PRINCIPAL COMPONENT ANALYSIS FOR HEAD-RELATED TRANSFER FUNCTION LOW-ORDER FILTER DESIGN

Michał Pec, Paweł Strumiłło

Institute of Electronics, Technical University of Łódź
211/215 Wólczańska, 90-924 Łódź, Poland
michal.pec@gmail.com, pawel.strumillo@p.lodz.pl
www.eletel.p.lodz.pl

Abstract: The Principal Component Analysis (PCA) is a technique that allows for dimensionality reduction in a data set. In this communication, PCA was applied to HRTF data to extract the frequency components that are most important for precise sound localization in both azimuth and elevation. This is a new approach to use PCA for low-order HRTF modelling. In the proposed method, analysis is performed for large overlapping sectors instead of the entire HRTF set. The reported method allows for reduction of HRTF filter order without decreasing virtual sound source localization accuracy.

1 Introduction

Filtering of sounds through Head Related Transfer Functions (HRTF) is a method of obtaining spatialized sound through stereophonic headphones.

In most applications utilizing this technique, HRTFs are modelled with Finite Impulse Response (FIR) filters. Studies conducted by authors, with personalized HRTFs modelled with FIR filters of 128 coefficients, show that virtual sound sources can be perceived by humans with accuracy reaching average errors of 6.36° in azimuth and 9.47° in elevation [1]. Although, implementation of a filter of even higher order is possible, finding a way to model HRTFs with filters of lower order can be beneficial.

Methods presented in this paper allow for reduction of filter size to the 48th order without decreasing the localization resolution of sound sources.

2 Head Related Transfer Functions

Head Related Transfer Function (HRTF) is a mathematical representation of the influence of the acoustic system formed by the pinnae, the head and the human torso on the deformation of an acoustic signal’s spectrum reaching the listener’s ear from a given direction.

Using HRTF measurement results to filter an audio signal can make a listener subconsciously identify the spatial properties of the sound, resulting in the externalization of the virtual sound source to the surrounding space.

The most reliable method of acquiring individual HRTFs is by measuring the response to a known acoustic signal with microphones placed at the entrance to the each ear channel. The measurement is conducted simultaneously for both ears and for a large number of directions around the listener.

3 Principal Component Analysis of HRTF

The magnitude responses of HRTFs are fairy flat up to approximately 2kHz, but for higher frequencies the magnitudes vary significantly with peak-to-peak values reaching 40dB [3]. The shape of the magnitude spectrum depends on the sound source’s location relatively to the listener and differs considerably for different angles. The aim of this study is to find frequency components, which vary most with a change of sound source direction. This allows to discard HRTF magnitude spectra components which are listener-independent, or their influence on localization process is negligible. As a result, such preprocessed HRTFs can be modelled with a filter of a much lower order.

3.1 Principal Component Analysis Theory

Principal Component Analysis (PCA) is a standard linear data analysis technique [4]. It involves a mathematical procedure which reduces data dimensionality by performing a covariance analysis between factors. PCA is very useful, in cases for which the variables in the data set are correlated with one another. Because of this redundancy it is possible to reduce the number of variables into a smaller number of principal (uncorrelated) that will cover most of the variance in the data under analysis.

3.2 HRTF Data

HRTF data was acquired with a measurement system [1], [2] installed in an anechoic chamber. The system allows for automated HRTF measurement for full azimuth range and elevation range from -45° to 90°. Signal acquisition is performed simultaneously for both ears and HRTFs are equalized on the fly against a reference measurement to remove the effect of the loudspeakers and microphones used in the measurement. Measured HRTFs are given in the form of impulse responses (HRIRs).

3.3 PCA Reconstruction of HRTF Data

Principal Component Analysis was performed on the magnitude spectrum of HRTFs. From the whole data set for each ear subsequent subsets corresponding to particular sectors in azimuth-elevation space were chosen and PCA was carried out on them.

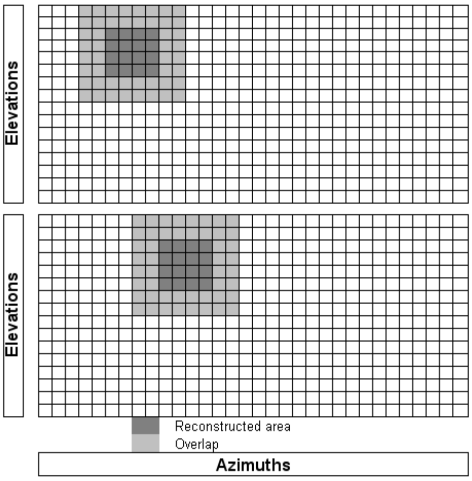

Figure 1. An idea of the overlapping sectors of input data

Because the input data covers a larger range of angles than the output data, an overlap factor was defined. The mechanism of this process is outlined in Figure 1. The main idea of such solution stems from the requirement to strengthen the differences between HRTFs for a narrow range of angles. Overlapping is used to avoid discontinuity in the reconstructed data that can be audible during 3D sound rendering.

Input data for each iteration covered full angle elevation range and 100° azimuth range. There was no need to use overlapping in elevation and the overlap in azimuth was 40°. This means, that azimuth range of the reconstructed data was 60° in each iteration. For each data set reconstruction with 6 principal components was performed.

Such preprocessed HRTFs were modelled with a 48th order FIR filter designed according to the least squares fit to the frequency response.

4 Listening Test

Perceptual verification of results was carried out by employing a subjective listening test. The test was conducted within software written in Matlab. Sounds were presented through Sennheiser HD-650 headphones. Five subjects (2 female, 3 male) volunteered in the study. HRTFs used for the test were measured personally for each participant. Each HRTF set was modelled with: a) a full FIR filter of 128 coefficients, b) with a reduced 48–coefficient filter without any preprocessing, and c) a 48-coefficient filter reconstructed with 6 Principal Components.

As a stimuli synthesized organ sound mixed with white noise (SNR=28dB) was used.

4.1 Procedure

Virtual sound sources were placed in 12 discrete locations in azimuth corresponding to full hours on a clock face and three positions in elevation for angles: φ_1=-27°, φ_2=0° and φ_3=27°. After a test sound was emitted, the test participant was to report the perceived location of the sound in terms of azimuth (clock hour) and elevation (up, down, or eye level).

4.2 Results

Since the perceived sound source’s position was reported with limited resolution, the experiment data cannot be analyzed by means of an absolute localization error. The impact of the preprocessing and reduction of the filter order was evaluated basing on the number and type of localization errors listed in Table 1.

Table 1. List and description of localization errors

	Error type	Description
Azimuth	Perfect indication	Perceived azimuth equal to the azimuth of the presented sound source
	Slight error	Perceived azimuth differs from the azimuth of the presented sound source by approx. 30°
	Serious error	Perceived azimuth differs from the azimuth of the presented sound source by more than 30° and not classified as front– back confusion
	Front–back error	Sound perceived in front of the listener instead of back and vice versa
Elevation	Perfect indication	Perceived elevation equal to the elevation of the presented sound source
	Slight error	Perceived elevation differs from the elevation of the presented sound source by “one step” (i.e. up instead of eye level , or eye level instead of down)
	Up–down error	Sound perceived as originating from below instead from above and vice versa

The trial results are shown in Figure 2 and in Table 2.

Table 2. Mean percentage share of all error types for the reference and the reconstructed HRTFs

Error type		Reference HRTF	Low-order w/o preprocessing	Low-order, 6 principal components
Azimuth	Perfect indication	44%	21%	40%
	Slight error	27%	22%	31%
	Serious error	15%	16%	12%
	Front – back error	14%	41%	17%
Elevation	Perfect indication	61%	39%	55%
	Slight error	34%	35%	37%
	Up – down error	5%	26%	8%

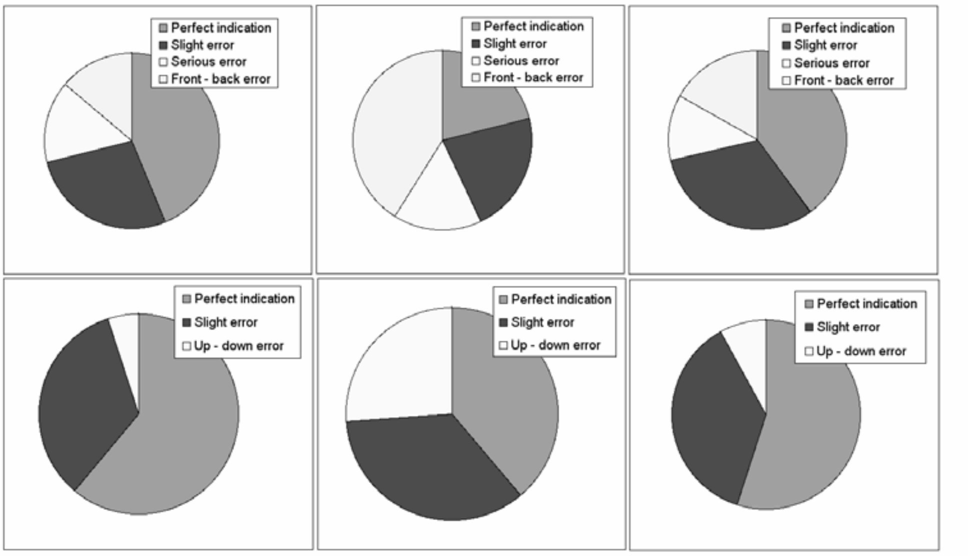

Figure 2. Share of azimuth (up) and elevation(down) errors for the reference HRTFs (left), HRTFs modelled with low-order filter w/o preprocessing (middle) and modelled with low-order filter after preprocessing (right).

4.3 Discussion

In spite of the relatively small group of participants, clear trends can be observed from the gathered data. Low-order filter models of HRTFs with no preprocessing stage significantly decrease the localization accuracy. Use of Principal Component reconstruction allows for reduction of filter order without noticeable negative consequences.

5 Conclusions and future work

In this work low-order modelling of HRTFs by means of the Principal Component Analysis was analyzed. The trials that were carried out with the preprocessed HRTFs and modelled with FIR filters of the 48th order showed that the reported method successfully reduces the computational load of HRTF filtering without significant losses in sound source localization accuracy.
In the near future, implementation of recursive filters for modeling HRTFs is considered.

Acknowledgements

This work has been supported by the Ministry of Science and Higher Education of Poland research grant no. N N516 370836 in years 2009-2010.

Referrences

[1] M. Pec, M. Bujacz, P.Strumiłło, A. Materka, “Individual HRTF Measurements for Accurate Obstacle Sonification in an Electronic Travel Aid for The Blind”, International Conference on Signals and Electronic Systems, Kraków 2008.

[2] M.Pec, M. Bujacz, P. Strumiłło, “Head Related Transfer Function measurement and processing for the purpose of creating a spatial sound environment”, Signal Processing Symposium, Jachranka, Poland, 2007

[3] H. Møller, M. F. Sorensen, D. Hammershøi, C. B. Jensen, “Head-Related Transfer Functions of human subjects”, J. of the AES, 43, 300–321, 1995

[4] J. Shlens, “A Tutorial on Principal Component Analysis”, 2009 http://www.snl.salk.edu/~shlens/pub/notes/pca.pdf