Simulating transfer functions in a reverberant room including source directivity and head-shadow effects

Martin Kompis and Norbert Dillier
ENT-Department, University Hospital, CH-8091 Zurich, Switzerland

(Received 30 January 1991; revised 15 December 1992; accepted 12 January 1993)

A new procedure for simulating the acoustical response of a receiver to a source in a reverberant room is proposed. Convolving such an impulse response with an input signal, e.g., speech, is useful in applications where an acoustical environment must be highly reproducible, easily controllable, or is unavailable for an actual recording. The proposed procedure is an extension of the image method by Allen and Berkley [J. Acoust. Soc. Am. **65**, 943–950 (1979)]. The new features include the possibility to simulate simple source and microphone directivity, as well as head-shadow effects, which can be considerable for subjects with microphones mounted close to the head. A rigid sphere diffracting a plane wave provides a good approximation for the effects of speaker as well as listener head shadow and allows simple yet realistic simulations of many relevant situations. Implementational issues are discussed and the method is verified by comparing simulation results with recordings in actual rooms. It was found that by including source directivity and head shadow, transfer functions become significantly more realistic.

PACS numbers: 43.55.Ka, 43.66.Pn, 43.66.Ts

INTRODUCTION

The evaluation of new algorithms for digital hearing aids requires careful consideration of the acoustical environment. Multimicrophone recordings in different rooms and situations however are not only time consuming, but also suffer from inherent constraints. The number of rooms available for recordings is limited, and often a considerable level of acoustic noise cannot be eliminated.

Simulating a room response and convolving it with test signals, rather than performing recordings in real rooms, overcomes these limitations. Among the advantages of this latter approach is the ability to establish a well-defined and easily controllable environment, which is reproducible and offers the possibility to explore a great variety of rooms.

Different methods for simulating room transfer functions are known.[1] Some of them enable the user to simulate acoustically complex environments, like concert halls, or to introduce directivity of sound sources and microphones, including microphones mounted in the ear canals of a subject.[2] Often, these sophisticated methods require extensive measurements of the acoustic environment prior to the simulation.

For many applications in hearing aid research, this effort is not justified. The efficient simulation of easily controllable, well defined, but still realistic environments can be more important than the design of a detailed model of one special room.

An efficient method of simulating impulse responses in small rooms, originally described by Allen and Berkley,[3] has been used for this purpose by Feder *et al.*[4] and Peterson.[5] In the context of digital hearing aids, this method is excellent in modeling some aspects, e.g., the reverberation time of the simulated room, but weak in others.

One major drawback of this method is that only omnidirectional sources and microphones can be modeled. In hearing aid applications, however, the microphones are often mounted close to a subject's head, resulting in a head-shadow effect, which, in most cases, is not negligible. Furthermore, the directivity of the source can also be important.

In this paper, an improved version of the simulation procedure by Allen and Berkley[3] is proposed. The influences of head-shadow and source directivity are introduced into the model, resulting in a considerably more realistic simulation when listeners are involved, while still maintaining a relatively simple simulation procedure. The simplifications used are discussed and some results of the new simulation method are validated using measurements in actual rooms.

I. MODELING THE HEAD SHADOW

A popular approach[6,7] to model the head shadow uses the diffraction of a plane harmonic wave by a rigid sphere, which can be calculated analytically. Simpler models, such as two omnidirectional microphones positioned at a slightly greater distance than the actual head diameter, have been proposed[8] and used for simulation purposes.[9] The rigid-sphere model, however, yields a more realistic model of interaural time difference, as demonstrated by Kuhn,[7] and interaural intensity difference, as will be shown in Sec. V. Furthermore, the diffraction of sound waves by the head of a speaker or listener introduces a spectral distortion to the signal, which in many cases may not be negligible. This effect is also modeled by the rigid-sphere model, as shown below.

The analytical computation of the diffraction of a plane harmonic wave by a rigid sphere has been treated before[10,11] and is not repeated here. Some implementa-

 0001-4966/93/052779-09$06.00

tional problems concerning our application will be discussed in Sec. III.

Besides the rigid sphere, more detailed models of a human listener are known. The transformational properties of the pinna and the ear canal of a listener can be included by a computational model of the outer ear,[12] or, instead of using an analytical model, the impulse response of microphones placed in the ears of a listener to an acoustic point source at different positions can be directly measured and included in the simulation procedure.[2] Both methods can account for the effects of the listener's body on the transfer functions, which are neglected in the simple rigid sphere model.

However, these sophisticated techniques include some disadvantages, which make the simpler rigid sphere model more suitable for a number of applications. Since it does not require complex measurements of transfer functions to microphones mounted at the head of a listener, simulations can be performed with virtually no preparation time. Furthermore, the use of a simple rigid sphere model overcomes problems caused by interindividual variability of transfer functions around the head of different listeners.[13]

For certain applications in the field of hearing aid research, the inclusion of pinna and ear canal properties into the model are not desirable. Here, the microphones are often mounted above the ears. If they are in the ear, the ear canal and often also the concha is occluded by the hearing aid itself, thus changing the transformational properties of the outer ear significantly. In an early stage of the research, the final placement of the microphones may not even be known. The simple rigid sphere model is a reasonable compromise for these cases. Whether the rigid sphere model is appropriate for a certain application or not needs to be determined based on real measured data. In this section, the rigid sphere model will be compared to measurements of other authors. In Sec. V, the complete simulation procedure including the head model will be evaluated using measurements in real rooms.

A. The listener's head shadow

When modeling the head shadow of a listener as the effect of a rigid sphere, Kuhn[7] found that a sphere with a radius of 9.3 cm should be chosen to match actual interaural time differences. As for the azimuthal dependence of the intensity at a listener's ear, data can be found in the literature. Shaw[14] compiled the results from 12 studies, covering the data of 100 subjects. The investigation of Mehrgardt and Mellert[15] is also useful.

Figure 1 shows a comparison between the rigid-sphere model and the data taken from these studies. For four frequencies, the intensity at a microphone placed in the ear of a listener is displayed as a function of the position of the sound source. The source is assumed to be at a great distance and to lie in a horizontal plane through the subject's ear axis. Thus no proximity effects will occur. The azimuth of the source, the angle between the front direction, and the vector pointing from the subject's head toward the source in the horizontal plane is varied. An azimuth of

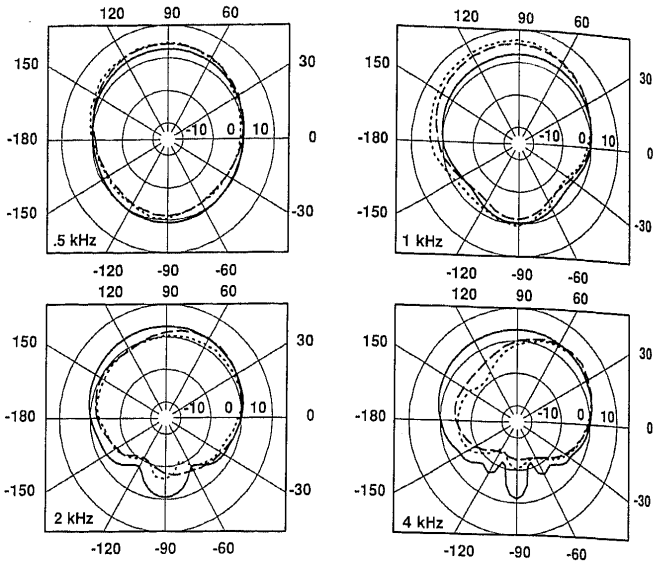

FIG. 1. Azimuthal dependence in dB. Solid lines are for the rigid sphere model, compared with data from Shaw[14] (dashed lines) and Mehrgardt and Meller[15] (dotted lines).

$+90°$ describes the situation in which the ear with the affixed microphone is irradiated directly, while at $-90°$ the source is at the contralateral side of the listener. For our model, the radius of the sphere was 9.3 cm and a speed of sound of 340 m/s was assumed. All data in Fig. 1 have been normalized to yield 0 dB for the front direction. Figure 1 shows that the rigid sphere is a good model at low frequencies, whereas at 4 kHz, the measured data deviate considerably due to pinna effects, which are not considered in our simple sphere model.

We assume the ears to be placed at $\pm 90°$ relative to the front direction. Blauert[6,16] prefers an angle of 100° which is not obvious from the data given in Fig. 1, especially when considering the lobing around $-90°$. However, with the procedure described in Sec. II, any angle can be modeled easily.

Our model does not account for body effects. While this does not seem to disturb the transfer functions in the horizontal plane, effects both in the upper and lower hemisphere can be expected.[12] Unfortunately, only very few measurements for the lower hemisphere are reported in the literature, where the influence might be serious.

B. The speaker's head shadow

The principle of reciprocity states that the response at a point B due to the excitation of a source at point A is the same as the response at point A due to the excitation at point B. Hence, the diffraction of a plane wave by a rigid sphere can be reinterpreted as a transfer function between a point source on the surface of the sphere and a receiver placed at a considerable distance. Therefore, the same approach as above may be used to model the speaker's head shadow.

Lacking any relevant data for an optimal radius of a

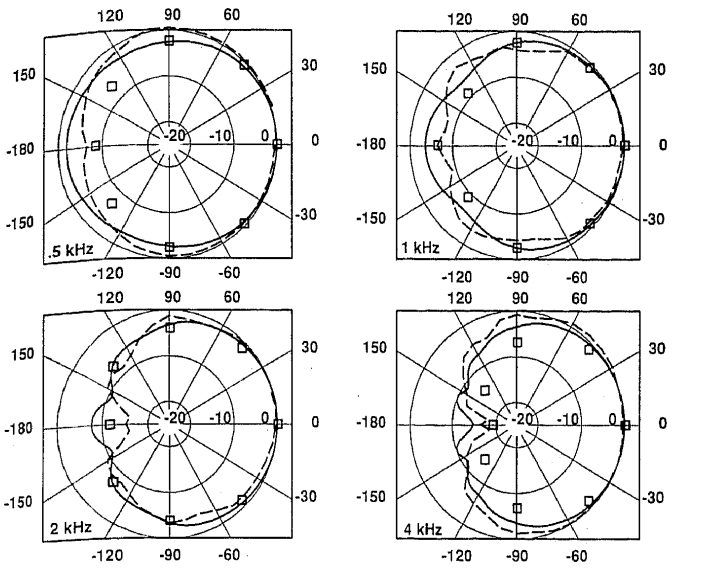

FIG. 2. Sound-pressure level in dB in the horizontal plane of a speaker. Solid lines are for the rigid sphere model, compared with data from Flanagan[18] (dashed lines) and Dunn and Farnsworth[17] (squares).

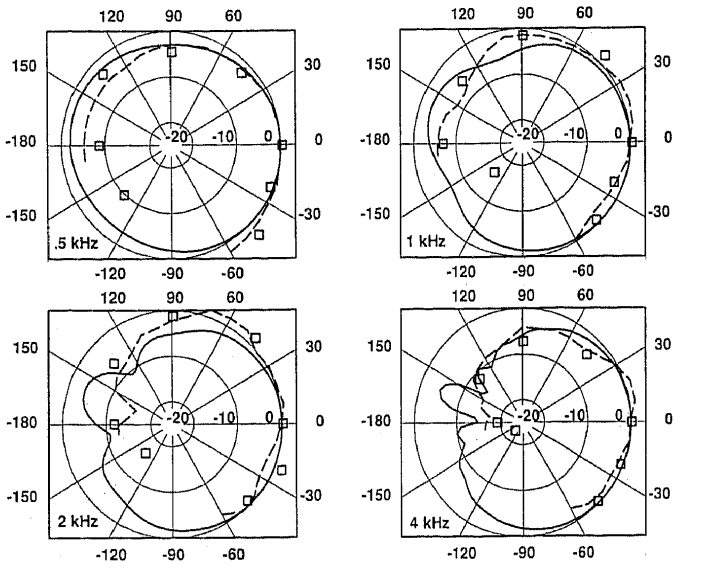

FIG. 3. Sound-pressure level in dB in the median plane for a speaker. Solid lines are for the rigid sphere model, compared with data from Flanagan[18] (dashed lines) and Dunn and Farnsworth[17] (squares).

sphere modeling a speaker's head shadow, the value of 9.3 cm found by Kuhn[7] was again used.

To compare the model with actual measurements, we used the data found by Dunn and Farnsworth[17] and Flanagan.[18] Dunn and Farnsworth measured the sound pressure at 76 positions around the head of a single speaker in octave and half-octave bands, whereas Flanagan used a life-size manikin equipped with a transducer in its mouth and measured the resulting sound-pressure level in the horizontal and the median plane trough the transducer at five different frequencies. The mouth-to-receiver distance was 30 cm in Flanagan's investigation, while Dunn and Farnsworth varied the distance between 5 and 100 cm. From the latter investigation, the measurements performed at a distance of 60 cm were taken. They compromise best between the prevention of proximity effects, number of measurements available, and error due to reverberation from the walls.

Figures 2 and 3 show the comparison between our head model and the series of measurements for the horizontal and median plane. Note that in our model the position of the mouth has been chosen to be 20° below the horizontal plane. It is difficult to estimate this angle exactly, since the deviations of the shape of a human head from a sphere are quite large. Consequently, the lobe on the opposite side of the mouth, which would be the best indicator, is smeared. Nonetheless, a correspondence between measurements and model can still be observed. For most positions of the receiver (except for the lower hemisphere at the back of the speaker) the differences between the simple spherical model and the measurements are in the same order of magnitude as between the results of the two investigations, to which they are compared. Again, as for the listener, no body effects were considered, leaving the quality of the model undertermined for a portion of the lower hemisphere.

II. PROCEDURE TO CALCULATE THE IMPULSE RESPONSE

The image method for simulating an empty, shoebox-shaped room[3] computes the transfer function between a given source and receiver in an enclosure as the sum of all given reflections within a certain period of time. Our method is based on the same principle. The procedure to compute the contribution of a single reflection is illustrated in Fig. 4. As in Ref. 3, for every reflection, an image R_0' of the source location R_0 is calculated. The length of the vector

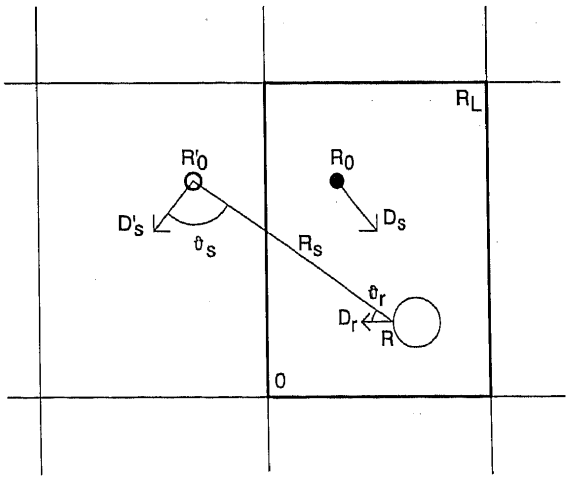

FIG. 4. Example for the calculation of a single reflection. The room with its lower left corner in the origin O and the opposite corner at RL is shown as a solid rectangle. Images of the room are drawn with thin lines. The solid dot R_0 within the room denotes the position of the source and the solid circle R_0' shows the image of the source for the reflection. R is the position of the receiver on the surface of a sphere (large circle). See text for the explanation of the other elements. Note that this is only a two-dimensional illustration of the process, while in a simulation all calculations must be carried out in three dimensions.

$$R_s = R_0' - R, \tag{1}$$

where R is the position of the receiver, determines the arrival time of the reflection

$$t = |R|/c, \tag{2}$$

where c is the sound velocity. Combined with the damping at the reflecting walls a factor

$$g = \frac{T_s}{t} \prod_{i=1}^{6} \beta_i^{N_i} \tag{3}$$

can be calculated where T_s is the sampling period, β_i is the reflection coefficient, and N_i is the number of reflections at the ith wall. Only real and frequency-independent reflection coefficients β_i are allowed in the simulation. The implications of this simplification will be discussed in Sec. IV.

Allen and Berkley now simply add g to the impulse response at a time T_I, which is the closest representation of t in multiples of T_s. Peterson[5] found it unsatisfactory to model the arrival time as multiples of the sampling period T_s only. He presented a method that is also mentioned in Ref. 3, whereby the echo arrival time is not distorted. This is accomplished by adding the impulse response of a low-pass filter at the exact arrival time, thus spreading the energy in a reflection over several samples rather than concentrating it into a single one. Since this is a better way to represent a reflection, the proposed improvement of the original method was optionally implemented as well, with minor modifications. However, only marginal benefit of this improvement could be obtained. Neither informal listening tests nor experiments with a two-microphone adaptive beamformer noise reduction system for hearing aids[19] revealed any noticeable differences between the original and the improved version of the room simulation method.

In Fig. 4, three additional vectors are shown. Here, D_r points to the direction of preference of the receiver, D_s does the same for the source, and D_s' is the image of D_s. Two angles, ϑ_r and ϑ_s, can be derived, giving the directions from which this specific reflection is received or emitted with respect to the directions of preference. For any angle ϑ and any frequency ω, the diffraction of a harmonic plane wave $P(\vartheta,\omega)$ which is a complex value, can be calculated. In the frequency domain, the values for source and receiver must be multiplied together to account for the convolution in the time domain. Thus we write

$$Y(\omega) = P(\vartheta_s,\omega) P(\vartheta_r,\omega) e^{-j\omega t_F}. \tag{4}$$

The last factor gives a correction for an arrival time which is not a multiple of T_s with

$$t_F = t - T_I. \tag{5}$$

Using an inverse discrete Fourier transform $F^{-1}[\ ]$, the representation of the reflection in the time domain is obtained:

$$y(n) = gF^{-1}[Y(\omega)]. \tag{6}$$

It consists of several samples and is added to the taps of the impulse response representing the time around T_I.

The simulation combines geometrical acoustics for the long distances and diffraction in the proximity of the head. If the distance between source and receiver is long with respect to the diameter of the sphere and the wavelength, a plane wave is approximated in the proximity of the listener, thus treating the diffraction around the head as a diffraction of a plane wave. Evidence for the appropriateness of this method, even for relatively small source-to-receiver distances, is given by experiment III in Sec. V.

A ray may pass through a sphere once or several times before reaching its destination. In this case, only the effects of the first diffraction by the speaker's sphere and the last diffraction by the listener's sphere are considered to be substantial.

Other directional sources or receivers can be chosen by substituting the diffraction P in Eq. (4) by a simple, real-valued, and frequency-independent function for the receiver and/or the source, $\mathrm{REC}(\vartheta_r)$ and $\mathrm{SRC}(\vartheta_s)$, respectively.

Finally, the resulting impulse response is passed through a low-frequency (0.01 of the sampling frequency), high-pass filter to remove an artificially large energy spike at zero frequency, which results directly from the implicitly assumed no-pressure-release boundary.[1]

III. IMPLEMENTATION

For the calculation of the diffraction of a plane harmonic wave by a rigid sphere $P(\vartheta,\omega)$, the formulas provided by Schwarz[10] were used, but others (e.g., Morse[11]) could have been used equivalently. As an infinite sum must be calculated, it is necessary to find a reasonable upper bound. It can be empirically derived[10] (considering also to the corrections provided by Cooper and Bauck[20]) that a number of summands of

$$n = 4 + 9(r/\lambda) \tag{7}$$

will result in an accuracy of approximately 10^{-3} for P. In Eq. (7), r is the radius of the sphere and λ is the wavelength of the incident harmonic wave.

As the calculation of P is very time consuming, we chose a simplified implementation. Instead of calculating P twice for every reflection and every frequency needed, a table was defined for all frequencies and a number of angles between 0 and π. While calculating the impulse response, values from this table were interpolated for any given angle ϑ. The stepsize for the table entries was chosen to be $\pi/45$, which results in an average error due to interpolation of approximately 10^{-3} at a sampling rate of 10 240 s^{-1} (this average error was estimated by weighting the individual errors at 256 linearly distributed angles between 0 and π were with their respective expected frequency of occurrence).

The inverse discrete Fourier transform in Eq. (5) was implemented as a fast Fourier transform of order 32. At a sampling frequency of 10 240 Hz, 32 samples span a time period during which sound can travel approximately twice around a head-sized object. In the time domain, the impulse response of a single reflection decays rapidly toward

the margin of the selected time window. Consequently, this rather short transform introduces an average error due to windowing below 10^{-3}, when compared to very long windows of 1024 samples for example. Again, the error was averaged over 256 different angles as above.

With simple directional speakers or microphones, as in experiment III, where $\mathrm{SRC}(\vartheta_s)$ is zero for many values of ϑ_s, computing time can further be reduced by skipping all calculations of Eqs. (3)–(6) for a reflection, as soon as $\mathrm{SRC}(\vartheta_s)$ is known to be zero.

IV. DISCUSSION OF THE SIMPLIFICATIONS USED

Besides using the rigid sphere model for the listener and speaker, as discussed in Sec. I, several other simplifying assumptions introduced in the last two sections deserve closer examination.

The procedure described above as well as the original method[3] do not allow the modeling of other than empty, shoebox-shaped rooms. While the shape of many small rooms is close to that of a simple box, hardly any room will be completely empty. In most cases, dimensions of furniture and equipment will exceed the wavelength of the acoustic sounds of interest significantly. Today, no simulation method is able to treat this problem accurately. A simple solution is to choose the reflection coefficients of the walls such that the reverberation time of the simulated empty box will be the same as that of the actual room. Examples for this approach will be given in Sec. V. Note however, that this procedure will inevitably cause an inaccuracy in the simulated impulse response.

As another simplification, only real and frequency-independent reflection coefficients of the walls are allowed. In reality, reflection coefficients will normally be complex and frequency dependent.[21] To justify the introduction of complex reflexion coefficients, the acoustical properties of the walls should either be known or measured beforehand. For the simulation of an average room, real coefficients can provide sufficiently realistic approximations, as will be shown in Sec. V.

In many, but not all rooms, the reverberation time does not strongly depend on the frequency, suggesting that at least the averaged absorption coefficients over all boundaries are more or less frequency independent. When rooms with significant differences in the reflection coefficients for different frequency bands are to be modeled, simulations with different coefficients can be prepared separately, adequately bandpass filtered and summed.

A similar problem may occur with the frequency-independent directivities $\mathrm{REC}(\vartheta_r)$ and $\mathrm{SRC}(\vartheta_s)$ introduced in the last section. Many acoustical sources will show more directivity at higher than at lower frequency. If this is a problem, several simulations with different directivities can again be bandpass filtered and added.

V. VERIFICATION OF THE SIMULATION RESULTS

In this section, three simple experiments carried out in actual acoustic environments are described. To demonstrate the quality of our simulation method, every situation is modeled as equal as possible and the results are compared to those of the respective experiment.

In all experiments, realistic and therefore complex acoustical environments with frequency-dependent reflection coefficients and source directivities as well as a dummy head including pinnae and a body substitute were used. Measurements were compared to the simulation procedure that yielded surprisingly good results despite its simplifications. Results with the new method were significantly superior than those with the original version which did not consider head-shadow and source directivity.

A. Overall sound pressure at a listener's ears

To demonstrate the important effect of head-shadow, two experiments were performed in an actual room. The room had a reverberation time (time required for reverberant energy to decay by 60 dB) of 0.37, 0.40, 0.42, 0.40, 0.44, and 0.46 s, when measured with octave-wide noise centered around the frequencies 125, 250, 500, 1000, 2000, and 4000 Hz, respectively. The listener was a dummy-head equipped with a stereo head-microphone, both from a Sennheiser MKE 2002 set. The listener's body was accounted for by using the body-substitute box provided by the set for mounting the dummy head. The directivity of a loudspeaker (Philips 22AHS86/16R) was measured beforehand using low-pass filtered white noise with a corner frequency of 5 kHz. For the angles 0°, 45°, 90°, 135°, and 180° from the front direction in the horizontal plane, the intensity of the emitted sound was measured outdoors. The results, normalized to the front direction, were 0, −2, −8, −11, and −13 dB. For both experiments, the loudspeaker was driven with white noise. The output of the microphones was low-pass filtered at 5 kHz and then digitized at a rate of 10 240 s^{-1}. For every position of the loudspeaker, the interaural sound-pressure level difference was derived for a signal of approximately 3-s duration.

While the position of the dummy-head remained the same during experiments I and II, the position of the loudspeaker was varied as follows. In experiment I, it was moved at a distance of 1 m from the center of the dummy-

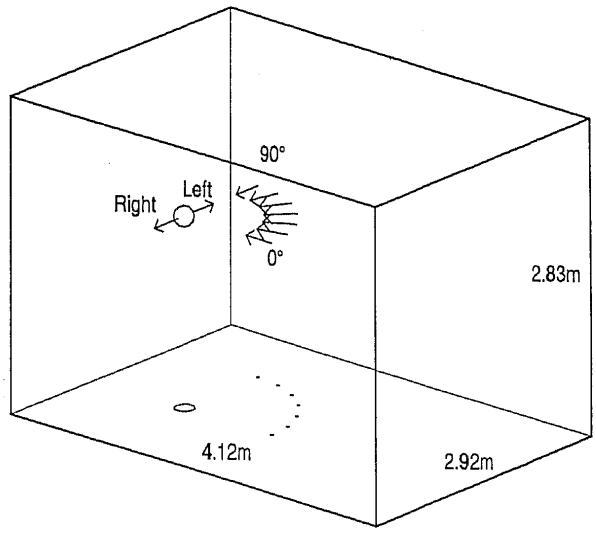

FIG. 5. Schematic of the setting for experiment I.

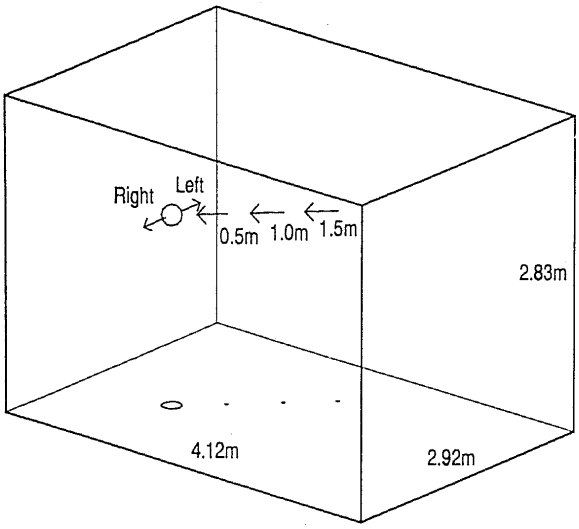

FIG. 6. Schematic of the setting for experiment II.

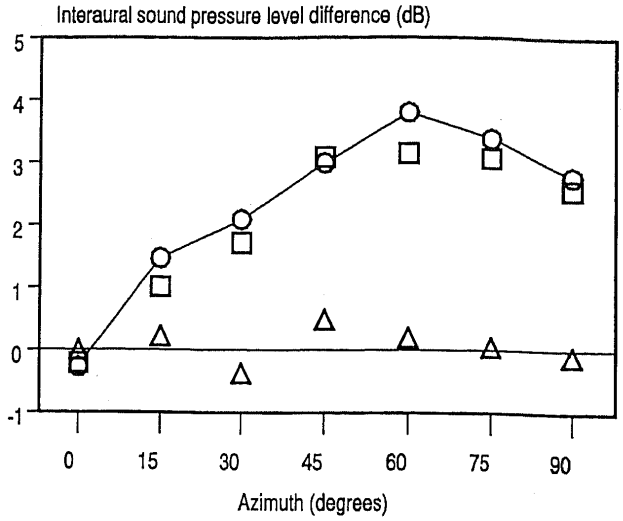

FIG. 7. Results of experiment I. Measured data (connected circles) in comparison to results obtained by simulation with head shadow and source directivity (squares) and with omnidirectional source and receivers (triangles).

head to seven different azimuth angles (0°, 15°, 30°, 45°, 60°, 75°, and 90°). In experiment II, the azimuth was kept constant at 45° and the distance from the dummy-head took the values of 0.5, 1, and 1.5 m. Figures 5 and 6 show the settings of experiments I and II, respectively.

The first 2048 taps of the impulse responses of all given situations were computed with our simulation procedure. The sampling frequency was kept constant at 10 240 Hz. All other relevant parameters are given in Table I. They were chosen to match the actual recording situation as closely as possible. The room was simulated as an empty box, although in reality it was neither empty nor exactly rectangular. The effect of furniture and equipment on the acoustical situation had to be transferred to the reflection coefficients β of the walls. Reflection coefficients were chosen to be equal for all walls such that Eyring's formula

$$\alpha = 1 - e^{\ln(10^{-6}) 4V/cT_r S} \tag{8}$$

revealed the same reverberation time T_r as the actual room. In Eq. (8), V is the volume of the room and S is the surface area of all six walls. From absorption coefficient α, we derive the reflection coefficient for all walls

$$\beta = (1-\alpha)^{1/2}. \tag{9}$$

The directivity of the loudspeaker was modeled roughly with the function SRC being

$$\mathrm{SRC}(\vartheta_s) = \begin{cases} 1, & 0 \leqslant \vartheta_s < \pi/8, \\ 0.79, & \pi/8 \leqslant \vartheta_s < 3\pi/8, \\ 0.39, & 3\pi/8 \leqslant \vartheta_s < 5\pi/8, \\ 0.28, & 5\pi/8 \leqslant \vartheta_s < 7\pi/8, \\ 0.22, & 7\pi/8 \leqslant \vartheta_s \leqslant \pi. \end{cases} \tag{10}$$

The same simulations were repeated without head-shadow and directivity effects of the loudspeaker. For every pair of simulations, the interaural pressure difference was derived directly from the sum of squares of the impulse responses.

Figure 7 shows the results of experiment I and the corresponding simulations. Clearly, the interaural pressure difference of up to 4 dB is accurately modeled with head shadow, while there is practically no difference without it. Note, that the maximum interaural intensity difference is found at an azimuth angle of about 60° and not at 90° as might have been anticipated; this is due to the lobing on the opposite side of the head, as shown in Fig. 1. Figure 8

TABLE I. Parameters used for the simulation in the experiments.

	Experiment I	Experiment II	Experiment III
Varying parameter	Azimuth of source ($\alpha = 0 \cdots \pi$)	Source to receiver distance (d=0.5–1.5 m)	Receiver directivity (head shadow or omnidirectional)
Source coordinates $R0$	$(1+\cos\alpha, 1.1+\sin\alpha, 1.7)$m	$(1+0.7d, 1.1+0.7d, 1.7)$m	(0.8,0.8,1.7)m
Preference direction of source D_s	$(-\cos\alpha, -\sin\alpha, 0)$	(−0.707,−0.707,0)	(−0.707, −0.707,0)
Receiver coordinates R	(1.00,1.10±0.093,1.70)m	(1.00, 1.10±0.093, 1.70)m	(0.3,0.5±0.093,1.7)m
Preference direction of receiver D_r	(0,±1,0)	(0,±1,0)	(0,±1,0)
Room size RL	(4.12,2.92,2.83)m	(4.12,2.92,2.83)m	(3.74,3.46,2.66)m
Reflection coefficients β (for all walls)	0.904	0.904	0.952
Reverberation time T_r	0.42 s	0.42 s	0.87 s

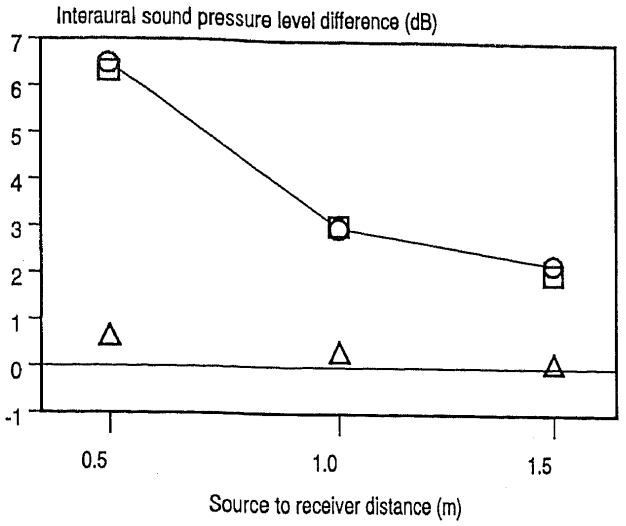

FIG. 8. Results of experiment II. Measured data (connected circles) in comparison to results obtained by simulation with head shadow and source directivity (squares) and with omnidirectional source and receivers (triangles).

shows the results of experiment II. Again, the model with head shadow and source directivity is clearly superior to the simple model.

B. Fine structure of the impulse response

The results of experiments I and II showed that the proposed simulation procedure accurately models the overall energy distribution around the head of a listener. To investigate the appropriateness of the model for individual reflections and for the frequency domain, experiment III was designed, in which another room was used and modeled.

In this room, the dummy head was placed in front of the loudspeaker. A schematic of the setting is shown in Fig. 9. The transfer function from the loudspeaker to each microphone was measured using a least-mean-squares (LMS) algorithm[22] for system identification. The reference signal required for the deconvolution process was recorded

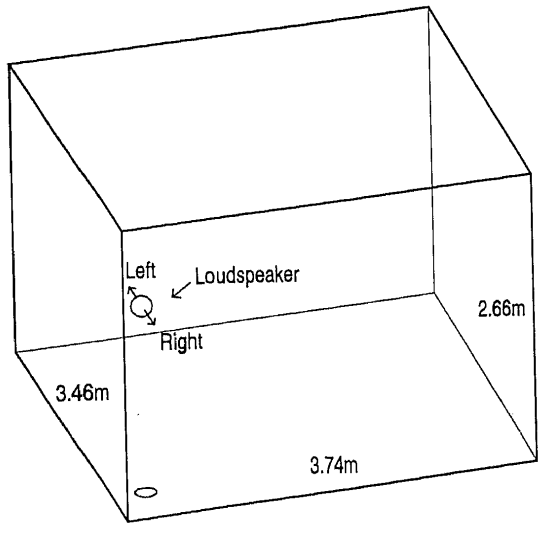

FIG. 9. Schematic of the setting for experiment III.

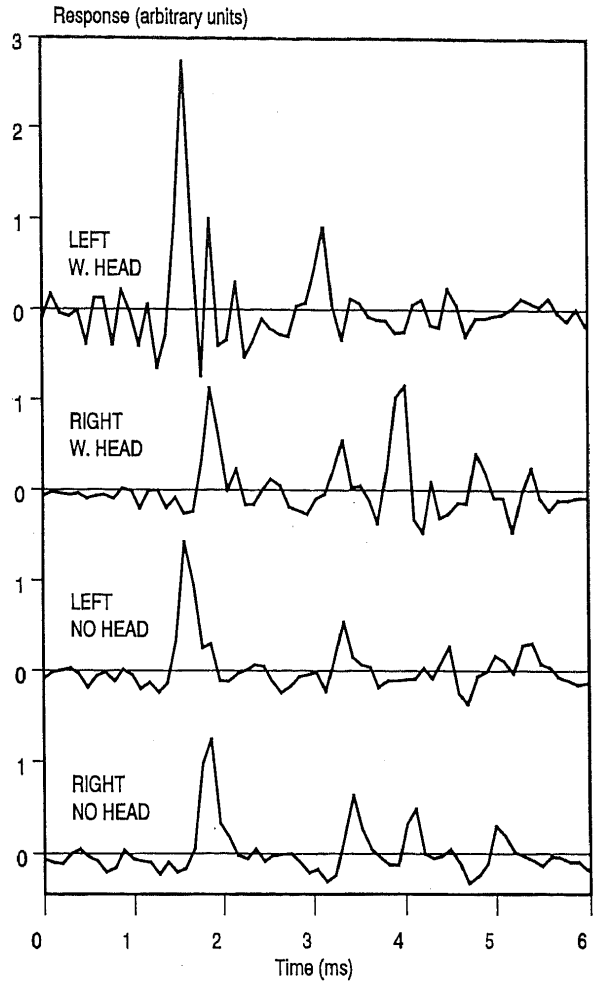

FIG. 10. Measured transfer functions from experiment III.

outdoors with the same microphones and loudspeaker, thus excluding from the results the direction-independent part of all transfer functions involved. The same measurements were also performed with the microphones alone, without dummy head. Figure 10 shows the first 6 ms of those four measured impulse responses.

As in the previous experiments, the situation was simulated with and without head shadow using the parameters shown in Table I. The directivity of the loudspeaker was modeled by a function

$$\mathrm{SRC}(\vartheta_s)=\begin{cases}1, & \vartheta_s<1.13,\\ 0, & \vartheta_s\geqslant 1.13,\end{cases} \tag{11}$$

which roughly represents the directional gain of the loudspeaker in a simple and implementationally efficient manner.

The setting was chosen such that three prominent reflections from the walls would follow the direct sound path within the first few milliseconds. The option for correct representation of arrival time was disabled during this simulation; i.e., the last factor in Eq. (4) was set to 1. This facilitates the comparison of the energy of the individual reflections. Figure 11 shows the first 62 points of the impulse response. For most reflections, a good agreement with the simulation results can be observed. Specifically, the differences between the two situations, with and with-

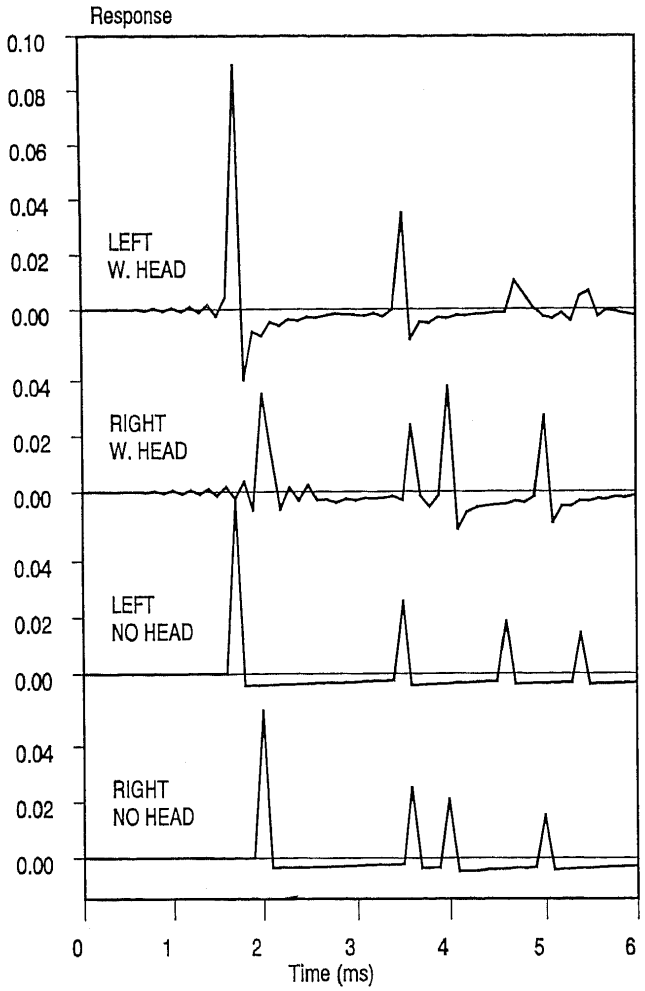

FIG. 11. Simulated transfer functions from experiment III.

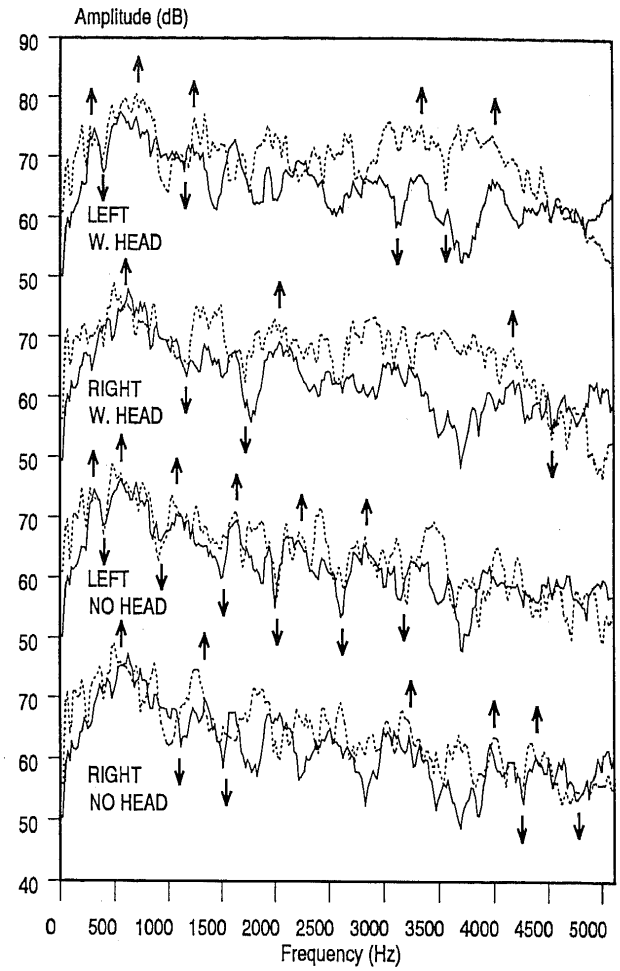

FIG. 12. Frequency-domain representation of the signals from experiment III in the simulated room (solid lines) and the real room (dotted lines).

out head shadow, are apparent and modeled correctly. The noise in the measurements in Fig. 10, which is most visible in the uppermost curve just before the first peak, is most probably caused by acoustic and electrical noise (one of the main reasons for room transfer function simulations) and its effect on the LMS algorithm for system identification.

The second issue which has been studied in experiment III is the behavior in the frequency domain. From each of the four signals recorded in the real room, 40 subsequent blocks of 512 samples were multiplied by a Hamming window and Fourier transformed. The magnitudes of all blocks belonging to one signal were averaged resulting in the four spectra (dotted lines) depicted in Fig. 12. They represent an averaged magnitude spectrum for the measured signals.

To produce comparable spectra for the simulated situation, the reverberation-free reference signal was filtered by the simulated room impulse responses of 2048 taps. The resulting signals were then processed in the same manner as the measured signals, leading to the spectra drawn as solid lines in Fig. 12.

The discussion of the different signals is more difficult in the frequency domain than in the time domain. However, a superficial examination already reveals reasonable agreement of the coarse shapes of the spectra for real and simulated rooms. Only above 2500 Hz in the spectra with the head shadow included, significant differences of more than about 5 dB occur. This is probably due to the missing pinnae in our model, which become more important for higher frequencies, as has been noted already in Sec. I.

A closer look at the spectra reveals typical sequences of peaks and valleys which can be identified both for the simulations and the real recordings. In Fig. 12, an arbitrary selection of these features have been marked by arrows. Although the spectra cannot be expected to match identically, distinct similarities can be seen even in their fine structure.

VI. SUMMARY

In this paper, a method was presented to simulate the transfer functions in a room. The main feature, the modeling of the head shadow, has been shown to be useful for modeling listeners in reverberant rooms. In Sec. V, it was shown that the simple procedure is able to reasonably model actual situations. This holds not only for interaural sound-pressure differences, but also for the fine structure of the impulse response in the time, and to some extent also in

the frequency domain, even if source and receiver are positioned close to the walls, as in experiment III.

ACKNOWLEDGMENTS

This work was supported by the Swiss National Research Foundation, Grant No. 4018-10864 and the Digital Speech Processing Department of Ascom Tech Ltd. We would like to thank Dr. Wai Kong Lai and Dr. C. W. Palaskas for helpful comments and their assistance in preparing this paper.

[1] J. P. Vian, "Different Computer modeling methods—their merits and their applications," in *Proceedings of the 12th International Congress on Acoustics*, Toronto, E4-10 (1986).
[2] J. Martin and J. P. Vian, "Binaural sound simulation of concert halls by a beam tracing method," in *Proceedings of the 13th International Congress on Acoustics*, Belgrade (1989), pp. 253–256.
[3] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," J. Acoust. Soc. Am. **65**, 943–950 (1979).
[4] M. Feder, A. V. Oppenheim, and E. Weinstein, "Maximum likelihood noise cancellation using the EM algorithm," IEEE Trans. Acoust. Speech Signal Process. **37**, 204–216 (1989).
[5] P. M. Peterson, "Simulating the response of multiple microphone to a single acoustic source in a reverberant room," J. Acoust. Soc. Am. **76**, 1527–1529 (1986).
[6] J. Blauert *Räumliches Hören* (Hirzel Verlag, Stuttgart, 1974), p. 58.
[7] G. F. Kuhn, "Model for the interaural time difference in the azimuthal plane," J. Acoust. Soc. Am. **62**, 157–167 (1977).
[8] E. M. von Hornbostel and M. Wertheimer, *Über die Wahrnehmung der Schallrichtung* (Sitzungsber. Akad. Wiss., Berlin, 1920), pp. 388–396.
[9] P. M. Peterson, N. Durlach, W. M. Rabinowitz, and P. M. Zurek, "Multimicrophone adaptive beamforming for interference reduction in hearing aids," J. Rehabil. Res. Dev. **24** (4), 103–110 (1987).
[10] L. Schwarz, "Zur Theorie der Beugung einer ebenen Schallwelle an der Kugel," Akust. Z. **8**, 91–117 (1943).
[11] P. M. Morse, *Vibration and Sound* (American Institute of Physics, New York, 1983), 2nd (paperback) printing, Chap. 27, pp. 311–326.
[12] K. Genuit, "Ein Modell zur Beschreibung von Aussenohrübertragungseigenschaften," Ph.D. thesis at the Rheinisch–Westfälische Technische Hochschule Aachen, Germany (1984).
[13] D. H. Cooper and J. L. Bauck, "Prospects for Transaural Recording," J. Audio Eng. Soc. **37**, 3–19 (1989).
[14] E. A. G. Shaw, "Transformation of sound pressure level from the free field to the eardrum in the horizontal plane," J. Acoust. Soc. Am. **56**, 1848–1861 (1974).
[15] S. Mehrgardt and V. Mellert, "Transformation characteristics of the external human ear," J. Acoust. Soc. Am. **61**, 1567–1576 (1977).
[16] J. Blauert, *Räumliches Hören, Nachschrift* (Hirzel Verlag, Stuttgart, 1985), pp. 22–23.
[17] H. K. Dunn and D. W. Farnsworth, "Exploration of pressure field around the human head during speech," J. Acoust. Soc. Am. **10**, 184–199 (1939).
[18] J. L. Flanagan, "Analog measurement of sound radiation from the mouth," J. Acoust. Soc. Am. **32**, 1613–1620 (1960).
[19] M. Kompis and N. Dillier, "Noise reduction for hearing aids: Evaluation of the adaptive beamformer approach," Proc. Ann. Int. Conf. IEEE Eng. Med. Biol. Soc. **13**, 1887–1888 (1991).
[20] D. H. Cooper and J. L. Bauck, "Corrections to L. Schwarz, 'On the theory of diffraction of a plane soundwave around a sphere' ['Zur Theorie der Beugung einer ebenen Schallwelle an der Kugel,' Akust. Z. **8**, 91–117 (1943)]," J. Acoust. Soc. Am. **80**, 1793–1802 (1986).
[21] H. Kuttruff, *Room Acoustics* (Applied Science, Ltd., London, 1979), 2nd ed. Chap. II, pp. 20–43.
[22] B. Widrow *et al.*, "Adaptive Noise canceling: Principles and applications," Proc. IEEE **63**, 1692–1716 (1975).