PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

Spatial Principal Component Analysis of Head-Related Transfer Functions using their complex logarithm with unwrapping of phase

Shouichi Takane[(1)]

[(1)]Department of Information and Computer Science, Faculty of Systems Science and Technology, Akita Prefectural University, 84-4 Ebinokuchi, Tsuchiya, Yurihonjo, Akita, 015-0055 Japan, E-mail: takane@akita-pu.ac.jp

Abstract

Head-Related Transfer Function (HRTF) changes depending on sound source position and has strong individuality. In order to compactly represent the spatial change of the HRTF in acceptable accuracy, the Spatial Principal Component Analysis (SPCA) has been adopted in many researches. Accuracy of the reconstructed HRTFs with the SPCA is different depending on which domain is used for the SPCA. In this article, a new procedure for the SPCA of the HRTFs using complex logarithm of the HRTFs with unwrapping of phase was proposed. Moreover, the performance of the procedure was evaluated via the SPCA of a single HATS. The results showed that the proposed SPCA procedure has the potential to reconstruct the HRTFs with the smaller distortion in frequency domain using the smaller number of principal component than those in the conventional ones.

Keywords: Spatial Principal Component Analysis, Head-Related Transfer Function, Complex logarithm, Unwrapping of phase

1 INTRODUCTION

Head-Related Transfer Function (HRTF) is defined as an acoustic transfer function from sound acquired at a center point when a listener is absent to that acquired at the listener's ear[1]. The HRTF varies due to the sound source position and has strong individuality in both objective and subjective senses. Some systems have been proposed in order to synthesize the sound at the listener's ears based on the synthesis of binaural sound signals involving his/her HRTFs [2, 3, 4, 5, 6]. In order to achieve the accurate synthesis in these systems, the HRTFs are ideally acquired in all directions around the listener and optimized for him/her. While a study considering the effecient sampling scheme of the HRTF measurement exists[7], data size of such set of HRTFs may become numerous. Therefore a requirement arises that data size of his/her HRTFs acquired in all directions must be as compact as possible with their synthesis accuracy accomplished to some extent.

One of effective methods for the compact representation of the HRTFs is the Principal Component Analysis (PCA)[8, 9], or the spatial feature extraction method[10, 11, 12]. All of them have their theoretical basis on the PCA or Singular Value Decomposition (SVD) of matrices assembled by using the HRTFs or their corresponding impulse responses called the Head-Related Impulse Responses (HRIRs). In such kind of methods, the spatial variation of HRTFs is modeled by using small number of principal components or eigenvectors. In this paper, this method is called the Spatial PCA (SPCA) of the HRTFs/HRIRs, as Xie did[11].

Although the researches on the compact representation of HRTFs based on the SPCA exist [8, 9, 10, 11, 12], there are some differences in which domain was used for the SPCA among these studies. Kistler *et al.* used the logarithm of the HRTF amplitude[9], and Xie used the linear amplitude of the HRTF[11]. In these two studies, the HRTFs were assumed to have only minimum-phase component and the phase component were reconstructed by using the Hilbert transform[13]. Chen *et al.* used the complex-valued frequency spectrum of the HRTFs[10], and Wu *et al.* used the HRIRs[12]. No minimum-phase approximation was required in these studies. The SPCA can be successively and commonly adopted with the use of each parameter, and the modelling accuracy of the HRTFs was regarded sufficient in both objective and subjective viewpoints in all of these studies, with linear combination of relatively small number of principal component. However, it is known that difference in the modelling accuracy by choosing each of the four domains. Liang *et al.* compared the modelling accuracy between the linear and the logarithmic magnitude of the HRTFs in the horizontal plane, assuming that the minimum phase approximation was satisfactory[14]. Takane compared the reconstruction accuracy of the HRTFs among the results of the SPCA with each of the four domains[15]. When the accuracy was compared in the same number of principal components, both researches reached the same conclusion that the SPCA using the linear amplitudes of the HRTFs has the relatively better accuracy than that using the logarithmic ampitude under the assumption that the minimum phase approximation is accepted. Takane also concluded, in the case that the minimum phase approximation is not accepted, that the SPCA using the complex frequency spectrum of the HRTFs brings about better accuracy. In other words, the more compact representation of the spatial variation of the HRTFs is possible with the SPCA using the complex frequency spectrum of the HRTFs.

On the other side, Takane also mentioned that the accuracy in frequency domain, observed from the Spectral Distortion (SD) between the reconstructed and the original HRTFs, is exceptionally small when the logarithmic amplitude was chosen as the domain of the SPCA[15]. This property has a potential that the HRTFs are reconstructed by using the relatively small number of principal components when only the accuracy in frequency domain is focused. Since cues in the frequency domain are well known to be important[17, 18], the SPCA using the logarithmic domain may have the property desirable for the sound localization. Based on such consideration, yet another SPCA using the complex logarithm of the HRTFs is proposed in this article. Complex logarithm of the HRTFs is used for assembling the covariance matrix, and obtained principal component are also complex. It should be noted that the imaginary component of the complex logarithm of the HRTFs must be unwrapped. Effectiveness of the proposed method is demonstrated via the modeling of a set of HRTFs covering all directions. The proposed procedure does not require the minimum-phase approximation, and the most attractive feature of the proposed procedure is its lower (and better) distortion characteristics in frequency domain. This may come from the usage of the logarithm of the HRTFs. Kistler *et al.* also used the logarithm of the HRTFs under the minimum-phase assumption[9], but they did not report such feature since they did not compare the results of the SPCA using various domains.

2 MODELING OF THE HRTFs/HRIRs BASED ON SPCA

2.1 Outline of the SPCA

The following is general procedures for the SPCA of the HRTFs/HRIRs:

1. Average of a certain set of M vectors $\mathbf{g}_m$ $(m = 1,\cdots,M)$ is calculated as follows:

$$\mathbf{g}_{\mathrm{av}} = \frac{1}{M}\sum_{m=1}^{M}\mathbf{g}_m. \tag{1}$$

2. Covariance matrix $\mathbf{R}$ is obtained by calculating the following equation:

$$\mathbf{R} = \frac{1}{M}\sum_{m=1}^{M}(\mathbf{g}_m - \mathbf{g}_{\mathrm{av}})\cdot(\mathbf{g}_m - \mathbf{g}_{\mathrm{av}})^{\mathrm{H}}. \tag{2}$$

   Note that $^{\mathrm{H}}$ indicates the Hermitian transpose. Size of the matrix $\mathbf{R}$ is $N \times N$, where N is denoted as the length of the vector $\mathbf{g}_m$.

3. The matrix $\mathbf{R}$ is decomposed into N pairs of principal component (eigenvectors) and eigenvalues by solving the following eigenvalue problem:

$$\mathbf{R}\cdot\mathbf{q}_k = \lambda_k\cdot\mathbf{q}_k. \tag{3}$$

   As a result, a set of the eigenvalues and principal components, λ_k and $\mathbf{q}_k$ $(k = 1,\cdots,N)$, is obtained. Note that λ_k $(k = 1,\cdots,N)$ are sorted from the largest to the smallest, *i. e.*, $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \cdots \geq \lambda_N$, and the principal components, $\mathbf{q}_k$, are also arranged to the corresponding eigenvalues.

4. By using the matrix, $\mathbf{Q}$, with $\mathbf{q}_k$ in its column vector, the weighting vector, $\mathbf{w}_m$, corresponding to the m-th vector $\mathbf{g}_m$ is calculated as follows:

$$\mathbf{w}_m = \mathbf{Q}^{\mathrm{H}} \left(\mathbf{g}_m - \mathbf{g}_{\mathrm{av}}\right). \tag{4}$$

As a result of the PCA, the weight $\mathbf{w}_m$ is approximated by using $\mathbf{q}_1 \sim \mathbf{q}_K (1 \leq K \leq N)$ as follows:

$$(\mathbf{w}_m)_K = \mathbf{Q}_K^{\mathrm{H}} \left(\mathbf{g}_m - \mathbf{g}_{\mathrm{av}}\right), \tag{5}$$

where $\mathbf{Q}_K$ is a matrix with its column vectors $\mathbf{q}_1 \sim \mathbf{q}_K$. The length of the vector $(\mathbf{w}_m)_K$ becomes K.

The vectors and matrices are assumed to have complex values in the procedures stated above. If their values are real, the Hermitian transpose is changed to the simple tranpose $^{\mathrm{T}}$. The value m reflects the sound source position, and also the subject if the HRTFs/HRIRs of multiple subjects are used for assembling the covariance matrix.

The parameters are reconstructed by using the principal components as follows:

$$(\mathbf{g}_m)_K = \mathbf{Q}_K \cdot (\mathbf{w}_m)_K + \mathbf{g}_{\mathrm{av}}. \tag{6}$$

The computed parameter $(\mathbf{g}_m)_K$ becomes approximation when $K < N$, but this may have acceptable accuracy in principle when the Cummulative Proportion of Variance (CPV) $R^2(K)$ is close to 1.0. The CPV is defined by using the eigenvalues of the covariance matrix, λ_k $(k = 1, \cdots, N)$, as follows:

$$R^2(K) = \frac{\sum_{k=1}^{K} \lambda_k}{\sum_{k=1}^{N} \lambda_k}, \tag{7}$$

where N is the total number of component, equals to the length of the vector $\mathbf{g}_m (m = 1, \cdots, M)$.

2.2 Domains used for the SPCA

As mentioned in the Introduction, some kinds of parameters are subject to the assembly of the covariance matrix. Kistler *et al.* used the logarithm of the HRTF amplitude[9], and Xie used the amplitude of the HRTF[11]. In these studies, the minimum-phase properties of HRTFs are assumed. Chen *et al.* used the complex HRTF spectrum[10]. While these studies extracted the parameters from the HRTFs, Wu *et al.* used the HRIRs[12]. Yet another parameter, complex logarithm of the HRTFs, is newly listed in this paper. According to these studies and the new proposal, such differences are treated as the modelling domains in this paper. It is called domain "I" when the HRIR is used for the PCA, domain "C" when the complex HRTF spectrum is used, domain "F" when the amplitude spectrum of the HRTF is used, domain "L" when the real logarithm of the amplitude spectrum of the HRTF is used, domain "CL" when the complex logarithm of the HRTF is used together with unwrapping, and domain "CLN" when the complex logarithm of the HRTF is used without unwrapping. These definitions are summarized in Table 1.

Considering the SPCA in each domain, it is obvious in the domains C, F, L, CL and CLN that the frequency spectrum has the following symmetric relations for $k = 1, \cdots, N$, where N indicates the length of a HRTF:

$$H_m(k) = H_m^*(N-k), \tag{8}$$

$$\log H_m(k) = \log H_m^*(N-k), \tag{9}$$

$$A_m(k) = A_m(N-k), \quad L_m(k) = L_m(N-k). \tag{10}$$

Table 1. Correspondence between modelling conditions and symbols.

Domain	Symbol
HRIR	I
HRTF	C
Amplitude of HRTF	F
Real log-amplitude of HRTF	L
Complex log-amplitude of HRTF (w/ unwrapping)	CL
Complex log-amplitude of HRTF (w/o unwrapping)	CLN

In Eqs. (8) to (10), $H_m(k)$ is the k-th component of $\mathbf{H}_m$, $A_m(k) = |H_m(k)|$, $L_m(k) = \log A_m(k)$, and * indicates the conjugation. This means that the vector length is almost halved in these domains.

When the covariance matrices assembled in the domains I, C, F, L, CL and CLN are respectively denoted as $\mathbf{R}^{(\mathrm{I})}$, $\mathbf{R}^{(\mathrm{C})}$, $\mathbf{R}^{(\mathrm{F})}$, $\mathbf{R}^{(\mathrm{C})}$, $\mathbf{R}^{(\mathrm{CL})}$, and $\mathbf{R}^{(\mathrm{CLN})}$, size of $\mathbf{R}^{(\mathrm{I})}$ is $N \times N$, while that of $\mathbf{R}^{(\mathrm{C})}$, $\mathbf{R}^{(\mathrm{F})}$, $\mathbf{R}^{(\mathrm{CL})}$ and $\mathbf{R}^{(\mathrm{CLN})}$ is $(N/2+1) \times (N/2+1)$. The domains C, F, L, CL and CLN have the advantage in the smaller size of the covariance matrix, but the data in the frequency domain become complex in the domains C, CL and CLN. Therefore data size per one component of the covariance matrix becomes twice in these domains. Using the amount of data of $\mathbf{R}^{(\mathrm{I})}$ as reference, that of $\mathbf{R}^{(\mathrm{C})}$, $\mathbf{R}^{(\mathrm{CL})}$ and $\mathbf{R}^{(\mathrm{CLN})}$ is about 1/2, and that of $\mathbf{R}^{(\mathrm{F})}$ and $\mathbf{R}^{(\mathrm{L})}$ is around 1/4.

2.3 The SPCA using complex logarithm of HRTFs

The m-th HRIR and HRTF are respectively expressed as $\mathbf{h}_m$ and $\mathbf{H}_m$, and $\mathbf{H}_m$ is further decomposed into its amplitude and phase components as follows:

$$\mathbf{H}_m = \mathbf{A}_m \exp\{j\Theta_m\}, \tag{11}$$

where $\mathbf{A}_m$ and Θ_m indicates vectors stored the amplitude and phase component of the m-th HRTF, respectively. The domains I, C, F, L, CL and CLN mean that $\mathbf{h}_m$, $\mathbf{H}_m$, $\mathbf{A}_m$, $\mathbf{L}_m \equiv \log \mathbf{A}_m (m = 1, \cdots, M)$, $\log \mathbf{H}_m$ and $\mathrm{Log}\,\mathbf{H}_m$ are respectively used for the SPCA. Obviously the following equation stands in the domain CLN:

$$\mathrm{Log}\,\mathbf{H}_m = \log \mathbf{A}_m + j\Theta_m, \tag{12}$$

where Log$\mathbf{H}$ denotes the complex logarithm and its imaginary part of is in the range $[-\pi, \pi)$. When the above equation is calculated in computer, the imaginary part of this equation is automatically truncated in the range $[-\pi, \pi)$, leading to its discontinuities along the frequency. This must be affected to the reconstruction of HRTF by using the SPCA. Therefore, this imaginary part is unwrapped in the domain CL as follows:

$$\log \mathbf{H}_m = \log \mathbf{A}_m + j\mathrm{W}[\Theta_m], \tag{13}$$

where $\mathrm{W}[\cdot]$ is unwrapping function.

2.4 Calculation of HRTFs from the SPCA in various domains

When the m-th vector is reconstructed by using the first K principal components in the domains I, C, F, L, CL and CLN, they are respectively denoted as $\left(\mathbf{h}_m^{(\mathrm{I})}\right)_K$, $\left(\mathbf{H}_m^{(\mathrm{C})}\right)_K$, $\left(\mathbf{A}_m^{(\mathrm{F})}\right)_K$, $\left(\mathbf{L}_m^{(\mathrm{L})}\right)_K$, $\left(\log\mathbf{H}_m^{(\mathrm{CL})}\right)_K$ and $\left(\mathrm{Log}\mathbf{H}_m^{(\mathrm{CLN})}\right)_K$. In order to obtain the corresponding reconstructed HRIR and HRTF from the SPCA results in each domain, the length of the vectors are almost doubled with the relations in Eqs. (8)~(10) except for the domain I. Then the FFT or IFFT is used with the additional computation of the real and the complex exponentials for the domains L, CL and CLN. For example, $\left(\mathbf{h}_m^{(\mathrm{CL})}\right)_K$ is obtained by: calculating $\left(\mathbf{H}_m^{(\mathrm{CL})}\right)_K$ from $\left(\log\mathbf{H}_m^{(\mathrm{CL})}\right)_K$, then the length of $\left(\mathbf{H}_m^{(\mathrm{CL})}\right)_K$ almost doubled, and calculating the IFFT of the resultant vector.

3 RELATION BETWEEN ACCURACY AND NUMBER OF PRINCIPAL COMPONENTS

A database of HRIRs of KEMAR HATS (Head And Torso Simulator) presented by Media lab. of MIT[19] was used. This database involves 710 pairs of HRIRs (total: 1420) with sampling frequency of 44.1 kHz. This value corresponds to M in Eqs. (1) and (2).

3.1 Conditions of analysis

The initial delay in each response was extracted, then 256 sample points were taken as the data for the analysis, windowing with latter half of 512-points Blackman-Harris window function adjusting its peak at that of the HRIR. The SPCA was executed by constructing the covariance matrices from HRIRs (called as domain I), HRTFs (domain C), amplitude of HRTFs (domain F), log-amplitude of HRTFs (domain L), complex log-amplitude of HRTFs with unwrapping (domain CL) and that without unwrapping (domain CLN). The HRTFs/HRIRs in all directions (710 directions×2 ears) were used, and the average vector (Eq. (1)) and the covariance matrix (Eq. (2)) were calculated in each domain.

3.2 Cumulative Proportion of Variance (CPV)

When the PCA is utilized for some data, the cumulative proportion of variance $R^2(K)$ defined as Eq. (7) is used for the reference indicating how much variance in data is covered by using the first K principal components. Change of the cumulative proportion of variance in each domain is plotted in Fig. 1. It is found out from this figure that the CPV is monotonically increased and converges to 1.0 as the number of component is increased in all domain. Among these six domains, the cases CLN and CL respectively have the slowest and the fastest increase as the number of principal components increases. This indicates that the the smallest number of principal components can cover a certain proportion of variance in the domain CL and the largest number is required for the domain CLN. Since the difference between the domain CL and CLN is whether the unwrapping of phase was executed or not, the difference in Fig. 1 certainly comes from the unwrapping processing.

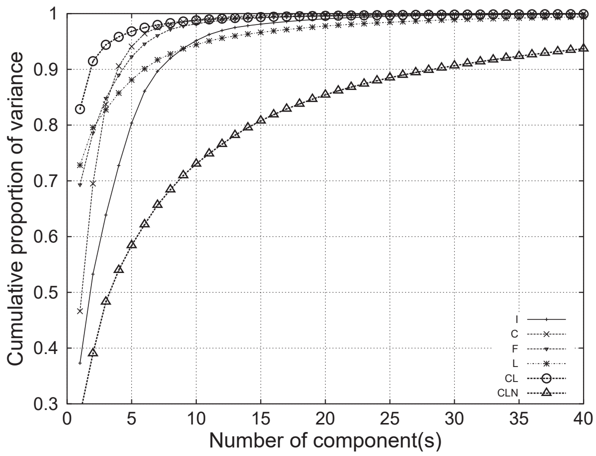

Fig. 1. Change in cumulative proportion of variance (CPV) with number of component in PCA. Lines and marks for the cases C, L, CL and CLN are relatively thicker and larger than those for the cases I and F.

The previous studies had various reference values in the proportion of variance, more than which the corresponding principal components are omitted. Kistler *et al.* set this value to 0.90[9], Chen *et al.* and Wu *et al.* set this value to 0.999[10, 12], and Xie set this to around 0.98[11]. Since the amount of data and analyzing purposes are different among these studies, direct comparison among these values is impossible. However, it can be said that the value more than 0.9 is set in these studies. After these studies, the least numbers of component to cover four values of the CPV, 0.90, 0.95, 0.99 and 0.999 are indicated in Table 2. Seeing Table 2, the domain CLN has the numbers of component largest in all of the CPV values. Since the dimension of the covariance matrix for this domain is 129×129, around 81% (105/129) of principal component are required to cover 0.999 of the variance in the domain CLN. On the contrary, the domain CL has the smallest values among six domains when the CPV value is 0.90 or 0.95, and the domain C has the smallest value when the CPV value is 0.99 or 0.999. From these comparison, it can be stated that the domains C and CL have almost the same characteristics. However, the domain C requires number of component less than that for the domain CL when the CPV value is raised to 0.999.

Table 2. The least number of component to cover the cumulative proportion of variance in each case.

Case	Variance			
	0.90	0.95	0.99	0.999
I	8	10	20	39
C	4	6	11	20
F	5	7	14	31
L	6	11	32	78
CL	2	4	12	40
CLN	29	47	84	105

3.3 Comparison of reconstruction accuracy

The CPV of the covariance matrix is an effective criterion for the coverage of variance with a certain number of component derived from the PCA. However, Takane indicated that the direct comparison of the CPV against a certain number of principal component among each domain is not possible[15]. Therefore, the reconstruction accuracy obtained for each domain is compared in this section. For this purpose, Spectral Distortion (SD) and Signal-to-Distortion Ratio (SDR) is used.

The SD is defined as standard deviation in log-amplitude of two frequency spectra as follows:

$$\mathrm{SD}[\mathbf{A},\mathbf{A}] = \sqrt{\frac{1}{N_f}\sum_{k=1}^{N_f} 20\log_{10}\left|\frac{A(k)}{\widehat{A}(k)}\right|} \quad [\mathrm{dB}], \tag{14}$$

where $\mathbf{A}$ and $\widehat{\mathbf{A}}$ are the frequency amplitude spectrum of the original and the estimated responses, respectively, and $A(k)$ and $\widehat{A}(k)$ are the k-th component of the vectors $\mathbf{A}$ and $\widehat{\mathbf{A}}$, respectively. Here the value of N_f is set so that it gets closest to the frequency of 20 kHz. The closer $\widehat{\mathbf{A}}$ is to $\mathbf{A}$, the smaller the SD is.

Signal-to-Distortion Ratio (SDR) is defined as the level difference between the energy of the original impulse response and that of the error as follows:

$$\mathrm{SDR}\left[\mathbf{h},\widehat{\mathbf{h}}\right] = 10\log_{10}\frac{||\mathbf{h}||}{\left\|\mathbf{h}-\widehat{\mathbf{h}}\right\|} \quad [\mathrm{dB}], \tag{15}$$

where $||\cdot||$ indicates the Euclid norm of the vector. The closer $\widehat{\mathbf{h}}$ is to $\mathbf{h}$, the larger the SDR is.

As for the SDRs in the domains F and L, the original impulse responses are ones constructed with its minimum phase component, which are different from the ones in the other domains. The SDRs in each case were computed as how much the reconstructed impulse response differs from the desired one. It is noted that this point is related to the calculation of SDRs, and not to that of SDs, since the SD is computed by using only magnitude of the reconstructed and the original HRTFs.

Accuracy in frequency domain with number of component set to K is computed in a certain domain X (X=I,C,F,L,CL,CLN), the overall average of SD and SDR weres calculated by using the following equations:

$$\mathrm{AvSD}(\mathrm{X},K) = \sqrt{\frac{1}{M}\sum_{m=1}^{M}\left\{\mathrm{SD}\left[\mathbf{A}_m,\left(\mathbf{A}_m^{(\mathrm{X})}\right)_K\right]\right\}^2}, \tag{16}$$

$$\mathrm{AvSDR}(\mathrm{X},K) = 10\log_{10}\left\{\frac{1}{M}\sum_{m=1}^{M}10^{\mathrm{SDR}\left[\mathbf{h}_m,\left(\mathbf{h}_m^{(\mathrm{X})}\right)_K\right]/10}\right\}. \tag{17}$$

Changes of the average SD and SDR with number of component(s) K in each domain are plotted in Fig. 2 and Fig. 3, respectively.

From Fig. 2, it is found out that the least average SD is obtained for the domain L when the number of component is less than 12. When the number of component is more than 12, the domains L and CL have almost the same average SD, and they acquire the lowest values of average SD among all domains. They correspond to the domains that the logarithm of the HRTFs is used for assembling the covariance matrix. The domain C has the SD smaller and almost equal to the domains C and CL when the number of component is increased. The average SD remains relatively large for the domains I and F.

From Fig. 3, the highest SDR is obtained for the domain C over all numbers of principal component. The characteristics of that for the domain CL is similar to that of the domain L, and the domain CLN has the lowest average SDR over all numbers of principal components. As a result, the proposed SPCA method does not have the best characteristics, compared with the other domains. One possible reason for this incorrespondence in the reconstruction accuracies between frequency and time domains is that the local difference in frequency domain may bring about a relatively large difference in time domain. However, it is found that the acceptable accuracy is achieved for the larger number of component in all domains except the domain CLN. At least the proposed method is appropriate in the sense that the reconstruction accuracy for the domain CL is in the similar level to that for all domains.

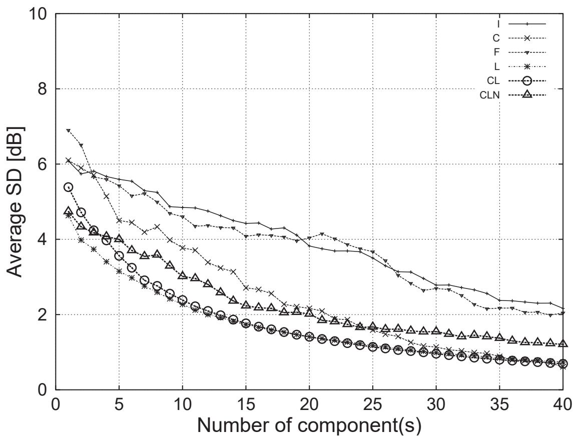

Fig. 2. Change of average SD with number of component(s) in six domains, respectively expressed as AvSD(I,K), AvSD(C,K), AvSD(F,K), AvSD(L,K), AvSD(CL,K) and AvSD(CLN,K) ($K = 1, \cdots, 40$).

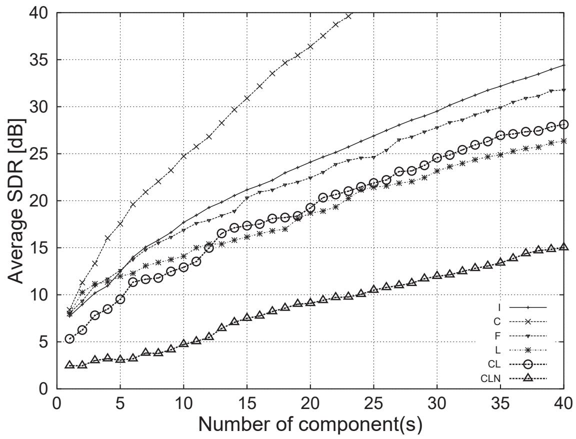

Fig. 3. Change of average SDR with number of component(s) in six cases, respectively expressed as AvSDR(I,K), AvSDR(C,K), AvSDR(F,K), AvSDR(L,K), AvSDR(CL,K) and AvSDR(CLN,K) ($K = 1, \cdots, 40$).

4 CONCLUDING REMARKS

In this paper, the SPCA of the HRTFs using complex logarithm of the HRTFs with unwrapping of phase was proposed. Although there exist the articles in which the logarithm of the HRTFs is applied to the SPCA, their phase component are assumed to have minimum-phase. On the contrary, the proposed procedure requires no such assumption. Moreover, the performance of this method was examined via the SPCA of a single HATS. Summarizing the above results, the proposed SPCA procedure has the potential to reconstruct the HRTFs with the smaller SD using the smaller number of principal components than those in the other domain.

The reconstruction accuracy was evaluated by mainly using the SD in this paper, but it is also necessary to confirm the above mentioned feature via subjective experiment. This is regarded as one of the important future work concerning the proposed procedure in this article.

ACKNOWLEDGMENTS

A part of this study was supported by JSPS KAKENHI (JP17K00242).

REFERENCES

[1] J. Blauert, Spatial Hearing (MIT Press, 1983).

[2] M. Otani and T. Hirahara, "A dynamic auditory display: its design, performance, and problems in HRTF switching," Proc. Japan-China Joint Conf. Acoust., SS-1-3, 2007.

[3] S. Yairi, Y. Iwaya and Y. Suzuki, "Estimation of detection threshold of system latency of virtual auditory display," Applied Acoustics, **68**(8), 851-863, 2007.

[4] J. D. Miller, "Slab: a software-based real-time virtual acoustic environment rendering system," Proc. 2001 ICAD, 279-280, 2001.

[5] S. Takane, S. Takahashi, Y. Suzuki, T. Miyajima, "Elementary real-time implementation of a virtual acoustic display based on ADVISE," Acoust. Sci. & Tech., **24**(5), 304-310, 2003.

[6] D. R. Begault, 3-D sound for virtual reality and multimedia (AP Professional, 1994).

[7] W. Zhang, M. Zhang, R. A. Kennedy and T. D. Abhayapala, "On high-resolution Head-Related Transfer Function Measurements: An efficient sampling scheme," IEEE Trans. Audio, Speech & Lang. Process., **20**(2), 575-584, 2012.

[8] W. L. Martens, "Principal component analysis and resynthesis of spectral cues to perceived direction," Proc. ICMC, 274-281, 1987.

[9] D. J. Kistler and F. L. Wightmann, "A model of head-related transfer functions based on principal component analysis and minimum-phase reconstruction," J. Acoust. Soc. Am., **91**(3), 1637-1647, 1992.

[10] J. Chen, B. D. Van Veen and K. E. Hecox, "A spatial feature extraction and regularization model for the head-related transfer functions," J. Acoust. Soc. Am., **97**(1), 439-452, 1995.

[11] Bo-Sun Xie, "Recovery of individual head-related transfer functions from a small set of measurements," J. Acoust. Soc. Am., **132**(1), 282-294, 2012.

[12] Z. Wu, F. H. Y. Chan, F. K. Lam, J. C. K. Chan, "A time domain binaural model based on spatial feature extraction for the head-related transfer functions," J. Acoust. Soc. Am., **102**(4), 2211-2218, 1995.

[13] A. V. Oppenheim and R. W. Schafer, Discrete-time signal processing, third edition (Prentice Hall, 2010).

[14] Z. Liang, B. Xie and X. Zhong, "Comparison of principal component analysis of linear and logarithmic magnitude of Head-Related Transfer Functions," Proc. 2nd IEEE International Congress on Image and Signal Processing, 2009.

[15] S. Takane, "Effect of domain selection for compact representation of spatial variation of Head-Related Transfer Function in all directions based on Spatial Principal Components Analysis," Applied Acoustics, **101**(1), 64-77, 2016.

[16] T. Nishino, S. Kajita, K. Takeda and F. Itakura, "Interpolating head related transfer functions in median plane," Proc. 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, NY, 17-20, 1999.

[17] F. Asano, Y. Suzuki and T. Sone, "Role of spectral cues in median plane localization," J. Acoust. Soc. Am., **88**(1), 159-168, 1990.

[18] K. Iida, M. Itoh, A. Itagaki and M. Morimoto, "Median plane localization using a parametric model of the head-related transfer function based on spectral cues," Applied Acoustics, **68**, 835-850, 2007.

[19] W. G. Gardner and K. D. Martin, "HRTF measurements of a KEMAR," J. Acoust. Soc. Am., **97**, 3907-3908, 1995.