AN INTERACTIVE TRAINING GAME USING 3D SOUND FOR VISUALLY IMPAIRED PEOPLE

Hsiao Ping Lee[1, 2], Yen-Hsuan Huang[2] and Tzu-Fang Sheu[*]

[1]Department of Medical Research, Chung Shan Medical University Hospital, Taichung, Taiwan, R.O.C.
[2]Dept. of Medical Informatics, Chung Shan Medical University, Taiwan, R.O.C.
[*]Dept. of Computer Science and Information Engineering, Providence University, Taiwan, R.O.C.

ABSTRACT

The number of visually impaired people is increasing year by year. Although attention has been given to the needs of people with disabilities, most of the discussion has focused on social welfare, while talk about assistive technology for people with disabilities is rare. The blind need training courses for reconstruction and rehabilitation. Orientation and mobility (O&M) training is important for visually impaired people, teaching safe, efficient and effective travel skills. Skills learned from O&M training courses can help the blind walk on the street safely. Crossing the street is especially dangerous, since blind people cannot see traffic lights, and rely mostly on sound for information about their environment. Thus, learning to recognize the varied sounds of vehicles and determining the direction and speed of moving vehicles is critical. In this paper, we propose an interactive game with 3D sound that simulates a busy street environment. The proposed game tries to build a virtual environment with 3D sound to help visually impaired people learn to cross the street safely. As the proposed training game is designed for the blind, the technologies of Kinect and Text-to-Speech (TTS) are used in the human-computer interface in the proposed game, so that they can use the game independently.

1. INTRODUCTION

Vision is the most important and natural way humans receive information from the environment. We rely on vision to handle most of our tasks every day. Those with adventitious blindness depended on vision to receive most messages from their environment before going blind, so it is difficult for them to adapt when they lose their eyes. For example, if we drop a key, we will look for the key with our eyes and pick it up. Visually impaired people, on the other hand, have to listen to the sound when a key falls. Today, the number of visually impaired people is increasing year by year. Although a great deal of attention has been given to the needs of people with disabilities, most of the discussion has been about social welfare rather than assistive technology for people with disabilities. Training courses in reconstruction and rehabilitation allow the blind to lead a more independent life.

Orientation and mobility (O&M) [1] is important training for visually impaired people. O&M is a profession that focuses on instructing individuals who are blind or visually impaired in safe and effective travel through their environment [2]. O&M training research is carried out by medical and special education researchers, however, most O&M training needs extensive space and expensive facilities. At the Institute for the Blind of Taiwan, for example [3], the trainee needs a large area in which to train. The instructor assists the blind person in learning to use a white cane or guide dog to walk on the road [4]. After training, the blind person can sense his or her location by noise, smell and sound direction. The limitation is that the visually impaired person cannot learn alone, an instructor must assist with the training. Our goal is to develop O&M training for blind people using assistive technology. We want the training to be entertaining, game-based, and easy for visually impaired people to use by themselves.

In this paper, we report the development of an Auditory Perception Re-establishment Training System (APRETS) for visually impaired people. APRETS can simulate the traffic on a busy street environment, assisting the blind to learn to cross the street using 3D sound. APRETS is simple to operate by motion detection, using the Kinect system. The rest of the paper is organized as follows. Previous work related to APRETS is presented in Section II. The state of the art in software in discussed in Section III. This is followed by the implementation of APRETS in Section IV. We had visually impaired people test our system, and report the results of the test in Section V. Finally, the conclusions of this work are given in Section VI.

2. BACKGROUND AND RELATED WORKS

APRETS uses 3D sound technology to simulate real 3D traffic sounds, which are synthesized using computerized techniques [5]. Most 3D sound techniques are based on head-related transfer function (HRTF) [6]. HRTF is based on the principles of Interaural Time Differences (ITD), Interaural Intensity Differences (IID) and the Pinna Effect to generate 3D sound [7]. In reality, if you hear one sound, the time and the intensity of the sound detected is slightly different in the two ears, allowing us to judge the sound direction. Software allows us to duplicate this effect using HRTF. As the technology improves, the sound cards already in personal computers will allow us to hear 3D sound.

APRETS uses a motion detection to operate the system. The motion detection system we used is "Kinect," introduced by Microsoft in 2010 [8, 9]. As opposed to traditional game control systems, Kinect uses body motion, allowing the player to control the game instinctively. Using the same properties, Kinect can be used in other domains, such as education and rehabilitation [10, 11]. The Kinect sensor is a horizontal bar connected to a small base with a motorized pivot and is designed to be positioned lengthwise above or below the video display. The device has an RGB camera, depth sensor and CMOS running, which provides full-body 3D motion capture.

In the Kinect development kit, we used "OpenNI" (Open Natural Interaction), a multi-language, cross-platform framework. Currently supported platforms are Windows, Mac, and Linux [12]. OpenNI has a three-layer concept. The first layer represents the interaction applications. The second layer represents the middleware components. The third represents the hardware device such as a microphone, color camera and 3D depth camera. In OpenNI, there are two types of Production Nodes; one is a Sensor Related Production Node, the other one is the Middleware Related Production Node. Therefore, OpenNI can run in three modes: (1) full body analysis, (2) hand point analysis, (3) gesture detection and (4) scene analyzer. We use these modes in our system.

Text-to-speech (TTS) is used in the human-computer interface in APRETS. The TTS system converts normal language text into speech [13]. Through many years of development of TTS, the output speech is becoming quite fluent. Different TTS engines can support different languages, but our TTS engine outputs English speech.

APRETS is an interactive game for visually impaired people, so our goal is "Game Accessibility" [14]. In general, intuitive operation in game design has three components [15]: (1) Continuous presentation, (2) Physical actions, (3) Reversible actions with immediate feedback. From the above, we can know the message present, the operating mode and the immediate feedback are game design points. For visually impaired people, our goal was to create an Audio Game, designed with the following capabilities [16]:

- Voice Navigation
- Voice can repeat
- Easy Control
- Different levels
- Objects Hint

The interactive game that we proposed has the following features:

- No special space is needed for training; APRETS can be used in anywhere.
- No expensive facilities are needed; it uses only headphones, Kinect and a PC.
- Easy to use; visually impaired people can operate it independently without help.
- 3D Sound technology is used in the game; 3D sound simulates the environment of busy street traffic, with many different vehicles simulated.
- Supports speech feedback.
- Visually impaired people can use Kinect as a special human interface.

Our system can reduce the expense of O&M training for blind people and give them higher motivation to participate in the training. APRETS uses a text-to-speech engine so blind people can accept and control APRETS easily.
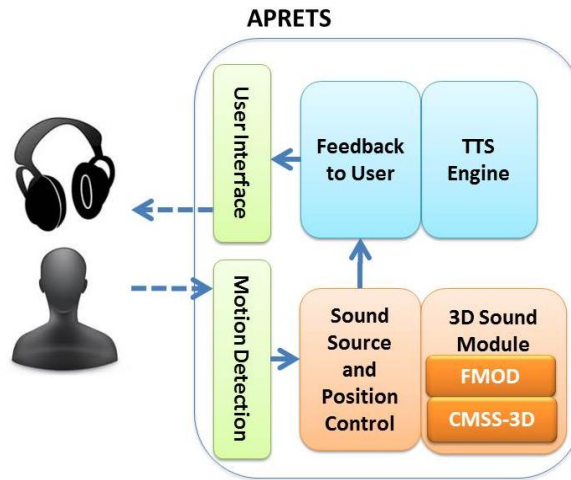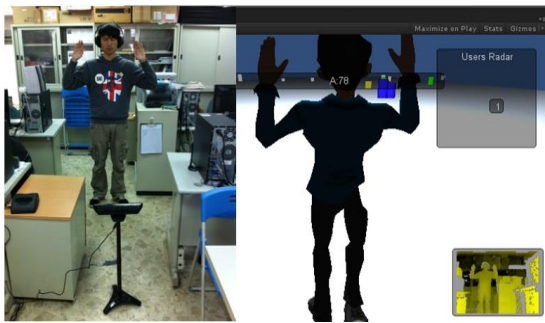

Figure 1. The system architecture.

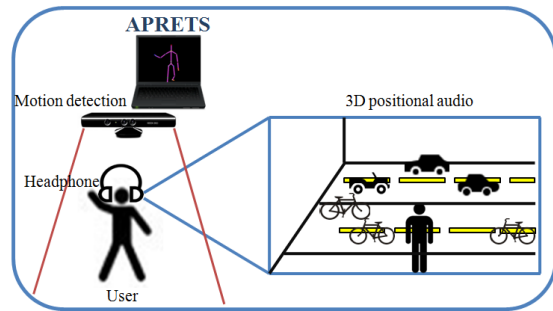
Figure 2. Hands up for calibration.

Figure 3. The game situation.

3. SYSTEM DESIGN

We developed an interactive game for visually impaired people to use Kinect and 3D Sound for O&M training. To make the game more interesting, we added music and speech.

3.1 Experimental Environment

1) Motion detection – Kinect: Nintendo's "Wii" made Motion Sensing Game gaming popular [17, 18], but Microsoft Kinect is more popular worldwide. Kinect doesn't need a controller, as your body is the controller, allowing you to play games more simply and more intuitively. We like this, so in our system we used Kinect to control the system.

2) Wireless Headphone: Our system requires 3D Sound and freedom of movement, so wireless headphones are necessary.

3) Sound Card: The sound card must support 3D sound. The Creative sound card uses CMSS-3D surround sound technology. CMSS-3D has been used to make spatial sound [19], so we chose to use the Creative sound card output for 3D sound.

4) Personal Computer: APRETS can run on any computer running a Windows platform. For convenience, we used a laptop when doing the test. *3.2 System Architecture*

In this paper, the computer is our data server. APRETS was developed on a Windows platform with four parts: User interface, 3D sound module, Motion detection driver and APRETS kernel module. The system architecture is presented in Figure 1.

1) User interface: Following voice hints, the user makes the correct pose to operate the system. Since the system is for visually impaired people, we don't need graphics on screen. In the game, we have sound hints and voice hints. The user can put his hands up and swing his arm to operate the system.

2) 3D Sound module: The system uses FMOD produced 3D Sound. FMOD is a proprietary audio library [20], a popular and powerful cross-platform interactive audio system. The FMOD software was used with the Creative sound card hardware [21].

3) Motion detection driver: The motion detection system receives skeleton information and position information. This information is interpreted by the APRETS kernel module. We exploited the software development kit, OpenNI to develop this module.

4) APRETS kernel module: The APRETS kernel module controls data transfer and training. Feedback hints from the user's motion and progression feed back to change the sound source and position.

3.2 The Flowchart of the Game

We designed a game like the classic videogame "Frog crossing." The traffic is different from "Frog crossing," in that the sounds of bike, motorcycle, car and truck are produced using 3D sound, and there are many vehicles in one game, since it is designed to help visually impaired people react appropriately in real situations. We hope to help blind people learn to recognize hints about traffic including each vehicle's direction, distance and speed using the APRETS system.

The game flow chart is simple, as we don't want to make it too hard to play the game. The game begins with calibration of the Kinect system with the user placing his or her hands up as shown in Figure 2. When calibration is completed, the system will explain verbally how to play the game.

The game situation is presented in Figure 3. There are three roadways in the game. The first roadway has only bikes, going at different speeds and appearing at different times. This is the easiest roadway, with the bikes looping three to four times. The user can cross safely after the bikes pass. In this roadway, we want the user to learn to judge the bikes' path and get used to the system. Cars and motorcycles appear in the second roadway, so the speeds are faster and it is harder to cross than the first roadway. In the second roadway, we want the user to judge both the speed of vehicles and type of vehicles. The hardest level is the third roadway, which has more cars and trucks, at faster speeds than the second roadway. Once the user is able to cross the third roadway, he or she has successfully completed the game.

When the game is over, the system will calculate the time spent playing the game and tell the user. If you are unfortunate and have a traffic accident in the game, there will be a sound of brakes, the game is over, and it will tell you how much time you spent playing the game. If you succeed, you can hear a short section of jazz music after completing the game. We want the user to feel exhilaration in playing APRETS as they learn how to navigate through traffic on the roadway as part of O&M training. Since people love games, we hope they will love this training.

4. CONCLUSION

In this paper, we report the development of an interactive training game for visually impaired people, and use of that interactive game for O&M training. The game can simulate a street environment with busy traffic. APRETS is a computer-based game that visually impaired people can play with Kinect and wireless headphones. We have designed a user-friendly interface which uses text-to-speech and voice to guide the blind user in using APRETS. After the game prototype was completed, we invited five visually impaired persons to test the proposed game. We completed a questionnaire and interview with the subjects after the experiment. Most subjects were satisfied with the training game. The overall satisfaction was 3.88. The users also gave some feedback and suggestions for system improvement. All users thought the game, played with Kinect and 3D sound, was really novel and interesting. In the future, we will extend the functions of the training game for other O&M training.

ACKNOWLEDGEMENT

REFERENCES

[1] Hill, Everett W, and M. Snook-Hill. "Orientation and mobility." *Foundations of Education for Blind and Visually Impaired Children and Youth*, New York, 1986.

[2] Orientation and mobility, http://en.wikipedia.org/wiki/Orientation_and_mobility, Retrieved 01, 2013.

[3] Institute for the Blind of Taiwan, http://www.ibt.org.tw/, Retrieved 10, 2011.

[4] Institute for the Blind – Orientation and mobility training, http://library.taiwanschoolnet.org/cyberfair2005/angel29989187/institute1-2.htm, Retrieved 10, 2011.

[5] FALLER, K. J, BARRETO, Armando. "Simulation and Real-Time Implementation for Teaching 3D Sound". *COMPUTERS IN EDUCATION JOURNAL*, 16.2: 36, 2006.

[6] head-related transfer function, http://en.wikipedia.org/wiki/Head-related_transfer_function, Retrieved 12, 2011.

[7] Duncan James Philip and Newman, "Head-Related Transfer Functions. And their application in: Real-Time Generation of Binaural Spatial Audio." October 2003.

[8] Micorsoft, http://www.microsoft.com/zh-tw/default.aspx, Retrieved 01, 2013.

[9] Kinect, http://en.wikipedia.org/wiki/Kinect, Retrieved 01, 2013.

[10] Da Gama, A., et al. "Poster: Improving motor rehabilitation process through a natural interaction based system using Kinect sensor." *3D User Interfaces (3DUI), 2012 IEEE Symposium* on. IEEE, 2012.

[11] Tölgyessy, Michal, and Peter Hubinský, "The Kinect Sensor in Robotics Education." *Proceedings of 2nd International Conference on Robotics in Education*, 2010.

[12] He, Guan-Feng, et al. "Real-time gesture recognition using 3D depth camera." *Software Engineering and Service Science (ICSESS), 2011 IEEE 2nd International Conference on.* IEEE, 2011.

[13] Prahaliad and J. Zhejiang, A Text to Speech Interface for Universal Digital Library Univ SCI, vol. 6A, no. 11, pp. 1229-1234, 2005.

[14] Game-Accessibility, http://www.accessibility.nl, Retrieved 01, 2012.

[15] Winberg, Fredrik, and Sten-Olof Hellström. "Qualitative aspects of auditory direct manipulation: a case study of the Towers of Hanoi." *Proc. of the 7th Int. Conf. on Auditory Display*. ICAD, 2001.

[16] M. France, "Audio Game Survey Results.", http://www.accessibility.nl/games/researchdocs/MarkFranceAudioGame2007.pdf, 2007.

[17] Nintendo of America Inc, http://www.nintendo.com/wii, Retrieved 01, 2012

[18] WII, http://zh.wikipedia.org/zh-hk/Wii, Retrieved 12, 2011.

[19] Sodnik, Jaka, et al. "Spatial sound localization in an augmented reality environment." *Proceedings of the 18th Australia conference on Computer-Human Interaction: Design: Activities, Artefacts and Environments.* ACM, 2006.

[20] Blommer, Mike, and Jeff Greenberg. "Realistic 3d sound simulation in the virttex driving simulator." *Proceedings of DSC North America*, 2003.

[21] Sound Blaster technologies in FMOD, http://connect.creativelabs.com/developer/Gaming/SoundBlasterFMOD.pdf, Retrieved 01, 2013.