Implementation issues of 3D audio in a virtual room

Jarmo Hiipakka[1,3], Tommi Ilmonen[2], Tapio Lokki[2], Matti Gröhn[2], Lauri Savioja[2]
Jarmo.Hiipakka@hut.fi

Helsinki University of Technology
[1]Laboratory of Acoustics and Audio Signal Processing
P.O.Box 3000, FIN-02015 HUT
[2]Telecommunications Software and Multimedia Laboratory
P.O.Box 5400, FIN-02015 HUT

[3]Nokia Research Center, Speech and Audio Systems Laboratory
P.O.Box 407, FIN-00045 NOKIA GROUP

ABSTRACT

This paper presents the audio system built for the virtual room at Helsinki University of Technology. First we discuss the general problems for multichannel sound reproduction caused by the construction of, and the equipment in virtual rooms. We also describe the acoustics of the room in question, and the effect of the back-projected screens and reverberation to the sound. Compensation of the spectral deficiencies and the problems with the large listening area and high frequency attenuation are introduced. The hardware configuration used for sound reproduction is shortly described. We also report the software applications and libraries built for sound signal processing and 3D sound reproduction.

Keywords: 3D sound, multichannel reproduction, virtual room, virtual acoustic environment.

1 INTRODUCTION

A virtual room is a fully immersive multiuser virtual reality system [1]. The visual display of the virtual room consists of large back-projected screens of size 3×3 m surrounding the users. Lightweight shutter glasses are used for a completely immersive stereoscopic view of the virtual world. A tracking device is used to monitor head position and orientation of one user so that correct stereo perspective can be obtained. If there are more than one user simultaneously present in the virtual room, the stereo perspective is exact only for the tracked observer. However, the perspective for other observers is not overly distorted, and real cooperation in the virtual world is possible. The user's movements in virtual reality and the interaction with the virtual world are usually handled using a wand (a 3D mouse) or a data glove.

Many virtual reality systems use only computer graphics to produce immersive virtual environments. In virtual reality research the main emphasis has always been on 3D graphics. However, sound is an important part of our everyday life and should be included in the creation of an immersive virtual environment. In virtual reality applications sound can either be environmental or background, or relate directly to the visualization. Background sound in virtual environment creation can be used to create a certain mood for the observer, e.g., background sound can be used for noise in a factory environment or for traffic noise in traffic simulations. With sound relating to the visualization we mean sound that is consistent with the created virtual world. The sound sources are often attached to corresponding objects and a realistic 3D soundscape is surrounding the observer. A good overview of 3D sound in virtual environments has been presented by Begault [2]. In this paper we concentrate on realistic soundscape creation by presenting the audio system built for the virtual room (Fig. 1) at Helsinki University of Technology (HUT) [3,4].

This paper is organized as follows: In section 2 we describe a generic sound rendering model for virtual reality applications. The section 3 concentrates on the acoustics of our virtual room and compensation of the spectral deficiencies introduced by loudspeaker placement and the back-projection screens. The audio reproduction system is introduced in section 4. The relevant applications are discussed in section 5 and section 6 concludes the paper.

Stereoscopic Displays and Virtual Reality Systems VIII, Andrew J. Woods, Mark T. Bolas
John O. Merritt, Stephen A. Benton, Editors, Proceedings of SPIE Vol. 4297 (2001)

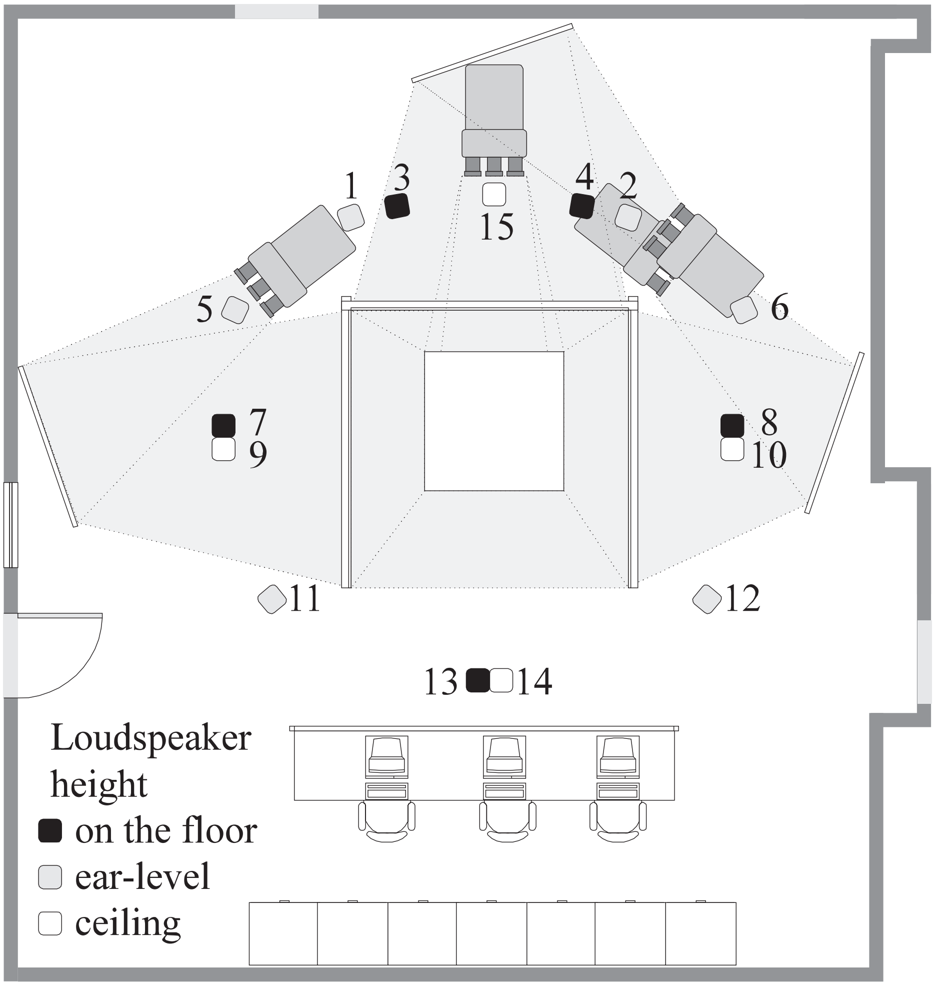


Figure 1: A schematic figure of the virtual room at HUT. Color indicates the height of the loudspeaker. Dark loudspeakers are on the floor, and light ones are ceiling-mounted.

2 AUDIO IN A VIRTUAL ENVIRONMENT

For an immersive virtual reality experience, a three-dimensional soundscape needs to be created; a sound reproduction system capable of positioning sound sources anywhere in the 3-D space is required, and for best level of fidelity, also the cues to the environmental context of the complete acoustic environment. For any spatial sound system used in virtual reality applications, there is of course an additional requirement: the system must interact with the graphics software so that the synchronization between sound and image is correct.

2.1 Sound Rendering Model for Virtual Reality

The process of producing immersive 3-D sound in virtual reality systems can be divided into three different stages. First of them deals with the stimulus, which can be natural or synthetic audio. In this context, natural audio means either pre-recorded or live digital audio. In virtual acoustic applications, recorded audio is not very useful because most of the recordings contain also the acoustical cues (e.g. reverberation) of the recording environment. A better choice is to use synthetic audio because it is more freely controllable. Although synthetic audio sounds unnatural in many cases it is useful in interactive applications where the properties of a sound source may vary according to the actions of the observer.

On the second stage, the propagation of the sound emitted from the sound sources is modeled. The rendering of the acoustical environmental context can be based on the physical properties of the virtual space or on the perceptual qualities of the desired soundscape [5].

For the physics-based approach, an acoustics simulation software has to model the propagation of the direct sound and the early reflections, taking into account the propagation delay, the absorption of the surface materials and air as well as the sound source directivity [6,7]. The spatial properties of the resulting sound can be controlled by moving the sound sources or the observer in the virtual space, by altering the size and shape of the space or by varying the acoustical properties of the surface materials.

The perceptual rendering model provides a more direct way of specifying the attributes describing the perceived acoustical properties of the sound sources and the acoustical space. Although the perceptual rendering model gives natural sounding results, it does not guarantee a physically feasible result related to the geometry of a specific space. An example of this approach is the Spat processor [8] developed at IRCAM.

The last of the three stages contains the spatial sound reproduction to the observer. There are two main spatial sound reproduction methods, namely binaural and multichannel reproduction. In binaural 3D sound reproduction techniques, the principle is to control the sound signal in the entrances of the listener's ear canals as accurately as possible. This requirement is easiest to fulfil with headphones but can be done also with a pair of loudspeakers. If headphones are used, the sound quality does not depend much on the acoustics of the surrounding room. However, one of the most important features in virtual rooms is that many users can simultaneously interact with the virtual environment and with each other. If headphones were used, this interaction could be disturbed. Another practical disadvantage with headphones is the dependency of computational capacity on the number of users. For every user, an individual soundscape has to be calculated and everyone needs his or her own head-tracking device to get a correct soundscape. Using a pair of loudspeakers, 3D sound can be created for one spot. Unfortunately, this does not suit the needs in a large scale multi-user virtual environment.

Multichannel sound reproduction is an attractive solution for 3-D sound in virtual rooms. With multichannel reproduction we avoid the need for individualized head-related transfer functions (HRTF) and head tracking. In multiuser applications the communication between users is easier than when using headphones. In multichannel reproduction, multiple loudspeakers are placed around the listener. With multichannel techniques, it is possible to reproduce sound signals naturally from correct directions. When the direction of the virtual sound source coincides with the direction of a real loudspeaker, the source direction is exact. When these directions do not coincide, different panning methods can be used. In our system we apply vector base amplitude panning (VBAP), which is a simple mathematical way to calculate the gain coefficients for the loudspeakers [9]. VBAP also allows for arbitrary loudspeaker placement which is a good feature in virtual rooms where mirrors and projectors hinder regular placement of the loudspeakers. Another often utilized multichannel spatial sound reproduction method is Ambisonics [10,11], which is suitable for creating background soundscapes because recording and coding methods are readily available for Ambisonics.

3 ACOUSTICS OF THE VIRTUAL ROOM

Spatial sound reproduction using loudspeakers in virtual rooms is far from being trivial. The quality of the resulting soundfield is subject to the quality of the equipment, to the applied processing techniques, and to the acoustics of the room. In virtual rooms there are back-projected screens and data projectors that hinder arbitrary positioning of loudspeakers, and typical back-projection screens are not acoustically transparent. Some deficiencies in the response of the reproduction system can be compensated for, but there are practical limits to this approach. Headphone reproduction for spatial audio solves these problems, but introduces other practical difficulties, as described previously.

To minimize the influence of the room acoustics, the room must be as anechoic as possible. This means that the walls and the ceiling must be covered with absorbent material. For the room to be silent enough, all noisy equipment should be placed in another room. For practical reasons, however, at least the data projectors are usually in the same room with the screens. The mirrors and the screens also produce reflections that may be disadvantageous for sound reproduction.

3.1 Room Acoustic Measurements

To investigate the acoustics of our virtual room at HUT, we conducted a series of impulse response measurements from all fifteen loudspeakers to nine microphone positions inside the virtual room. The height of all microphone positions was 160 cm, and the measuring points formed a square-shaped grid with 150 cm long sides with the central point in the middle of the virtual room. In the measurements we used the multi-channel IRMA measurement system [12] with MLS sequences as the source signal. At the time of the measurements, there was a data projector only for one of the walls, so we will have to repeat our measurements, when the system is fully constructed.

According to our measurements, the reverberation time measured using a broadband excitation in our virtual room is in the order of 400 ms. Figure 2 depicts the frequency contents calculated from short segments of the impulse responses from two different loudspeaker positions to the middle of the virtual room. Both loudspeakers are on the floor. In the first case (Fig. 2a), there is no screen between the loudspeaker and the microphone (loudspeaker number 13 in Fig. 1), but in the second case (Fig. 2b), there is a screen (loudspeaker number 8). Before analysis, the impulse responses have been truncated so that the direct sound is in the beginning of the response, i.e., the propagation delay has been removed from the responses. The distance from the loudspeaker to the microphone is approximately the same in each case. The different line styles in the figures denote the segment of the impulse response that has been used for response calculation. For clarity, the frequency response has been smoothed with ERB scale [13] resolution.

From the figure, the influence of the screen on sound can be easily seen. The higher frequencies of the direct sound are attenuated more than 10 dB when there is a screen between the loudspeaker and the microphone. The reverberant sound field will bypass the screen, i.e., the level of the reverberant sound is hardly at all influenced by the presence of the screen. The attenuation of the direct sound affects the localization of the sound sources and blurs directional cues. Also, the level of the highest frequencies is rather much reduced, and this makes the sound dull. The dampening effect can also be seen in the corresponding impulse responses presented in Fig. 3.

3.2 Spectral Compensation

When the simulated acoustics of a virtual environment is reproduced using loudspeakers, the room response has to be equalized so that desired acoustical conditions are faithfully reproduced. Room equalization in the whole audible frequency range

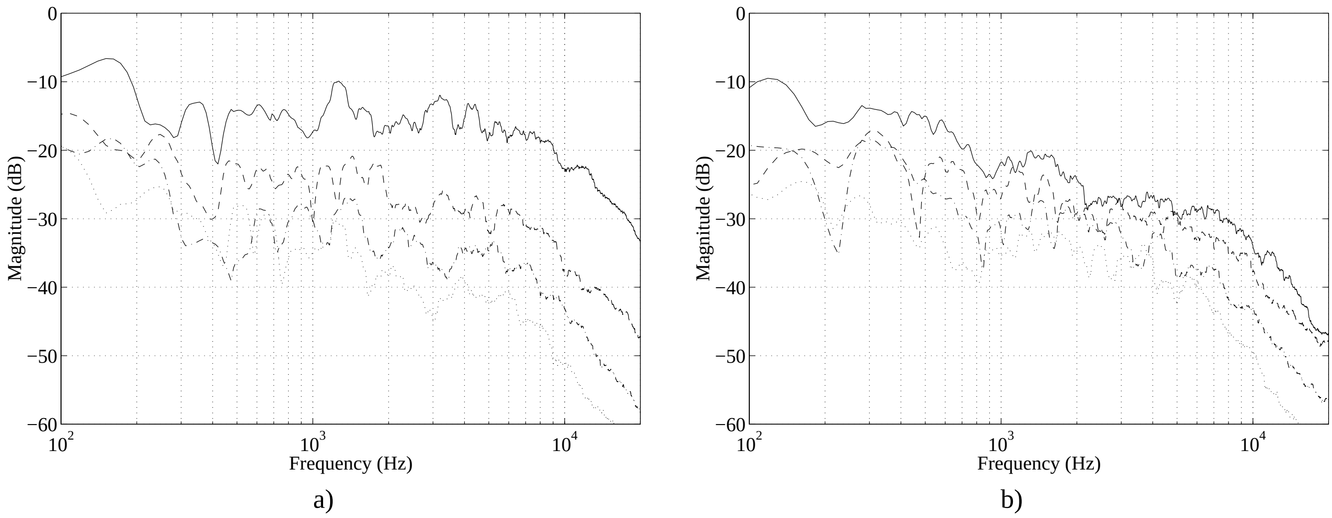
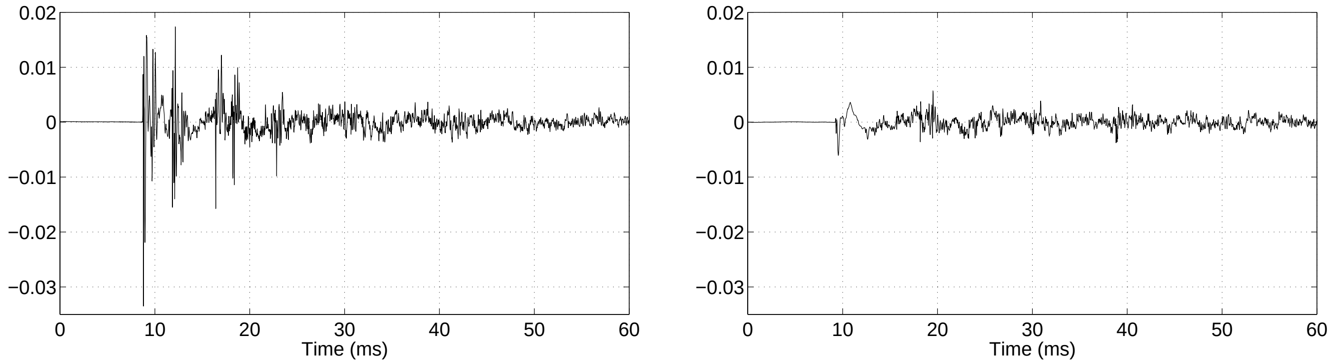


Figure 2: Frequency responses calculated from several 23.2 ms segments of two impulse responses. The starting times of the segments are: 0 ms for the solid line, 30 ms for the dashed line, 60 ms for the dashdot line, and 90 ms for the dotted line. The responses in figure a) are for the loudspeaker on the floor in the back of the virtual room (without a screen), and those in figure b) are for the loudspeaker on the floor to the right of the virtual room.



a) b)

Figure 3: Impulse responses from two different loudspeakers to the center of the virtual room. Note the dampening effect of the back-projected screen in the beginning of the response presented in (b).

is an impossible task when we have multiple loudspeakers and the listener is moving. (Room equalization has been discussed in more detail, e.g., in [14,15]).

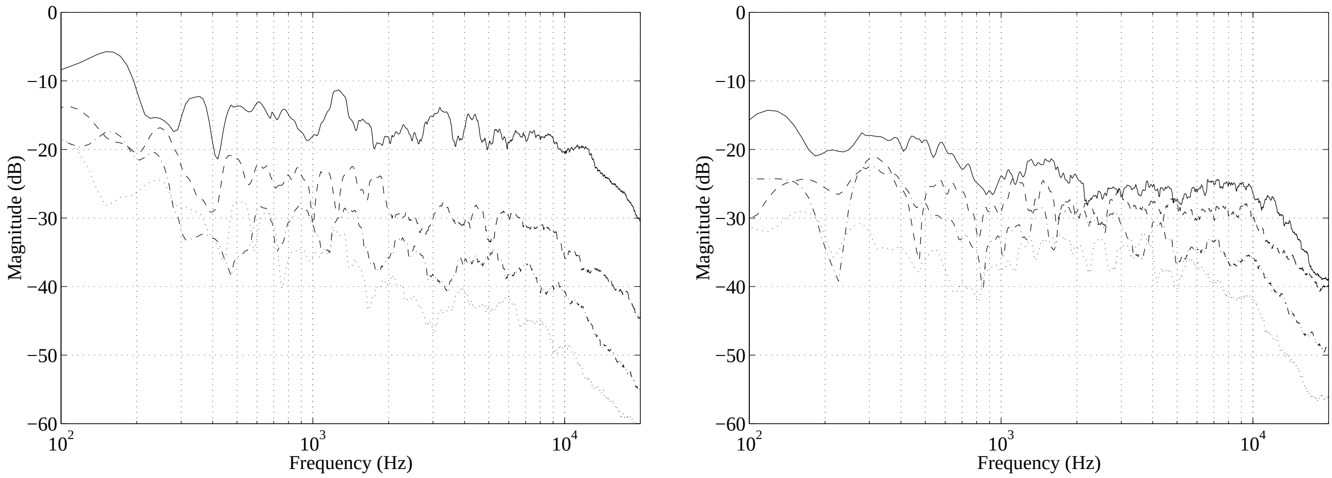
Filters that perform spectral compensation for the screens can be designed automatically by using, e.g., linear prediction. With automated design methods, it is in principle possible to produce a very flat frequency response from the loudspeakers to the listening area. In our case, the difficulty comes from the fact that high frequencies are rather much attenuated. If we attempt to achieve a flat frequency response, the level of the high frequencies needs to be very high even when using moderate overall sound pressure levels. This is not practical with normal loudspeakers; that is why we want only to partially compensate for the high-frequency attenuation.

For compensation filter design, we first power-averaged the frequency responses from one loudspeaker to all nine microphone positions;

$$|H_{k,\,avg}(\omega)| = \sqrt{\frac{1}{M}\sum_{i=1}^{M}|H_{k,\,i}(\omega)|^2}\,, \tag{1}$$

where M is the number of microphone positions, k is the index of the loudspeaker, and $H_{k,\,i}$ is the complex frequency response from loudspeaker k to microphone position i. Additionally, the averaged response was smoothed using ERB scale resolution, and the response was resampled using 100 frequency points equidistant on the ERB scale from 100 Hz to 10 kHz. Naturally, this is done for all fifteen loudspeakers individually. As a next step we defined the desired frequency response that is attenuated by 10 dB at 10 kHz with respect to the level at 100 Hz, and descends linearly (on dB scale and logarithmic frequency scale) in between. After determining the averaged responses and defining the desired response, we calculated the error between the two and used this error as an input for filter design.

After trying out different filter design techniques, we chose IIR filters of order 10 designed using the Yule-Walker method (function yulewalk in MATLAB [16]) for practical implementation of spectral compensation. There are a few reasons for selecting such a small filter order. The computational requirements for the filters must be very modest; the compensation has to be done for fifteen channels, and the computer must also be able to simultaneously run the auralization engine in real-time. Additionally, the listening area inside the virtual room is so large that it is not feasible to try to compensate the response very accurately. The chosen equalization method provides us with sufficiently uniform timbre across the whole listening area. Figure 4 shows equalized responses that correspond to the unequalized responses presented in Fig. 2. It can be seen that the effect of the compensation filter is quite small on the responses without a screen between the loudspeaker and the microphone (Fig. 4a), whereas the responses with the screen (Fig. 4b) are considerably altered.



a) b)

Figure 4: Frequency responses calculated from segments of the compensated impulse responses. The loudspeakers and the segment boundaries are the same as those in figure 2.

In our case, it is practically impossible to implement the digital compensation so that the level of the direct sound would be increased with respect to the reflections; the only way to achieve this goal is to decrease the energy of the reflections by adding more absorbent material to the walls of the room.

4 SYSTEM IMPLEMENTATION

The main sound processing blocks used for spatial audio reproduction in the virtual room are depicted in Fig 5. The control input for the audio system comes from the visualization and user interaction. The input from the visualization is used to calculate the locations of the sound sources and the early reflections in the virtual world. The sources are processed using digital filters for material and air absorption, and directivity (see [6] for further details). Then the sources and reflections are panned for the specific loudspeaker arrangement. In addition to the early reflections we have implemented the late reverberation using efficient recursive methods. Both the early reflections and the reverberant sound are equalized for the influence of the screens, as described in section 3.2. The input sound to the room acoustic modeling comes from a real-time multichannel input or from one or more audio files, and the output is sent via the multichannel sound card to digital-to-analog converters and speakers. In addition to the auralized sound, we have a consumer grade sound card in our system for voice command-and-control purposes, so that the user in the virtual room can use spoken command phrases to control the system.

4.1 Sound Reproduction Hardware

The hardware for sound reproduction in the virtual room is built around one dual-processor PC computer running Linux operating system. This computer acts as a central sound processing unit. The computer runs special software that is used for room acoustic modeling and auralization, sound source panning, and equalization filtering.

The problems with using PC hardware and Linux for sound processing are mostly due to the fact that the support for advanced multi-channel sound cards on Linux is on quite an immature stage. First we planned to use a Sonorus Studi/o card, but the Linux drivers for this card were unusable for our purposes. As the second alternative we have successfully used RME's Digi9652 audio interface card with ALSA [17] drivers. Despite all the problems, the overall performance on inexpensive low-latency Linux systems outperforms many other computers and operating systems.

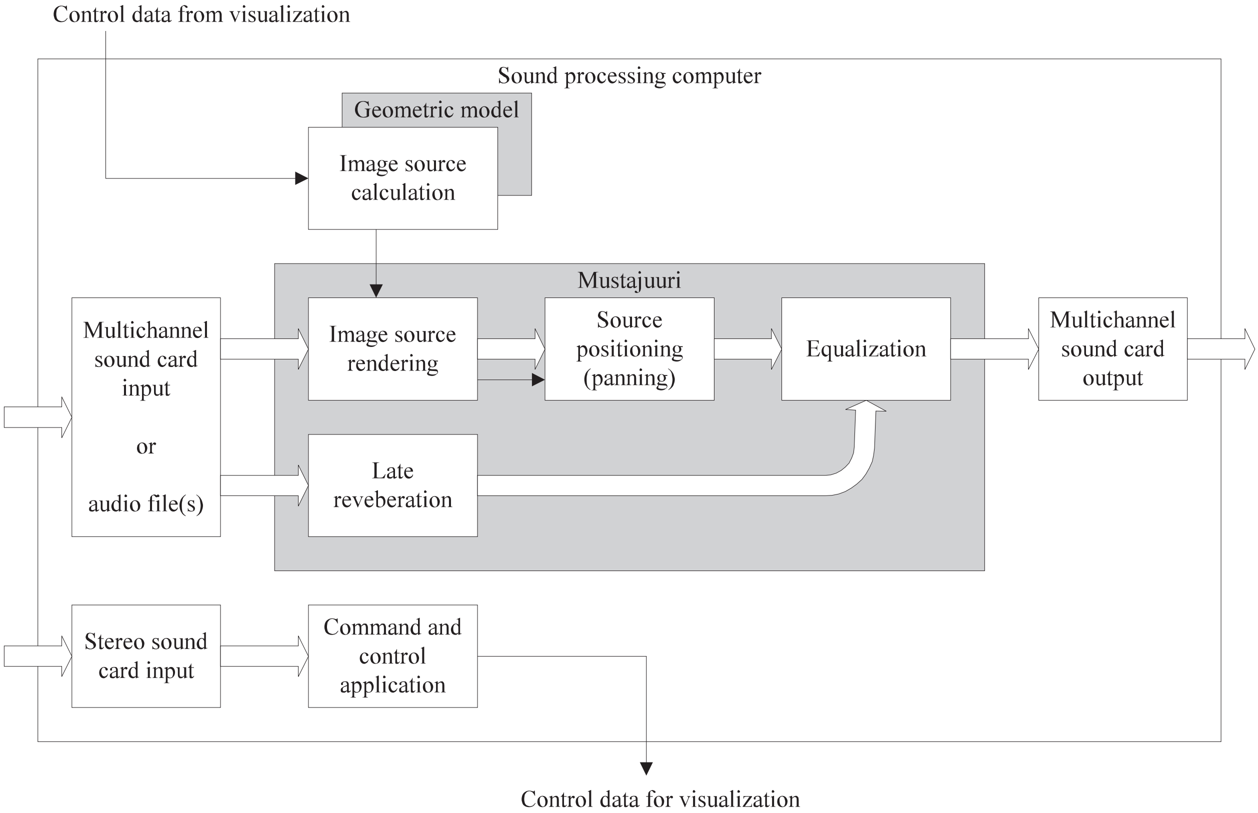


Figure 5: Sound processing blocks and data flow for spatial audio reproduction in the virtual room at HUT. Large arrows are used for sound signals, and smaller for control data.

As sound sources we currently use audio files and sound synthesis applications running on the central audio processing PC or on other hosts. The interconnections between computers and other audio equipment can be changed in a flexible manner, because we use only standard audio cables to connect the equipment together, and we have written all the software ourselves.

Sound output from the Linux PC is taken from two eight-channel ADAT interfaces that are connected to two eight-channel digital-to-analog converters. The current loudspeaker system in the virtual room consists of fifteen active Genelec monitoring loudspeakers. We have plans to include subwoofer(s) in the system, because high-quality bass reproduction is vital for the immersion in many applications.

The demands that the command-and-control application sets to the sound quality are not nearly as strict as they are for the rest of the system. In fact, the speech recognition engine runs with a lower sampling rate (max. 22,050 Hz) than the auralization (44,100–48,000 Hz). Because the voice input of the command-and-control system need not be in sample-accurate synchronization with the auralization system, we use another, lower quality sound card for the voice input. This way the speech recognition software can be run fully isolated. This also frees us from using processing power for unnecessary sampling rate conversion.

4.2 Software Implementation

All software used for signal processing has been written with C++ language using object-oriented programming approach. The software is split to efficient low-level signal processing library and higher level dynamic signal processing/routing toolkit.

For efficient signal processing needed for the virtual room, we have implemented a low-level signal processing library DIVADSP. The library contains generic signal processing blocks, and optimized versions of several DSP structures. The units are combined on code level, and new combinations generally require recompilation of the application or the DIVADSP library. This design choice favors performance over flexibility. The DSP library contains implementations for several types of typical filter structures. Generic filter classes can be used as filters of any order, and for sake of efficiency, special implementations have been written for the most used filters of first, second, and third order. Auxiliary functions can be used to calculate filter coefficients using a few common filter design methods.

To complement the low-level signal processing library and to make the sound system as versatile as possible, a flexible high-level signal processing tool is required. An in-house application—called Mustajuuri—is a generic plugin-based real-time tool that suits this purpose. With proper plugins this application is able to run the audio processing for nearly any 3D sound application. Using Mustajuuri, we have implemented the complete audio signal processing chain containing sound synthesis, room acoustic modeling, panning, and equalization for our virtual room. A version of Mustajuuri is available at <URL:`http://www.tml.hut.fi/~tilmonen/mustajuuri/`>.

Mustajuuri is a highly modular signal processing platform. It is built of plugin authoring libraries and a real-time engine. The base libraries act as a software development kit (SDK), while the real-time engine takes care of running the signal processing with low latency in a separate DSP thread. Plugin is a signal processing block. There are no hard limits to how simple or complex a plugin can be. The system has been tuned for plugins that are either medium in complexity (compressor, equalizer) or complex (synthesizers, MIDI sequencers). Modularity is also a natural way to handle platform-dependent features such as audio I/O and MIDI I/O.

A DSP network is created by connecting plugins together in run-time. Since most of the functionality comes from the plugins, the system can be easily extended. New plugin types can be written in C++ and they are loaded by the application without recompilation. By nature, all plugins can deal with audio signals and with events (for example MIDI events). The event interface is also used to transmit parameters to the plugins. While the most common parameters such as numeric values and text strings are directly supported, the user can create new parameter types if needed.

For this project, the customizability of Mustajuuri has been a deciding factor. We have been able to incorporate plugins for flexible auralization into the software. By creating proper plugins one can control the software in a novel way—it is even possible to directly access the internal data structures of the application. Since all the source code is available, one has total control over the internals of the software. As the software is plugin-based, we do not need to do modifications to the application when adding or modifying features.

By nature, Mustajuuri is an application that is used locally (the user and the software on the same machine). In virtual reality applications, graphics and sound are often processed in different computers. For this reason, system components must communicate via network. For the needs of the virtual room, we created an auralization control plugin, that adds auralization server features to Mustajuuri. This server is controlled via a socket interface.

We have minimized the work that is needed to add auralized sound to applications. A lightweight C++ library is used to hide the socket interface from the client application. This library contains the logic that is needed to control the signal processing. The normal way to use this library is like one would use a local sound-processing library: First the system is configured and initialized for the task at hand and then the system is controlled by sending parameter changes to it.

5 APPLICATIONS

A virtual room can provide high degree of immersion for a wide spectrum of virtual environments. Eckel [18] has listed the following application areas which can gain benefit from the use of a virtual room:

- scientific visualization and sonification [19],
- industrial data exploration,
- product presentation and marketing,
- product design and evaluation,

- architectural planning and walk through,
- exhibition design and evaluation,
- media art installations.

All the application areas listed above could also use related 3D sound for greater impact. The main application for spatial audio in the HUT virtual room has been auralization combined with architectural visualization [20]. In the field of media art installations, we have created real-time virtual orchestra performances [21]. In the future we will continue the research on voice input in virtual reality environments, visualization and sonification of building services data [22], and the use of sound in navigation in virtual environments [23].

6 CONCLUSIONS

In this paper we have described implementation issues of 3D audio for a virtual room. We have presented a sound rendering model for immersive audio used in virtual rooms. Also the acoustical issues including the non-optimal loudspeaker positioning and the effect of the screens on sound quality were overviewed. The virtual room is not an optimal listening environment but with measurement and compensation the problems can be reduced. In this paper we have presented responses measured in the virtual room at HUT. These responses clearly show the effect of the back-projection screens on sound. In the virtual room at HUT, we have chosen to improve the responses by digital equalization filters. The techniques we have used allow us to compensate for the spectral deficiencies, but the ratio between the direct sound and the reflections can only be increased by adding more absorption to the room.

We have also presented the system implementation for spatial audio reproduction in the HUT virtual room. The core component in our system is a dual-processor computer running Linux and custom software for real-time room acoustic modeling and auralization. The software implementation is divided into low and high level signal processing components. The low level implementation emphasizes efficiency, whereas the main feature of the higher level system is flexibility. The flexibility is attained using a plugin-based approach. Future research will consist of both fundamental research on spatial audio and its applications.

ACKNOWLEDGMENTS

This work has been partly supported by the Helsinki Graduate School in Computer Science, the Academy of Finland, and Tekes, the National Technology Agency.

REFERENCES

1. C. Cruz-Neira, D. Sandin, T. DeFanti, R. Kenyon, and J. Hart, “The Cave – Audio Visual Experience Automatic Virtual Environment”, *Communications of ACM*, 35(6), pp. 64–72, 1992.
2. D. Begault, *3-D Sound for Virtual Reality and Multimedia*, Academic Press, Cambridge, MA, p. 293, 1994.
3. J. Jalkanen, *Building a Spatially Immersive Display: HUTCAVE*, Licentiate Thesis, Helsinki University of Technology, 4 March 2000, p. 132.
4. J. Jalkanen, and H. Napari, “How to build a virtual room”, *In Proc. SPIE Vol. 4297B*, San Jose 2001.
5. J.-M. Jot, “Synthesizing Three-Dimensional Sound Scenes in Audio or Multimedia Production and Interactive Human-Computer Interfaces”, *In Proc. 5th International Conference: Interface to Real & Virtual Worlds*, Montpellier, France, May 1996.
6. L. Savioja, J. Huopaniemi, T. Lokki, and R. Väänänen, "Creating Interactive Virtual Acoustic Environments", *Journal of the Audio Engineering Society*,47(9), pp. 675-705, September 1999.
7. J. Huopaniemi, L. Savioja, and M. Karjalainen, “Modeling of Reflections and Air Absorption in Acoustical Spaces: A Digital Filter Design Approach”, *In Proc. 1997 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA’97)*, Mohonk, New Paltz, New York, October 19-22, 1997.
8. J.-M. Jot, “Efficient Models for Reverberation and Distance Rendering in Computer Music and Virtual Audio Reality”, *In Proc. 1997 International Computer Music Conference (ICMC97)*, pp. 236-243, Thessaloniki, Greece, September 1997
9. V. Pulkki, “Virtual Sound Source Positioning Using Vector Base Amplitude Panning”, *Journal of the Audio Engineering Society*, 45(6), pp. 456-466, 1997.

10. M. Gerzon, “Periphony: With-Height Sound Reproduction”, *Journal of the Audio Engineering Society*, 21(1/2), pp. 2-10, 1973.
11. D. Malham, and A. Myatt, “3-D Sound Spatialization Using Ambisonics Techniques”, *Computer Music Journal*, 19(4), pp. 58-70, 1995.
12. T. Peltonen, *A Multichannel Measurement System for Room Acoustics Analysis*, Master’s thesis, Helsinki University of Technology, 2000, p. 119.
13. B. C. J. Moore, R. W. Peters, and B. R. Glasberg, “Auditory filter shapes at low center frequencies”, *Journal of Acoustical Society of America*, vol. 88, pp. 335–345, 1990.
14. R. Walker, “Equalisation of Room Acoustics and Adaptive Systems in the Equalisation of Small Room Acoustics”, *In Proc. AES 15th Int. Conf. Copenhagen*, pp. 32–47, Denmark, 1998.
15. J. N. Mourjopoulos, “Digital Equalization of Room Acoustics”, *Journal of the Audio Engineering Society*, 42(11), pp. 884-900, 1994.
16. Mathworks, *MATLAB Signal Processing Toolbox*. User’s Manual, 1994.
17. Advanced Linux Sound Architecture, see <URL:`http//www.alsa-project.org/`>.
18. G. A. Eckel, “Spatial Auditory Display for the CyberStage”, *In Proc. Int. Conf. Auditory Display (ICAD’98)*, Glasgow, Scotland, November 1-4, 1998.
19. G. Kramer, B. Walker, T. Bonebright, P. Cook, J. Flowers, N. Miner, J. Neuhoff, R. Bargar, S. Barrass, J. Berger, G. Evreinov, M. Gröhn, S. Handel, H. Kaper, H. Levkowitz, S. Lodha, B. Shinn-Cunningham, M. Simoni, W. Tecumseh Fitch, and S. Tipei, *Sonification Report: Status of the Field and Research Agenda*, ICAD, 1999, p. 38.
20. T. Takala, E. Rousku, T. Lokki, L. Savioja, J. Huopaniemi, R. Väänänen, V. Pulkki, and P. Salminen, “Marienkirche – A Visual and Aural Demonstration Film”, *Electronic Art and Animation Catalogue (SIGGRAPH’98)*, pp. 149, Presented in SIGGRAPH’98 Computer Animation Festival (Electronic Theater), Orlando, Florida, USA, Jul. 19-24, 1998.
21. T. Lokki, J. Hiipakka, R. Hänninen, T. Ilmonen, L. Savioja, and T. Takala, “Real-time audiovisual rendering and contemporary audiovisual art”, *Organised Sound*, 3(3), pp. 219–233, 1998.
22. M. Gröhn, M. Laakso, M. Mantere, and T. Takala, "3D Visualization of Building Services in Virtual Environment", *In Proc. SPIE Vol. 4297B*, San Jose 2001,
23. T. Lokki, M. Gröhn, L. Savioja, and T. Takala, “A Case Study of Auditory Navigation in Virtual Acoustic Environments”, *In Proc. Int. Conf. Auditory Display (ICAD’2000)*, pp. 145-150, Atlanta, GA, USA, April 2-5, 2000.