Individualization of head-related transfer functions

Florian Klein[1] and Stephan Werner[1]
[1] TU Ilmenau, Institut für Medientechnik, Fachgebiet Elektronische Medientechnik, 98693 Ilmenau,
E-Mail: florian.klein@tu-ilmenau.de, stephan.werner@tu-ilmenau.de

Introduction

Head-related transfer functions (HRTF) are of major importance for binaural hearing tasks. Non-individual HRTFs, such as those measured with artificial heads, show weaknesses when it comes to the precise localization or externalization of sound sources. Applications like virtual auditory displays and binaural hearing aids can profit from HRTF individualizations.

An algorithm which tries to estimate HRTFs from measured anthropometric data is presented. Based on the *CIPIC* database, regression analysis can be applied to predict HRTFs for an arbitrary direction. Crucial points are the choice of meaningful anthropometric parameters and the evaluation of the prediction results.

Individualization methods for HRTFs

To obtain individualized HRTFs, time-consuming and financially expensive measurements are commonly performed [6]. Because of their drawbacks, several alternative methods have been developed:

- Simulation of HRTFs based on computer models (for example see [2])
- Subjective test to obtain the best-suited HRTF for an individual [5]
- Estimation of HRTFs based on a reduced set of anthropometric data [3] and [7]

The simulation of HRTFs may yield good results. An accurate model of the head is required to achieve an acceptable accuracy in high frequencies. Again, special devices are necessary for creating such a model. Hearing tests for finding the best-suited HRTF are able to reduce the errors introduced by non-individualized HRTFs. Fast and effective approaches are available, but the success of these methods is always limited by the "best" available HRTF. Estimation methods can be further divided into sub-categories, e.g. geometrically, structurally or empirically motivated models (see overview in [6]).

The *CIPIC* database

The implemented algorithm is empirically motivated and uses the publically available *CIPIC* database [1] for ground truth. This database contains HRTF measurements for 45 individuals. For 37 of these 45 subjects, all 37 anthropometric parameters are available. 17 parameters describe the head and torso and another 10 parameters are available for each ear of the subjects. Figure 1 shows the whole set of anthropometric measures.

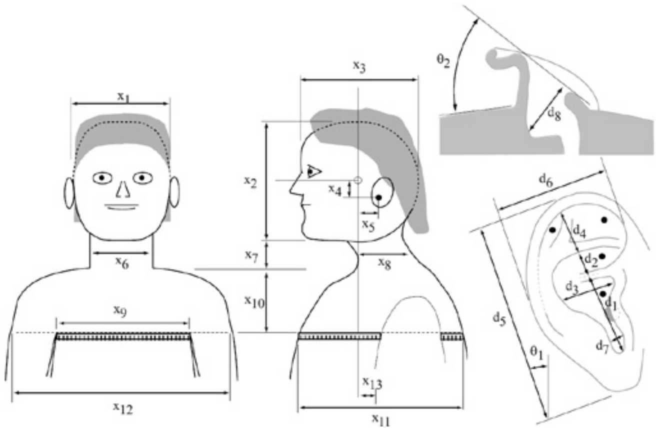

Figure 1: The anthropometric measures used in *CIPIC* [1]

Implementation

The following approach is informed by the works of Rodríguez et al. [3] and Hugeng et al. [7]. A brief overview is given in Figure 2.

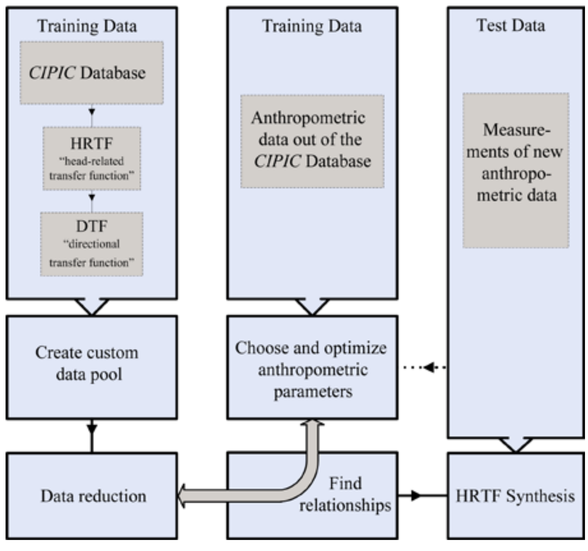

Figure 2: Overview of the individualization process (also see [3])

Create custom data pool: After preprocessing the HRTFs, a pool of HRTFs needs to be created for the further steps. The goal is to reduce the amount of data or to focus on specific features of the data. In this work, the focus is on the HRTFs in the horizontal plane. Therefore, only the HRTFs from this plane are used.

Data reduction: In the next step, principal component analysis is performed to decorrelate the data and to reduce the dimensions of data set. Based on the variance of the given data, an estimation about the information content of each principal component can be made. The first 20 principal components contain over 98% of the information.

Choose and optimize anthropometric parameters: The anthropometric data also needs to be processed. Not all of the parameters shown in Figure 1 have a significant influence on the acoustical transmission. Therefore, correlation analysis between these parameters and the principal component weights is performed. The best predictor for each principal component weight is found this way. The average correlation for a predictor set is between 0.27 and 0.3, which is a weak starting point for the regression step. Investigations have shown that the set of the most important parameters depends on the direction. The parameter set is built upon these findings and on further improvements proposed by Rodríguez et al. [3].

Find relationships: After these preparations, regression analysis is applied. For this task, multiple linear regression (MLR) is commonly used. Previous work suggested that the use of non-linear regression could improve the prediction results. Because of this, support vector regression (SVR) with a radial basis function kernel was also used for comparison.

HRTF synthesis: With these found relationships, a HRTF prediction can be done. Figure 3 shows some exemplary prediction results. For each calculation, the data of the test subject was removed from the data set and the remaining subjects where used for the training.

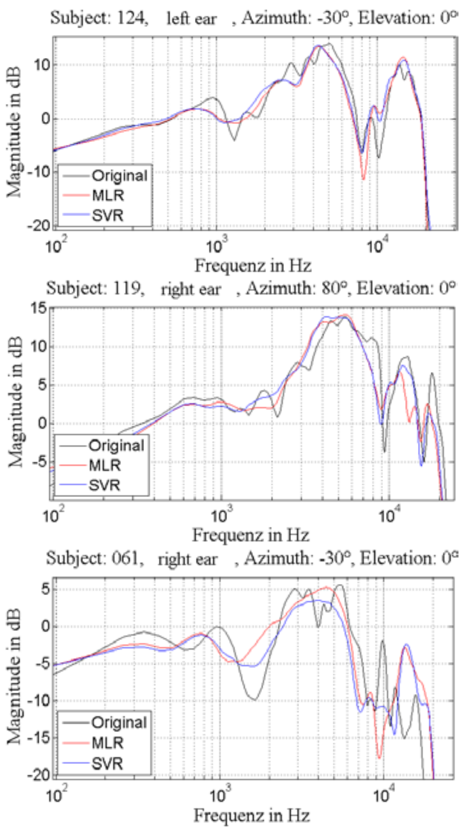

Figure 3: Exemplary prediction results for different subjects, ears and directions. From top to bottom: good results to bad results.

The predictions results over all subjects show a significant fluctuation in accuracy.

Evaluation

Additionally, the HRTFs of six persons were measured with probe microphones and an individual headphone equalization [4] was done for each of them. Anthropometric data was measured according to the *CIPIC* specifications. A MUSHRA-based listening test was used to compare artificial head HRTFs with the predicted HRTFs based on localization and coloration differences. The individually measured HRTFs acted as a reference. For 1 out of 6 subjects, an improvement over the artificial head HRTFs could be measured in both cases. A possible cause of these results could be the difference between the prediction accuracy of the left and right ears of the subjects. Furthermore, an ABX-based listening test could produce more comprehensive results.

Conclusion

The results lead to the conclusion that the given anthropometric data may not be sufficient to fully describe HRTF magnitude spectra. Both regression algorithms deliver comparable results. Furthermore, it is difficult to judge the prediction based on the difference of the magnitude alone. A comparison based on the difference of perceptually relevant features would be more meaningful. Future work could investigate perceptually relevant features more closely. Based on the results, HRTF estimation methods could be improved.

Literature

[1] V. R. Algazi, R. O. Duda and D. M. Thompson. The CIPIC HRTF Database. In IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 2001, p. 99-102.

[2] Y. Kahana. Numerical modeling of the head-related transfer function. PhD Thesis. University of Southampton. 2000.

[3] S. G. Rodríguez and M. A. Ramírez. HRTF Individualization by Solving the Least Squares Problem. In Audio Engineering Society 118th Convention, 2005.

[4] Z. Schärer. Kompensation von Frequenzgängen im Kontext der Binauraltechnik. Magister Thesis. TU Berlin. 2008.

[5] B. U. Seeber and H. Fastl. Subjective Selection of Non-Individual Head-Related Transfer Functions. In Proceedings of the 2003 International Conference on Auditory Display, Boston, 2003, p. 259-262.

[6] S. Xu, Z. Li and G. Salvendy. Individualization of Head-Related Transfer Function for Three-Dimensional Virtual Auditory Display: A Review. Virtual Reality, Lecture Notes in Computer Science 2007 Vol. 4563, 2007, p. 397-407.

[7] Hugeng, W. Wahab and D. Gunawan. Improved Method for Indivdualization of Head-Related Transfer Functions on Horizontal Planes Using Reduced Number of Anthropometric Measurements. Journal of Telecommunications 2(2), May 2010, p. 31-41.