

# COMMON ACOUSTICAL-POLES/ZEROS MODELING FOR 3D SOUND PROCESSING

M. C. Chen and S. F. Hsieh  
Department of Communications Engineering  
National Chiao Tung University  
Hsinchu, Taiwan 300  
email: sfhsieh@cc.nctu.edu.tw

## ABSTRACT

HRTFs(Head Related Transfer Functions) are a set of transfer functions from different spatial locations to both ears. They constitute an important database in 3-D sound signal processing. However, such an enormous data set makes real-time implementations difficult. Inspired by the resonant characteristic of pinnae, we seek a common-pole/zero model to reduce this vast data set. A Shanks common-poles model is proposed to improve previous Prony work. Based on a linearized least-squares criterion, a novel approach using iterative prefiltering and recursive updating the common poles is proposed and outperforms previous works.

## 1. INTRODUCTION

It is known that head-related transfer functions (HRTFs) convey important cues for human spatial hearing[1]. Hence, the position-dependent HRTFs are widely used in 3-D sound signal processing. Nevertheless, the vast size of the HRTFs data set makes real-time implementation difficult owing to the memory size as well as the computational cost needed in convolution.

Two main types of models, all-zero minimum-phase FIR model [2] and Pole/Zero IIR model [3][4][5] have been used to reduce the huge data. In general, pole/zero model exhibits better performance in tracking the resonant part of HRTFs than the FIR model. Recently, there have been many IIR-related researches such as a logarithmic error measure ARMA formulations [4], and a balanced model truncation method [8]. However, a common-pole model for these two approaches appear to be difficult.

Our work here is based on the Common Acoustic Pole/Zero model (CAPZ), which is inspired by the resonant structure of pinnae[5]. That means all HRTFs of some group share one set of poles (resonant poles) and possess their own zeros individually. This obviously reduces the amount of poles with respect to an individual-Pole/Zero model. Previous work on the CAPZ model uses the Prony method to decide the model parameters [6]. We will extend the common-pole Prony model to the Shanks

method [8]. By formulating a linearized least-squares problem, a novel approach using iterative prefiltering and recursive updating the common poles is proposed and outperforms previous two methods. Computer simulations will also be performed to justify their effectiveness.

## 2. COMMON POLES MODELING

### 2.1 Common Acoustic Pole/Zero (CAPZ) Model

Denote one group of  $k$  HRTFs as

$$h_1(n), h_2(n), \dots, h_k(n)$$

with each represents the impulse response from a specific location to the listener's ear. Since a CAPZ model means that every HRTF in this group shares the same poles and possesses its zeros individually, we can formulate the transfer function of the  $i$ -th synthesized HRTF as

$$\hat{H}_i(z) = \frac{B_i(z)}{A(z)} = \frac{b_{0,i} + b_{1,i}z^{-1} + \dots + b_{q,i}z^{-q}}{1 + a_1z^{-1} + a_2z^{-2} + \dots + a_pz^{-p}}$$

where  $B_i = \{b_{0,i}, \dots, b_{q,i}\}$  is its zero coefficients and  $A = \{a_0, \dots, a_p\}$  is the common pole coefficients. We aim to seek  $A(z)$  and  $B_i(z)$ 's so that they approximate  $H_i(z)$ 's under some criterion.

### 2.2 Prony method

Define the error between the  $i$ -th actual HRTF and the synthesized HRTF from the CAPZ model as

$$E_i(z) = H_i(z) - \hat{H}_i(z) = H_i(z) - \frac{B_i(z)}{A(z)}$$

However, it is well known to be difficult in finding pole/zero coefficients that directly minimizes the squared error. As a remedy, we replace  $E_i(z)$  with the filtered error,  $\tilde{E}_i(z)$ , defined as

$$\tilde{E}_i(z) = A(z)E_i(z) = A(z)H_i(z) - B_i(z)$$

Since the pole coefficients are shared by every HRTF in one group, the determination of common pole coefficients should be based on the minimization of all of the filtered

errors in this group. Hence, we define a *Group Filtered Error* as

$$GFE = \sum_{i=1}^k \sum_{n=0}^{N-1} |\tilde{e}_i(n)|^2$$

where  $\tilde{e}_i(n)$  is inverse Z transform of  $\tilde{E}_i(z)$  and  $N$  is the length of  $\tilde{e}_i(n)$ .

To find the poles' coefficients  $A$  and the zeros coefficients  $B$ 's that minimize the group filtered error, we have

$$\begin{aligned} & \min_A \left( \sum_{i=1}^k \sum_{n=0}^{N-1} |\tilde{e}_i(n)|^2 \right) \\ &= \min_A \left( \sum_{i=1}^k \sum_{n=0}^{N-1} |a(n) * h_i(n) - b_i(n)|^2 \right) \\ &= \min_A \left( \sum_{i=1}^k \sum_{n=0}^{N-1} \left| h_i(n) - \left( - \sum_{k=1}^p a(k) h_i(n-k) + b_i(n) \right) \right|^2 \right) \end{aligned}$$

Rewrite it in matrix form, we have a set of least squares equations:

$$\begin{bmatrix} I & H_{o,1} \\ 0 & H_1 \end{bmatrix} \begin{bmatrix} b_1 \\ -a \end{bmatrix} \equiv \begin{bmatrix} r_{o,1} \\ r_1 \end{bmatrix}$$

$$\vdots$$

$$\begin{bmatrix} I & H_{o,k} \\ 0 & H_k \end{bmatrix} \begin{bmatrix} b_k \\ -a \end{bmatrix} \equiv \begin{bmatrix} r_{o,k} \\ r_k \end{bmatrix}$$

where  $I$  is an identity matrix;

$$g_{o,i} = [h_i(0), \dots, h_i(q)]^T$$

$$g_i = [h_i(q+1), \dots, h_i(N-1)]^T, \quad i = 1, 2, \dots, k$$

and  $H_{o,i}$  and  $H_i$  are the convolution matrices by cascading the impulse response of  $h_i(n)$ :

$$H_{o,i} = \begin{bmatrix} 0 & 0 \\ h_i(0) & 0 \\ h_i(1) & h_i(0) \\ \vdots & \vdots & \ddots \\ h_i(q-1) & h_i(q-2) & \vdots & h_i(q-p) \end{bmatrix}$$

$$H_i = \begin{bmatrix} h_i(0) & \dots & h_i(q-p-1) \\ \vdots & \ddots & \vdots \\ h_i(0) & \dots & h_i(N-1-p) \end{bmatrix}$$

The Prony method computes the common pole coefficients  $a$  from the normal equation:

$$a = -(H^T H)^{-1} H^T R \quad \text{and} \quad b_i = H_{o,i} a + g_{o,i}$$

where

$$R = \begin{bmatrix} r_1 \\ \vdots \\ r_k \end{bmatrix}, \quad H = \begin{bmatrix} H_1 \\ \vdots \\ H_k \end{bmatrix}$$

## 2.3 Shanks method

The difference between the Prony and Shanks methods lies in the determination of individual zero coefficients. For any specific zero coefficients, in Prony method, the  $i$ -th zero coefficients are the optimal solution to minimize *Group Filtered Error*.

However, minimizing the filtered error, after all, is a compromised approach. Therefore we propose to use the Shanks method in zero determination for our CAPZ model. After using the Prony method to determine the poles, the Shanks method returns to our initial goal of minimizing the group error. It should be noted that the group error here is now defined as  $\sum_{i=1}^k \sum_{n=1}^{N-1} e_i(n)$  where  $e_i(n)$  is the inverse Z transform of  $E_i(z)$ .

Hence, the zero determination for some specific HRTF in one group, for instance, the  $i$ -th HRTF can be explained as following plot.

Therefore, we have following equations in obtaining the zeros:

$$\begin{bmatrix} g(0) & 0 & \dots & 0 \\ g(1) & g(0) & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \dots & g(0) \end{bmatrix} b_i \equiv \begin{bmatrix} h_i(0) \\ h_i(1) \\ \vdots \\ \vdots \end{bmatrix}$$

or,

$$Gb_i \equiv h_i$$

where  $b_i = [b_{0,i}, b_{1,i}, \dots]^T$ , and  $g(n)$  is the impulse response of the inverse system  $1/A(z)$ .

## 2.4 Iterative Prefiltering method

Rewrite the impulse error  $E_i(z) = H_i(z) - \hat{H}_i(z)$  as

$$E_i(z) = \frac{A(z)H_i(z) - B_i(z)}{A(z)}, \quad i = 1, 2, \dots, k$$

Suppose the denominator of  $E_i(z)$  at iteration  $j$  is known as

$$A(z) = A^{(j)}(z)$$

Then we can find an optimum  $A(z) = A^{(j+1)}(z)$  and  $B(z) = B^{(j+1)}(z)$  in the numerator of  $E_i(z)$  so that  $E_i(z)$  is minimized. That is, an *iterative linearized least-squares problem* can be written as follows:

$$E_i^{(j+1)}(z) = \frac{A^{(j+1)}(z)H_i(z) - B_i^{(j+1)}(z)}{A^{(j)}(z)}, \quad i = 1, \dots, k$$

Define

$$G^{(j)}(z) = \frac{1}{A^{(j)}(z)} \text{ and } F_i^{(j)}(z) = H_i(z)G^{(j)}(z)$$

With inverse Z transforms  $g^{(j)}(n)$  and  $f_i^{(j)}(n)$ , respectively. After some algebraic manipulations, we can formulate the iterative prefiltering algorithm as

$$\begin{bmatrix} F_1^{(j)} & G^{(j)} & 0 & \dots & 0 \\ F_2^{(j)} & 0 & G^{(j)} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ F_k^{(j)} & 0 & 0 & \dots & G^{(j)} \end{bmatrix} \begin{bmatrix} a^{(j+1)} \\ -b_1^{(j+1)} \\ \vdots \\ -b_k^{(j+1)} \end{bmatrix} \equiv \begin{bmatrix} -f_1^{(j)} \\ -f_2^{(j)} \\ \vdots \\ -f_k^{(j)} \end{bmatrix}$$

The iterative prefiltering algorithm can start with the poles coefficients  $A(z)$  solved from the Prony method. From our computer simulations, it converges quickly (no more than 5 iterations) to an optimum pole/zero solution.

### 3. COMPUTER SIMULATIONS

Here we will use the HRTF data set measured by University of Wisconsin. Before starting to simulate, the HRTFs were normalized and the initial delays are removed. In our simulations, we will use 15 HRTFs with the azimuth 90 degree and elevations ranging from +90 to -50 degree with 10-degree step. The CAPZ model we simulate here has  $q=20$  zeros and  $p=3$  common poles. To see the synthesis performance of the  $i$ -th HRTF on the

CAPZ model, we define an error index as  $\epsilon_i = \frac{\|h_i - \hat{h}_i\|}{\|h_i\|}$

where  $h_i$  and  $\hat{h}_i$  are the  $i$ -th actual and synthesized HRTFs, respectively.

We first examine if the IIR model is suitable for HRTF modeling. Take the HRTF at azimuth 90 degree and elevation 0 degree as an example. Fig. 1 shows the impulse responses for one specific HRTF using 3 pole/zero models, from which we can see that the iterative prefiltering method has the smallest modeling error and the Shanks method's performance is better than the Prony method as we expect. Fig. 2 compares the error indices of 3 IIR and all-zero FIR models, under the assumption of the same filter order. Clearly, the iterative prefiltering can use the least number of zeros to closely approximate the impulse response of the HRTF. From Fig 1 and 2, we can see that a pole/zero model is indeed suitable for HRTF modeling.

Now we use 15 HRTFs to show that the CAPZ model is appropriate for modeling this group of HRTFs. The iterative filtering method is used and it needs about 5 iterations for the coefficients to converge. We focus on one specific HRTF and compare the difference between a common-pole model (over a group of 15 HRTFs) and an individual pole/zero model. That is, a CAPZ model only needs one set of common poles and 15 sets of zeros, while

an individual pole/zero model needs 15 sets of poles and 15 sets of zeros. From Fig 3, we can see that the impulse response of the common-pole/zero model is very close to that of the individual pole/zero model as well as the actual HRTF.

When all of the HRTFs within one group are considered, Table 1 compares the group error indices (sum of squared errors for all synthesized impulse responses) for 3 IIR models. We note that iterative prefiltering method is the best one, the Shanks follows, and the Prony method is the worst one.

### 4. CONCLUSION

We have shown that the CAPZ model is an efficient model for one group of HRTFs, which is consistent with human hearing physiology. Extending a previous Prony model, we have proposed the Shanks CAPZ method that seeks a better solution of zero coefficients than the Prony method. In addition, the iterative prefiltering method, by recursively updating the common poles, outperforms them and is very promising for 3D sound signal processing.

### 5. REFERENCES

- [1] R. Durand and Begault, *3-D Sound for virtual reality and multimedia*, AP Professional, 1994.
- [2] J. J. Huang, *Realization of Low-ordered 3D Sound System*, Master thesis, National Chiao Tung University, 1996.
- [3] C. P. Brown and R. O. Duda, "An efficient HRTF model for 3-D sound," *IEEE WAAPAA'97 Workshop Committee* Oct., Mohonk Mountain House New Paltz, New York., 1997.
- [4] M. A. Blommer and G. H. Wakefield, "Pole-zero approximations for head-related transfer functions using a logarithmic error criterion," *IEEE Trans. on Speech and Audio Processing*, Vol. 5, No. 3, May 1997.
- [5] R. O. Duda, "Modeling head related transfer functions," *Conference Record of Asilomar Conference on Signals, Systems and Computers*, IEEE Computer Society Press, Vol 2, p 996-1000, 1993.
- [6] Y. Haneda, S. Makino, Y. Kaneda, and N. Kitawaki, "Common-acoustic-pole and zero modeling of head-related transfer functions," *IEEE Trans. on Speech and Audio Processing*, Vol. 7, No. 2, p 188-195, March 1999.
- [7] J. Mackenzie, J. Huopaniemi, V. Valimaki, and I. Kale, "Low-order modeling of head-related transfer functions using balanced model truncation," *IEEE Signal Processing Letters*, Vol. 4, No. 2, p 39-41, February 1997.

[8] M. C. Chen, *Common-pole/zero model on HRTFs data sets*, Master thesis, National Chiao Tung University, 1998.

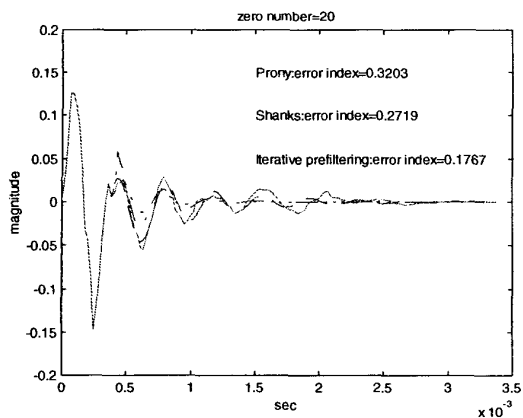


Figure 1. Comparison of impulse responses for three IIR models.

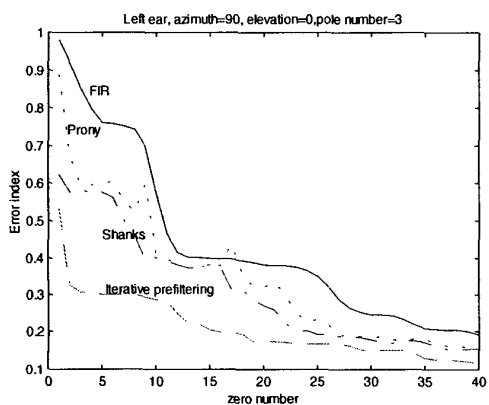


Figure 2. Error index comparison for FIR and three IIR models.

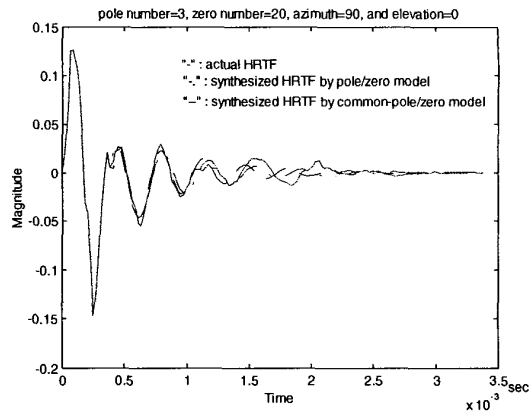


Figure 3. Comparison of FIR and three IIR models.

Table 1. Group error index comparison of individual- and common-pole/zero models.

	$\left\{ \frac{B_i(z)}{A_i(z)} \right\}$	$\left\{ \frac{B_i(z)}{A(z)} \right\}$
Prony	5.187	5.327
Shanks	4.851	5.167
Iterative prefiltering	2.949	3.213