Approximation of Head Related Transfer Functions

Jaka Sodnik, Rudolf Susnik and Saso Tomazic

University of Ljubljana, Faculty of Electrical Engineering
Trzaska 25, Ljubljana, Slovenia
jaka.sodnik@fe.uni-lj.si

Abstract

The set of Head Related Transfer Functions (HRTF) enables the generation of spatial sound [1]. Two FIR filters with 512 coefficients are required for any spatial position. This paper suggests a method for the description of HRTFs which reduces the total number of necessary coefficients for the factor 4.5 – 14 and enables the construction of new filters for nonmeasured positions. The method exploits the high cross correlation among HRTFs and high significance of Inter-Aural Time Difference (ITD) factor for space sound generation. The reduction is performed in four steps: interpolation, HRTF alignment, decimation and Principal Component Analysis (PCA).

1 Introduction

HRTF filters are empirically acquired FIR filter sets that describe the changes of the sound as it travels from its source towards the human eardrum. The use of HRTFs enables listening of the virtual sound source at any position in space through the headphones. The two main parameters for establishing the azimuth of the sound source location are ITD and Inter-Aural Level Difference (ILD). They are included in HRTFs. ITD defines the time difference between the arrivals of sound signals to the left and right eardrum. ILD defines the level difference between these two signals.

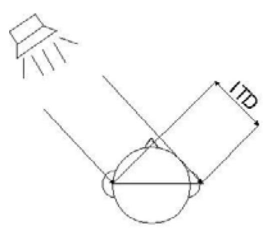

Figure 1: ITD efect

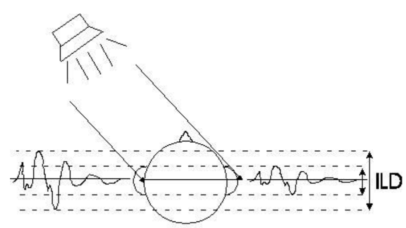

Figure 2: ILD efect

2 Methods

The main objective of our research is the ortogonalization of the 72 measured HTRFs in the horizontal plain, each containing 512 samples. The first part includes the interpolation of functions by factor 10 to obtain 5120 samples for each function. The calculation of cross correlation between separate functions defines their similarity. The maximal cross correlation enables the elimination of ITD parameter from all functions. Without the appropriate delay, HRTFs for different values of azimuth differ only in amplitude gain. With decimation, the length of filters can be reduced back to 512 samples. At this point, the Principal Component Analysis (PCA) is performed.

2.1 PCA

PCA is a statistical method based on Gaussian distribution of random variable [2]. The starting point of the method is autocorrelation matrix (2), composed of input vectors (1) (HRTFs for different values of azimuth). These are called the training set.

$$p_i = \begin{bmatrix} p_{i,1} \\ p_{i,2} \\ \\ p_{i,n} \end{bmatrix} \quad (1) \qquad \Phi_{pp} = \frac{1}{N}\sum_{i=1}^{N} p_i \cdot p_i^T \quad (2)$$

In order to find the directions of maximum variance, we calculate the eigenvalues of autocorrelation matrix and corresponding eigenvectors. To extract most of the variation, we seek for the biggest eigenvalues and prepare a linear transformation matrix consisting of the corresponding eigenvectors (3). Its elements are orthonormal and written with the minimum of required parameters. The projection of training data into eigenspace is necessary to obtain the matrix of appropriate weights for all azimuth angles (4).

$$V = [v_1, v_2, \ldots v_J] \quad (3) \qquad w_i = V^T \cdot p_i. \quad (4)$$

All the information about HRTF filters is now represented with a small number of eigenvectors and a set of weights for all HRTFs. The main objective of the method is the possibility of generation of filters for new spatial positions, using eigenvectors and the calculated set of weights.

3 Results

From autocorrelation matrix Φ (512 x 512 samples) 14 orthogonal eigenvectors can be calculated. Each HRTF filter is defined with the set of eigenvectors and the appropriate weight. The number of coefficients required to describe all eigenvectors and all weighs is 4.5 times smaller than the number of coefficients in the original HRTF set. Further analysis and subjective criteria show that only the first 120 coefficients from each filter are needed to generate quality space sound. If those irrelevant samples are cut off before the PCA method is performed, the number of coefficients reduces 14 times. The set of weights for different azimuth angles enables us to manipulate the reconstructed functions. The linear interpolation of two proximate functions (5) gives a satisfactory approximation for intermediate azimuth angles.

$$w_{new} = \frac{w_1 + w_2}{2} \quad (5)$$

The quality of the new approximated HRTF filters was verified by using the subjective criteria of several testees who tried to determine the spactial position of the interpolated sound sources.

References

[1] Bill Gardner and Keith Martin, **HRTF Measurements of a KEMAR Dummy-Head Microphone**, MIT Media Lab Perceptual Computing – Technical Report #280, May, 1994
[2] Stanford University, Image Systems Engineering Program, **PCA Principles**, 1996
Url: http://ise0.stanford.edu/class/ee368a_proj00/project2/node7.html