J. Acoust. Soc. Jpn. (E) 15, 4 (1994)

ACOUSTICAL LETTER

Common acoustical poles independent of sound directions and modeling of head-related transfer functions

Yoichi Haneda, Shoji Makino, and Yutaka Kaneda
NTT Human Interface Laboratories,
3-9-11, Midori-cho, Musashino, 180 Japan
(*Received 22 September 1993*)

Keywords: Head-related transfer functions, Pole/zero model, Modeling, Sound directions
PACS number: 43. 66. Ba, Qp

1. Introduction

A head-related transfer function (HRTF) represents the transmission characteristics from a sound-source to an ear of a listener. Modeling of HRTFs is important for such applications as the reproduction of virtual sound images. A conventional all-zero model (FIR filter) or pole/zero model (IIR filter) is usually used for modeling the HRTFs, but all the parameters of the conventional models change when the source direction changes. In this paper, we first extract the common factors from HRTFs corresponding to the different sound directions. Then, we model the HRTFs with the estimated common factors and individual parameters, which depend on the sound directions, as shown in Fig. 1.

2. Common-acoustical-pole and zero model

HRTFs have resonance systems composed of the pinna and ear canal.[1] The poles in these resonance systems can be considered as common to all HRTFs corresponding to the different sound directions. Therefore, the common-acoustical-pole and zero model (CAPZ model)[2] can be used for modeling the HRTFs. The CAPZ model represents the acoustic transfer function by common acoustical poles estimated from multiple transfer functions and different zeros. In this paper, the HRTFs on the horizontal plane are considered as shown in Fig. 2.

Now, the HRTF $G(\theta, z)$ (where $G(\theta, z)=G_L(\theta, z)$ or $G_R(\theta, z)$ as shown in Fig. 2) from a source located in direction θ at distance r to the left or right ear can be expressed by the CAPZ model

$$G(\theta, z)=\frac{B(\theta, z)}{A(z)}=\frac{C\prod_{n=1}^{Q}(1-q_n(\theta)z^{-1})}{\prod_{n=1}^{P}(1-p_n z^{-1})}=\frac{\sum_{n=0}^{Q} b_n(\theta)z^{-n}}{1+\sum_{n=1}^{P} a_n z^{-n}}. \quad (1)$$

Here, p_n denote common acoustical poles, which are independent of the sound directions θ, $q_n(\theta)$ denote zeros, which depend on the sound directions θ, a_n denote common AR coefficients corresponding to the common acoustical poles p_n, $b_n(\theta)$ denote MA coefficient corresponding to zeros $q_n(\theta)$, and P and Q are the orders of poles and zeros, respectively. The distance r between the source and the listener is set constant.

3. Measurement of the HRTFs

The impulse responses of the HRTFs on the horizontal plane were measured in an anechoic room using a dummy head (B & K: TYPE4128). Sound direction θ was set in intervals of 15 degrees. The frequency band was 100 Hz to 20 kHz, and the sampling frequency was 48 kHz. Figure 3 shows the frequency responses of the measured HRTFs of the left ear from 0 to 330 degrees every 30 degrees. The peaks, for example, at 3 kHz, 9 kHz, and 13 kHz are common to all HRTFs, so they can be estimated as common acoustical poles.

4. Estimation of the common acoustical poles

The common acoustical poles p_n are estimated as the common AR coefficients a_n that minimize the following index I:

$$I=\sum_{i=1}^{M}\sum_{k=0}^{\infty}(\epsilon_{eq}(\theta_i, k))^2. \quad (2)$$

Here, $\epsilon_{eq}(\theta_i, k)$ is the equation error[3] for each direction θ_i, given by

$$\epsilon_{eq}(\theta_i, k)=g(\theta_i, k)-\sum_{n=1}^{P} a_n g(\theta_i, k-n)-\sum_{n=0}^{Q} b_n(\theta_i)\delta(k-n), \quad (3)$$

$$(\delta(k)=1\ (k=0),\quad \delta(k)=0\ (k\neq 0).)$$

and $g(\theta_i, k)$ is the measured impulse response for each direction θ_i $(i=1, 2, \ldots, M)$.

P and Q were selected to minimize the sum of P and Q that satisfies the following modeling error $E(\theta_i)$ of -25 dB.

$$E(\theta_i)=10\log_{10}\left(\sum_{k=0}^{N}(\epsilon_{out}(\theta_i, k))^2\Big/\sum_{k=0}^{N}(g(\theta_i, k))^2\right) \quad (4)$$

Here, $\epsilon_{out}(\theta_i, k)$ is the output error between the measured impulse response $g(\theta_i, k)$ and model impulse response $y(\theta_i, k)$.

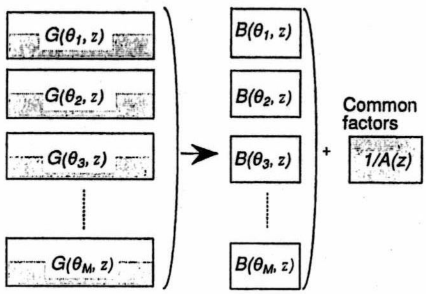

Fig. 1 Effective modeling of multiple head-related transfer functions (HRTFs).

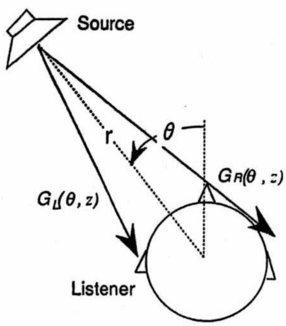

Fig. 2 Head-related transfer functions between a source and a listener's ears.

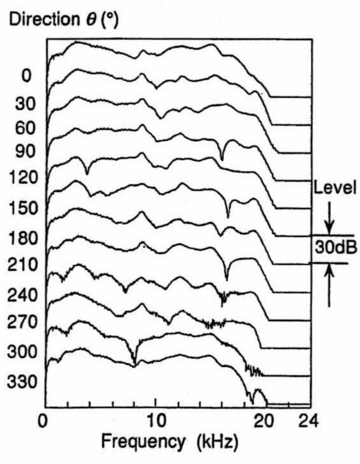

Fig. 3 Frequency responses of the measured HRTFs of the left ear.

$$y(\theta_i, k) = \sum_{n=1}^{P} a_n y(\theta_i, k-n) - \sum_{n=0}^{Q} b_n(\theta_i)\delta(k-n) \quad (5)$$

Finally, we chose $P=20$ and $Q=40$.

Figure 4 shows the frequency response of the transfer function with common AR coefficients that were estimated with 12 HRTFs of the left ear corresponding to the sound directions from 0 to 330 degrees every 30 degrees. Common peaks at 3, 9, and 13 kHz shown in Fig. 3 are extracted as shown in Fig. 4. These peaks correspond to the resonances of the outer ear.

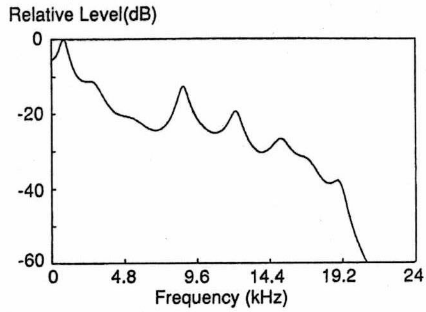

Fig. 4 Frequency response of the transfer function with the estimated common AR coefficients.

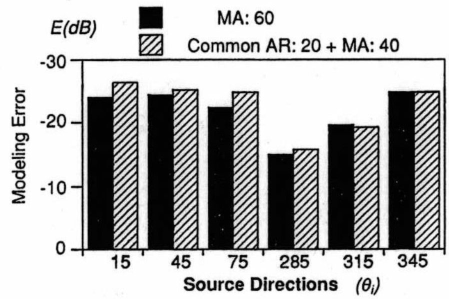

Fig. 5 Modeling errors of the conventional all-zero model (60 MA coefficients, black bars) and the common-acoustical-pole and zero model (20 common AR coefficients +40 variable MA coefficients, shaded bars) for HRTFs.

5. Evaluation of the CAPZ modeling of HRTFs

The effectiveness of the CAPZ modeling of HRTFs was evaluated by comparing the modeling error to the conventional all-zero model having the same number of coefficients. Since the CAPZ model had 20 common AR coefficients and 40 MA coefficients, the conventional all-zero model had 60 MA coefficients. The HRTFs of the sound directions 15, 45, 75, 285, 315, and 345 degrees, which were not used for estimating common AR coefficients, were used for the evaluation.

Figure 5 shows the modeling errors of both models. The modeling errors of the proposed model are as good as those of the conventional all-zero model, even though it uses a smaller number of variable parameters. These results confirm that the common acoustical poles can be estimated from HRTFs corresponding to the different directions, and that the CAPZ model successfully

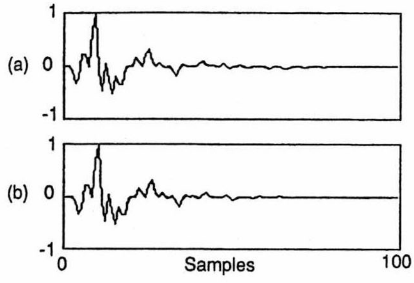

Fig. 6 Impulse responses of (a) the actual HRTF and (b) the common-acoustical-pole and zero modeled HRTF at 15 degrees.

represents the HRTFs. Figure 6 shows the measured impulse response and the CAPZ modeled impulse response at 15 degrees. Both impulse responses are equivalent within a modeling error of −27 dB.

6. Conclusion

The common acoustical poles of HRTFs have been estimated from HRTFs corresponding to the different sound directions. The estimated poles correspond to the resonances of the outer ear. When HRTFs are modeled by the common-acoustical-pole and zero model, the differences in the HRTFs for the different sound directions can be expressed by the change in zeros that are free from the effect of the common acoustical poles.

References

1) J. Blauert, *Spatial Hearing* (M.I.T. Press, Cambridge, 1983).
2) Y. Haneda, S. Makino, and Y. Kaneda, "Modeling a room transfer function using common acoustical poles," Proc. 1992 Int. Conf. Acoust. Speech Signal Process. Vol. 2, pp. 289–292 (1992).
3) L. Ljung and T. Söderström, *Theory and Practice of Recursive Identification* (M.I.T. Press, Cambridge, 1983).