Springer

Journal of Mechanical Science and Technology 23 (2009) 3340~3348

Journal of
Mechanical
Science and
Technology

www.springerlink.com/content/1738-494x
DOI 10.1007/s12206-009-0925-z

Analysis on perceptual sensitivity to head-related impulse responses in the median plane †

Sungmok Hwang[1,*], Youngjin Park[2] and Youn-sik Park[2]

[1]*Vibration and Noise Research Part, Marine Research Institute, Samsung Heavy Industries, Geoje, 656-710, Korea*
[2]*Center for Noise and Vibration Control, Department of Mechanical Engineering, KAIST, Daejeon, 305-701, Korea*

(Manuscript Received June 25, 2009; Revised September 8, 2009; Accepted September 8, 2009)

Abstract

This study deals with the perceptual sensitivity to Head-Related Impulse Responses (HRIRs) in the median plane based on a series of subjective listening tests using a pair of headphones. First, the non-individualized HRIRs were modeled from 12 principal components (PCs) extracted from Principal Components Analysis (PCA) of the CIPIC HRTF database. The Just Noticeable Difference (JND) in weight of PCs (PCWs) at each elevation was estimated. It was not observed the common elevation-dependent tendency or PCW-dependent tendency of JND in PCWs across the five subjects who participated in the tests, and the inter-subject variation of JND in PCWs was large. The JND in HRIRs can be estimated indirectly from the JND in PCWs because the HRIRs can be represented by a linear summation of the PCs weighted by PCWs. The common elevation-dependent tendency of JND in Directional Impulse Responses (DIRs), which are the mean-subtracted HRIRs, across the five subjects can be found. The change in PCWs does not seem to contribute to our perception of sound source characteristics; however, the resulting change in HRIRs due to the change in PCWs seems to contribute. The subjects showed larger JND in DIRs in the frontal region than in the rear region. This means that our perception of sound source characteristics is more sensitive for frontal sources than rear sources.

Keywords: Head-related transfer functions; Head-related impulse response; Just noticeable difference; Principal components analysis

1. Introduction

Humans can perceive a sound direction as one of the crucial auditory abilities with Head-Related Transfer Functions (HRTFs), which is defined as the sound pressure at the eardrum divided by the sound pressure measured at the position of the head center with the head absent [1, 2]. The physical structures of a listener, such as head, external ear (pinna), shoulder, and torso, transform the spectrum of sound waves when they reach the listener's eardrum. This physical transform of sound waves is well encrypted in the HRTFs. If sounds are filtered with HRTFs and delivered to a listener through a pair of headphones, then a virtual acoustic environment can be produced and the listener feels the spatialized sounds appear to originate from the desired directions in the 3-dimensional space surrounding him/her [3, 4]. This also can be achieved by using 2-channel or multi-channel loudspeakers [5, 6]. Systems or techniques generating spatialized sounds and conveying them to a listener are referred to as Virtual Auditory Display (VAD) [7-9]. VAD has many promising applications such as entertainment including virtual reality, PC games, virtual audio including home theatre, auditory navigation, teleconference, and so on [10-16].

Since VAD generates convincing 3-dimensional sounds by real-time convolution of audio source with

† This paper was recommended for publication in revised form by Associate Editor Yeon June Kang
*Corresponding author. Tel.: +82 42 350 3076, Fax.: +82 42 350 8220
E-mail address: sungmok.Hwang@samsung.com

the HRTFs corresponding to the desired source positions, VAD requires a large library containing the HRTFs corresponding to source positions densely distributed in the 3-dimensional space. In other words, many HRTFs must be empirically measured and stored in order to generate well-spatialized sounds by VAD, and this makes real-time implementation difficult because of the large memory size requirement. Thus, it is necessary to model the HRTFs using only a few parameters while keeping the perceptual relevant features of the HRTFs for practical applications. Several methods for modeling HRTFs or Head-Related Impulse Responses (HRIRs), which are the time domain counterparts of the HRTFs, have already been developed, and they are mainly based on Pole-Zero approximation, State-Space model, and Principal Components Analysis (PCA) [3, 4, 17-29].

However, whatever methods are used for HRIR modeling, an error is inevitably introduced in the modeling process. The modeling error might be reduced when the HRIR is modeled as high order FIR/IIR filters in Pole-Zero modeling method, as high order state-space model, or with many Principal Components (PCs) in PCA-based method. In this case, however, the memory size must also be large. The modeling accuracy and the memory size are in a trade-off relationship. Then, one of the basic questions is how accurately the HRIRs should be modeled to provide no perceptual differences between the original and modeled HRIRs. To resolve this problem, investigation of the perceptually acceptable threshold in the change in HRIR is necessary. In other words, it is necessary to investigate the Just Noticeable Difference (JND) in HRIRs. This study deals with this problem based on a series of subjective listening tests using a pair of headphones. The research objective in this study is to estimate the JND in HRIRs.

Only the median-plane HRIRs are dealt with in this study because the azimuthal difference can be mainly dealt with the time delay or phase shift alone [1]. The PCA-based modeling method is used to model HRIRs, and the non-individualized HRIR, which is the mean of HRIRs of 45 subjects in the CIPIC HRTF database, is used in this study. The JND in weights of PCs used to model HRIRs is estimated, and the JND in HRIRs is estimated indirectly from the results. In section 2, the general methodology and detailed description of the apparatus and procedure used in the JND test are provided. Section 3 provides the detailed results and analysis, and section 4 contains a brief summary and conclusion.

2. Method

2.1 PCA of median-plane HRIRs

In this section, the general procedure of PCA of the median-plane HRIRs is briefly described. More specific details can be obtained from the authors' previous work [4].

We obtained a set of PCs from PCA of the median-plane HRIRs in the CIPIC HRTF database [30]. Before PCA was performed, we carried out a pre-processing on the median-plane HRIRs to remove the initial time delay and to extract the early response that lasts for 1.5 ms since the arrival of direct pulse [31]. The response of 1.5 ms includes the effects of pinna, head, shoulder, and torso. Thus, we can reduce the size of the dataset to be analyzed in PCA without loss of meaningful information by the pre-processing.

The original data matrix, $\mathbf{Y}(\in \mathbb{R}^{N\times M})$, is composed of the pre-processed median-plane HRIRs at the left ear. The response of 1.5 ms corresponds to 67 samples (sampling frequency: 44.1 kHz), and the number of median-plane HRIRs at each ear is 2205 (45 subject × 49 elevations from -45° to 225° with 5.625° intervals). Thus, the dimension of $\mathbf{Y}$ is 67×2205 in our case. The empirical mean of $\mathbf{Y}$ is needed to obtain the mean-subtracted HRIRs, and the empirical mean vector, $\mathbf{u}(\in \mathbb{R}^{N\times 1})$, is computed by

$$\mathbf{u}[n] = \frac{1}{M}\sum_{m=1}^{M}\mathbf{Y}[n,m]\,. \tag{1}$$

The mean-subtracted data matrix, $\mathbf{B}$, is computed by

$$\mathbf{B} = \mathbf{X} - \mathbf{u}\cdot\mathbf{h}\,, \tag{2}$$

where $\mathbf{h}(\in \mathbb{R}^{1\times M})$ is a row vector of all 1's. The PCs can be obtained from the covariance matrix of the mean-subtracted data matrix, and the k dominant PCs are the k eigenvectors of the covariance matrix corresponding to the k largest eigenvalues. The set of HRIRs can be reconstructed by a weighted linear combination of the PCs. A reasonable measure to determine the number of PCs is the percentage reconstruction error in the least-square sense, which is defined as

$$\%\,\text{error}(k) = \left\|\mathbf{Y}-\hat{\mathbf{Y}}_k\right\|_F^2 \Big/ \left\|\mathbf{Y}\right\|_F^2 \times 100\ (\%)\,, \tag{3}$$

where $\mathbf{Y}$ and $\hat{\mathbf{Y}}_k$ indicate the original data matrix, which is composed of the 2205 median-plane HRIRs, and the reconstructed data matrix from the k PCs, respectively. Subscript F indicates the matrix Frobenius norm. As the number of PCs increases, the reconstruction error decreases exponentially. In this study, the reconstruction error bound is set to be 5%, yielding 12 PCs. Detailed descriptions can be obtained in [4].

2.2 Apparatus and procedure

Eight male subjects (subject ID: CH, KB, LS, CY, KY, LY, PJ, KD) with normal hearing sensitivity participated in the experiment. The experiment was carried out in an audiometric room, where sound insulation is above approximately 50 dB in the frequency range above 1 kHz. The apparatus of the listening tests is shown in Fig. 1. Signals were generated by the MATLAB™ program, and an open-air headphone (AKG K1000) was driven by them through the computer sound card (Creative SB X-Fi Elite Pro) and the audio amplifier (Audio Analog Verdi Settanta). However, the headphone to ear-carnal transfer function (HpTF) should be equalized correctly because the correction of HpTF is very closely related to the localization performance [32]. Møller *et al.* [33] concluded that the inter-subject variability of the HpTF is significant and individualized equalization of the HpTF is recommended. Kulkarni and Colburn [34] claimed that the variability of HpTF is also high across headphone placements for a given listener. Thus, to achieve an accurate equalization, the HpTF had to be measured and equalized whenever each listener wore the headphone. Before the listening test was performed, each subject wore a headphone and a binaural microphone. The microphone was mounted inside a subject's pinna. Then, a broadband white noise (20 Hz ~ 20 kHz) was emitted through the headphone, and the output signals from the microphone were received. After the HpTF between the electronic input and output signals was computed, the inverse filter of HpTF was obtained using the FIR least squares inverse (length: 512 samples, delay: 256 samples).

The listening tests were carried out at nine elevations in the median plane from -30° to 210° with 30° intervals. The 0° is ahead of the listener, the 90° is above, and the 180° is behind. At each elevation, the subjects listened to two stimuli, the reference and perturbed stimuli, and then they judged whether the two stimuli were the same or different. At j-th elevation (j=1,2,…,9), the reference stimulus, s^r_j, was computed as

$$s_j^r[n] = \left\{ s[n] * h_j^r[n] \right\} * HpTF^{-1}[n] \ , \tag{4}$$

where * is the convolution operator. s is the bandpass-filtered white noise (1 kHz ~ 18 kHz) of duration 300 ms including the cosine-squared rise and fall times of 20 ms, respectively, and $HpTF^{-1}$ is the inverse filter of HpTF. h^r_j is the reference HRIR computed as

$$h_j^r[n] = \sum_{i=1}^{12} w_{j,i} \cdot \mathbf{v}_i[n] + \mathbf{u}[n] \, , \tag{5}$$

where $w_{j,i}$ is the mean of i-th PCWs of 45 subjects in the CIPIC HRTF database at j-th elevation. $\mathbf{v}_i$ is i-th PC, and $\mathbf{u}$ is the empirical mean in Eq. (1). At j-th elevation, the perturbed stimulus for k-th (k=1,2,…,12) PCW, $s^p_{j,k}$, was computed as

$$s_{j,k}^p[n] = \left\{ s[n] * h_{j,k}^p[n] \right\} * HpTF^{-1}[n] \, , \tag{6}$$

where $h^p_{j,k}$ is the perturbed HRIR for k-th PCW at j-th elevation computed as

$$h_{j,k}^p[n] = \sum_{\substack{i=1 \\ i \neq k}}^{12} w_{j,i} \cdot \mathbf{v}_i[n] + \left(w_{j,k} + \Delta_{j,k} \right) \cdot \mathbf{v}_k[n] + \mathbf{u}[n] \, , \tag{7}$$

where $\Delta_{j,k}$ is the perturbation for k-th PCW at j-th elevation. The perturbation (Δ) for the (m+1)th per

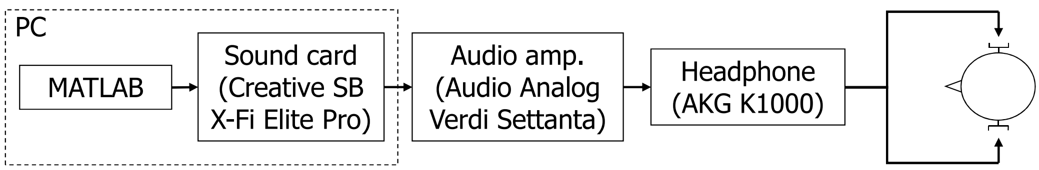

Fig. 1. Apparatus for the listening tests.

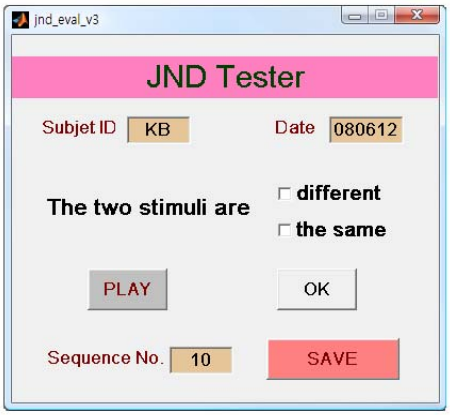

Fig. 2. MATLAB™ GUI for JND test.

turbed stimulus is set to be

$$\Delta^{m+1} = \begin{cases} 0.5\Delta^m, & \text{if } s^p \text{ is distinguishable} \\ 1.5\Delta^m, & \text{if } s^p \text{ is not distinguishable} \end{cases} \quad (8)$$

The initial perturbation, Δ^0, was set to be sufficiently large, 0.3. The JND is determined as Δ^m if the absolute difference between Δ^m and Δ^{m-1} is smaller than a threshold ($\Delta^0 \times 2^{-5}$).

The left and right channels of the headphone are driven by the same signal with the assumption of bilateral symmetry of ear shape. For convenient test procedure, MATLAB™ GUI depicted in Fig. 2 was used for the listening test. When the subject registered his ID and date, a block of test signals was generated. At each elevation, JNDs for 12 PCWs were determined, thus each block contains 108 sequences (9 elevations × 12 PCWs) in random order. The consecutive stimuli, $\{s^r \; s^p \; s^p \; s^r\}$ or $\{s^r \; s^p \; s^r \; s^p\}$, were presented by pushing the "PLAY" button. For each sequence, the task of the subject was to detect the difference between two stimuli independently of the nature of the differences (sound direction, timbre, etc.). Thus, the test represents a worst-case situation, since any possible difference may be detected. In this way it is possible to estimate the lowest threshold necessary to ensure that the reference and perturbed stimuli cannot be distinguished. After listening to the consecutive stimuli, the subject judged whether the consecutive stimuli were the same or different by using the check box, and pushed the "OK" button. Then, the sequence number was increased by one. When the number hits 108, the experiment was completed and the result was saved by pushing the "SAVE" button. Four blocks were given for each subject and test of each block was conducted on a different day. Approximately 40 minutes was needed to complete the experiment for one block on average.

Table 1. Statistically significant difference in four JND blocks for each subject.

		1st block	2nd block	3rd block
Subject CH	2nd block	**	N/A	N/A
	3rd block	**	-	N/A
	4th block	**	-	-
Subject KB	2nd block	-	N/A	N/A
	3rd block	-	-	N/A
	4th block	**	*	*
Subject LS	2nd block	-	N/A	N/A
	3rd block	-	**	N/A
	4th block	-	-	-
Subject CY	2nd block	*	N/A	N/A
	3rd block	**	**	N/A
	4th block	**	**	**
Subject KY	2nd block	-	N/A	N/A
	3rd block	**	**	N/A
	4th block	**	**	-
Subject LY	2nd block	-	N/A	N/A
	3rd block	-	-	N/A
	4th block	**	**	**
Subject PJ	2nd block	-	N/A	N/A
	3rd block	-	-	N/A
	4th block	**	**	**
Subject KD	2nd block	**	N/A	N/A
	3rd block	**	**	N/A
	4th block	**	**	**

* $p<0.05$, ** $p<0.01$

3. Result and analysis

First, the repeatability of the results obtained by each subject should be checked because the data without repeatability must be excluded in the analysis to obtain reliable results. Thus, a statistically significant difference was investigated based on the *t-test* for each subject. Table 1 summarizes the *t-test* results. In the case of subject CH, for example, a statistically significant difference was observed between the 1st block and the other three blocks ($p<0.01$), whereas no statistically significant difference was observed between the other blocks. Thus, the 1st block was not included in the JND analysis for him. By the same reason, the 4th block was excluded in the JND analysis for subjects KB, LY, and PJ. For subject LS, the 2nd block was excluded in the analysis. In the case of subjects CY, KY, and KD, a statistically significant difference was observed for all blocks, thus they were

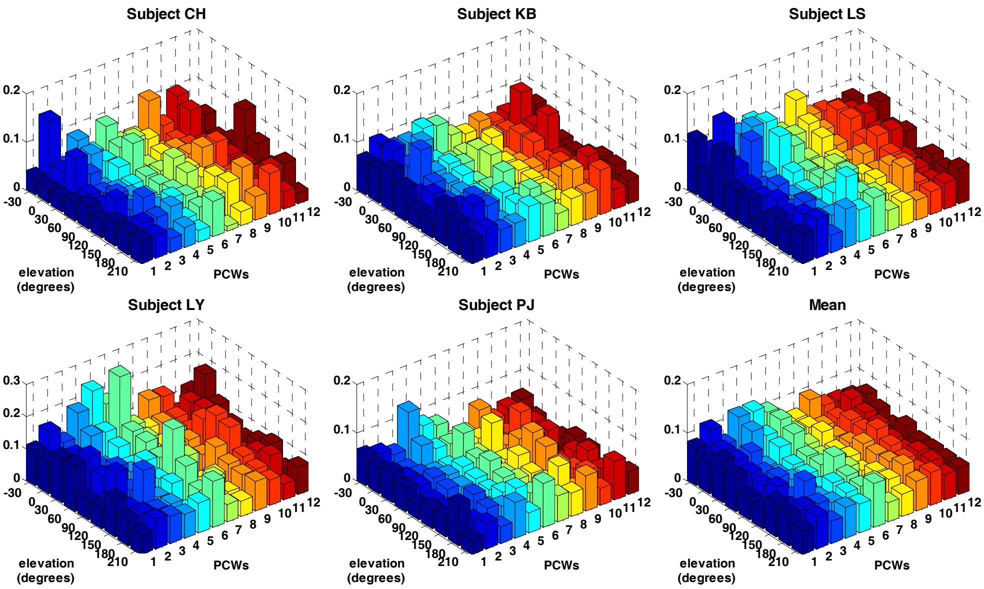

Fig. 3. JND in PCWs of each subject (left-top: CH, middle-top: KB, right-top: LS, left-bottom: LY, middle-bottom: PJ) and mean of five subjects (right-bottom).

excluded in the JND analysis. As a result, only five subjects' data were included in the following analysis.

The JND in PCWs is first estimated at each elevation in section 3.1. From the results, the JND in HRIRs is indirectly estimated in section 3.2.

3.1 JND in PCWs

Fig. 3 shows the mean of JNDs in PCWs of each subject (left-top panel: CH, middle-top panel: KB, right-top panel: LS, left-bottom panel LY, middle-bottom panel: PJ) and the mean of the five subjects (right-bottom panel). Common elevation-dependent or PCW-dependent tendencies are not observed across all subjects. This means that the inter-subject variation of JND in PCWs is larger than the inter-elevation variation or the inter-PCWs variation. To investigate the dependency of JND in PCWs on elevations and PCWs in detail, an Analysis of Variance (ANOVA) was performed at a level of significance 0.05 for each subject. Especially, the two-way ANOVA, which is a way of studying the effects of two factors separately, was applied because elevation and PCW were considered as factors in this study. Tables 2-6 show the two-way ANOVA table for each subject. In the case of subjects CH, KB, and LS, the *p*-value (Prob>F) of elevations is less than the level of significance (0.05) whereas the *p*-value of PCWs shows the opposite results.

Table 2. Two-Way ANOVA table for subject CH (Grand mean=0.0487, significance level=0.05).

Source of variation	SS	df	MS	F	*p-value* (Prob>F)
Elevations	0.00805	8	0.00101	2.19	**0.0358**
PCWs	0.00648	11	0.00059	1.28	0.2482
Error	0.04047	88	0.00046		
Total	0.05501	107			

Table 3. Two-Way ANOVA table for subject KB (Grand mean=0.0550, significance level=0.05).

Source of variation	SS	df	MS	F	*p-value* (Prob>F)
Elevations	0.00908	8	0.00113	3.15	**0.0035**
PCWs	0.00585	11	0.00053	1.47	0.1554
Error	0.03171	88	0.00036		
Total	0.04664	107			

Table 4. Two-Way ANOVA table for subject LS (Grand mean=0.0666, significance level=0.05).

Source of variation	SS	df	MS	F	*p-value* (Prob>F)
Elevations	0.01761	8	0.00220	5.06	**0.0000**
PCWs	0.00572	11	0.00052	1.20	0.3016
Error	0.03827	88	0.00043		
Total	0.06160	107			

Table 5. Two-Way ANOVA table for subject LY (Grand mean=0.1006, significance level=0.05).

Source of variation	SS	df	MS	F	*p-value* (Prob>F)
Elevations	0.02184	8	0.00273	2.24	**0.0319**
PCWs	0.04728	11	0.00430	3.52	**0.0004**
Error	0.10742	88	0.00122		
Total	0.17654	107			

Table 6. Two-Way ANOVA table for subject PJ (Grand mean=0.0467, significance level=0.05).

Source of variation	SS	df	MS	F	*p-value* (Prob>F)
Elevations	0.00493	8	0.00062	1.96	0.0602
PCWs	0.00689	11	0.00063	2.00	**0.0377**
Error	0.02759	88	0.00031		
Total	0.03941	107			

Therefore, it can be said that their JNDs in PCWS are mainly dependent on elevations. On the other hand, JND in PCWs of subject PJ is dependent not on elevations but on PCWs. In the case of subject LY, JND in PCWs is dependent on both elevation and PCWs. From the results of the two-way ANOVA, it can be concluded again that JND in PCWs is significantly different from subject to subject.

The grand mean of each subject, which is also noted in the title of each table, ranged from 0.0467 to 0.1006, and the overall mean, which is defined as mean of the five subjects' grand means, is 0.0635. Especially, the grand mean of subject LY is larger than the ones of the other subjects, whereas the difference of grand means of the other subjects is almost insignificant. This can also be seen in Fig. 3. The reason why his JND in PCWs is larger than JNDs of other subjects might be that his hearing sensitivity related with detection of the difference between two stimuli used in this study is slightly lower than other subjects, although he has no problem in real-life hearing.

3.2 JND in HRIRs

In section 3.1, JND in absolute values of PCWs was investigated and the common tendency of JND in PCWs across all subjects was not observed. In this section, JND not in absolute values of PCWs but in relative values of PCWs is dealt with in detail. The normalized perturbation (NP) for k-th PCW at j-th elevation can be defined as

$$\mathrm{NP}_{j,k} = \Delta_{j,k}^2 \Big/ \sum_{i=1}^{12} w_{j,i}{}^2 . \tag{9}$$

JND in NPs represents the square of JND in PCWs normalized by the sum of square of 12 mean PCWs. The NP has a special physical meaning for the HRIRs. The difference in PCWs is directly related to the difference in HRIRs because the HRIR is formed by a linear combination of PCs weighted by PCWs. The NP in Eq. (9) can be rewritten as

$$\mathrm{NP}_{j,k} = \frac{\left|\mathbf{h}_j^r - \mathbf{h}_{j,k}^p\right|^2}{\left|\mathbf{h}_j^r - \mathbf{u}\right|^2} = \frac{\left|\left(\mathbf{h}_j^r - \mathbf{u}\right) - \left(\mathbf{h}_{j,k}^p - \mathbf{u}\right)\right|^2}{\left|\mathbf{h}_j^r - \mathbf{u}\right|^2} . \tag{10}$$

The equivalence of Eq. (9) and Eq. (10) results from the orthogonality of the PCs. Eq. (10) indicates that the NP directly represents the power of the difference between the mean-subtracted reference and perturbed HRIRs normalized by the power of mean-subtracted reference HRIRs. The mean-subtracted HRIR is called Directional Impulse Response (DIR) as the mean-subtracted HRTF is called Directional Transfer Function. Thus, it can be said that JND in NPs represents how sensitive we are to differences in normalized DIRs.

Fig. 4 shows the mean of JNDs in the normalized DIRs, i.e. NPs, of each subject and mean of the five subjects with respect to varying elevations. In case of subject LY, JND in DIRs is significantly larger than JNDs of other subjects for all elevations. He also showed the largest grand mean of JNDs in PCWs, as described in section 3.1. Although he shows significantly larger JNDs in DIRs than the ones of other subjects, the common elevation-dependent tendency across all subjects can be observed. For all subjects, JNDs in the rear region (i.e., elevation > 90°) are larger than the ones in the frontal region (i.e., elevation < 90°). All subjects except for subject LY show significantly large JNDs in DIRs at elevations of 150° and 180°. In the case of subject LY, JND in DIRs at elevation of 150° is significantly large. On average across all subjects, it can be concluded that the subjects can distinguish the change in DIRs if the change in the power of difference in DIRs after normalization is larger than approximately 3~4 % except for some rear sources. For rear sources at elevations of 150° and 180°, the subjects can distinguish the change in DIRs if the power of difference in DIRs after normalization is above approximately 7~8 %.

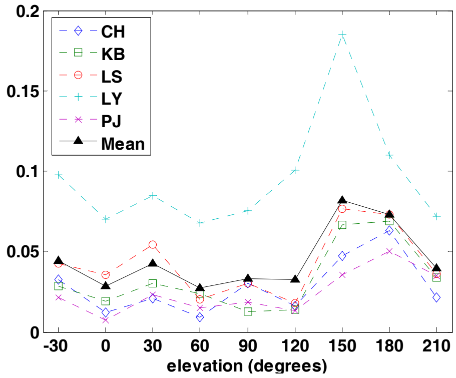

Fig. 4. JND in normalized DIRs.

4. Concluding remark

Perceptual sensitivity to HRIRs in the median plane was investigated based on a series of subjective listening tests using a pair of headphones. First, the non-individualized HRIRs were modeled from 12 PCs extracted from PCA of the CIPIC HRTF database. The JND in PCWs at each elevation was estimated. From the results, no common elevation-dependent tendency or PCW-dependent tendency of JND in PCWs was observed across the five subjects who participated in the tests, and JND in PCWs was significantly different from subject to subject. In other words, the inter-subject variation of JND in PCWs was larger than the inter-elevation variation or the inter-PCWs variation.

Since the HRIRs can be represented by a linear summation of the PCs weighted by PCWs, the JND in HRIRs was indirectly estimated from the JND in PCWs. The common elevation-dependent tendency of JND in DIRs, which is the mean-subtracted HRIRs, across the five subjects can be found. The subjects showed larger JND in DIRs in the frontal region than in the rear region. On average across the five subjects, it can be concluded that the change in DIRs after normalization is audible when it is larger than approximately 3~4 % except for some rear sources. The change in DIRs after normalization larger than approximately 7~8 % is audible for the rear sources at elevations of 150° and 180°.

From the results obtained in this study, the change in PCWs does not seem to contribute directly to our perception of sound source characteristics; however, the resulting change in HRIRs due to the change in PCWs seems to contribute. In addition, our perception of sound source characteristics is more sensitive for frontal sources than rear sources.

Most previous methods model the HRIRs for all source positions with the same accuracy. Based on the results obtained in this study, however, it is reasonable to model HRIR with variable accuracy according to the source position. Especially, in the median-plane, the HRIRs for frontal sources needed to be modeled more accurately than the ones for rear sources.

Acknowledgment

This study was supported by the BK21 program and the National Research Laboratory program through the Korea Science and Engineering Foundation funded by the Ministry of Education, Science, and Technology (R0A-2005-000-10112-0).

References

[1] J, Blauert, *Spatial hearing*, MIT Press, Cambridge, MA, USA, (1983).
[2] D. R. Begault, *3-D sound for virtual reality and multimedia*, Academic Press, Cambridge, MA, USA, (1994).
[3] D. J. Kistler and F. L. Wightman, A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction, *Journal of the Acoustical Society of America*, 91 (1992) 1637-1647.
[4] S. Hwang, Y. Park and Y. Park, Modeling and customization of head-related impulse responses based on general basis functions in time domain, *Acta Acustica united with Acustica*, 94 (2008) 965-980.
[5] J.-W. Choi, HRTF-field reproduction for robust virtual source imaging, *Journal of the Acoustical Society of Korea*, 23 (2004) 324-332.
[6] H.-Y. Kim, On the physical and perceptual precision of the multi-point control method in HRTF simulation, *Journal of the Korean Society for Noise and Vibration Engineering*, 18 (2008) 199-207.
[7] R. D. Shilling and B. Shinn-Cunningham, Virtual auditory displays. In K. M. Stanney (Ed.), *Handbook of virtual environments: Design, implementation, and applications*. Mahwah, NJ: Erlbaum, USA (2002).
[8] Y. Iwaya, Individualization of head-related transfer functions with tournament-style listening test: Listening with other's ears, *Acoustical Science and Technology*, 27 (2006) 340-343.
[9] A. Honda, H. Shibata, J. Gyoba, K. Saitou, Y.

Iwaya and Y. Suzuki, Transfer effects on sound localization performances from playing a virtual three-dimensional auditory game, *Applied Acoustics*, 68 (2007) 885-896.

[10] D. R. Begault and E. M. Wenzel, Techniques and applications for binaural sound manipulation in human-machine interfaces, *International Journal of Aviation Psychology*, 2 (1992) 1-22.

[11] J. M. Loomis, R. G. Golledge, and R. L. Klatzky, Navigation System for the blind: Auditory display modes and guidance, *Presence: Teleoperators and Virtual Environments*, 7 (1998) 193-203.

[12] S. Holland, D. Morse and H. Gedenryd, AudioGPS: Spatial audio navigation with a Minimal attention interface, *Personal and Ubiquitous Computing*, 6 (2002) 253-259.

[13] A. B. Oving, J. A. Veltman and A. W. Bronkhorst, Effectiveness of 3-D audio for warnings in the cockpit, *International Journal of Aviation Psychology*, 14 (2004) 257-276.

[14] R. S. Bolia, Special issues: Spatial audio displays for military aviation, *International Journal of Aviation Psychology*, 14 (2004) 233-238.

[15] T. Lokki and M. Gröhn, Navigation with auditory cues in a virtual environment, *IEEE Multimedia*, 12 (2005) 80-86.

[16] S.-J. Kwon, J.-H. Chun, S, Jang and M.-W. Suh, Driving performance analysis of the adaptive cruise controlled vehicle with a virtual reality simulation system, *Journal of Mechanical Science and Technology*, 20 (2006) 29-41.

[17] J. Chen, B. D. Van Veen and K. E. Hecox, External ear transfer function modeling: A beamforming approach, *Journal of the Acoustical Society of America*, 92 (1992) 1933-1944.

[18] M. A. Blommer and G. H. Wakefield, Pole-zero approximations for head-related transfer functions using a logarithmic error criterion, *IEEE Transactions on Speech and Audio Processing*, 5 (1997) 278-287.

[19] J. Mackenzie, J. Huopaniemi, and V. Välimäki, Low-order modelling of head-related transfer functions using balanced model truncation, *IEEE Signal Processing Letters*, 4 (1997) 39-41.

[20] Y. Haneda, S. Makino, Y. Kaneda and N. Kitawaki, Common-acoustical-pole and zero modeling of head-related transfer functions, *IEEE Transactions on Speech and Audio Processing*, 7 (1999) 188-196.

[21] E. A. Durant and G. H. Wakefield, Efficient model Fitting using a genetic algorithm: Pole-zero approximations of HRTFs, *IEEE Transactions on Speech and Audio Processing*, 10 (2002) 18-27.

[22] A. Kulkarni and H. S. Colburn, Infinite-impulse-response models of the head-related transfer function, *Journal of the Acoustical Society of America*, 115 (2004) 1714-1728.

[23] D. W. Grantham, J. A. Willhite, K. D. Frampton and D. H. Ashmead, Reduced order modeling of head related impulse responses for virtual acoustic displays, *Journal of the Acoustical Society of America*, 117 (2005) 3116-3125.

[24] W. L. Martens, Principal components analysis and resynthesis of spectral cues to perceived direction, *Proc. of International Computer Music Conference*, San Francisco, CA, USA, (1987) 274-281.

[25] J. C. Middlebrooks and D. M. Green, Observations on a principal components analysis of head-related transfer functions, *Journal of the Acoustical Society of America*, 92 (1992) 597-599.

[26] J. Chen, B. D. Van Veen and K. E. Hecox, A spatial feature extraction and regularization model for the head-related transfer function, *Journal of the Acoustical Society of America*, 97 (1995) 439-452.

[27] Z. Wu, F. H. Y. Chan, F. K. Lam and J. C. K. Chan, A time domain binaural model based on spatial feature extraction for the head-related transfer function, *Journal of the Acoustical Society of America*, 102 (1997) 2211-2218.

[28] K. Shin and Y. Park, Enhanced vertical perception through head-related impulse response customization based on pinna response tuning in the median plane, *IEICE Transactions on Fundamentals*, E91-A (2008) 345-356.

[29] S. Hwang and Y. Park, Interpretations on principal components analysis of head-related impulse responses in the median plane, *Journal of the Acoustical Society of America*, 123 (2008) EL65-71.

[30] V. R. Algazi, R. O. Duda, D. M. Thompson and C. Avendano, The CIPIC HRTF database, *Proc. of the 2001 IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, NY, USA, (2001) 99-102.

[31] S. Hwang and Y. Park, HRIR customization in the median plane via principal components analysis, *Proc. of the 31st International Audio Engineering Society conference*, London, UK, (2007).

[32] H. Møller, Fundamentals of binaural technology, *Applied Acoustics*, 36 (1992) 171-218.

[33] H. Møller, D. Hammershøi, C. B. Jensen and M. F.

Sørensen, Transfer characteristics of headphones measured on human ears, *Journal of the Audio Engineering Society*, 43 (1995) 203-217.
[34] A. Kulkarni and H. S. Colburn, Variability in the characterization of the headphone transfer-function, *Journal of the Acoustical Society of America*, 107 (2000) 1071-1074.

Sungmok Hwang received his B.S. degree in Mechanical Engineering from Yonsei University, South Korea, in 2004. He then received his M.S. and Ph.D. degrees from KAIST in 2006 and 2009, respectively. Dr. Hwang is currently a senior research engineer at Marine Research Institute, Samsung Heavy Industries. His research interests include sound and vibration control, audio signal processing, robot auditory system, binaural synthesis.

Youngjin Park received his B.S. and M.S. degrees in Mechanical Engineering from Seoul National University, South Korea, in 1980 and 1982, respectively. He then received his Ph.D. degree in Mechanical Engineering from University of Michigan, USA, in 1987. From 1987 to 1988, he worked as a research fellow at University of Michigan. He also worked as an assistant professor at NJIT, NJ, from 1988 to 1990. He joined the faculty of KAIST in 1990, where he is a Professor of Mechanical Engineering. His research interests include general control theories, virtual audio synthesis, active control of noise and vibration, system identification.

Youn-sik Park received his B.S. degree in Mechanical Engineering from Seoul National University, South Korea, in 1972. He then received his M.S. and Ph.D. degrees in Engineering Mechanics from Iowa State University, USA, in 1979 and 1981, respectively. From 1981 to 1984, he worked as a senior researcher at KAIST. He joined the faculty of KAIST in 1984, where he is a Professor of Mechanical Engineering. His research interests include vibration dynamics, system identification, experimental and analytical modal analysis, machine and manipulator dynamics.