

!"#\$% "& "'#" )&\*+& "\*, -& ". (/ -&0+#+&1+(-&!  
\*2+(344.,1 "\*, -&5(-0(!+12& -.- '6("&7(8& ',&++#,&'(9:; !

!"#\$!%&

# CERTIFICATE

## OF ACHIEVEMENT

!2,5(1+#\*,0,1 " "+(5(4#+5+&\*+7\*(-

+~~XXXX~~

"

---

0-#(\*2+(1-&\*#,\$\*, -&("5(

25\$/ " # \$ % \$ & ' \$ # !

**E W !**

**W**

&' A2 - B2+

= E"=>?! F"#.3A2+:31\*: ) \*+, - \* . \* / \* + \* E G \* H \* + 1 \*

PAPER

## Deployment of models of HRIRs using PCA in digital signal processing board

To cite this article: Hugeng Hugeng and Tommy Sutiawan 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **508** 012079

View the [article online](#) for updates and enhancements.



**IOP | ebooks™**

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the collection - download the first chapter of every title for free.

# Deployment of models of HRIRs using PCA in digital signal processing board

**Hugeng Hugeng<sup>1\*</sup> and Tommy Sutiawan<sup>2</sup>**

<sup>1</sup>Electrical Engineering Department, Universitas Tarumanagara, Jl. Letjen S. Parman No. 1, Jakarta 11440, Indonesia.

<sup>2</sup>Computer Engineering Department, Universitas Multimedia Nusantara, Jl. Scientia Boulevard, Gading Serpong, Tangerang 15811, Indonesia

\*hugeng@ft.untar.ac.id

**Abstract.** Three-dimensional sound effect has been around us when we are playing video games and watching movies in theatres. This happens because many researches have been developed to deploy head-related impulse responses (HRIRs) for 3-D sound reproduction. In reproducing 3-D sound, monaural sound is convolved with a pair of HRIRs. In this research, PKU-IOA HRTF Database has been used. An original HRIR consists of 1024 samples with 65,536 Hz sampling rates. Pre-processing is performed on the HRIRs, so that we obtain five data types including the original ones, i.e. original HRIRs (type 1); original HRIRs chopped to first 512 samples (type 2); and 349 samples (type 3); HRIRs with 256 samples and 48 kHz sampling rate (type 4) that originating from original HRIRs with 349 samples that are down-sampled to 48 kHz; and the last data type is minimum phase HRIRs with first 200 samples that are reconstructed from original HRIRs (type 5). Each data type is then modelled using principal components analysis (PCA). The averages of percentage errors are respectively 24.259%, 24.244%, 24.242%, 21.080%, and 5.422% for type 1 to type 5. The best-resulted model is minimum phase HRIRs (type 5); these HRIRs are then reconstructed and deployed in Digital Signal Processor (DSP) Board TMS320C5535 to reproduce 3-D sound from monaural one.

## 1. Introduction

Research and development for 3-D sound aim to get users a more engaging experience while watching a movie or playing video games with virtual reality technology because the user feels like coming into the world that being watched or played [1]. It is based on the knowledge of the human ability to determine sound position in 3-D space, from distance or length, the direction from above and below, in front and back, and to both sides of the sound positions even though human has only two ears. This is possible because our brain, inner ear and outer ear (pinna) work together to determine the position of the sound [2][3].

HRIR has been the focus of acoustic and psychoacoustic research. Explicitly HRIR can be referred to as an acoustic filter of human beings that filters out the spatial sound that falls on the head, torso, and pinna. HRIRs can be used to simulate spatial or 3-D sound through headphones and loudspeakers. The HRIR data set is estimated to contain all the sonic cues that affect the perception of spatial sound positions. However, massive HRIR measurements need intensive effort from an individual to stay still for a long time period to measure HRIRs from all points which sound may come. With HRIR modelling, we can model other non-existing HRIRs from a set of HRIRs in a database. Thus, less HRIRs are only needed to be measured which reduces the effort and cost needed.



In this research, we implement modelled HRIRs conducted by principal components analysis (PCA) on digital signal processor (DSP) board. Previously, similar researches had been done before that HRIR data can be implemented in a DSP board to produce 3-D sound sources [4][5]. However, HRIR data implemented in [4] used the interpolation of original measured HRIRs from CIPIC HRTF Database and was not processed first using PCA modelling which was then reconstructed to be implemented into the DSP board. The researchers in [5] used DSP board TMS320C6713DSK development kit that served as a hardware platform for the designed spatial audio synthesis system. This DSP board has more capacity than the DSP board used in our research, but it is more complicated.

The modelling HRIRs using PCA was used because PCA is the most effective way of describing HRIRs with reduced data. The statement was made by Marten who had implemented the PCA on HRTFs by describing the original data sets with some orthogonal components and corresponding weight variables [6]. The weight variations described with appropriate functions can present HRTF very effectively with only a few data coefficients. They are also possible to produce spatial sounds for arbitrary positions that are not measurable since each weight variation is represented by a continuous function whose value is readable for all azimuth positions [7].

The rest of the paper is organized as follows. In Section 2, we discuss some basic concepts about PKU-IOA HRTF Database and DSP board used. Section 3 consists of the research methodology. Section 4 provides the results of the research and the discussion about them. The conclusion is explained in Section 5.

## 2. Literature review

In this section, some details of DSP board used and PKU-IOA HRTF Database are explained.

### 2.1. PKU-IOA HRTF database

A pair of Head-Related Impulse Responses (HRIRs) from a single spatial point is recorded using a Knowles Electronics Mannequin for Acoustics Research (KEMAR) mannequin in one file in the PKU-IOA-HRTF Database and stored as a dat file where the length of each file is 2048 samples, and the data type is a double data type. The first 1024 samples is HRIR from the left ear (large ear), and then last 1024 samples is HRIR from the right ear (small ear). The sample rate of this database is 65,536 Hz. This database can be downloaded from <http://www.cis.pku.edu.cn/auditory/Staff/Dr.Qu.files/Qu-HRTF-Database.html> [8].

The PKU-IOA HRTF Database is a high-spatial resolution Head-Related Transfer Function (HRTF) database with spacing between 20 cm to 160 cm, including 20, 30, 40, 50, 75, 100, 130, and 160 cm; elevation from -40 to 90 degrees at a step of 10 degrees. The angles of the azimuth range from 0 to 360 degrees with different steps which depend on the elevations. On elevation from -40 to 50 degrees, the azimuth step is 5 degrees; at the step of 10 degrees on elevation of 60 degrees; at the step of 15 degrees on 70 degrees elevation; at the step of 30 degrees at elevation 80 degrees, and at the step of 360 degrees at 90 degrees elevation.

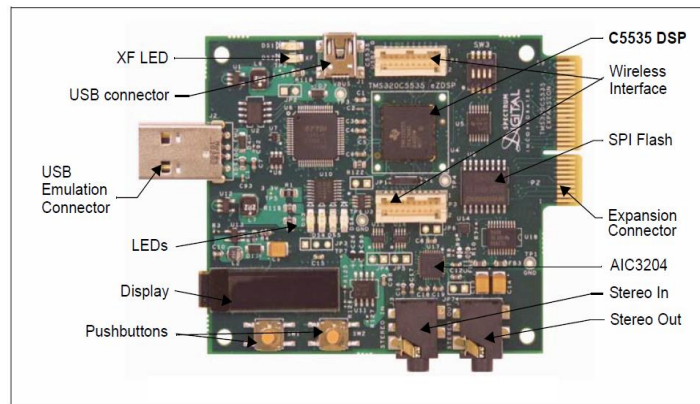
### 2.2. Digital signal processor board

After evaluation of distinguish digital signal processor (DSP) architectures, we have decided to use C5535 eZdsp architecture of Texas Instruments. The following features of his architecture are very important and useful in the presented task:

- Texas Instrument's TMS320C5535 Digital Signal Processor
- Texas Instruments TLV320AIC3204 Stereo Codec (stereo in, stereo out)
- USB 2.0 interface to C5535 processor
- 8 Mbytes SPI flash
- Embedded USB XDS100 JTAG emulator
- Compatible with Texas Instruments Code Composer Studio

A 16-bit fix point EZDSP TMS320C5535 was chosen as the signal processing platform in the realized project. It features sufficient processing power while being enough current consumption efficient to build a mobile device based on this DSP. Texas Instruments released a DSP optimized C/C++ compiler

embedded in Code Composer Studio programming environment, that is capable of very efficient machine code generation. The key features of eZdsp TMS320C5535 can be seen at Figure 1.



**Figure 1.** Key features of eZdsp TMS320C5535 [9].

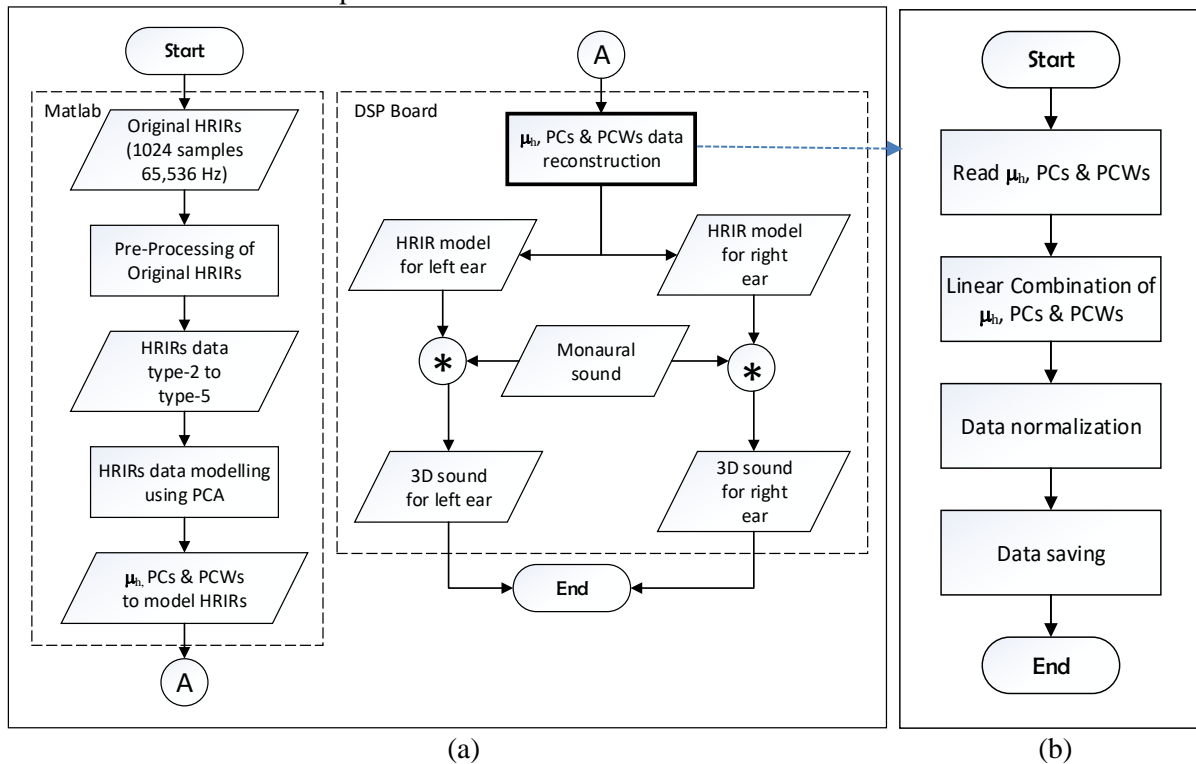
### 3. Research methodology

This research is divided into 2 main parts, those are modelling the HRIRs from the PKU-IOA HRTF Database using PCA in Matlab; and the implementation of the resulted PCA models of HRIRs into reconstructed HRIR data in DSP Board, as seen in Figure 2. Figure 2a describes the research scheme performed in our work. For each distance of sound source from head centre, there are 793 positions in sphere, that means there are 793 HRIRs for each ear. In performing PCA modelling, we use two kinds of HRIR data; i.e. HRIR data from 3 distances (50 cm, 75 cm, and 100 cm) and HRIR data only from distance 75 cm. All HRIR data come from original HRIRs in the used database with sampling rate of 65,536 Hz and 1024 samples in each HRIR. The goal of pre-processing of original HRIRs is to investigate which HRIR data type after a certain pre-processing results in the best HRIR model. Pre-processing is performed on the HRIRs, so that we obtain five data types including the original ones, i.e. original HRIRs (type 1); original HRIRs chopped to first 512 samples (type 2); and chopped to first 349 samples (type 3); HRIRs with 256 samples and 48 kHz sampling rate (type 4) that originating from first 349 samples of original HRIRs which are down-sampled to 48 kHz; and the last data type is minimum phase HRIRs with first 200 samples that are reconstructed from original HRIRs (type 5) [10]. Each data type is then modelled using PCA, which is described in [10][11]. The output of PCA is eigen vectors,  $\mathbf{v}_i$  (PCs), their weights,  $\mathbf{w}$  (PCWs), and mean of all HRIRs,  $\boldsymbol{\mu}_h$ . The resulted HRIR model is the linear combination between a few eigen vectors and their corresponding weights, added by mean of all HRIRs. Only a few eigen vectors are used in order to reduce the usage of data, but 90% variability in the whole data is maintained. Here we use 43 eigen vectors for HRIR data type 1 to type 4 and only 10 eigen vectors to HRIR data type 5.

The flowchart of data reconstruction for modelling HRIRs using PCA can be seen on Figure 2b. We modelled and reconstructed HRIR data from  $\boldsymbol{\mu}_h$ , PCs, and PCWs with distance of 75 cm, elevations from  $-40^\circ$  to  $50^\circ$  with an elevation step of  $10^\circ$ , and azimuths from  $0^\circ$  to  $355^\circ$  with  $5^\circ$  azimuth jump. The PCA variables ( $\boldsymbol{\mu}_h$ , PCs, and PCWs) are then exported to Code Composer Studio (CCS) program in the form of text files. These files are moved into the project folder for CCS. The variables are combined to produce HRIR model for left ear and HRIR model for right ear. Each HRIR model is convolved with monaural sound that comes from a particular sound position, so that it outputs a pair of 3D sound for left and right ear.

Figure 2b shows the data reconstruction process of HRIR models. It begins with reading all PCA variables in the form of text files from Matlab. A few number of  $L$  eigen vectors in matrix  $A_L$  are linearly combined with their corresponding weights,  $\hat{\mathbf{w}}$ , then its result is added by  $\boldsymbol{\mu}_h$ . In the purpose of data normalization, we search the maximum absolute sample of all resulted models,  $\hat{\mathbf{h}}$ . The HRIR models are divided by this maximum, and at the end the output is multiplied by 32,767 and rounded to the

nearest integer. This kind of data normalization is needed because we use a 16-bit fixed points DSP board. The result of data normalization is then stored and saved in the form of text files. We call these models for DSP later to compare them with the resulted PCA models from Matlab.



**Figure 2.** Research scheme in implementing HRIRs into DSP Board.

#### 4. Results and discussion

In this section, we explain the parameter used to show the performance of PCA modelling and data reconstruction in DSP board. Analysis of the performance of deployment HRIR models into DSP Board is also performed.

##### 4.1. Performance parameter

A goal of PCA is to keep as much of the variation of the original database as possible, while minimizing the number of eigen vectors and weights needed to model the original HRIRs. The percentage error between an original HRIR,  $\mathbf{h}$ , from the database and its model,  $\hat{\mathbf{h}}$ , is defined as

$$e(\theta, \phi) = 100\% \times \frac{\|\mathbf{h}(\theta, \phi) - \hat{\mathbf{h}}(\theta, \phi)\|^2}{\|\mathbf{h}(\theta, \phi)\|^2} \quad (1)$$

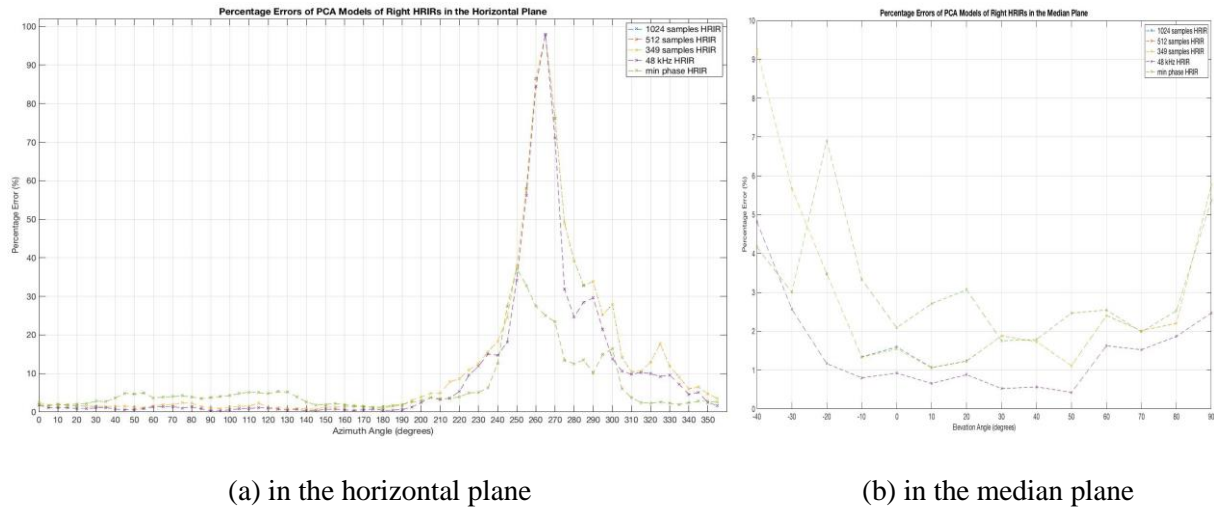
where  $\theta, \phi$  are the azimuth and elevation angle of the position of sound source, respectively.

##### 4.2. Performance of HRIRs modelling

Figure 3a shows the percentage errors of right ear HRIRs modelling along the horizontal plane. In this case, PCA is performed by using all HRIRs from distance 75 cm only. The worst error for each HRIR data type 1 to 4 happens at azimuth angle of  $265^\circ$  with error of about 98%, where the sound position is located in the opposite of right ear. This result is as expected because the amplitudes of HRIRs in the contralateral side of hearing ear are very small compared to amplitudes of HRIRs in the ipsilateral side, so that PCA has difficulties to approximate these HRIRs well.

The percentage errors of right ear HRIRs modelling along the median plane can be seen in Figure 3b. As observed, the percentage errors are relatively small and almost the same in positions along the median plane, i.e. between 0 and 10% for all data types. This happens because the HRIRs have similar shapes and relatively large amplitudes at those positions. Table 1 provides the averages of percentage errors of right ear HRIR models along the horizontal plane and the median plane, for all data types. The average of percentage errors of data type 1, type 2, and type 3 are relatively the same, i.e. about 2.90% in the

median plane and 11.72% in the horizontal plane, for right ear HRIR models. This fact has validated that the discarded samples of original HRIRs have no contribution to the performance of HRIRs modelling using PCA. Using HRIR data type 5 or the minimum phase HRIRs, it results in the smallest average error of 6.38% for positions along the horizontal plane. Although the average error along median plane for minimum phase HRIRs is not the smallest, it is almost the same with other average errors of remaining data types. This result is in line with our previous work in [12].



**Figure 3.** Percentage errors of right ear HRIR models using PCA.

**Table 1.** Average of percentage errors of right ear HRIR models

HRIR data	No. of samples used	Horizontal plane	Median plane
Type 1	1024	11.73	2.91
Type 2	512	11.72	2.90
Type 3	349	11.72	2.90
Type 4	256	9.66	1.49
Type 5	200	6.38	3.12

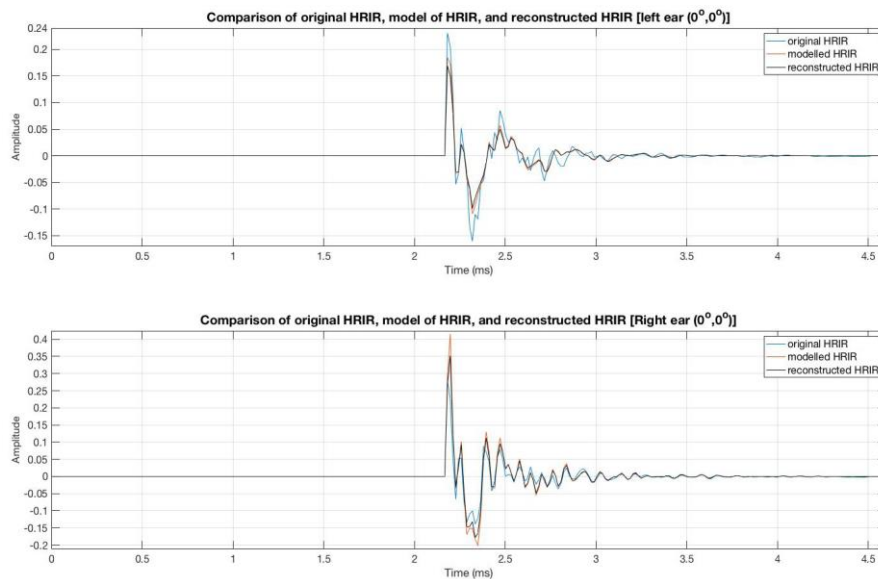
**Table 2.** Global average of percentage errors of HRIR models

Pre-processing	HRIR data	No. of samples used	No. of eigen vectors used	Avg. of % errors (3 distances)	Avg. of % errors (75 cm)
Original 65,536 Hz	type 1	1024	43	24.26	13.48
Chopped type 1	type 2	512	43	24.24	13.46
Chopped type 1	type 3	349	43	24.24	13.45
Resampled type 3 to 48 kHz	type 4	256	43	21.04	9.79
Minimum phase	type 5	200	10	5.45	5.91

Table 2 gives the information about the global average of percentage errors of HRIR models for all positions. Our first experiment used HRIR data from 3 distances (50 cm, 75 cm, and 100 cm) with targeted models using PCA for HRIRs in all positions of distance 75 cm. Second experiment using PCA was conducted by using only HRIR data from distance 75 cm. Overall, the averages of percentage errors of models that are delivered from HRIR data only with distance 75 cm are much better than those from

HRIR data with three distances. It is plausible because the HRIR data from a distance have relatively large amplitude variations compared to HRIR data from other distances.

The averages of percentage errors of models from HRIR data type 1 to type 3 are relatively the same for both cases. The models of HRIR data type 4 are slightly better than the models of HRIR data type 1 to type 3. However, the best models come from HRIR data type 5 which are minimum phase HRIRs. The average of percentage errors of the models of minimum phase HRIRs is the smallest compared to other models, i.e. 5.91% for HRIR data from 75 cm and 5.45% for HRIR data from three distances. It shows that the reconstructed minimum phase HRIR data using PCA almost resemble the original HRIR data because the resulting global average is very small. That is why, we further use the models of minimum phase HRIRs extracted from only HRIR data with 75 cm distance in the data reconstruction for implementation in the DSP board.



**Figure 4.** Reconstructed minimum phase HRIRs for position directly in front of a listener

The HRIR data type used in Figure 4 is minimum phase HRIR. As it can be seen, the blue graphs are the original HRIRs, the red graphs are the modelled HRIRs using PCA, and the black graphs are the reconstructed HRIRs in DSP board. The position of sound source for these HRIR data is directly in front of a listener, which has azimuth  $0^\circ$  and elevation  $0^\circ$  or  $(0^\circ, 0^\circ)$ . The reconstructed HRIRs are very similar visually as the modelled HRIRs that were used to make the reconstructed HRIRs. The percentage error between the reconstructed HRIR and the modelled HRIR for left ear (Figure 4 above) is 1.03%, whereas between the original HRIR and the modelled HRIR is 11.32%; and between the original HRIR and the reconstructed HRIR is 14.29%. For right ear (Figure 4 below), the percentage error between the reconstructed HRIR and the modelled HRIR is 1.21%, but that between the original HRIR and the modelled HRIR is 28.80%; and that between the original HRIR and the reconstructed HRIR is only 18.40%. This result indicates that the reconstructed HRIRs in DSP board have a very high degree of similarity to their original HRIR data.

The next test is to calculate the time required by DSP Board TMS320C5535 eZdsp to conduct the process of reconstruction of HRIR data from modelled HRIRs using PCA with C programming language in simulating 12 sound positions moving around the horizontal plane. The test is performed by debugging the C program. Debug programs that run only from the beginning and end of the HRIR data reconstruction process are completed in order to calculate the clock cycle required to process the modelling using PCA. The test gives the result of clock cycle calculation of 76,622,703 clock cycle. To change the clock cycle count into units of time, the calculation is done by multiplying the cycle count with periods in the DSP board [9]. In this research, the frequency used is 300 MHz. From the calculation, the result of HRIR data reconstruction is 255.41 ms.



## 5. Conclusion

Based on the results of this research, implementation the HRIR models using PCA in DSP board TMS320C5535 eZdsp™ can be performed by first using Matlab program in processing original HRIR data and then modelling them by PCA. We found that using the minimum phase HRIRs resulted in the best models using PCA method compared to original HRIRs, chopped original HRIRs, and resampled HRIRs, that were tested in our experiments. The global average of percentage errors of these models is only 5.91%. Using minimum phase HRIRs, the reconstructed HRIRs in DSP board have a very high degree of similarity compared to the modelled HRIRs in Matlab. The reconstruction process in DSP board to simulate 12 sound sources moving around the horizontal plane takes only 255.41 ms.

## Acknowledgments

The authors gratefully thank the Peking University and Chinese Academy of Sciences for providing the PKU-IOA HRTF Database.

## References

- [1] Freeland F P, Biscainho L W P and Diniz P S R 2002 Efficient HRTF interpolation in 3D moving sound *Proc. AES Conf.: 22nd Int. Conf. on Virtual, Synthetic, and Entertainment Audio* (Espoo, Finland, June 15-17) 000232
- [2] Kendall G S 1995 A 3-D sound primer: directional hearing and stereo reproduction *Comput. Music J.* **19**(4) 23-46
- [3] Hugeng H, Anggara J and Gunawan D 2017 Enhanced three-dimensional HRIRs interpolation for virtual auditory space *Proc. IEEE Int. Conf. on Signals and Systems 2017* (Bali, Indonesia, May 16-18) 35-39
- [4] Hugeng H, Laya F, Wahab W and Gunawan D 2015 Implementation of HRIR interpolations on DSP board TMS320C5535 eZdsp™ *Proc. The 14<sup>th</sup> Int. Conf. on QiR (Quality in Research)* (Lombok, Indonesia, August 11-12) C15.3
- [5] Pelczynski P and Bochenska M 2012 Real time implementation of dynamic 3D sound sources in a digital signal processor *Proc. 11<sup>th</sup> IFAC/IEEE Int. Conf. on Programmable Devices and Embedded Systems* vol 45(7) (Brno, Czech Republic, May 23-25) 277-282
- [6] Martens W 1987 Principal components analysis and resynthesis of spectral cues to perceived direction *Proc. The 1987 Int. Comput. Music Conf., ICMC 1987* (Champaign/Urbana, USA, August 23-26) 274-281
- [7] Sodnik J, Umek U, Susnik R, Bobojevic G and Tomazic S 2004 Representation of head related transfer functions with principal component analysis *Proc Annu. Conf. Aust. Acoust. Soc.*, (Gold Coast, Australia, November 3-5) 603-608
- [8] Qu T, Xiao Z, Gong M, Huang Y, Li X and Wu X 2009 *IEEE Transactions on Audio, Speech, and Language Processing* **17**(6) 1124-1132
- [9] Spectrum Digital, Inc. 2011 TMS320C5535 eZdsp™ Technical Reference
- [10] Hugeng H, Wahidin W and Gunawan D 2015 A new selection method of anthropometric parameters in individualizing head-related impulse responses *TELKOMNIKA* **13**(3) 1014-1020
- [11] Hugeng H, Wahab W and Gunawan D 2011 The effectiveness of chosen partial anthropometric measurements in individualizing head-related transfer functions on median plane *ITB J. ICT* **5**(1) 35-56
- [12] Hugeng H, Wahab W and Gunawan D 2010 Effective preprocessing in modeling head-related impulse responses based on principal components analysis *Signal Processing: An Int. J. (SPIJ)* **4**(4) 201-212