**TELKOMNIKA**, Vol.13, No.3, September 2015, pp. 1014~1020
ISSN: 1693-6930, accredited A by DIKTI, Decree No: 58/DIKTI/Kep/2013
DOI: 10.12928/TELKOMNIKA.v13i3.1792 

# A New Selection Method of Anthropometric Parameters in Individualizing Head-Related Impulse Responses

**Hugeng*[1], Wahidin Wahab[2], Dadang Gunawan[3]**
[1]Department of Computer Engineering, Universitas Multimedia Nusantara (UMN),
Jl. Scientia Boulevard, Gading Serpong, Tangerang 15810, Indonesia, Ph./Fax: +6221-54220808/00
[2,3]Department of Electrical Engineering, Universitas Indonesia (UI),
Kampus Baru UI, Depok 16424, Indonesia, Ph./Fax: +6221-7270078/77
*Corresponding author, e-mail: hugeng@umn.ac.id*[1], wahidin.wahab@ui.ac.id[2], guna@eng.ui.ac.id[3]

***Abstract***

*A trend issue in modeling head-related impulse responses (HRIRs) is how to individualize HRIRs models that are convenient for a particular listener. The objective of this research is to show a robust selection method of eight anthropometric parameters out of all 27 parameters defined in CIPIC HRTF Database. The proposed selection method is systematically and scientifically acceptable, compared to 'trial and error' method in selecting the parameters. The selected anthropometric parameters of a given listener were applied in establishing multiple linear regression models in order to individualize his / her HRIRs. We modelled the entire minimum phase HRIRs in horizontal plane of 35 subjects using principal components analysis (PCA). The individual minimum phase HRIRs can be estimated adequately by a linear combination of ten orthonormal basis functions.*

***Keywords***: *HRIR modeling, HRIR individualization, multiple regression analysis, principal components analysis*



## 1. Introduction

Head-related impulse response is the impulse response of a human ear that functions as an acoustic filter of human auditory system from a sound source to the entrance of ear canal. Two main cues in localizing the directions of sound sources on the horizontal plane are Interaural Time Difference (ITD) and Interaural Level Difference (ILD) [1]. ITD and ILD are almost undistinguishable on the median plane. However, localization of sound direction on this plane is possible by spectral modification, mainly due to reflection and diffraction on pinnae folds [2]. Filtering monaural sound using human binaural HRIRs creates Virtual Auditory Space (VAS) in virtual reality. This depends on the human psychoacoustic characteristics. Human tends to find a convincing spatial sound sufficiently using two ear channels. Control of ITD, ILD, and spectral modification is significant in providing information of sound source direction to a listener in order to create VAS. All these three primary sound cues are encrypted in HRIR. The Fourier-pair of HRIR in frequency domain is known as head-related transfer function (HRTF). Many researches had shown that HRTF varies among subjects due to inter-individual differences in anthropometric parameters and changes in sound sources' directions [2, 3].

A series of empirical measurements of individual HRTFs for a specific listener are required in synthesizing perfect VAS system. These measurements will inevitably grow prohibitive, taking into account the requirements of specialized and expensive equipment and the measurement time spent. Commercial VAS systems are recently synthesized usually in an inexpensive way by using non-individualized or generic HRTFs that ignore inter-individual differences. The works in [3, 4], however, showed that non-individualized HRTFs, i.e. unsuitable HRTFs applied to a listener, suffer from distortions such as in-head localization when using headphones, poor vertical effects, inaccurate lateralization, and weak front-back distinction. Thus, it is essential to develop an individualization method to estimate proper HRIRs for a listener which is able to provide adequate sound cues without necessitating a measurement of the individual HRIRs.

A vast body of researches is devoted to the individualization of HRTF in frequency domain or HRIR in time domain. A number of HRTF individualization methods have been

*Received March 26, 2015; Revised May 20, 2015; Accepted June 10, 2015*

proposed, such as HRTF clustering and selection of a few most representative ones [5], HRTF scaling in frequency [6], a structural model of composition and decomposition of HRTFs [2], HRTF database matching [7], the boundary element method [8], HRIR subjective customization of pinna responses [9] and of pinna, head, and torso responses [10] on median plane, and HRTF personalization based on multiple regression analysis (MRA) on horizontal plane [11]. Shin and Park [9] proposed HRIR customization method based on subjective tuning of only pinna responses (0.2 ms out of entire HRIR) on median plane using PCA of the CIPIC HRTF Database. They attained the customized pinna responses by letting a subject tune the weights on three basis functions. Hwang and Park [10] followed the similar method as in [9], but they fed PCA with the entire median HRIRs; i.e. each HRIR was 1.5 ms long (67 samples) since the arrival of direct pulse. This HRIR included the pinna, head, and torso responses. They tuned the weights of three dominant basis functions subjectively according to the three largest standard deviations at each elevation. Hu et al. [11] personalized the estimated log-magnitude responses of HRTFs by MRA. Firstly, the log-magnitude responses were approximated using PCA as linear combination of weighted basis functions. The weights of the basis functions were subsequently approximated using anthropometric parameters by MRA. Our individualization method was better than the method in [11], because we used the minimum phase HRIRs ($HRIRs_{mp}$) in time domain to be modelled in PCA, and our anthropometric parameters selection method was different. Modeling of $HRIRs_{mp}$ by PCA was based on the fact that modeling minimum phase HRIRs provided best results among other preprocessings of HRIRs in time domain, as shown by Hugeng et al. in [12].

This research was a comprehensive research in fulfilling and validating the goal to develop parametric models of HRTFs that can be tuned based on few number of listener's own anthropometries. These anthropometries should provide crucial perceptual psychoacoustic effects on spatial sound. At first, for PCA modeling, a best preprocessing and data type of HRIR in time domain; and a best preprocessing and data type of HRTF in frequency domain were found as published in [12]. The best data types were minimum phase HRIRs and magnitude HRTFs.

The individualization of $HRIRs_{mp}$ for sound sources on the horizontal plane was explained and investigated in [13] that used the same individualization method as one that is explained in this paper. In [14], individualization of magnitude HRTFs for sound sources on horizontal plane was investigated. The individualization method used was similar to the method used here, except that the individualization of magnitude HRTFs was done in frequency domain. After individualization process in frequency domain was finished, the individualized magnitude HRTFs should be reconstructed back to time domain to yield individualized HRIRs.

In this research, the entire horizontal $HRIRs_{mp}$ from the original HRIRs in the CIPIC HRTF Database were included in a single analysis. Thus, all horizontal $HRIRs_{mp}$ share the same set of basis functions, which comprise the inter-individual variation as well as the inter-elevation variation. The responses of first 1.5 ms of $HRIRs_{mp}$, which contain the effects of pinna, head, and torso, were included in PCA, as proposed by [10]. This paper presents an individualization method by developing the statistical PCA model between the selected anthropometric parameters and the $HRIRs_{mp}$ in a different and novel way compared to [9-11]. Section 2 describes details of the algorithm of individualization method. Section 3 explains the selection method of independent and dependent variables of multiple regression models and then elaborates the performance of the proposed method through calculation of the error between each measured HRIR and corresponding estimated HRIR.

## 2. Research Method

In this paper, we underlined the method for selecting eight anthropometries out of all 27 anthropometries. These selected anthropometric parameters, together with minimum phase HRIRs were established into multiple regression models in order to individualize a given listener's HRIRs. Figure 1 shows the algorithm structure of our HRIR individualization method. The database used was measured and provided by CIPIC Interface Laboratory of University of California at Davis. From Figure 1, we can see that at first, the entire original HRIRs on horizontal plane of 35 subjects are processed by cepstral analysis to be converted to their corresponding $HRIRs_{mp}$. The mean of the entire $HRIRs_{mp}$ is then calculated. In order to achieve a kind of data from $HRIRs_{mp}$ that have zero mean, the mean is subtracted from each $HRIR_{mp}$ to obtain corresponding minimum phase direct impulse response ($DIR_{mp}$). These $DIRs_{mp}$ are a set of minimum phase HRIRs data with zero empirical mean which is required to obtain a basis that minimizes mean square error of the approximated data. The whole set of $DIRs_{mp}$ are then inputted to PCA, which then results in basis functions or principal components (PCs) together with their weights. The linear combination of weighted PCs forms estimated $DIR_{mp}$. We applied multiple linear regression (MLR) as the method to individualize the estimated $DIR_{mp}$. MLR utilizes weights of PCs (PCWs) and anthropometric parameters in order to provide regression coefficients that later can be applied to model $DIR_{mp}$ of a new listener.

The process of reconstruction to the desired HRIR model is shown by dashed lines in Figure 1. A model of $DIR_{mp}$ that results from MLR analysis, based on the anthropometric parameters of the listener, is added with the mean of $HRIRs_{mp}$ to yield a model of $HRIR_{mp}$. Initial left- and right-ear time delay (Onset in Figure 1) due to distance from sound source to ear drum are inserted back to the $HRIR_{mp}$ model, resulting in the desired HRIR model. More details about the minimum phase HRIR, and MLR modeling are explained in the following subsections.

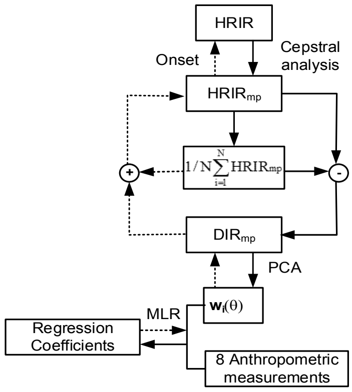


Figure 1. Proposed HRIRs Individualization Method [15]

**2.1. Minimum Phase HRIR**

Kulkarni et. al [16] suggested that the phase of HRIR can be approximated by minimum phase. A system function, H(z), of an HRIR, h(n), is said to be minimum phase if all poles and zeros of H(z) lie inside the unit circle |z| =1. The minimum phase HRIR, $h_{mp}(n)$, can be obtained through the calculation of real cepstrum of its original HRIR, which has arbitrary phase. It can be said that the $h_{mp}(n)$ is the removed initial time delay version of HRIR, but both kinds of HRIR have the same magnitude spectrum in the frequency domain. The real cepstrum, v(n), of HRIR, h(n), is calculated as follow,

$$v(n) = \mathrm{Re}\{F_D^{-1}\{\ln|F_D\{h(n)\}|\}\}, \quad (1)$$

Where ln and Re{} denote respectively natural logarithm and the real part of a complex variable, $F_D\{\}$ and $F_D^{-1}\{\}$ are the discrete Fourier transform and its inverse respectively. This real cepstrum is then weighted by the following window function, w(n), which is given by:

$$w(n) = \begin{cases} 0 & \text{if } n < 0, \\ 1 & \text{if } n = 0, \\ 2 & \text{if } n > 0. \end{cases} \quad (2)$$

For a rational H(z), the window function is a complex conjugate inversion of the zeros outside the unit circle, so that a minimum phase HRIR is provided. Hence the desired minimum phase HRIR, $h_{mp}(n)$, is yielded by:

$$h_{mp}(n) = \mathrm{Re}\{\exp(F_D\{w(n).v(n)\})\}. \quad (3)$$

**2.2. Multiple Linear Regression Modeling**

Suppose that the relation between the weights vector of the $PC_i$ (i=1,2,…,10) in azimuth θ of all subjects, $\mathbf{w}_i(\theta)$ (35x1), and the corresponding anthropometric parameters, **X** (35x9), as follows,

$$\mathbf{w}_i(\theta) = \mathbf{X} \, . \, \boldsymbol{\beta}_i(\theta) + \mathbf{E}_i(\theta), \quad (4)$$

Where **X** is the matrix composed of a 35x1 column vector with all 1's and 8 anthropometric parameters of all subjects being analyzed, $\boldsymbol{\beta}_i(\theta)$ is the regression coefficients column vector (9x1), and $\mathbf{E}_i(\theta)$ is the estimation errors column vector (35x1). The regression coefficients are found by least-square estimation. This is performed by solving the optimization problem $\min\{E_{i,n}(\theta)\}$, where $E_{i,n}(\theta)$ is the n-th dependent variable's estimation error. PCWs and anthropometric parameters are respectively the model's dependent and independent variables. From Equation (4), the regression coefficients due to i-th PCW in azimuth θ, $\boldsymbol{\beta}_i(\theta)$, can be written as,

$$\boldsymbol{\beta}_i(\theta) = (\mathbf{X}^T . \mathbf{X})^{-1} . \mathbf{X}^T \, \mathbf{w}_i(\theta). \quad (5)$$

From Equation (5), in order to enhance the performance of the multiple linear regression models, it is needed to select both dependent and independent variables carefully. Correlation analysis is used to select the independent variables in obtaining more accurate and simpler MLR model, as explained in [13].

**3. Experiments' Results and Discussion**

The CIPIC HRTF Database used contains not only the measured HRIRs, but also some anthropometric parameters for 45 subjects, including the KEMAR mannequin with both small and large pinnae. The detail definitions of the all 27 anthropometric parameters are given in [1, 17]. Estimation of the listener's own HRIRs via his or her own anthropometric parameters will directly affect the feasibility and complexity of the system. It is clearly not advisable to introduce all parameters into the model. Some useful information will be concealed by the unnecessary parameters, which results in a worse regression model. Besides, many parameters are very difficult to be measured correctly. However, some parameters of 8 subjects are not available in the database. According to our anthropometric parameters selection, 8 selected parameters are included only in 35 subjects.

The performances of the estimated HRIRs on the horizontal plane were evaluated by the comparison of mean-square error of the differences between the estimated HRIRs and the measured HRIRs to the mean-square error of the measured HRIRs in percentage, which is defined by:

$$e_j(\theta) = 100\,\% \times \|\, \mathbf{h}_j(\theta) - \hat{\mathbf{h}}_j(\theta)\|^2 \,/\, \|\, \mathbf{h}_j(\theta)\|^2 \quad (6)$$

Where $\mathbf{h}_j(\theta)$ is the j-th measured $HRIR_{mp}$ with azimuth θ in horizontal plane, $\hat{\mathbf{h}}_j(\theta)$ is the corresponding estimated $HRIR_{mp}$ of $\mathbf{h}_j(\theta)$. If the error is larger, the performance of the estimated $HRIR_{mp}$ is worse, where better localization results will be achieved with small $e_j(\theta)$. The average errors are different among subjects in the database. The good performance of the estimated left-ear $HRIR_{mp}$ of a subject is not always followed by small error of the right-ear ones.

We have accomplished the modeling of $HRIRs_{mp}$ from 35 subjects, for sources on the horizontal plane using PCA with 10 basis functions. Here the average error, as defined by Equation 6, attained from this PCA modeling across 35 subjects and sources on the horizontal plane is 8.11%. In the following paragraphs, we will discuss a systematic and scientifically acceptable selection process of 8 anthropometric parameters from all 27 parameters defined in [16], for individualization of HRIRs.

As the first step, we individualized PCA models of $HRIRs_{mp}$ on the horizontal plane with all 27 anthropometric parameters using the MLR. By looking carefully in the CIPIC HRTF Database, these parameters were only completed for 35 subjects. The average error obtained in this condition is 11.85%. A series of experiments were then performed using each of 26, 25, 24, and 23 parameters. We selected k parameters out of 27 parameters using combinations of k out of 27 parameters, where k < 27. The resulted average errors are 12.46%, 13.04%, 13.60%, and 14.10%, respectively. The average error of using 23 parameters could be considered not dropping significantly compared to the use of 27 parameters (14.10% compared to 11.85%).

Observing the anthropometries database, there are many parameters, namely $x_4$, $x_5$, $x_{13}$, $d_8$, $\theta_1$, and $\theta_2$ that could not be simply measured because of the difficulty in determining the reference points and the parameters are so small (only several millimeters) that one needs a very precise instruments. These 6 parameters also have weak correlations with maximum ITD ($ITD_{max}$). The correlation coefficients between $ITD_{max}$ and each of these parameters, $\rho$, are 0.161, 0.098, 0.222, 0.397, 0.243, and 0.284 respectively. The experiments of using 21 parameters by excluding $x_4$, $x_5$, $x_{13}$, $d_8$, $\theta_1$, and $\theta_2$, resulted in average error of 15.29%.

We made an observation on the regression coefficients, ß, from MLR. The regression coefficients, ß, have very small near zero values and very small variations from MLR between PCWs; $\mathbf{w}_i(\theta)$, i = 1,2,...,10 where θ = azimuth angles, and each of $x_{11}$, $x_{14}$, $x_{15}$, $x_{16}$, and $x_{17}$. Thus, these parameters were also not included in the MLR model. At this point, we achieved average error of 17.97%.

The experiments of MLR modeling were continued further by ignoring the neck and torso items (i.e. $x_7$, $x_8$, $x_9$, and $x_{10}$). This is due to our observation on the responses of neck and torso that were represented by the last few samples in the HRIR, which have very small near zero values. Thus, their contributions to overall HRIR could be neglected. Nevertheless, torso top width, $x_9$, is represented by neck width, $x_6$, and shoulder width, $x_{12}$, from the chosen anthropometrie, while neck depth, $x_8$, is represented by head depth, $x_3$, for the depth characteristic. Individualization of $HRIRs_{mp}$ using MLR with 12 parameters, ignoring $x_7$, $x_8$, $x_9$, and $x_{10}$, resulted in average error of 20.13%.

In the last step, we reserved 8 anthropometric parameters, i.e. $x_1$, $x_3$, $x_6$, $x_{12}$, $d_1$, $d_3$, $d_5$, and $d_6$, which were explained before in [13]. Correlation analysis is applied to determine parameters that have strong correlation with $ITD_{max}$. Applying the $HRIRs_{mp}$ of 35 subjects for sources on the horizontal plane, individualization using MLR between the chosen parameters and PCWs, $\mathbf{w}_i(\theta)$, provides the average error of 22.22 %. This result is close to when we employed data from 37 subjects, that is 22.50% [13]. Here we came to the opinion that our selection method of anthropometric parameters before as in [13] is fully confirmed with the selection method explained in this paper. Table 1 summarizes the average errors caused by MLR models using various numbers of anthropometries explained above.

Table 1. Average Errors of MLR Models Using Various Numbers of Anthropometries

| No. of Anthropometries | Average Error (%) |
|---|---|
| 27 | 11.85 |
| 23 | 14.10 |
| 21 | 15.29 |
| 16 | 17.97 |
| 12 | 20.13 |
| 8 | 22.22 |

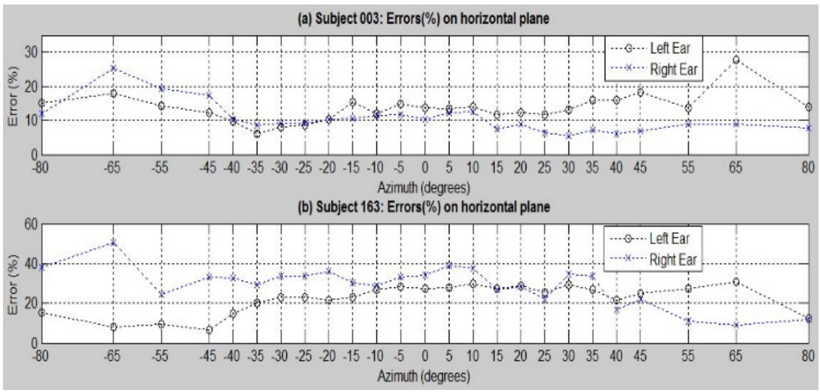


Figure 2. Left- and Right-Ear Errors of Subject 003 and Subject 163 in the Frontal Horizontal Plane

In the following subsections, the results of $HRIRs_{mp}$ individualization are shown using data from subject 003 and subject 163, to be compared to the results attained in [11]. Here eight anthropometric parameters, $x_1$, $x_3$, $x_6$, $x_{12}$, $d_1$, $d_3$, $d_5$, and $d_6$ were fed into MLR modeling in order to calculate regression coefficients. The regression coefficients then were applied in estimating the PCWs of $DIR_{mp}$ at each direction on the horizontal plane of each subject.

Figure 2 shows the left- and right-ear errors of azimuth angles in the frontal horizontal plane of subject 003 and subject 163. The left- and right-ear errors of subject 003 in the frontal horizontal plane are generally good, numbering below 20% except at 2 azimuth angles. The average error obtained from the data of HRIRs of left ear of subject 003 is 13.61%, and that of right ear is 10.56%. As can be seen from Figure 2(a), the errors are commonly larger for contralateral sources than for ipsilateral sources. The errors of subject 163 in the frontal horizontal plane can be said worse than those of subject 003. The average error obtained from the data of HRIRs of left ear of subject 163 is 22.28%, while of right ear is 29.22%. The sources of extreme contralateral for right ear, i.e. $-80^o$ and $-65^o$, result in errors of 37.82% and 50.79%, respectively. The unsystematic behavior of weights of PCs in the PCA across subjects and across directions causes difficulty for MLR to estimate them.

The estimated minimum phase HRIR of subject 003 can well approximate corresponding measured minimum phase HRIR particularly at the first 20 samples. Figure 3 shows the estimated and measured $HRIRs_{mp}$ of both left and right ear in the extreme locations in the frontal horizontal plane. The top, middle, and bottom panel corresponds to azimuth angle -80°, 0°, and 80° respectively.

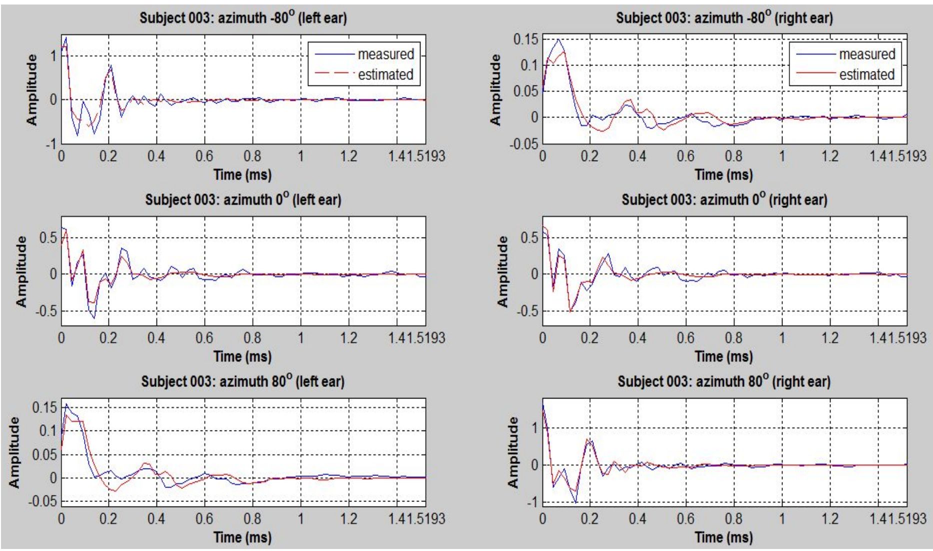


Figure 3. Measured and Estimated Minimum Phase HRIRs of Subject 003 in the Frontal Horizontal Plane

## 4. Conclusion

Our proposed selection method of anthropometric parameters explained here, which is simple, systematically and scientifically acceptable, is fully confirmed with the selection method as in [13]. The resulted eight anthropometries were incorporated in a simple and efficient individualization method of the model of minimum phase HRIRs based on multiple regression analysis. This proposed individualization method showed better performance in the objective simulation experiments than the performance in [11] which has been discussed in [13].

**Acknowledgements**

The authors gratefully thank to CIPIC Interface Laboratory of California University at Davis, USA for providing the CIPIC HRTF Database.

## References


[1] Algazi VR, Duda RO, Thompson DP, Avendano C. *The CIPIC HRTF Database*. Proceedings of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA). New Paltz. 2001: 99-102.
[2] Algazi VR, Duda RO, Morrison RP, Thompson DM. *Structural Composition and Decomposition of HRTF.* Proceedings of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA). New Paltz. 2001: 103-106.
[3] Wenzel EM, Arruda M, Kistler DJ, Wightman FL. Localization Using Nonindividualized Head-Related Transfer Functions. *The Journal of the Acoustical Society of America (JASA).* 1993; 94: 111-123.
[4] Blauert J. Spatial Hearing: The Psychophysics of Human Sound Localization. Revised Edition. Cambridge: The MIT Press. 1997.
[5] Shimada S, Hayashi M, Hayashi S. A Clustering Method for Sound Localization Transfer Functions. *Journal of Audio Engineering Society*. 1994; 42(7/8): 577-584.
[6] Middlebrooks JC. Virtual Localization Improved by Scaling Non-Individualized External-Ear Fransfer Functions in Frequency. *The Journal of the Acoustical Society of America (JASA).* 1999; 106(3): 1493-1510.
[7] Zotkin DN, Hwang J, Duraiswami R, Davis LS. *HRTF Personalization Using Anthropomentric Measurements.* IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA). New Paltz. 2003: 157-160.
[8] Katz BFG. Measurement and Calculation of Individual Head-Related Transfer Functions Using a Boundary Element Model Including the Measurement and Effect of Skin and Hair Impedance. PhD Dissertation. Pennsylvania: Graduate Program in Acoustics - Pennsylvania State University; 1998.
[9] Shin KH, Park Y. Enhanced Vertical Perception Through Head-Related Impulse Response Customization Based on Pinna Response Tuning in The Median Plane. *IEICE Transactions Fundamentals*. 2008; E91-A(1): 345-356.
[10] Hwang S, Park Y. *HRIR Customization in the Median Plane via Principal Components Analysis*. Proceedings of AES 31st Int. Conf. on New Directions in High Resolution Audio. London. 2007.
[11] Hu H, Zhou J, Ma H, Wu Z. *Head-Related Transfer Function Personalization Based on Multiple Regression Analysis*. Proceedings of IEEE Int. Conf. on Computational Intelligence and Security. Guangzhou. 2006; 2: 1829-1832.
[12] Hugeng, Wahab W, Gunawan D. Effective Preprocessing in Modeling Head-Related Impulse Responses Based on Principal Components Analysis. *Signal Processing: An Intl. Journal (SPIJ).* 2010; 4(4): 201-212.
[13] Hugeng, Wahab W, Gunawan D. *Enhanced Individualization of Head-Related Impulse Response Model in Horizontal Plane Based on Multiple Regression Analysis*. Proceedings of IEEE 2$^{nd}$ Int. Conf. on Computer Engineering and Applications (ICCEA). Bali Island. 2010; 2: 226-230.
[14] Hugeng, Wahab W, Gunawan D. Improved Method for Individualization of Head-Related Transfer Functions on Horizontal Plane Using Reduced Number of Anthropometric Measurements. *Journal of Telecommunications*. 2010; 2(2): 31-41.
[15] Hugeng. Individualization of Parametric Models of Head-Related Transfer Functions. PhD Dissertation. Depok: Dept. Elect. Eng. - University of Indonesia; 2011.
[16] Kulkarni A, Isabelle SK, Colburn HS. *On the Minimum-Phase Approximation of Head-Related Transfer Functions.* Proceedings of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA). New Paltz.1995: 84-87.
[17] CIPIC HRTF Database Files. Release 1.2. September 23, http:// interface.cipic.ucdavis.edu/CIL_html/CIL_HRTF_database.htm. 2004.