Proc.1st IEEE Benelux Workshop on Model based Processing and Coding of Audio (MPCA-2002), Leuven, Belgium, November 15, 2002

MODELLING ACOUSTIC SPACES FOR AUDIO VIRTUAL REALITY

U. Peter Svensson

Acoustics group, Dept. of Telecommunications, Norwegian University of Science and Technology (NTNU), NO-7491 Trondheim, Norway

svensson@tele.ntnu.no

ABSTRACT

The computational modelling of acoustic spaces is fundamental to many applications in auralization/virtual acoustics. The demands vary widely, from real-time simulation in multimedia and computer games, to non-real time situations with high accuracy needs, such as prediction of room acoustic conditions in music performance spaces. Acoustic spaces include single room or multi-room spaces, with simple or complex geometries and boundary conditions. Outdoor spaces can range from city environments to open landscapes. Sound transmission through partitions is an important issue in some cases. This presentation gives an overview of techniques used in the various auralization applications.

1. INTRODUCTION

The generation of audio for virtual reality has many equivalent terms, such as auralization [1], [2], virtual acoustics [3], binaural room simulation [4], auditory display etc depending on application. Some applications require auditory and visual components; the requirements for accuracy might differ widely and the need for real-time processing, or interactivity, depends on application.

The most common approach to auralization or virtual acoustics is a two-stage process: the computation of an impulse response (IR) representing an acoustic space, and the convolution of this impulse response with a dry (anechoically recorded or synthetically generated) source signal, see Fig. 1. Depending on how the source directivity is modelled and which reproduction technique that is chosen, the impulse response must be two-channel or multi-channel.

The IR computation relies on a calculation method for the sound field modelling. These methods could be divided into two classes: those based on sound field decomposition and those based on the numerical solution of the wave equation. Methods of the second type are rigorous and have the advantage that they reliably give a correct response. On the other hand, this comes at a high computational price and an inflexibility when it comes to scalability of accuracy etc. The two classes are described separately in chapters 2 and 3.

1.1. Time- or frequency domain calculations

Auralization requires that the IR of an acoustic space is computed. This can be done directly in the time-domain or via the inverse Fourier transform of a frequency-domain solution. For some methods, frequency domain formulations might be more easily available but they have the disadvantage that the time-axis is "collapsed" into the frequency domain information, i.e., no separation or decomposition of the early part and late part is possible.

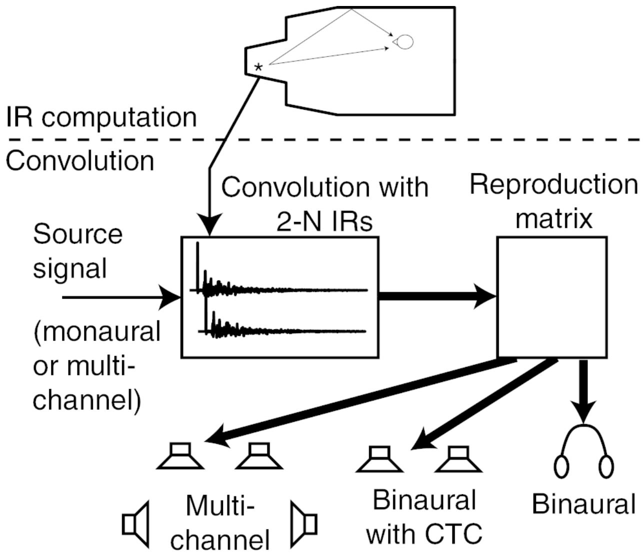

Figure 1: *Illustration of an auralization system with the two stages: impulse response (IR) computation and Convolution.*

From a perceptual viewpoint, the IR of an acoustic space does not need to be calculated with the same detail for the late part as for the early part. This is because the larger the number of sound waves that contribute around the same arrival time, the smaller are our possibilities to perceive details of the compound sound field. Also, the limited frequency resolution of our hearing makes it unimportant to calculate the detailed interference effects in the frequency response correctly at high frequencies. It is large-scale parameters that determine the perception, such as energy-ratios in frequency bands.

Finally, the spatial details are critical for the direct sound and the very early part of the impulse responses, but for the later part where many sound waves interact, quantities such as interaural cross-correlation might determine the perception.

These observations lead to the conclusion that accurate methods waste much computational efforts on calculating unimportant details correctly. Still, the "right amount" of detail must be computed; the spectrum, the energy decay envelope, and the spatial distribution should be computed with the right amount of accuracy, with the potential strong echo/reflection resolved. Also, in special cases where there are very few waves interacting, such as in outdoor cases, or with the typical flutter echo of two flat hard walls facing each other, a listener might be quite sensitive to the details.

2. METHODS BASED ON SOUND FIELD DECOMPOSITION

A number of calculation methods use an approach which views the sound field, and hence the IR, as a sum of elementary components that are represented by, e.g., image sources, or rays. This is illustrated in Fig. 2 where secondary sources are indicated. They are of the image source, edge source or surface source type. A great advantage of this approach is that it can be possible to assign priorities to the individual contributions and the use of a detailed early part and a simplified late part is an example of this.
The total impulse response $h_{tot}(t)$ could be written as a sum of secondary source contributions,

$$h_{tot}(t) = \sum_{i=0}^{\infty} h_i(t) * \delta(t - t_i) \tag{1}$$

where $h_i(t)$ is the IR that represents secondary source i and the propagation delay t_i is handled separately by the Dirac function $\delta(t)$. Various approaches are used to compute the IRs $h_i(t)$. Since the number of secondary-source contributions grows rapidly for fully or partly enclosed spaces (indoor cases or city-like outdoor environments) a common approach is to calculate a few contributions, N_{detailed}, by some more accurate method and the late part of the IR, $h_{\text{simplified}}$, by some simplified method:

$$h_{tot}(t) = \sum_{i=0}^{N_{\text{detailed}}} h_i(t) * \delta(t - t_i) + h_{simplified}(t) \tag{2}$$

For single-room indoor cases, the simplified part can quite easily by generated by an exponential-decay model [3],[5]. Whenever non-exponential decays are encountered (e.g., single-room cases with subspaces such as under balconies, double-room cases with smaller openings between, out-door cases etc) more refined methods must be used for $h_{\text{simplified}}$.
The secondary-source IRs/filters h_i can be further decomposed as illustrated in Fig. 3 into a source filter, a propagation filter, a receiver filter and a cascade of reflection filters [3]. Such a decomposition can allow a pre-calculation, and even a parametric form, of these source filters etc for more efficient updating of the total IR computation for cases where IR updating speed is crucial such as in interactive cases.
A common technique for the calculation of h_i is to design an FIR-filter based on the desired magnitude function, which is given by interpolation between octave-band values. The octave-band values are found by calculating the intensity

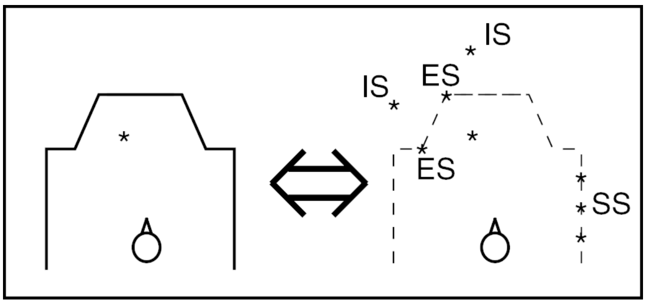

Figure 2: *Various types of secondary sources, IS = image source, ES = edge source, SS = surface source, represent an acoustic space.*

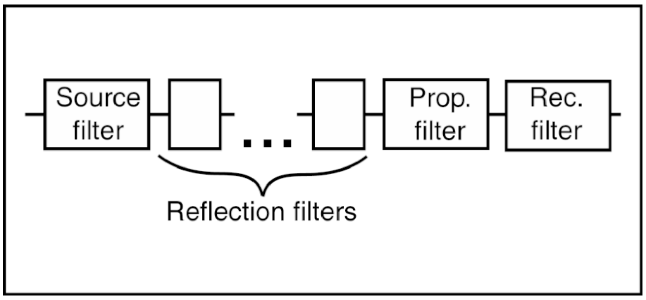

Figure 3: *The filter that represents a secondary source, h_i, can be decomposed into source, reflection, propagation, and receiver filters.*

contributions from a secondary source. The potential phase shift that is introduced at a wall reflection is usually neglected and a linear-phase or minimum-phase filter design is used. This neglection of phase shifts is sometimes erroneously interpreted as a neglection of interference effects in the total IR, h_{tot}, but it should be noted that the phase shift caused by the propagation delay is handled correctly and consequently, interference effects will be present. Neglecting the phase shifts is motivated by two factors: firstly, the phase shift are usually not known and secondly, the phase details are inaudible in a diffuse sound field [6].

2.1. The image source concept / specular reflection

The image source concept replaces walls and other reflecting surfaces with mirror image sources. The amplitudes of the image sources are adjusted so that the boundary conditions, i.e., the wall impedances at the positions of the walls, are fulfilled. This can be done exactly for just a few idealized cases, e.g., a rectangular room with rigid walls. However, this image source method yields solutions that are asymptotically correct at high frequencies, even for walls with specified impedances of locally reacting surfaces. Fig. 4 illustrates that an image source will have a strength that depends on the angle θ, i.e., the image source must be directional. Furthermore, the figure illustrates that image sources are active only for receivers that can see the image source through a finite reflecting plane. Receiver R2 in Fig. 4 is not reached by a specular reflection and this illustrates one problem with the image source concept: a discontinuous sound field will occur around zone boundaries such as the dashed line in Fig. 4.
For the simulation of outdoor cases, some cases might have only direct sound and the ground reflection so the ground reflection might be very important for a realistic impression.
A factor which leads to inaccuracy is that both the source and the receiver usually are close to the ground, and in addition the source-to-receiver distance might be long.
For an image source, the filter subdivision in Fig. 3 can be represented by an intensity-based form that gives the intensity, I_i, of a image-source contribution,

$$I_i = P_S e^{-mr_i} \frac{DF_i}{4\pi r_i^2} \prod_{j=1}^{n_{refl,i}} \left(1 - \alpha_{i,j}\right)\left(1 - \delta_{i,j}\right), \tag{3}$$

where P_S is the source power, r_i is the distance from the image source number i to the receiver, m is the air attenuation coefficient, DF_i is the directivity factor for the relevant radiation angle, $\alpha_{i,j}$ and $\delta_{i,j}$ are the absorption coefficients and

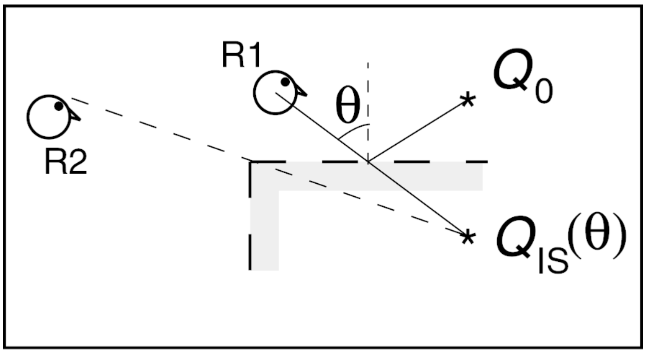

Figure 4: *A source above a finite plane, represented by an image source, and two receiver positions.*

scattering coefficients for image source *i* and wall reflection *j*, respectively. The form in Eq. (3) neglects atmospheric refraction effects that are important in long distance sound propagation and that are more easily modelled with ray tracing [7]. Furthermore, the form in Eq. (3) assumes that the image source strength is independent of the angle θ, and the diffuse field value for $\alpha_{i,j}$ is used. In order to get a realistic direcitivity of the image source, the complex wall impedance, i.e., including phase information, must be used. It can also be noted that surface scattering reduces the intensity of the specular reflection.

2.1.1. Properties of the image source method

The image source method has some strengths and weaknesses:

- The finding of valid image sources can be a computationally very demanding task for higher order reflections. Various methods are available: repeated calculation of the image source positions by mirroring and validation check (slow) [8], using ray tracing (fast but less accurate) [9], or using beam tracing with beam split-up (fast and accurate) [10], [11], [12].

- Once a list of valid image sources is established, a very flexible tool for the simulation of acoustic spaces results since source directivity/rotation, wall impedance, receiver directivity/rotation, image source priority etc can be changed without recalculating the image source positions [3].

- The method can not handle surface scattering/diffuse reflection which is a big drawback since real surfaces always have a mix of specular and scattering properties and the balance between them can have large effects on the sound field in, e.g., room acoustics [13].

- The method is more inaccurate for impedance surfaces the closer the source or receiver is to the wall, expressed in wavelengths. This means that the image source method is inaccurate at low frequencies and/or for very grazing incidence. For more accurate calculations, the image source can be complemented by a surface integral term [14].

- The method can not by itself handle the effect of the finite size of reflecting surfaces correctly, which is exemplified in Fig. 4 by that the reflection disappears in a discontinuous way, when the receiver is moved across an image source visibility zone boundary. For rigid surfaces, this can, however, be corrected exactly by introducing edge sources [15], see Section 2.5.

- Curved surfaces can not be handled by the image source concept so piecewise flat approximations must be employed. Such flat approximations can not represent the spreading effect of a convex surface correctly without applying some artificial scattering effect.

- If the receiver moves, the image source positions don't have to be recalculated but their visibilities must be checked again.

2.2. Ray tracing / diffuse reflection

At real surfaces, reflection will yield both a specular component and scattered components as illustrated in Fig. 5. The specular components can be represented by image sources but the scattered components must be handled by other methods, such as ray tracing or radiosity methods. The ray tracing typically lets each ray that hits a surface be reflected either specularly or diffusively, the choice of which is determined by the scattering coefficient of the surface [16]. It could be viewed as more realistic/accurate if each ray generated a number of secondary rays but such an approach would lead to large computational efforts for high reflection orders [17].

One important issue for any model implementing diffuse reflection is what model to use for this. The classic Lambert diffusion model has been suggested because of its tradition in optics and ease of implementation. However, more realistic modelling differs quite much from Lambert diffusion [18]. It should be noted that the Lambert model implies that the strength of a reflected ray in a certain direction is completely independent of the direction of the incident ray. More realistic modelling leads to that both the incidence angle and scattering angle affects the strength. The access to accurate input data has been very limited for the diffuse reflection properties of surfaces but recent international standards will improve this situation [19], [20].

2.2.1. Properties of the ray tracing method

- Easily handles the mix of specular and scattered reflection which is very important for realistic IR computation.

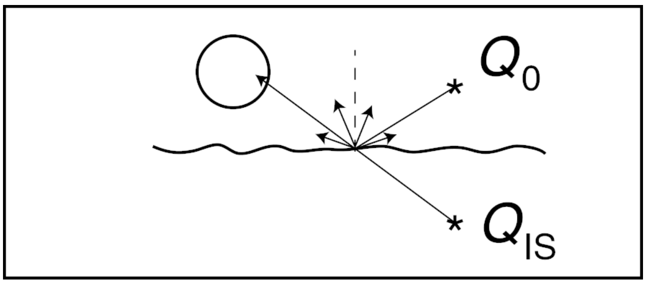

Figure 5: *A surface reflects both specularly, represented by an image source, and diffusively, illustrated by scattered components.*

- Easily scalable for accuracy/speed trade-offs via the choice of number of rays.
- Gives a decreasing accuracy for later times of the IR, which might be perceptually suitable.
- A drawback is that frequency-dependent scattering coefficients might require separate ray tracing processes for each octave band, which is computationally inefficient.
- Another drawback is that if the receiver moves, the ray tracing must be recalculated unless large amounts of intermediate data is stored.
- Ray tracing can not handle the edge diffraction effect of finite surfaces and is not suitable for handling surface impedance (unless the ray tracing is used for finding the valid image sources). It could, however, handle curved surfaces even if most implementations use plane surfaces only.

2.3. Radiosity

Instead of letting the rays that are emitted from the source sample the boundaries of an acoustic space, the radiosity method predetermines the wall reflection parts in the form of larger wall elements. The contributions between all wall-elements, and from the source to the wall-elements, and finally from all wall-elements to the receiver are determined via so-called form-factors. This can be described by an integral equation for the normal intensity at each wall element, $I_{n,i}$,

$$I_{n,i}(t) = I_{n,0\to i}(t) + \sum_{j:j\neq i} k_{j\to i}\rho_j I_{n,j}\left(t - \frac{r_{j\to i}}{c}\right)\Delta S_j \quad (4)$$

where $k_{j\to i}$ is the form factor from element j to element i, ρ_j is the energy reflection coefficient at element j, and ΔS_j is the size of wall element j. $I_{n,0\to i}(t)$ is the contribution directly from a source to element i. Most versions of radiosity employ scattering only, i.e., no specular reflection. With scattering only, and with a Lambert model for the scattering, the form factors have a simple form,

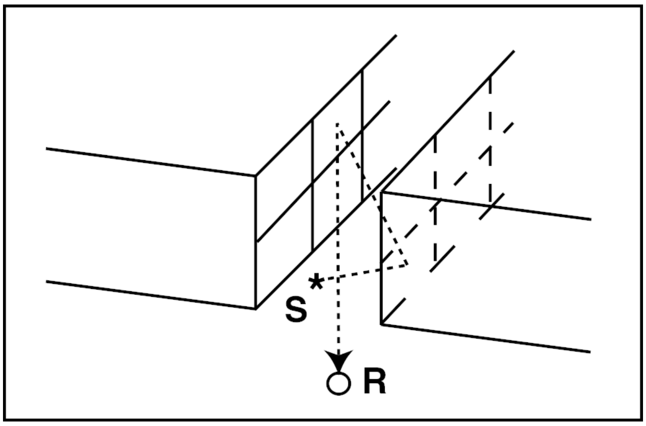

Figure 6: Radiosity modelling of a street corner case.

$$k_{j\to i} = \frac{\cos\theta_i \cos\theta_j}{\pi r_{ij}^2} \quad (5)$$

In the same way as was discussed in Section 2.2, more realistic modelling of scattering would require triplet-form factors: from one wall element via a second wall element to a third wall element [21].
Fig. 6 illustrates the radiosity principle for a street corner case. One sound path is illustrated and only the mid-points of the involved wall elements are indicated even though the form factors actually represent an integration over both wall ele ments. It should be noted that the wall elements do not need to be smaller than the wavelength, which makes this a computationally efficient method [22].

2.3.1. Properties of the radiosity method

- Easily scalable for accuracy/speed trade-offs via the choice of number of wall elements.
- The wall element source strength is independent of the receiver position so it is easy to handle a moving receiver; only the last stage of contribution strengths from wall elements to the receiver must be recalculated.
- The extension to specular reflection [21] has been suggested.
- Radiosity can not directly handle the edge diffraction effect of finite surfaces, even if an extension could be envisioned. Radiosity could in principle handle curved surfaces.

2.4. Beam tracing

Beam tracing is similar to ray tracing in that it emits a number of rays from the source. However, adjacent rays are viewed as forming a beam, i.e., a cut-out of the wavefront. This facilitates that edges can yield beam-splitting [10], [11], see Fig. 7, and furthermore, that edge sources can be introduced in addition to the beams [12]. However, surface scattering is not easily introduced so a combination with radiosity has been suggested for this purpose [23].

2.4.1. Properties of the beam tracing method

- Efficient method for finding valid image sources.

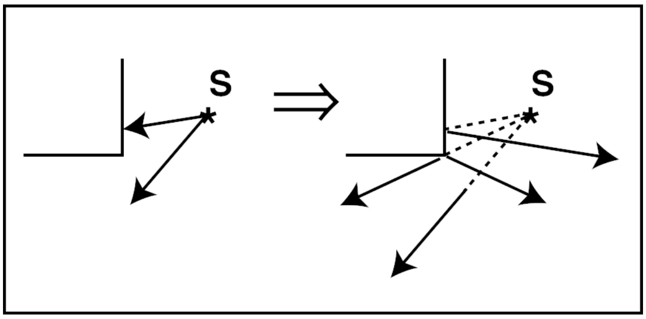

Figure 7: The beam split-up at an edge.

- Can be combined with edge diffraction. but can not easily handle surface scattering.

2.5. Image sources plus edge diffraction

As mentioned in Section 1.1, the concept of edge sources can be introduced to give additional terms to the specular terms in the total IR. These additional terms make the wavefronts continuous at the zone boundaries where specular reflections otherwise would lead to discontinuities. Edge sources are secondary sources that are positioned along edges and radiate in all directions, with analytical directivity functions β that are given by, [15],

$$\beta = \sum_{i=1}^{4} \frac{\sin\left(\nu\varphi_i\right)}{\cosh\left(\nu\cosh^{-1}\eta\right) - \cos\left(\nu\varphi_i\right)}, \qquad (6)$$

where

$$\varphi_i = \pi \pm \theta_S \pm \theta_R, \quad \eta = \frac{1+\sin\alpha\sin\gamma}{\cos\alpha\cos\gamma}. \qquad (7)$$

The factor ν is the so-called wedge index, which is defined as $\nu = \pi/\theta_w$, where θ_w is the wedge angle (on the open side). Whenever the wedge index is an integer value, the directivity function is zero. As an example, for an interior 90-degree corner, $\theta_w = \pi/2$, so $\nu = 2$ and thus, there is no edge diffraction. Furthermore, in eq. (7), the term φ_i is created by the four possible combinations of plus and minus signs. Finally, the angles θ_S, θ_R, α, and γ represent the angles for a sound path incident on, and reradiating from, an edge point, as detailed in [15]. For numerical implementations, edges need to be subdivided into small elements and the contributions from edges lead to IRs that display a time smearing caused by the different path lengths. Fig. 8 shows an example of a stage house geometry, from [24], with an array of receivers positioned laterally in front of a stage house with a single source. By plotting the IRs in a stacked fashion, as in [25], the wavefronts are clearly visible. When only specular reflections are modelled, as in Fig. 8(b), the truncated wavefronts are visible whereas the corrected wavefronts are shown in Fig. 8(c) by the inclusion of edge diffraction.

2.5.1. Properties of the edge diffraction method

- It omplements the image source method directly and together they give teh exact solution for rigid-surface cases.
- No formulation is available for impedance surfaces.
- The computation time grows rapidly with reflection order since a large number of diffraction/specular combinations exist [24], [26].

(a)

(b)

(c)

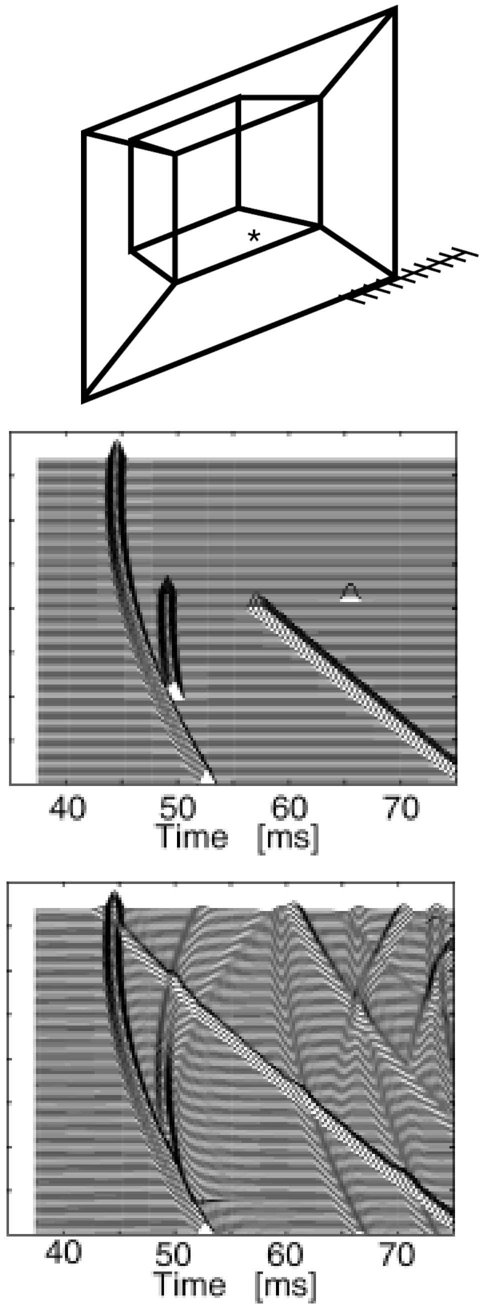

Figure 8:(a) A stage house model with a source and a receiver array indicated. Stacked impulse responses for (b) specular reflection only, and (c) specular reflection + edge diffraction.

3. METHODS BASED ON SOLVING THE WAVE EQUATION

The numerical solution of the wave equation can be done with methods that subdivide a geometry into volume elements or surface elements. For all methods, the computational effort increases very rapidly with the studied bandwidth which is the price to pay for getting "unimportant" but very detailed