

A moving listener in a virtual audio environment with particularly irregular reverberation decay

Jean-Jacques Embrechts, Stephane Lesoinne and Nicolas Werner

University of Liege

Dep. Electrical Engineering and Computer Science, Intelsig group

Sart-Tilman Building B28, B-4000 Liege

Belgium

`jjembrechts@ulg.ac.be`

Abstract

Auralization in a virtual room or space is realised by the convolution of an anechoic signal (i.e. music or speech signal) with the acoustical room impulse response (RIR). The RIRs (possibly left and right RIRs in the case of binaural reproduction) are constant, unless the source and/or the listener are moving in the virtual space. In that case, it becomes difficult to perform real-time auralization, since a periodic re-initialisation of the RIRs is required, in addition to the convolution task. To speed up the process of auralization and to reach real-time performances, the late part of the RIRs can be approximated by an exponential decay. However, this assumption cannot be justified in some particular audio environments, for example in disproportionate rooms and/or rooms with asymmetrical reverberation properties. In this paper, a method allowing the displacement of the listener in a virtual audio space is presented, which respects the characteristics and variability of the "late reverberation" decay. This method is based on the interpolation between echograms computed at some pre-defined locations in the virtual space.

1 Introduction

Auralization in a virtual room or space is realised by the convolution of an anechoic signal (i.e. music or speech signal) with the acoustical room impulse response (RIR). Headphone auralization can be monaural (the same convolved signal is sent to both ears) or binaural (left and right signals are reproduced separately, giving an impression of sound localization and reverberation) [1, 2].

If the source and the listener are static in the virtual space, the RIRs (possibly left and right RIRs in the case of binaural reproduction) are constant. However, if one of them is moving, then the RIRs constantly change. It becomes therefore difficult to perform real-time auralization, since a periodic re-initialisation of the RIR is required, in addition to the convolution task.

A solution consists in the identification of three components in the RIR : the direct sound, the early reflections and the "late reverberation". To speed up the process of auralization and to reach real-time performances, the early reflections are modelled by mirror sources (up to a limited order of reflection) and the "late reverberation" contribution is approximated by an exponential decay, independent of the source and the listener's locations [1, 2]. This assumption can be justified in diffuse sound fields,

but this cannot be the case in some particular audio environments, for example in disproportionate rooms and/or rooms with asymmetrical reverberation properties [3].

In this paper, we present a method allowing the displacement of the listener in a virtual audio space, while respecting the characteristics and variability of the "late reverberation" decay.

2 Auralization

Auralization aims at creating virtual acoustic environments. In particular, when the RIRs are obtained by room acoustics modeling, this is known as (fully) computed auralization [4].

Room acoustics programs (e.g. ray-tracing programs) are used to compute echograms, for a fixed position of the sound source and some receiver's locations. Each source-receiver pair gives rise to one echogram, which is itself transformed into a room impulse response (RIR) : see [1, 2, 5] for more details.

A variant to this classical method consists in the computation of directional echograms [3] : in that case, the acoustical contributions at the receiver are distinguished by their direction of incidence, which gives rise to several directional RIRs (DRIRs) for each source-receiver pair (in [3], 26 DRIRs are computed at each receiver position). This is particularly useful in asymmetrical acoustical spaces, to reveal flutter echoes or the influence of asymmetrical distribution of absorbing materials.

For binaural auralization (reproduction with headsets), the DRIRs are then convolved with an anechoic signal $s(t)$ and with the head-related transfer functions (HRTF) of the listener's head. If p_{left} is the *ear drum* sound pressure to be reproduced at the left ear and $h_{SR}(\theta, \phi, t)$ is the DRIR computed in the direction of incidence (θ, ϕ) , then :

$$p_{left}(t) = s(t) * \iint_{4\pi} h_{SR}(\theta, \phi, t) * HRTF(left, \theta, \phi, t) \sin\theta d\theta d\phi \quad (1)$$

For multiple-loudspeaker auralization, the anechoic signal must be convolved with the DRIR (or with the combination of DRIRs) corresponding to each loudspeaker.

3 Irregular reverberation decays

Acoustical spaces with disproportionate geometries and/or asymmetrical distributions of absorbing materials can lead to very different directional echograms around the listener's head, which in turn creates significant non-exponential decays in the late part of the RIR.

One such room has already been described in [6] : this was an *academic* example. Here follows another example, which has been rather inspired by a real room acoustics project. The acoustical space is an ancient horse-riding school which will be transformed into a cultural hall. The dimensions are 21.5m x 37.6m x 14.9m . The left and right walls (when looking to the scene) have been treated with acoustical absorbents and diffusors, but the back and front walls (which are parallel) create specular reflections which lead to a clear flutter echo in figure 1. Moreover, the ceiling is fully covered with absorbing materials. As a consequence, the "up" directional echogram is clearly shorter than the others in figure 1. For all these reasons, the echogram at this receiver's position (and integrated for all angles of incidence) has a non-exponential decay.

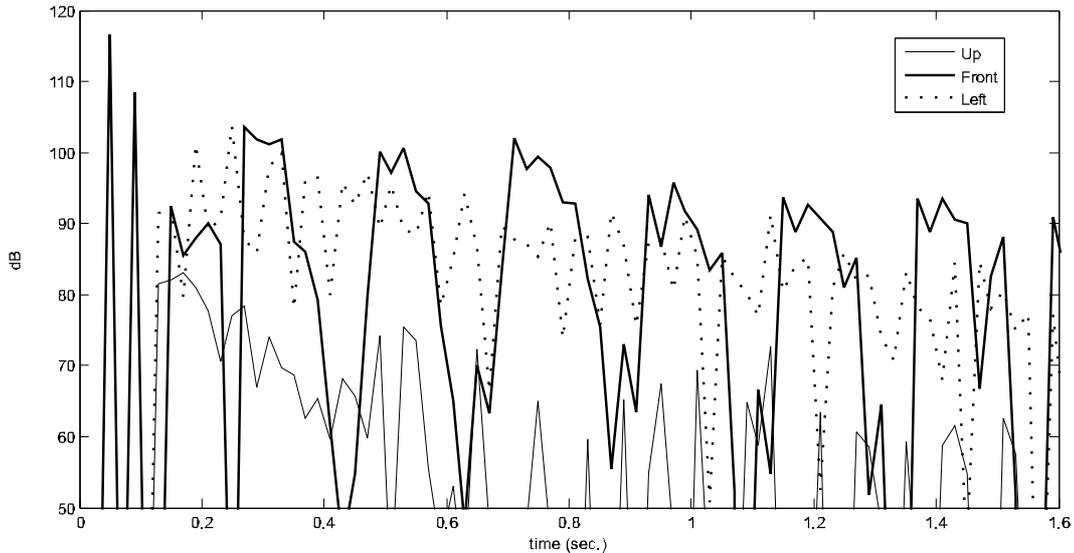


Figure 1: Three directional echograms computed at the same position in a cultural hall, showing a flutter echo in the front-back direction and a steeper decay in the up-down direction

4 The problem of the moving listener

The method we have implemented tries to get rid of the assumption of an exponential decay for the late reverberation part of the RIR, irrespective of the receiver's position.

To do this, it is first necessary to predict how the echogram is modified when the receiver moves (the source remaining at the same location in our study). This prediction is rather easy for the direct sound component and the first-order mirror sources contributions. Indeed, as the corresponding *source* point is known, it is sufficient to compute the new distance to the receiver and the new direction of incidence, to update these contributions (the problem of viewed/non viewed sources must however be carefully investigated). And this computation could be quite fast, as long as the number of mirror sources (which is related to their maximum order) is limited, allowing real-time implementations [1].

A first problem arises when diffuse reflections must be accounted for. These non-specular reflections occur on rough surfaces in the virtual room. Mirror sources cannot be defined to model these reflections, which impedes the updating of their contributions.

The second problem arises for the late part of the RIR, which results from a combination of a great number of high-order specular as well as diffuse reflections. It is therefore impossible to simply predict the modifications induced by a displacement of the listener.

The solution that we adopted was to compute the echograms at some pre-defined listener locations in the virtual space. The echogram at any new position will then be estimated by an interpolation between the closest pre-defined locations. This interpolation is restricted to the diffuse contributions and the *late reverberation* part, or in other words to the whole echograms, except the direct sound and first-order mirror sources components which are still updated as explained before.

Figure 2 shows three echograms computed in the long disproportionate room described in [6]. Clearly, the decay is not simply exponential (not linear in a dB-scale)

and the echograms reveal a significant flutter echo between two parallel walls. These echograms are computed at three receiver's locations situated in the same vertical plane, perpendicular to the horizontal direction in which the flutter echo takes place.

Even if the three receiver's locations are 22.5 meters apart, it is shown that a linear interpolation is sufficient in that case, as the echograms are very similar.

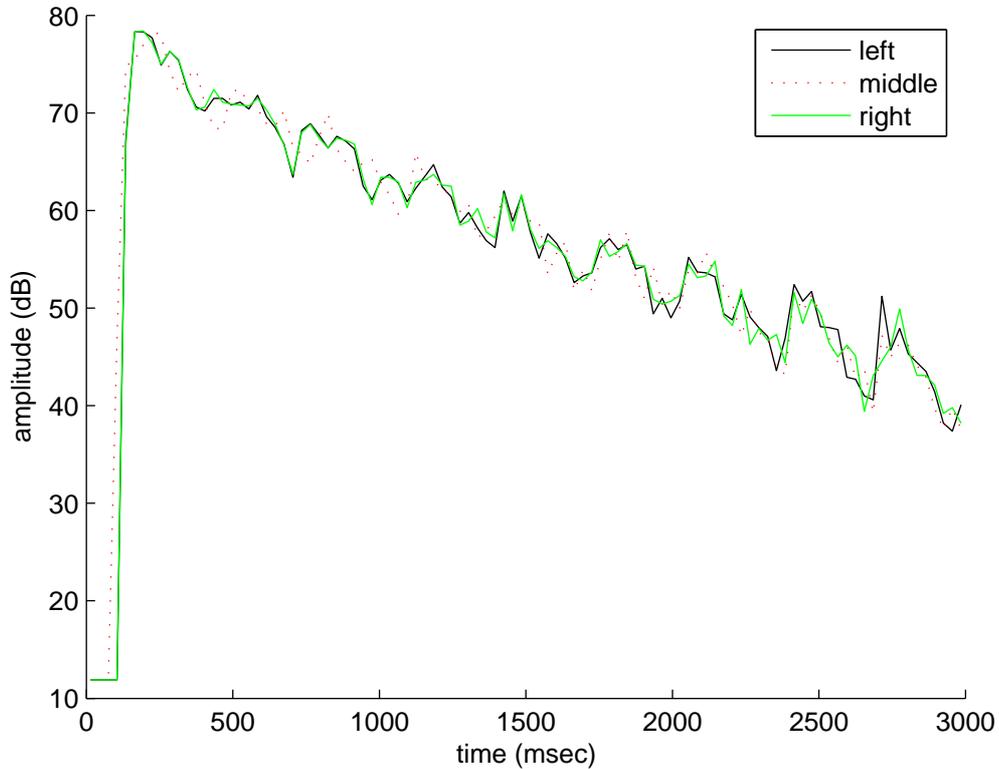


Figure 2: Three echograms (amplitude in dB vs time in msec) computed at three receiver's locations in a long disproportionate room. All receivers lie in the same vertical plane perpendicular to the flutter echo main direction (octave band 125 Hz) : *left*, *middle* and *right* refer to the receiver's location relative to the main room's axis.

Figure 3 shows the echograms computed at five receiver's locations, aligned *along* the direction of the flutter echo and situated 10 meters apart. The flutter echo structure still appears, but the local extrema of the individual echograms do not coincide anymore. This example shows that the late reverberation part cannot be considered as constant (and so, neither is the RIR). In that case, a linear interpolation would certainly respect the general decay slope of the echograms, but certainly not their fine details (unless the pre-defined locations are very close to one another).

Another difference between the echograms in figure 3 is the initial *start-up* time, which is more and more delayed as the receiver's location moves away from the source. This effect is accounted for in the particular treatment of the direct sound and the mirror sources contributions, but not in the diffuse contributions. It has been shown in [7] that this could lead to interpolated (or predicted) echograms starting earlier than their true initial time. However, this effect is well limited and not audible, if the pre-defined locations are not too far from one another.

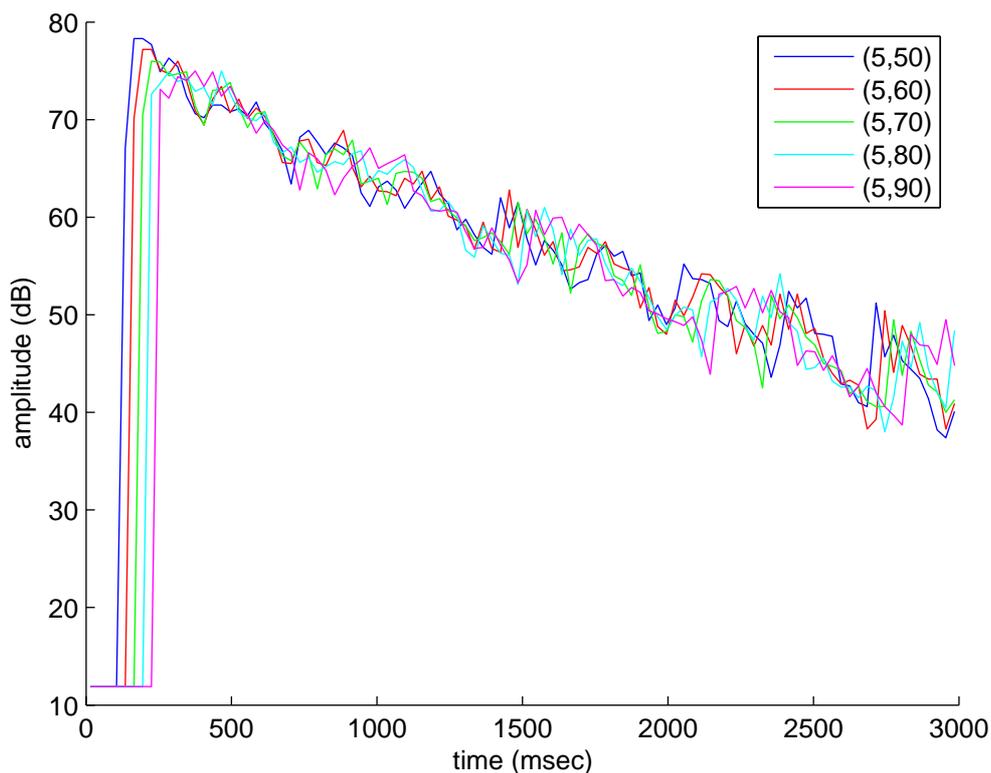


Figure 3: Five echograms (amplitude in dB vs time in msec) computed at five receiver’s locations in a long disproportionate room. All receivers lie in the same vertical plane parallel to the flutter echo main direction (octave band 125 Hz).

5 Tests and discussion

Tests of the method have been reported in [7]. In each virtual room, the listener’s displacements were restricted to an horizontal plane (about 1m75 above the floor). The listener can also rotate his head and the binaural cue is updated accordingly (a head-tracker sensor is fixed to the headset and continuously records the listener’s head orientation). An anechoic speech signal (source) has been used for the cue tests.

In the computer model, the mirror sources up to the third order were treated separately. The computing time needed to update the RIR (including the contributions of the mirror sources) is of course critical. In this work, it could be as high as one or two seconds. Therefore, the application is not real-time for very fast displacements in the room, i.e. for which the virtual speed of the listener exceeds 1m/s.

In order to avoid blanks in the cue tests, the old RIR remains active as long as the updated one is not available : so, the listener can perceive a small delay between the displacement action and its audible effect.

The auralization with the interpolated RIRs was judged of good quality by the listeners, compared to the auralization without interpolation, at the same virtual location. Only in one very long room, it could happen that the interpolated echogram starts *before* the direct sound (see section 4). This effect could be perceived by very acute listeners, but a simple way to avoid it would be to delete from the interpolated

echogram all these contributions coming earlier than the direct sound.

Another (more objective) test consisted in computing some room acoustics parameters from the interpolated and the corresponding non-interpolated echograms. These parameters were the following : the reverberation time T30, the early decay time EDT, the definition D50 and the clarity C80. These two last parameters were the most affected by the interpolation : the deviations could be as high as 10 percents for D50 and 0.5 dB for C80. But, again, these deviations were not audible by most listeners.

An important issue is the number and the definition of the pre-defined listener's locations. It is first recommended to define a rectangular mesh of receivers, in order to reduce the complexity of interpolation calculations. It is also recommended to choose at least one receiver in each zone of interest.

It is interesting to reduce the number of pre-defined locations as far as possible, in order to restrict the size of the echograms' database. A compromise has to be found between this size and the quality of the auralization resulting from the interpolation. This compromise will depend on the virtual room itself : diffuse sound fields and simple rooms will require less pre-defined receivers than complex rooms with specular surfaces. In the more simple rooms, the distance between these locations could be as high as 40m, without affecting the quality of auralization.

Acknowledgment

The authors wish to thank Sebastien Geron who realized the user interface for the tests and who analyzed their results.

References

- [1] L. Savioja, J. Huopaniemi, T. Lokki and R. Vaananen, "Creating Interactive Virtual Acoustic Environments", J. Audio Eng. Soc. Vol. 47, No. 9, 675-705, September 1999.
- [2] L. Savioja, "Modeling Techniques for Virtual Acoustics", PhD Thesis, Helsinki University of Technology, 2000.
- [3] J.J. Embrechts, S. Lesoinne and N. Werner, "Computation of directional impulse responses in rooms for better auralization", Convention paper No.6498 of the 118th Convention of the Audio Eng. Soc. (AES), pp5, Barcelona (2005).
- [4] M. Kleiner, B.-I. Dalenback, and P. Svensson, "Auralization-An Overview", J. Audio Eng. Soc. Vol. 41, No. 11, 861-875, November 1993.
- [5] D. Archambeau, "Realization of an auralization software based on echograms calculated by a sound ray technique", Proc. 2nd Forum Acusticum, Berlin (1999).
- [6] S. Lesoinne, N. Werner and J.J. Embrechts, "3D real-time auralization with separate rendering of direct sound, reflection and directional late reverberation", Proc. 12th Int. Conf. on Auditory Display, pp4, London (2006).
- [7] S. Geron, "Modification des reponses impulsionnelles acoustiques d'une salle lorsque le recepteur se deplace. Realisation du module de calcul", Student Thesis (in french), University of Liege, 2006.