A FAST AND ACCURATE "SHOEBOX" ROOM ACOUSTICS SIMULATOR

Steven M. Schimmel, Martin F. Müller, and Norbert Dillier

Laboratory for Experimental Audiology, University of Zurich, CH-8091 Zurich, Switzerland
{steven.schimmel,martin.mueller.orl,norbert.dillier}@usz.ch

ABSTRACT

We present a new "shoebox" room acoustics simulator that is designed to support research into signal processing algorithms that are robust to reverberation. It is an improvement over existing room acoustics simulators because it is computationally fast, portable to many kinds of research environments, and flexible to use. The proposed simulator is also perceptually accurate because it models both specular and diffuse surface reflections. An efficient implementation of the simulator is made freely available for download from the open source project ROOMSIM on SourceForge.

Index Terms**—** Room acoustics, simulation, reverberation, diffuse reflections, binaural room impulse response.

1. INTRODUCTION

Many signal processing applications require algorithms that are robust to reverberation. For example, beam forming and noise cancelation algorithms in hearing aids must function in a wide variety of reverberant conditions. Similarly, it is often desirable that the performance of automatic speech recognition systems does not suffer in the presence of reverberation. Robust algorithms are typically tested in a number of reverberant conditions during their design to evaluate their robustness. These reverberant conditions can be obtained by recording the impulse response of several source and receiver configurations in a number of rooms. Such real measurements, however, are time-consuming to make, and are easily contaminated by acoustic noise. Alternatively, these reverberant conditions can be generated using a room acoustics simulator, a computer program that simulates the acoustics of a given room configuration with high accuracy. This alternative is typically much faster than taking real measurements, thus allowing many more reverberant conditions to be evaluated, and avoids acoustic noise.

For these purposes, a room acoustics simulator should satisfy several requirements, as we have discussed in more detail in [1]. First, it should be accurate. The impulse response it generates for a virtual room should contain all the elements that would be present if the impulse response was recorded in a similar real room. Moreover, in the special case of a binaural room impulse response, the impulse response should be perceptually convincing. That is, when the impulse response is convolved with a source signal and played back to a human listener, it should give the listener the impression of being present in the room, and the source and the room should be perceived as they were specified in the simulation. Second, the simulator should be fast, so that many reverberant scenarios can be quickly evaluated. Third, it should be portable to a wide variety of research environments.

There are several types of room acoustics simulators available today. The freely available ROOMSIM program for MATLAB [2] is designed for simulation of shoebox rooms for educational and research purposes. The MATLAB implementation of this simulation is accessible and easy to understand, but also renders the simulator computationally slow. Moreover, it does not simulate diffuse surface reflections, which are considered to be an important component of reverberation in real rooms [3]. Its main interface of interactive menus also makes it less suitable for batch processing of reverberant scenarios. Another class of simulators are commercial architectural acoustics software programs such as Odeon [4]. These programs are designed to simulate the acoustics of geometrically complex rooms such as churches, theaters and concert halls for accurate prediction and diagnosis of room acoustic properties. They require precise specification of the room geometry, their algorithms are optimized for the computation of room acoustic properties rather than for a room impulse response, and they are not necessarily tuned for speed. A third kind of simulators is designed for virtual acoustics (for example, the CAVE simulator [5]). Their purpose is to compute the impulse response of a given room setup in real time, which is then used to augment the images of a virtual reality system with a plausible acoustic room impression. These simulators typically run on dedicated hardware with enough computing power to meet the real time requirement, and are therefore less portable to other research environments.

Since none of the existing simulators satisfy all of the stated requirements, we propose a new room acoustics simulator that meets all of the needs. This simulator is similar to the ROOMSIM program for MATLAB [2], but it is optimized for speed and supports diffuse surface reflections. In this paper, we first give a brief introduction into physical room acoustic phenomena in section 2 to define some terminology.

This research was jointly funded by Phonak AG and the Swiss NSF on KTI grant #9646.1/PFLS-LS. Part of this work was presented at the 8th ITG conference on Speech Communication, Aachen, 2008 [1].

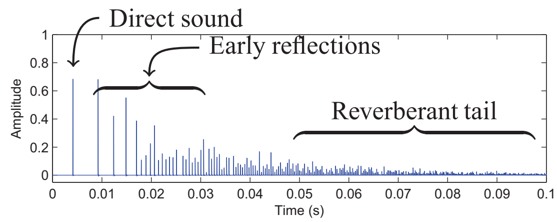

Fig. 1: Components of a room impulse response.

We then describe the simulator in section 3, and evaluate its performance in section 4. Finally, we conclude with a discussion of future research directions for the proposed room acoustics simulator.

2. ROOM ACOUSTICS

The intensity of sound emitted by most sound sources varies with direction and frequency. As the sound propagates in a room, its intensity decreases proportional to the square of the distance it has traveled, because the initial energy is spread over a sphere of continually increasing radius. The sound intensity will also drop due to air absorption. When sound hits a surface, part of the sound energy is absorbed by the surface, part of it is reflected specularly (i.e., the angle of incidence equals the angle of reflection), and part of it is reflected diffusely (i.e., scatter in every direction). The amount of absorption, specular reflection, and diffuse reflection are frequency dependent properties of the surface. Sound is also absorbed and reflected by objects in the room, and diffracted around edges. When sound reaches a receiver, the receiver's directional sensitivity as well as the physical presence of the receiver in the sound field alter the sound intensity that is finally perceived or recorded.

The multiple sound propagation paths from a source to a receiver in a room can be modeled accurately by a linear time-invariant system defined by the so-called *room impulse response*. A room impulse response contains several perceptually relevant components, as shown in Figure 1: direct sound, early specular reflections, and the reverberant tail. The direct sound corresponds to the propagation path that involves no surface reflections. It is most important perceptually because it is arrives straight from the source and, in the binaural case, is primarily responsible for the perceived location of the sound source. The early reflections are the propagation paths that reflect off only a few surfaces and arrive at the receiver roughly within 50 ms after the direct sound. Their arrival rate is usually low enough that the auditory system can resolve each of the reflections individually. They provide a sense of presence to a listener and in many cases enhance the perception of the sound. The reverberant tail consists of all propagation paths that reflect off many surfaces before reaching the receiver. They arrive so rapidly that the auditory system

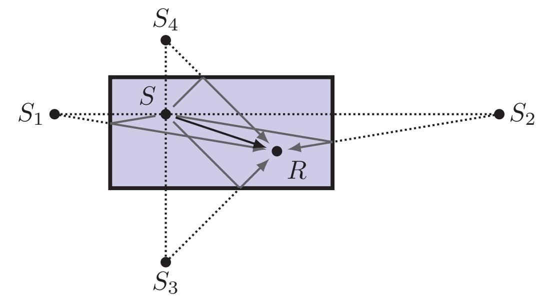

Fig. 2: Virtual image sources correspond to multiple sound propagation paths in the image source method.

does not resolve them individually, but rather integrates them temporally and spatially into a combined percept. The reverberant tail characterizes the surface absorption and size of the room, and the ratio of direct sound to reverberation energy translates into a sense of distance of the sound source.

It is well known that diffuse surface reflections are another perceptually significant component of room impulse responses (see for example [3, 6]). In rooms with diffusely reflecting surfaces, sound energy generally reaches the receiver earlier and from more directions than in an identical room with purely specularly reflecting surfaces. As a consequence, diffuse reflections reduce the spectral comb-filter effect of strong specular reflections, shorten the reverberation time, and produce a smoother exponential energy decay of the reverberant tail [6].

3. THE ROOM ACOUSTICS SIMULATOR

3.1. Room model

The room acoustics simulator that we propose is restricted to empty "shoebox" rooms, which implies that it can handle neither geometrically complex rooms nor objects in a room. This restriction constitutes a considerable reduction in the computational complexity of the simulator, while still permitting an important class of rooms. A "shoebox" room is modeled as a three-dimensional space bounded by six rectangular faces. The faces, or surfaces, are described by frequency dependent absorption and scattering coefficients that can be manually configured, or that can be selected from a database of common surface materials. The location and orientation of the sound source and receiver are specified by x, y and z coordinates and yaw, pitch and roll angles. The sound source directivity and receiver sensitivity can be set to an idealized pattern, for example omnidirectional or cardoid, or a measured (binaural) response from the publicly available MIT KEMAR [7], CIPIC [8], or LISTEN [9] head-related transfer function (HRTF) datasets. Additionally, the simulator supports a large number of options to control details of the room model and the exact output of the simulator.

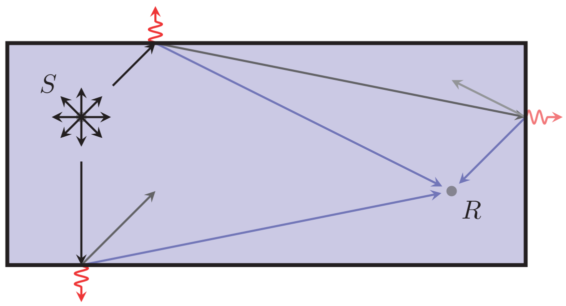

Fig. 3: In the diffuse rain algorithm, sound rays (black) are emitted from the source and traced trough the room. When a ray hits a surface, part of its energy is absorbed (red) and the contribution of a diffuse secondary ray (blue) to the receiver is computed before the ray is continued in a random direction (gray).

3.2. Simulation methods

The room acoustics simulator uses two techniques to generate a room impulse response: the *image source* method [10] and the *diffuse rain* algorithm [5, 11]. In the image source method, the sound source is mirrored in the room's surfaces to create virtual image sources, as illustrated in Figure 2. A straight line from a virtual source to the receiver corresponds to a sound propagation path in the room involving one or more surface reflections. By knowing the line's length and the absorption of the surfaces it intersects, the contribution of the corresponding sound propagation path to the room impulse response can be computed. The image source method is guaranteed to find all the propagation paths in a room exactly. This makes it ideally suited to simulate direct sound and low order reflections. However, the method is computationally inefficient for high order reflections, as the number of virtual sources grows rapidly with reflection order. Moreover, it is only suited to model specular surface reflections.

The room acoustics simulator uses the diffuse rain algorithm to generate the higher order reflections, the reverberant tail and the diffuse reflections. This algorithm was chosen because it is able to generate these elements of the room impulse response efficiently. In the diffuse rain algorithm, sound energy rays are emitted from the source and traced throughout the room, as illustrated in Figure 3. When a sound ray hits a surface, the sound energy carried by the ray is reduced according to the absorption of the surface. Next, a direct contribution of the ray to the room impulse response is determined by shooting a secondary ray from the point of impact to the receiver and logging this ray's angle of incidence, arrival time, and remaining sound energy at the receiver. The original ray is then continued from the point of impact in a random direction, and traced further until its sound energy drops below threshold. All rays are handled in this way, and the process is repeated for all frequency bands because absorption and diffuse reflection are frequency dependent phenomena.

At the receiver, the energy of a ray arriving at time t for frequency f at an angle $\psi = (\theta, \phi)$ is accumulated in the time-frequency histogram $E_i(n, k)$ as illustrated in Figure 4, where t, f and ϕ are discretized to bin indexes as follows

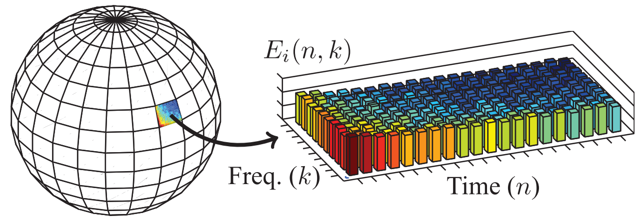

Fig. 4: Time-frequency sound energy histogram for one spherical bin at the receiver (after [5]).

$$n = \underset{n'}{\operatorname{argmin}} |t - n'\Delta t|,$$
$$k = \underset{k'}{\operatorname{argmin}} |f - f_{k'}|,$$
$$i = \underset{i'}{\operatorname{argmin}} ||\psi - \psi_{i'}||,$$

with Δt the time resolution of the histogram, f_k the center of the k^{th} frequency band, and ψ_i the angle of the center of spherical bin i. For each spherical bin, the time-frequency histogram is converted to an impulse response in a few steps. First, a Poisson noise process

$$p_i(t) = \sum_r \delta(t - \tau_r)$$

is generated such that the number of arrivals $A_i(n)$ per time Δt,

$$A_i(n) = \int_0^{\Delta t} p_i(t - n\Delta t)dt,$$

follows a Poisson distribution, i.e.,

$$\Pr[A_i(n) = m] = \frac{[\lambda_i(n)]^m e^{-\lambda_i(n)}}{m!},$$

where the rate parameter $\lambda_i(n)$ is proportional to the sound energy at time n, $\lambda_i(n) \sim \sum_k E_i(n, k)$. Next, the spectrogram of the Poisson noise signal $p_i(t)$, which is spectrally white, is shaped according to time-frequency histogram $E_i(n, k)$. Then, the shaped noise signal is convolved with the receiver's response for the bin's center angle ψ_i. Finally, the resulting signals of all spherical bins are summed into an impulse response, which is superimposed on the output from the image source method to yield the full room impulse response.

3.3. The ROOMSIM project

We have developed an efficient C implementation of the algorithms described in section 3.2. This implementation of the proposed room acoustics simulator is freely available for

Room	
Dimensions (m)	$6.25 \times 3.75 \times 2.5$
Inverse square law	Yes
Air absorption	Yes
Humidity	50%
Temperature	20°C
Surfaces	
Frequency bands (Hz)	125 250 500 1000 2000 4000
Absorption	0.1 0.3 0.8 0.85 0.75 0.65
Scattering	0 0 0 0 0 0
Source	
Position	$(3.5, 1.8, 1.1)$
Orientation	$0, 0, 0$
Type	Omnidirectional
Receiver	
Position (m)	$(1.5, 1.8, 1.1)$
Orientation	$0, 0, 0$
Type	Omnidirectional

Table 1: Settings of the simulated room. Scattering (i.e., diffuse reflections) is disabled because it is not supported by the MATLAB version of ROOMSIM.

download from `roomsim.sourceforge.net`. The implementation is platform independent, and currently has interfaces for MATLAB and C/C++.

4. EVALUATION

To demonstrate the speed improvement of the proposed room acoustics simulator over the ROOMSIM program for MATLAB [2], we generated the impulse response of a room with both simulators using the settings given in Table 1. The resulting room has an RT_{60} of 0.245 seconds (averaged over frequency). Its maximum RT_{60} (at 125 Hz) is 0.924 second, which the MATLAB version truncates to 0.736 seconds in order not to run out of memory. The maximum reflection order for the x, y, and z dimension are 21, 34, and 51, requiring the computation of 1,150,510 image sources.

The average running time of the MATLAB simulator on this task was 486.87 seconds ($N = 10$, $\sigma = 0.96$), whereas that of the simulator proposed in this paper was 11.69 seconds ($N = 10$, $\sigma = 0.02$). Hence, the proposed simulator constitutes a conservatively estimated $40\times$ speed up.

5. CONCLUSIONS AND DISCUSSION

In this paper, we have presented a new room acoustics simulator that is fast, accurate, and flexible. It models low-order specular reflections using the image source method, and models high-order and diffuse reflections using the diffuse rain algorithm. An efficient C implementation of the simulator is freely available for download from the ROOMSIM project on SourceForge at `roomsim.sourceforge.net`.

A manuscript that describes the room acoustics simulator in detail is in preparation. In this manuscript, we fully describe the simulation algorithms and evaluate the simulator's computational performance and perceptual accuracy.

For our research, an important aspect of room acoustics simulation is the generation of binaural room impulse responses. The proposed room acoustics simulator is currently able to generate these using one of the publicly available HRTF datasets [7–9]. Binaural room impulse responses generated with an HRTF selected from one of these datasets are useful to evaluate binaural signal processing algorithms in reverberant conditions. However, such impulse responses may not be entirely perceptually convincing to a human listener, because non-individualized HRTFs are known to cause front-back confusion and incorrectly perceived elevation. To incorporate individualized HRTFs in the room acoustics simulator, we are considering the work by Duda and colleagues [12, 13] to model an individual's HRTF based on anthropomorphic measurements.

6. REFERENCES

[1] S. M. Schimmel, M. Müller, and N. Dillier, "Efficient and accurate room acoustics simulation for binaural signal processing," in *ITG-Fachtagung Sprachkommunikation (Aachen, Germany), accepted for publication*, 2008.

[2] D. R. Campbell, K. J. Palomäki, and G. J. Brown, "A MATLAB simulation of "shoebox" room acoustics for use in research and teaching," *Computing and Information Systems Journal*, vol. 9, 2005.

[3] M. Hodgson, "Evidence of diffuse surface reflections in rooms," *JASA*, vol. 89, pp. 765–761, 1991.

[4] G. M. Naylor, "ODEON—Another hybrid room acoustical model," *Appl. Acoustics*, vol. 38, pp. 131–143, 1993.

[5] T. Lentz, D. Schröder, M. Vorländer, and I. Assenmacher, "Virtual reality system with integrated sound field simulation and reproduction," *EURASIP J. Adv. in Sig. Proc.*, pp. 1–19, 2007.

[6] B.-I. Dalenbck, M. Kleiner, and P. Svensson, "A macroscopic view of diffuse reflection," *J. AES*, vol. 42, pp. 793–807, 1994.

[7] W. Gardner and K. Martin, "HRTF measurements of a KEMAR dummy head microphone," Tech. Rep. 280, MIT Media Lab, Perceptual Computing, 1994.

[8] V. R. Algazi, R. O. Duda, D. M. Thompson, and C. Avendano, "The CIPIC HRTF database," in *WASSAP*, 2001.

[9] O. Warusfel, "The LISTEN HRTF database," `http://recherche.ircam.fr/equipes/salles/listen/`.

[10] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *JASA*, vol. 65, no. 4, pp. 943–950, 1979.

[11] R. Heinz, "Binaural room simulation based on an image source model with addition of statistical methods to include the diffuse sound scattering of walls and to predict the reverberant tail," *Appl. Acoustics*, vol. 38, pp. 145–159, 1993.

[12] V. R. Algazi, C. Avendona, and R. O. Duda, "Estimation of a spherical-head model from anthropometry," *J. AES*, vol. 6, pp. 472–479, 2001.

[13] P. Satarzadeh, V. R. Algazi, and R. O. Duda, "Physical and filter pinna models based on anthropometry," in *Proc. of the 122nd AES convention*, 2007.