A Survey of Vision Aids for the Blind," in Proceedings of the 6th World Congress on Intelligent Control and Automation, June 21 - 23, 2006, Dalian, China
7. American Foundation for the Blind. Available: http://www.afb.org/
8. K. Ito, et al. "CyARM: an alternative aid device for blind persons," in Proceedings of CHI Extended Abstracts 2005. pp.1483-1488
9. N. Bourbakis, "Sensing surrounding 3-D space for navigation of the blind." in IEEE Eng Med Biol Mag. 2008 Jan-Feb;27(1):49-55.
10. M. Hirose, T. Amemiya, "Wearable Finger-Braille Interface for Navigation of Deaf-Blind in Ubiquitous Barrier-Free Space," In Proc. of 10th International Conference on Human-Computer Interaction (HCI International 2003), Universal Access in Human Computer Interaction, Vol. 4, HCII 2003, pp. 1417-1421, Crete, Greece, June 2003.
11. Wicab, Inc., "Brainport Vision Technology", http://vision.wicab.com/technology/
12. M. A. Torres-Gil, O. Casanova-Gonzalez, and J. L. Gonzalez-Mora, Applications of virtual reality for visually impaired people, W. Trans. on Comp., Vol 9 Issue 2, February 2010, pg. 184-193
13. M. S. Beauchamp, N. Yasar, R. Frye, and T. Ro (2008). Integration of touch, sound and vision in human superior temporal sulcus. NeuroImage, 41, 1011-1020.
14. B.I. Prilutsky, M. G. Sirota, R. J. Gregor, and I. N. Beloozerova, Quantification of motor cortex activity and full-body biomechanics during unconstrained locomotion. Journal of Neurophysiology 2005, 94: 2959-2969
15. Microsoft Robotics Developer Studio. Available: http://www.microsoft.com/robotics/
16. T. Amemiya, K. Hirota, and M. Hirose, "Wearable Tactile Interface for Way-Finding Deaf-Blind People using Verbal and Nonverbal Mode", Trans. of VRSJ, Vol. 9, No. 3, pp. 207-216, 2004. (in Japanese)
17. H. Tang and Z. Zhu, "A Segmentation-based Stereovision Approach for Assisting Visually Impaired People", the 13th International Conference on Computers Helping People with Special Needs (ICCHP), July 11-13, 2012, Linz, Austria.
18. F. Palmer, Z. Zhu and T. Ro, "Wearable Range-Vibrotactile Field: Design and Evaluation", the 13th International Conference on Computers Helping People with Special Needs (ICCHP), July 11-13, 2012, Linz, Austria.
19. A. Khan, J. Lopez, F. Moideen, W. L. Khoo and Z. Zhu, "KinDectect: Kinect Detecting Objects", the 13th International Conference on Computers Helping People with Special Needs (ICCHP), July 11-13, 2012, Linz, Austria.