ACOUSTIC RENDERING OF BUILDINGS

Rudolf Rabenstein, Oliver Schips and Alexander Stenger

Universität Erlangen-Nürnberg, Lehrstuhl für Nachrichtentechnik
Cauerstrasse 7, D-91058 Erlangen, Germany
e-mail: rabe@nt.e-technik.uni-erlangen.de

ABSTRACT

While visual rendering of buildings is the state of the art in today's design programs, acoustic or auditory rendering is still in its infancy. This paper reviews some promising approaches to the computer simulation of sound propagation and perception in buildings. The range of methods spans from the numerical solution of the wave equation to advanced geometric methods based on ray tracing and radiosity algorithms. Furthermore concepts for modelling the human sound perception are discussed. Finally some issues of practical implementation will be addressed.

1. INTRODUCTION

It is a common feature of modern computer aided design programs to reward the user with a photorealistic rendering of the prospective building. However despite all visual perfection, the auditory component is completely missing. There is no impression of the room acoustics, no assessment of the noise level in offices or workshops and no evaluation of the effect of sound absorbing materials by computer simulation. It has to be emphasized, that all these topics are well studied subjects and that computer simulations of individual acoustic effects are common practice. The problem with acoustic (or auditory) rendering of buildings is that a realistic simulation of the acoustic properties requires a real time computation of complex wavefields. In order to provide a Compact Disk sound quality, it would be necessary to update a 3D sound pressure field with complex boundaries every 23 microseconds. To provide such a feature as an addition to a visual rendering capability is more than even modern desktop computers can handle.

Nevertheless, there have been successful attempts from the virtual reality community to assign sounds to visual objects and to supply real time implementations of virtual scenes [6]. While such audio-visual simulations may sound plausible to the viewer of an animated movie, they do not satisfy the requirements of engineering practice.

However it seems worthwhile to step ahead by combining proven visual rendering algorithms with elements of scientific simulation. This paper reviews some promising approaches in this direction. It starts with a short introduction into the basics of acoustics and systems theory and presents in some detail the concept of the room impulse response. The core of the paper is the treatment of four approaches to the simulation of sound propagation in rooms. Finally modelling of human sound perception by head related transfer functions is discussed and some implementation issues will be addressed.

2. SOUND PROPAGATION

2.1. Basic Theory

From a scientific standpoint it is always pleasing to establish a computer model from first principles. In acoustics (and many other fields) this amounts to the numerical solution of partial differential equations. From a practical standpoint, this is not attractive at all and as a consequence, simpler models are preferred.

We present here two general modelling concepts: the wave equation and an input-output model from linear systems theory. Both concepts are based on the assumption of linearity: The sound field from two sources in parallel is a superposition of the individual sound fields of each single source.

2.1.1. Wave Propagation and Acoustics

Sound propagation can be described by velocity and pressure as functions from space and time. It is governed by the wave equation

$$c^2 \Delta p = \frac{\partial^2 p}{\partial t^2} \qquad (1)$$

where $p(\mathbf{x}, t)$ is the sound pressure, $\mathbf{x}$ is the vector of space coordinates, t is the time, and c is the speed of sound. The derivation of the wave equation from first principles is given e.g. in [17]. In order to calculate the sound field emanating from a source in a specific room, we need an additional source term in (1) and boundary conditions, which describe sound reflection and absorption at the walls. A numerical technique for the approximate solution of the wave equation is presented in section 2.3.1.

2.1.2. Linear Systems Theory

A completely different approach to the description of sound propagation (and other linear and time-invariant systems) is the general input-output relation from linear systems theory

$$y(t) = h(t) * v(t) = \int_{-\infty}^{t} h(t-\tau)v(\tau)d\tau \ . \qquad (2)$$

$v(t)$ is the source signal (e.g. sound pressure) at the location of the sound source (see fig. 1), $y(t)$ is the resulting signal at the location of some receiver (e.g. microphone). The sound propagation from the source to the receiver is described by the function $h(t)$. It is called the *impulse response*, because it is the response at the receiver to a very short pulse at the source.

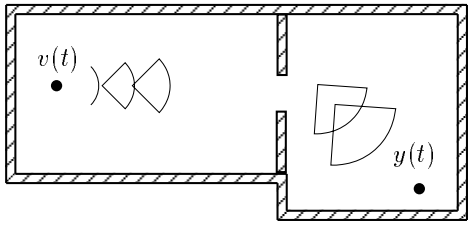

Figure 1: Source and receiver position within a building

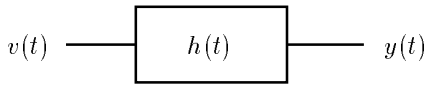

Figure 2: Input-output description of sound propagation within a building

For fixed source and receiver locations, all the spatial information is contained in the impulse response $h(t)$. Thus, the knowledge of the impulse response is sufficient to calculate the response $y(t)$ to a known source signal $v(t)$. The sound propagation in a room according to fig. 1 is thereby reduced to the black-box-model given in fig. 2.

The operation (2) is called a *convolution* and denoted by $*$. A simpler expression results in the frequency domain

$$Y(\omega) = H(\omega)V(\omega) \qquad (3)$$

where $Y(\omega) = \mathcal{F}\{y(t)\}$ is the Fourier transform of $y(t)$ and similar for $V(\omega)$. $H(\omega) = \mathcal{F}\{h(t)\}$ is called the *transfer function*. It is a complex valued function of the real frequency parameter ω. The meaning of the transfer functions is best described by the response to a sine-wave of angular frequency ω

$$v(t) = \sin(\omega t) \ . \qquad (4)$$

The response of a linear and time-invariant system is again a sine-wave, however, in general with different amplitude and phase shift

$$y(t) = |H(\omega)| \sin(\omega t + \phi(\omega)) \qquad (5)$$

Amplitude $|H(\omega)|$ and phase $\phi(\omega)$ are just the magnitude and phase of the complex transfer function $H(\omega)$

$$H(\omega) = |H(\omega)| \exp(j\phi(\omega)) \qquad (6)$$

Here $j = \sqrt{-1}$ is the imaginary unit. The transfer function $H(\omega)$ is a function of the frequency ω, since the response of a system to sine-waves with different frequencies will in general depend on ω. Transfer function and impulse response are equivalent descriptions of the black-box-system in fig. 2.

Neither convolution nor Fourier transformation of continuous-time signals can be implemented directly on a digital computer. Instead one works with sampled versions of $v(t)$ and $y(t)$. This turns (2) into a *discrete convolution* and the Fourier transformation into a discrete Fourier transformation, which can be efficiently computed with fast algorithms (*Fast Fourier transformation, FFT*). Discrete convolution and FFT are the building blocks for the digital processing of sound signals (and others). A more detailed description of digital signal processing techniques in acoustics is found in [9, 17]. In this paper, we will keep the continuous-time notation (2,3) for simplicity.

The black-box-description (2) is simple and effective because all the spatial information is concealed in the impulse response $h(t)$. The only remaining problem is how to obtain $h(t)$. There are three possibilities:

- The impulse response can be calculated from the Green's function of the wave equation (1). However, this is only feasible for very simple geometries and very simple boundary conditions.
- There are reliable and quick methods to measure the impulse response of a room. Examples are presented in section 2.2. Of course, measurements are restricted to existing buildings.
- The impulse response can be estimated in the planning stage from the room geometry and other data. Several methods are presented in sections 2.3.2, 2.3.3, and 2.3.4.

2.2. Measurement of Room Impulse Reponses

Methods for the measurement of room impulse responses are based on the transfer function description (3) of a room according to fig. 1. Note, that the general form of (3) is also valid for more complex geometries than the one depicted in fig. 1. Since the source signal $v(t)$ is known and the response $y(t)$ can be obtained by measurement, we can compute discrete time approximations to $V(\omega)$ and $Y(\omega)$ by the FFT-techniques discussed above. The impulse response follows by division and inverse Fourier transformation (also performed by FFT)

$$h(t) = \mathcal{F}^{-1}\left\{\frac{Y(\omega)}{V(\omega)}\right\} \qquad (7)$$

Practical measurement procedures have to consider the selection of appropriate source signals, background noise during the measurement and possible nonlinearities of the electro-acoustical equipment. A method for this purpose which has also been successfully applied to the analysis of other weakly nonlinear systems has been described in [7, 8, 14]. Its performance is shown here by some measurement examples. Fig. 3 displays the source function $v(t)$, which has been used in these measurements. It is a broadband signal with bounded amplitude and is called a *chirp signal*.

The impulse response of an 8 m^3 anechoic chamber is shown in fig. 4. The rapid decay is due to the highly absorbing walls. The impulse response of an office room with a typical duration of 0.1 s is shown in fig. 5. Note the different time scales in figs. 4 and 5. Fig. 6 displays the impulse response of two interconnected corridors with a total of 250 m^3 and no furniture. One of the corridors is connected to a staircase. The direct sound path was blocked, so that only reverberation occurs. Note again the time scale.

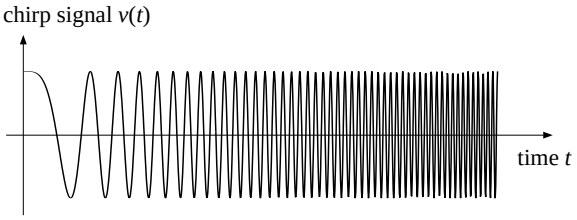

Figure 3: Chirp signal as source function

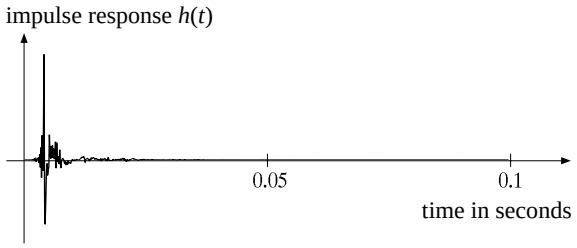

Figure 4: Impulse response of an anechoic chamber

Once obtained by measurements, we can use any of these impulse responses to mimic the response of the corresponding rooms by computing the convolution with an arbitrary source signal $v(t)$ according to (2). Although this is easily done on a small desktop computer, it is not true rendering, because it requires the building to be already available for measurements.

2.3. Modelling of Sound Propagation

To perform acoustic rendering of virtual buildings, we have to find means to compute the room impulse response from nothing more than the building geometry and the acoustic properties of the building materials. Four methods to perform this task will be discussed in some detail. The first method is based on the numerical evaluation of the wave equation (1) and is capable of a faithful simulation of all acoustical effects. A severe limiting factor is the computing power to perform these computations in real time. The other three methods are based on geometrical acoustics, i.e. they model sound propagation by sound rays which travel similarly to light rays in optics. Diffraction effects are neglected by this assumption. However, sound ray propagation can be treated with the same principle methods which are used with great success in visual rendering.

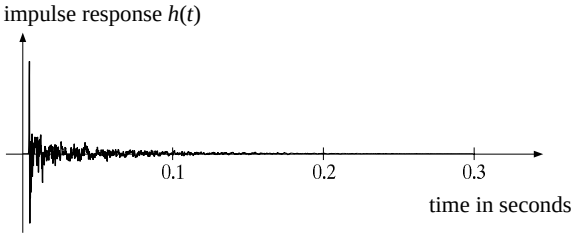

Figure 5: Impulse response of an office room

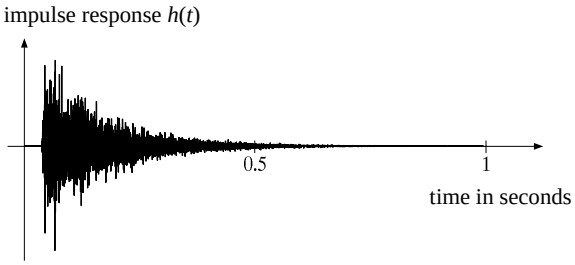

Figure 6: Impulse response of a system of interconnected corridors

2.3.1. Numerical Simulation

Conceptually, the most straightforward way to the simulation of the acoustic behaviour of a building is the numerical solution of the wave equation (1). We present a simple finite difference approach for a rough sketch of the idea. Detailed accounts on the numerical solution of partial differential equations can be found elsewhere [11]. The basic steps are the representation of the sound pressure $p(\mathbf{x}, t)$ on a discrete-time, discrete-space mesh with a finite number of nodes and the approximation of the differential operators in (1) by difference operators defined in terms of the mesh nodes.

For simplicity, we restrict the derivation to one spatial dimension, i. e. $p(\mathbf{x}, t) = p(x, t)$ with the scalar space variable x. The nodes of the mesh are given by $x = nh$ and $t = kT$, where h and T are the step sizes in space and time, respectively, and n and k are the corresponding discrete variables. The node values of $p(x, t)$ at these locations are denoted by $p_{n,k}$. A simple second order approximation of the differential

operators is obtained by

$$\frac{\partial^2 p}{\partial t^2} \approx \frac{1}{T^2}\left[p_{n,k+1} - 2p_{n,k} + p_{n,k-1}\right] \quad (8)$$

$$\Delta p = \frac{\partial^2 p}{\partial x^2} \approx \frac{1}{h^2}\left[p_{n+1,k} - 2p_{n,k} + p_{n-1,k}\right] \quad (9)$$

When the values of the sound pressure $p_{n,k}$ for the time steps k and $k-1$ are already calculated, then we obtain an explicit expression for the approximate value $p_{n,k+1}$ at the new time step $k+1$ by replacing both sides of (1) with their discrete time approximations (8,9)

$$p_{n,k+1} = \lambda^2 p_{n-1,k} + 2(1-\lambda^2)p_{n,k} + \lambda^2 p_{n+1,k} - p_{n,k-1}\,, \quad (10)$$

where

$$\lambda = \frac{cT}{h} \quad (11)$$

is the ratio between the distance cT that the sound wave travels within one time step of length T and the spatial step size h. The most simple formula results for $\lambda = 1$

$$p_{n,k+1} = \left(p_{n-1,k} + p_{n+1,k}\right) - p_{n,k-1}\,. \quad (12)$$

A graphical representation of this discretization scheme is shown in fig. 7. The value $p_{n,k+1}$ is obtained by the sum of the neighbours of $p_{n,k}$ and by subtracting $p_{n,k-1}$.

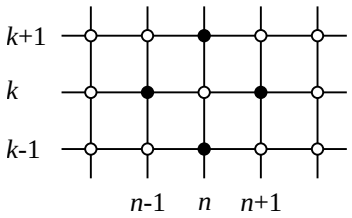

Figure 7: Second order discretization scheme for the wave equation

The generalization of this discretization scheme to three spatial dimensions is straightforward. With the vector $\mathbf{n} = [n_1, n_2, n_3]$ of discrete space variables in each direction and with $\lambda^2 = 1/3$ follows

$$p_{\mathbf{n},k+1} = \frac{1}{3}\sum_{\mathbf{m}} p_{\mathbf{m},k} - p_{\mathbf{n},k-1}\,. \quad (13)$$

Here $\sum_{\mathbf{m}}$ denotes summation over all spatial neighbours of $p_{\mathbf{n},k}$.

This discretization scheme has been used in [12, 13] for the simulation of air pressure wave propagation in a model of three rooms separated by open doorways. It has been shown that the effects of reflection at the walls and diffraction at a pillar could be numerically reproduced.

A major drawback of this method is the numerical expense associated with the update of the 3-D mesh in every time step. To illustrate the situation, we estimate the computing power in floating point operations per second (flops) which is necessary for a real time simulation of the propagation of a sound wave with a frequency f_0 of 3 kHz in a building with the volume of 100 m^3. The maximum time step size according to the sampling theorem is $T = 1/(2f_0)$. In practice, T should be smaller to avoid numerical dispersion. Since the update of one node value according to (13) requires 7 floating point operations, the computing power for one node is given by $N_0 = 14 f_0 = 42$ kiloflops. Furthermore, the step sizes in time and space are linked by the requirement $\lambda^2 = 1/3$, which yields the spatial step size ($c \approx 340\mathrm{m/s}$)

$$h = \sqrt{3}\,cT = \frac{\sqrt{3}}{2}\frac{c}{f_0} \approx 10\,\mathrm{cm}$$

The number M of nodes in the volume V is $M = V/h^3 = 100/(0.1)^3 = 10^5$, which amounts to a total computing power of $N = M N_0 = 4.2$ Gigaflops for this rather modest situation. Since the number of flops grows linear with the volume V and proportional to the fourth power of the frequency f_0, we cannot expect real time rendering of buildings by numerical simulation in the near future. However, numerical simulation is feasible for low frequencies, and it is the only method which describes diffraction correctly.

2.3.2. Image Sources

We have just seen, that real time simulation by finite difference approximation of the wave equation has a price that is hard to pay. We can cut this price, if we are willing to sacrifice some of the benefits of physical modelling. In a description by the wave equation, any point in the sound field can act as a source and vice versa, we can pick up the sound at any point of the wave field. Such a flexibility is not always required. A minimal approach would be, to consider only one source and one listener. Also some acoustical effects can be neglected under certain circumstances. For example, if the room sizes are large compared to the wavelength, diffraction plays only a minor role. In most practical situations, this is true for frequencies above 1 kHz. If also reflection at the walls is considered in a simplified way, then a geometric model of wave propagation can be adopted.

The great advantage of this approach is, that geometric models are well established in visual rendering and consequently some proven algorithms have been developed. However, acoustical rendering requires some additional consideration, which is not necessary in visual rendering. The most important topic is the finite propagation speed of sound waves compared to the almost instantaneous spread of light. Furthermore, damping of light intensity is negligible for small distances, whereas sound waves undergo severe attenuation. The adaption of visual rendering schemes for acoustics has to take these effects into account.

The simplest way of geometric modelling is the method of image or mirror sources. It is a classical

method in physics for the treatment of boundary value problems. An early implementation as a computer program for acoustic simulation has been described e.g. by [1].

Fig. 8 shows a rectangular room with one source and one receiver. The propagation is modelled as the superposition of direct, single and multiple reflected rays from the source to the receiver. Rather than constructing all the different propagation paths within the room, the reflections are modelled as direct paths from additional imaginary sources outside the room. They are positioned in such a way, that their associated direct paths coincide with the reflections inside the room. These positions can be obtained by symmetry considerations. The arrangement in fig. 8 models only a few of the first reflections. For a realistic simulation, it would have to be extended to an infinite continuation in three dimensions.

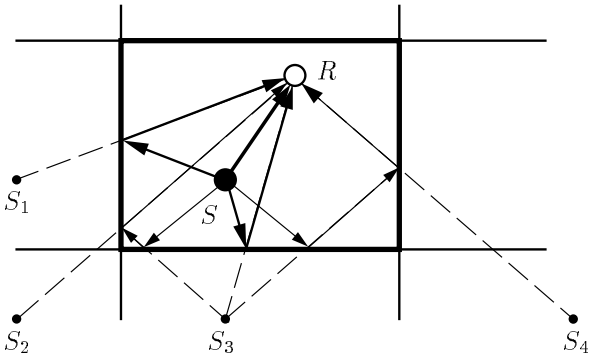

Figure 8: Modelling of wave propagation by image sources (receiver R, source S, image sources S_i, $i = 1, 2, 3, 4, \ldots$)

For a simple visual model, we can imagine the room in fig. 8 to be a mirror cabinet with ideally reflecting walls. The response to an impulse at the source is then simply the sum of delayed impulses, where the delay time is given by the total length of the reflection path divided by the speed of sound. The length of the reflection path is obtained from fig. 8 as the distance of the corresponding image source to the receiver position.

For acoustic modelling, we have to consider the mirror sources as emitters of spherical pressure waves. The sound pressure at the receiver caused by the source with number i at distance d_i emitting a pulse $\delta(t)$ is then given by (see [1])

$$p_i(t) = \frac{1}{4\pi d_i}\delta\big(t - \frac{d_i}{c}\big) \tag{14}$$

When all walls are ideal reflectors then the total impulse response in terms of sound pressure is given by

$$h(t) = \sum_i p_i(t) \tag{15}$$

where the sum is taken over all surrounding image sources within a certain region (ideally infinitely many). The response to an arbitrary source signal $v(t)$ follows from (2) and (15)

$$y(t) = \sum_i \frac{1}{4\pi d_i} v(t - \frac{d_i}{c}) \tag{16}$$

as the sum of attenuated echos with delay d_i/c.

Angle and frequency independent wall absorption can be modelled by inclusion of appropriate absorption coefficients $b_{i\ell}$ which describe the absorption of the signal from source i at the ℓ-th reflection.

$$h(t) = \sum_i b_i p_i(t)\,, \qquad b_i = \prod_{\ell=1}^{L_i} b_{i\ell} \tag{17}$$

L_i is the total number of reflections for source i. The model can also be extended to frequency dependent reflection, when the reflection coefficients $b_{i\ell}$ are replaced by the impulse responses $b_{i\ell}(t)$ of systems which describe the reflection processes

$$h(t) = \sum_i b_i(t) * p_i(t)\,, \tag{18}$$

where

$$b_i(t) = b_{i1}(t) * b_{i2}(t) * \cdots b_{iL_i}(t) \tag{19}$$

is a multiple convolution model for the frequency dependent reflection at a total of L_i surfaces. The description of these multiple reflections is simpler in the frequency domain. We obtain by Fourier transformation the transfer function

$$H(\omega) = \sum_i B_i(\omega) P_i(\omega)\,, \qquad B_i(\omega) = \prod_{\ell=1}^{L_i} B_{i\ell}(\omega) \tag{20}$$

with the transfer functions $B_{i\ell}(\omega)$ of each frequency dependent reflection at the ℓ-th surface.

The image source model is conceptually simple, but it requires a high number of image sources in order to model the complete impulse response of a room. If we want to compute the impulse response of a room with volume V up to a length of t_0 seconds, we have to extend the room in either directions until the image sources fill a sphere of radius ct_0. The number of image sources is roughly given by the relation of the volume of the sphere to the volume of the room

$$N = \frac{\frac{4}{3}\pi(ct_0)^3}{V} \tag{21}$$

For $t_0 = 1\,\mathrm{s}$ and $V = 100\,\mathrm{m}^3$, a total of 1.7 million image sources results. This number can be handled for frequency independent reflections, but the consideration of frequency dependent reflections by multiple convolutions becomes impractical. The numerical expense increases also when multiple sources have to be considered. Additionally, any change of position of

the receiver or one of the sources requires a recalculation of the impulse response.

Furthermore, obtaining the locations of the image sources is a tedious procedure if the building geometry is more complicated than the rectangular box model shown in fig. 8. This procedure is described in [3] for polyhedra with arbitrary shape. The locations of possible image sources are computed by a tree-structured algorithm and tested for validity, proximity and visibility. Only valid sources which are visible and sufficiently close to the listener contribute to the impulse response. However, the computing time required to carry out this procedure is a severe limiting factor.

On the other hand, the image source method is capable of modelling the interference between reflected sounds, if the model is formulated in terms of pressure rather than intensity [4].

2.3.3. Ray tracing

The ray tracing method is also based on geometrical acoustics. It is similar to the image source method, since it also considers the possible reflection paths by which the sound travels to the receiver.

The source model emits a number of sound rays in all directions. A certain amount of energy is assigned to each ray. The propagation path of each ray is traversed and the energy losses due to absorption in the air and at each reflection are recorded. All rays that hit the receiver within a certain amount of time contribute to the impulse response.

This procedure is very similar to the well known application of ray tracing in visual rendering, except for the different physical effects, which account for the energy losses along the path. The advantage of the ray tracing method over the image source method lies in the treatment of arbitrary complex 3D shapes without requiring excessive computing power.

The numerical modelling of the reflections is done similar as in the image source method. The effect of ideal reflections, frequency independent and frequency dependent absorption can be calculated by the same principal algorithms as in (15,17,18).

2.3.4. Radiosity

The image source method and the ray tracing method depend on the position of the receiver. Whenever the receiver changes the position within a building, the impulse response has to be computed anew. The radiosity method is a view independent algorithm. It requires knowledge about the building geometry and about the sound sources. From this input, an approximation of the sound field within a building can be constructed.

The radiosity method has been used for a long time in thermodynamics to calculate the radiation heat transfer between surfaces [16]. Later, it became one of the working horses of computer graphics for visual rendering [5]. Its adaption to auditory rendering is described in [15].

When applied to light waves, the radiosity method describes an equilibrium for the energy distribution within an enclosure, that is constructed from a finite set of surfaces (patches). This finite set of patches covers the surface of the enclosure completely, so that all energy flows are considered. Any openings are also described as a partial surface with appropriate properties. The rate B_ℓ at which energy leaves the surface ℓ is derived from the conservation of energy. Since the energy exchange between the surfaces is assumed to happen instantaneously, B_ℓ is always given as the sum of the energy E_ℓ, which is generated by the surface ℓ itself and the reflection of the irradiations from the other surfaces.

$$B_\ell = E_\ell + \rho_\ell \sum_i B_i F_{\ell i} \tag{22}$$

B_i is the rate at which energy leaves the other surfaces, $F_{\ell i}$ is the fraction of B_i which arrives at surface ℓ and ρ_ℓ is the reflection coefficient at surface ℓ. $F_{\ell i}$ depends on the size and the relative orientation of the surfaces ℓ and i. It is also called *form factor*. We will not describe the meaning of (22) and the calculation of the form factors in more detail, because these topics have been discussed thoroughly in [5, 16] and many other references on heat radiation and computer graphics.

The two basic differences between radiation transfer and sound propagation are

- damping of the sound intensity by the air,
- finite propagation speed of sound.

Since we are not dealing with real or virtual point sources as in the image source or ray tracing method, there is no attenuation with increasing distance. The damping considered here is due to conversion of kinetic to thermal energy by friction and molecular effects in the air. The relation between the sound pressure $p(d)$ at distance d from the source and the source pressure p_0 is given by an exponential law

$$\frac{p(d)}{p_0} = \exp(-\alpha d) \; . \tag{23}$$

α is the *absorption* or *extinction* coefficient and depends on the air temperature and humidity. We may describe this effect concisely by the air attenuation factor [15] between the surfaces ℓ and i with the distance $d_{\ell i}$

$$\Phi_{\ell i} = \exp(-\alpha d_{\ell i}) \; . \tag{24}$$

$\Phi_{\ell i}$ describes the fraction of the rate of energy $B_i F_{\ell i}$ which arrives at surface ℓ in the presence of transmission losses. Thus the energy balance for sound propagation in a lossy medium is given by

$$B_\ell = E_\ell + \rho_\ell \sum_i B_i F_{\ell i} \Phi_{\ell i} \; . \tag{25}$$

We see that damping of the sound intensity by the air can be modelled by the air attenuation factor, which acts on B_i in the same way as the form factor $F_{\ell i}$ does.

A more fundamental change to the equilibrium equation (22) is necessary to take the finite propagation speed into account. Equations (22) and (25) show no time dependence, since it is understood that any changes in the system lead instantaneously to a new equilibrium. No dynamical or memory effects have to be considered with respect to the speed of light.

This is no longer true for sound propagation. The energy transport from distant surfaces is not only subject to transmission losses but also to time delay. Sound energy which leaves a surface does not arrive in the same time instant at other surfaces. To account for the travel time of the sound waves, we have to formulate the conservation of energy with time dependent quantities. The time delays have to be expressed by the distance between the surfaces and the speed of sound. This leads to the following time dependent formulation of the energy balance, which includes propagation losses and delay

$$B_\ell(t) = E_\ell(t) + \rho_\ell \sum_i B_i\big(t - \frac{d_{\ell i}}{c}\big) F_{\ell i} \Phi_{\ell i} \,. \quad (26)$$

The energy exchange rates $B_\ell(t)$ and the emission rates $E_\ell(t)$ are now functions of time.

A rough sketch of the algorithm is given in fig. 9. The input data are the building geometry and the reflection coefficients of the surfaces. They serve to calculate the form factors and the air attenuation factors. This is the most time consuming task of the whole procedure. However, for a given geometry, these factors are calculated only once since they do not change with time. Once they are available, the energy balance (26) can be solved repeatedly by updating the time variable t. We can use the radiosity method for the calculation of impulse responses, if we place the receiver as an additional small patch into the sound field. Alternatively, we can also calculate the response to source signals at the receiver position directly.

When the receiver position changes with time, then the form factors and air attenuation factors from the building surfaces to the receiver have to be updated as well. All the factors between the building surfaces remain unchanged. The source functions enter this model via the emission rates $E_\ell(t)$. Since any surface can act as a source, also moving sources can be easily realized without updating any factors.

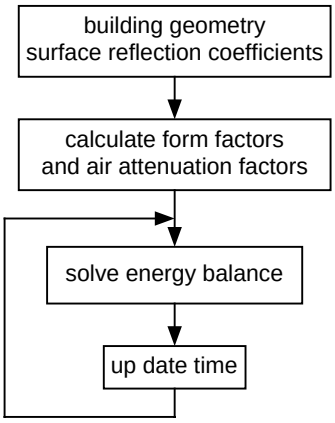

Figure 9: Algorithm for the radiosity method

3. SPATIAL HEARING

So far, we were only concerned with sound propagation in pure technical terms. The receiver was nothing more than a point in space, where the sound pressure is recorded, e.g. by a microphone. However, acoustic rendering is more than just the computer simulation of the physical effect of sound wave propagation in complex environments. If we aim at a true threedimensional impression of a virtual sound field, then also questions of human sound perception have to be considered. At this point, we leave the secure basis of physical science and enter the realm of psychoacoustics. A tremendous amount of research has been conducted in this area. For an introduction, see [2] and the references contained therein. We will quote only some of the common concepts to describe what is called spatial hearing. This term can be roughly defined by some questions, which naturally arise, when we try to evaluate the human sound perception capabilities:

- How can we tell the direction of a sound source?
- How can we tell the distance of a sound source?
- How can we separate different sound sources, even in the presence of strong background noise?

3.1. Interaural Time and Intensity Differences

The ability to tell the direction of a sound source can be traced to two basic differences in the sound perception with both ears [2]:

- The interaural time difference (ITD) is the difference in the arrival time of a wave front between the two ears.
- The interaural intensity difference (IID) is the difference in perceived intensity between the two ears.

Both kinds of differences act together to provide spatial information about the direction of arrival of a wave front. The human hearing system can resolve time differences between both ears for signals up to frequencies of 1 kHz. The relation between the relative time difference of two signals and the direction of arrival is based on a simple geometrical argument. In this sense, the ears act as a two-microphone-array. For

low frequencies, there is no noticeable intensity difference between both ears, due to diffraction at the head. For frequencies with a wavelength shorter than the diameter of the head (typically 1.5 kHz and higher) diffraction becomes negligible and an intensity difference between both ears is perceived.

There is evidence [2], that the human hearing system combines ITD and IID for the location of sound sources. ITD dominates for low frequencies, while IID works for high frequencies.

Application of these concepts to acoustic rendering calls for post processing of the receiver signal. Once the response to a sound source is determined by the methods of the previous section, the receiver signal has to be split into two signals for the left and for the right ear respectively. These signals are delayed and filtered in order to produce the interaural time and intensity differences which correspond to the direction of the source signal.

3.2. Head-Related Transfer Functions

The ITD and IID just considered are only the starting point for an interpretation of the human hearing system. They allow to discriminate the source position in terms of left – right, but they do not explain why we can also locate sound sources above and below or in front or in the back of our head.

These effects can only be explained when the shape of the pinnae is taken into account. They act as a filter between the sound field outside the ears and the eardrum. Technically, their effect on spatial hearing can be described by a transfer function, the so called *head-related transfer function (HRTF)*. Of course there are two HRTFs, one for each ear. They are defined as the frequency dependend relation of the sound pressure at the eardrum to the sound pressure at the receiver location when the listener is absent. Much effort has been devoted to measurements of the HRTFs and corresponding data sets for typical HRTFs are available [2].

Their application to acoustical rendering is straightforward. After the receiver signal has been split into a left and right ear signal and both signals have been treated with the appropriate ITD and IID information, the final step is the filtering with the corresponding HRTF for each ear. When these signals are delivered by ear phones directly to the listener's eardrum, they carry the necessary spatial cues to enable spatial hearing of virtual sound sources.

4. IMPLEMENTATION

Finally, we will briefly address some issues in the practical implementation of the methods and concepts described above. The first topic is the proper choice of the sound propagation method. The most general method with the highest potential of accuracy is the numerical simulation of wave propagation by a suitable algorithm like the finite difference method. Unfortunately, the numerical expense increases with the fourth power of the highest signal frequency, which renders this method suitable for low frequencies only.

The required number of floating point operations for the geometrical methods is independent from the signal frequency. However, they rely on the assumption of sound ray propagation and do not account for diffraction. This effect is only negligible for wavelengths which are small compared to the room dimensions (roughly above 1 kHz).

These arguments suggest to split the source signal into a low frequency and a high frequency part. The low frequency part can be modelled by a numerical method, which considers diffraction. The high frequency part can be treated with a geometrical method or with a combination of the ray-tracing and the radiosity method [10] (see fig. 10).

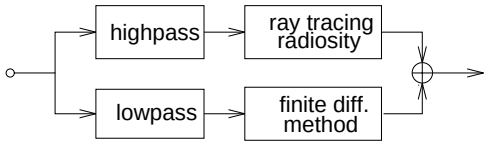

Figure 10: Combination of numerical simulation for low frequencies and geometric simulation for high frequencies

Finally, we put the pieces together to form a complete acoustic rendering system. We have not mentioned sound modelling, because it is not peculiar to our application. Sounds may be prerecorded or synthesized according to some model. The propagation from the source to the receiver is traced with a system according to fig. 10 or a part thereof. The spatial hearing capabilities of a human listener are taken into account by splitting the received signal into two channels for the left and the right ear, respectively. Both channels receive the proper delay and filtering to produce the desired psychoacoustic effects (see fig. 11).

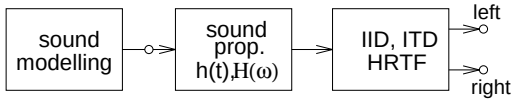

Figure 11: Combination of sound propagation and human preception modelling

5. CONCLUSION

As yet, there is no single method which is capable of all aspects of acoustical rendering in complex environments as they appear in buildings. Sound propagation

simulation by numerical methods with desktop computers is restricted to low frequencies. The situation may change in the future with the use of more efficient numerical algorithms and the availability of more computing power at competitive prices.

The image source method is a straightforward and computational feasible method for very simple room geometries. Advanced geometric methods like the ray tracing or the radiosity method can be implemented on modern computers as long as the number of rays in the ray tracing method or the number of patches in the radiosity method are chosen reasonably small. However, geometric methods neglect diffraction effects which are present for large wavelengths in the order of the room dimensions. Since numerical and geometrical methods are complementary, it seems worthwhile to combine both methods.

In addition to the simulation of sound propagation, also an impression of spatial hearing has to be achieved. A proper modelling of the interaural time and intensity differences and of the effects of the pinnae by head related transfer functions provide the cues to binaural hearing.

Some aspects of acoustical rendering have not been considered. So far, only the effect of sound wave reflection at walls has been modelled, but no transmission of sound was taken into account. This would be the prerequisite for the assessment of sound absorbing materials and for the determination of the total sound level within a building. Also the question of accuracy was not raised. Are these real time methods for acoustic rendering only crude approximations to reality or do they have the potential to make quantitative predictions? Is the sound quality good enough for a real time judgment of the properties of a concert hall or an auditorium?

The methods and concepts, which were reviewed in this paper, provide an exciting extension to the visual rendering capabilities of modern design programs. Their usefulness as a design tool in its own right has yet to be established. Not only further laboratory research but also user experience from a growing number of implementations will promote this process.

6. REFERENCES

[1] J. B. Allen and D.A. Berkley. Image method for efficiently simulating small-room acoustics. *J. Acoust. Soc. Am*, 65(4):943–950, 1979.

[2] D. R. Begault. *3-D Sound for Virtual Reality and Multimedia*. Academic Press, Cambridge, USA, 1994.

[3] J. Borish. Extension of the image model to arbitrary polyhedra. *J. Acoust. Soc. Am*, 75(6):1827–1836, 1984.

[4] S. M. Dance, J. P. Roberts, and B. M. Shield. Computer prediction of sound distribution in enclosed spaces using an interference pressure model. *Applied Acoustics*, 44:53–65, 1995.

[5] C. M. Goral, K. E. Torrance, D. P. Greenberg, and B. Battaile. Modeling the interaction of light between diffuse surfaces. *Computer Graphics*, 18(3):213–222, 1984.

[6] J. H. Hahn, H. Fouad, L. Gritz, and J. W. Lee. Integrating sounds and motions in virtual environments. *Presence: Teleoperators and Virtual Envirnoments*, to appear.

[7] F. Heinle, R. Rabenstein, and A. Stenger. Measuring the linear and non-linear properties of electro-acoustic transmission systems. In *Proc. of Int. Workshop on Acoustic Echo and Noise Control (IWAENC'97)*, London, 1997. to appear.

[8] F. Heinle and H. W. Schüßler. Measuring the performance of implemented multirate systems. In *Proc. Int. Conf. Acoustics, Speech and Signal Proc. (ICASSP 96)*, 2754–2757. IEEE, 1996.

[9] S. M. Kuo and D. R. Morgan. *Active Noise Control Systems – Algorithms and DSP Implementations*. John Wiley & Sons, New York, 1996.

[10] T. Lewers. A combined beam tracing and radiant exchange computer model of room acoustics. *Applied Acoustics*, 38:161–178, 1993.

[11] A. Quarteroni and A. Valli. *Numerical Approximation of Partial Differential Equations*. Springer-Verlag, Berlin, 1994.

[12] L. Savioja, J. Backman, A. Järvinen, and T. Takala. Waveguide mesh method for low-frequency simulation of room acoustics. In *Proc. Int. Congress on Acoustics (ICA '95)*, 637–641, 1995.

[13] L. Savioja, T. J. Rinne, and T. Takala. Simulation of room acoustics with a 3-d finite difference mesh. In *Proc. Int. Computer Music Conference (ICMC '94)*, 463–466, 1994.

[14] H. W. Schüßler and F. Heinle. Measuring the properties of implemented digital systems. *FREQUENZ*, 48(3-7), 1994.

[15] J. Shi, A. Zhang, J. Encarnacao, and M. Göbel. A modified radiosity algorithm for integrated visual and auditory rendering. *Computers & Graphics*, 17(6):633–642, 1993.

[16] R. Siegel and J. R. Howell. *Thermal Radiation Heat Transfer*. Hemisphere Publishing Corporation, Washington, 1981.

[17] L. J. Ziomek. *Fundamentals of Acoustic Field Theory and Space-Time Signal Processing*. CRC Press, Boca Raton, 1995.