An interactive and real-time based auralization system for room acoustics, implementing directional impulse responses and multiple audio reproduction modules for spatialization (the AURALIAS project)

L. Bos[1], J.J. Embrechts[2]

[1] *University of Liege, Intelsig group, Sart-Tilman B28, 4000 Liege, Belgium, Email: bos@montefiore.ulg.ac.be*
[2] *Same affiliation, Email: jjembrechts@ulg.ac.be*

Introduction

Auralization is "the technique of creating audible sound files from (simulated, measured, or synthesized) data" [1]. In the field of room acoustics predictions, the simulated data consist in room impulse responses (RIR) or directional room impulse response (DRIR) [2].

Several audio reproduction systems can be used for creating these virtual sound fields: some are "lightweight", attached-to-the-user systems, like the reproduction through headphones including the effect of HRTFs (head-related transfer functions), while some others are more expensive in terms of resources and materials. Among this last category are the systems based on multiple loudspeakers arranged in a specific environment, like an anechoic room, an immersion studio or a CAVE (see [1,3] for more details).

At the beginning, auralization systems were mainly developed conjointly with room acoustics programs. From the 90's, the most famous of them were equipped with an auralization module based on headphones' reproduction. But, as the computers became more and more powerful, the feeling of immersion was improved by the coupling of the auralization with the visualization of the virtual room, allowing the displacement of the listener and the sound source [1,4]. This enhanced complexity requires *real-time auralization*, including very fast convolution algorithms between RIRs and the input anechoic signals.

Also, the quality of auralization was improved, from the crude single channel reproduction of the reverberation in the virtual room, to more accurate localizations of the direct sound, the early reflections and (when possible) the late part of the reverberation [2].

It is the objective of the AURALIAS project (2007-2010), to develop an interactive and real-time auralization system for room acoustics.

The AURALIAS research project [5]

AURALIAS aims at developing an interactive tool for architects, acousticians and more generally all people involved in a room acoustics project (including, why not, the final user). The system must be able to immerse the "listener" in a virtual sound field, looking at an image of the virtual room and interacting with it through an appropriate interface.

It can be a single-user application, but it has more preferably been thought as a collaborative experience between a restricted number of persons (for example, between the acoustician and the architect). In this respect, the immersion studio is presently equipped with six loudspeakers in a horizontal plane (the stereo pair at ± 30 deg. from the frontal viewing direction and four more loudspeakers every 60 deg., in a nearly circular arrangement). It is intended to complete this horizontal structure with some loudspeakers in elevation, in front of the users. All these loudspeakers are fed with their own auralized signal computed by the VBAP technique [6].

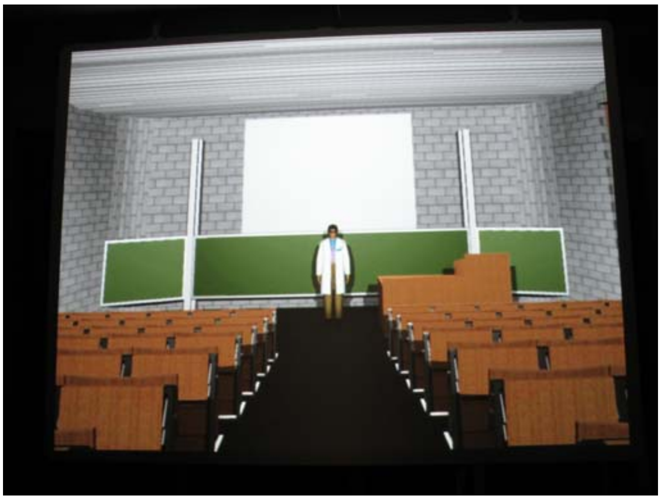

Figure 1: View of the screen in the immersion studio. The users are looking to an image of the virtual room. The front loudspeakers (not apparent on this figure) are situated immediately on the left and right of the screen

The users of the studio are placed in the middle of the loudspeakers' circle and they look at the screen in front of them, on which an image of the virtual room is projected: see figure 1. They can interact with this projected view, for example by displacing their virtual position in the room. The auralized signals are then changed accordingly, in real-time.

As the system is conceived for professional applications, especially for architects and acousticians, it is intended to keep the quality of auralization as high as possible, for the largest category of rooms that can be met in practice. In this respect, an accurate localization of the direct sound and the early specular contributions is of course an important issue, but also the rendering of all directional characteristics of the sound field. Complete directional impulse responses are therefore computed around the listener and each of them is distributed to the appropriate loudspeakers for reproduction. This allows for the auralization of rooms with special geometries, such as long disproportionate rooms in which flutter echoes in certain directions can dominate the late part of the RIR.

Finally, interaction with the general system is provided by an interface, which must be as intuitive and user-friendly as possible. One of them being tested at the moment is based on a ultra-mobile PC (UMPC) interacting with the auralization and the visualization systems through a wireless link.

The room acoustics model

Auralization systems are based on simulated or measured (D)RIRs. The room acoustics' model used in AURALIAS is a ray-tracing program developed by the Applied Acoustics' team of the university of Liege [7,8]. This program includes most options offered by modern room acoustics' programs (including wideband diffusion [8]) and it has been recently updated to compute directional echograms [2].

In this respect, each receiving position in the modelled room is surrounded by a transparent spherical receptor, divided into a given number of solid angles covering the entire sphere. Usual choices are 6 (up-down, left-right, and front-rear) or 26 solid angles. For this last choice, the solid angles' extension is 45 degrees in azimuth and elevation, including one cone ($\pm$22.5 deg.) at the top and one similar at the bottom. Each solid angle of each receptor records the energy of the incident sound rays, which leads to an echogram per solid angle. This process is repeated for each source position, which finally gives one echogram per solid angle, per frequency (octave) band and per source position.

Finally, these directional echograms (see figure 2) are converted into directional RIRs. Echograms (which are energy-time responses) give the impulse responses' envelope, and the conversion algorithm is adding phase information through "an adequate fine structure representing the actual reflections statistics" [1, p.225].

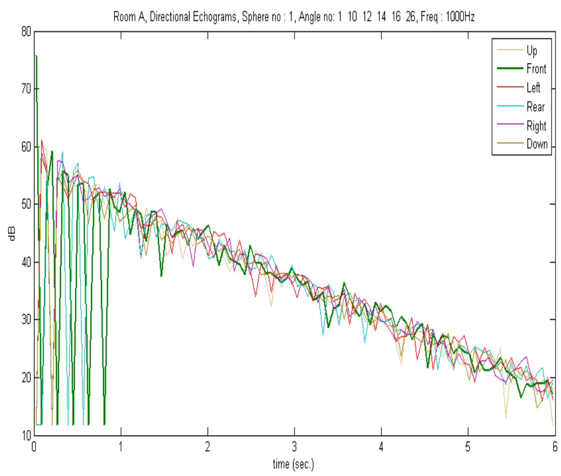

Figure 2: Directional echograms computed in the 1 kHz octave-band. All reflections are specular.

To improve localization cues, image sources can be computed up to a specified order, for reflecting surfaces being fully or partially specular. Their contribution is therefore removed from the directional echograms, while they are used to derive specific "filters" or impulse responses which can be precisely oriented (in azimuth and elevation) around the virtual listener.

Initially, the ray-tracing program was not conceived for real-time applications. However, as real-time interaction is an important issue for AURALIAS, the algorithm is presently updated to allow for fast modifications of some characteristics of the virtual room:

- the displacement of the sources and the receiver are presently based on pre-computed echograms, at pre-defined locations specified by the user. It would be desirable to provide a fast re-evaluation of these echograms for any position of sources and receiver;
- it is also intended to allow for the real-time modification of the surfaces' acoustical properties and of some geometrical parameters.

The auralization filter

The auralization filter deals both with sound reverberation and sound spatialization.

Convolution and reverberation

The reverberation part is currently implemented with a frequency block segmented convolution based on the overlap-add method. According to this method, the filter response is split into several uniform blocks. These blocks are then transformed into the frequency domain by the FFT method, multiplied with a block of input samples (previously converted in the frequency domain too), and finally transformed back to the time domain. This method is less computational intensive than the direct "temporal" convolution, but it introduces an input-output delay at least equal to the length one block. Reducing this latency by using smaller blocks leads to increase the computational load. To deal with this dilemma, the method can be improved by splitting the filter response into non uniform block lengths, using smaller blocks at the beginning of the filter to reduce the latency and larger blocks at the end to keep the computational load acceptable[9, 10]. It's also possible to obtain a zero delay convolution by convolving the first block in the time domain [10].

However, it appears nowadays that the computational load associated to a specific reverberation algorithm may not be the main criteria for real time applications any more. Indeed, the multi-core architecture of modern Central Processor Units (CPU) allows for parallel tasks and offers more and more powerful instructions. These methods have then to be also judged on their ability to take advantage of recent CPUs architectures. This remains of course a large field of investigations, but it seems quite obvious that it's easier to parallelize a simplest algorithm than a more complicated one!

Spatialization

The sound spatialization is achieved by using the Vector Based Amplitude Panning (VBAP) method [6, 11]. This method allows creating virtual sound sources by distributing the same monophonic signal on several loudspeakers. The position of a virtual sound source then depends on the amplitude of the signal applied to each loudspeaker. As previously mentioned in this paper, a two dimensional spatialization is currently obtained with a 6 loudspeakers configuration and is about to be completed soon to obtain a three dimensional restitution.

Acoustical properties of the immersion studio

It was decided at the beginning of the project not to auralize in an anechoic room, but in a listening studio. The reverberation time recommended for such a space is about 0.3 to 0.4 s, with little dependence on frequency. It is also recommended to attenuate early reflections reaching the listener, such that their difference with the direct field is at least 10 dB [12,13].

A volume of at least 40m³ should be used. A rectangular room with the proportions 1:1.6:2.4 is known to offer an adequate distribution of modal frequencies. Listening rooms for multichannel sound should further fulfil the following conditions (Length, Width, Height):

$1,1\ W/H \leq L/H \leq 4.5\ (W/H)-4$,

$L/H < 3$ and $W/H < 3$.

All these geometrical conditions are fulfilled by our immersion studio. Furthermore, the impacts on the walls of the first order reflections from the six loudspeakers to the central position of the listener have been identified. Reflecting panels have been placed along the walls and oriented in such a way as to redirect the first reflections outside the listening area, where possible. If not possible, absorbing panels have been used to cover the corresponding parts of the walls (mainly the lateral and front walls) and diffusing elements have been placed on the rear wall (figure 3).

Figure 3: Diffusers placed on the rear wall.

The floor has been covered with a thick carpet, and the micro-perforated thin screen (on which the images are projected) has not shown any significant acoustical influence.

The reverberation times measured in the studio are shown in figure 4. The clarity C80 is comprised between 15 and 20 dB. The definition D50 is greater than 80% at 125 Hz and greater than 90% at all other frequencies. Finally, the measured RASTI is comprised between 0.85 and 0.90.

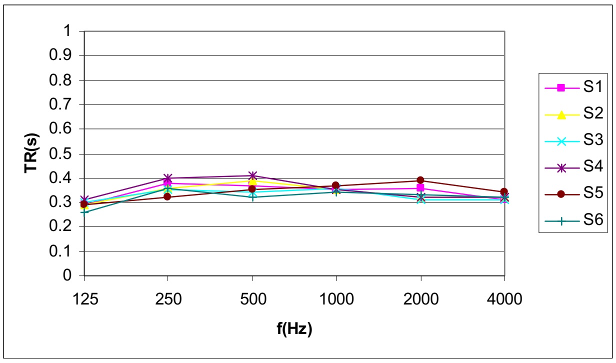

Figure 4: Reverberation times measured in the studio at the central position of the listener, and for the six positions of the loudspeakers.

Features of the first prototype

Hardware

The first prototype runs on the standard PC under Windows XP 32 bits whose main specifications are:

- CPU : Intel 2.4GHz core 2 quad
- GPU : NVIDIA GeForce 8500 GT
- Ram : 3Go

A *Toshiba TDP – EX 20* data projector is used for the projection of virtual rooms in front of the users.

The sound restitution is provided by six *Far XMD range Digital active three way* loudspeakers, those loudspeakers being driven by an *EDIROL AudioCapture FA – 101* sound card.

The user interface is deported on a *Samsung Q1 Ultra* UMPC. To improve the quality of immersion, this device should be completed (or replaced) in the near future by a joystick.

Possible actions by the user

The first prototype allows the auralization of an isotropic and monophonic sound source. The user can turn this source on and off, he can control its audio volume and choose its anechoic message. The virtual listener can move to pre-defined positions in the virtual room. Those positions (and that of the source) have been chosen during a previous step of room acoustics simulation, since it's not yet possible to compute long DRIRs in real time. During the auralization session, the adequate DRIRs are then extracted from a filterbank according to the current position of the listener. This also allows to switch between different virtual rooms during the auralization. Finally, it's also possible to choose the orientation of the listener in the horizontal plane and to record the audio output for playback.

Further developments

The next prototype will offer new features concerning the user's interaction in the auralized rooms and both the sound spatialization and reverberation methods.

It will then be possible to add several sound sources and to move them to pre-defined positions. It should also be possible to auralize the user own speech and to partially modify the auralized room on the fly. Room's modifications could affect both the acoustic properties of materials and the geometry of the room.

The current two dimensional VBAP sound spacialization module will be first completed by a three dimensional one and then by some others spatialization methods like headphones HRTF sound reproduction. In order to compare them, the software architecture will enable to switch from one spatialisation mode to another at running time.

As most sound spatialization methods suffer from a quite small sweet spot, a user tracking device will also be implemented to adapt the sound to the real user's position in the immersion studio.

The reverberation will also be more deeply investigated by implementing different convolution modules and by allowing the user to switch between them at running time.

Acknowledgements

The authors wish to thank Ir Fabienne Duthoit for her contributions in the acoustical study of the immersion studio.

AURALIAS is a project funded by the Region of Wallonia (Belgium) under contract n. 616416.

References

[1] M. Vorländer: Auralization. Springer-Verlag, Berlin, 2008

[2] J.J. Embrechts, N. Werner, S. Lesoinne: Computation of Directional Impulse Responses in Rooms for Better Auralization. Convention paper n.6498 of the 118th convention of the Audio Eng. Soc., Barcelona, 2005

[3] M. Kleiner, B.-I. Dalenbäck, P. Svensson: Auralization-An Overview. J. Audio Eng. Soc. **41** (1993), 861-875

[4] Reference to the DIVA and EVE homepage. URL: `http://eve.hut.fi/`

[5] Reference to the AURALIAS homepage. URL: `http://www.auralias.be`

[6] V. Pulkki: Virtual sound source positioning using vector base amplitude panning. Journal of the Audio Engineering Society **45** (1997), 456-466

[7] J.J. Embrechts: Sound field distribution using randomly traced sound ray techniques. Acustica **51** (1982), 288–295

[8] J.J. Embrechts: Broad spectrum diffusion model for room acoustics ray-tracing algorithms. Journal of the Acoustical Society of America **107** (2000), 2068-2081

[9] W. Gardner: Efficient convolution without input-output delay. Journal of the Audio Engineering Society **43** (1995), 127-136

[10] G. Garcia: Optimal Filter Partition for Efficient Convolution with Short Input/Output Delay. Proceedings of the 108th International Conference of the Audio Engineering Society, n.5660, 2002

[11] V. Pulkki, T. Lokki: Creating auditory displays with multiple loudspeakers using VBAP: A case study with DIVA project. International Conference on Auditory Displays, ICAD, Glasgow, 1998

[12] P. Newel: Recording Spaces. Focal Press, 1998.

[13] B. Walker: Room Acoustics for Multichannel listening: Early reflection control. Proceedings of the 22nd Audio Eng. Soc. UK Conference, Cambridge, 2007