INTERNATIONAL WORKSHOP
ON THE PRINCIPLES AND APPLICATIONS OF SPATIAL HEARING
IWPASH
NOVEMBER 11-13, 2009, ZAO, MIYAGI, JAPAN

Development of virtual auditory display as middleware with high accuracy

Yukio IWAYA[*1], Makoto OTANI[*2], Wataru Sato[*1], and Yôiti Suzuki[*1]
1 Research Institute of Electrical Communication, Tohoku University
2 Faculty of Engineering, Shinshu University

1 Introduction

By the late 1980s, simple real time Virtual Auditory Display (VAD) systems became possible with a DSP chip. Recently, signal processing for VAD has become quite easy with a native CPU in a personal computer. For that reason, several research groups have developed VAD systems based on PCs, for examples, SLAB at NASA [1] and DIVA at Helsinki University of Technology [2]. Moreover, some commercial or GNU software systems [3] are available. We also have put a lot of efforts into developing a high performance VAD software engine as middleware on personal computers. This VAD works with a native CPU of a personal computer on Microsoft Windows and outputs sounds for headphones installed with a three-dimensional position sensor. Consequently, this VAD is responsive to a head movement to realize high precision. In this paper, the overview and applications of the VAD middleware are introduced.

2 Overview of the middleware (SifASo)

Our VAD engine is called as Simulation environment for three-dimensional Acoustic Software (SifASo). It has high performance to render three-dimensional auditory space, such as locating several sound sources by convolving individualized [4] or non-individualized HRTFs, the Doppler effects [5], first order reflections, and reverberations. The HRTFs are smoothly interpolated according to head movements and sound source movements. Another advantage of SifASo is its total system latency of only around 30 ms, including the latency of the trace data of listener's head movements, which is shorter than most of existing engines and much shorter than the detection threshold of the delay [6]. Because of these advantages, SifASo realizes stable, precise, and natural positioning of rendered sound images, especially moving sounds. SifASo was developed as middleware in a form of a set of Dynamic Link Libraries (DLLs) so that it can be called from application software systems. Therefore, it facilitates a development of programs including advanced signal processing techniques. The class diagram of the main part of SifASo is shown in Fig. 1.

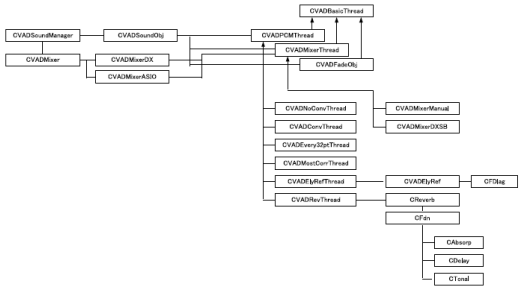

Fig. 1. Structure of class diagram of SifASo

3 Example of applications

We intended to apply SifASo to develop a system to train space cognition for visually impaired people. Visually impaired people must recognize spaces without visual information, but especially with auditory three-dimensional information. For that reason, training of the ability of spatial cognition is regarded as an extremely important subject to learn for pupils and those who have lost their sight. Based on the middleware engine, we developed several application software systems that are apparently intended as games for entertainment. Furthermore, they are useful not only for the training purposes, but also for improvement of quality of life (QoL).

* E-mail address: iwaya@riec.tohoku.ac.jp

‘Bee Bee Beat’ is an action game that trains the ability of sound localization (Fig.2 (a)). This game resembles the popular ‘whack-a-mole game.’ In this game, sounds of honeybees appear instead of pesky moles. A player has a plastic hammer to hit honeybees. A three-dimensional position sensor and a vibration unit are installed in the hammer. The player should localize a sound of honeybee quickly and hit it with the hammer. When the position of the hammer and that of the bee coincide within a certain present range, the hammer vibrates and some points are given. From playing the game, we found several positive transfer effects for practical life [7], such as face-communication ability, ability of avoidance from sound coming, and so on.

‘Sound formula (Fig. 2 (b))’ is a car race game that can be played using only sounds provided by an auditory display. In this game, the Doppler effects are added to the race machines.

‘Mental Mapper (Fig. 2 (c))’ is a kind of maze game and aims directly at training the acuity to draw mental/cognition maps. The player’s task is to reach a goal, walking in a maze built with the virtual sound space. Walking in the maze is virtually controlled by a joy pad. In this game, the first order reflections from walls are added. If a player reaches some specified places in a maze, one animal cry begins as a *sound mark,* which signals that the player has passed the place. A maze editor was also prepared to allow users to create arbitrary mazes.

(a) Bee Bee Beat (b) Sound formula (c) Mental Mapper

Fig. 2. Example of three-dimensional sound games for visually impaired people

4 Future works

We are continuously developing software VAD system with high accuracy and more number of rendering methods, such as high-order reflections, ambient sounds, expression of distance, and so on. Individualization and compensation of HRTFs are extremely important for enhancing the realism of the virtual sound space.

Acknowledgements

This work was supported by Grant-in-Aid for Scientific Research (C) No.20500110, Grant-in-Aid for Specially Promoted Research No.19001004 from MEXT Japan, and the Cooperative Research Project Program from the Research Institute of Electrical Communication, Tohoku University for Dr. Seki, Y. The authors thank Prof. M. Ohuchi for his efforts in development of these applications.

References

[1] http://human-factors.arc.nasa.gov/SLAB/

[2] L. Savioja, J. Huopaniemi, T. Lokki and R. Väänänen, “Creating interactive virtual acoustic environments”, J. Audio Eng. Soc, 47, 9, 675–705, 1999.

[3] For example, http://www.3dsoundsurge.com, http://www.ircam.fr.

[4] Yukio Iwaya, “Individualization of head-related transfer functions with tournament-style listening test: Listening with other’s ears,” Acoust. Sci. & Tec, 27, 6, 340-343, 2006.

[5] Yukio Iwaya and Yôiti Suzuki, “Rendering moving sound with the Doppler effect in sound space,” Applied Acoustics, 68, 8, 916 - 922, 2007.

[6] Satoshi Yairi, Yukio Iwaya, and Yôiti Suzuki, “Estimation of detection threshold of system latency of virtual auditory display,” Applied Acoustics, 68, 8, 851-863, 2007.

[7] Akio Honda, Hiroshi Shibata, Jiro Gyoba, Yukio Iwaya, and Yôiti Suzuki, “Transfer effects on communication and collision avoidance behavior from playing a three-dimensional auditory game based on a virtual auditory display,” Applied Acoustics, 70, 6, 868-874, 2008.