Principal Component Analysis of Pinna's Structural Details on HRTF

Na QI[1], Xuefei FENG

Abstract. Principal component analysis (PCA) was used to reduce dimension and simplify the analysis of HRTF amplitude spectra of dummy head installed with different artificial pinna model. In addition to no-pinna case, 4 kinds of structural model of pinna were discussed. Directional mean of the corresponding frequency bands and the basic functions of spectral shape of HRTF were given to illustrate the effect of different pinna details act on HRTF. It proves, besides the no-pinna case, No.2 pinna structure (with connected cymba conchae and cavitas conchae), No.4 pinna structure(with thickened antitragus) affect the HRTF most obviously in the horizontal plane. However, the effect of helix alone on HRTF is rather little. As a whole, different pinna structural detail has different effect on HRTF amplitude spectrum.

Keywords: *pinna structure; HRTF; principal component analysis*

1.1 Introduction

Pinna plays an important role in sound transmission from source to eardrum. The sound spectrum features caused by pinna reflection and scattering is a significant cue for localization, which appear in the spectrum cues of Head-related transfer functions (HRTF) [1]. As one of the highly personalized organ in the human body, individualized pinna structure details will affect the sound transmission to a certain extent. So, it is necessary to analysis the spectrum cues of HRTF caused by the details of pinna structure. Even ignoring these individualized features, as a function of a multi-variable, the amount of data contained by the HRTFs is still large. Principal component analysis (PCA) is an effective method that removes the correlations among data, thereby reducing data dimensionality. To date, PCA has been applied to HRTFs and head-related impulse responses (HRIRs) in order to reduce dimension and simplify the analysis[2] [6] [7]. In this paper, it is used to analyse the HRTF amplitude spectrum of

---

[1] Na QI

Communication Acoustics Lab, Communication University of China, Beijing 100024, China
e-mail: qina@cuc.edu.cn

dummy head installed with 4 kinds of pinna model and an additional no-pinna case, and the influence of different pinna structure on HRTF is also discussed.

1.2 Artificial pinna and HRTF Measurement

1.2.1 Design of artifical pinna models

Based on the morphological investigation of Chinese adult pinna structure[3],morphological details of pinna were classified, and the most common pinna structure was obtained on the basis of statistical results, i.e. the so-called "typical pinna". An artificial "typical pinna" model was designed, as shown in Fig.1.1. The "typical pinna" is the basis of designing artificial pinna models with different structure details. As an important localization cues of the HRTF, pinna shape, size, structure are closely related to the individualized HRTF[3][4]. Listening test shows that pinna with blocked cavum such as concha、fossa helicis、 fossa triangularis and groove between tragus and antitragus does affect the subjective localization of a sound source in some extent [4].Besides the "typical pinna" and no-pinna case, 3 other kinds of artificial pinna model are listed in table.1.1, which are obtained by manually modifying certain details of the "typical pinna" structure model. The "typical pinna" model is labeled as No.1 pinna structure. Photo of the first four kinds of pinna model is given in Fig.1.2. This paper analyzes the HRTF of dummy head with these 5 kinds of artifical pinna model mentioned above.

1.2.2 HRTF measurement with different pinna structure

A dummy head was used in measuring HRTF, which was installed with the above 5 kinds of pinna model. It was designed according to the national code of China GB/T2428-1998 in the Communication Acoustic laboratory of Communication University of China[5],As shown in Fig.1.1.

A spherical coordinate with the head center as the origin was used in the measurement and analysis. A sound source position was specified by(r, θ, φ) to denote the distance relative to the origin, azimuth, and elevation, respectively. The azimuth ranges from 0° to 360° with $\theta = 0^o$ and 90° corresponding to the right and front directions, respectively. The elevation ranges from -90° to 90° with $\varphi = 0^o$ and 90° corresponding to the horizontal and the above respectively.

The loudspeaker was placed in the front of the dummy head, 1.2 meters above the ground. Measurements were made with 45° azimuth interval uniformly in the horizontal plane. 5 kinds of artificial pinna model were installed on the dummy head and measured in turn. First, an 8191-point MLS signal with sampling frequency 44.1 kHz was generated by computer. Then, an 8-period repeated MLS signal was fed to speaker through the computer sound card and amplifier. The signal was picked up by a pair of miniature microphones (DPA 4060BM) fixed on the entrance of the blocked ear canal. The microphone outputs were then passed though microphone conditioning amplifier, sound card and a computer. After deconvolution with matlab program by computer, the head-related impulse response(HRIR) was obtained, and it was windowed to remove the room reflections and equalized to compensate for the loudspeaker and microphone transfer functions. Transforming the HRIR into frequency domain, a 256 points HRTF was obtained.
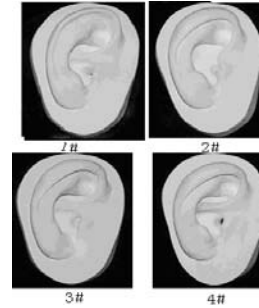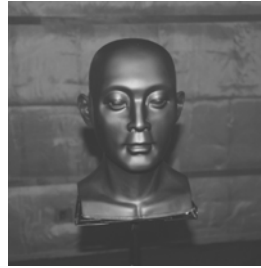

Fig.1.1 picture of artificial "typical" pinna model and dummy head

Fig.1.2 Pinna structure model

Table. 1.1 The structural details and first three basis functions' percentage of energy of 5 kinds of pinna

pinna structure classification		First n basis vectors' percentage of energy		
Label	Structural feature	1	2	3
No.1	the "typical pinna"	91.4%	96.9%	98.6%
No.2	Cymba conchae and cavitas conchae is connected	91.7%	95.5%	98.1%
No.3	Helix is minified	88.8%	96.3%	98.7%
No.4	Antitragus is thickened	90.5%	97.6%	99.1%
No.5	no-pinna	98.2%	99.3%	99.8%

1.3 principal component analysis of HRTF

1.3.1 Principle

HRTF is defined as a pressure ratio

$$H(r,\theta,\varphi,f) = \frac{P(r,\theta,\varphi,f)}{P_o(r,f)} \qquad (1.1)$$

Where f is the frequency, $P(r,\theta,\varphi,f)$ denotes the sound pressure at the left or right ear, and $P_o(r,f)$ denotes the sound pressure at the origin with the head absent. So, HRTF is a multi-variable function. For the convenience of operation it can be regarded as a continuous function of direction and frequency in the far field. Furthermore, it can be represented by a weighted combination of a small set of basis functions, as shown in Eq.(2)[7]. It's an effective method to reduce data size and simplify the calculation.

$$\mathbf{H}(\theta,\varphi,f) = \sum_q \mathbf{d}_q(f)\mathbf{w}_q(\theta,\varphi) \qquad (1.2)$$

Where $\mathbf{d}_q(f)$ denotes the basis function (or basis vector) of HRTF spectra, and $\mathbf{w}_q(\theta,\varphi)$ denotes the weighted coefficient related to the spatial direction.

The HRTF amplitude spectra measured from M directions can be regarded as M column vectors of length N, thus HRTF from M azimuth angles can be treated as a matrix with N rows and M columns. Each row $\mathbf{H}(\theta_i)$ can be divided into:

$$\mathbf{H}(\theta_i) = \sum_q \mathbf{d}_q\mathbf{w}_q(\theta_i) + \mathbf{H}_{av} \qquad (1.3)$$

Where $\mathbf{d}_q$ ($q = 1, 2, \cdots, N$) is frequency spectra basis function(shorted as basis functions); $\mathbf{w}_q(\theta_i)$ are spatial weighting coefficients that vary as functions of direction and individual; $\mathbf{H}_{av} = \frac{1}{M}\sum_{i=1}^{M}\mathbf{H}(\theta_i)$ is a $N \times 1$ column vector, it's the mean spatial function of $\mathbf{H}(\theta_i)$, and is the independent part with direction in HRTF amplitude.

According to the method mentioned in[2], N orthogonal basis functions of the spectral shape d_q and the corresponding spatial weighting coefficient $\mathbf{w}_q(\theta_i)$ can be solved. After subtracting $\mathbf{H}_{av}$ from each $\mathbf{H}(\theta_i)$, $\mathbf{H}_{\Delta i}$ is obtained. $\mathbf{H}_{\Delta i}$ from all M

directions compose a $N \times M$ matrix $\mathbf{H}_\Delta$, thus an $N \times N$ covariance matrix $\mathbf{R}$ can be constructed as:

$$\mathbf{R} = \frac{1}{M}[\mathbf{H}_\Delta][\mathbf{H}_\Delta]^T \qquad (1.4)$$

where the superscript T denotes the transpose of the matrix. The frequency basis vectors can be obtained from eigenvectors $\mathbf{w}_q(\theta_i)$ of matrix $\mathbf{R}$ for the preceding Q largest and positive eigenvalues λ_q, then using the orthogonality of $\mathbf{d}_q$ in turn, $w_q(\theta_i)$ is represented as

$$\mathbf{w}_q(\theta_i) = \mathbf{d}_q^T \mathbf{H}_{\Delta i} \qquad (1.5)$$

If q<N is chosen, then Eq.(2) is an approximate representation of the original HRTF magnitude spectra, and the data quantity of HRTF is simplified. The cumulative percentage variance of energy represented by preceding Q basis vectors is given as

$$\eta = \sum_{q=1}^{Q} \lambda_q \bigg/ \sum_{q=1}^{N} \lambda_q \times 100\% \qquad (1.6)$$

1.3.2 Different pinna structure analysis

In this paper, the HRTF of left ear in the horizontal plane is analyzed for all 5 kinds of pinna structure. The discussing frequency range is from 0 to 20kHz.Because the different pinna structure model is obtained based on changing the details of the "typical pinna" model, so analysis the amplitude spectra difference between them, the corresponding changes in pinna details impact on HRTF will be found.

1.3.2.1 Analysis of directional mean $\mathbf{H}_{av}$

$\mathbf{H}_{av}$ is the average of HRTF from all directions, so it reflects the average characteristics of HRTF spectrum. It shows the general difference of different pinna structure effect on HRTF. $\mathbf{H}_{av}$ of different pinna structure are presented in Fig.1.3. First, we can find that the difference of these five curves begin from 3kHz. It proves that pinna is an important localization cue for high frequency. The curve of no-pinna case(No.5) is relatively smooth, there is no valley caused by pinna, it is obviously distinguished from the "typical pinna" case (symbolized as No.1 pinna) and others. For No.2 pinna structure, only the conchae section is modified——cymba conchae

and cavitas conchae is artificially connected. There is an obvious peak at 13kHz. For No.3 pinna structure, where helix is minified, there is little difference with No.1 pinna structure, it is likely to infer that the impact of the helix on HRTF is indistinctively. Antitragus is thickened in No.4 pinna structure. The difference is obvious compared No.4 with No.1 pinna sttructure,it has a significant peak at 4.5kHz,and has a valley at 9kHz.

1.3.2.2 Spectral basis vector and directional weighting coefficient analysis

The structure features and energy percentage of first three basis functions of different pinna structure are listed in table.1.1. The energy percentage reaches above 98%, when first three basis functions are chosen. Thus, the first three basis functions can basically represent the original HRTF energy information, which greatly simplifies the amount of data, and provides a simple and effective analysis of the influence of the pinna structural changes on HRTF data.

According to the result shown in Fig.1.3, besides no-pinna case(No.5), No.2 and No.4 pinna structure have a relatively distinct difference in directional mean compared with the "typical pinna"(No.1). Further analysis need to be carried out for these types of pinna model. The first three basis vector $\mathbf{d}_1$ 、 $\mathbf{d}_2$ and $\mathbf{d}_3$ of No.1, No.2 and No.5 pinna are presented in Fig.1.4.

In Fig.1.4(a), it indicates that the difference between no-pinna case and the "typical pinna" is the most obvious. For "typical pinna" structure there is a clear peak in the curve of $\mathbf{d}_1$, whereas, a smooth curve appear for the no-pinna case. Comparing these two types of structure, $\mathbf{d}_2$ and $\mathbf{d}_3$ also present complex and clear distinction, which are all brought out due to the complex structure of pinna.
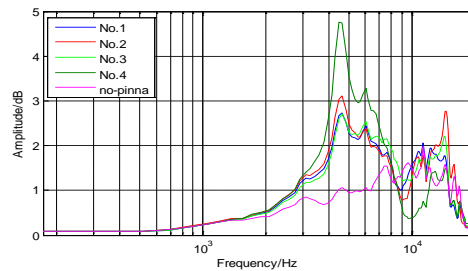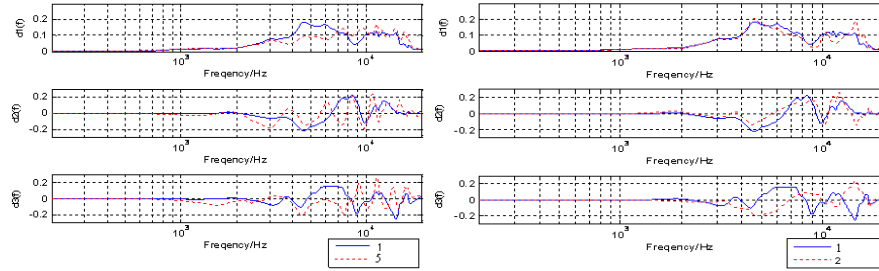

Fig.1.3 Comparison of H_{av} among 5 kinds of pinna model

(a) Comparison of No.1 with No.5 pinna (b) Comparison of No.1 with No.2 pinna

Fig.1.4 The first three basis functions of pinna structure

Comparing $\mathbf{d}_1$ of No.2 pinna with the "typical pinna" in Fig.1.4(b), in the range of frequency above 9kHz the difference is more obvious, and a peak appear at 15kHz;when f>3kHz,visible difference is existed in $\mathbf{d}_2$, especially an valley presents at 15kHz for No.2 pinna. The variation in $\mathbf{d}_3$ is more complex, it's ignored for a smaller proportion of the energy.
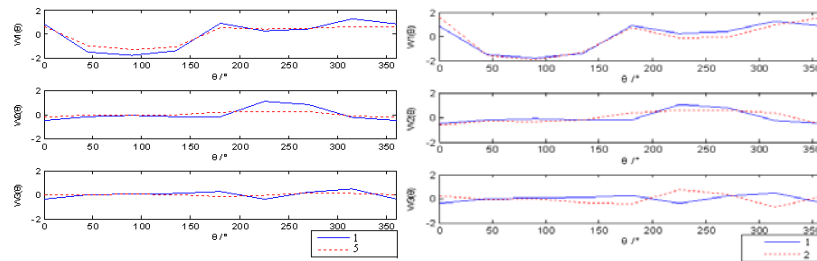
The directional weighting coefficients of these pinna structures are given in Fig.1.5. From these curves, it shows that the fluctuation of $\mathbf{w}_1$ is the biggest which indicates the largest contribution in reconstruction of HRTF. Both of $\mathbf{w}_2$ and $\mathbf{w}_3$ are all varying around 0. $\mathbf{w}_1$ is a minus value for the contralateral direction, which lead to damping in high frequency coupled with $\mathbf{d}_1$; On the contrary, it's a plus value for the ipsilateral side, and enhance the HRTF spectra in high frequency. This is the result of the head scattering and diffraction, and it is also consistent with the investigation of Xie[6]. Comparing the no-pinna case with the "typical pinna" case, it shows that significant deviation appears mainly in the ipsilateral direction, especially at $\theta = 180°$and 310° for $\mathbf{w}_1$ and around 220°~270°for $\mathbf{w}_2$ .It indicates that pinna effect is obvious in these directions.

In summary, no matter how much the difference exist in basis vectors and directional weighting coefficients, effect of $\sum_q \mathbf{w}_q \mathbf{d}_q$ is reinforcing the sound source located in the ipsilateral ear and attenuation of high frequency in the contralateral side, and this is the main role of the pinna.

1.4 Conclusions

This paper presents the influence of different pinna structure on HRTF limited to the horizontal plane. Through the principal component analysis on HRTF of different

pinna structure, the presented results show that the specific structure of pinna has some influence on HRTF, and it's an important cue for sound localization. In the horizontal plane, besides the no-pinna case, No.2 pinna structure (with connected cymba conchae and cavitas conchae),No.4 pinna structure(with thickened antitragus) affect the HRTF most obviously. However, the effect of helix alone on HRTF is rather little. To sum up, pinna as a whole, different structural detail of pinna impact on HRTF in varying degrees and also play a different role on localization.

(a) Comparison of No.1 with No.5 pinna (b) Comparison of No.1 with No.2 pinna

Fig.1.5. The weighting coefficients of the first three basis functions

1.5 References

1. Xie,B.S.(2008)Head-related transfer function and virtual auditory.National defense industry press, Beijing, China
2. Kister D, Wightman F(1992)A model of head-related transfer function based on principal components analysis and minimum-phase reconstruction. J. Acoust. Soc. Am, vol.91 (3), pp.1637-1647
3. Qi,N.,Li.L.,Zhao,W(2010)The measurement and classify of Chinese pinna structure. Technical Acoustics, vol.29(5),pp.518-522.
4. Zhao,W.,Meng,Z.H(2010)A test to study the modified pinna for sound localization. vol.34(6),pp.38-44
5. Qi,N.,etc.(2011)A kind of dummy head. China, Patent. 201120555311.6
6. Xie,B.S.(2008)Principal components analysis on HRTF and the effect of pinna. Technical Acoustics,vol.27(5), pp. 374-375
 7. Zhang,T.T.,Xie,B.S.(2008)Principal component analysis and spatial fourier reconstruction of head-related transfer functions. Audio Engineering,vol.32(8), pp.48 -56