

# Phonon Tracing for Auralization and Visualization of Sound

Martin Bertram\*  
TU Kaiserslautern

Eduard Deines  
TU Kaiserslautern

Jan Mohring  
ITWM Kaiserslautern

Jevgenij Jegorovs  
ITWM Kaiserslautern

Hans Hagen  
TU Kaiserslautern

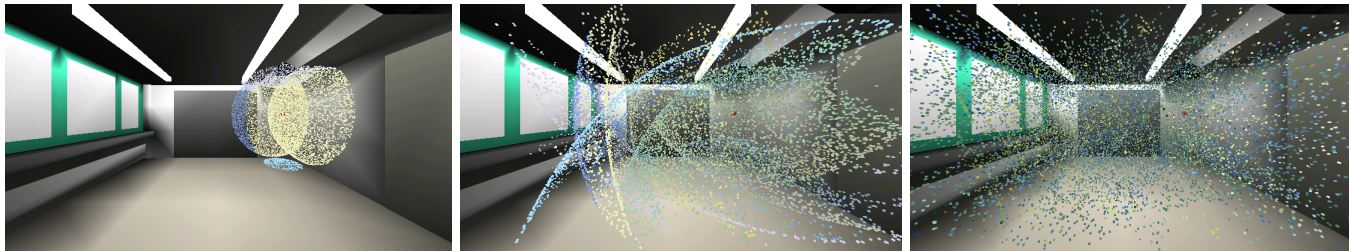


Figure 1: Wave-front propagation from a spherical sound source. The visualization shows phonons with color-coded spectral energy due to reflection at different materials.

## ABSTRACT

We present a new particle tracing approach for the simulation of mid- and high-frequency sound. Inspired by the photorealism obtained by methods like photon mapping, we develop a similar method for the physical simulation of sound within rooms. For given source and listener positions, our method computes a finite-response filter accounting for the different reflections at various surfaces with frequency-dependent absorption coefficients. Convoluting this filter with an anechoic input signal reproduces a realistic aural impression of the simulated room. We do not consider diffraction effects due to low frequencies, since these can be better computed by finite elements. Our method allows the visualization of a wavefront propagation using color-coded blobs traversing the paths of individual phonons.

**CR Categories:** I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Raytracing; I.6.8 [Simulation and Modeling]: Types of Simulation—Monte Carlo; J.2 [Computer Applications]: Physical Sciences and Engineering—Physics;

**Keywords:** acoustics, auralization, raytracing, photon mapping

## 1 INTRODUCTION

Auralization is concerned with simulating and reproducing the acoustic properties of a scene. It can be considered as an *auditive variant of visualization*, providing insight into the “good” and “bad” acoustic properties of classrooms, theaters, congress rooms, churches, etc. In this work, we contribute a simple, yet powerful auralization method, providing a combination of acoustic simulation and visualization. Sound is hereby represented as particles called *phonons*, emanating from a sound source, being reflected at different surfaces and finally collected at a listener’s position. Figure 1 contains a visualization of this process showing the different frequency decompositions due to reflection at materials with versatile absorption properties. The spectral energy of phonons within low, mid and high-frequency bands is mapped onto the blue, green and red color components, respectively.

\*e-mail: bertram@informatik.uni-kl.de

The objective of our auralization is the optimization of acoustic properties based on virtual models, for example of class rooms. For a given sound source (or a set of sources) and multiple listener positions, our simulation computes a finite impulse response (FIR) filter for each listener position, such that convolution with an anechoic input signal will provide a realistic impression of the sound perceived at this position. The progress of the simulation can be visualized by sliding through time and rendering all phonons corresponding to a unit impulse (Dirac) sent out from the source at time zero. Due to distinct absorption of different frequency bands at the surfaces of a scene, the spectral energy distribution of the phonons is altered at each reflection.

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. After a large time period, a large number of reflections has occurred, such that individual wave fronts cannot be identified, anymore. The color mix of the phonons, however, shows the overall acoustic quality. The darker the color of phonons gets, the more energy is absorbed by the different materials.

Applications of our work include (but are not limited to) the acoustic improvement during architecture, design, and equipment of class- and congress rooms. Compared to commercial products designed for home entertainment, our method is more precise for highly detailed scenes, since a much greater number of reflections can be considered for complex scenes. The limitations are low-frequency simulations where diffraction dominates the process. These are frequencies in the order  $f = c/\lambda \sim c/l$ , where  $c$  is the speed of sound and  $l$  denotes the diameter of the simulated room. Since only linear propagation of sound is considered by our method, diffraction needs to be simulated by different methods, like finite elements. For acoustic qualities, such as clarity of sound, however, lowest frequencies can often be neglected.

The remainder of our paper is structured as follows: In section 2 we review related work and summarize the differences with respect to our approach. Section 3 describes our simulation, composed of phonon emission, phonon collection, filtering, and visualization. Numerical examples and a qualitative evaluation are provided in section 4, followed by a brief conclusion.

## 2 RELATED WORK

In the theory of acoustics there are two main approaches simulating the propagation of sound. The first approach is based on wave equations that are numerically solved, for example with use of finite element methods (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 and the raytracing method.

The *image-source method* [1, 2] models specular reflections by inserting additional sound sources obtained by mirroring the location of audio source over polygonal surfaces inside the scene. The key idea here is, that a sound ray from a sound source  $S$  reflected on a surface has the same acoustic effect as one from the mirrored source  $S'$ , reduced by the absorbed energy part. In a rectangular box-shaped room it is easy to construct image sources up to a certain order of reflections. Following the image source calculations, for a receiver position  $P$ , each of the virtual sources must be considered to determine whether it can be heard in  $P$ . All sources visible from  $P$  are summed up, reproducing the acoustical impression at  $P$ , where in addition to energy the individual time delays are considered. The advantage of the image-source method is, that it is very accurate, but it becomes very complicated for non-box-shaped rooms and curved surfaces cannot be simulated, at all. The number of the mirrored sources increases exponentially with reflection order. Hence, this approach is suitable only for simple room geometry and low reflection orders.

In the *raytracing method* [12, 13] several rays are traced from the sound source to receiver that is typically represented as a sphere. The reflections over the surfaces in the scene are simulated according to specular laws or Lambert's law of diffuse reflections. In addition to absorption properties of considered surfaces, timely delay and air absorption are taken in account. This approach is general and easy to implement, but it is very computation expensive and can cause aliasing artifacts due to discrete number of rays. When changing the receiver's position, the tracing of all rays needs to be performed, again.

Due to the shortcomings of the two classical approaches described above, 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 method and raytracing is introduced in [21]. Here the visibility check of the image-source algorithm are performed via raytracing. Beam-tracing methods [4, 5, 16] overcome the aliasing problem of classical raytracing 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 [20, 15]. To overcome the dependency of the simulation on the receiver position the radiosity method was extended to be used in room acoustics [17, 11]. Due to the computation complexity these methods do not seem practical for large environments. Newer approaches try to cope with complexity by exploiting GPU hardware accelerating the simulation calculations [6].

The aim of audio simulation is the estimation of acoustic properties, such as reverberation time and the reproduction of acoustic benchmarks. Auralization is the process of producing audible impression of a room. Therefore an anechoic signal is convolved with

the room impulse signal and displayed over a loudspeaker system or head phones, often using head related transfer function (HRTF) for directional tracking. A survey of existing auralization systems and explicit description of the DIVA auralization system is given in [18, 14]. In the visualization context, only few approaches to acoustics have been described [19, 10].

Our method is inspired by a technique known as *photon mapping* [7, 8, 9] used for rendering photo-realistic images. The idea of this method is to trace photons from a light source into a scene, where their path is reflected and refracted at specular / transparent surfaces and ends on surfaces (or in volumes) with a diffuse component. These photons, composed of color intensity and outgoing direction, are associated with the surfaces ending their paths. When rendering these surfaces, for example by raytracing, a photon map is evaluated by inserting the  $n$  closest photons into a bi-directional reflection distribution function (BRDF) determining the local light intensity.

In the present work, we contribute a novel simulation technique for auralization, also applicable to large scenes. Our method is coupled with an interactive visualization, showing the impact of certain materials on the frequency decomposition and intensity of reflected wave fronts. With few restrictions, it is possible to alter acoustic material properties at interactive rates, re-using the particle paths computed by the simulation.

## 3 ALGORITHM

Photon mapping [7, 8] 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, where a couple of issues need to be re-considered. For example, the human ear is not capable of locating the origin of a sound source as precise, as the human eye is capable of locating the origin of light. In the case of high frequencies where diffraction is negligible, one can locate directional sound with a precision of about 30 degrees. Thus, a lower number of particles (phonons) may be sufficient for auralization.

Another difference is that the human ear can distinguish a lot more frequency bands than the eye. Since the eye has only three different receptors for red, green, and blue light, any continuous distribution within the visible spectrum between 400 and 700  $\mu\text{m}$  of wavelength is mapped onto three variables. Consequently, only three wavelengths need to be considered for rendering images. The human ear, however, can distinguish frequencies in the spectrum  $\Omega = [20\text{Hz}, 20000\text{Hz}]$  fairly precise. In addition, the spectrum of an audio signal is highly time-dependent in contrast to most illumination problems that require only a static solution.

**Problem specification.** Our algorithm requires the following input:

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

Examples for absorption functions provided by CARA<sup>1</sup> are illustrated in figure 2. The output of our method is a FIR filter  $f_i$  for each listener's position  $\mathbf{l}_i$  corresponding to the pulse response with respect to the sound source and a visualization of the simulation process.

<sup>1</sup>www.cara.de

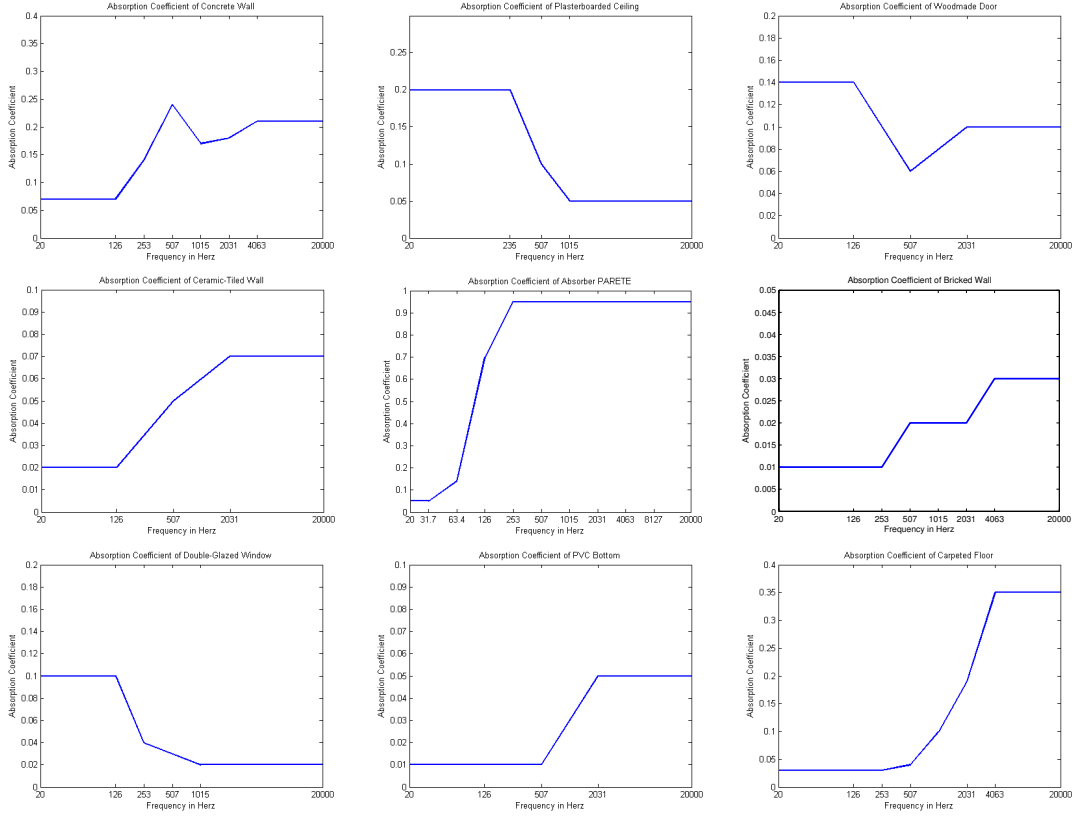


Figure 2: Absorption functions  $\alpha_i$  for different materials, provided by CARA.

### 3.1 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  $\mathbf{p}_p$
- its outgoing direction  $\mathbf{v}_p$ .

Additional information, such as the number of reflections and the material index of the last reflection may also be attached for visualization purposes.

Our absorption and energy functions  $\alpha_j$  are represented by  $n_e = 10$  coefficients associated with the frequencies 40, 80, 160, ..., 20480 Hz. The corresponding basis functions for the energy spectrum are wavelets  $\psi_i$  ( $i = 1, \dots, n_e$ ) adding up to a unit impulse (discretized Dirac). We will define these wavelets later. For now, we assume that every phonon is composed of different frequencies, which is much more efficient than tracing a single phonon for each individual frequency band.

Phonons are emitted from the source  $\mathbf{s}$  according to the emission probability distribution  $E$ . In the case of a sound source with equally distributed emission over the unit sphere, we use three independent random variables  $x, y, z$  with normal (gaussian) distribution using the probability function

$$P(x) = \frac{2}{\sqrt{\pi}} \exp(-x^2). \quad (1)$$

This is obtained by inserting a random variable  $\xi$  equally distributed on  $[0, 1]$  into  $x = \Phi^{-1}(\xi)$ , where  $\Phi$  is the normal distribution function

$$\Phi(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2) dt = \text{erf}(x). \quad (2)$$

Since the variable  $(x, y, z)^T$  forms a rotationally symmetric gaussian distribution, its normalized direction is equally distributed on the unit sphere [3] and can be used to emit phonons at a spherical sound source.

Our phonons are started at the source with a (discretized) unit energy spectrum  $e_{p,i} = 1$  ( $i = 1, \dots, n_e$ ). At the first intersection of the phonon ray with the scene, the phonon direction  $d_p$  is reflected with respect to the local surface normal and the absorbed energy is subtracted according to the local material  $m_j$ ,

$$e_{p,i} := (1 - \alpha_{j,i}) e_{p,i}, \quad (3)$$

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 (whose evaluation is discussed later). For efficient computation of ray-surface intersection, we use an octree-based partitioning of the scene.

If the maximal energy of the phonon exceeds the energy threshold, i.e.  $\max\{e_{p,i}\}_{i=1}^{n_e} > \epsilon_e$ , the next phonon re-uses the path and energy of the preceding one, saving computation time. It is started at the current position with respect to the outgoing direction  $d_p$  and contributes to the phonon map at the next surface intersection. If the threshold  $\epsilon_e$  is not exceeded and a minimum number of reflections have been computed, then a new phonon is started from the sound source  $\mathbf{s}$ . After a prescribed number  $n_p$  of phonons have contributed to the global phonon map, the tracing is terminated.

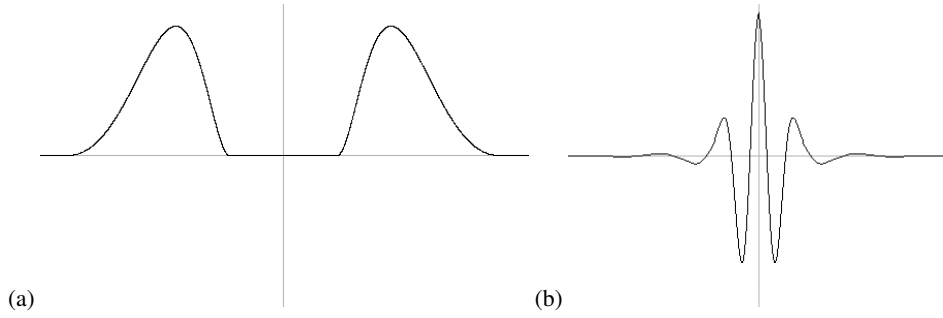


Figure 3: Bandpass filter, (a) in spectral domain; (b) in time domain.

The approach can be extended by considering bi-directional reflection distribution functions (BRDF's), for example when modeling corrugated materials by plain surfaces. A BRDF provides the reflection factor for a pair of incoming and outgoing directions. For any fixed incoming direction, it can be used as probability function for a random variable determining the reflection vector. We note that the material's corrugation must be very coarse to justify this approach, since the wave lengths of sound are much wider than those of light. Due to the strong dependency of the wavelengths and reflection distribution, individual BRDF's may be defined for several bandwidths, making it necessary to trace multiple succeeding phonons in different directions, carrying distinct parts of the energy spectrum.

### 3.2 Phonon Collection and Filtering

The remaining task of our method is collecting the phonons' contribution to a FIR filter  $f$  for every listener's position  $\mathbf{l}$ . This filter  $f$  corresponds to the pulse response from the source, recorded at  $\mathbf{l}$ , such that convolution with an anechoic source signal  $g_s$  reproduces the perceived signal  $g_l$ , where

$$g_l = g_s * f = \int_{\mathbb{R}} g_s(\tau) f(\cdot - \tau) d\tau. \quad (4)$$

In the case of uniform absorption for all frequencies, the contribution of a phonon visible from the listener is simply a scaled, translated Dirac (unit impulse),

$$\delta_{p,l,i}(t) = \exp(-k h_{pl}^2) e_{p,i} \delta(t - (d_p + d_{pl})/c), \quad (5)$$

where  $k$  is a constant,  $h_{pl}$  denotes the shortest distance of the continued phonon path  $\mathbf{p}_p + t\mathbf{v}_p$  and the listener  $\mathbf{l}$ ,  $e_{p,i}$  denotes the phonon's energy (at frequency  $\omega_i$ ),  $d_{pl}$  denotes the distance between phonon  $\mathbf{p}_p$  and listener  $\mathbf{l}$ , and  $c$  is the speed of sound. The term  $(d_p + d_{pl})/c$  corresponds to the time elapsed between emission and reception of a phonon. The Dirac  $\delta$  is shifted by this time interval 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 raytracing [12, 13], a sphere is used to collect rays at a listener's position. Using a gaussian, however, provides much smoother filters, since more phonon rays contribute to the filter, weighted by their shortest distance. The constant  $k$  determines the radius of this gaussian and can be compared with the shininess in the specular term of the Phong illumination model. For now, we use  $k = 200m^{-2}$ , keeping in mind that this constant needs to be adapted to the phonon density.

In the more general case of frequency-dependent absorption, the Dirac is subdivided into wavelets  $\psi_i$  representing the individual frequency bands,

$$\delta(t) = \sum_{i=1}^{n_e} \psi_i(t). \quad (6)$$

With the different energy coefficients  $e_{p,i}$ , the filter becomes

$$f(t) = \sum_p \exp(-k h_{pl}^2) \sum_{i=1}^{n_e} e_{p,i} \psi_i(t - (d_p + d_{pl})/c) \quad (7)$$

For a large number of phonons, it is more efficient to pre-compute  $n_e$  different filters, one for each bandwidth, based on equation (5). Their convolutions with the individual wavelets are then efficiently computed employing the Fast Fourier Transform (FFT). By the convolution theorem, equation (7) is equivalent to

$$f(t) = \sum_{i=1}^{n_e} \psi_i * \sum_p \delta_{p,l,i}(t), \quad (8)$$

where the convolution translates and scales the wavelets properly for each individual phonon.

**Filter design.** Our band-pass filters  $b_i = \hat{\psi}_i$  are constructed as follows. To obtain quickly decaying wavelets, the filters need to be smooth in the spectral domain. Therefore, we use the shape of cosine functions,

$$c_i(x) := \begin{cases} \frac{1}{2} + \frac{1}{2} \cos(\pi(x-i)) & \text{if } x \in [i-1, i+1]; \\ 0 & \text{else.} \end{cases} \quad (9)$$

With the re-definition  $c_1(x) := 1$  for  $x < 1$  and  $c_{n_e}(x) := 1$  for  $x > n_e$ , we obtain the property  $\sum_{i=1}^{n_e} c_i(x) = 1$ , everywhere.

In our implementation, we use  $n_e = 10$  and absorption coefficients for the frequencies  $\omega_i = 20 \cdot 2^i$  Hz ( $i = 1, \dots, 10$ ), covering the range from 40 to 20480 Hz. On the logarithmic scale, our filters are defined as

$$b_i(\omega) = \begin{cases} c_i(\log_2(\omega/\omega_0)) & \text{if } \omega > 0; \\ \delta_{i1} & \text{if } \omega = 0; \\ c_i(\log_2(\omega_0/-\omega)) & \text{if } \omega < 0, \end{cases} \quad (10)$$

where  $\omega_0 = 20$  Hz and  $\delta_{i1}$  denotes the Kronecker delta. The wavelets  $\psi_i$  are obtained by the inverse Fourier Transform.  $b_0$  and  $b_{10}$  are low- and highpass filters, respectively, whereas the remaining  $b_i$  are bandpass filters with support  $[\omega_{i-1}, \omega_{i+1}]$ , see figure 3. Our filters have the following properties:

- compact support in  $\Omega$
- symmetry and smoothness in both domains
- their sum is one in  $\Omega$  and a Dirac in the time domain.

**Discretization.** The 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. Due to the uncertainty principle,

the pulse response may become non-causal. To limit this effect, we combine the widest filters for 40, 80, 160, and 320 Hz into one (decaying much faster) and cut off all filters after 314 samples (corresponding to 6.5 msec) in both directions. Hence, the low-frequency portions of an echo, filtered due to different absorption coefficients, may be received up to 6.5 msec in advance of the main peak, which in fact cannot be recognized by the human ear. In the case of uniform absorption, the discretized wavelets still form a partition of the unit impulse, providing perfect echos.

Once, the response filter  $f$  has been computed, it needs to be normalized with respect to the number of phonons used and the radius of the gaussian located at the listener. Therefore, we normalize the filter such that the signal's intensity at  $1m$  distance from the source  $s$  (without any reflections) corresponds to the intensity of the anechoic source signal. The normalization factor is simply one over the sum of all gaussian terms collected from the outgoing phonon paths in  $1m$  distance from the source.

An additional extension we did not consider here is the absorption of air,

$$1 - \exp(-\alpha_{air} d), \quad (11)$$

where  $\alpha_{air}$  is the absorption coefficient (depending on air humidity and temperature) and  $d$  is the traversed distance. This term is extremely small, compared to the absorption of materials, such that it is not considered in our simulation.

### 3.3 Visualization

The computationally most expensive part of our method is the computation of ray intersections for phonon tracing and the visibility determination in the collection phase. Once the phonon map composed of phonon paths with instances at each reflection has been computed, it can be used for interactive visualization. When dropping the visibility check in the collection phase (or substituting a crude approximation), the filters for a varying listener position can be estimated at nearly interactive rates.

Our visualization focuses on the spacial propagation of a pulse response from the sound source. The corresponding wave front traverses the room and is reflected at surfaces, altering its intensity and energy spectrum. We visualize this wave front by rendering a small sphere for every phonon path with color coded spectral energy. Therefore, we use the RGB components, such that blue corresponds to the average of  $e_{p,1}, \dots, e_{p,4}$  (40, 80, 160, 320 Hz), green corresponds to  $e_{p,5}, \dots, e_{p,7}$  (640, 1280, 2560 Hz), and red to  $e_{p,8}, \dots, e_{p,10}$  (5120, 10240, 20480 Hz).

When sliding through time, the spheres follow the phonon paths. At small time values, the reflected wave fronts are clearly visible. When the number of reflections increases, however, it becomes more difficult to recognize individual fronts. In order to tackle this problem, we integrated the following function into 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 / traversed 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  $\mathbf{p}_p$  at a reflection point, and its outgoing direction  $\mathbf{v}_p$ , according to section 3.1. 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  $t_{p_i} \leq t < t_{p_{i+1}}$  and to draw a sphere on the line segment  $\mathbf{p}_{p_i}, \mathbf{p}_{p_{i+1}}$ , corresponding to a phonon's location at time  $t$ . In addition, certain predicates can be used, for example selecting all phonon whose path was reflected at the first (or at any) reflection on a selected material.

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.

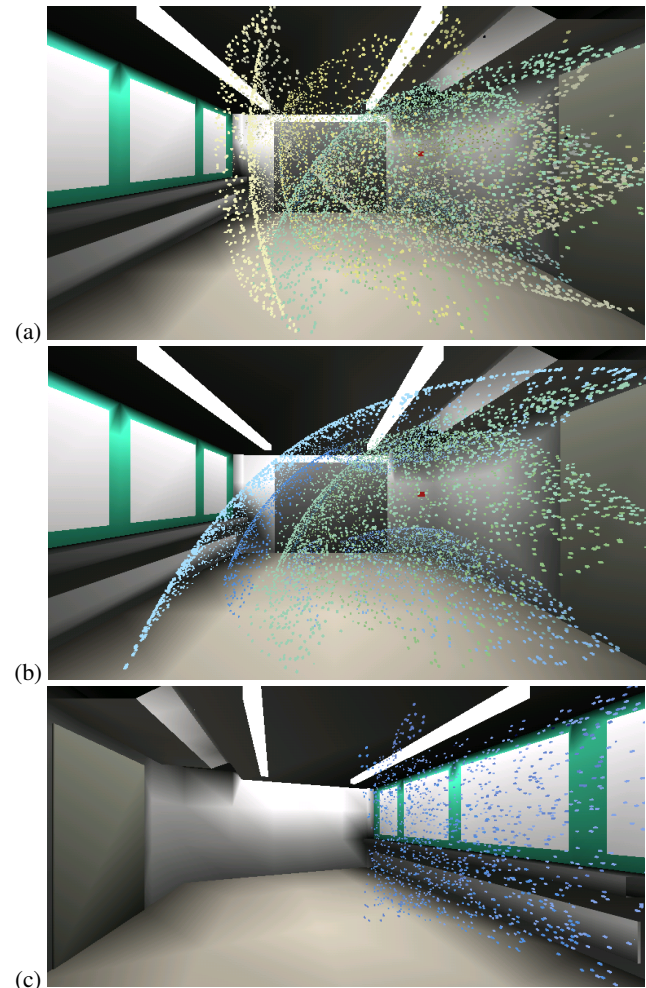


Figure 4: Selecting wave fronts by material reflection. (a) Phonons reflected at least once by a wall; (b) phonons reflected from bottom; (c) phonons reflected from canvas (behind viewpoint).

## 4 RESULTS

### 4.1 Visualization Application

Our visualization was used to examine the acoustic properties of a virtual reality (VR) laboratory that is also used for auralization purposes. Figure 1 shows the wave front propagation based on 1M phonons (33043 paths), 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

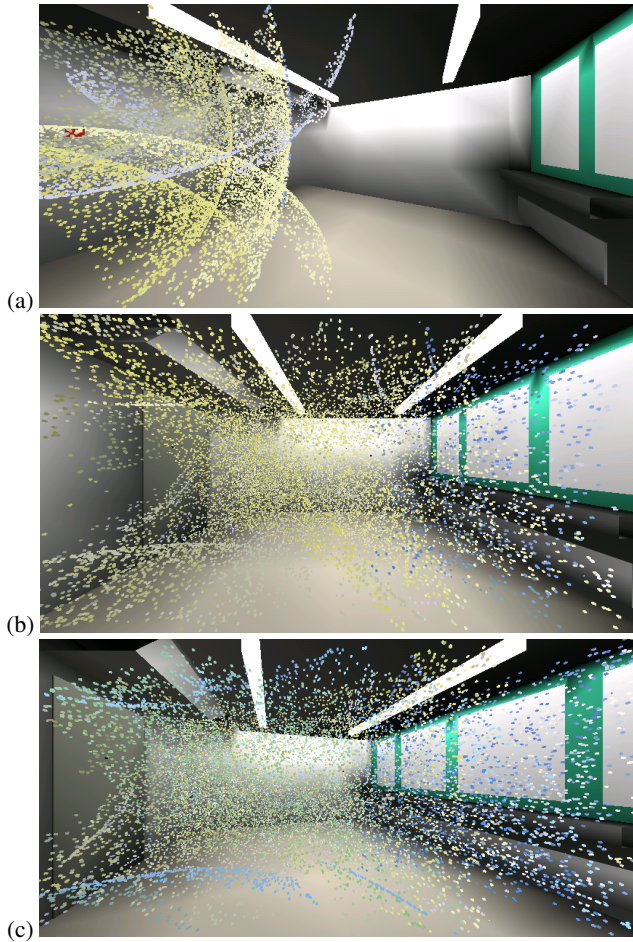


Figure 5: Changing material of bottom. (a) all phonons at  $d = 1.5m$ ; (b) phonons at  $d = 4.5m$ ; (c) same as (b) with old material.

towards lower frequencies, since the phonons' color is dominated by the blue component after a number of reflections.

To identify the reason for the frequency shift, we look at phonons reflected by a selected material (at least once) in their earlier paths. Figure 4 shows these phonons for reflections from walls, bottom, and canvas, respectively, at a distance of 4.5 m where the front direction can still be recognized. We observe that phonons reflected from walls carry mostly a yellowish color, despite of their potential reflection from additional materials. Hence, the energy of these phonons is shifted towards the mid and high frequencies. The bottom 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, see figure 4b. While a potential frequency shift can already be seen in a material's absorption coefficients, their impact on room acoustics can be studied much better with the aid of our visualization.

In figure 5 we replaced the carpet by a material with similar absorption as the walls. When comparing figures 5b and 5c, it becomes evident that this change is sufficient for increasing the intensities of mid and high frequencies. The acoustic properties of this room are significantly improved by this modification, since low frequencies are also propagated by diffraction and thus do not depend on linear reflections that much. 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 mi-

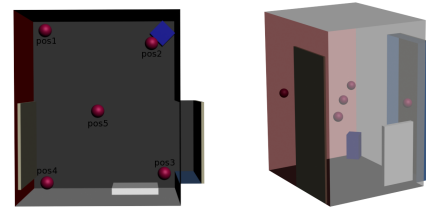


Figure 6: Geometry of closet.

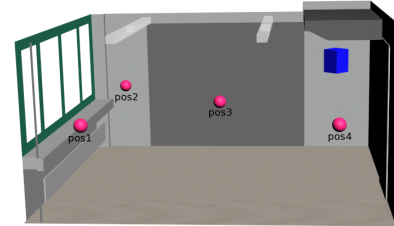


Figure 7: Geometry of the VR-laboratory.

crophone or improve the auditive quality of concerts.

## 4.2 Simulation

We have applied our simulation and visualization method to the virtual models of three different rooms, a closet, the VR-laboratory, and a medium sized classroom. The computation times for tracing 1M phonons on a 2.4GHz Processor are listed in table 1. Due to the large number of reflections, the number of different paths is much smaller than the number of phonons. The times for collecting phonons and filtering are about the same as the tracing times, due to the visibility test. If the latter is dropped, the filters can be constructed at nearly interactive rates. The table shows that the time complexity of our method does not depend very much on the complexity of the scene, due to octree partitioning.

## 4.3 Evaluation

In order to validate our simulation results, illustrated in figures 9 and 10, we compared them to measurements and simulations obtained by a commercial room-acoustics program CARA

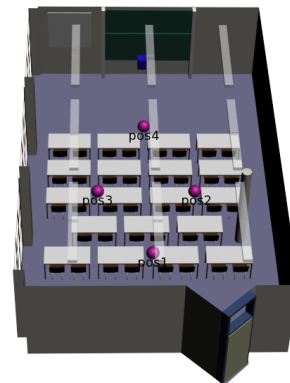


Figure 8: Geometry of the classroom.

room model	points	triangles	materials	phonon paths	avg. reflections	time [sec]
closet	150	156	8	23555	42.5	261
laboratory	521	546	17	33043	30.3	204
classroom	23413	16489	16	25024	40.0	308

Table 1: Computation times for constructing the phonon map based on one million phonons.

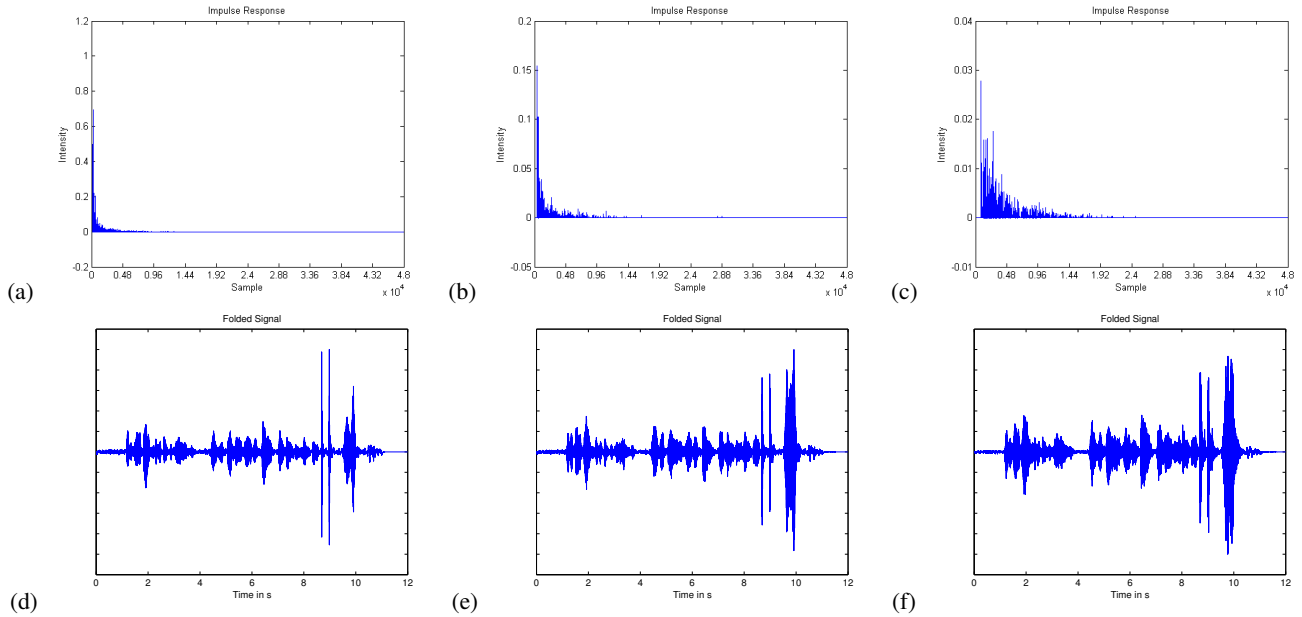


Figure 9: Room impulse response and convolution with the anechoic signal of the three simulated rooms. (a+d) closet; (b+e) laboratory; (c+f) classroom.

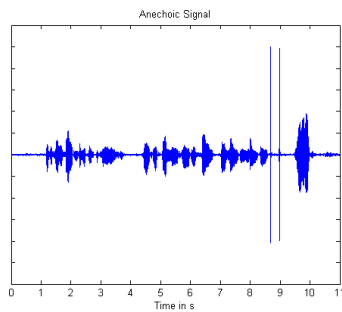


Figure 10: Anechoic signal.

(www.cara.de) based on the image-source method. All signals have a sampling rate of 48 kHz. To keep modeling errors low we chose a simple box-shaped closet room ( $1.9m \times 1.61m \times 2.49m$ , see figure 6) with a small number of faces and different absorption coefficients according to figure 2.

The sound field was produced by a single loudspeaker (blue box) and measurements were taken at five positions, see figure 6. The box was fed with sinusoidal signals of different frequencies ranging from 46.9 Hz to 24000 Hz. Each signal was displayed long enough (0.75 sec) to damp out other frequencies. The same input signal was convolved with the impulse responses simulated with both CARA and our algorithm. Figure 11 shows a comparison of

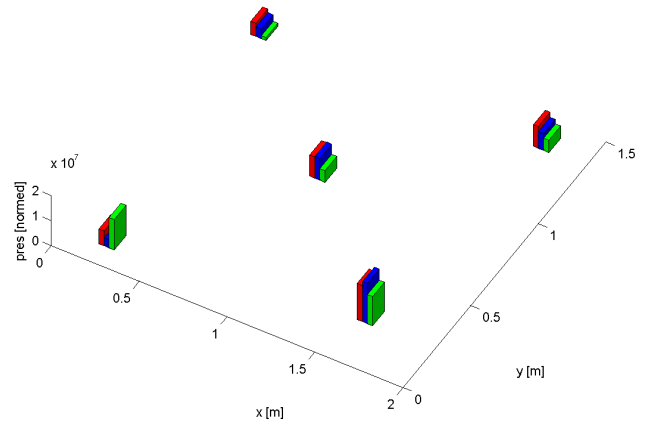


Figure 11: Comparison of the recorded amplitude at 750Hz (red) with results from phonon tracing (blue) and CARA (green).

the measured and simulated pressures at a frequency of 750 Hz. At frequencies lower than 400 Hz, the results of both simulations differ significantly from our measurement, which may be due to diffraction effects and errors in the absorption coefficients (used by both simulations). At medium and high frequencies, our simulated pressures were mostly closer to the measured pressures than those obtained by CARA, due to the greater number of computed reflections. We note, however, that these measurements were very sen-

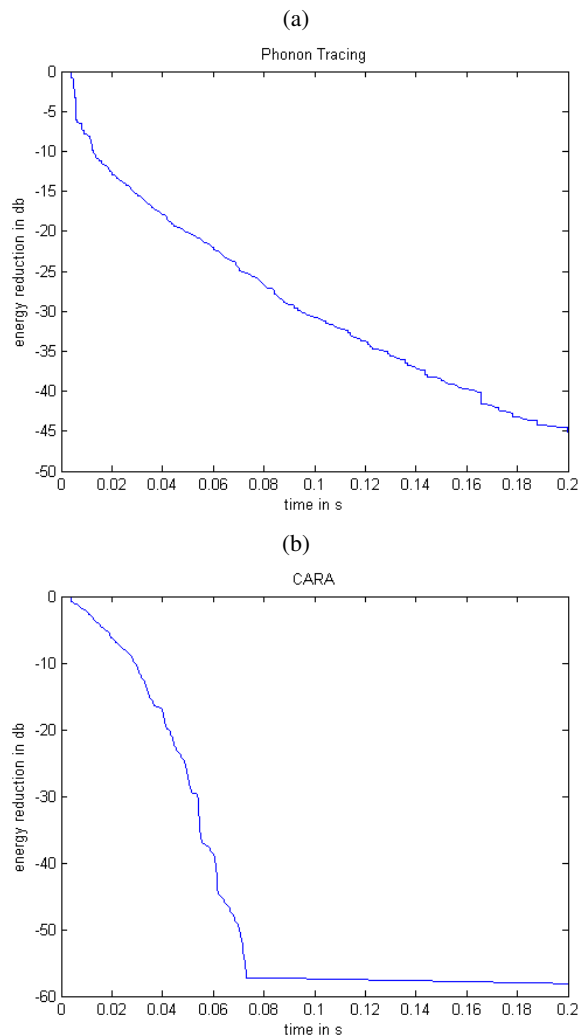


Figure 12: Comparison of residual energy descent in db. (a) phonon tracing; (b) CARA.

sitive with respect to the microphone position, particularly in the high-frequency range.

In addition, we compare the descent of the residual energy, calculated from the impulse response simulated with both CARA and our algorithm. The residual energy descent is illustrated in figure 12. Ideally, the descent is exponential, implying a constant slope in the logarithmic figure. Due to the limited number of reflections considered by CARA, their reverberation is cut off after a short time period, whereas the (logarithmic) decay in our solution is nearly linear.

## 5 CONCLUSIONS

We presented a new particle tracing method for the auralization and visualization of room acoustics. The results obtained by our simulation compare favorably to existing methods. Our interactive visualization helps to optimize acoustic properties of class- and concert rooms. It is particularly useful for determining the impact of certain materials, allowing their interactive exchange. Future work will be directed at the integration of our method into a unified virtual-reality framework.

## ACKNOWLEDGMENTS

This work was supported by the Stiftung Rheinland Pfalz für Innovation under contract no. 15202-386261/644. We thank Lars Kurandt and Norbert Siedow for their helpful comments and discussions.

## 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] J. Borish. Extension of the image model to arbitrary polyhedra. *J. Acoust. So. Amer.*, 75(6):1827–1836, 1984.
- [3] M. Devroye. *Non uniform random variate generation*. Springer, New York, 1986.
- [4] T. Funkhouser, I. Carlbom, G. Elko, G. Pingali, M. Sondhi, and J. West. A beam tracing approach to acoustic modeling for interactive virtual environments. In *Computer Graphics (SIGGRAPH 98)*, pages 21–32, Orlando, FL, July 1998.
- [5] T. Funkhouser, P. Min, and I. Carlbom. Real-time acoustic modeling for distributed virtual environments. In *Computer Graphics (SIGGRAPH 99)*, pages 365–374, Los Angeles, August 1999.
- [6] M. Jedrzejewski and K. 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.
- [7] H.W. Jensen. Global illumination using photon maps. In *Rendering Techniques '96 (Proceedings of the 7th Eurographics Workshop on Rendering)*, pages 21–30, 1996.
- [8] H.W. Jensen and P.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.
- [9] A. Keller and I. Wald. Efficient importance sampling techniques for the photon map. In *Vision, Modelling and Visualization*, pages 271–279. IOS Press, 2000.
- [10] 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.
- [11] N. Korany, J. Blauer, and O.A. Alim. Acoustic simulation of rooms with boundaries of partially specular reflectivity. *Applied Acoustics*, 62:875–887, 2001.
- [12] U. Krockstadt. Calculating the acoustical room response by the use of a ray tracing technique. *Journal of Sound and Vibrations*, 8(18), 1968.
- [13] U. Kulowski. Algorithmic representation of the ray tracing technique. *Applied Acoustics*, 18:449–469, 1984.
- [14] T. Lokki. *Physically-based Auralization*. PhD thesis, Helsinki University of Technology, 2002.
- [15] T. Lokki, U. Svensson, and L. Savioja. An efficient auralization of edge diffraction. In *Audio Engineering Society, 21th International Conference*, pages 166–172, St. Petersburg, Russia, June 2002.
- [16] 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.
- [17] R. Rabenstein, O. Schips, and A. Stenger. Acoustic rendering of buildings. In *5th Int. IBPSA Conference Building Simulation*, Prag, Czech Republic, September 1997.
- [18] L. Savioja, T. Lokki, and J. 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.
- [19] 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.
- [20] 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.
- [21] 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.