Sound Spatialization Framework: An Audio Toolkit for Virtual Environments

Jens Herder
Spatial Media Group
University of Aizu
Fukushima-ken 965-8085
Japan
voice: [+81](242)37-2579; fax: [+81](242)37-2706
e-mail: herder@u-aizu.ac.jp
www: http://www.u-aizu.ac.jp/~herder

Abstract

The Sound Spatialization Framework is a C++ toolkit and development environment for providing advanced sound spatialization for virtual reality and multimedia applications. The Sound Spatialization Framework provides many powerful display and user-interface features not found in other sound spatialization software packages. It provides facilities that go beyond simple sound source spatialization: visualization and editing of the soundscape, multiple sinks, clustering of sound sources, monitoring and controlling resource management, support for various spatialization backends, and classes for MIDI animation and handling.

Keywords: sound spatialization, resource management, virtual environments, spatial sound authoring, user interface design, human-machine interfaces

1 Introduction

Computer graphics applications, including virtual reality environments, use spatial sound as an user interface. Good animation requires impressive and immersive sound. Also, audio-visual equipment like an ordinary PC which can produce such effects is becoming increasingly available. The Sound Spatialization Framework [Her98], developed at the University of Aizu, adds to graphics systems a sound management runtime environment and a development environment for spatial sound applications [Her97]. The toolkit has the following features:

- Sound spatialization API, conforming to the VRML97 [BCM97] standard
- Efficient resource management, including clustering of sound sources
- Flexible control and monitoring of resource management processes
- Visualization of sound objects in a virtual environment

- Editing of sound objects
- Soundscape manipulation
- Support for various sound spatialization backends
- Dynamically extensible for new spatialization backends
- Full integration into Open Inventor
- Easy control of animations using MIDI streams
- Handling of MIDI streams
- Active lights for enhancing a visual scene
- Active cameras to focus on events
- Camera animations
- Simple scripting

The sound spatialization resource management [HC97] plays the central role of the system. The context shown in Figure 1 shows as the top layer a virtual reality application accessing the system via a sound spatialization application programmer interface. The resource management process itself is hidden to the VR environment. Different sound spatialization backends are accessible through an abstract sound spatialization backend interface. This abstraction makes it easy to add support for new devices without the necessity of changing the management module or application.

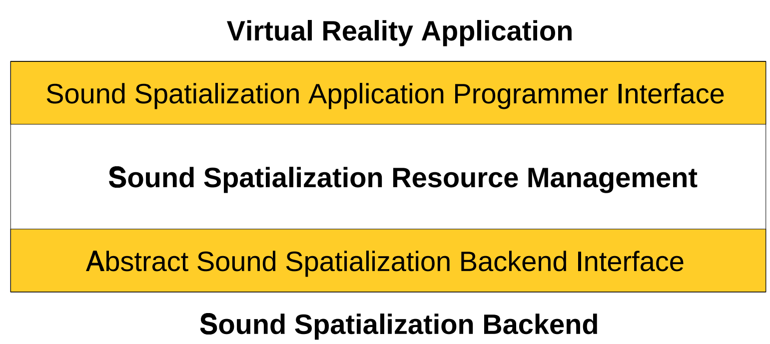

Figure 1: Resource management context

An example application for spatial sound is a chatspace, a place in cyberspace allowing people to meet and communicate. Spatial sound adds the following benefits [BCR97]:

- spatial awareness — attendance position or orientation relative to others
- understanding of the space — ambience: size and reverberation characteristics of a room

- ability to focus on events — someone entering a room or commencing talking
- separation of voices for better audibility — the cocktail party effect
- speaker identification via position

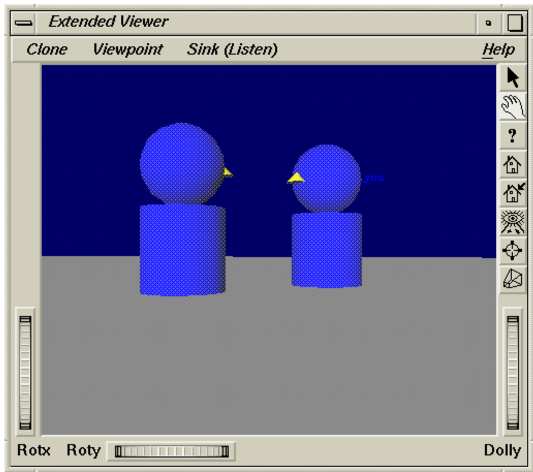

Figure 2: Chatspace – people meet and communicate in cyberspace

Without a resource management system like that provided with the toolkit, application programmers must anticipate a lot of different configurations, with a consequent burden in programming spatial audio sources. During development of a systems which use spatial audio, the required spatialization resources and available spatialization resources are hard to predict (e.g., the number of participants in a chatspace application might vary, or the available spatialization resources on a certain computer system might be limited). Resource management eases the implementation by taking over the resource management processes, gradually scaling down the resource demand, depending on resource availability in a perceptually optimal form.

2 Visual Debugger for Sound Objects in a Virtual Environment

Design and implementation of sound for virtual environments can be simplified and improved using a visual debugger, as shown in Figure 3. The left shows a running VR application; on the right side, sound objects (i.e., sound source and sinks) and parameters like location, orientation, priority, intensity or sensitivity, and radiation pattern are visualized using 3D icons. All objects can be instantly modified and any scene change by an application or editor is directly visible.

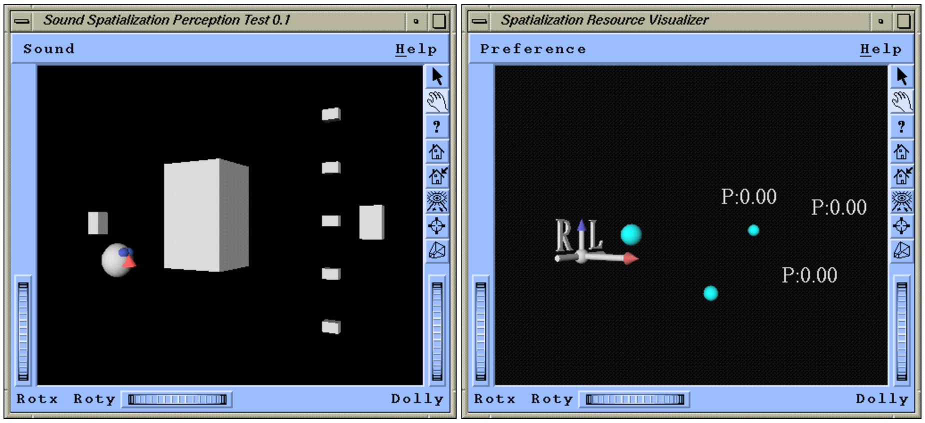

Figure 3: Visual debugger for sound objects

3 Audio Rendering Process

A scenegraph defines a scene with graphical, interactive, acoustical (see Figure 4), and behavior nodes and can be constructed either by class instantiation or external files which again can be created using other authoring tools. Each node is defined in its own local coordinate system. During an audio rendering pass all transformations are resolved and necessary audio control data passed to the resource manager. The resource management and final rendering step in a spatialization backend are done in world coordinates. Resource management involves mapping from sound sources to available spatialization channels, including a scheme to predict the perceptual relevance of a sound source in a given configuration. Resources are used economically by applying a clustering technique which mixes spatially proximate sound sources, representing them as a single sound (virtual) source.

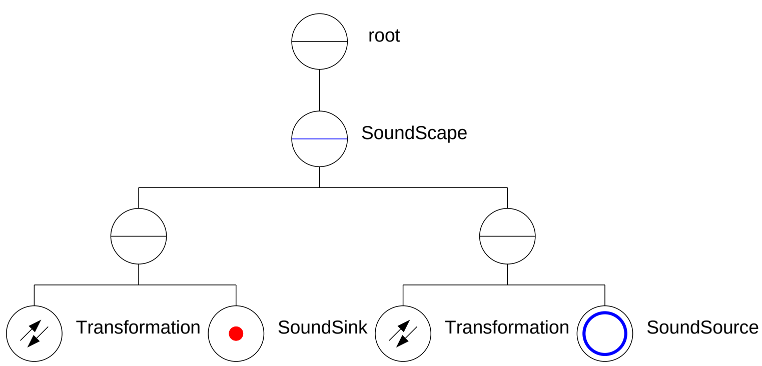

Figure 4: Scenegraph with sound objects

4 Implementation

The toolkit was developed on SGI workstations, using spatialization backends like the Acoustetron II [Beg94, p. 205–208] from Aureal/Crystal River Engineering and Pioneer Sound Field Control system [AMY+98]. Also a minimal spatialization backend for MIDI is included using simple panning and distance-dependent intensity scaling. Support for other spatialization backends can be easily added using a defined interface.

The Open Inventor graphics toolkit [Wer94] was expanded for classes (nodes) to support the spatial sound extensions, which are used for our virtual reality applications. Open Inventor is a superset of the VRML 1.0 standard [BPP95], which does not support sound or dynamic behavior of objects. For the sound extensions of Open Inventor, we followed the VRML97 standard [BCM97], but added a node for sound sinks. This allows scenes to have multiple sinks as well as a sink which can be separated from the viewpoint. The binary dynamic shared library comprising 62 classes and example programs (i.e., Open Inventor file viewer which supports the spatial audio extension, runtime environment for perceptual tests controled by external script and scene file, MIDI animator, MIDI filter which doubles and transposes a note, ..) are freely available [Her98].

5 Conclusion and Future Work

Besides development of VR applications, we use the toolkit as a testbed for different spatialization backends, studies of spatialization resource management processes, and experiments with unique concepts like multiple listener instantiations. We plan to enable interactive editing of sound objects in a graphically (using manipulators for the visual representatives), additional to the implemented textual form. Additional sound objects like sound occluders will be integrated [MH99]. Also we plan to improve the simple MIDI spatialization algorithm.

Acknowledgments

The author thanks Michael Cohen and William L. Martens for fruitful discussions. This work was started as part of the Helical Keyboard project, funded by the Fukushima Prefectural Foundation for the Advancement of Science and Education. The author acknowledges Minefumi Hirose, Soushi Hiruta, Kuniaki Honno, Kimitaka Ishikawa, Junichi Suda, Taku Suzuki, and Junichi Yoshiba for their help in doing some of the implementation as part of their senior and master projects at the University of Aizu.

References

[AMY+98] Katsumi Amano, Fumio Matsushita, Hirofumi Yanagawa, Michael Cohen, Jens Herder, William Martens, Yoshiharu Koba, and Mikio Tohyama, *A Virtual Reality Sound System Using Room-Related Transfer Functions Delivered Through a Multispeaker Array: the PSFC at the University of Aizu Multimedia Center*, TVRSJ: **Trans.**

of the **V**irtual **R**eality **S**ociety of **J**apan **3** (1998), no. 1, 1–12, ISSN 1342-4386.

[BCM97] Gavin Bell, Rikk Carey, and Chris Marrin, ISO/IEC *14772-1:1997: The Virtual Reality Modeling Language (*VRML97*)*, 1997, http://www.vrml.org/Specifications/VRML97/.

[BCR97] Woodrow Barfield, Michael Cohen, and Craig Rosenberg, *Visual, Auditory, and Combined Visual-Auditory Displays for Enhanced Situational Awareness*, The International Journal of Aviation Psychology **7** (1997), no. 2, 123–138.

[Beg94] Durand R. Begault, *3-D Sound for Virtual Reality and Multimedia*, Academic Press, 1994, ISBN 0-12-084735-3.

[BPP95] Gavin Bell, Anthony Parisi, and Mark Pesce, *The Virtual Reality Modeling Language, Version 1.0 Specification*, May 1995, http://www.vrml.org/Specifications/VRML1.0/.

[HC97] Jens Herder and Michael Cohen, *Sound Spatialization Resource Management in Virtual Reality Environments*, ASVA'97 — Int. Symposium on Simulation, Visualization and Auralization for Acoustic Research and Education (Tokyo, Japan), The Acoustical Society of Japan (ASJ), April 1997, pp. 407–414.

[Her97] Jens Herder, *Tools and Widgets for Spatial Sound Authoring*, Compugraphics '97, Sixth International Conference on Computational Graphics and Visualization Techniques: Graphics in the Internet Age (Vilamoura, Portugal) (Harold P. Santo, ed.), GRASP, December 1997, ISBN 972-8342-02-0, pp. 87–95.

[Her98] Jens Herder, *Sound Spatialization Framework*, web site, University of Aizu, Japan, 1998, http://www-ci.u-aizu.ac.jp/SF/.

[MH99] William L. Martens and Jens Herder, *Perceptual criteria for eliminating reflectors and occluders from the rendering of environmental sound*, ICAD/ASA/EAA Workshop on Auditory Display (Berlin) (Durand R. Begault, ed.), Int. Community for Auditory Display (ICAD), Acoustical Society of America (ASA), and European Acoustics Association (EAA), March 1999, in press.

[Wer94] Josie Wernecke, *The Inventor Mentor*, Addison-Wesley, 1994, ISBN 0-201-62495-8.