

A spatial feature extraction and regularization model for the head-related transfer function^{a)}

Jiashu Chen^{b)}

Departments of Neurology and Neurophysiology, University of Wisconsin-Madison, H6/573 CSC, 600 Highland Avenue, Madison, Wisconsin 53792

Barry D. Van Veen

Department of Electrical and Computer Engineering, University of Wisconsin-Madison, 1415 Johnson Drive, Madison, Wisconsin 53706

Kurt E. Hecox

Department of Neurology, University of Wisconsin-Madison, H6/573 CSC, 600 Highland Avenue, Madison, Wisconsin 53792

(Received 12 July 1993; accepted for publication 10 August 1994)

A functional representation is proposed for complex valued (amplitude and phase) head-related transfer functions (HRTFs), including both frequency and spatial dependence. The frequency variation is spanned by a set of eigentransfer functions (EFs) that are generated using the Karhunen–Loève expansion. Any HRTF is represented as a weighted combination of the EFs where the weights are functions of the HRTFs spatial location and are termed spatial characteristic functions (SCFs). Samples of the SCFs are obtained by projecting the measured HRTFs onto the EFs. A regularization framework is employed to obtain a functional representation for the SCFs by fitting each set of SCF samples with a two-dimensional spline. Acoustic validation of the model's fidelity and predictive capability is provided using 2188 measured HRTFs from a KEMAR manikin and 1816 measured HRTFs from an anesthetized live cat. Errors between measured and modeled HRTFs are generally less than one percent. Larger errors occur in the contralateral regions for KEMAR and lower back regions for the cat as a consequence of the relatively small HRTF amplitudes resulting from head shadowing. Methods for reducing these errors are discussed.

PACS numbers: 43.66.Ba, 43.64.Bt, 43.66.Qp [HSC]

INTRODUCTION

Recent extensions of the classical duplex theory of binaural hearing have focused on measurement and study of the external ear's free-field-to-eardrum transfer functions or head-related transfer functions (HRTFs). The magnitude and phase spectra of measured HRTFs vary as a function of sound source location. Hence, it is commonly acknowledged that the HRTF introduces important cues in spatial hearing. Measured HRTFs have been used to study the transformation characteristics of the external ear and to synthesize virtual direction-dependent stimuli over headphones (Blauert, 1983; Middlebrooks *et al.*, 1990; Carlile, 1990; Wightman and Kistler, 1989a,b; Musicant *et al.*, 1990; Wenzel *et al.*, 1990).

Important foundations for understanding HRTFs are due in large part to Shaw and his associates (Shaw, 1974a,b, 1975, 1979). In a series of sophisticated acoustic experiments, they recorded multimodal wave propagation on a simplified pinna model and on real human ears. Shaw concluded that the cartilage structure of the external ear forms a number of acoustic resonators. The resonant frequencies and their magnitudes vary with sound source location. As a result, the spectrum of incoming sounds is altered by the direction-

dependent transformation characteristics of the external ear. These physical findings are supported by behavioral studies that suggest the direction-dependent cues provided by the pinna are important for determination of source elevation (Herbrank and Wright, 1974; Butler, 1975) and azimuth (Blauert, 1983; Oldfield, 1984). Another series of behavioral experiments conducted by Batteau (see Tobias, 1972) indicate that pinna cues are primarily responsible for sound image externalization.

Techniques to measure the HRTFs at given sound source locations using broadband stimuli were developed by Mergardt and Mellert (1977) and Blauert (1983). These techniques have been recently refined by several authors (Wightman and Kistler, 1989a; Carlile, 1989; Pralong *et al.*, 1992; Middlebrooks *et al.*, 1989, 1990).

Advances in computer and digital signal processing technology have enabled researchers to synthesize directional stimuli for headphone presentation using measured HRTFs. This capability allows complete and independent control of localization cues in synthesizing virtual auditory signals and eliminates the difficulties of setting up experiments in an anechoic chamber. For example, this approach has been used to study human localization behavior (Wightman and Kistler, 1989a,b) as well as single unit virtual receptive fields at the brain stem and cortex (Reale *et al.*, 1994; Brugge *et al.*, 1992). Another application of HRTF synthesis is 3-D virtual auditory display. Virtual 3-D auditory display

^{a)}Portions of this work were presented at the 124th ASA meeting [J. Acoust. Soc. Am. **92**, 2333(A) (1992)] and IEEE ICASSP-93 Conference.

^{b)}Present address: Nicolet Instrument Corp., 5225 Verona Rd., Madison, WI 53711-4495.

is used in applications such as advanced human-computer interfaces, navigation aids for the visually impaired, telepresence, and virtual reality (Wenzel, 1992).

Use of measured HRTFs presents serious limitations in many synthesis applications. First, the number of measurements can be very large if the entire auditory space is to be represented on a fine grid. This presents formidable experimental obstacles with human or animal subjects because of the corresponding long data collection times. Second, it is very difficult, if not impossible, to visualize and study the characteristics of such a vast data set. This is especially true when investigating differences across subjects or species. Third, the measured HRTFs only represent discrete samples of the underlying continuous auditory space. Synthesis of sounds coming from unmeasured spatial locations is not possible without resorting to an interpolation procedure. As noted by Wightman *et al.* (1992), how to best accomplish interpolation is a difficult question. Furthermore, when multiple sound sources or room acoustics are included in the simulation of complex acoustical environments, the computational burden of approaches depending exclusively on measured HRTFs becomes prohibitive because of the large number of HRTFs required.

These limitations have stimulated interest in functional representations of the HRTFs. That is, one seeks a mathematical model or equation that represents the HRTF as a function of frequency and direction. Simulation of 3-D auditory display is then performed by using the model or equation to obtain the HRTF. Perhaps Batteau's work (Batteau, 1967, 1968) represents the first functional approach. Batteau conjectured that the external ear could be modeled as a three-channel two-delay and sum acoustic coupler. One delay varies with the sound source elevation and the other with the source azimuth. Wright *et al.* (1974) and Watkins (1979) used this model in localization research to synthesize crude ear drum signals. A rather recent effort is that of Genuit (1986). He devised a filter-bank model that has 16 time-delayed channels. Using classic acoustics he established the relationship between the filter parameters and external ear geometry. This is an attractive approach because it avoids the need for ear canal recordings. Thorough acoustic and behavioral validation of this model is not yet reported, however.

More recently, a functional model of the external ear derived from measured HRTFs was proposed by Chen *et al.* (1992). Here the external ear is modeled as a multisensor broadband beamformer, with the sensor geometry and beamformer weight vector set chosen to represent the physical characteristics of the external ear. Similar to Batteau's model, the beamforming model provides an explicit mathematical relationship between the HRTF and source location. Hence, it can interpolate HRTFs at arbitrary directions. However, the number of weights must increase rapidly to maintain model accuracy as the size of the solid angle to be modeled increases (Chen, 1992). This results in excessive computational burden and numerical instability. These problems are mainly due to the fact that the weight vectors do not form an orthogonal basis for the modeled HRTF space. Application of the beamforming model is thus limited to repre-

sentation of external ear characteristics over relative small solid angles.

Low-dimensional and orthogonal representations for sets of HRTFs have been generated (Martens, 1987; Kistler and Wightman, 1992; Middlebrooks and Green, 1992) by applying principal components analysis (PCA) to the logarithms of the HRTFs magnitudes after the directionally independent frequency dependence is removed. Martens (1987) considered HRTFs measured at 36 directions in the horizontal plane. Kistler and Wightman (1992) applied this method to HRTFs measured at 256 spatial directions over both azimuth and elevation from a set of 10 subjects. From the PCA representation, HRTFs for each direction were reconstructed using phase estimates based on the minimum phase characteristics corresponding to each magnitude function. Kistler and Wightman also assessed the perceptual validity of their approach by comparing human listeners' judgements of direction based on HRTFs as measured empirically with judgements based on HRTFs as reconstructed from the PCA. Agreement was generally good. Note that these studies do not provide a functional representation since the weights are available only for the measured directions.

In this paper, we propose an alternate functional model for the complex valued (magnitude and phase) HRTFs. In this model the HRTFs are expressed as weighted combinations of a set of complex valued eigentransfer functions (EFs). The EFs are an orthogonal set of frequency-dependent functions; the weights applied to each EF are functions only of spatial location and are thus termed spatial characteristic functions (SCFs). Estimates of the EFs are obtained from a discrete Karhunen-Loève expansion procedure applied to a sample covariance matrix constructed from the complex valued measured HRTFs represented on a linear scale. Samples of the SCFs at the measurement locations are obtained by projecting each of the EFs onto the measured HRTFs. A functional representation or mathematical equation for each SCF is obtained by fitting the SCF samples with a thin-plate (two-dimensional) spline using procedures of regularization theory (Tikhonov and Arsenin, 1977; Wahba, 1990). We hence adopt the descriptive terminology "the spatial feature extraction and regularization" (SFER) model.

An extensive acoustical validation of the SFER model establishes its effectiveness. We show that over large angular sectors the relative approximation and interpolation mean squared errors are less than one percent. Some relationships between the functional behavior of the SCFs and physical external ear geometry are conjectured.

The paper is organized as follows. Section I develops the SFER model and procedures for estimating the EFs and SCFs from the measured HRTFs. Section II describes the acoustic validation of the model for two data sets: One collected from a KEMAR manikin head and the other from an anesthetized live cat. The paper concludes with a discussion in Sec. III. Notationally, lower and upper case bold face characters represent vectors and matrices, respectively. Superscript * and H denote complex conjugate and Hermitian transpose, respectively.

I. SFER MODEL

The basic approach of the SFER model is to approximate the complex HRTF for a given direction (θ, ϕ) as a weighted sum of M (most important) complex EFs $q_i(\omega)$ combined with a directionally independent component $q_0(\omega)$ as indicated in Eq. (1)

$$\hat{h}(\omega, \theta, \phi) = \sum_{i=1}^M q_i(\omega) \hat{w}_i(\theta, \phi) + q_0(\omega). \quad (1)$$

The complex weights $\hat{w}_i(\theta, \phi)$ are termed spatial characteristic functions SCFs. The SCFs are obtained by fitting a thin plate spline to a discrete set of weights obtained by projecting the measured HRTFs onto the EFs. Since the frequency range of interest is bandlimited, the complex EFs $q_i(\omega)$ are represented as vectors \mathbf{q}_i of complex samples in frequency.

A. Eigentransfer functions (EFs)

The Karhunen–Loève expansion (KLE) is used here to express the HRTF in terms of an orthogonal function expansion. The KLE is usually used to expand a random process (Van Trees, 1968; Papoulis, 1991) in terms of a set of orthogonal functions such that the randomness of the process is contained within the expansion coefficients. Particular realizations of the process are represented as weighted sums of a fixed set of time functions. The weights are uncorrelated random variables. Note that the KLE is a doubly orthogonal expansion since the basis functions are orthogonal and the expansion coefficients are uncorrelated. Here we propose using the KLE to obtain an orthogonal expansion that separates the temporal or frequency variation of the HRTFs from the spatial variation. The spatial variation is treated as “randomness” in this application of the KLE, although it is in fact quite deterministic. This leads to the functional model for the HRTFs Eq. (1).

In our application, the HRTFs are not available as continuous functions of θ and ϕ . Therefore, a discrete analog of the KLE must be employed using measured HRTFs. Sampling in frequency, ω is only necessary for computational purposes. Provided the data collection procedure used to measure the HRTFs satisfies the conditions of the sampling theorem and has sufficient time extent to completely capture the longest HRTF impulse response, the HRTFs can be regarded as continuous functions of frequency since an arbitrary frequency resolution is obtained by zero padding the impulse responses and taking the discrete Fourier transform.

Let $\mathbf{h}(\theta_j, \phi_j)$, $j=1, 2, \dots, P$ be an N -by-one complex valued vector representing the N frequency samples of the HRTF measured at the j th location, (θ_j, ϕ_j) . As noted above, N can be chosen arbitrarily large to obtain any desired frequency resolution. Applying the KLE methodology to this N -dimensional representation for the spatially sampled HRTFs replaces the EF basis functions $q_i(\omega)$, $i=0, 1, \dots, M$, with the EF vectors \mathbf{q}_i , $i=0, 1, \dots, M$, where

$$\mathbf{q}_0 = \frac{1}{P} \sum_{j=1}^P \mathfrak{z}(\theta_j, \phi_j) \mathbf{h}(\theta_j, \phi_j) \quad (2)$$

is the sample mean and \mathbf{q}_i , $i=1, 2, \dots, M$ are the eigenvectors corresponding to the M largest eigenvalues of the sample covariance matrix

$$\mathbf{R}_h = \frac{1}{P} \sum_{j=1}^P [\mathfrak{z}(\theta_j, \phi_j) \mathbf{h}(\theta_j, \phi_j) - \mathbf{q}_0] \times [\mathfrak{z}(\theta_j, \phi_j) \mathbf{h}(\theta_j, \phi_j) - \mathbf{q}_0]^H. \quad (3)$$

The weighting function $\mathfrak{z}(\theta_j, \phi_j)$ is a sample of a non-negative real-valued function that can be used to emphasize some directions over others. Let the eigenvalues of \mathbf{R}_h be arranged in decreasing order $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$. The mean squared error (MSE) associated with use of M EFs to represent the P HRTFs is

$$\text{MSE} = \sum_{i=M+1}^N \lambda_i. \quad (4)$$

The eigenvectors of a sample covariance matrix corresponding to the largest eigenvalues are often referred to as principal components in the context of statistical data analysis. Martens (1987) adopted this terminology, motivated in part by the desire to use multivariate analysis as a data reduction procedure for the spectral variation in the logarithm of the direction-dependent transfer function (DTF) magnitude associated with changing direction. Kistler and Wightman (1992) also applied principal component analysis to logarithm of the DTF magnitude. Note that although our driving philosophy is functional approximation, not statistical data analysis, the mathematics of KLE in the discrete finite case and PCA are the same. However, in the SFER model the sample covariance matrix in Eq. (3) is computed using the complex valued HRTFs, not the logarithm of the DTF magnitudes. We use the Karhunen–Loève vocabulary instead of the PCA vocabulary because the communication theory approach is the basis of our work. A more thorough discussion of the differences between the PCA and SFER approaches is provided in Sec. III. We refer to the eigenvectors of \mathbf{R}_h as EFs because of their physical significance as transfer functions in the SFER model.

B. Spatial characteristic functions

The sampling in space is more problematic than sampling in frequency because the spatial bandwidth is unknown. Hence, the required spatial sampling density and mathematics for reconstructing the HRTFs as a continuous function of θ and ϕ from samples are unknown. Here we describe a reconstruction procedure based on fitting two-dimensional splines to spatial samples of the SCFs. Samples of the M SCFs at the P measurement locations are obtained from the M EFs as

$$w_i(\theta_j, \phi_j) = \mathbf{q}_i^H [\mathbf{h}(\theta_j, \phi_j) - \mathbf{q}_0], \quad i=1, 2, \dots, M; \quad j=1, 2, \dots, P. \quad (5)$$

Reconstruction of the underlying continuous SCFs from the samples obtained in (5) is an inherently ill-posed problem because the samples do not uniquely define the SCFs in the absence of additional assumptions. Furthermore, the samples

in (5) are usually corrupted by the presence of noise. Regularization theory (Tikhonov and Arsenin, 1977) offers a general framework for transforming ill-posed problems to well-posed problems through the use of smoothness constraints. In the present application the smoothness constraints imply that a small change in covariates θ and ϕ induce a small change in the SCFs. The smoothness constraints are quite reasonable since the HRTFs originate from a physical system of limited spatial extent.

The software package developed by Gu (1989) is used here to obtain the parameters of the thin plate spline approximating the real and imaginary components of each SCF. The regularization parameter that controls the trade-off between the fidelity and the roughness of the solution is determined using the generalized cross validation method (Craven and Wahba, 1979; Wahba, 1977).

Let $\hat{w}_i(\theta, \phi)$, $i=1, 2, \dots, M$ be the SCFs reconstructed from the corresponding real and imaginary thin plate splines. The HRTF at N frequency points is then modeled as

$$\hat{\mathbf{h}}(\theta, \phi) = \sum_{i=1}^M \mathbf{q}_i \hat{w}_i(\theta, \phi) + \mathbf{q}_0. \quad (6)$$

Note that alternate methods, such as ones based on linear interpolation, could be used to reconstruct continuous SCFs from measured samples. The effectiveness of alternate approaches has not yet been evaluated.

II. MODEL VALIDATION

A. The free-field and ear-canal recordings

Two sets of data are employed in evaluation of the SFER models in this paper: One is acquired using a KEMAR manikin head and the other one is from an anesthetized live cat. A female pinna was used on the KEMAR manikin head. A brief summary of the data acquisition setup is given here. A more detailed description is provided in Musicant *et al.* (1990).

Data acquisition is conducted in a double-walled, sound-insulated room (IAC), with inner dimensions of 2.25 \times 2.15 \times 2 m. The subject (either the cat or the KEMAR head model) is positioned in the center of a spherical coordinate system. The inner walls of the room are lined with 4-in.-thick reticulated wedged foam (Sonex) to approximate an anechoic environment. The conventional (azimuth, elevation) coordinate system for localization research is used. The frontal direction at ear level is defined as 0° azimuth and 0° elevation. The left and right sides at ear level are defined as -90° and +90° azimuth, respectively. The back direction at ear level is defined as \mp 180° azimuth. Direction on the median plane spans the range of \pm 90° elevation.

The stimulus is presented through a movable loudspeaker system controlled by a computer, allowing the speaker to move on a 75-cm-radius sphere with a minimum step size of 0.9°. Both the KEMAR and cat data are collected at 4.5° intervals over the sphere, except the portion below -36° elevation and the portion behind the subject below 0° elevation and from -162° to +162° azimuth. For each spatial position one ear-canal recording is made as an average of

TABLE I. Data acquisition set up for KEMAR and cat.

Parameters	Subject	
	KEMAR	Cat
Pulse duration	20 μ s	10 μ s
Speaker	Realistic mid range	Radio-Shack Super Tweeter
Speaker freq. range	2-25 kHz	2-37.5 kHz
Microphone	B&K 1/2 in. with coupler	Knowles EA 1934
Time window	10.0 ms	6.4 ms
Sampling rate	102.4 kHz	160 kHz
Frequency resolution	100 Hz	156 Hz

300 responses to a broadband speaker signal obtained in response to a pulse input. A calibration or free-field signal is collected with a microphone located at the virtual position of the subject's eardrum. The cat ear-canal recordings are collected with a probe microphone surgically inserted into the ear canal as close as possible to the eardrum. For the KEMAR case, the microphones are located in the couplers representing the ear canal. Table I lists additional information concerning the data acquisition parameters.

The signal-to-noise ratio is difficult to define precisely because the signal power varies with azimuth and elevation. After averaging 300 responses, the ratio of the average post-stimulus power in the raw data has average values over azimuth and elevation of 61.6 and 56.0 dB for the KEMAR recordings and the cat recordings, respectively. In the KEMAR recordings the noise has a flat spectral shape with level of approximately -60 dB. The spectral level of the KEMAR ear-canal recordings is generally well above the noise floor, even in the spectral notches. For example, at zero degrees azimuth and elevation the highest peak is at 20 dB and deepest notch at approximately -25 dB. Relative to zero degrees azimuth and elevation, the overall spectral level tends to increase in the ipsilateral direction and decrease in the contralateral direction. The spectral levels for the cat ear-canal recordings are similar to those of KEMAR, although the noise characteristics differ. The noise in the cat data has a low-pass characteristic with peak level of -45 dB. The level is generally low (\cong -60 dB), however, at higher frequencies where spectral notches occur in the ear-canal recordings.

B. Estimation of HRTFs based on free-field and ear-canal recordings

A variable acoustic delay is present in each ear-canal recording because the distance from the loudspeaker to the ear varies slightly as a function of spatial location. An onset based time alignment procedure was used to remove this variable delay and to provide a consistent time origin across all spatial locations. The first data point that exceeded 10% of the maximum response over the entire recording was chosen as the onset of the response. Each recording was then shifted backward or forward to a common time origin. The onset criterion was preferred over determining the relative delay from the peak of the cross correlation between the free-field and ear-canal recordings because of potential errors in the cross-correlation method associated with HRTFs hav-

ing very large group delays. The time shift associated with each individual recording can be stored for later incorporation in a binaural model.

The spectral analysis method of system identification (Ljung, 1987) was then used to estimate the HRTFs from the free-field and ear-canal recordings. The spectral analysis method first estimates the cross-correlation function between the free-field and ear-canal recordings and the auto-correlation function of the free-field recordings. A window function is applied to the estimated cross- and auto-correlation functions to deemphasize unreliable correlation estimates at large time lags (Kay, 1987). The HRTF is estimated by dividing the cross spectrum by the autospectrum, where the cross and autospectra are given by the discrete Fourier transforms (DFT) of the corresponding windowed correlation estimates. This procedure results in an HRTF estimate of lower variance than the estimate obtained by dividing the DFT of the ear-canal recording by the DFT of the free-field recording (Ljung, 1987). The spectral analysis procedure was implemented using the Matlab 4.0 System Identification Toolbox (Mathworks, Inc., Natick, MA) SPA routines.

Although the primary purpose of applying the window to the correlation functions is to deemphasize less reliable correlation estimates, here it also served to remove the effects of echoes present in the sound insulated room. Previous investigators (Hiranaka and Yamasaki, 1983; Pössl *et al.*, 1986) have reported the human HRTF impulse response to be about 0.5 ms in length. These findings are consistent with our data. The speaker pulse response measured in free space for the KEMAR data was approximately 0.5 ms long so the measured external ear response was expected to last about 1.0 ms. Therefore, the significant portion of cross- and auto-correlation functions are contained within 2 ms of the zero-lag point. The cat impulse responses have slightly shorter duration because the cat external ear is smaller than KEMAR's. A Hanning window with window width of 256 data points representing 2.5 ms in duration was applied to the cross- and autocorrelation functions. Hence, the windowing procedure eliminated echoes and noise occurring 1.25 ms or later after the onset of the free-field and ear-canal recordings. A similar window was applied to the cat data.

The effect of varying window width is illustrated by comparing estimated HRTFs and the corresponding impulse responses in Figs. 1 and 2, respectively, for window widths of 64, 128, 256, and 512 data points. The results obtained using 256 and 512 point windows are very similar. The shorter window widths (64 and 128 points) overly smooth the cross- and auto-correlation functions and introduce a symmetric distortion in the recovered impulse response. This effect is not evident in the 256 and 512 point cases. The 256 point window was chosen to minimize the effects of echoes and keep the variance as small as possible.

Recall that zero padding can be used to estimate the HRTFs at an arbitrary frequency resolution. Here 256 frequency samples were used to represent the full positive frequency range, 51.2 kHz for KEMAR and 80 kHz for the cat. Only the first half of these samples were retained to represent the effective frequency ranges, 25.6 kHz for KEMAR and 40

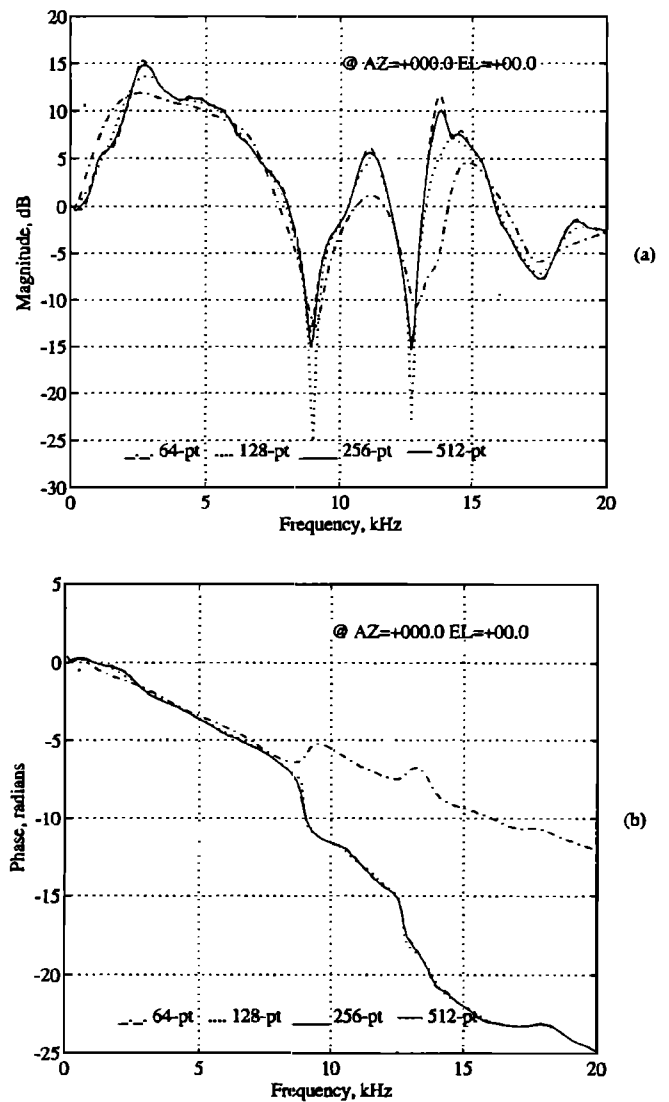


FIG. 1. Magnitude and phase of the HRTF at $\theta=0^\circ$, $\phi=0^\circ$ obtained using different window widths.

kHz for the cat. A linear phase component corresponding to a time delay of 0.5 ms was added to each HRTF to prevent noncausal impulse responses that could result from small errors in the time alignment procedure. The result was stored as a 129 point complex valued data record.

C. Covariance analysis of the measured HRTFs

The eigenvalues λ_i , $i=1, \dots, N$ represent the sample variance of $\mathbf{h}(\theta_j, \phi_j) - \mathbf{q}_0$, $j=1, 2, \dots, P$ projected onto each EF \mathbf{q}_i , $i=1, \dots, N$. This is the variability in the data that is represented by the corresponding EF. Hence, the number of eigenvectors, M , required to achieve a given mean squared error is determined by the relative size of the eigenvalues. The results below indicate that a very large percentage of variability in the measured HRTFs is represented using a relatively small value of M .

Recall that there are 80 HRTFs sampled on each elevation circle. The circles at higher and lower elevations span smaller solid angles than the circles near the equator. As an extreme example of this phenomena, at the vertex or north

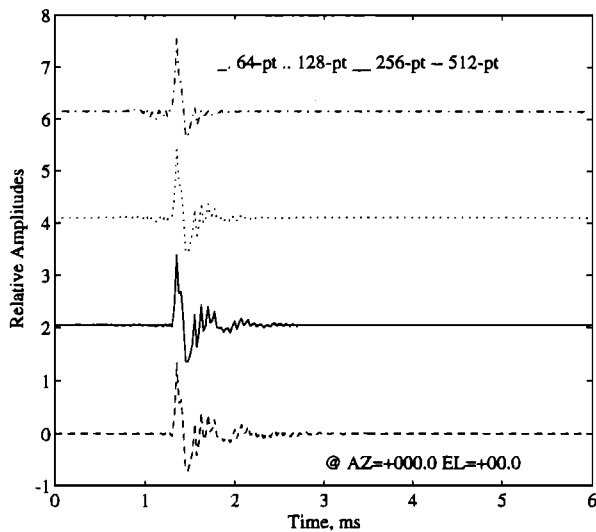


FIG. 2. Impulse responses of the HRTF at $\theta=0^\circ$, $\phi=0^\circ$ obtained using different window widths.

pole 80 HRTFs are measured at one single location. To prevent these higher elevation HRTFs from dominating the sample covariance matrix, the HRTFs are weighted proportional to the solid angle spanned by the corresponding elevation circles with the weighting function

$$\mathfrak{Z}(\theta, \phi) = 1 - |\sin(\phi)|. \quad (7)$$

A convenient metric for determining an appropriate value M is the percentage variance represented by the first M components. This is expressed as

$$\% \text{ var}(M) = \frac{\sum_{i=1}^M \lambda_i}{\sum_{i=1}^N \lambda_i} \times 100\%. \quad (8)$$

A natural question that arises is the spatial sampling density required to adequately describe HRTF behavior. Table II presents $\% \text{ var}(M)$ of the measured HRTFs for KEMAR and the cat for $M=1$ to 20. The left half gives $\% \text{ var}(M)$ when the sample covariance matrix is constructed from the HRTFs measured on a 9° grid, that is, 561 HRTFs for KEMAR and 593 HRTFs for cat. The right half gives the $\% \text{ var}(M)$ for measurements on a 4.5° grid, that is, 2188 HRTFs for KEMAR and 1816 HRTFs for cat, that is represented by the EFs determined from the HRTFs on the 9° grid. The percent variance of the 4.5° data set represented by the EF's determined from HRTFs on a 9° grid and that of the 9° data set becomes nearly identical as M increases. This suggests that sampling at 9° intervals captures nearly all the variability in the HRTFs, provided that M is properly chosen.

For the KEMAR data 99.9% of the variance is represented by 12 EFs while 17 EFs are required to represent 99.9% of the cat data variance. This indicates that the HRTFs have most of their energy distributed in a low-dimensional subspace. The increased number of EFs required by the cat data to account for the same amount of variance is a consequence of the broader bandwidth used for the cat data. If the cat HRTF bandwidth is reduced from 40 to 20 kHz, then only 11 EFs are required to represent 99.9% of the variance in the

TABLE II. Percent variance as a function of M for KEMAR and cat data ($N=128$).

M	% var(M) on a 9.0° grid		% var(M) on a 4.5° grid	
	KEMAR	Cat	KEMAR	Cat
1	68.58	55.78	60.74	49.02
2	87.34	75.17	84.18	69.29
3	95.37	84.26	94.40	80.58
4	98.09	92.03	97.63	89.45
5	98.90	95.61	98.68	93.79
6	99.36	96.75	99.23	95.39
7	99.56	97.62	99.47	96.48
8	99.69	98.30	99.61	97.38
9	99.79	98.88	99.74	98.21
10	99.85	99.23	99.81	99.72
11	99.90	99.44	99.87	99.03
12	99.92	99.61	99.91	99.26
13	99.95	99.73	99.93	99.46
14	99.96	99.83	99.95	99.62
15	99.98	99.88	99.97	99.74
16	99.99	99.92	99.98	99.83
17	99.99	99.95	99.98	99.90
18	99.99	99.97	99.98	99.93
19	99.99	99.98	99.99	99.95
20	99.99	99.98	99.99	99.97

cat data. Unless otherwise noted, in the remainder of this paper $M=12$ and $M=17$ EFs are used for the KEMAR and cat models, respectively.

D. Mean-square error between the measured and modeled HRTFs

The effectiveness of the spatial feature extraction and regularization procedure is established by comparing measured and modeled HRTFs. Overlaid plots of measured and modeled HRTFs qualitatively illustrate the model performance while average mean square error in 32 subregions quantitatively validate the model. An emphasis is given to the KEMAR data in the plots due to space constraints. A more detailed error analysis can be found in (Chen, 1992).

In order to use one set of measurements to validate both the fidelity and predictive power of the SFER model, HRTFs on a 9° grid (561 for KEMAR and 593 for cat) are used to determine the EFs and SCFs of the model. The average SNR of KEMAR data on the 9° grid is 53.2 dB with only 3 of the 561 measurements having SNR less than 30 dB. For the cat, the average SNR on the 9° grid is 57.7 dB and all 593 measurements have SNR greater than 30 dB. The model fidelity is evaluated by comparing measured and modeled HRTFs at the locations used to determine the model's EFs and SCFs. The predictive power is demonstrated by comparing measured and modeled HRTFs at the locations not used to determine the EFs and SCFs of the model. These locations are at the midpoints of those used to determine the model parameters and thus represent likely locations of maximum interpolation error.

Figures 3–6 compare measured and modeled responses on the median plane and equator for the KEMAR HRTFs. The solid lines represent measured HRTFs and dashed lines represent modeled HRTFs. Panel (a) depicts the magnitudes and panel (b) the phases. In each case the linear phase com-

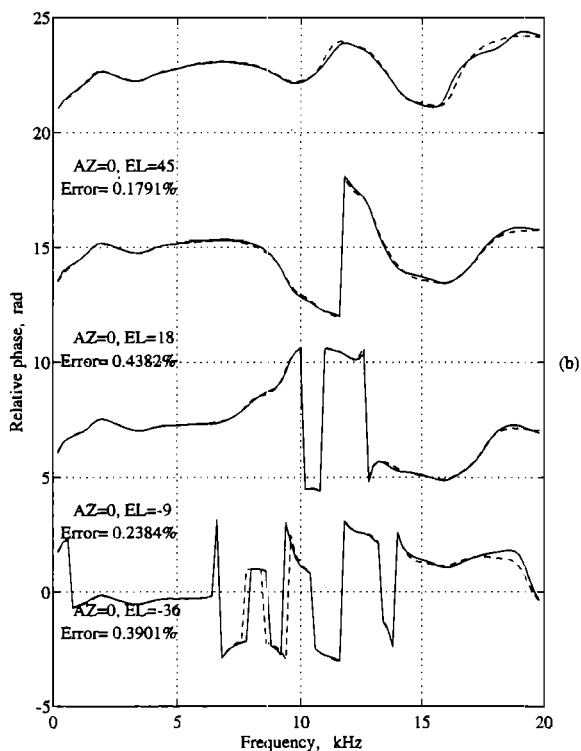
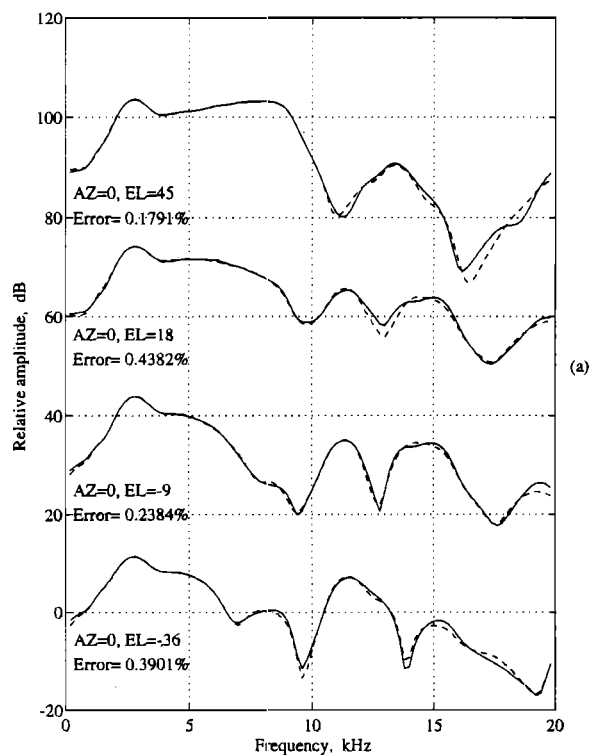
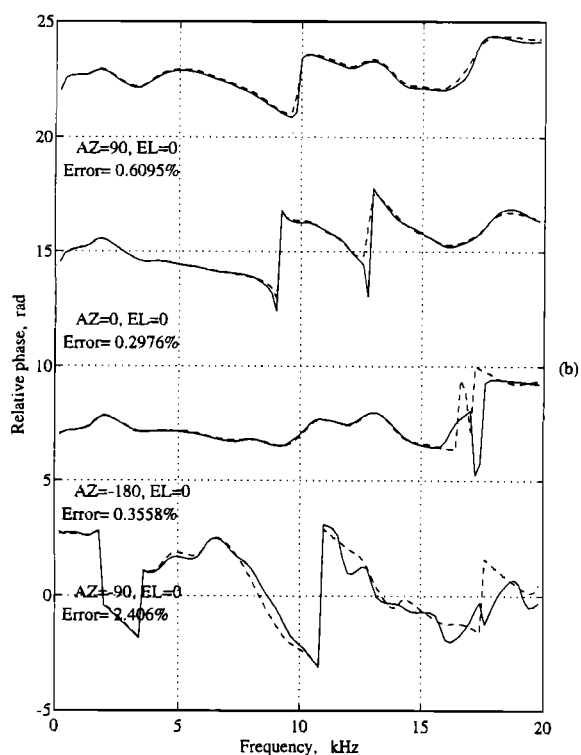
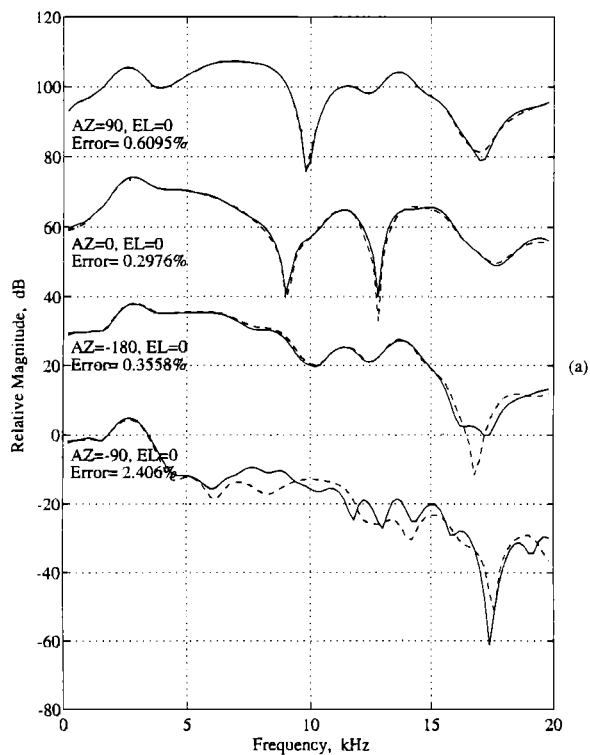


FIG. 3. Comparison between measured and modeled HRTFs at four locations in the horizontal plane. Panel (a) depicts the magnitude and panel (b) depicts the phase. Solid lines represent measured responses and the dashed lines represent modeled responses.

FIG. 4. Comparison between measured and modeled HRTFs at four locations in the median plane. Panel (a) depicts magnitude and panel (b) depicts phase. Solid lines represent measured responses and the dashed lines represent modeled responses.

ponent is removed to aid visualization of the phase characteristics. The azimuth and elevation of each HRTF is noted by symbols “AZ=” and “EL=,” respectively. The percent mean square error e_j at each location is noted with the symbol “Error.” It is defined as

$$e_j = \frac{\|\mathbf{h}_j - \hat{\mathbf{h}}(\theta_j, \phi_j)\|^2}{\|\mathbf{h}_j\|^2} \times 100\%, \quad (9)$$

where \mathbf{h}_j is the j th measured HRTF and $\hat{\mathbf{h}}(\theta_j, \phi_j)$ is given by Eq. (6).

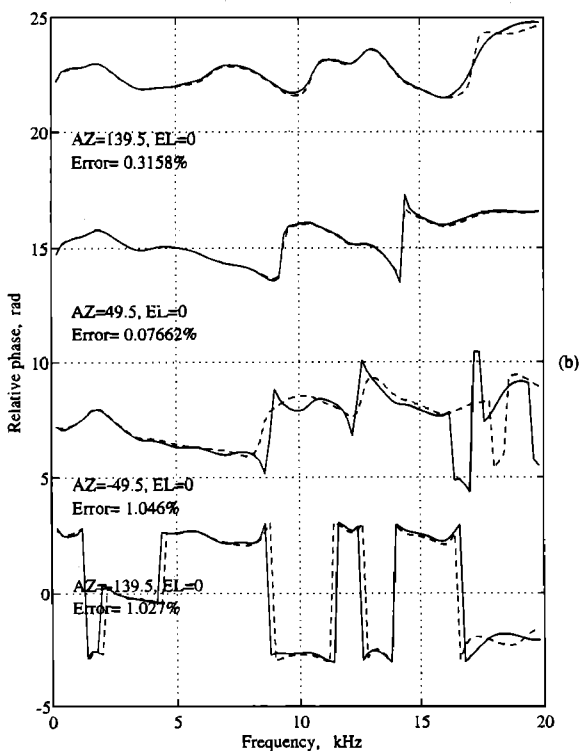
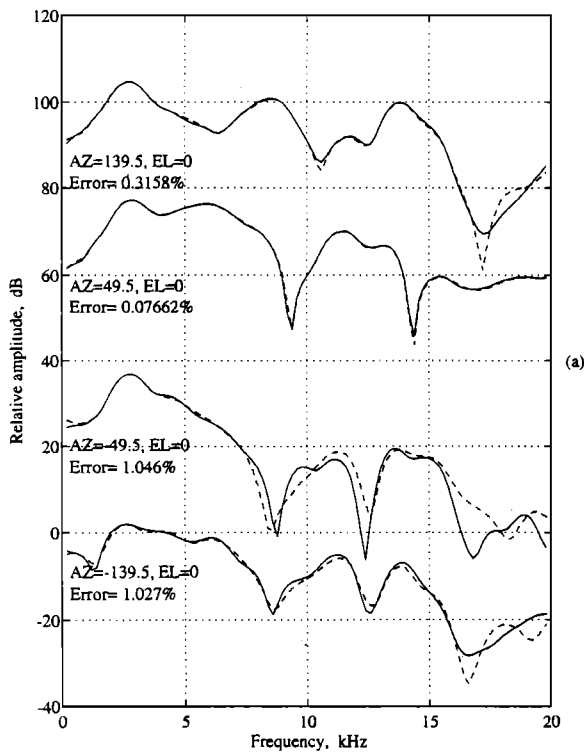


FIG. 5. Comparison between measured and model interpolated HRTFs at four locations in the horizontal plane. Panel (a) depicts magnitude and panel (b) depicts phase. Solid lines represent measured responses and the dashed lines represent model interpolated responses.

Figure 3 compares measured and modeled HRTF magnitude and phase at four typical directions in the horizontal plane. Figure 4 compares measured and modeled responses at four directions in the median plane. The directions illustrated in Figs. 3 and 4 are in the set used to determine the model parameters. Figures 5 and 6 illustrate the predictive

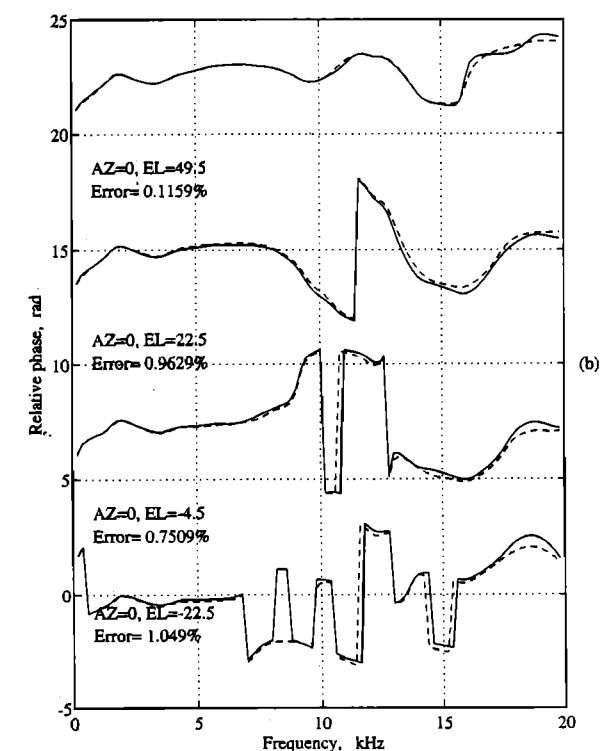
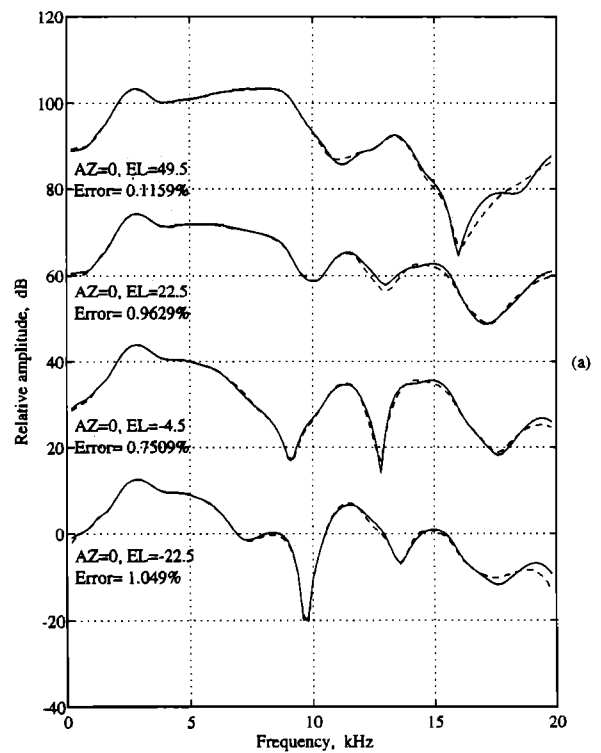


FIG. 6. Comparison between measured and model interpolated HRTFs at four locations in the median plane. Panel (a) depicts magnitude and panel (b) depicts phase. Solid lines represent measured responses and the dashed lines represent model interpolated responses.

power of the model by comparing measured and modeled HRTF magnitude and phase at representative directions not used to determine the model parameters.

Observe that the interpolation errors are slightly greater than approximation errors. The smoothing parameter λ controls the tradeoff between interpolation and approximation

TABLE III. Average error in 32 subareas on the sphere for KEMAR.

Azimuth in degrees	Elevation in degrees			
	-36 to 0	0 to 30	30 to 60	60 to 90
-180 to -135	7.5530	0.9848	0.3866	0.6013
-135 to -90	5.5541	6.5346	0.8624	0.6737
-90 to -45	2.1699	2.8521	0.8732	0.5820
-45 to 0	0.5981	0.6260	0.4446	0.4751
0 to 45	0.5153	0.6324	0.4929	0.4235
45 to 90	0.5413	0.5037	0.4117	0.7242
90 to 135	0.6337	0.6978	0.4528	0.5728
135 to 180	0.7382	0.5455	0.3015	0.4974

errors. A larger value for λ would increase the smoothing emphasis and potentially equalize the interpolation and approximation errors.

Table III lists the average errors over 32 subregions on the sphere for KEMAR using 2188 locations. The average errors over most of the frontal and ipsilateral regions are less than 1%, indicating an excellent “goodness-of-fit” in these areas. Larger errors are obtained over five contralateral and lower elevation areas. Several factors contribute to this increased error. The most significant is the interaural level difference associated with head shadowing.

Figure 7 depicts the distribution of HRTF broadband gain as a function of azimuth and elevation. The levels on the ipsilateral side between 50 to 100° azimuth and 0 to 54° elevation are near the maximum of 35 dB while levels on the contralateral side ranging from -180 to -75° azimuth and -36 to 45° elevation are near the minimum of 10 dB. This indicates that the energy in the contralateral HRTFs is less than 0.5% of that in the ipsilateral HRTFs. Therefore it is likely that the contralateral HRTFs contribute relatively little to the first 12 EFs. A second factor is the likelihood that contralateral sounds excite more vibration modes on the external ear due to diffraction of sound around the head. This results in spectral shapes that are more complicated and consequently more difficult to model. A third possible factor is the smoothing of SCF fine structure associated with the spline fit. The effects of head shadowing and the relative complexity of contralateral HRTFs can be minimized by giv-

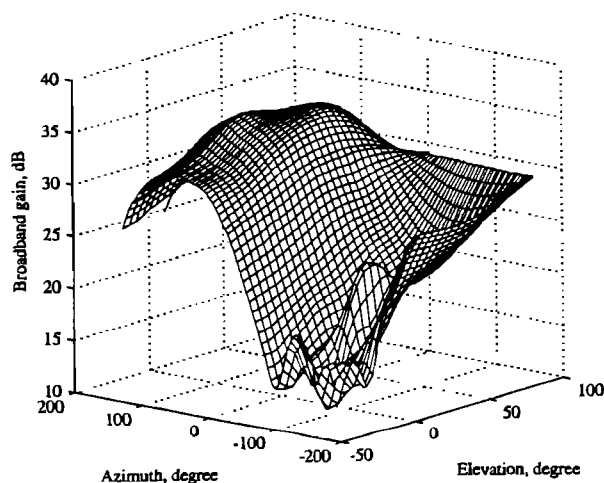


FIG. 7. HRTF Broadband gain as a function of azimuth and elevation.

TABLE IV. Average error in 32 subareas on the sphere for cat.

Azimuth in degrees	Elevation in degrees			
	-36 to 0	0 to 30	30 to 60	60 to 90
-180 to -135	16.4067	0.7370	0.5270	0.5014
-135 to -90	0.5606	0.4991	0.4718	0.3589
-90 to -45	1.2550	2.9693	1.8120	1.8897
-45 to 0	0.5135	1.0907	0.8938	0.6892
0 to 45	0.6622	2.1584	1.4731	0.7854
45 to 90	1.1545	3.6854	1.6455	1.5719
90 to 135	0.4738	0.9128	1.0418	0.8972
135 to 180	15.4327	0.7157	0.6435	0.5764

ing the contralateral HRTFs increased weight during the construction of the covariance matrix and by increasing the number of model components. The degree of smoothing in the spline fitting procedure can be changed by varying the smoothing parameter λ . Note that optimizing model performance is not the focus of the current paper.

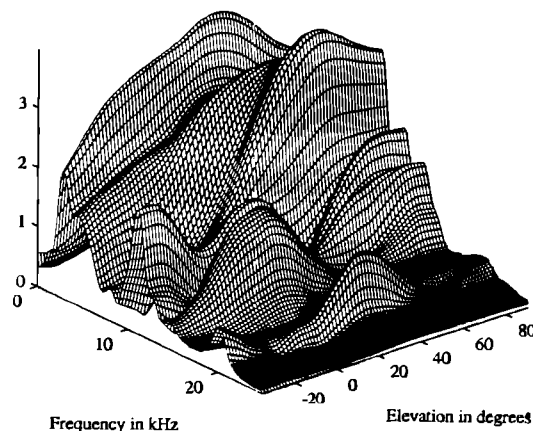
In contrast, the cat mean squared errors over all 1816 locations shown in Table IV do not show evidence of significant lateral difference. Here the poorest performance is in the lower back region. This is most likely because the cat's ears are positioned at the top of the head. The interaural level difference is not significant; head shadowing occurs only in the lower back region.

The SFER model's interpolation capabilities are further illustrated in Fig. 8, which depicts a mesh plot of the HRTF magnitude on the median plane. Here the model is used to obtain the HRTFs at 1.5 degree elevation intervals. Measured HRTFs are not available on this fine scale so quantitative performance evaluation is not possible. However, the qualitative characteristics are those of a well behaved physical system. This figure illustrates the ability of the SFER model to represent the HRTFs as a continuous function of direction.

III. DISCUSSION AND SUMMARY

A. Characteristics of EFs and SCFs

Plots of the KEMAR and cat percent mean-squared error as a function of M for the ipsilateral, contralateral, and entire modeling regions are depicted in Figs. 9 and 10. The ipsilat-

FIG. 8. Mesh plot of the modeled HRTF magnitude in the median plane, $\phi = -36^\circ$ to 90° , $\theta = 0^\circ, 1.5^\circ$ intervals.

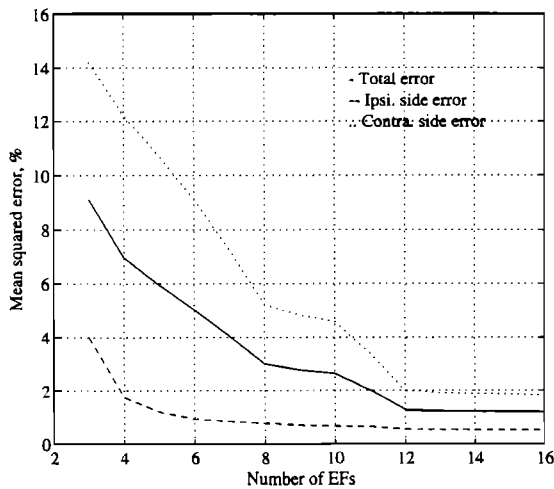


FIG. 9. The total mean-squared error, ipsilateral mean-squared error, and contralateral mean-squared error as a function of number of EFs used in the KEMAR model. Solid line represents the total mean error. Dashed and dotted lines represent the ipsilateral and the contralateral error, respectively.

eral region is defined as azimuths 0° to 180° and elevations from -36° to 90° while the contralateral region spans azimuths 0° to -180° and elevations -36° to 90° . Figure 9 indicates that the total error as a function of the number of EFs for KEMAR flattens out somewhat after eight EFs. Increasing from four to eight EFs reduces the mean error by 4%, while increasing from eight to twelve only yields an additional 1.8% reduction. The total error is not significantly reduced by using more than 12 EFs. A similar phenomena is evident in the percent variance represented (see Table II). The mean and first seven EFs are plotted in Fig. 11, with relative magnitude on the left and relative phase on the right. The labels next to each curve denote the number of the corresponding EF with zero denoting the average HRTF. The magnitudes are offset by 25 dB and the phases by 15 radians for visualization purposes. The EFs are the eigenvectors of

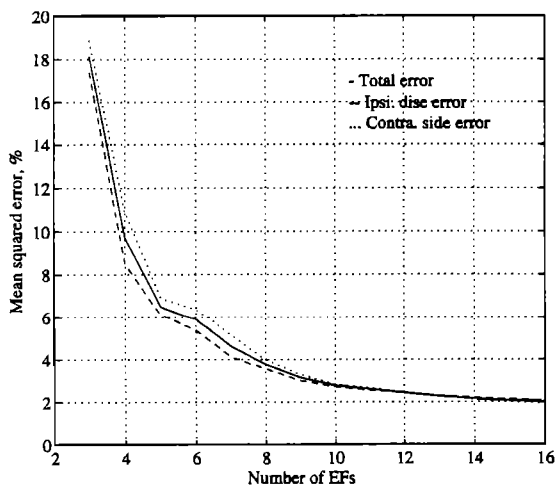


FIG. 10. The total mean-squared error, ipsilateral mean-squared error, and contralateral mean-squared error as a function of number of EFs used in the cat model. Solid line represents the total mean error. Dashed and dotted lines represent the ipsilateral and the contralateral error, respectively.

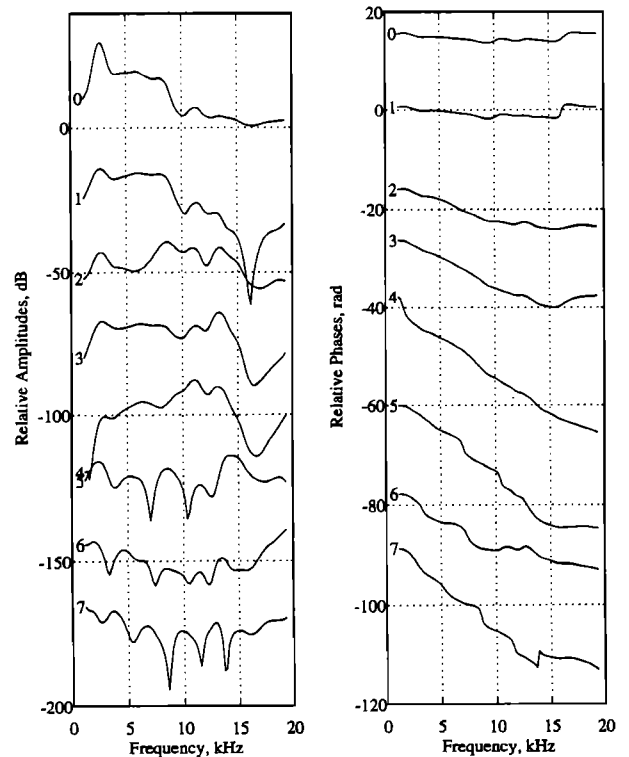


FIG. 11. The mean HRTF and the first through seventh-order EFs for KEMAR. The left panel depicts magnitudes and the right panel phases.

\mathbf{R}_h [see Eq. (3)] and have identical norms, that is, the average magnitudes as a function of frequency are identical. A common linear phase term corresponding to a pure time delay is removed from all phase curves.

Each EF magnitude is characterized by resonant peaks and valleys; the level of relative detail increases with increasing order. Relatively sharp notches are found only in the fifth and higher order EFs. The phase lag increases with increasing order. Plots of the percent error in the contralateral and ipsilateral regions as a function of the number of EFs given in Fig. 9 indicate that the fourth through eighth EFs make a more significant contribution to the contralateral HRTFs than they do to the ipsilateral HRTFs. This is consistent with the increasing level of magnitude detail and phase lag in both the higher order EFs and contralateral HRTFs. The larger phase lags correspond to the increased time extent of the contralateral HRTF impulse responses. Shaw (1979) discovered that resonances at higher frequencies are generally excited by sounds arriving from the front and top directions. Sounds arrive at the contralateral ear from the front, back, and top directions because they are diffracted around the surface of the head. Hence, contralateral sounds excite the largest number of resonances. This behavior is evident in both the HRTF and higher order EF magnitudes.

In contrast, all EFs contribute approximately equally to the contralateral and ipsilateral percent errors in the cat as depicted in Fig. 10. The difference between the cat and KEMAR error performance is most likely associated with the position of the ears on the head. The cat ears are located on the top of the head. Thus contralateral sounds are not shadowed by the cat's head.

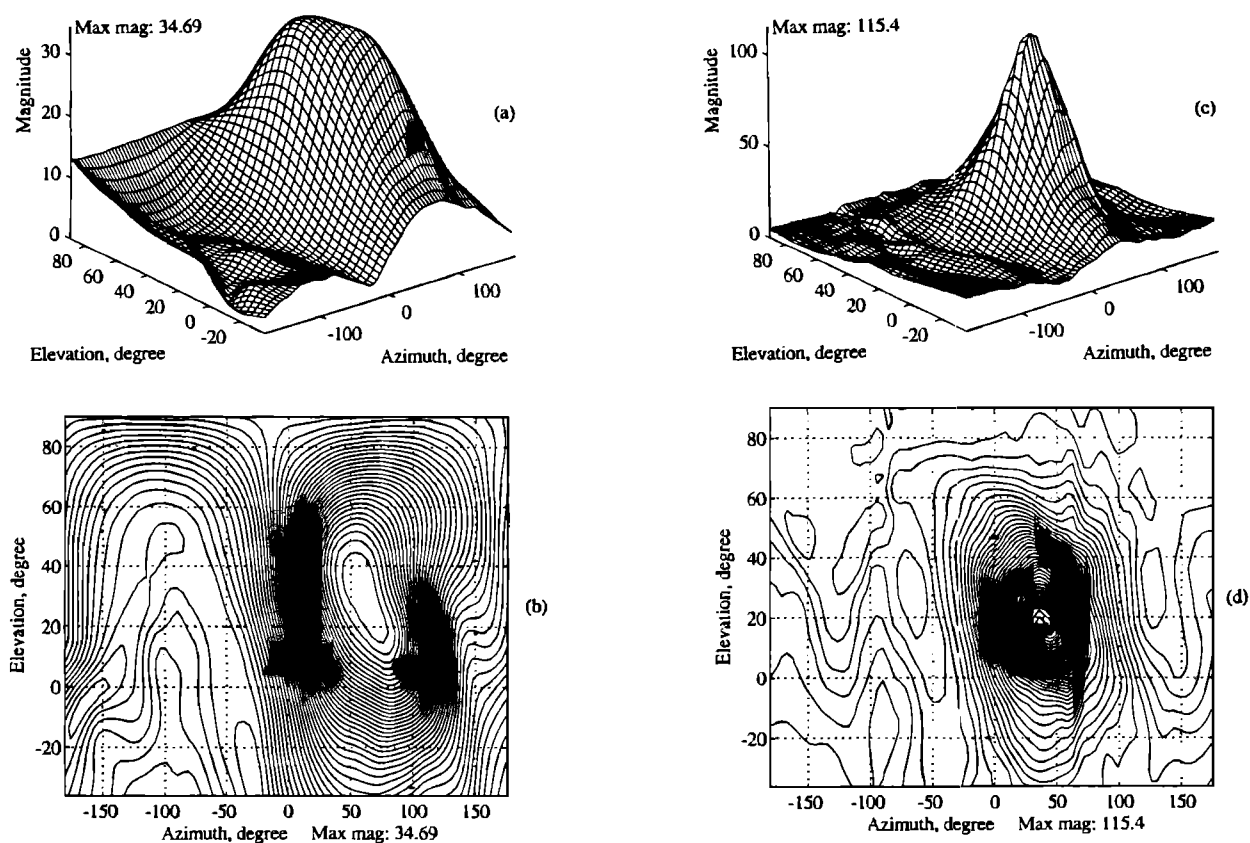


FIG. 12. Mesh and contour plots of the first-order SCF for KEMAR and cat. All plots are magnitudes only. (a) KEMAR's regularized first order SCF mesh plot; (b) KEMAR's regularized first-order SCF contour plot; (c) cat's regularized first-order SCF mesh plot; (d) cat's regularized first-order SCF contour plot.

It is difficult to directly relate specific EF features to features in the HRTF magnitudes because the EFs and SCFs are complex valued. However, it is interesting to note that the collection of resonant frequencies in the EFs include the group of eigen frequencies reported by Shaw (1979) for a mechanical model and average human ear. This is not surprising since the EFs can be expressed as a linear combination of HRTFs. More frequency peaks are observed in the EFs that perhaps account for the spectral details of the HRTFs. The ear canal resonance at 2.7 kHz is dominant in the average EF and is present to a lesser extent in the other EFs. The average and first order EFs have broadband gain over the low and middle frequency ranges (up to 8 kHz). This broadband gain is consistent with Shaw's description of the average human ear response (Shaw, 1979).

Mesh and contour plots of the first order SCF magnitudes of right ear of KEMAR and cat are depicted in Fig. 12. Mesh plots of KEMAR's second through seventh SCF magnitudes are depicted in Fig. 13. The lower order KEMAR SCFs exhibit large magnitude differences between ipsilateral and contralateral regions at lower elevations indicating that head shadowing causes the largest variation in HRTF magnitudes. The fifth through seventh KEMAR SCFs have approximately equal magnitude in both contralateral and ipsilateral regions, consistent with our previous observation that the fifth through seventh EF's make relatively more significant contributions to the contralateral HRTFs than the ipsilateral HRTFs. As with the EFs, the complexity of the SCFs increase as the order increases.

A major peak is evident in the first order KEMAR SCF

[Fig. 12(a) and (b)] at azimuth $\theta=85^\circ$ and elevation $\phi=18^\circ$. This is the direction of maximum broadband gain for KEMAR's right ear. In contrast, the peak location for the cat is at azimuth $\theta=50^\circ$ and elevation $\phi=18^\circ$ [Fig. 12(c) and (d)]. These peak positions are closely related to the opening direction of the pinnae and the position of the ears on the head of KEMAR and the cat.

A rigid head model (see Blauert, 1983) predicts a moderate magnitude elevation when sound is presented from the contralateral side as a consequence of diffraction around the head. This phenomena is reflected in the first order KEMAR SCF at elevations from -36° to 50° , the region where the head interacts most with wave propagation. In contrast, the cat's first order SCF does not reflect significant magnitude elevation for contralateral sounds. The cat's ear is on the top of the head and thus head shadowing effects are not a factor for contralateral sounds. Conjectures of additional relationships between SCF behavior and external ear geometry are posed in Chen (1992).

While there are consistent parallels between certain of Shaw's findings and our representations the EFs described herein bear no resemblance to the PCA basis vectors presented by Kistler and Wightman (1992) or those of Middlebrooks and Green (1992). This is due to fundamental differences between the PCA and SFER models as discussed below.

B. Comparison to PCA models

Both the SFER model and PCA representations (Martens, 1987; Kistler and Wightman, 1992; Middlebrooks and

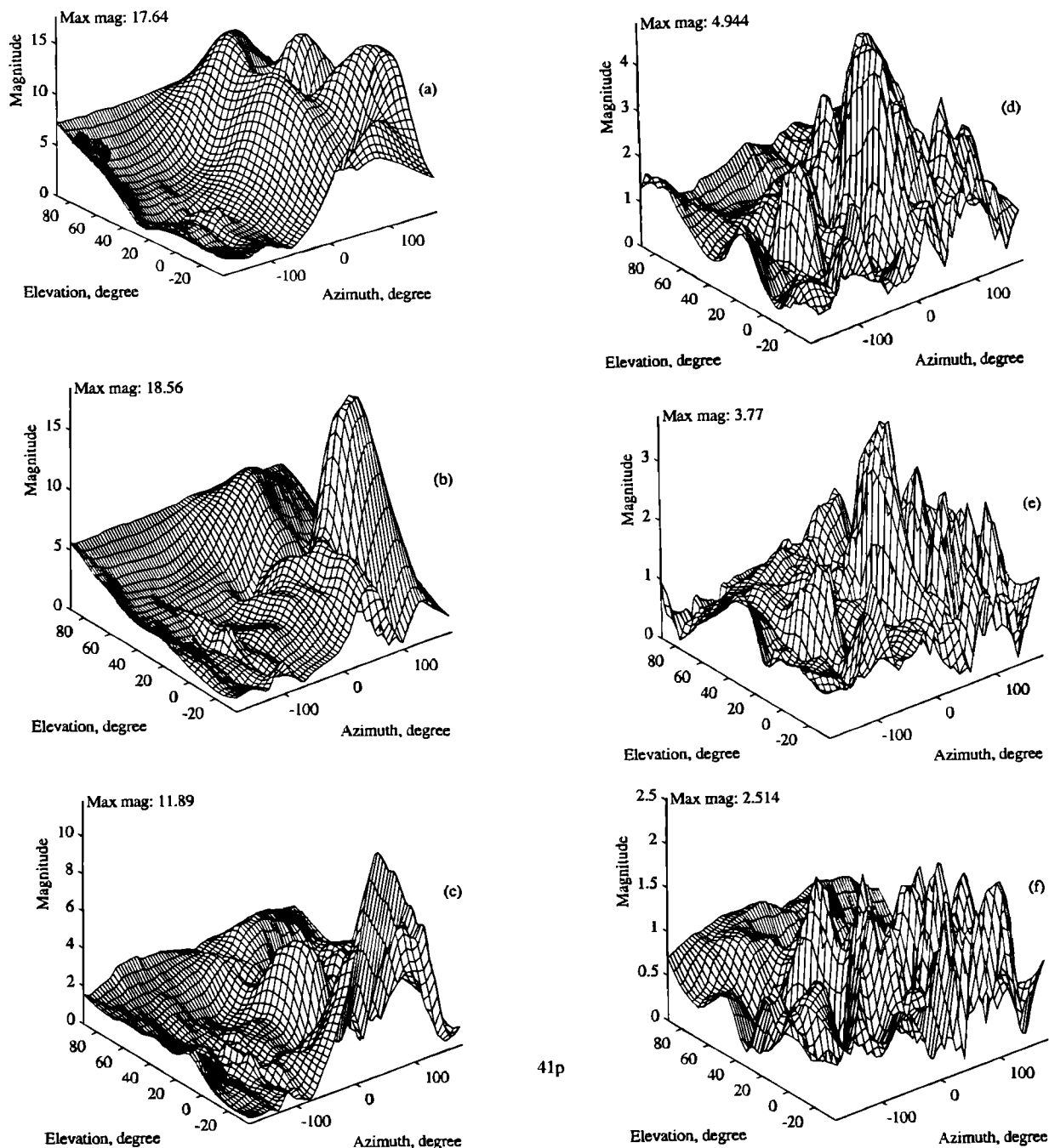


FIG. 13. Magnitude mesh plots of second- through seventh-order SCFs of KEMAR. (a) 2nd order; (b) 3rd order; (c) 4th order; (d) 5th order; (e) 6th order; (f) 7th order.

Green, 1992) involve eigenanalysis of a sample covariance matrix constructed from the measured HRTFs. The SFER model constructs the sample covariance matrix from the complex valued (magnitude and phase) HRTFs while the PCA representations use the logarithm of HRTF magnitude after the directionally independent frequency dependence is removed.

Neither the logarithm nor the magnitude operation commute with either mean computation or sample covariance matrix formation and eigenanalysis. Hence, it is impossible to directly relate the PCA basis vectors and the EFs in the SFER model. Consider the mean HRTF. In the PCA approach the mean is actually the log of the product of all

measured HRTF magnitudes. Subtraction of the mean corresponds to dividing each measured HRTF magnitude by the product of the magnitudes of all HRTFs, resulting in the so-called DTF (Martens, 1987; Kistler and Wightman, 1992; Middlebrooks and Green, 1992). In the SFER model the mean is simply the complex valued average of all measured HRTFs. Note that here subtraction of the mean does not generally remove the direction-independent effects. This is why the ear-canal resonance is evident in each EF while it is not present in the PCA basis vectors.

The PCA representations ignore the measured phase. Kistler and Wightman (1992) used a minimum phase assumption to derive a phase characteristic from the PCA rep-

resentation of the HRTF magnitude. The SFER model avoids minimum phase approximation by directly representing the measured HRTF phase.

C. Summary

A multivariate function approximation approach is used to develop a model for the HRTF magnitude and phase. The KLE separates the frequency and spatial dependence of the HRTFs into a set of EFs and SCFs, respectively. The SCF samples obtained from the measured HRTFs are fit with a thin plate spline. This results in a functional HRTF model that can be evaluated to predict the HRTF at arbitrary azimuth and elevation.

Measured HRTFs from KEMAR and an anesthetized live cat are used to establish both the fidelity and predictive capabilities of the model. Percent mean squared errors on the order of one percent are typical, with larger errors at the contralateral ear of KEMAR. This increased error most likely results from head shadowing effects and can be reduced by increasing the relative weight given to the contralateral HRTFs during model construction. No effort was made in this work to optimize model performance by customizing the HRTF covariance matrix or the smoothing parameter used in fitting the SCF splines.

Further work includes optimizing model performance, behavioral validation, and exploration of the relationships between external ear geometry and model parameters.

ACKNOWLEDGMENTS

The authors express appreciation to our colleagues Drs. A. Musicant, J. Chan, and J. Hind in the Dept. of Neurophysiology at University of Wisconsin for providing the KEMAR and cat ear recordings. We also thank the reviewers and editor for their criticisms which led to a significantly improved presentation. This work was supported by the National Institute of Health NINCDS Grant No. NSR 0116436.

Batteau, D. W. (1967). "The Role of the Pinna in Human Localization," *Proc. R. Soc. London Ser. B* **168**, 158–180.

Batteau, D. W. (1968). *Listening with the Naked Ear: The Neuropsychology of Spatially Oriented Behavior* (Dorsey, Homewood, IL).

Blauert, J. (1969). "Investigations of Directional Hearing in the Median Plane with the Head Immobilized," Ph.D. thesis Technische Hochschule, Aachen.

Blauert, J. (1983). *Spatial Hearing* (MIT, Cambridge, MA).

Brugge, J. F., Chan, J. C. K., Hind, J. E., Musicant, A. D., Poon, P. W. F., and Reale, R. A. (1992). "Neural Coding of Virtual Acoustic Space," *J. Acoust. Soc. Am.* **92**, 2333.

Butler, R. A. (1975). "The Influence of the External Ear and Middle Ear on Auditory Discrimination," in *Handbook of Sensory Physiology, Auditory System*, edited by W. D. Keidel and W. D. Neff (Springer-Verlag, Berlin), Vol. 2.

Carlile, S. (1989). "Directional Properties of the Outer Ear in the Ferret," *Br. J. Audiol.* **23**, 154.

Carlile, S. (1990). "The Auditory Periphery of the Ferret. I: Directional Response Properties and the Pattern of Interaural Level Differences," *J. Acoust. Soc. Am.* **88**, 2180–2195.

Chen, J., Van Veen, B. D., and Hecox, K. E. (1992). "External Ear Transfer Function Modelling: A Beamforming Approach," *J. Acoust. Soc. Am.* **92**, 1933–1944.

Chen, J. (1992). "Auditory Space Modeling and Virtual Auditory Environment Simulation," Ph.D. thesis, University of Wisconsin, Madison.

Craven, P., and Wahba, G. (1979). "Smoothing Noisy Data With Spline Functions: Estimating the Correct Degree of Smoothing by the Method of Generalized Cross Validation," *Numer. Math.* **31**, 377–403.

Genuit, K. (1986). "A Description of the Human Outer Ear Transfer Function by Elements of Communication Theory," *Proceedings of 12th International Congress on Acoustics* (Toronto, Canada), ADSTR. B6–8.

Gu, C. (1989). "Rkpack and Its Applications: Fitting Smoothing Spline Models," Tech. Rep. 857, Department of Statistics, University of Wisconsin, Madison.

Hebrank, J., and Wright, D. (1974). "Spectral Cues Used in the Localization of Sound Source on the Median Plane," *J. Acoust. Soc. Am.* **56**, 1829–1834.

Hiranaka, Y., and Yamasaki, H. (1983). "Envelope Representations of Pinna Impulse Responses Relating to Three-dimensional Localization of Sound Sources," *J. Acoust. Soc. Am.* **73**, 291–296.

Kay, S. (1987). *Modern Spectrum Estimation: Theory and Application* (Prentice-Hall, Englewood Cliffs, NJ).

Kistler, D. J., and Wightman, F. L. (1992). "A Model of Head-related Transfer Functions Based on Principal Components Analysis and Minimum-phase Reconstruction," *J. Acoust. Soc. Am.* **91**, 1637–1647.

Ljung, L. (1987). *System Identification: Theory for the User* (Prentice-Hall, Englewood Cliffs, NJ).

Martens, W. L. (1987). "Principal Components Analysis and Resynthesis of Spectral Cues to Perceived Direction," in *The International Computer Music Conference*, edited by J. Beauchamp (International Computer Music Association, San Francisco, CA, 1987), pp. 274–281.

Mehrgardt, S., and Mellert, V. (1977). "Transformation Characteristics of the External Human Ear," *J. Acoust. Soc. Am.* **61**, 1567–1576.

Middlebrooks, J. C., and Green, D. M. (1990). "Directional Dependence of Interaural Envelope Delays," *J. Acoust. Soc. Am.* **87**, 2149–2162.

Middlebrooks, J. C., and Green, D. M. (1992). "Observations on a Principal Components Analysis of Head-related Transfer Functions," *J. Acoust. Soc. Am.* **92**, 597–599.

Middlebrooks, J. C., Makous, J. C., and Green, D. M. (1989). "Directional Sensitivity of Sound-pressure Levels in the Human Ear Canal," *J. Acoust. Soc. Am.* **86**, 89–108.

Musicant, A. D., Chan, J. C., and Hind, J. E. (1990). "Direction-dependent Spectral Properties of Cat External Ear: New Data and Cross-species Comparisons," *J. Acoust. Soc. Am.* **87**, 757–781.

Oldfield, O. (1984). "Acuity of Sound Localization: a Topograph of Auditory Space. II. Pinna Cues Absent," *Perception* **13**, 601–607.

Papoulis, A. (1991). *Probability, Random Variables, and Stochastic Processes* (McGraw-Hill, New York), 3rd ed.

Pössl, C., Schröter, J., Opitz, M., Divenyi, P. L., and Blauert, J. (1986). "Generation of Binaural Signals for Research and Home Entertainment," *Proc. of 12th Int. Cong. Acoust.*, B1–6.

Pralong, D., Phillips, J. R., and Carlile, S. (1992). "Acoustic Validation of an 'In Ear' System for Recording the Head-related Transfer Functions (HRTFs) in the Human," *J. Acoust. Soc. Am.* **92**, 2334(A).

Reale, R. R., Brugge, J. F., Chan, J. C. K., Musicant, A., Hind, J. E., and Poon, P. W. F. (1993). "Virtual-space Receptive Fields of Neurons in Primary Auditory Cortex of Cat," *Hear. Res.* **73**, 67–84.

Shaw, E. A. G. (1974a). "Acoustic response of external ear replica at various angles of incidence," *J. Acoust. Soc. Am.* **55**, 432.

Shaw, E. A. G. (1974b). "Wave Properties of the Human Ear and Various Physical Models of the Ear," *J. Acoust. Soc. Am. Suppl.* **1** **56**, S3.

Shaw, E. A. G. (1975). "The external Ear: New Knowledge," in *Earmolds and Associated Problems, Scand. Audiol. Suppl.*, edited by S. Daalgaard, **5**, 24–48.

Shaw, E. A. G. (1979). "The Elusive Connection: 1979 Rayleigh Medal Lecture," *Ann. Mtg. Institute of Acoustics (U.K.)* (Southampton, England).

Tikhonov, A. N., and Arsenin, V. Y. (1977). *Solutions of Ill-Posed Problems* (Wiley, New York).

Tobias, J. V. (1972). *Foundations of Modern Auditory Theory* (Academic, New York).

Van Trees, H. L. (1968). *Detection, Estimation, and Modulation Theory* (Wiley, New York).

Wahba, G. (1977). "Practical Approximate Solutions to Linear Operator Equations When the Data Are Noisy," *SIAM J. Numer. Anal.* **14**, 651–667.

Wahba, G. (1984). "Surface Fitting with Scattered, Noisy Data on Euclidean D-spaces and on the Sphere," *Rocky Mountain J. Math.* **14**, 281–299.

Wahba, G. (1990). *Spline Models for Observational Data* (Society for Industrial and Applied Mathematics, Philadelphia, PA).

- Wahba, G., and Wendelberger, J. (1980). "Some New Mathematical Methods for Variational Objective Analysis Using Splines and Cross-validation," *Monthly Weather Rev.* **108**, 1122–1145.
- Watkins, A. J. (1979). "The Monaural Perception of Azimuth: A Synthesis Approach," in *Localization of Sound: Theory and Applications*, edited by R. W. Gatehouse (The Amphora, Groton, CT).
- Wenzel, E. M. (1992). "Issues in the Development of Virtual Acoustic Environments," *J. Acoust. Soc. Am.* **92**, 2332(A).
- Wenzel, E. M., Stone, P. K., Fisher, S. S., and Foster, S. F. (1990). "A System for Three-dimensional Acoustic 'Visualization' in a Virtual Environment Workstation," in *Proceedings of the IEEE Visualization Conference* (San Francisco, CA), pp. 329–337.
- Wenzel, E. M., Wightman, F. L., and Foster, S. H. (1988). "A Virtual Display System for Conveying Three-dimensional Acoustic Information," *Proc. Hum. Fac. Soc.* **32**, 86–90.
- Wightman, F. L., and Kistler, D. J. (1989a). "Headphone Simulation of Free-field Listening: I: Stimulus Synthesis," *J. Acoust. Soc. Am.* **85**, 858–867.
- Wightman, F. L., and Kistler, D. J. (1989b). "Headphone Simulation of Free-field Listening: II: Psychophysical Validation," *J. Acoust. Soc. Am.* **85**, 868–878.
- Wightman, F. L., Kistler, D. J., and Arruda, M. (1992). "Perceptual Consequences of Engineering Compromises in Synthesis of Virtual Auditory Objects," *J. Acoust. Soc. Am.* **92**, 2332(A).
- Wright, D., Hebrank, J. H., and Wilson, B. (1974). "Pinna Reflections as Cues for Localization," *J. Acoust. Soc. Am.* **56**, 957–962.