

An auralization system for real time room acoustics simulation

PACS 43.55.Ka

Diogo Alarcão; David Santos; J. L. Bento Coelho
CAPS – Instituto Superior Técnico, Technical University of Lisbon
Av. Rovisco Pais 1
P-1049-001 Lisboa
Portugal
Tel. +351 21 841-9367
e-mail: diogo.alarcao@ist.utl.pt

ABSTRACT

A real time room acoustics auralization system for arbitrarily shaped rooms is presented. The system is based on an accelerated image source method and a time-dependent radiosity method for the computation of the complete binaural room impulse response. The auralization module is implemented as a wide band frequency dependence system. Elements such as wall and air absorption, source and receiver directivities are taken into account. The system permits the auralization of suitable audio signals by including the direct sound, the early specular and diffuse reflections as well as the reverberant tail of the room response. Accelerating techniques and algorithmic improvements are used in this system allowing real time auralization through binaural technology.

RESUMEN

Neste artigo é apresentado um sistema para a simulação e auralização em tempo real da acústica de salas. O módulo de cálculo do sistema baseia-se num método das imagens acelerado e num método de radiosidade hierárquica dependente do tempo para a computação da resposta binaural de um dado espaço fechado, de geometria arbitrária. O módulo de auralização utiliza uma implementação espectral de banda larga. As directividades da fonte e receptor são consideradas, bem como a absorção das paredes e a atenuação no ar. O sistema permite a auralização de sinais sonoros anecóicos gravados tendo em conta o sinal directo, as primeiras reflexões, especulares e difusas, bem como a cauda reverberante da resposta da sala em simulação. Várias técnicas de aceleração e de optimização dos algoritmos são utilizadas neste sistema permitindo a auralização em tempo real através da tecnologia binaural.

1. INTRODUCTION

For most practical purposes in room acoustics it is only necessary to obtain the prediction of how the sound energy propagates inside the enclosed space.

In addition, under the simplifying assumption that the enclosure's walls reflect sound energy as a mixture of specularly and diffusely reflected components, it can be shown that solutions for the propagation of sound energy inside arbitrary spaces can be achieved by a combined method [1, 3]. This new combined method resorts to an extended mirror image source method solving for the propagation of the specularly reflected sound energy components and to a time-dependent hierarchical radiosity approach in order to solve for the propagation of the diffusely reflected sound energy components. This combined method was used for the development of a real time auralization system at the Group of Acoustics and Noise Control, at CAPS-IST.

2. DESCRIPTION OF THE SYSTEM

The module for solving the specularly reflected sound components was implemented by using an extended image source method, while the module for solving the diffusely reflected components is based on a Hierarchical Radiosity [2] method allowing the error threshold to be defined by the user in order to ensure maximum accuracy of the solution and low computation times. Both modules run on independent calculation threads.

2.1 Extended Image Source Method

Owing to the exponential growth of the number of potential image sources with reflection order, given by:

$$M + M(M-1) + \dots + M(M-1)^{K-1} = \frac{M(M-1)^K - 1}{M-2} \approx (M-1)^K \quad (1)$$

with M being the number of considered walls and K the reflection order, the exact geometrical calculation of image sources is performed only until some given maximum geometrical reflection order, typically 3 to 5.

2.1.1 Accelerated image source method

In order to obtain the list of visible mirror image sources in the least time possible, four accelerating techniques are used.

- *Back-Face Culling*

“Back-face culling” is of straightforward implementation and consists on taking into account the orientation of the input polygons according to their inward-orientated normals. Then, mirroring takes place only in the case that the sources (be it the original sound source or the constructed mirror images) are facing the inward face of some polygon.

- *Impossible Wall Combinations*

This technique has been suggested in references [4, 5] and permits recognising a priori non-physical mirror images representing impossible combinations of input polygon pairs. The original geometric representation of the enclosure under study is used in order to build a list of impossible polygons combinations, meaning that sound reflected by some particular polygon can not reach directly other polygons.

The information stored in a pre-process phase in the complete list of impossible reflections combinations is then used during the geometric construction phase of the potential mirror images in order to discard immediately these combinations. With this accelerating technique, one spares greatly on the costly validity and visibility tests.

- *View-Frustum*

The use of a so-called view-frustum [6] allows a third type of branch to cut from the complete tree of potential mirror images. These branches represent mirror images, which are constructed in a valid manner, but which are not visible from any position within the enclosure and therefore do not take part in the room's impulse response.

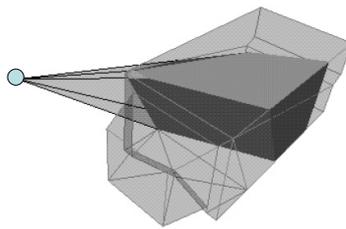


Figure 1 - Example of a view frustum for discarding higher order images impulse response.

This accelerating technique was used in a slightly different manner in [4], while in [5] the authors suggest the use of radiation angles, which in practical terms fulfil the same purpose.

Start with some mirror image and construct its view frustum as indicated in the example of Figure 3. In practice, one uses the fact that valid reflection surfaces for descendants of an image source (yielding higher order mirror images) are the ones whose inward-orientated faces are seen within the view frustum, whose apex coincides with the mirror image.

This accelerating technique yields a good performance increase in the MISM since it allows, especially for high orders of reflection, to cut immediately many branches of the tree of potential images.

•Clustering of Input Polygons

This technique uses the fact that in the majority of the enclosures found in practice all the input polygons (which can be in great number) normally lie on a small set of tri-dimensional planes. Therefore, in a pre-processing phase, a clustering algorithm builds up a single “parent” polygon in a similar way as the so-called convex hull of a set of points is constructed. This clustering algorithm greatly reduces the complexity of the problem, because the number of wall combinations can be greatly reduced in a wide variety of practical cases, with the apparent consequence that the number of potential mirror images is enormously reduced due to the reduced basis of the exponential law.

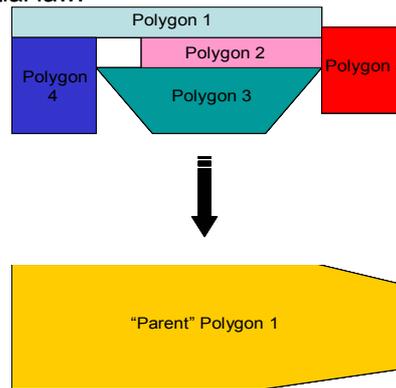


Figure 2 - Example of clustering of input polygons into one single “parent” polygon

The accelerated image source method with the clustering algorithm uses a simplified representation of the enclosure given by all the “parent” polygons (which are convex) for the geometric construction of the potential images. Only when a valid and visible mirror image is found, in relation to this set of “parent” polygons, does one need the complete set of original input polygons for determining which of them actually are responsible for the construction of the determined mirror image. This determination is not a very costly operation, since one need only to determine inside which of the input polygons the intersection points are located. It is obvious that a valid and visible image relative to the reduced set of “parent” polygons can be an invalid or invisible mirror image afterwards, relative to the finer set of input polygons. However, these cases constitute normally a minority, since the “parent” polygons are constructed in a way similar to the convex hull of a set of points, which is the convex set with smallest area that contains all the original input points.

In Table 1 we indicate some calculation times for the implemented accelerated image source method for different rooms. Calculation times are for reflection order 5. As can be seen, the calculation times are very small, suitable for real time auralization purposes. The hardware used consists of a PC with an Intel Core Duo 1.86 MHz processor and 1 GB RAM running Windows XP Pro.

Table 1 - Calculation times for the accelerated image source method

Room	Polygons	Calc Time [ms]
Simple	26	1
Congress Centre	55	15
PTB Music Studio	77	20

Figure 3 shows the room visualization module with the calculated sound paths between source and receiver for order 3 for the auditorium of the congress centre of IST.

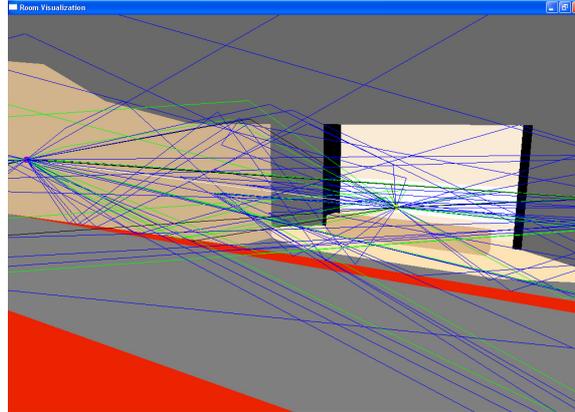


Figure 3 - Room visualization module showing sound paths between source and receiver

2.1.2 Extension to high orders of reflection

Information supplied by the accelerated image source method is then used for a statistical and deterministic extrapolation to higher reflection orders in order to obtain the reverberation tail of the room's specular impulse response.

The extrapolation step resorts to the following parameters, obtained during the exact geometrical construction phase:

- Number of visible images per reflection order
- Mean distance receiver - visible images per reflection order
- Standard deviation of the distances per reflection order
- Mean reflection coefficient per reflection order
- Standard deviation of reflection coefficients per reflection order
- Mean diffusivity coefficient per reflection order
- Standard deviation of the diffusivity coefficients per reflection order.

The room impulse response from specular reflections is thereby calculated from two terms:

$$I_{\text{specular}}(t) = \sum_{K=0}^{\text{maxgeoorder}} \sum_{i=1}^{N(K)} \frac{\Pi}{4\pi d_i^2} \rho_i e^{-md_i} (1 - \Delta_i) \delta\left(t - \frac{d_i}{c}\right) + \sum_{K=\text{maxgeoorder}+1}^{\text{maxorder}} \sum_{i=1}^{N(K)} \frac{\Pi}{4\pi D_i(K)^2} \bar{\rho}(K) e^{-mD_i(K)} \overline{1 - \Delta(K)} \delta\left(t - \frac{D_i(K)}{c}\right) \quad (6)$$

where "maxgeoorder" means the maximum reflection order for which the exact geometrical calculation of the visible images is done.

2.2 Hierarchical Time Dependent Radiosity Method

Solving for the diffuse reflected components by a finite element approach, raises two problems:

Problem 1: For the image sources method it is optimum that the walls constituting the geometry of the enclosure be as large as possible, whereas for a finite element approach one desires that the input polygons constitute a mesh of several smaller "patches".

Problem 2: If M initial walls are split into n patches, then the number of energy links, i.e. the approximate number of form factors, will be proportional $O(n^2)$, and therefore the calculation effort will be high.

One solution for both problems is to adopt a multi-resolution approach through hierarchical linking [2], whereby the patches are generated adaptively from the geometry's input walls and the number of links (form factors) is $O(M^2+n)$, therefore permitting a considerable computational saving.

The hierarchical subdivision stop criterion is based in first place on the absolute minimum area of patches (defined by the user) and in second place on a threshold condition for the form factors between pairs of patches (also defined by the user). Form factors are calculated by using the analytic formula of the form factor from a differential area to a parallel disc, and the visibility factor for the occluded form factor is calculated by ray-casting.

The solution of the time-dependent equations is done by a Gauss-Seidel relaxation scheme, using an energy "push-pull" operation throughout each hierarchy of patches. In order to limit the

exponential growth of diffuse reflections with increasing reflection orders, a condensing algorithm with an internal sampling rate is used.

At each iteration, energy is gathered at predefined receiving points, calculated through the well known Lambertian cosine law.

2.3 Auralization Module

The auralization module of the system runs independently on asynchronous threads, loading periodically the list of visible images as calculated by the accelerated image sources module and the energy echograms as calculated by the hierarchical time-dependent radiosity module. The auralization module implements a wide band approach for the treatment of the frequency dependence. It is currently adapted for rendering sound through binaural technology.

2.3.1 Auralization of the specular components

The list of visible images calculated by the accelerated image source method comprises the sources location as well as the list of walls that generated them. The incidence angles regarding the particular receiver point are also recorded in the list, as well as the image source's directivity. With this information, equivalent time invariant linear filters are constructed online by the auralization system.

Elements such as source directivity, wall and air absorption are represented through linear filters in logarithmic frequency dependence with an octave band scale with 10 values ranging from 31.5 Hz up to 16 kHz. These filters possess only magnitude values, whereby their phase is determined by a minimum phase reconstruction step done through an FFT Hilbert Transform.

In order to obtain the binaural pressure impulse responses, the result of the previous filters have to be convolved (in the dual frequency domain after interpolation) with the set of HRIR for each ear corresponding to the correct direction of incidence at the receiver.

The auralization module uses currently the HRIR database of IRCAM, which were measured at a distance of 2.0 m. This database has a resolution of 15° in the azimuth angle. The vertical resolution consists of 10 elevation angles starting at -45° ending at +90° in 15° steps. The steps per rotation vary from 24 to only 1 (90° elevation). Measurement points are always located at the 15° grid, but with increasing elevation only every second or fourth measurement point is taken into account. As a whole, there are 187 HRIR pairs. Each HRIR data has a sampling rate of 44.1 kHz and 24 bits of quantization, consisting of 8192 samples long. We pre-processed each HRIR data set in order to have three new data sets: the first has 512 samples, the second has 128 samples, and the third has 32 samples.

The 512 samples long set is used for the processing of the direct sound, while the 128 samples long set is used for processing of the specular components contained within the list of visible images. The third set is used for the processing of the extrapolated images.

All the convolutions are done in the dual frequency domain using radix-2 FFTs.

The final result of this stage corresponds to two binaural impulse responses containing all the information about the specularly reflected sound components.

2.3.2 Auralization of the diffuse components

The result of the hierarchical time-dependent radiosity method consists of an energy echogram for the 10 octave bands. This echogram is considered as an envelope of the squared pressure impulse response for each frequency band.

In order to reconstruct a pressure related impulse response a procedure using a filtered white noise is implemented. In detail, a 10 ms width Gaussian window is applied successively together with an equivalent magnitude filter determined from the calculated energy echograms to the white noise. This procedure is repeated for increasing time steps until the time window slides out of the time interval of interest.

For the binaural simulation two uncorrelated responses are created which are equalised by means of an averaged HRTF according to diffuse incidence.

2.3.3 Auralization of all components

The binaural impulse responses for the specular and for the diffuse components are combined into a single, total binaural impulse response set. This final set is then convolved with a streaming 44.1 kHz, 16 bit audio signal through a Gardner like low-latency filtering approach with an overlap-add implementation. Sound reproduction is done using binaural technology through conveniently equalised headphones.

3. Real Time System

In Figure 4 we present the block diagram of the complete real time simulation and auralization system implemented on an Intel I7 Extreme processor platform. In Table 2 we present some performance data of the whole implemented real time system.

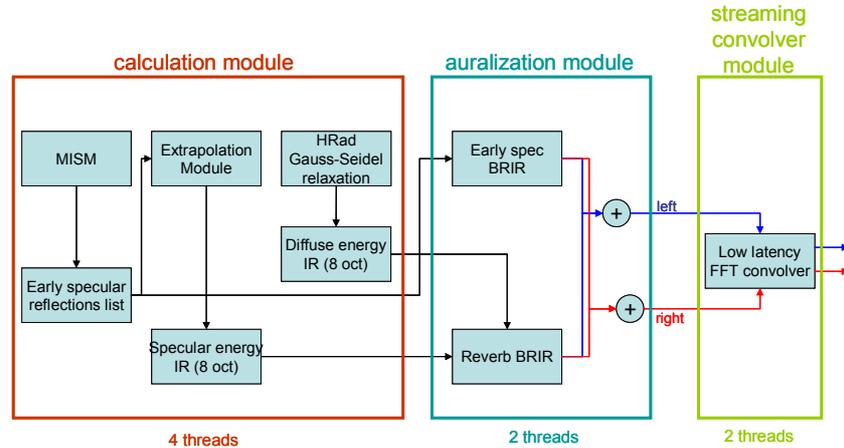


Figure 4 – Block diagram of the complete real time simulation and auralization system

Table 2 – Performance figures of the real time system

Task	Calc Time [ms]
MISM early images (order 4)	20
Extrapolation module	38
HR Gauss Seidel relaxation	1100
Diffuse energy IR at receiver	6
Early specular refl BRIR	5
Reverb BRIR	10
FFT Convolver (10s audio)	62

4. Conclusions and Future Work

The described system allows for both the real time movement of receiver and source. The accelerated image source method permits the real time calculation of all visible images for some source-receiver pair, while the hierarchical time-dependent radiosity method calculates the diffusely reflected sound components through octave band energy echograms. The auralization system is based on a wide band simulation approach. Elements such as wall and air absorption, source and receiver directivities are taken into account. Future work will access the possibility of using some space partitioning hierarchy for speeding up the accelerated image source method and the hierarchical time dependent radiosity method for rooms with greater complexity.

4. Acknowledgements

This work was financially supported by the Portuguese Science and Technology Foundation (Project PTDC/EEA-ELC/80910/2006).

REFERENCES

- ¹ D. Alarcão and J. L. Bento Coelho; *A combined method for room acoustics integrating specular and diffuse reflections*. In Proceedings of 18th. International Conference on Acoustics, Kyoto, Japan, April, 2004.
- ² P. Hanrahan, D. Salzman and L. Aupperle; *A rapid hierarchical radiosity algorithm*. Computer Graphics, 25(4):197-206, 1991.
- ³ D. Alarcão, *Acoustic Modelling for virtual spaces*; PhD Thesis, IST, 2005.
- ⁴ H. Lee; B. Lee; *An efficient algorithm for the image model technique*. Applied Acoustics, Vol. 24, p. 87-115, 1988.
- ⁵ U. R. Kristiansen, A. Krokstad, T. Follestad; *Extending the image method to higher-order reflections*. Applied Acoustics, Vol. 38, n. ° 2-4, p. 195-206, 1993.
- ⁶ A. Watt, M. Watt; *Advanced animation and rendering techniques – Theory and practice*. ACM Press, Addison-Wesley Publishing Company, 1992.