Audio-visual Virtual Reality System for Room Acoustics

Eduard Deines[1], Martin Hering-Bertram[2], Jan Mohring[2], Jevgenijs Jegorovs[2], and Hans Hagen[3]

1 **University of California, Davis, USA**
edeines@ucdavis.edu
2 **Fraunhofer ITWM, Kaiserslautern, Germany**
{martin.hering-bertram,jan.mohring,jegorovs}@itwm.fraunhofer.de
3 **Department of Computer Science, University of Kaiserslautern**
hagen@informatik.uni-kl.de

Abstract

We present an audio-visual Virtual Reality display system for simulated sound fields. In addition to the room acoustic simulation by means of phonon tracing and finite element method this system includes the stereoscopic visualization of simulation results using a 3D back projection system as well as auralization by use of a professional sound equipment. For auralization purposes we develop a sound field synthesis approach for accurate control of the loudspeaker system.

1998 ACM Subject Classification I.3.5 Computational Geometry and Object Modeling, J.2 Physical Sciences and Engineering

Keywords and phrases Special Relativity, Minkowski, Spacetime, Virtual Camera

Digital Object Identifier 10.4230/DFU.SciViz.2010.303

1 Introduction

For architectural planning of class rooms, theaters or concert halls the auditive impression of resulting rooms is very important. Auralization provides such a feasibility. It can be considered as the auditive variant of visualization, providing insight into acoustic properties of a room and its reasons. Virtual Reality environments enable an immersive representation of computer-generated scenes. Integrating acoustic simulation, visualization, and auralization into the design process aims at interactive design and immediate exploration of the virtual model.

In this work we present an audio-visual Virtual Reality system which integrates computer-aided simulation, visualization and auralization of acoustics in a room model. For auralization purposes we combine a FEM based approach and the phonon tracing algorithm [5] in order to obtain a realistic impression of the sound perceived at given listener positions. The wave field synthesis approach enables a correct auditive rendering of the convolved signals on a professional sound equipment. With our system, walkthrough of the listener positions is possible, such that a visual and auditive impression of the scene can be provided.

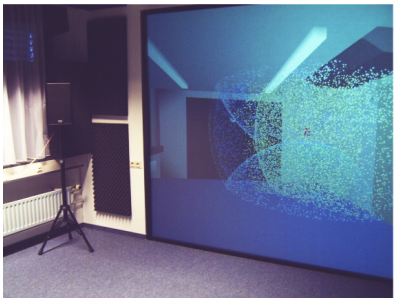


Figure 1 Room acoustic visualization on the VR system.

Scientific Visualization: Advanced Concepts.
Editor: Hans Hagen; pp. 303–320
Dagstuhl Publishing
Schloss Dagstuhl – Leibniz Center for Informatics (Germany)

From our visualization, the effect of different materials on the spectral energy distribution can be observed. The first few reflections already show whether certain frequency bands are rapidly absorbed. The absorbing materials can be identified and replaced in the virtual model, improving the overall acoustic quality of the simulated room. A wave front is the pressure wave corresponding to a unit pulse. In order to visualize the wave fronts spread out from the sound source we use both particles (spheres) and surface elements, color coded using the energy spectra of the corresponding virtual sound particles. After a sufficient time period, a great number of reflections has occurred, such that individual wave fronts cannot be identified, anymore. The darker the color of spheres gets, the more energy has been absorbed by the different materials. Applications of our work include (but are not limited to) the acoustic improvement during architectural design, and equipment phases of class- and congress rooms. sec Our paper is organized as follows. In Section 2 we review related work. Section 3 describes the acoustic virtual reality system. Section 4 outlines our simulation algorithm, implying the phonon tracing approach and the finite element method. Thereafter, we present our synthesis algorithm in Section 5, follows by visualization of the simulation results (Section 6). In Section 7 we show an application of our sound visualization approach, before we conclude our work.

2 Related Work

In room acoustics there are two main approaches simulating the propagation of sound. The first approach is based on the wave equation which is solved numerically, for example by use of the finite element method (FEM). The simulation results are very accurate, but the complexity increases drastically with the highest frequency considered, since a volume grid with $O(n^3)$ cells needs to be constructed where n is proportional to the highest frequency. The time complexity for solving this is typically $O(n^3\ log\ n^3)$. Hence, the wave model is suitable for low frequencies only.

The second approach, known as geometric acoustics, describes the sound propagation by sound particles moving along a directed ray. There exists a variety of such methods for simulating room acoustics. They are mostly based on optical fundamentals, and make use of approaches developed there. Two classical methods for acoustic simulation are the image-source method [1, 6] and the ray tracing method [25, 26]. Due to the shortcomings of the two classical approaches, continuative approaches have been developed in recent years. Mostly, they employ parts of the classical schemes or a combination of them. One approach that makes use of advantages of image-source and ray tracing is introduced in [46]. Here, the visibility check of the image-source algorithm is performed via ray tracing. Beam-tracing methods [12, 13, 32] overcome the aliasing problem of classical ray tracing by recursively tracing pyramidal beams, implying the need for highly complex geometric operations, still ignoring diffraction effects at low frequencies. An approach for calculation of edge diffraction in room acoustics is presented in [45, 30]. To overcome the dependency of the simulation on the receiver position the radiosity method was extended to be used in room acoustics [40, 24]. Due to the computation complexity these methods do not seem practical for large environments. Newer approaches cope with complexity by exploiting GPU hardware accelerating the simulation calculations [17]. Approaches utilizing the photon mapping [18] also exist [20, 5].

The aim of audio simulation is to estimate acoustic properties, such as reverberation time and the reproduction of acoustic benchmarks. Auralization is the process of producing audible impression of a room. It can be realized in two different ways. The first one is the ”direct room impulse response rendering” approach, where the room impulse response (RIR) is measured or modelled and afterwards convolved with the anechoic signal. In the DIVA auralization system [41, 29] a different approach is proposed, the "parametric room impulse response rendering". Here, the RIR is not calculated before the auralization process. Instead, a set of either perception-based or physic-based parameters for auralization is defined. In [34] an audio-rendering system for use in immersive virtual environments

is introduced. It is optimized for efficient rendering of moving sound sources. A survey of existing auralization systems is given in [41].

There are two major sound rendering approaches: wave field synthesis [4] and binaural synthesis [3]. The latter approach uses headphones to reproduce the sound field simulated at the ears of the virtual listener. The advantages are cheap equipment and low number of channels to be computed. However, the listener is bothered by the headphones and synthesis requires deep knowledge of the acoustics of the human head and psycho-acoustics. Therefore, we apply wave field synthesis. The underlying theory is based on Huygens' principle, or more mathematically, on the fact that solutions of the Helmholtz equation correspond to distributions of sound sources on the surface of the domain (Kirchhoff-Helmholtz integral, [14]). This continuous distribution is approximated by a finite number of loudspeakers. The higher the number of speakers and the lower the frequency the larger is the sweet spot where the approximation is accurate. Usually, only the sound field of a 2D listening plane is synthesized as this requires already quite large number of speakers. For instance, Fraunhofer IDMT furnished a cinema with 192 speakers to achieve good reproduction at all seats [22]. For our purpose it is enough to have a good reproduction close to the proband's head the position of which is, except for small movements, fixed. The present system comprises currently 4 full-band speakers and a bass. In a future system additional mid and high-band speakers will be added.

The existing acoustic visualization techniques can be classified into two groups. The methods of the first group consider the propagation of the sound waves or rays inside the room independent of the listener position. Whereas the algorithms in the second group base on the measured or simulated room impulse response for individual listener. Yokota et. al. [47] visualized sound propagation of 2-D sound fields using a difference time domain method. Petrausch and Rabenstein [37] introduce a program for simulation and OpenGL based visualization of 2-D sound wave propagation in real time. Tokita and Yamasaki [44] depicted particle displacements resulted from a wave-based simulation on a rigid 3-D grid in a cube shaped room. Lokki [29] presented a visualization using the image-source method. Lokki and Nenonen [28] utilized cave equipment for immersive visualization of trace paths and particle paths propagating inside closed rooms. In [39] Pulkki and Lokki present an approach visualizing edge diffraction. Funkhouser et. al. [12] used visualization of source points, receiver points, pyramidal beams, reverberation path etc. in order to understand and evaluate their acoustic modeling method. Lauterbach et. al. [27] showed the sound propagation resulting from frustum tracing. Sound visualization approaches utilizing the phonon tracing algorithm as well as comparative visualization of the phonon tracing and an FEM based solver are presented in [10] and [9] respectively. Khoury et. al. [21] represented the sound pressure levels inside the room by means of color maps. Additionally, the precedence effect (or "law of the first wave front") was analyzed by using isosurfaces. Stettner et. al. [43] visualized acoustic metrics such as clarity and definition as well as spatial impression by use of specific icons. Monks et al. [33] introduced an interactive optimization system for acoustic design. The results are presented by means of icons depicting early decay time (EDT), interaural cross-correlation coefficient (IACC), and bass ratio (BS). Furthermore, the sound strength was displayed at room surfaces using color to indicate the sound-level data at different time moments. Several approaches for visualization of measured sound intensity also exist [36, 11, 31].

For room acoustic modeling we combine in our system a FEM based method (for low frequency) and the phonon tracing approach [5] (for middle and high frequency). By this means we can consider the diffraction and interference, which can not be neglected for low frequencies, and manage the complexity problem of FEM by middle and high frequencies. For sound rendering purposes we introduce a wave field synthesis approach. In addition to the auralization we provide a visualization of sound wave propagation inside the virtual room.

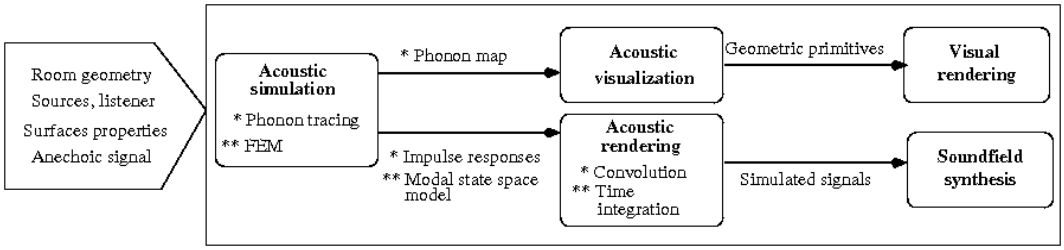


Figure 2 Modules of our audio-visual Virtual Reality system.

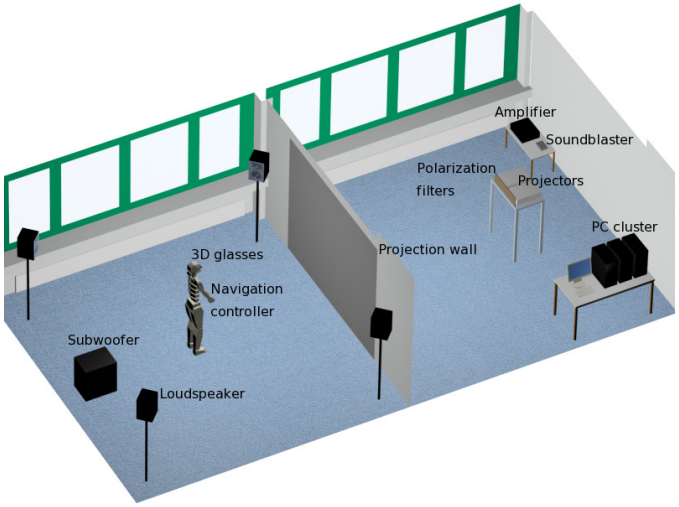


Figure 3 Visual and audio Virtual Reality display system.

3 Acoustic Virtual Reality System

In this section we present our visual and auditive virtual reality display system. A schematic overview is given in Fig. 2. As input the system requires the geometry model of the room, the absorption or reflection properties of the room surfaces, sound sources characteristics, and listener positions. The module ”Acoustic simulation” computes the modal state space model (low frequencies) by use of the Finite Element Method (FEM) described in Section 4.2, and the room impulse responses (RIR) and the phonon map (middle and high frequencies) by means of the phonon tracing algorithm introduced in Section 4.1. On the basis of this information the anechoic source signal is modified for sound synthesis in the ”Acoustic rendering” module. Thereafter, the soundfield synthesis is performed utilizing the acoustic hardware. The ”Acoustic Visualization” module provides the visualization of wave propagation from the sound source. The graphical rendering ("Visual rendering" module) is implemented utilizing a stereoscopic back projection system. By use of our VR system, a walkthrough of the listener positions is possible, such that a visual and auditive impression of the scene can be provided.

The hardware of the Virtual Reality system includes the 3D back projection system and the acoustic system (see Fig. 3). The 3D back projection system is composed of two high resolution digital D-ILA projectors for displaying mono and passive stereo signals, two circular mechanical shiftable polarization filters, and one projection wall, suitable for polarization filters. By means of

Table 1 Technical details of the VR System.

3D Powerwall	
2 projectors	1400x1050, 1500 ANSI Lumen
1 screen	plastic film suitable for polarization filters (2.88m x 2.30m)
2 filters	circular polarization filters
glasses	plastic glasses with circular polarization inserts
4.1 Acoustic system	
1 sound card	7.1 USB sound blaster
4 loudspeakers	two-way universal loudspeakers
1 subwoofer	active subwoofer with integrated amplifier
1 amplifier	four-channel amplifier with integrated programmable FIR Filters
PC cluster	
1 master	Dual Xeon 3.0 GHz, 2GB RAM, PCI-Express, NVidia Quadro FX1300
2 render nodes	Single Xeon 3.0 GHz, 2GB RAM, PCI-Express, NVidia Quadro FX5400

this system a stereoscopic rendering of virtual scenes is possible, observed by wearing polarizing eyeglasses. The acoustic system contains a 7.1 soundblaster and a professional surround sound equipment. This surround sound system on its part includes four two-way loudspeakers and one subwoofer facilitating the localization of the virtual sound source. The system is driven by the PC-cluster composed of three connected DELL computers. Two of them are the render nodes, and one acts as master controlling the navigation and audio-visual output. Technical specifications of the hardware of the Virtual Reality system are summarized in table 1. To reduce the intensity of reflections from the walls we attached acoustic absorber plates inside our virtual reality lab.

The software application for the visual stereoscopic and acoustic rendering of the simulation results on the hardware described above is implemented in C++ using QT[1] and OpenGL[2] APIs for GUI programming and rendering. Particularly, we used QT's client-server concept for controlling the graphical representation. The render nodes are are responsible for drawing the view for the left and right eye, respectively. To ensure that both rendered images are displayed simultaneously for a flicker-free representation of the scene, synchronization of the render nodes is required. Some APIs for parallel rendering already exist, for example Chromium[3], which is a system for interactive rendering on clusters.

To avoid restrictions imposed by existing synchronization software packages, a simple method based on sockets and full hand-shake synchronization is implemented. Nevertheless, since pure OpenGL is used for graphic rendering the application of Chromium or another VR API's (e.g. VR Juggler[4], DIVERSE[5]) in order to be flexible to use another VR hardware configuration is straightforward. The full hand-shake synchronization is realized as follows. The master sends command messages to the render nodes. When both renderer replied on a successful execution of the instruction, a new command can be send. In Fig. 4 the loading process of a room geometry from a file is conceptually shown. The master sends the command for loading the scene from file (*load*

[1] http://trolltech.com/
[2] http://www.opengl.org/
[3] http://chromium.sourceforge.net/
[4] http://www.vrjuggler.org
[5] http://diverse-vr.org/

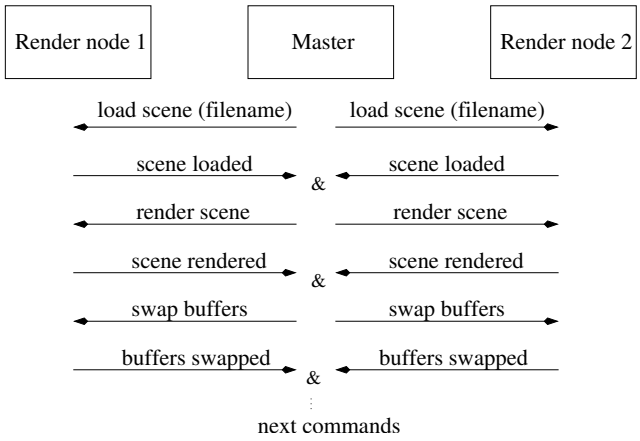


Figure 4 Full hand-shake synchronization for rendering

scene). After the affirmative reply from both render nodes, instruction for rendering follows (*render scene*). Is this successfully executed and confirmed, the buffers can be swapped now for the final representation of the left eye and right eye views on the display. Before the master can send new instructions to render nodes, it needs to wait for the response to the previous message, even in the case of single operations like swapping buffers. Otherwise, two consecutive messages sent by the master may collide causing a serious time delay.

For auralization purpose we used the synthesis method described in Section 5. In order to drive the loudspeaker of the audio system separately with the corresponding signals the PortAudio library[6] for acoustical output is utilized. The audio playback is administrated by the master computer and is started in a new process (*thread*).

4 Room Acoustic Simulation

4.1 Phonon Tracing Approach

Photon mapping [18, 19] is often used for rendering photo-realistic images, supplementing uni-directional raytracing by a variety of visual effects, like color bleeding and caustics. We adopt a similar approach to the simulation of sound, named phonon tracing [5], which is summarized in following.

Problem Specification

Our simulation algorithm requires the following input information:

- position of sound source s,
- emission distribution E of sound source,
- one ore more listener positions l_i,
- a triangulated scene with tagged material m_j,
- an absorption function $\alpha_j : \Omega \mapsto (0,1]$ for each material,
- an acoustic BRDF for each material (if applicable),
- an energy threshold ε for terminating the phonon paths.

[6] www.portaudio.com

The output of our approach is a FIR filter f_i for each listener's position l_i corresponding to the impulse response with respect to the sound source and the phonon-map containing for each phonon the energy spectrum e_p, the traversed distance d_p, the phonon's position p_p at the reflection point, its outgoing direction v_p, number of reflections r_p, and the material m_p at the current reflection.

Our simulation algorithm is executed in two stages, the *phonon tracing* step constructs the phonon map, and the *phonon collection and filtering* step collects the phonon's contribution to a FIR filter for every listener position.

Phonon Tracing

Every phonon p emitted from the sound source carries the following information:

- an energy spectrum $e_p : \Omega \mapsto \mathbb{R}^+$,
- the distance d_p traversed from the source,
- the phonon's current position p_p,
- the outgoing direction v_p.

Our absorption and energy functions α_j are represented by $n_e = 10$ coefficients associated with the frequencies $40, 80, 160, ..., 20480$ Hz. The basis function for the energy spectrum are wavelets adding up to a unit impulse. Every phonon is composed of different frequencies, which is more efficient than tracing a single phonon for each individual frequency band.

Phonons are emitted from the source s according to the emission probability distribution E and have at starting point a unit energy spectrum $e_{p,i} = 1$ $(i = 1, ..., n_e)$. At the intersection of the phonon ray with the scene, the phonon direction d_p is reflected with respect to the surface normal and the absorbed energy is subtracted according to the local material m_j, and the distance d_p is set to the traversed distance. The phonon is fixed at the intersection point, contributing to a global phonon map.

If the maximal energy of the phonon exceeds the energy threshold, i.e. $max\{e_{p,i}\}_{i=1}^{n_e} > \varepsilon$ and a minimal number of reflections is achieved, the next phonon re-uses the path and energy of the preceding one, saving computation time. It started at the current position with respect to the outgoing direction d_p and contributes to the phonon map at next surface intersection. If the threshold is not exceeded and a minimum number of reflections has been computed, the a new phonon is stated from the source. After a prescribed number n_p of phonons have contributed on the global phonon map, the tracing is terminated. The phonon map is used for further visualization purposes on our virtual reality system.

Phonon Collection and Filtering

The remaining task of the phonon tracing method is collecting the phonon's contribution to a FIR filter f for every listener's position l. This filter corresponds to the impulse response from the source, recorded at l, such that convolution with an anechoic signal, reproduces the perceived signal.

In the case of uniform absorption for all frequencies, the contribution of a phonon visible from the listener is simply a scaled, translated unit pulse (Dirac). The Dirac is shifted by the time elapsed between emission and reception of a phonon and scaled by the phonon's energy $e_{p,i}$ multiplied by a gaussian weighting the distance of the ray to the listener. In classical acoustic ray tracing [25, 26], a sphere is used to collect rays at listener position. Using a gaussian, however, provides much smoother filters, since more phonon rays contribute to the filter, weighted by their shortest distance.

In the more general case of frequency-dependent absorption, the unit impulse is subdivided into wavelets representing the individual frequency bands. The filter becomes the a sum of this wavelets scaled by $e_{p,i}$ and shifted by the elapsed time. In our implementation we use 10 frequency bands and absorption coefficients for the frequencies $\omega_i = 20 \cdot 2^i$ Hz $(i = 1, ..., 10)$. We construct band-pass filters in spectral domain by means of cosine functions in order to obtain quickly decaying wavelets.

The wavelets are computed by the inverse Fourier transform. Our band-pass filters have the following properties:

- compact support in the frequency domain Ω,
- symmetry and smoothness in both domains,
- their sum is one in Ω and a Dirac in the time domain.

The impulse response filters produced by our implementation are sampled at a rate of 48 kHz. For generating the filter bank, we used 2^{14} samples, employing the inverse FFT for computing the discretized wavelets. After the computation the response filter is normalized. A complete description of our phonon tracing algorithm, particularly addressing the filter design, can be read in more detail in [5].

4.2 Finite Element Method (FEM)

Phonon tracing or any other method based on geometric acoustics (ray tracing, mirror image) fail in the low frequency range for two reasons:

1. Wavelengths are of the order of typical dimensions of the room. Hence, diffraction can no longer be neglected.
2. Damping is typically low at low frequencies and reverberation times become too long to be represented by a convolution kernel of reasonable length.

Therefore, we have to fall back on wave acoustics to simulate the low frequency part of the sound field. For closed rooms this is preferably done by the finite element method (FEM), which approximates the wave equation by a large system of ordinary differential equations (ODEs) the unknowns of which are the pressures at grid points covering the room. In general, there are by far too many unknowns to solve these systems of ODEs in real time. Hence, we need to reduce the system to a concise state-space model with similar input-output behavior in the frequency range of interest.

There are many different approaches to model reduction [2]. The common observation is that system dynamics can often be represented quite well by a superposition of a few (generalized) eigen-modes. The coefficients of these modes are the unknowns of the new reduced system. Finally, assuming samplewise constant input (e.g. acceleration of the loudspeaker membrane), the continuous state-space model is transformed into a discrete one, which can be solved in real time. In practice, a low pass filter is used to split the input signal into a low frequency part and a remainder. The low frequency part (< 300 Hz) is handled in the way described here and the remainder by phonon tracing.

In the following we list the steps to get from the wave equation (1) to a reduced discrete state-space model (5) describing the transient response of a room to an excursions of a loudspeaker membrane. The wave equation and associated boundary conditions read:

$$\begin{aligned} \frac{\partial^2 p}{\partial t^2} - c_0^2 \Delta p &= 0 \quad \text{on} \quad G \\ c_0 \frac{\partial p}{\partial n} &= -\frac{1-R}{1+R} \frac{\partial p}{\partial t} \quad \text{on} \quad \Gamma_w \\ \frac{\partial p}{\partial n} &= -\rho_0 \frac{\partial^2 x_m}{\partial t^2} \quad \text{on} \quad \Gamma_m \,. \end{aligned} \tag{1}$$

$p = p(t,x)$ denotes pressure, c_0= 343 m/s the velocity of sound, and ρ_0 =1.2 kg/m the density of air at room temperature, G the interior of the room, and Γ_w and Γ_m the surfaces of walls and membrane, respectively. x_m is the excursion of the membrane and R is a reflection coefficient. It may depend on the particular wall, but is constant for all frequencies. This is a minor problem, as we use the model only for a small frequency band.

Approximating the pressure distribution by a superposition of, for instance, piecewise quadratic ansatzfunctions $p(t,x) = \sum_{i=0}^{N} p_i(t)\, \varphi_i(x)$ and integrating (1) with respect to the φ_i gives a FE model

of the form:

$$\begin{aligned} M\ddot{p} + D\dot{p} + Kp &= Fu \\ y &= Pp\,. \end{aligned} \tag{2}$$

The real $N \times N$ matrices M, D, K are called mass, damping, and stiffness matrix. $p = p(t)$ is a vector composed of the coefficients p_i. $u = u(t)$ is the input, e.g. the acceleration of the membrane. F transforms this input into a force. P is a projection matrix extracting certain interesting pressures y_i. Setting

$$\hat{x} = \begin{bmatrix} p \\ \dot{p} \end{bmatrix}, \quad \hat{E} = \begin{bmatrix} I & 0 \\ 0 & M \end{bmatrix}, \quad \hat{A} = \begin{bmatrix} 0 & I \\ -K & -D \end{bmatrix}, \quad \hat{B} = \begin{bmatrix} 0 \\ F \end{bmatrix}, \quad \textit{and} \quad \hat{C} = \begin{bmatrix} P & 0 \end{bmatrix}$$

the FE model may be rewritten as a state-space model:

$$\begin{aligned} \hat{E}\dot{\hat{x}} &= \hat{A}\hat{x} + \hat{B}u \\ y &= \hat{C}\hat{x}\,. \end{aligned} \tag{3}$$

Assuming that $\hat{x}$ is essentially composed of the columns of a matrix $U \in \mathbb{R}^{N \times n}$, $n \ll N$, and projecting the equations on the columns of $V \in \mathbb{R}^{N \times n}$ we end up with a reduced state-space system where

$$\tilde{x} = U\hat{x}, \quad \tilde{E} = V^t \hat{E} U\,, \quad \tilde{A} = V^t \hat{A} U\,, \quad \tilde{B} = V^t \hat{B}\,, \quad \tilde{C} = \hat{C}U\,. \tag{4}$$

The columns of U and V may be found by expanding the associated transfer function $H(s) = \hat{C}(s\hat{E} - \hat{A})^{-1}\hat{B}$ about some shifts $s_j = \mathrm{i}\omega_j$. Here, we used the *rational dual Arnoldi* algorithm described in [35]. Finally, dividing the first equation of the reduced version of (3) by $\tilde{E}$ and performing a balanced truncation [48] we end up with a state-space system of typically a few hundred unknowns rather than several 10,000 degrees of freedom. Integrating the reduced version of (3) over the length Δt of one sample for constant input u_n, i.e.

$$x_n = \mathrm{e}^{\tilde{A}\Delta t} x_{n-1} + \int_0^{\Delta t} \mathrm{e}^{\tilde{A}(\Delta t - \tau)}\,\mathrm{d}\tau\, \tilde{B} u_{n-1},$$

leads to a discrete state-space system

$$\begin{aligned} x_n &= Ax_{n-1} + Bu_{n-1} \\ y_n &= Cx_n\,. \end{aligned} \tag{5}$$

Switching to a suitable basis $x_n = T\xi_n$ it is always possible to turn the system into *companion canonical form* [23], where the first $n-1$ rows of A coincide with the last $n-1$ rows of the identity matrix of order n, the last row contains the negative coefficients of the characteristic polynomial of A, and B is the n-th unit vector. Hence, updating the state vector x_n and evaluating the pressure at a certain position requires $2n$ multiplications and $2n-1$ additions. Below, we will present an example where $n = 149$, i.e. 595 floating point operations are needed per new sample. On the other hand, if the sampling rate is 48,000 and we are using a convolution kernel instead which can represent at least a full wave of a 20 Hz signal, we need at least 4799 floating point operations.

Numerical Example

We consider a small room with two doors, a heating covered by absorbing material, walls made of bricks or concrete, an absorbing ceiling and a loudspeaker in one of the corners, cf. Fig.5. We assume a reflection coefficient $R = 0$ at the heating (total absorption), $R = 0.8$ at the ceiling, and $R = 1$ (total refection) at all the other surfaces. The original FE model contains 29272 unknowns and was created

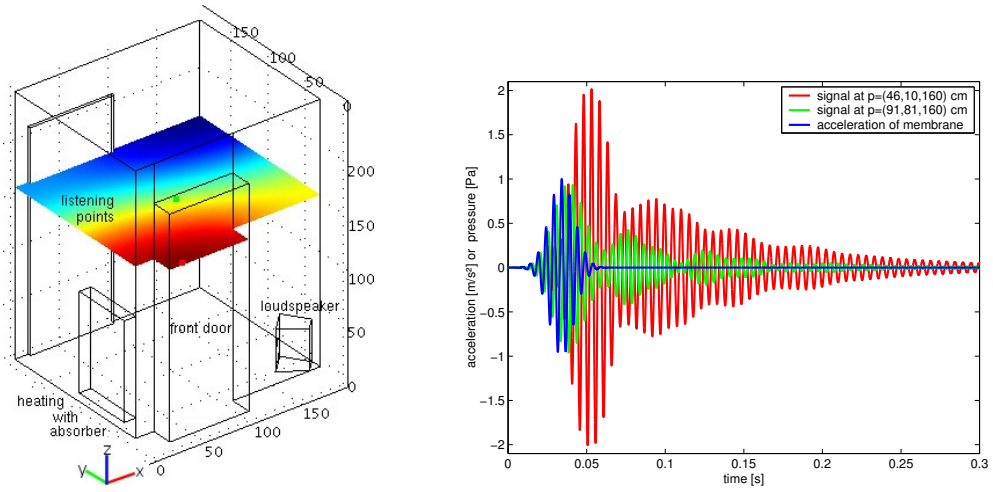


Figure 5 Response to a low frequency beep in a small weakly absorbing room .

with FEMLAB. Model reduction was done in MATLAB and yielded a model of order 145. Virtual measurements are taken in a plane 160 cm above the floor. The fourier transform of the input signal to the loudspeaker is a Gaussian centred about 200 Hz with standard deviation of 20 Hz. Responses are illustrated for a virtual microphone in the small indentation at the front door and in the middle of the room. Note that the indentation acts as a resonator. A standing wave is excited which oscillates between front door and opposite corner while taking low values in the center of the room. Close to the front door reflections accumulate in such a way that the maximum amplitude is reached when the loudspeaker has already stopped playing.

5 Acoustic Synthesis

Let $\tilde{p}_i \in \mathbb{R}^{2n_b}$ a section of the digital signal simulated in point $\tilde{x}_i \in \mathbb{R}^2$ of the horizontal measuring plane cutting the virtual room. n_b is the block length. We choose a power of 2 to speed up later Fourier transforms. The n_p points are distributed equally on a small circle of radius r about the virtual head position. The radius has to be of the same order as the wave length corresponding to the highest frequency to be reproduced:

$$r \approx \lambda_{\min} = \frac{c_0}{f_{\max}} \,. \tag{6}$$

For $f_{\max}$= 20,000 Hz and a sound velocity of c_0= 343 m/s the radius is only 12 mm. The exact radius depends on the number of loudspeakers n_l. To get an idea, imagine that an arbitrary sound field of frequency $f_{\max}$ has to be approximated by a polynomial with n_l coefficients which, for small n_l, will be valid only within a small radius. Choosing r this small we avoid artefacts in reproducing the high frequency part of the sound field while, at lower frequencies, the approximation will still be good also for larger radii.

The system for sound field synthesis is made to reproduce sound in n_p fixed positions on a ring of radius r about a point in the projection room. As the proband may change the line of vision in the virtual room these points cannot be directly associated with the n_p points from the fixed grid where sound is simulated. Hence, the simulated samples $\tilde{p}_{i,n}$ of time step n are interpolated and evaluated at the rotated real positions x_i using cubic splines in the angle.

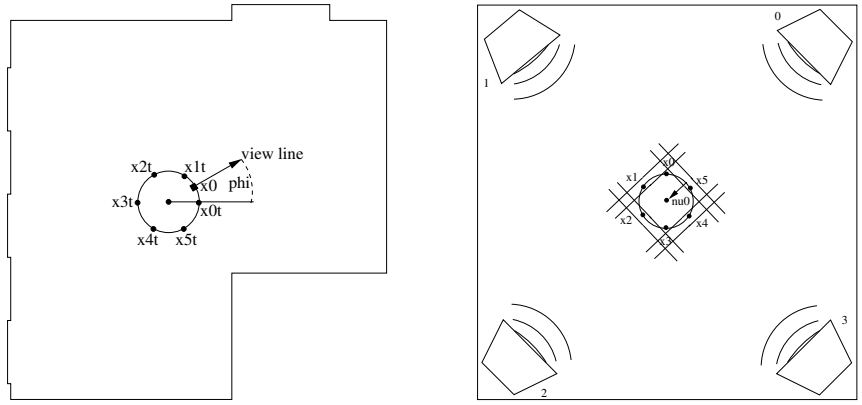


Figure 6 Virtual and projection room.

The resulting signals $p_i \in \mathbb{R}^{2n_\mathrm{b}}$ are first windowed by a lifted cosine and then subjected to a fast Fourier transform:

$$\hat{p}_i = \mathrm{fft}(p_i^w), \quad p_{i,n}^w = \tfrac{1}{2}\left[1 - \cos\frac{\pi(n+\frac{1}{2})}{n_\mathrm{b}}\right] p_{i,n} \,. \tag{7}$$

Synthesizing the sound field we assume that the loudspeakers induce plane waves in the central listening position and that time delay and amplification are the same for all loudspeakers. These conditions are met fairly well as we use an anechoic projection room and the speakers are furnished with FIR filters and placed on a circle. However, in a later version the plane-wave assumption will be replaced by using measured impulse responses. So far, the Fourier coefficients $\hat{q}_{j,k}$ of the signal of the j-th speaker are computed to satisfy:

$$\hat{p}_{i,k} = \sum_{j=1}^{n_\mathrm{l}} \hat{q}_{j,k} \exp\left(-\mathrm{i}\frac{\omega_k}{c_0}\langle v_j, x_i\rangle\right) \,. \tag{8}$$

$\omega_k = \frac{\pi f_\mathrm{s}}{n_\mathrm{b}} k$ is the k-th angular frequency, f_s the sampling rate, and v_j the radiation direction of the j-th speaker. This leads to a linear system for each frequency:

$$A_k \hat{q}_{.k} = \hat{p}_{.k} \,. \tag{9}$$

For stability reasons the number n_p of reproduction points is chosen greater than the number n_l of speakers and (9) becomes overdetermined. Moreover, in order to keep speaker signals bounded we introduce a small regularization parameter α:

$$\hat{q}_{.k} = T_k \hat{p}_{.k}, \quad T_k = [A_k^* A_k + \alpha \mathrm{I}]^{-1} A_k^* \,. \tag{10}$$

The matrices T_k are computed once for each frequency when initializing the system. As the speaker signals will be real we have $\hat{q}_{.,k} = \hat{q}^*_{.,n_\mathrm{b}-k}$ and only half of the systems need to be solved. Next, the $\hat{q}_j$ are retransformed into time domain by the inverse Fourier transform to get q_j^w. The superscript w indicates that these speaker signals will reproduce only the windowed measurements. To get the final speaker signals, the first half of q_j^w is added to the last half of the q_j^w computed for the last block:

$$q_j = q_{j,1\ldots n_\mathrm{b}}^{w,\mathrm{new}} + q_{j,n_\mathrm{b}+1\ldots 2n_\mathrm{b}}^{w,\mathrm{old}} \,. \tag{11}$$

Finally, $q^{w,\text{new}}_{j,n_b+1\ldots 2n_b}$ is stored for adding in the next step. The following table summarizes the algorithm:

1. *Get next block of simulated measurements.*
2. *Per time step: interpolate simulated measurements at reproduction points of projection room.*
3. *Scale measurements in projection room by lifted cosine-function.*
4. *Fast Fourier transform of scaled measurements in projection room.*
5. *Per frequency: compute loudspeaker signals to optimally reproduce measurements in a regularized least squares sense.*
6. *Inverse fast Fourier transform of loudspeaker signals.*
7. *Add first half of loudspeaker signals to last half stored for last block.*
8. *Store last half of loudspeaker signals to add in the next step.*
9. *Goto 1.*

6 Visualization

Our first visualization method described in [5] focuses on the spatial propagation of a sound wave from the source. The corresponding wave front traverses the room and is reflected on surfaces, altering its intensity and energy spectrum. We visualized this sound waves by rendering small spheres representing the sound particles. These are color coded by means of their spectral energy. Therefore, we use the RGB components, such that blue corresponds to the average of the energy by 40, 80, 160, 320 Hz, green corresponds to the average of energy by 640, 1280, 2560 Hz, and red to the average by 5120, 10240, 20480 Hz. When sliding through time, the spheres follow the simulated phonon paths. We integrated following functions in our interactive visualization system:

- varying the percentage of phonons to be rendered,
- rendering only the phonons reflected from a selected material,
- exchanging selected materials,
- varying time / transversed distance.

The data structure supporting this visualization is an array of phonons carrying their energy spectrum e_p, the traversed distance d_p, the phonon's position p_p at a reflection point, and its outgoing direction v_p, according to section section 4.1. In addition, we record the number of reflections r_p and the material m_p at the current reflection. Since all phonons sharing the same path are listed consecutively in the array, it is simple, for example, to select all consecutive pairs p_i, p_{i+1} where the current time t satisfies $t_{pi} \leq t < t_{pi+1}$ and to draw a sphere on the line segment p_{pi}, p_{pi+1}, corresponding to a phonon's location at time t.

The exchange of a certain material requires only the phonons' energy to be re-evaluated, where the phonon paths remain fixed. To allow the exchange, it is necessary to enforce a minimum number of reflections for every path in advance, since otherwise materials with high absorption coefficients cannot be replaced. An application scenario of this kind is provided in the next section.

To improve the visualization of sound propagation to be able to better distinguish between different reflected wave fronts and still preserving the scalability we use surface elements (surfels) [38] for particle representation. In Fig. 7 a sketch of surfels is depicted. The surfel representation can be adjusted according to users needs on accuracy as well as graphics hardware capabilities. The level of detail can be changed by prescribing the number of points describing the circumcircle of the disk (from quad to circle in Fig. 7).

In order to depict the propagated sound wave front, a surfel is rendered on each particle position at given time. The size of the surfel is given by the radius r_i and the orientation by the normal n_i

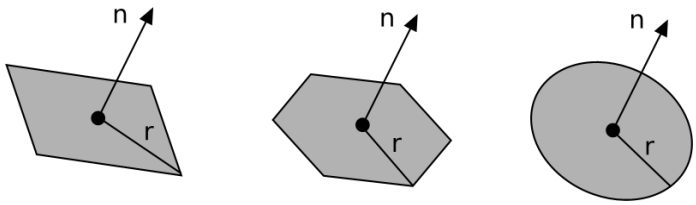


Figure 7 Sketch of surface elements (surfels) at different levels of detail.

which corresponds to the sound particle traveling direction. The radius r_i varies depending on the number n_{ph} of phonons to be rendered (given by the user) as well as on the traversed distance l of the sound wave (time step given by the user):

$$r_i = s\sqrt{\frac{l}{n_{ph}}}\,. \tag{12}$$

The radius r_i can also be expanded by the scaling factor s in order to obtain a surface rendering of the sound wave front. A small radius r_i results in a visualization on the wave front similar to that using spheres, whereas a large r_i leads to a surface-like representation. An example of sound wave propagation visualization is depicted in Fig. 8.

7 Visualization Application

The visualization approaches described in the previous section were used to examine the acoustic properties of a virtual reality (VR) laboratory that is also used for auralization purposes. Fig. 10 shows the wave front propagation based on 30000 paths traced from the sound source resulting in 910610 phonons in the phonon map, at a traversed distance of 1.5, 4.5, and 10 m. At small distances / traversal times, the individual wave fronts can be recognized, whereas large distances provide insight into the frequency decomposition of the various reflections. We observe a shift towards lower frequencies, since the particles color is dominated by the blue component after a number of reflections.

In order to identify the reason for the frequency shift, we look at sound particles reflected from a selected material in their earlier paths. Fig. 9 shows these particles for reflections from walls, floor, and canvas, respectively. We observe that particles reflected from walls carry mostly a yellowish color, despite of their potential reflection from additional materials. Hence, the energy of these particles is shifted towards the mid and high frequencies. The floor and the canvas reduce high frequencies. Reflections from the canvas affect mostly the right side of the room, whereas the impact of the carpet is much greater. While a potential frequency shift can already be seen in material absorption coefficients, their impact on room acoustics can be studied much better with the aid of our visualization approach.

In Fig. 11 the carpet is replaced by a material with similar absorption as the walls. When comparing Fig. 11(a) and 11(b), it becomes evident that this change is sufficient for increasing the intensities of mid and high frequencies. This interactive visual observation of intensities changes because of material modifications adds significantly to the acoustic properties improvement during the design process. While the acoustics of the laboratory are not much of importance, it may have a greater impact on the design of larger classrooms. Optimizing the acoustics of such larger rooms may, for example, eliminate the need of using a microphone or improve the auditive quality of concerts.

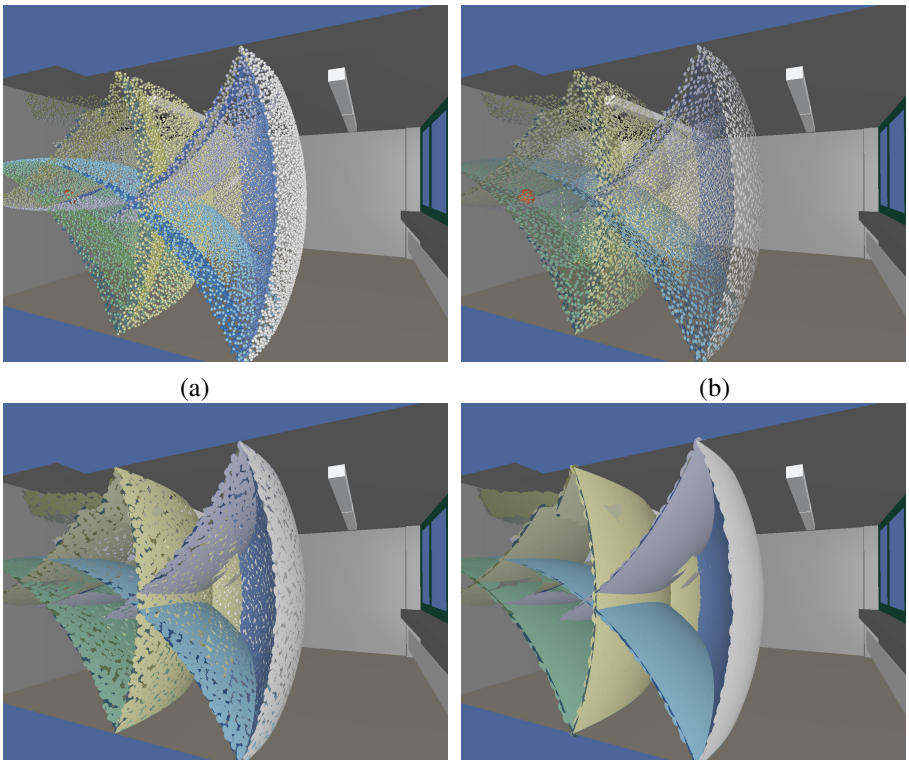


Figure 8 Wave front representation using surface elements (surfels) in contrast to spheres. Spheres used for rendering of phonons (a). Surfels with increased radius from (b) to (d).

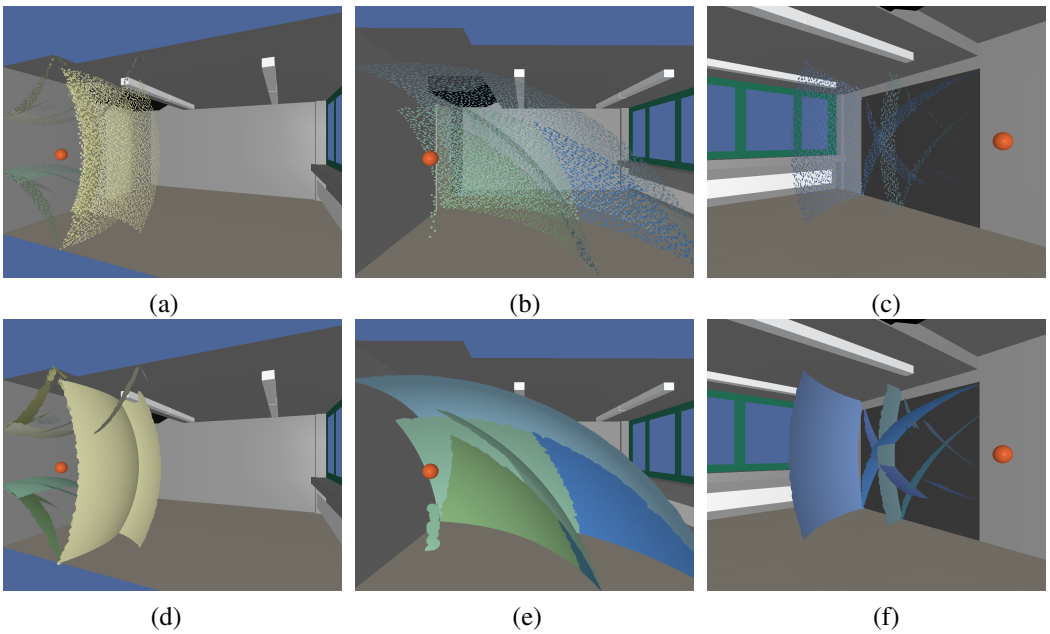
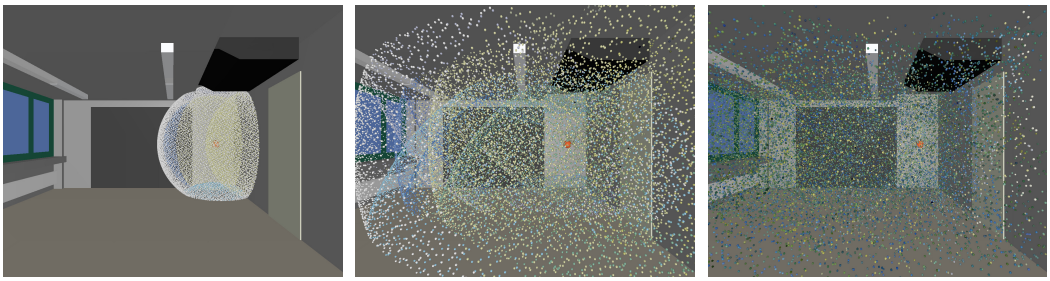
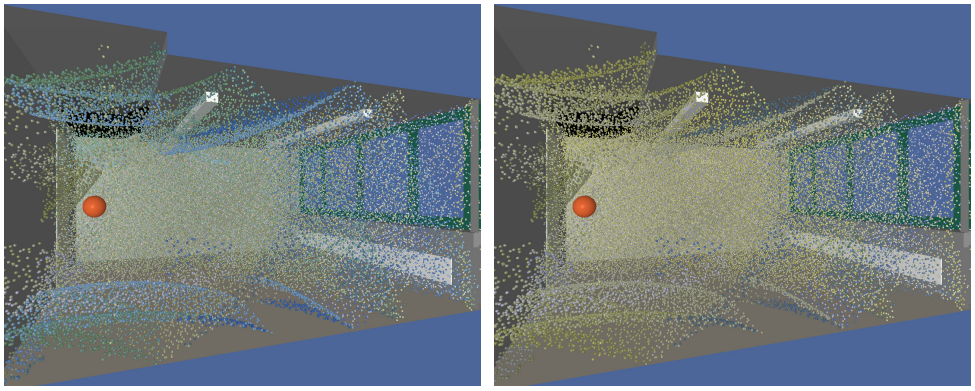


Figure 9 Wave fronts reflected from walls (a+d), floor (b+e), and projection wall (c+f) using smaller (a,b,c) and greater (d,e,f) radius.



(a) (b) (c)

Figure 10 Wave front propagation from a spherical sound source. The visualization shows phonons with RGB color-coded spectral energy due to reflection at different materials at 1.5m (a), at 4.5m (b), and at 10m (c).



(a) (b)

Figure 11 Changing material of the floor. (a) all phonons at d = 4.5m; (b) same as (a) with new material.

8 Conclusions

In this work we have presented an audio visual Virtual Reality system for room acoustics. For acoustic simulation purposes we have combined a geometric approach, the phonon tracing, and a FEM based solver. Sound rendering is performed utilizing a professional sound equipment using our synthesis approach. Furthermore we have introduced visualization techniques for sound wave propagation from the sound source. For visual rendering a stereoscopic 3D back projection display is used. The system enables the exploration of acoustical behavior inside a room visual as well as aural allowing the feasibility to improve the acoustics of the room during the design process. Future work includes continuative methods visualizing well defined room acoustic metrics as well as the visualization of the low frequency part simulated by means of FEM. Further work on integration of human machine interfaces like 3D mouse, game port or haptic devices would improve the interaction in the virtual environment.

Acknowledgments

This work was supported by the Stiftung Rheinland Pfalz für Innovation under contract no. 15202-386261/644.

References

1 J.B. Allen and A. Berkeley. Image method for efficiently simulating small-room acoustics. *J. Acoust. So. Amer.*, 65(4):943–950, Apr. 1979.

2 A.C. Antoulas and D.C. Sorensen. Approximation of large-scale dynamical systems: An overview. Technical report, Rice University, 2001.

3 I.and T. Lentz Assenmacher, T. Kuhlen, and M. Vorländer. Integrating real-time binaural acoustics into VR applications. In *Eurographics Symposium on Virtual Environments*, pages 129–136, Grenoble, June 8–9 2004.

4 A.J. Berkhout, D. de Vries, and P. Vogel. Acoustic control by wave field synthesis. *J. Acoust. Soc. Am.*, 93(5):2764–2778, 1993.

5 M. Bertram, E. Deines, J. Mohring, J. Jegorovs, and H. Hagen. Phonon tracing for auralization and visualization of sound. In *IEEE Visualization*, pages 20–30, Minneapolis, MN, October 2005.

6 J. Borish. Extension of the image model to arbitrary polyhedra. *J. Acoust. So. Amer.*, 75(6):1827–1836, 1984.

7 P. Bourke. 3d stereo rendering using opengl (and glut). `http://astronomy.swin.edu.au/~pbourke/opengl/stereogl/`, 1999.

8 L. Cremer and H.A. Müller. *Die wissenschaftlichen Grundlagen der Raumakustik, Band I*. S. Hirzel Verlag Stuttgart, 1978. 2. völlig neubearbeitete Auflage.

9 E. Deines, M. Bertram, J. Mohring, J. Jegorovs, F. Michel, H. Hagen, and G.M. Nielson. Comparative visualization for wave-based and geometric acoustics. In *IEEE Visualization 2006*, pages 1173–1179, Baltimore, Maryland, USA, October 29 – November 3 2006.

10 E. Deines, F. Michel, M. Bertram, H. Hagen, and G. Nielson. Visualizing the phonon map. In *Eurographics / IEEE-VGTC Symposium on Visualization*, Lisbon, Portugal, Mai 8-10 2006.

11 Y. Fukushima, H. Suzuki, and A. Omoto. Visualization of reflected sound in enclosed space by sound intensity measurement. *Acoust. Sci. & Tech.*, 27(3):187–189, 2006.

12 Thomas A. Funkhouser, Ingrid Carlbom, Gary Elko, Gopal Pingali Mohan Sondhi, and Jim West. A beam tracing approach to acoustic modeling for interactive virtual environments. In *Computer Graphics (SIGGRAPH 98)*, pages 21–32, Orlando, FL, July 1998.

13 Thomas A. Funkhouser, Patrick Min, and Ingrid Carlbom. Real-time acoustic modeling for distributed virtual environments. In *Computer Graphics (SIGGRAPH 99)*, pages 365–374, Los Angeles, August 1999.

14 W. Hackbusch. *Integralgleichungen, Theorie und Numerik*. Teubner Verlag, 1989.

15 L. Harrison, D. McAllister, and M. Dulberg. Stereo computer graphics for virtual reality. In *ACM SIGGRAPH '97*, pages 20–30, Minneapolis, MN, October 1997. Course Notes 6.

16 Larry F. Hodges. Tutorial: Time-multiplexes stereoscopic computer graphics. In *IEEE Computer Graphics & Applications*, pages 20–30, Orlando, FL, March 1992.

17 Marcin Jedrzejewski and Krzysztof Marasek. Computation of room acoustics using programable video hardware. In *International Conference on Computer Vision and Graphics ICCVG'2004*, Warsaw, Poland, September 22-24 2004.

18 Henrik Wann Jensen. Global illumination using photon maps. In *Rendering Techniques '96 (Proceedings of the 7th Eurographics Workshop on Rendering)*, pages 21–30, 1996.

19 Henrik Wann Jensen and Per H. Christensen. Efficient simulation of light transport in scene with participating media using photon maps. In *Computer Graphics (SIGGRAPH 98)*, pages 311–320, July 1998.

20 B. Kapralos, M. Jenkin, and E. Millios. Sonel mapping: Acoustic modeling utilizing an acoustic version of photon mapping. In *IEEE International Workshop on Haptics Audio Visual Environments and their Applications (HAVE 2004)*, Ottawa, Canada, October 2-3 2004.

21 S. Khoury, A. Freed, and D. Wessel. Volumetric visualization of acoustic fields in cnmat's sound spatialization theatre. In *Visualization '98*, pages 439–442 & 562. IEEE, 1998.

22 B. Klehs and T. Sporer. Wave field synthesis in the real world part 1: In the living room. In *114th AES Convention*, Amsterdam, March 22 - 25 2003.

23 H.W. Knobloch and H. Kwakernaak. *Lineare Kontrolltheorie*. Springer-Verlag, 1985.

24 N. Korany, J. Blauer, and O. Abdel Alim. Acoustic simulation of rooms with boundaries of partially specular reflectivity. *Applied Acoustics*, 62:875–887, 2001.

25 U. Krockstadt. Calculating the acoustical room response by the use of a ray tracing technique. *Journal of Sound and Vibrations*, 8(18):118–125, 1968.

26 U. Kulowski. Algorithmic representation of the ray tracing technique. *Applied Acoustics*, 18:449–469, 1984.

27 C. Lauterbach, A. Chandak, and D. Manocha. Interactive sound propagation in dynamic scenes using frustum tracing. In *In Proceeding of IEEE Visualization*, Sacramento, CA, USA, October 28 – November 1 2007.

28 T. Lokki and V. Nenonen. Imersive visualization of room acoustics. In *Joint Baltic-Nordic Acoustics Meeting*, Gothenburg, Sweden, November 8–10 2006.

29 Tapio Lokki. *Physically-based Auralization*. PhD thesis, Helsinki University of Technology, 2002.

30 Tapio Lokki, Peter Svensson, and Lauri Savioja. An efficient auralization of edge diffraction. In *Audio Engineering Society, 21th International Conference*, pages 166–172, St. Petersburg, Russia, June 2002.

31 J. Merimaa, T. Lokki, T. Peltonen, and M. Karjalainen. Measurements, analysis, and visualization of directional room responses. In *Proceedings of the 111th Audio Engineering Society (AES) Convention*, New York, NY, USA, September 21–24 2001.

32 M. Monks, B.M. Oh, and J. Dorsey. Acoustic simulation and visualization using a new unified beam tracing and image source approach. In *Convention of the Audio Engineering Society*. ACM, 1996.

33 M. Monks, B.M. Oh, and J. Dorsey. Audioptimization: Goal-based acoustic design. *IEEE Computer Graphics and Applications*, 20(3):76–91, 2000.

34 M. Naef, O. Staadt, and M. Gross. Spatialized audio rendering for immersive virtual environments. In *Symposium on Virtual Reality Software and Technology*, pages 65–72. ACM, November 2002.

35 K.H.A. Olsson. Model order reduction in FEMLAB by dual rational Arnoldi. Master's thesis, Department of Mathematics, Chalmers University of Technology and Göteborg University, 2002.

36 A. Omoto and H. Uchida. Evaluation method of artificial acoustical environment: Visualization of sound intensity. *Journal of Physiological Anthropology and Applied Human Science*, 23:249–253, 2004.

37 S. Petrausch and R. Rabenstein. Highly efficient simulation and visualization of acoustic wave fields with the functional transformation method. In *Simulation and Visualization*, pages 279–290, Otto von Guericke Universität, Magdeburg, March 2005.

38 H. Pfister, M. Zwicker, J. van Baar, and M. Gross. Surfels: Surface elements as rendering primitives. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pages 335–342, New Orleans, LA, USA, 2000.

39 V. Pulkki and T. Lokki. Visualization of edge diffraction. *Acoustics Research Letters Online (ARLO)*, 4(4):118–123, 2003.

40 Rudolf Rabenstein, Oliver Schips, and Alexander Stenger. Acoustic rendering of buildings. In *5th Int. IBPSA Conference Building Simulation*, volume 2, pages 181–188, Prag, Czech Republic, September 1997.

41 Lauri Savioja, Tapio Lokki, and Jyri Huopaniemi. Auralization applying the parametric room acoustic modeling technique - the diva auralization system. In *International Conference on Computer Auditory Display*, pages 219–224, Kyoto, Japan, July 2-5 2002.

42 A. Sontacchi and R. Hoeldrich. Enhanced 3D sound field synthesis and reproduction system by compensating interfering refelexions. In *COST G-6 Conference on Digital Audio Effects*, Verona, December 7–9 2000.

43 A. Stettner and D.P. Greenberg. Computer graphics visualization for acoustic simulation. In *International Conference on Computer Graphics and Interactive Techniques*, pages 195–206. ACM, 1989.

44 Y. Tokita and Y. Yamasaki. Visualization of 3-dimensional sound fields by numerical solutions of particle displacement. *Acoust. Sci. & Tech.*, 26(2):215–217, 2005.

45 R.R Torres, U.P. Svensson, and M. Kleiner. Edge diffraction in room acoustics computations. In *EAA Symposium on Architectural Acoustics*, Madrid, Spain, Oct. 16-20 2000.

46 M. Vorländer. Simulation of the transient and steady-state sound propagation in rooms using a new combined ray-tracing/image-source algorithm. *J. Acoust. So. Amer.*, 86(1):172–178, 1989.

47 T. Yokota, S. Sakamoto, and H. Tachibana. Visualization of sound propagation and scattering in rooms. *Acoust. Sci. & Tech.*, 23(1):40–46, 2002.

48 K. Zhou and J.C. Doyle. *Essentials of robust control*. Prentice-Hall inc., 1998.