Dependence of Subject and Measurement Position in Binaural Signal Acquisition*

V. RALPH ALGAZI, CARLOS AVENDANO, AND DENNIS THOMPSON, *AES Student Member*

Center for Image Processing and Integrated Computing, University of California, Davis, CA 95616, USA

A detailed evaluation of the dependence of the head-related transfer functions (HRTFs) on the position of the recording microphone along the ear canal for blocked and unblocked conditions is presented. Measurements at 250 source locations for a KEMAR mannequin and at 125 locations for two human subjects were performed. Conditions under which the magnitude transfer function from the HRTF acquisition position to the eardrum is essentially independent of the source location are shown quantitatively. These results are in agreement with and provide an extensive validation of other studies. The cause and magnitude of measurement errors are also analyzed in detail, and it is shown that these errors are subject dependent.

0 INTRODUCTION

The binaural synthesis of a three-dimensional auditory space requires the replication by a sound system of the signals (or of important features of these signals) that would reach the ears of a listener under free-field conditions.

A key problem that arises in the design of a binaural synthesis system is the measurement of the head-related transfer function (HRTF). For such a measurement it is quite desirable not to have the source-location-dependent results be critically sensitive to the position of the point within (or outside) the ear canal where the HRTF is measured.

Measurements near the eardrum will account for all individual localization characteristics of the listener, including the ear canal resonance. However, such measurements are somewhat intrusive and may not be necessary. It has been reported by many authors [1]–[3], and it will be verified and complemented in this engineering report, that all localization information can be obtained at a number of points within the ear canal (and possibly a few millimeters outside).

In this engineering report we present extensive results and an analysis of HRTF data collected at different points along the ear canal, with blocked and unblocked meatus. The HRTFs of a KEMAR dummy and of two human subjects were measured for 250 and 125 source locations, respectively. These results were analyzed to assess the errors or uncertainty of the measurements for all sound locations. We then evaluated the possible location-dependent contributions to the ear canal transmission due to the source position and found that up to 12 kHz this contribution is small. In addition we found that all location-dependent characteristics of the HRTF may be acquired with a blocked meatus at any of a range of positions along the ear canal.

In binaural synthesis a compensation that accounts for the response and loading of the acoustic transducer has to be applied. This compensation will depend on the measurement point and on the acoustic transducer used (most commonly headphones), as well as on the individual. If the HRTF is measured at any position with an open ear canal (including close to the eardrum), the headphone-to-measurement transmission has to be compensated. For blocked ear canal measurements, both the headphone-to-open and the headphone-to-blocked transmission at the position of measurement may be needed to devise the correct compensation. The results in this investigation suggest that the scheme for binaural synthesis through headphones based on blocked meatus HRTF measurements and the sound transmission model of Møller [4] is adequate to convey all localization cues available in free-field hearing.

1 SOUND TRANSMISSION MODEL

We will refer in this work to the sound transmission model proposed in Møller [4]. The model consists of an acoustic transmission line where the sound pressure and the acoustic impedance are measured at different points

* Manuscript received 1998 April 27; revised 1999 October 1.

along the line (Fig. 1). The model is applicable to free-field listening as well as to binaural synthesis through headphones.[1] The two situations are illustrated in Fig. 1. Measurement points shown in the figure correspond physically to the positions of the sound source (P_2, P_5) to the entrance of the ear canal (P_3, P_6) and to the eardrum (P_4, P_7). The acoustic impedances associated with these locations are included in the model. These are the source radiation impedance (free field or headphone) Z_{rad}, the ear canal entrance impedance $Z_{ear\ canal}$, and the eardrum impedance $Z_{eardrum}$. Note that this model also applies to measurements at other positions.

Binaural signals at the eardrum can be synthesized from blocked meatus HRTF measurements along the ear canal if

- the transmission can be separated into its location-dependent and location-independent parts, where the location-dependent part can be measured with a blocked meatus
- blocked ear canal HRTF measurements with proper compensation can reproduce the correct binaural signals at the eardrum.

In order to verify these assumptions, we carried out several experiments.

[1] Synthesis through loudspeakers is not considered here but can also be modeled in a similar way.

2 EXPERIMENTS

A set of experiments was performed with a dummy head and with two human subjects. The HRTFs of one ear were measured for a large number of source locations distributed mainly in the ipsilateral hemisphere.

2.1 Experimental Setup

Experimental conditions were designed to keep measurement errors as low as possible. For the open ear canal measurements, two points along the ear canal were recorded simultaneously, thus eliminating any physical reset of experimental conditions. For the blocked ear canal HRTF measurements, the experimental setup was minimally disturbed between runs.

Five sources (Bost Acoustimass Cube loudspeakers) where located on a motor-driven computer-controlled hoop with a radius of 1.07 m, and driven by a Symetrix A-220 amplifier. The axis of rotation of that hoop was a line passing through the ear canals of the subject. We used an interaural-polar coordinate system, where azimuth θ is the angle between a ray to the source and the median plane. Thus $-90° \leq \theta \leq 90°$, where $\theta = -90°$ corresponds to the left side of the subject and $\theta = 90°$ to the right side. Elevation is the angle described by the plane of the hoop with the horizontal, and its range is $-90° \leq \phi < 270°$, where $\phi = -90°$ is below the subject, $\phi = 0°$ is when the hoop is on the horizontal plane (frontal hemisphere), and $\phi = 90°$ is directly

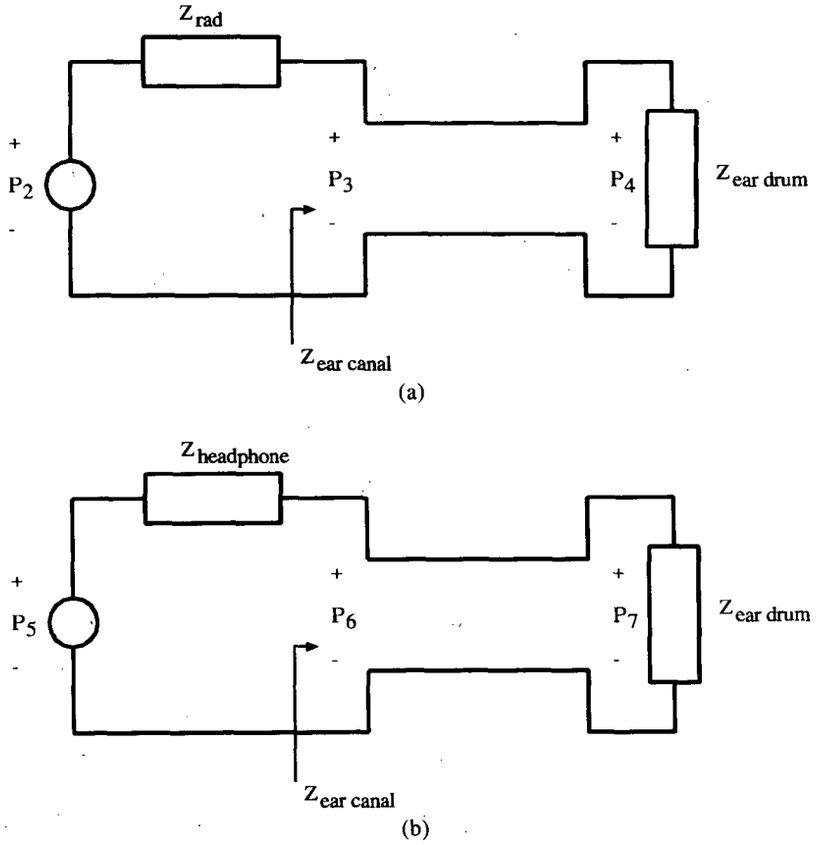


Fig. 1. Sound transmission model. (a) Free-field condition. (b) Headphone condition (adapted from [4, figs. 3 and 6]. P_2, P_5—source pressures for blocked meatus condition (free-field and headphone listening respectively); P_3, P_6—pressures at entrance to open ear canal; P_4, P_7 pressures measured at eardrum.

above of the subject's head. Due to mechanical limitations, elevations directly below the subject could not be measured and elevation angles were confined to $-62° \leq \phi \leq 245°$.

A block diagram of the system is shown in Fig. 2. Head-related impulse responses (HRIRs) were measured in the time domain with a Crystal River Snapshot system, using a Golay code of length 2048 at a 44.1-kHz sampling rate. The recordings were made with Etymotic ER-7C probe microphones (with their battery-powered preamplifiers), and the interface to the PC was through a Momentum 56 sound card. The response of the system components is shown in Fig. 3, where the contributions of the loudspeakers and microphones can be observed separately. The response is mainly determined by the loudspeaker, with the microphone contributing to the high-frequency rolloff. The overall SNR at the output (HRIR) is better than 40 dB.

MATLAB was used to control the position of the hoop, the firing of the loudspeakers, and the generation and recording of test signals. The recording position was at the tip of a plastic probe tube attached to the microphone capsules, which was set at the desired measurement point. The recordings were made in a quiet room with dimensions 6 by 6 by 3.5 m. Reflections from the walls, ceiling, and floor were eliminated by windowing the measured impulse responses at 4.5 ms.

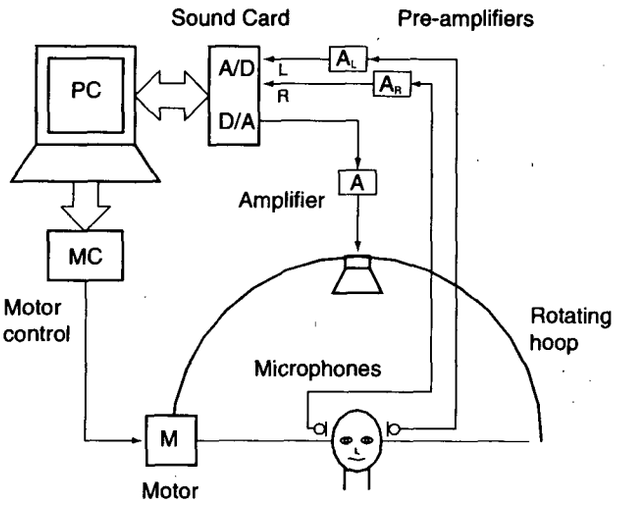


Fig. 2. Apparatus for HRTF measurement. HRIRs are measured using a modified Crystal River Snapshot system running in MATLAB. Loudspeakers are placed along a rotating hoop controlled from PC.

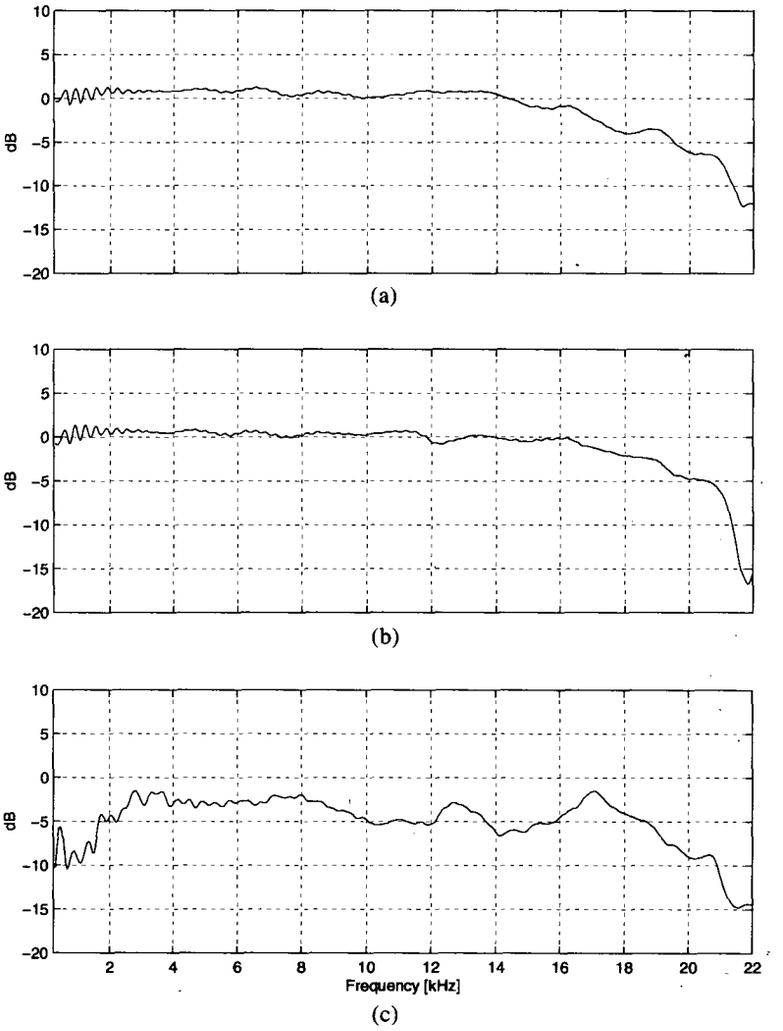


Fig. 3. Frequency responses. (a) Left microphone. (b) Right microphone. (c) Loudspeaker (referenced to an Earthworks M30 calibration microphone).

2.2 KEMAR Head Measurements

The KEMAR (Knowles Electronics Manikin for Acoustic Research [5]) dummy head measurements were used as a baseline so that movement errors encountered in human subject data would not appear. The KEMAR data were collected at five different azimuths (θ = $-80°$, $-45°$, $-20°$, 0°, and 40°) and 50 elevations, from -47.8 to 227.8° in steps of 5.6°, for a total of 250 locations. Two sets of data were collected, one set with an open ear canal and another with a blocked ear canal. For the blocked ear canal condition we measured HRTFs at five positions: 4 mm outside the ear canal, at the entrance of the ear canal (0 mm), and 4, 11, and 22 mm inside the ear canal. The "eardrum" for the KEMAR is approximately at 22 mm from the entrance of the ear canal. For the open ear canal condition we measured HRTFs at one point approximately 4 mm outside the ear canal, at the entrance (0 mm), and at the eardrum (22 mm).

For measurements at the "eardrum" of the KEMAR, the microphone capsule was placed inside the head, and microphone probes were passed through the center of a silicon rubber ear plug and positioned so that the probe tip was flush with the surface of the plug. The ear plug was then positioned at a depth of 22 mm within the ear canal. The custom Zwislocki coupler in the KEMAR was used to simulate eardrum impedance and is discussed in [5]. Measurements at other positions along the ear canal were made by just pushing the ear plug to the desired depth.

For measurements made outside the ear canal, an extremely fine strand of copper wire was wrapped around the probe tube, from its base to within 10 mm of its tip. This additional rigidity allowed the probe tube to be bent into the desired position in space at the entrance to the ear canal. The microphone capsule was mounted on the exterior of KEMAR's head, just under the jaw line.

2.3 Human Subject Measurements

Data for two human subjects were also collected, but less densely than for the KEMAR, so as to minimize the data collection time, reducing fatigue and the associated risk of involuntary motion. Both subjects had similar pinna dimensions as the KEMAR (pinna length 6.5 mm), and their ear canal depths were 27 mm for CA and 19 mm for DT, compared to 22 mm for the KEMAR.

Data were collected for five different azimuths (θ = $-80°$, $-45°$, $-20°$, 0°, and 40°) and 25 elevations, from -47.8 to 227.8° in steps of 11.2°, for a total of 125 locations. Two sets of data were collected for each subject. In the first set recordings of the sound pressures near the eardrum and at the entrance to the open ear canal were made simultaneously for all source locations. The second set of measurements was made at the blocked entrance to the ear canal.

Subject motion was reduced with the use of a chin rest. The interference of the chin rest with the sound field was neglected since our interest was in comparative measurements of HRTFs under alternate experimental conditions.

The measurements near the eardrum were performed by carefully inserting the flexible plastic probe tube into the ear canal until the subject felt it barely touching the eardrum. The tube was then pulled out by a couple of millimeters. (This procedure was suggested by a clinical audiologist at the UC Davis Medical Center.) All the measurements were performed as for the KEMAR.

3 Analysis and Results

3.1 Methods

In the analysis and presentation of the results we will use the acoustic transmission line model outlined in the preceding. Let $\psi_k = (\phi, \theta)_k$ denote the source location, where $k = 1, 2, \ldots, K$ and K is the total number of locations. Referring to Møller, Fig. 1(a) shows the model for free-field listening and Fig. 1(b) that for headphone listening. For free-field listening we let $p_4(t, \psi_k)$ be the sound pressure at the eardrum [referenced to a free-field location at the center of the head $p_1(t, \psi_k)$], $p_3(t, \psi_k)$ the pressure at a point along the open ear canal, and $p_2(t, \psi_k)$ the pressure at the blocked point along the ear canal. Pressure $p_2(t, \psi_k)$, which does not exist in normal listening conditions, models the blocked meatus measurement. Z_{rad} is the radiation impedance seen from the entrance to the ear canal for free-field listening. Fourier transforms are denoted by $P_4(\omega, \psi_k)$, $P_3(\omega, \psi_k)$, and $P_2(\omega, \psi_k)$, respectively, where ω is the frequency variable. Similar pressures for headphone listening are denoted by $P_7(\omega, \psi_k)$, $P_6(\omega, \psi_k)$, and $P_5(\omega, \psi_k)$. $Z_{headphone}$ is now the impedance seen from the entrance the ear canal for headphone listening [4].

For our experiments the sound pressures in Møller's model correspond to our HRIR measurements (that is, the sound source is effectively an impulse). We illustrate some of our measurements in Fig. 4. Here we show HRTFs measured at the eardrum [$P_4(\omega, \psi_k)$] and at the blocked meatus [$P_2(\omega, \psi_k)$] for all elevations, and at an azimuth of θ = 45°. The magnitude frequency response is given as a function of elevation, with the gain (in decibels) represented by the gray-scale level. Note that we limited our analysis to the range from 250 Hz to 15 kHz.

Qualitatively both HRTFs are quite similar, except for location-independent effects indicated by horizontal bands across all elevations. Such location-independent differences are due to the ear canal resonance and acoustic impedance differences caused by differences in the anatomy of the subjects.

Quantitatively our objective is to determine whether all three pressures have the same directional information when $p_2(t, \psi_k)$ and $p_3(t, \psi_k)$ are measured within, at the entrance, and possibly a few millimeters outside the ear canal. In such a case pressure ratios $P_4(\omega, \psi_k)/P_3(\omega, \psi_k)$, $P_4(\omega, \psi_k)/P_2(\omega, \psi_k)$, and $P_3(\omega, \psi_k)/P_2(\omega, \psi_k)$ are independent of direction, and represent the respective transmissions between points along the ear canal.

Alternatively we can estimate the direction-independent transmission between points along the ear

canal, and compare the response at the eardrum with measurements carried out at other points.

3.2 Computation of Ear Canal Tranmission

A first approach would be to compute the ratio of the Fourier transforms of the measured sound pressures. Throughout this engineering report we will consider only the magnitude of the Fourier transforms, which we denote by $P(\omega, \psi_k)$ for simplicity. The magnitude of the transmission can be computed as

$$T_{i,j}(\omega) = \frac{P_i(\omega, \psi_k)}{P_j(\omega, \psi_k)}, \qquad i, j = 1, 2, 3, 4, \qquad i \neq j. \tag{1}$$

Note that we wrote $T_{i,j}(\omega)$ as being independent of ψ_k, that is, location independent. However, low-magnitude values in the HRTF, such as pinna notches, or the frequency characteristics of the measuring equipment at very high or very low frequencies, result in unreliable pointwise estimates of the ear canal transmission.

Thus we first evaluate a location-independent ear canal transmission $M_{i,j}$. Using this transmission we can then estimate the sound pressure $P_i(\omega, \psi_k)$ from $P_j(\omega, \psi_k)$ as

$$\hat{P}_i(\omega, \psi_k) = M_{i,j}(\omega) P_j(\omega, \psi_k). \tag{2}$$

We then evaluate the error between the estimate $\hat{P}_i(\omega, \psi_k)$ and the target pressure $P_i(\omega, \psi_k)$. If we write the error as

$$E_{i,j}(\omega, \psi_k) = P_i(\omega, \psi_k) - \hat{P}_i(\omega, \psi_k) \tag{3}$$

we can formulate a least-squares minimization problem to estimate $M_{i,j}(\omega)$ at each frequency. This results in a set of N independent equations, where N is the number of frequency bands,

$$E_{i,j}(\omega_n, \psi_k) = P_i(\omega_n, \psi_k) - M_{i,j}(\omega_n) P_j(\omega_n, \psi_k), \qquad n = 1, 2, \ldots, N. \tag{4}$$

Minimizing the expected value, or average, with respect to the source location (index k) of the squared error [Eq. (4)] we get

$$\min_{M_{i,j}(\omega_n)} \frac{1}{K} \sum_{k=1}^{K} E(\omega_n, \psi_k)^2 = \min_{M_{i,j}(\omega_n)} \frac{1}{K} \sum_{k=1}^{K} [P_i(\omega_n, \psi_k) - \hat{P}_i(\omega_n, \psi_k)]^2 \tag{5}$$

which has for solution

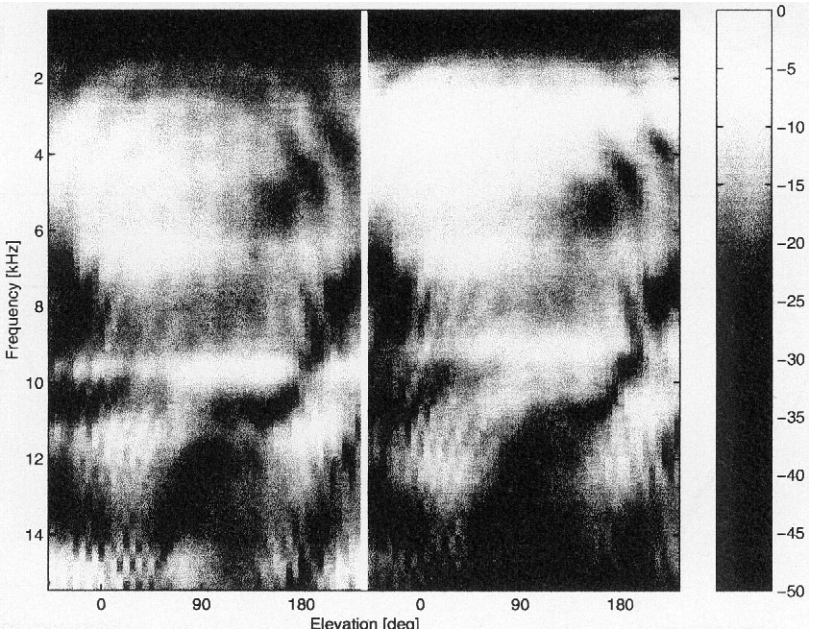
$$M_{i,j}(\omega_n) = \frac{C_{i,j}}{c_{j,j}} \tag{6}$$

where

$$c_{i,j} = \frac{1}{K} \sum_{k=1}^{K} P_i(\omega_n, \psi_k) P_j(\omega_n, \psi_k)$$

$$c_{j,j} = \frac{1}{K} \sum_{k=1}^{K} P_j(\omega_n, \psi_k)^2.$$

The previous analysis is essentially a linear regression on the data distributions for each frequency band. A similar observation was made by Middlebrooks and Makous [2], fig. 4] for only two frequency bands and in a



(a) (b)

Fig. 4. HRTF magnitude for $\theta = -45°$. (a) Blocked entrance to ear canal. (b) Eardrum. Images show magnitude in dB (gray scale) as a function of frequency and elevation.

decibel scale.

By such a method we have estimated the location-independent transmission that provides the best match between HRTFs measured at several positions in the ear canal. By analyzing the goodness of the fit, or the residual errors as a function of frequency and location, we resolve whether location-dependent effects occur, and we also gain insight into the causes of inaccuracies in the HRTF measurements. We first discuss global or aggregate results.

3.3 Global Results

A global measure of directional dependence that has been used in the past is the "directional spread." Such measure is the standard deviation of $T_{i,j}(\omega)$ [computed from Eq. (1)] in decibels as a function of frequency [2], [3], [6]. Note that such a directional spread is a characterization that combines measurement errors due to several causes with possible location-dependent effects. As we shall show, location-dependent contributions are within the limits of measurement accuracy.

We thus compute the directional spread for our data, with the slight modification that $M_{i,j}(\omega)$ is used as the mean, and the sum of squared errors in decibels between $M_{i,j}(\omega)$ and $T_{i,j}(\omega)$ estimates the variance.

The results for the KEMAR are presented in Fig. 5 for the four different positions along the blocked ear canal. These results indicate that the spread is small—less than 1 dB up to 12 kHz for all cases, except for the measurement at a point 4 mm outside the ear canal.

Directional spreads for our two human subjects are shown in Fig. 6 for measurements taken at the entrance of the ear canal under either blocked or unblocked conditions. Recall that open ear canal measurements close to the eardrum were taken simultaneously, whereas blocked ear canal measurements require a new experimental setup. By comparing the results of Fig. 6 side by side, we note that results of different experimental setups may differ by as much as 4 or 5 dB. By contrast, errors are much smaller for simultaneous measurements. Thus errors for human subjects that do occur are likely due to motion, during or between experiments. We will analyze such errors in more detail in the next subsection.

In general there is a trend for the error to increase with frequency, but the 1- or 2-dB deviations in KEMAR's ear canal transmission indicate very small errors or location contributions for frequencies below 15 kHz.

However, the "directional" spread provides a simple summary of several effects such as error magnitude in the computation of the ear canal transmission, but it does not give any insight into the cause of these errors. Additional information can be obtained by examination of the residual error in the computation of $M_{i,j}$.

We first demonstrate that large errors are correlated with frequency ranges that exhibit low response amplitudes. For open meatus measurements we show in Fig. 7 the normalized residual error as a function of frequency for one human subject (CA) and for the KEMAR, together with the average power at each frequency at the point of measurement. Normalization is performed by dividing the error [Eq. (3)] by the average power of the eardrum pressure P_i at each frequency. We observe that large global errors occur at frequencies where either sound pressure P_i or pressure P_j exhibits low energy.

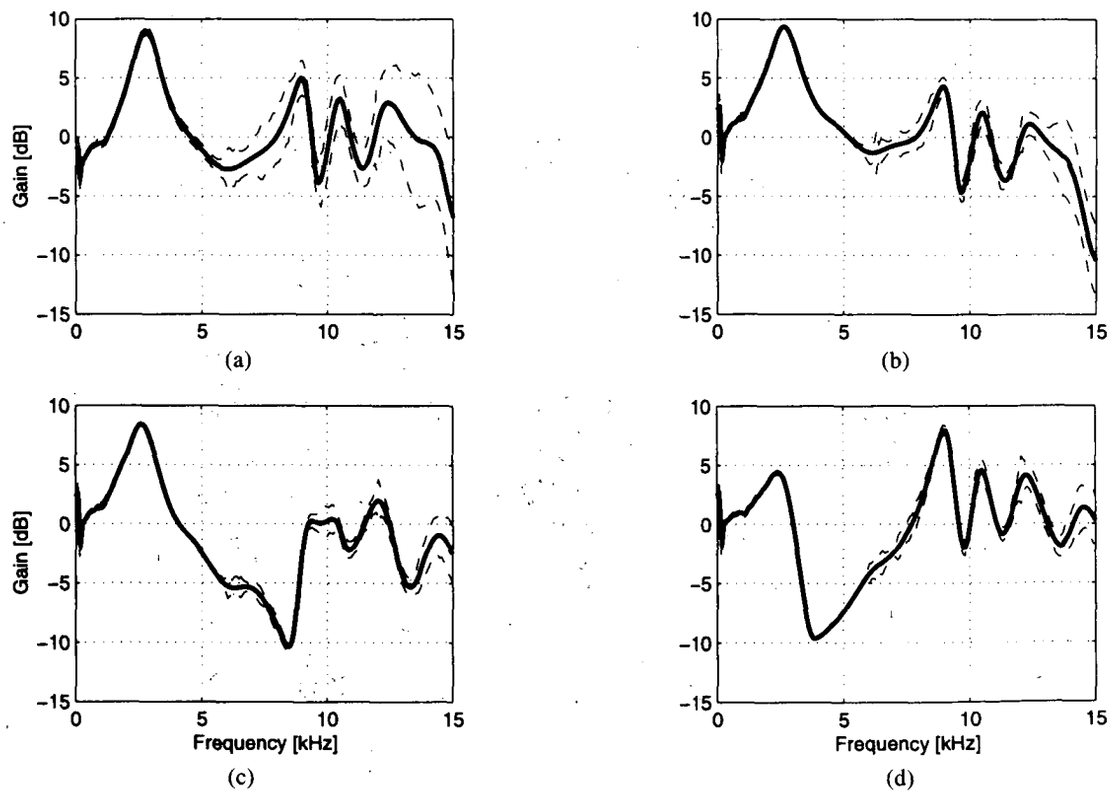


Fig. 5. Sound transmission within ear canal of KEMAR. Blocked meatus condition. (a) 4 mm outside. (b) Entrance (0 mm). (c) 4 mm inside. (d) 11 mm. —— "optimal" transmission; --- ±1 standard deviation.

Note that these errors are subject dependent. We also analyzed these subject-dependent global errors in more detail for the entire set of locations.

Our results observed for global behavior are in agreement with those obtained previously by other researchers [2], [3], [6], [7].

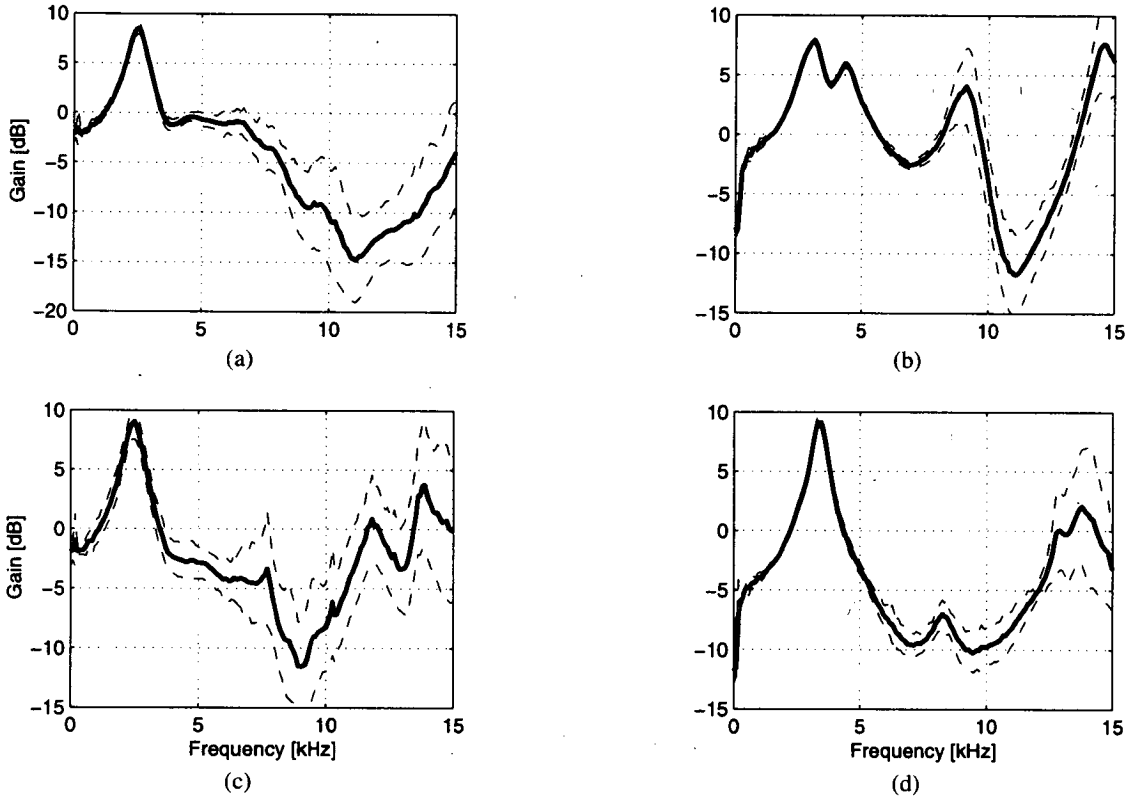


Fig. 6. Sound transmission between eardrum and entrance to ear canal for two subjects. (a) Blocked ear canal, subject CA. (b) Open ear canal, subject CA. (c) Blocked ear canal, subject DT. (d) Open ear canal, subject DT. —— mean for all source locations; --- ±1 standard deviation of log magnitudes.

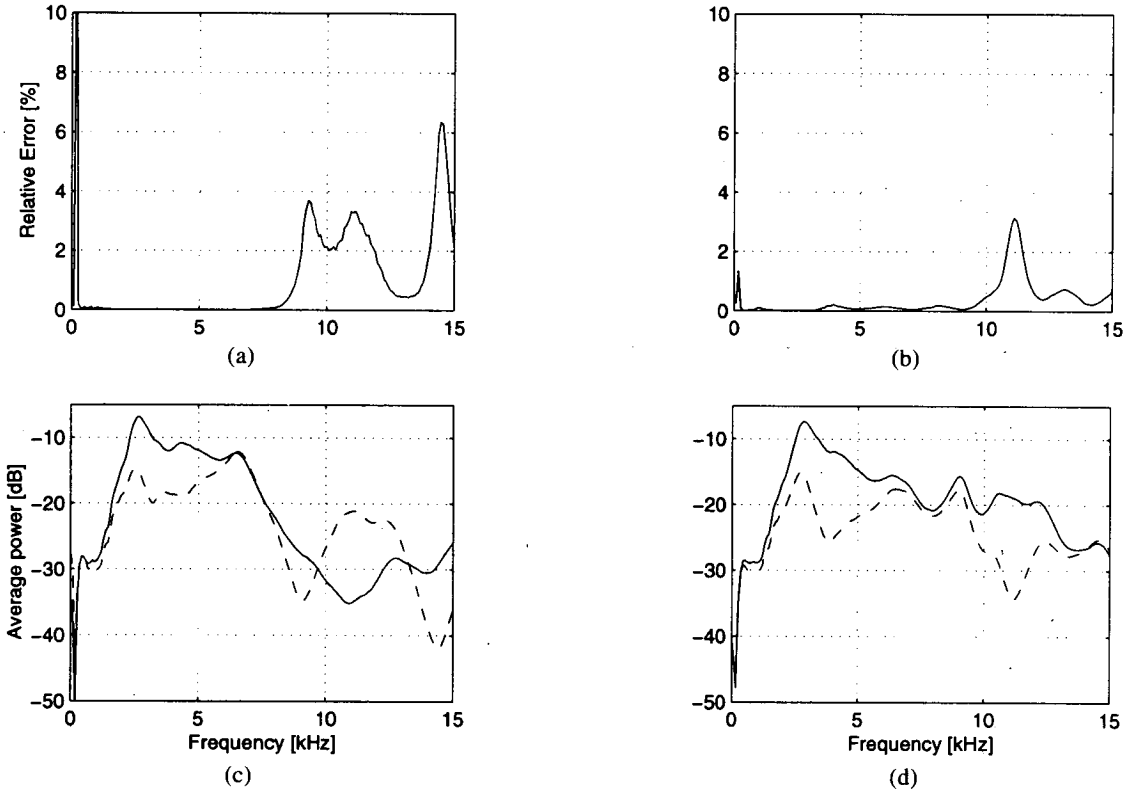


Fig. 7. Normalized residual errors for open meatus data. (a) Subject CA. (b) KEMAR. Corresponding averaged power of measured pressures. (c) Subject CA. (d) KEMAR. —— eardrum; --- entrance.

3.4 Detailed Error Analsis of KEMAR Data in the Spectral Domain

The location-independent transmission $M_{i,j}(\omega)$ is used to determine, in the frequency domain, the residual errors at various locations. The spectral distortion in the estimate $\hat{P}_i(\omega, \psi_k)$, that is, the error between the ith sound pressure and its estimate ased on $M_{i,j}(\omega)$, is plotted in Fig. 8 together with the sound pressure $P_i(\omega, \psi_k)$. We chose two locations that illustrate the detailed behavior of such spectral errors. These are the HRTFs recorded for azimuth $\theta = -45°$ and two elevations, $\phi = -90°$ and $\phi = 0°$.

As noted previously, the errors increase with frequency and with decreasing energy level of $P_i(\omega, \psi_k)$, but they are specially large around pinna notches, where the exact response is most uncertain. An overall view of the frequency-domain errors for a dense set of elevations at a single azimuth is shown in Fig. 9. This relative error is computed as

$$Er_{i,j}(\omega, \psi_k) = \frac{20 \log_{10}[P_i(\omega, \psi_k) + D_{i,j}(\omega, \psi_k)] - 20 \log_{10}[P_i(\omega, \psi_k)]}{|20 \log_{10}[P_i(\omega, \psi_k)]|} \quad (7)$$

where

$$D_{i,j}(\omega, \psi_k) = |P_i(\omega, \psi_k) - \hat{P}_i(\omega, \psi_k)| . \quad (8)$$

We note again that large relative errors, up to 25% at a limited set of frequencies, are concentrated above 10 kHz and near spectral notches. Thus we attribute global errors, discussed in the previous section, to the limited accuracy in the positioning of the sound source that is most critical near spectral notches, and to noise and uncertainty in the measurement. The largest errors were found at spectral notches, whereas high-frequency errors are in the 5–10% range. As a whole, errors are very small, and are seldom more than 5%.

From this set of detailed measurements we have confirmed the hypothesis that the measured blocked meatus HRTF at the entrance to the ear canal can be transformed into the HRTF of the KEMAR at the eardrum by a fixed, that is, location-independent, ear canal transmission. We have also shown that measurements at a number of different positions along the ear canal are also usable. Furthermore we have elucidated the major sources of errors in such measurements. These are concentrated at low-energy frequency bands or at the specific frequencies where spectral notches occur. Before discussing the practical importance of these, we also analyze the measurements for human subjects.

3.5 Detailed Error Analysis of Human Subject Data in the Spectral Domain

In the examination of the data collected for the human subjects our goal is twofold. First we wish to complete the experimental verification of the effect of location on the HRTF measurement for human subjects. Second we would like to extend our observations or analysis of

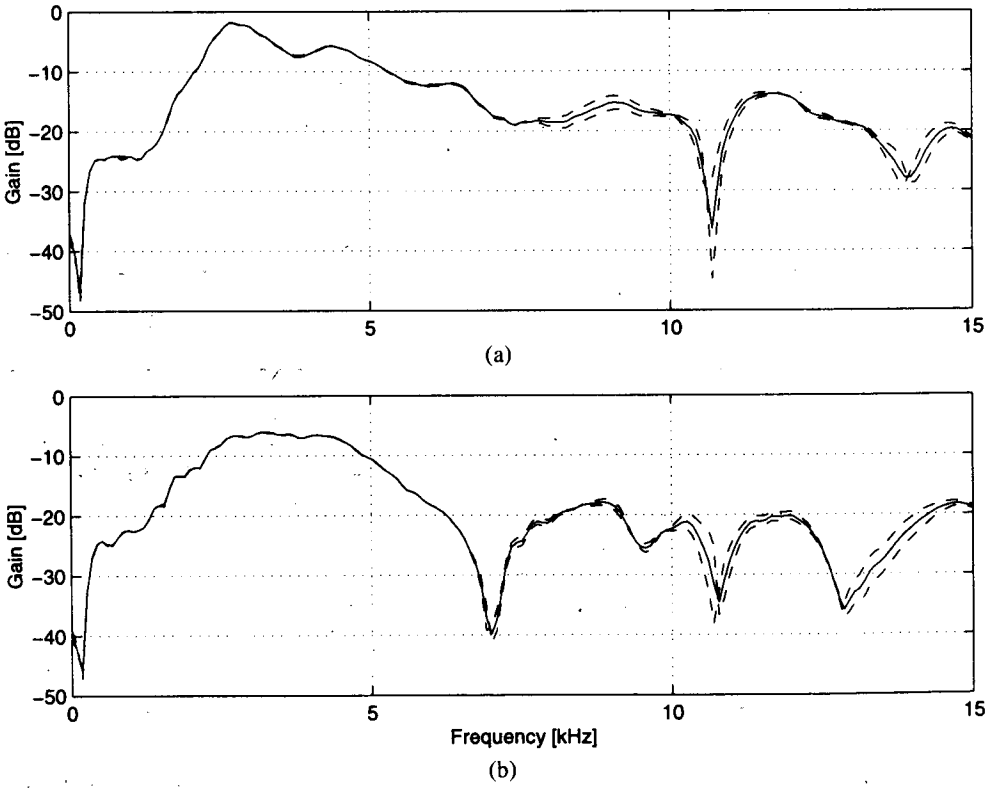


Fig. 8. Sound pressure $P_i(\omega, \psi_k)$ in dB for selected representative locations. --- ± error in dB.

errors to understand the effect and magnitude of other unavoidable uncertainties that occur in human subject measurements. The results shown in Fig. 6 indicate that global errors could be as high as 5 dB at some frequencies below 10 kHz, and somewhat larger above this frequency.

We are mainly interested in the effect of unavoidable small movements by subjects during an experiment or from one experimental setup to the next. To minimize the effects of movement, the open ear canal measurements were performed simultaneously at the two points, whereas blocked meatus measurements had to be taken after an experimental reset. This would explain the larger deviations observed in Fig. 6(a) and (c), as compared to Fig. 6(b) and (d).

Even for the open entrance-to-eardrum measurements that were made simultaneously at two positions, head motion during the procedure may cause the two probe tubes to move within the ear canal, changing the relative position between the two microphone probes, and that would be indicated by slightly different ear canal transmissions for different source locations.

As for the KEMAR, we computed the relative spectral error for both subjects at all locations (Fig. 10). The magnitude of the errors is in the same range as for the KEMAR, but some regions of the frequency-elevation plane exhibit anomalies. We examined some of these regions (e.g., for example, low elevations at 9 kHz for subject CA), where the error did not correlate with the presence of localized spectral notches. Our analysis revealed two main causes for these errors: First, monitoring the HRIRs for each elevation, we detected an onset discontinuity that is due to motion of the subject at around $\phi = 0°$. The computation of $M_{i,j}$, which is carried out by a least-squares fit over all elevations, will show a larger residual error for $\phi \leq 0°$ since $M_{i,j}$ was not constant for the entire experimental session. This effect is seen as a vertical discontinuity in the data at around $\phi = 0°$ [Fig. 10(a)].

Such an error is compounded in the low-energy frequency bands. Referring again to Fig. 7(a) for subject CA, we observe three peaks in the normalized residual error, at around 9, 11, and 14 kHz, that correspond to low-energy regions in either P_i or P_j. Data for subject DT (not shown) have a peak in the residual error at 13.5 kHz, which also corresponds to an average low-energy frequency band. These subject-dependent bands of errors can be observed readily in Fig. 10. Another error band at 170 Hz is apparent for both subjects, and is caused by a spectral notch in the measuring equipment (see Fig. 3).

The results of the global analysis of Section 3.3 can now be elaborated on. For example, in Fig. 6(b) the global analysis indicates that the largest absolute deviations are in the 11-kHz region, whereas the largest relative spectral errors occur at 9 kHz. As with the KEMAR, errors indicated by departure from a fixed location-independent transmission occur in regions of low energy, irrespective of the frequency band. The increase of 10% in error for human subjects can be attributed to subject movement.

4 DISCUSSION

This work is part of a project to characterize and model the individual HRTFs of subjects, with special importance given the individual differences. We are collecting a substantial number of HRTFs so as to develop a common parametric model that will fit all or most of the experimental data.

While our experimental facilities and methodology

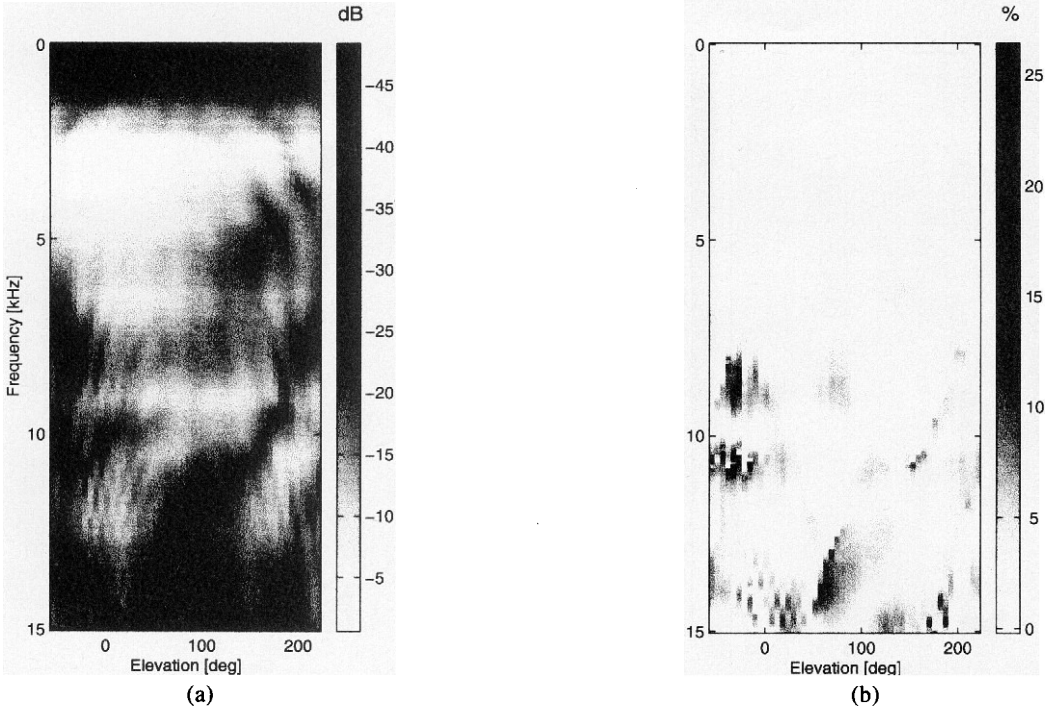


Fig. 9. (a) $P_i(\omega, \psi_k)$ in dB. (b) Relative spectral distortion for blocked meatus condition. Showing only data for $\theta = -45°$.

are adequate for the task, the measurement position within the ear canal was an issue that we had to resolve. Our results show that there is substantial flexibility in the position of the recording probe microphone within the ear canal, in that a location-independent transmission correction will transform such measurements to the equivalent pressure at the eardrum.

The analysis of errors in the experimental data indicates that individual differences also affect measurement accuracy since these are concentrated at low-energy frequency bands or at the specific frequencies where spectral notches occur. These subject-dependent errors may require repeated measurements and signal averaging to reduce them if necessary. It is our view that the larger errors observed at or near spectral notches will not be significant in auditory simulations. More serious are the errors due to subject movement. For our experiments, where we use a chin rest to limit motion, the errors due to motion are smaller than 12% at frequencies below 13 kHz. With less physical restraints on subjects, these errors could be large. We note, first, that the use of time-domain measurements allows the detection and estimation of head motion by monitoring the changes of recorded signal onset during the measurement session. As an alternative, for some critical or reference measurements, the use of a head tracker will facilitate the correction of changes in the head position and orientation.

5 SUMMARY

A detailed study of the effects of the measurement conditions on experimental HRTFs has been conducted. We determined that the position within the ear canal at which the HRTF is measured is not critical in that the HRTF at the eardrum can be determined from any such measurements by a simple transfer function compensation that is independent of the position of the sound source in three-dimensional space. The common and least intrusive technique of performing such measurements at the entrance of the blocked ear canal is thus validated. By analyzing and comparing the results of HRTF measurements for a KEMAR dummy and for two human subjects we determine that the errors are subject dependent. They are concentrated in subject-dependent frequencies of low pressure, and generally will occur at frequencies above 12 kHz. Errors will also be concentrated near spectral notches where small equipment or subject motion result in significant changes in response. For human subjects, motion can be detected from the measurements and corrected or repeated if necessary. Except for gross subject motion, measurement errors are concentrated in low-response regions or lead to small mispositioning of spectral notches. These errors are not expected to be critical for any spatial sound simulations.

6 ACKNOWLEDGMENT

This work was supported by the National Science Foundation under grant IRI-96-19339. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The authors would like to thank Richard Duda for his suggestions and comments.

7 REFERENCES

[1] H. Møller, D. Hammershøi, C. B. Jensen, and M. F. Sorensen, "Transfer Characteristics of Headphones Measured on Human Ears," *J. Audio Eng. Soc.*,

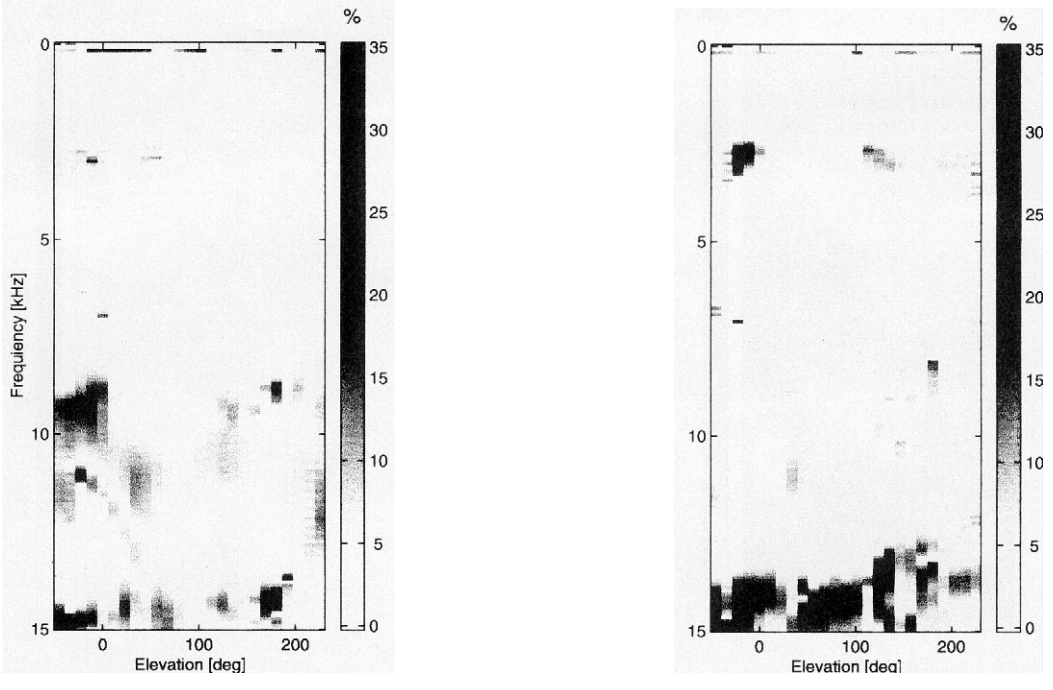


Fig. 10. Relative spectral distortion for open meatus condition. (a) Subject CA. (b) Subject DT. Showing only data for $\theta = -45°$.

vol. 43, pp. 203–217 (1995 Apr.).

[2] J. C. Middlebrooks and J. C. Makous, "Directional Sensitivity of Sound-Pressure Levels in the Human Ear Canal," *J. Acoust. Soc. Am.*, vol. 86, pp. 89–107 (1989 July).

[3] S. Mehrgardt and V. Mellert, "Transmission Characteristics of the External Human Ear," *J. Acoust. Soc. Am.*, vol. 61, pp. 1567–1576 (1977 June).

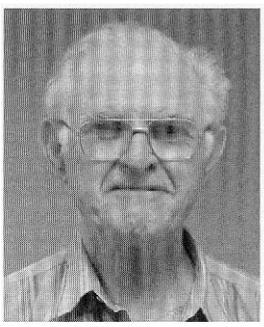
[4] H. Møller, "Fundamentals of Binaural Technology," *Appl. Acoust.*, vol. 36, pp. 171–218 (1992).

[5] M. D. Burkhard and R. M. Sachs, "Anthropometric Manikin for Acoustic Research," *J. Acoust. Soc. Am.*, vol. 58, pp. 214–222 (1975 July).

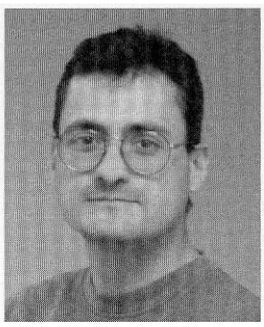
[6] D. Hammershøir and H. Møller, "Sound Transmission to and within the Human Ear Canal," *J. Acoust. Soc. Am.*, vol. 100, pp. 408–427 (1996 July).

[7] E. A. G. Shaw, "External Ear Response and Sound Localization," *Localization of Sound: Theory and Applications*, R. W. Gatehouse, Ed. (Amphora Preso, Groton, CT, 1982), pp. 30–41).

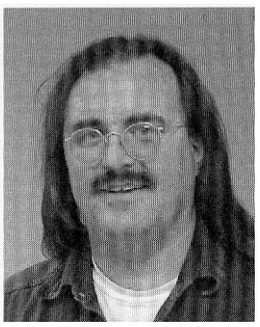
THE AUTHORS



V. R. Algazi



C. Avendano



D. M. Thompson

V. Ralph Algazi received the degree of Ingénieur Radio from l'Ecole Supérieure d'Electricité (ESE), Paris, France, and the M.S. and Ph.D. degrees from the Massachusetts Institute of Technology, Cambridge, in 1952, 1955, and 1963, respectively.

He was at M.I.T. from 1959 to 1965 as a Research and Teaching Assistant and then as a Post-Doctoral Fellow and Assistant Professor. On the faculty of the University of California, Davis, since 1965, he was chairman of the Department of Electrical and Computer Engineering from 1975 to 1986. He founded CIPIC, the Center for Image Processing and Integrated Computing, in 1989 and served as its director until 1994. He is now a Research Professor at CIPIC, pursuing research interests in signal processing, engineering applications of human perception for both speech and images, and image and video processing and coding.

Dr. Algazi is a life senior member of the IEEE and a member of SPIE and AAAS.

•

Carlos Avendano was born in Mexico City in 1968. He received the B.S. degree from the Instituto Tecnologico y de Estudios Superiores de Monterrey (ITESM CEM) Mexico City, in 1991, and the M.S. and Ph.D. degrees from the Oregon Graduate Institute of Science and Technology, Portland, in 1993 and 1997, respectively, all in electrical engineering.

While in Oregon he pursued research in speech processing with applications to enhancement and robust automatic recognition. He is currently a Post-Doctoral Fellow at the CIPIC Interface Lab at the University of California, Davis, working on several aspects of spatial hearing, binaural technology, and virtual auditory displays.

Dr. Avendano is member of the IEEE.

•

Dennis M. Thompson was born in Bradenton, FL, in 1958. He studied Electronic Technology at the College of the Redwoods, Eureka, CA. He is currently working toward a degree in electrical engineering at the University of California at Davis.

In 1985 he started Yknot Sound, a regional sound company specializing in live music PA systems. Currently he is working at the CIPIC Interface Lab, where he designs hardware and software. His main research interest is high-quality 3D sound reproduction. He still enjoys working in concert sound reinforcement, with an emphasis on quality over quantity.

Mr. Thompson is a student member of the Audio Engineering Society.