STATISTICAL METHOD TO IDENTIFY KEY ANTHROPOMETRIC PARAMETERS IN HRTF INDIVIDUALIZATION

M. Zhang, R. A. Kennedy, and T. D. Abhayapala

Research School of Engineering, CECS
The Australian National University
Canberra, 0200 ACT, Australia

W. Zhang

CSIRO
Process Science and Engineering
Sydney, 2234 NSW, Australia

ABSTRACT

This paper identifies the main anthropometric parameters which strongly influence the head-related transfer functions (HRTFs) in a direct physical way using statistical analysis on HRTF measured data. Principle component analysis is separately performed on the head-related impulse responses of all subjects at each direction for each ear to extract the individual information. Then the individual information, along with all anthropometric parameters, is introduced in the multiple linear regression analysis, where F statistic and t statistic are used to characterize the key parameters having strong direct physical effect on the HRTFs. Ultimately, combining with the results of the analysis of the inter-parameter correlations, only eight anthropometric parameters out of 27 are identified as the crucial elements in the role of spatial localization, which provides a guide for efficient HRTF individualization using key anthropometric parameters.

Index Terms— HRTF, Anthropometric Parameters, PCA, Multiple Linear Regression Analysis

1. INTRODUCTION

Head-related transfer functions (HRTFs), or the equivalent head-related impulse responses (HRIRs) in the time domain, describe how a sound is filtered by the head, torso and pinna of a listener as it propagates from the source to the listener's ear drum in free space. As empirically measured, interaural time differences (ITD) and interaural level differences (ILD) can be easily detected in HRIR, which play a significant role in horizontal plane localization, and complicated characteristics with multiple peaks and notches exist in the magnitude spectra of the HRTFs, which is known to provide primary cues for localization in the median plane [1, 2, 3]. As known, these prominent spatial cues depend not only on the sound source directions, but also on the listener's head, torso, and pinnae, captured by the anthropometric parameters [1], which induce many kinds of physical distortions, such as reflection, diffraction, dispersion, shadowings and resonance. [2, 4].

Individualized HRTFs can be obtained directly and accurately by measuring them on each listener. However, it is a greatly time-consuming and complicated procedure. Hence it is important and necessary to explore a mathematical individualization method to estimate the individual HRTFs. Jin *et al.* [5] developed a functional model relating morphological measurements of the external ear and head to HRTFs using PCA related techniques and multivariable linear regression analysis. Zotkin *et al.* [6, 7] proposed a database matching method which was accomplished by taking a picture of subject's ears, measuring the parameters from the image, and then finding the best match in the CIPIC HRTF database.

More than the matching method [7], Xu *et al.* [8] utilized the factor analysis to identify the most representative physical measurements from the 17 measurements in the U.S. Army Anthropometric Survey database. Then only three identified measurements were introduced in the population grouping and the resulting group mean HRTFs were found to be acceptable at most sampled positions. Differently, Hu *et al.* [9] introduced correlation analysis into anthropometric parameters selection, by which eight parameters were determined as the independent variables. Then the selected eight parameters were imported in the regression model to estimate principle components, which were obtained from principal component analysis. Further, in their later work [10], they analyzed the correlation between direction transfer function and anthropometric parameters. And finally, the same eight parameters as in the work [9] were selected as the inputs of back-propagation artificial neural network for HRTF individualization.

However, the previous parameter selection methods for HRTF individualization are limited in the correlation analysis either among anthropometric parameters [8, 9] or between anthropometric parameters and the raw HRTF data [10]. The questions for the first method are: 1) when few parameters are correlated with each other, which one should be introduced into the model and which one should be deleted from the model? 2) should all those uncorrelated parameters be introduced into the model? The concern for the second method is that the common information in the raw HRTF data might obscure the real effect of some anthropometric parameters.

The aim of this paper is to determine a minimal set of anthropometric parameters which most strongly influence the variations we see in the HRTF as we vary from individual to individual. The distinction of this research from other related researches is that we extract the individual information from the raw HRTF data and analyze the correlation between this extracted information and the anthropometric parameters. Our approaches are based on principal component analysis (PCA) [11] Pearson's product-moment correlation coefficient analysis [12], and multiple linear regression analysis [13].

The rest of the paper is organized as follows. In Section 2, PCA is employed to extract the individual information from the raw HRTF data. We perform PCA on the HRIRs [14] of all subjects at each direction for each ear, in contrast to performing PCA over all directions for all subjects. This procedure expresses the original data in a way that highlights both of the similarities expressed by principal components, and differences carried by weights. Thereupon, weights, the individual HRTF information, are central to the analysis described below.

[1]The anthropometric parameters are the physical measurements of human's head, torso, and pinnae.

Since the anthropometric parameters of head, torso, and pinna might not be independent, In Section 3, Pearson correlation coefficient analysis is introduced to measure the strength of any linear dependence between any two anthropometric parameters, that is, an inter-parameter correlation. In Section 4, to identify the smallest number of significant anthropometric parameters, we use multiple linear regression analysis to analyze the relationship between the extracted individual HRTF information from PCA procedure, the weights, and the anthropometric parameters. Then we use F statistics to test the adequacy of the regression model and t statistics to test the significance of the regressors in the adequate model. Finally, in Section 5, combining the results from inter-parameter correlation analysis and the multiple linear regression analysis, we identify the key anthropometric parameters strongly correlated with HRIRs in a direct physical way, which guides the simplification to more efficient individualized HRTFs modeling.

2. PRINCIPAL COMPONENT ANALYSIS

Principal component analysis (PCA) is a classical statistical method involving a mathematical procedure that transforms a number of possibly correlated variables into an effectively smaller number of uncorrelated variables [11].

2.1. Data Structure

We perform PCA on the CIPIC HRTF Database [14]. The database consists of 45 subject HRIRs measured at 1250 directions (25 different azimuths and 50 different elevations) along with the anthropometric measurements for each subject. For each impulse response, there are 200 samples in the time domain. For anthropometric parameters, there are 37 measurements per subject — 15 for head and torso and 22 for two pinnae. However only 27 subjects have the full complement of measurements of 37 anthropometric parameters. Since we do not know which parameters strongly affect the profile of the HRIRs, we choose the 27 subjects that have complete data both in HRIRs and the anthropometric measurements to do the analysis. Hence, the data set under analysis would be a 200 by 1250 by 27 matrix.

The HRIRs measured at different directions for a particular subject often exhibit dependency while those of different subjects measured at a particular direction tend to be more independent. We perform PCA on HRIRs of all subjects at each direction separately. Each impulse response comprises a time series of 200 variables, which is a set of measurements from one particular experiment that can be viewed as a stochastic process. Then for each analyzing procedure, the data are organized as a set of I column vectors, $\vec{h}_{\hat{\mathbf{x}}1} \cdots \vec{h}_{\hat{\mathbf{x}}I}$, where $I = 27$ subjects, with each vector representing a HRIR of N variables, where $N = 200$, of a subject measured at a particular direction. Therefore, the data set for each direction, $\mathbf{H}_{\hat{\mathbf{x}}}$, is a 200 by 27 matrix, and there are 1250 data sets in total corresponding to 1250 directions.

2.2. Interpretation of the Principal Component Analysis Result

The statistical model for reconstructing the measured HRIRs can be given by

$$\hat{h}_{\hat{\mathbf{x}},i}(n) \cong \sum_{m=1}^{M'} w_{\hat{\mathbf{x}},i}^{m} e_{\hat{\mathbf{x}},m}(n) + u_{\hat{\mathbf{x}}}(n), \tag{1}$$

where n is the time index ($n = 1, 2, \cdots, 200$), i is the subject index ($i = 1, 2, \cdots, 27$), $\hat{\mathbf{x}}$ is the direction index, m is the PC index, $\hat{h}$ denotes the reconstructed HRIR, w denotes the weights, e denotes the principal components (PCs), and $u_{\hat{\mathbf{x}}}$ denotes the mean value of the measured HRIR at a particular direction.

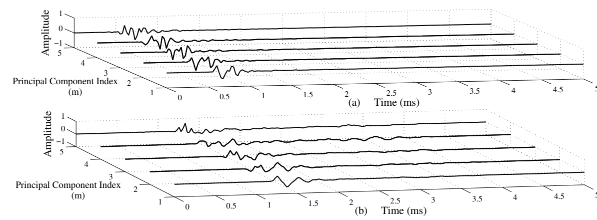

Fig. 1. PCs at the direction of $\phi = -80^\circ$ and $\theta = 0^\circ$. (a) First 5 PCs for reconstructing the left ear HRIRs; (b) First 5 PCs for reconstructing the right ear HRIRs.

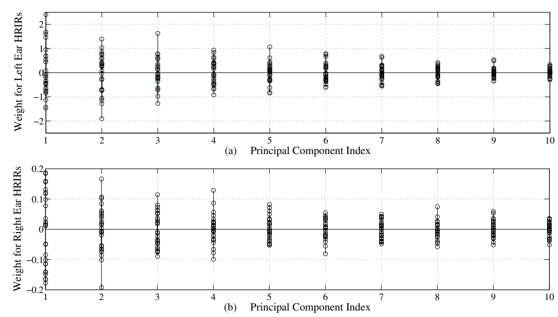

Fig. 2. Sample distribution of weights across 27 subjects over all PCs at the direction of $\phi = -80^\circ$ and $\theta = 0^\circ$. (a) Weights for reconstructing the left ear HRIRs; (b) Weights for reconstructing the right ear HRIRs (We use a different weight scale from the upper one to reveal the interaural time difference details).

In our application, we extract the number of PC to cover approximately 95% of the variance in HRIRs. Then 10 to 16 PCs are used for reconstructing the original signals with the relative mean-squared error less than 5% when the direction varies from the ipsilateral angles (the azimuth from -90 degrees to 0 degree for the left ear while 90 degrees to 0 degree for the right ear) to the contralateral angles (the azimuth from 0 degrees to 90 degree for the left ear while 0 degrees to -90 degree for the right ear). In this section, as an illustration, we present the results of PCA on 27 subjects at the direction of the azimuth ϕ of -80° and the elevation θ of 0°.

Ten PCs are used to reconstruct the left ear HRIR at the direction mentioned above with the relative mean-squared error lower than 5% while 16 PCs are used for the right ear HRIR reconstruction with the same relative mean-squared error. In Figure 1, we show the first 5 PCs. Notice that all PCs for the same ear have nearly the same time delay. But the time delay of the left ear's PCs, shown in the upper panel, is much shorter than that of the right ear's ones, shown in the lower panel, because the sound source is placed at the direction of $\phi = -80^\circ$ and $\theta = 0^\circ$, the left ear side. This phenomenon explains that the PCs also carry the interaural time difference information.

The weights of the PCs represent the contribution of each principal component to the reconstructed HRIR. They are different from person to person and from direction to direction. Figure 2 shows the distribution of weights of all PCs at the direction mentioned above with the trace of each PC indicating the variations of weights among 27 subjects. It is the degree of deviation that demonstrates the influence of personal anthropometric information on the HRIRs. As expected, the spread of weights is the largest for the first PC and the smallest for the 10th PC, which means that the weight of the first PC is more important than others in representing the variation of the original data. Moreover, the value of weights for left ear shown in the upper panel is much greater than that for the right ear in the lower panel since the sound source is placed at the left ear side, which strongly indicates that the weights are the interaural level difference information carriers.

Since the PCs are common to all subjects, the weights carry the individual information of the HRIRs. They are essential elements which are introduced to the following analysis to reveal which anthropometric parameters have strong direct physical effect on HRIRs.

3. CORRELATION IN ANTHROPOMETRIC PARAMETERS

There are 37 anthropometric parameters in the CIPIC database. But they might not be independent, for example, people with a large head tent to have a thick neck, which implies that not all anthropometric parameters are necessary. In this section, we will characterize the inter-parameter correlations (correlations among the anthropometric parameters).

3.1. Anthropometric Parameters

We reorganize the order of the parameters, as shown in Table. 1. The first 10 parameters are left pinna measurements followed by 10 right pinna measurements, the next 2 parameters are general pinna measurements, and the following 15 parameters are torso and head measurements.

3.2. Characterizing Inter-Parameter Correlation

In our application, the Pearson Correlation Coefficient [12] is introduced to measure the strength of the linear dependence between two anthropometric parameters. It is defined as the sample covariance of the two parameters divided by the product of their sample standard deviations,

$$r_{jj'} = \frac{\text{cov}(\mathbf{a_j}, \mathbf{a_{j'}})}{\sigma_{a_j}\sigma_{a_{j'}}} = \frac{\sum(\mathbf{a_j} - \bar{a_j})(\mathbf{a_{j'}} - \bar{a_{j'}})}{\sqrt{\sum(\mathbf{a_j} - \bar{a_j})^2 \sum(\mathbf{a_{j'}} - \bar{a_{j'}})^2}}, \quad (2)$$

where $\mathbf{a}$ is a vector of an anthropometric parameter with 27 elements, $j, j' = 1, 2, \cdots, 37$, and $j \neq j'$. The value of $r_{jj'}$ is between $+1$ and -1 inclusive. The higher the absolute value the stronger the correlation. But there is no objective judgment on what is the exact value for strong correlation [15]. In this paper, we determine that the correlation is "strong" if $|r_{jj'}| > 0.7$ and "weak" otherwise.

Figure 3 highlights 22 pairs of anthropometric parameters with strong linear correlation. It can be seen from this figure that there is no strong correlation between any two pinna parameters of the same ear while, undoubtedly, strong correlations exist between most of the parameters of different ears. Only one pinna parameter, pinna offset down, indicates a strong inversely proportional linear correlation with a head parameter, head height. There are 14 pairs of head and torso measurements having strong correlation. They all present positive proportional linear relationship.

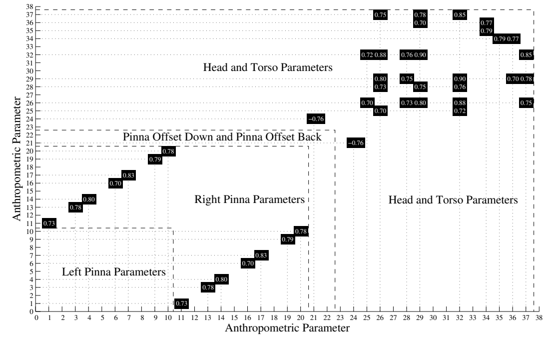

Fig. 3. Correlations among 37 anthropometric parameter measurements. The numbers are the Pearson correlation coefficients between two anthropometric parameters.

4. CORRELATION OF ANTHROPOMETRIC PARAMETER WITH WEIGHT

In this section we aim at revealing which anthropometric parameters have the direct physical effect on HRIRs. Recall Section 2.2, M' weights for M' principal components are required to represent the HRIR at a particular direction for a particular subject and they are the individual information carriers as well. Thus, our aim is to characterize the correlation of all anthropometric parameters with the weights. Multiple linear regression analysis [13] is applied to address this problem due to its ability to identify which regressor variables have greatest effect on the response variable.

4.1. Methodology

In our application, the regressors are the anthropometric parameters, $\mathbf{a_j} = \{a_{1j}, a_{2j}, \cdots, a_{Ij}\}$, and the response is a particular weight, $w^m_{\hat{\mathbf{X}},i}$. Then the multiple linear regression model takes the form

$$\mathbf{w}^{\mathbf{m}}_{\hat{\mathbf{X}}} = \mathbf{A}\boldsymbol{\Upsilon}^{\mathbf{m}}_{\hat{\mathbf{X}}} + \epsilon^{\mathbf{m}}_{\hat{\mathbf{X}}}, \quad (3)$$

where $\boldsymbol{\Upsilon}^{\mathbf{m}}_{\hat{\mathbf{X}}}$ are the regression coefficients, and $\epsilon^{\mathbf{m}}_{\hat{\mathbf{X}}}$ are the errors. Then the least-squares estimator of the regression coefficients is given by

$$\hat{\boldsymbol{\Upsilon}}^{\mathbf{m}}_{\hat{\mathbf{X}}} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{w}^{\mathbf{m}}_{\hat{\mathbf{X}}}, \quad (4)$$

where T represents transposed operator.

Once we have estimated the regression coefficients in the model, we face two important questions: (i) Is there at least one of the regressors in the model contributing significantly to the response? and [ii] How to determine which specific regressors are of real value in explaining the response?

To address the first question, we use an F statistic, which is defined as

$$F^m_{\hat{\mathbf{X}}} = \frac{(A\hat{\boldsymbol{\Upsilon}}^{\mathbf{m}}_{\hat{\mathbf{X}}}\mathbf{w}^{\mathbf{m}}_{\hat{\mathbf{X}}} - (\sum_{i=1}^{I} w^m_{\hat{\mathbf{X}},i})^2/I)/J}{(\mathbf{w}^{\mathbf{m}\prime}_{\hat{\mathbf{X}}}\mathbf{w}^{\mathbf{m}}_{\hat{\mathbf{X}}} - A\hat{\boldsymbol{\Upsilon}}^{\mathbf{m}}_{\hat{\mathbf{X}}}\mathbf{w}^{\mathbf{m}}_{\hat{\mathbf{X}}})/(I - J - 1)}, \quad (5)$$

where $F^m_{\hat{\mathbf{X}}}$ follows the F distribution. This is an overall test of the model adequacy. If the value of $F^m_{\hat{\mathbf{X}}}$ is greater than $F_{\alpha,J,I-J-1}$, where $100 * (1 - \alpha)\%$ indicates the confidence level and J and $I - J - 1$ indicate the degrees of freedom, at least one of the regressors in the model has great influence on the response.

Table 1. Anthropometric Parameters

Order	Measurements	Order	Measurements	Order	Measurements
1,11	Cavum concha height (L/R)	21	Pinna offset down	31	Torso top depth
2,12	Cymba concha height (L/R)	22	Pinna offset back	32	Shoulder width
3,13	Cavum concha width (L/R)	23	Head width	33	Head offset forward
4,14	Fossa height (L/R)	24	Head height	34	Height
5,15	Pinna height (L/R)	25	Head depth	35	Seated height
6,16	Pinna width (L/R)	26	Neck width	36	Head circumference
7,17	Intertragal Incisure width (L/R)	27	Neck height	37	Shoulder circumference
8,18	Cavum concha depth(L/R)	28	Neck depth		
9,19	Pinna rotation angle (L/R)	29	Torso top width		
10,20	Pinna flare angle (L/R)	30	Torso top height		

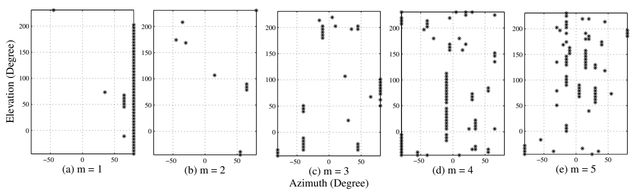

Fig. 4. The overall test results of the model adequacy for the left ear weights of 10 PCs at all directions with 15 head and torso parameters. m is the PC index.

After determining that at least one of the regressors is important, we need to pick up which one(s). A t statistic [13] is helpful to this question. It is defined as

$$t^m_{\hat{\mathbf{X}},j} = \left| \frac{\hat{\gamma}^m_{\hat{\mathbf{X}},j}}{\sqrt{(\mathbf{w}^{\mathbf{m}T}_{\hat{\mathbf{X}}}\mathbf{w}^{\mathbf{m}}_{\hat{\mathbf{X}}} - A\hat{\mathbf{\Upsilon}}^{\mathbf{m}}_{\hat{\mathbf{X}}}\mathbf{w}^{\mathbf{m}}_{\hat{\mathbf{X}}})(\mathbf{A}^T\mathbf{A})^{-1}}} \right|, \qquad (6)$$

where $t^m_{\hat{\mathbf{X}},j}$ follows the t-distribution. This is the test of the regressor significance. If $t^m_{\hat{\mathbf{X}},j} > t_{\alpha/2,I-J-1}$, the regressor $\mathbf{a}_j$ can be deemed as the significant element to the response. The above analysis does not include the error term in (3) because we focus on the significant regressor identification.

Therefore, in the following analysis, we employ this hypothesis method to characterize the significant anthropometric parameters with direct physical effect on HRIRs. We analyze the head and torso parameters and the pinna parameters separately.

4.2. Head and Torso Parameter characterization

There are 15 head and torso parameters and 1250 by 10 by 2 weights corresponding to 10 PCs of HRIRs at 1250 directions for both ears. We will analyze the correlation of each weight with 15 head and torso parameters.

Following the procedures discussed in Section 4.1, at first, we built the multiple linear regression model, in which regressor is the weight and regressors are the 15 head and torso parameters, to calculate the regression coefficient for each head and torso parameter.

Then we compute the F statistic to test the model adequacy. Totally we have 1250 by 10 values of F statistic for each ear. Here $I = 27, J = 15$, and we set $\alpha = 0.05$. Therefore, $F_{0.05,15,11} = 2.72$. If the value of $F^m_{\hat{\mathbf{X}}}$ is greater than 2.72, at least one head and torso parameter has a significant effect on the particular weight. Figure 4 plots the overall adequacy test results of the regression model for the left ear weights of 1st PC to 5th PC at all directions with 15 head and torso parameters. Stars indicate $F^m_{\hat{\mathbf{X}}}$ with a value greater than 2.72. In the panel of $m = 1$, most values of F statistic at azimuth angle of 80 degrees, where the sound is coming from the right ear side, are higher than 2.72. It means that the head and torso parameters are enough to model the first PC weights at these contralateral directions. For the panel of $m = 2$ and $m = 3$, the value of F statistic at only few directions are greater than 2.72. And the studies of other higher PCs show that the adequacy distributions are scattered all around, which is weak in conveying helpful information. Therefore, our following analysis focuses on the 1st PC weights at large contralateral directions.

Next, in the test of the regressor significance, we compute t statistic to evaluate each regression coefficient $\hat{\gamma}^m_{\hat{\mathbf{X}},j}$ where the sound source is coming from the large contralateral directions. In this case, the degrees of freedom is 11 ($I = 27$ and $J = 15$) and we set $\alpha = 0.05$. Therefore, $t_{0.025,11} = 2.201$. If the value of t statistic is greater than 2.201, the corresponding regressor a_j, the jth head and torso parameter, makes significant contribution to the response, the weight for the mth principal component at a particular direction of $\hat{\mathbf{X}}$. In Figure 5, only the results of 3 head and torso parameters are presented since their t statistic values are greater than the threshold at many directions while the rest parameters show little or no contribution to the responses. As we can see in Figure 5, the t values of the head depth are greater than the threshold at most elevations, which means it has the most significant effect on the weights, while for the other two parameters in Figure 5, only few t values are greater than the threshold. Thus, the head depth is deemed as the parameter having significant effect on the contralateral weights.

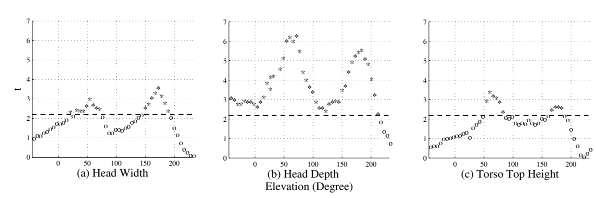

Fig. 5. The effect of 4 head and torso parameters on the 1st principal component weights of left ear at azimuth angle of 80 degrees and all elevation angles. Stars indicate that the value of t statistic is greater than 2.201 while circles indicate that the value of t statistic is less than 2.201. The dash line is t = 2.201.

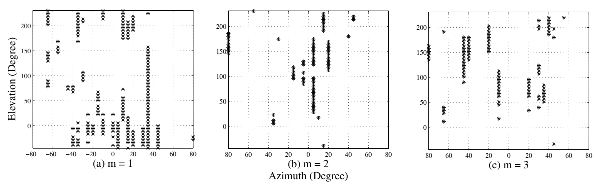

Fig. 6. The overall test results of the model adequacy for right ear weights of 3 principal components at all directions with 12 pinna parameters.

4.3. Pinna Parameter characterization

As in Section 4.2, we analyze the correlation of the pinna parameters with weights in the same way.

Firstly, we built the multiple linear regression model. For the right ear model, the response is the right ear weight and the regressors are the 10 pinna parameters of right ear and 2 general pinna parameters. For the left ear model, the response is the left ear weight and the regressors are the 10 pinna parameters of left ear and 2 general pinna parameters. So, there are 12 regressors in each model.

Then we test the model adequacy. Here $I = 27, J = 12$, and we set $\alpha = 0.05$. Therefore, $F_{0.05,12,14} = 2.53$. Figure 6 plots the overall adequacy test results of the regression model for right ear weights of the first 3 principal components at all directions with 12 pinna parameters. Stars indicate $F^m_{\mathbf{X}}$ with a value greater than 2.53. Notice that most F statistics with value greater than 2.53 are centered at the azimuth angles between -50 degrees to 50 degrees, especially for the 1st principal component (the panel with m = 1). It means that at least one pinna parameter has a significant effect on the weights at these middle directions. For other principal components, the adequacy distributions are scattered all around which fail to deliver useful information.

Finally, in regressor significance test, we compute t statistic for the weights of the first 3 principal component. In this case, the degrees of freedom is 14 ($I = 27$ and $J = 12$) and we set $\alpha = 0.05$. Therefore, $t_{0.025,14} = 2.145$. Figure 7 presents the value of t statistic for 6 pinna parameters with the weights of the right ear for the 1st principal component at all directions, which show the significant effect on the weights at many directions. They are cavum concha width, fossa height, pinna height, pinna width, and pinna rotation angle. Notice that pinna flare angle is also important when the azimuth angle is between 0 to 20 degrees. Therefore, 6 parameters are identified from the 1st PC.

For the 2nd principal component, we focus on the azimuth angle of 5 degrees since at this azimuth angle the regression models are adequate, as shown in the middle panel of Figure 6. Figure 8 plots t values of five pinna parameters that show their significance at many elevation angles. Cavum concha width, Fossa Height, pinna width, and pinna offset back can be found to be the significant regressors on the weights at most elevation angles. Pinna rotation angle shows to have great influence on the weights at lower or higher elevation angles too. Unfortunately the regression models at these directions are not adequate. Therefore it should not be included in the model. Then, from the 2nd PC, we identify 4 parameters in which three of them are also show their significant to the 1st PC. For the 3rd PC, we focus on the azimuth angle of 20 degrees for the same reason (see the lower right panel of Figure 6). Then pinna width and pinna flare angle are picked out, as shown in Figure 9. Both of them exhibit the significance at around over head directions and low front direction.

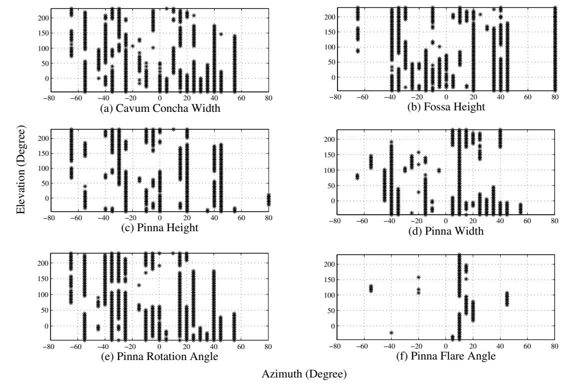

Fig. 7. The effect of 8 pinna parameters on the 1st principal component weights of right ear at all directions. Stars indicate that the t value is greater than 2.145 at those directions.

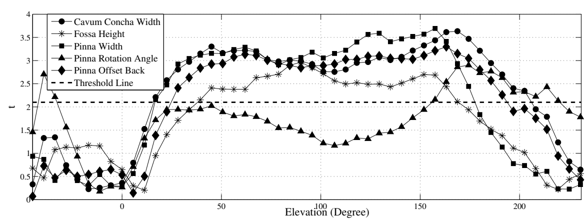

Fig. 8. The effect of 5 pinna parameters on the 2nd principal component weights of right ear at azimuth angle of 5 degrees and all elevation angles. The dash line is t = 2.145.

In summary, 7 pinna parameters are identified as having significant effects on the weights at azimuth angle between -50 degrees and 50 degrees, in which pinna height and pinna rotation angle are derived from the 1st PC, cavum concha width and fossa height from both the 1st and 2nd PC, pinna width from all the 3 PCs, and pinna flare angle from the 3rd PC.

5. KEY ANTHROPOMETRIC PARAMETER IDENTIFICATION

Some anthropometric parameters have direct physical effect on HRTFs while others have statistical influences with HRTFs by the virtue of being correlated with previous ones. Our goal is to identify the anthropometric parameters which are strongly correlated with HRTFs in a direct physical way.

From Section 4.2, we know that only one parameter, the head depth, is the crucial factor in influencing the weights at large contralateral angles. Recall Figure 3, it shows high correlation with the neck width and the shoulder width, which means neck width and shoulder width might have influences with HRTFs in a statistical way. But it is the head depth that definitely presents the effect on HRTFs in a direct physical way. Also, this can be proved to be reasonable by the spherical head model [16], in which only the radius of the sphere with respect to the radius of the head is used to theoretically calculate the low frequency HRTFs.

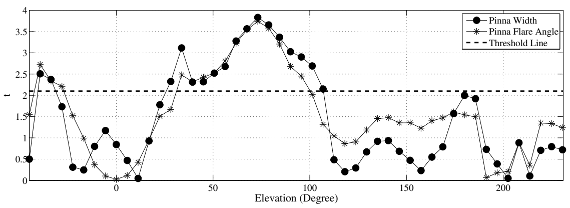

Fig. 9. The effect of pinna width and pinna flare angle on the 3rd principal component weights of right ear at azimuth angle of 20 degrees and all elevation angles. The dash line is $t = 2.145$.

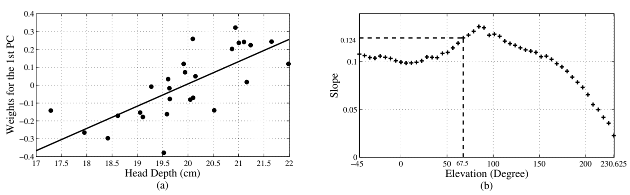

Fig. 10. The linear relationship between head depth and the weights of left ear at azimuth of 80 degrees.

Figure 10 illustrates the linear relationship between the head depth and the weight at larger azimuth angles. The left panel of Figure 10 is the scatter plot for two variables with a straight line. The slope of the straight line can be interpreted as the change in weight for a unit change in head depth. The right panel of Figure 10 describes that the slope varies along with the elevation. When the elevation is around 90 degrees, where the sound source is located over head, the head depth has great influence on weights, in other words, HRTFs. With the position of sound source going down, the influence of the head depth on HRTFs decreases.

From Section 4.3, 7 pinna parameters extracted from the first 3 principal components are the crucial elements. Recalling the inter-parameter correlation analysis results, no correlation exists between any two pinna parameters of the same ear. Reasonably all these 7 pinna parameters are remained. Their actions on HRTFs are interrelated.

Combining the above analysis, 8 anthropometric parameters are identified as having strong correlation with HRTFs in a direct physical way. They are head depth, cavum concha width, fossa height, pinna height, pinna width, pinna rotation angle, pinna flare angle, and pinna offset back.

6. CONCLUSION

In this paper, we proposed a statistical method to study which anthropometric parameters most strongly influence individualized HRTFs. The unique feature of the method investigated here is the introduction of the individual information embedded in HRTFs into the analysis of the correlation between HRTFs and the anthropometric parameters. The results suggest that the head depth has a significant effect on the HRTFs at large contralateral directions and the seven pinna parameters have a significant effect on the ipsilateral HRTFs, which provides a guide for efficient HRTF individualization using key anthropometric parameters.

7. REFERENCES

[1] V. C. Raykar, R. Duraiswami, and B. Yegnanarayana, "Extracting the frequencies of the pinna spectral notches in measured head related impulse response," *J. Acoust. Soc. Am.*, vol. 118, no. 1, pp. 364–374, July 2005.

[2] J. Blauert, *Spatial Hearing*, The MIT Press, Cambridge, Massachusetts, London, England, revised edition, 1997.

[3] D. W. Batteau, "The role of the pinna in human localization," *Biological Sciences*, vol. 168, no. 1011, pp. 158–180, Aug. 1967.

[4] Y. Huang and J. Benesty, Eds., *Audio Signal Processing For Next - Generation Multimedia Communication Systems*, Kluwer Academic Publishers, Assinippi Park, Norwell, Massachusetts, 02061 USA, 2004.

[5] C. Jin, P. Leong, J. Leung, A. Corderoy, and S. Carlile, "Enabling individualized virtual auditory space using morphological measurements," in *Proc. of the First IEEE Pacific - Rim Conference on Multimedia*. Citeseer, 2003, pp. 235–238.

[6] D. N. Zotkin, R. Duraiswami, and L. S. Davis, "Rendering localized spatial audio in a virtual auditory space," *IEEE Trans. Multimedia*, vol. 6, no. 4, pp. 553–564, August 2004.

[7] D.N. Zotkin, J. Hwang, R. Duraiswami, and L. S. Davis, "HRTF personalization using anthropometric measurement," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, NY, USA, October 2003, pp. 157–160.

[8] S. Xu, Z. Li, and G. Salvendy, "Individualized head-related transfer functions based on population grouping (L)," *J. Acoust. Soc. Am.*, vol. 124, no. 5, pp. 2708–2710, Nov. 2008.

[9] H. Hu, L. Zhou, H. Ma, and Z. Wu, "Head related transfer function personalization based on multiple regression analysis," in *Proc. IEEE Int. Conf. on Computational Intelligence and Security*, Guangzhou, China, November 2006, vol. 2, pp. 1829–1832.

[10] H. Hu, L. Zhou, H. Ma, and Z. Wu, "HRTF personalization based on artificial neural network in individual virtual auditory space," *Applied Acoustics*, vol. 69, pp. 136–172, 2008.

[11] I. T. Jolliffe, *Principal Component Analysis*, Springer series in statistics. Springer verlag, 175 Fifth Avenue, New York, NY 10010, USA, 2nd edition, 2002.

[12] W. Mendenhall and T. Sincich, *Statistics for Engineering and the Sciences*, Pearson Prentice-Hall, Upper Saddle River, N.J., 5th edition, 2007.

[13] D.C. Montgomery, E. A. Peck, and G. G. Vining, *Introduction to linear regression analysis*, A Wiley Interscience Publication, New York, USA, 3rd edition, 2001.

[14] V. R. Algazi, R. O. Duda, D. M. Thompson, and C. Avendano, "The CIPIC HRTF database," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Electroacoustics*, New Paltz, NY, USA, October 2001, pp. 99 – 102.

[15] W. C. Rinaman, *Foundations of probability and statistics*, Saunders College Publishing, Orlando, Florida, USA, 1993.

[16] R. O. Duda and W. L. Martens, "Range dependence of the response of a spherical head model," *J. Acoust. Soc. Am.*, vol. 105, no. 5, pp. 3048–3058, November 1998.