Efficient Modeling of the Head-Related Transfer Functions

Charles Lueck
Iowa State University
Ames, IA 50011
Email: cdl@iastate.edu

Abstract

Efficient representations of the head-related transfer functions (HRTFs) based on autoregressive moving average (ARMA) modeling are explored in this chapter. A relatively new technique for ARMA order estimation, based on the minimum eigenvalues of a covariance matrix, is used to estimate efficient model structures. A major advantage of this method is that no prior estimation of the model parameters is required. Examples will be given which indicate that the HRTFs are primarily autoregressive (AR) systems.

1 Introduction

The head-related transfer functions (HRTFs) describe the position-dependent transformation in sound pressure which occurs between a source in anechoic space and the eardrums of a listener. HRTF measurements are commonly obtained as a series of time-domain impulse responses using techniques similar to those found in Wightman and Kistler [1, 2], with a pair of responses being measured for each desired source position. Simulation of virtual acoustic environments typically involves finite impulse response (FIR) filtering based on measured HRTF impulse responses. Such implementations are, however, generally not computationally efficient.

In this chapter, autoregressive moving average (ARMA) systems will be employed to identify more efficient representations of the HRTFs. An ARMA system can be described by its input-output relationship, given by the following linear, constant-coefficient difference equation of order (p, q):

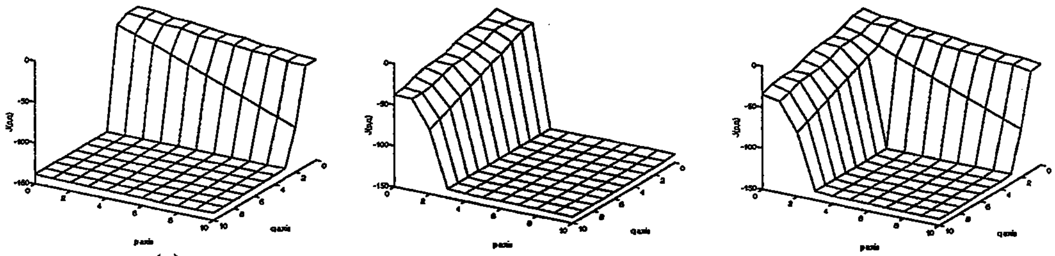
$$\sum_{i=0}^{p} a_i y(n-i) = \sum_{j=0}^{q} b_j x(n-j) \qquad (1)$$

where $x(n)$ is the system excitation and $y(n)$ is the system response.

Given excitation-response records for an unknown system, the system identification problem essentially becomes a two-part process. Firstly, one needs to determine the order (p, q) of the system and, secondly, one must estimate the system parameters (a_k's and $b_k^{\prime}s$). Estimating the order (p, q) will be the primary concern of this chapter. A recently developed technique for model order estimation based on minimum eigenvalues of a covariance matrix will be used to estimate appropriate model orders. Results will be presented which indicate that HRTFs are primarily AR systems for most source positions.

2 Minimum Eigenvalue Model Order Estimation

Assume for the moment that the observed excitation-response data satisfies Eq. (1) exactly for some undetermined orders p and q. Note that for finite-length data records, Eq. (1) can be



(a) (b) (c)

Figure 1: Minimum eigenvalue plots (in dB) for (a) an MA(3) system, (b) an AR(3) system, and (c) an ARMA(3,3) system.

rewritten in matrix form as

$$\begin{bmatrix} y(0) & y(-1) & \dots & y(-p) \\ y(1) & y(0) & \dots & y(1-p) \\ \vdots & \vdots & \vdots & \vdots \\ y(N) & y(N-1) & \dots & y(N-p) \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} x(0) & x(-1) & \dots & x(-q) \\ x(1) & x(0) & \dots & x(1-q) \\ \vdots & \vdots & \vdots & \vdots \\ x(N) & x(N-1) & \dots & x(N-q) \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_q \end{bmatrix} \quad (2)$$

or simply as

$$\mathbf{Y}_p \mathbf{a}_p = \mathbf{X}_q \mathbf{b}_q \quad (3)$$

in which $\mathbf{Y}_p$ is a $(N+1) \times (p+1)$ response matrix, $\mathbf{X}_q$ is a $(N+1) \times (q+1)$ excitation matrix, and $\mathbf{a}_p$ and $\mathbf{b}_q$ are respectively $(p+1) \times 1$ and $(q+1) \times 1$ column parameter vectors. Rearranging Eq. (3) yields

$$\begin{bmatrix} \mathbf{Y}_p & \vdots & -\mathbf{X}_q \end{bmatrix} \begin{bmatrix} \mathbf{a}_p \\ \dots \\ \mathbf{b}_q \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Let the data matrix $\mathbf{D}_{p,q}$ be defined to be

$$\mathbf{D}_{p,q} = \begin{bmatrix} \mathbf{Y}_p & \vdots & -\mathbf{X}_q \end{bmatrix}$$

and the corresponding covariance matrix $\mathbf{R}_{p,q}$ to be

$$\mathbf{R}_{p,q} = \mathbf{D}_{p,q}^T \mathbf{D}_{p,q} . \quad (4)$$

Using eigenvalue decomposition, the covariance matrix can be decomposed into the form

$$\mathbf{R}_{p,q} = \mathcal{Q} \Lambda \mathcal{Q}^T \quad (5)$$

in which the eigenvector matrix $\mathcal{Q}$ has as its columns the set of orthonormal eigenvectors of $\mathbf{R}_{p,q}$ and the diagonal eigenvalue matrix

$$\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_{\min}) \quad (6)$$

has as its elements the corresponding eigenvalues. Typically, the eigenvalues are ordered such that $(\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{\min})$.

It can be shown [3] that if the "true" order of the unknown system is(n_p, n_q) and the system order (p, q) in Eq. (2) is selected such that $p \geq n_p$ and $q \geq n_q$, then the corresponding covariance matrix $\mathbf{R}_{p,q}$ will have at least one zero eigenvalue (the minimum eigenvalue) since there will exist at least one exact solution for $\mathbf{a}_p$ and $\mathbf{b}_q$ in Eq. (4). Thus, the true order of the system can be selected as the lowest order (p, q) at which the minimum eigenvalue is zero.

Typically, however, the observed excitation and response data will not exhibit a perfect ARMA relationship as indicated by (1). In such cases, the minimum eigenvalue will not equal zero once the true system order has been reached but will equal some small, nonzero value. For example, Figure 1 shows typical minimum eigenvalue mesh plots for ARMA(3,3), MA(3), and AR(3) systems. For each system, a 200 coefficient impulse response is obtained to which a low-variance, white noise is added. The excitation is assumed to be a noise-free discrete-time impulse. The covariance matrix is then formed for all model orders (p, q) such that $0 \leq p \leq 10$ and $0 \leq q \leq 10$. For each order (p, q) the minimum eigenvalue of the corresponding covariance matrix is computed and plotted in dB relative to unity, i.e., $10 \times \log_{10}(\lambda_{\min})$. Since noise has been added to the observed system response of all three systems the minimum eigenvalues do not drop to zero once the true model order has been reached, but rather to some small, nonzero value. It can be shown [3] that for large N, the increment in minimum eigenvalues will be upper bounded by a value proportional to the noise variance, although an in-depth discussion of this topic is beyond the scope of this chapter.

3 Application to HRTF Data

This technique can be applied to HRTFs to estimate the appropriate ARMA model orders. For HRTF data, the system excitation $x(n)$ is taken to be the discrete-time impulse, and the observed system response $y(n)$ is taken to be the measured HRTF impulse response, converted to a minimum-phase sequence before applying the order estimation procedure.

Figure 2 shows the minimum eigenvalue plots for left ear HRTFs of a single subject for several positions on the horizontal plane. Here, an azimuth angle of 0° indicates a position directly in front of the subject, negative azimuth angles indicate positions to the subject's left, and positive azimuth angles indicate positions to the subject's right.

Comparing the theoretical mesh plots of Figure 1 to the HRTF plots of Figure 2, it appears as though the HRTFs are primarily AR systems. Indicative of an AR system is the fact that the minimum eigenvalues drop sharply in the p axis direction and slowly in the q axis direction. This effect is observed for the HRTFs shown in Figure 3, and for HRTFs in general. To select ARMA model order estimates, an eigenvalue threshold can be used. The order at which the minimum eigenvalue drops below this threshold is selected as the estimate.

To demonstrate the AR nature of the HRTFs, a -15 dB eigenvalue threshold was used. AR model order estimates were computed using this -15 dB threshold for 450 minimum-phase, left ear HRTFs of a single subject for positions distributed evenly about a spherical shell by considering minimum eigenvalues along the p axis. Figure 4a shows a histogram of the AR model order estimates for all 450 transfer functions. The mean order estimate is 9.8 and the highest order estimate is 14. Similarly, MA model order estimates for the -15 db threshold were also computed for the same 450 HRTFs by examining minimum eigenvalues along the q axis. The results are illustrated by the histogram of Figure 4b. It is clear that the MA model order estimates are significantly higher at this threshold than the AR model order estimates.

At lower thresholds, however, the difference between AR and MA model order estimates is reduced due to the fact that the MA model shows a gradual decrease in minimum eigenvalues

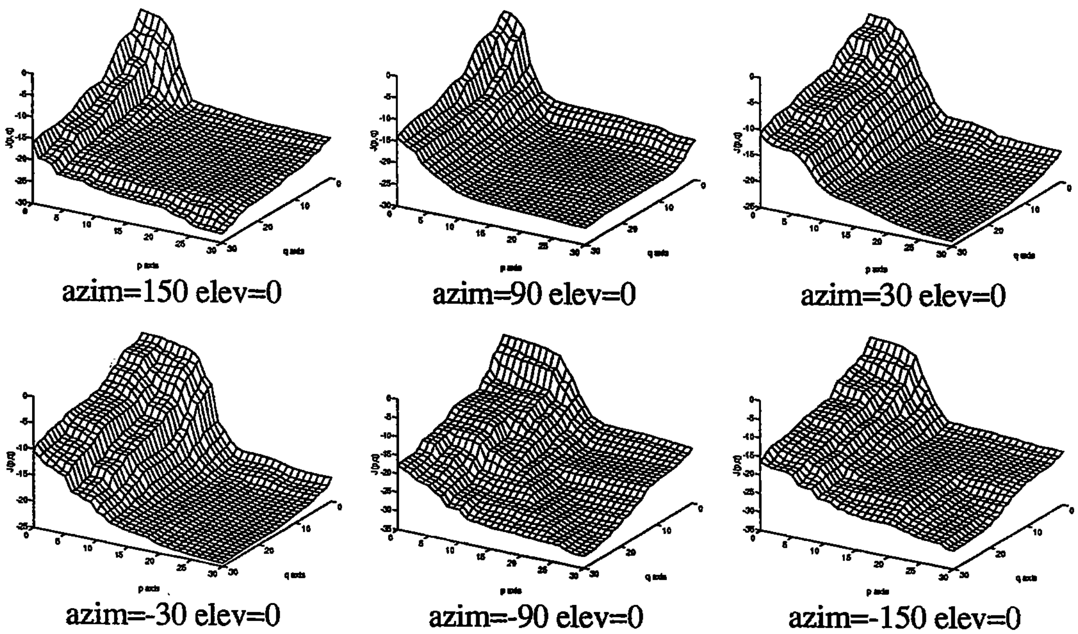


Figure 2: Log-scale minimum eigenvalue ploits for HRTF functions.

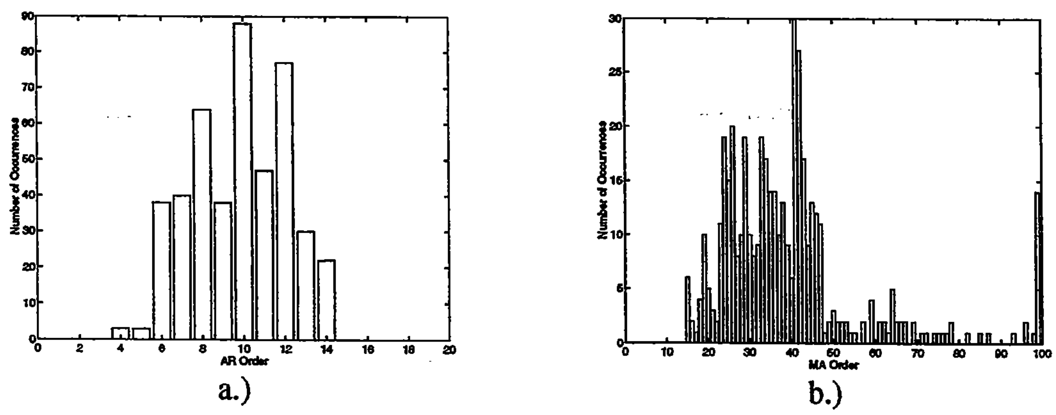


Figure 3: Histograms of −15 dB (a) AR order estimates (b) MA order estimates.

(along the q axis) while the AR model shows little decrease above 15th order (along the p axis). Thus, the AR model appears to have its greatest advantage over the MA model at low model orders. The question arises as to the perceptual contribution of the AR model for sound localization. Lowering the eigenvalue threshold corresponds roughly to lowering the allowable modeling error variance, an objective criteria which may not directly reflect the perceptual performance of the model. Consequently, perceptual listening tests are needed to fully assess the significance of these results.

Conclusions

The use of autoregressive moving average models for efficient representation of the head-related transfer functions has been explored in this chapter. Using a minimum eigenvalue model order estimation technique, the HRTFs have been shown to be primarily AR systems. These results suggest the implementation of HRTFs using autoregressive filters or possibly the use of an autoregressive prefilter for efficient computation, although the validity of such an implementation has yet to be subjectively demonstrated.

Acknowledgments

This work was supported, in part, by a Rockwell International Corporation Fellowship. The author would like to thank Fred Wightman and Doris Kistler at the University of Wisconsin for providing the HRTF data used in this chapter.

References

[1] Wightman, F., and D. Kistler. "Headphone Simulation of Free-Field Listening. I: Stimulus Synthesis." *J. Acoust. Soc. Am.* **85** (1989): 858–867.

[2] Wightman, F., and D. Kistler. "Headphone Simulation of Free-Field Listening. II: Psychophysical Validation." *J. Acoust. Soc. Am.* **85** (1989):868-878.

[3] Liang, G., D. Wilkes, and J. Cadzow. "ARMA Model Order Estimation Based on the Eigenvalues of the Covariance Matrix." *IEEE Transactions on Signal Processing* **41** (1993): 3003–3009.