

HRTF Customization using the LDV HRTF-database

**Martin Rothbucher, Alexander Kuhn
and Klaus Diepold**
Institute for Data Processing,
Technische Universität München
Arcisstr. 21, 80333 München, Germany
Email: {martin.rothbucher,
alexander.kuhn, kldi}@tum.de



HRTF Customization using the LDV HRTF-database

Martin Rothbucher, Alexander Kuhn and Klaus Diepold
Institute for Data Processing, Technische Universität München
Arcisstr. 21, 80333 München, Germany
Email: {martin.rothbucher, alexander.kuhn, kldi}@tum.de

01.03.2014



Martin Rothbucher, Alexander Kuhn and Klaus Diepold
Institute for Data Processing, Technische Universität München
Arcisstr. 21, 80333 München, Germany
Email: {martin.rothbucher, alexander.kuhn, kldi}@tum.de. *HRTF Customization using the LDV HRTF-database.*
Technische Universität München, Munich, Germany, 2014.

Abstract

In recent years, the number of researchers dealing with HRTFs and 3D sound has been growing. Since the HRTFs depend on the individual geometry of a listener, every individual needs his or her own set of HRTFs in order to perceive the best possible 3D sound. In this paper, we compare regression based approaches to customize a set of HRTFs with respect to the users. The regression algorithms are initially conducted with the well-known CIPIC HRTF data base in order to find parameters for the respective approaches. Then, the customization methods are conducted for the institute's HRIR-database (LDV-database).

Contents

1	Introduction	5
2	HRTF Customization Methods	6
2.1	Numerical approximation - HAT Model	6
2.1.1	Numerical Approximation - Fine Mesh	6
2.2	Tournament Style	6
2.3	HRTF Generation using Artificial Neural Networks	6
3	Customization by Regression	8
4	Experimental Set Up	11
4.1	Used Data	11
5	Results	13
6	Discussion	15
7	Figures and Tables (Appendix)	16

1 Introduction

Spatial 3D sound can be computed by using Head Related Transfer Functions (HRTFs). The HRTFs describe how sound from any given source is filtered by a person's anthropometry due to reflexions and diffractions of the human body. Since barely any two persons look exactly the same, barely any two sets of HRTFs are identical either. In literature, it is suggested to utilize an individual set of HRTFs to obtain highly immersive 3D sound synreport. However, individually measuring the personal HRTFs requires complex, expensive and time consuming methods seeing as they have to be acquired in an anechoic room with very accurate equipment.

To avoid these problems, researchers have developed several alternative approaches.

In this report, HRTF customization methods are overviewed and customization based on regression is described in detail.

2 HRTF Customization Methods

This section describes several possible solutions for the problem of HRTF Customization.

2.1 Numerical approximation - HAT Model

The HAT model is based on a numerical approximation of the HRTFs. It uses a so-called snowman model to represent a human instead of using the correct anthropometric features. It consists of two balls: a large one representing the body and a smaller one representing the head with some space between them representing the neck. The size of these balls and the distance between them has to be determined individually based on measured anthropometric parameters.

The HRTFs themselves are then calculated using the fast multipole method (FMM) as demonstrated by [7]. Several authors have found this model to be very effective. According to [6], the HAT model performs better in terms of localization than results based on customization by regression. However, it only works well for signals limited to 3 kHz as shown by [1]. Please refer to [1] for more information on the HAT model.

2.1.1 Numerical Approximation - Fine Mesh

Instead of using the HAT model, there is also the possibility of using the boundary element method (BEM) together with a fine mesh grid simulating a human body. Otani et al. propose a method that could run on a regular consumer computer. However, it does require a fine mesh 3D model of a human head. So even if this method makes HRTF customization less complex in terms of computational costs, the prerequisites are hard to come by. Refer to Otani et al. [17] for more information on customization by BEM.

2.2 Tournament Style

Some sources like [3] and [23] suggest choosing fitting HRTFs from existing data bases instead of creating new ones. The main idea of this approach is for the subject to listen to several sounds created with different HRIRs. Based on a tournament style comparison, the ones with better localization move on. Whenever the subject is not able to correctly localize a sound based on a HRIR, this HRIR is eliminated. In the end, only the most suited set of HRIRs remains. This approach, while not needing a lot of equipment still needs some time and bears the risk of not being able to find an ideal set of HRIRs, even if it is the best fitting one from a data base.

2.3 HRTF Generation using Artificial Neural Networks

Another way of trying to gain personal HRTFs is by using artificial neural networks. In this case, an artificial neural network (ANN) is trained, so it basically is supervised learning. Due to the potential

2.3 HRTF Generation using Artificial Neural Networks

of ANNs, this method sounds promising. However, ANNs are rather complex which is why the run time can be very long. This approach was not further pursued due to the high complexity of the approach. For further information please refer to [9] and [8].

3 Customization by Regression

Apart from the above methods, HRTF Customization can also be achieved by using regression as is shown for example by [10]. The idea is to use a regression model to determine the influence of the anthropometric data of a person on the individual HRTFs. Using this influence, the new, customized HRTFs can be easily calculated. The HRTFs of a set of subjects can be gathered in a tensor $\mathcal{H} \in \mathbb{R}^{N_d \times N_f \times N_p}$, where N_d is the number of directions, N_f is the number of frequency samples and N_p is the number of persons in the data set. In order to center the data and remove any non-directional information, the directional transfer functions (DTF), computed by

$$\mathcal{D}(i, j, k) = 20 \log_{10} |\mathcal{H}(i, j, k)| - \frac{1}{N_p} \sum_{k=1}^{N_p} 20 \log_{10} |\mathcal{H}(i, j, k)|,$$

can be used.

There are several different regression methods. A selection of the most promising ones is described in the following.

PCA

The PCA is one of the most common regression methods and is widely used.

The PCA uses the eigenvectors of the covariance matrix as regression weights. For more information on PCA, refer to [19], [20], [16], [12], [26], or [5]. The way the PCA was applied in this report is described in the following.

Step 1: Calculate the DTFs

Step 2: Calculate the eigenvectors W from the covariance matrix $\mathcal{C} = \mathcal{D}(i, :, :)^T \mathcal{D}(i, :, :)$

Step 3: Use eigenvectors W to solve $B = W \tilde{A} (\tilde{A}^T \tilde{A})^{-1}$ **Step 4:** Multiply the new anthropometric data \tilde{a}_{new} on B to obtain $w_{new} = B \tilde{a}_{new}$

Step 5: Calculate $\hat{\mathcal{D}} = \mathcal{D} W$ where \mathcal{D} is the DTF of the LDV HRIRs.

Step 6: Calculate $D_{new} = \hat{\mathcal{D}} w_{new}^T$

Step 7: Reverse the DTF by adding the logarithmic mean:

$$H_{new}(i, j, k) = 10^{\frac{1}{20} (D_{new}(i, j, k) + \frac{1}{N_p} \sum_{k=1}^{N_p} 20 \log_{10} |\mathcal{H}(i, j, k)|)}$$

Step 8: Calculate ifft and add the initial delay to be able to listen to results In this paper only the 10 largest principal components were used as described by [16]. Further explanation on the PCA and the use of principal components can be found in [25].

2DPCA

As it was shown in [16], the 2DPCA should deliver better results than the regular PCA. A big advantage of the 2DPCA is that the whole tensor is used in the eigenvalue decomposition whereas the PCA can only be applied to

one matrix at a time.

The scatter matrix S_p is calculated by

$$S_p = \frac{1}{N_d} \sum_{i=1}^{N_d} \mathcal{D}(i, :, :)^T \mathcal{D}(i, :, :), \in \mathbb{R}^{N_p \times N_p},$$

as described by [13] with N_d being the number of directions, in case of the LDV data this means 2160, and N_p being the number of subjects used for training. For more information on 2DPCA, refer to [13]. The 2DPCA can be formulated as follows:

Step 1: Calculate the DTFs)

Step 2: Calculate the eigenvectors W from the Scatter matrix **Step 3:** Use eigenvectors W to solve $B = W\tilde{A}(\tilde{A}^T\tilde{A})^{-1}$

Step 4: Multiply the new anthropometric data \tilde{a}_{new} on B to obtain $w_{new} = B\tilde{a}_{new}$

Step 5: Calculate $\hat{D} = \mathcal{D}W$ where \mathcal{D} is the DTF of the LDV HRIRs.

Step 6: Calculate $D_{new} = \hat{D}w_{new}^T$

Step 7: Reverse the DTF by adding the logarithmic mean:

$$H_{new}(i, j, k) = 10^{\frac{1}{20}(D_{new}(i,j,k) + \frac{1}{N_p} \sum_{k=1}^{N_p} 20 \log_{10} |\mathcal{H}(i,j,k)|)}$$

Step 8: Calculate ifft and add the initial delay to be able to listen to results

Tensor SVD

The SVD - much like the PCA - decomposes a signal into its main components.

The main advantage of performing calculations using tensors is that the three dimensional data from the data base does not need to be split up into two dimensional matrices only to be put back together again later on.

The n-way Toolbox [4] provides algorithms for that task. *tucker()* that calculates the desired singular values of a tensor was used in this paper. Apart from it using tensors instead of matrices, its main functionality is still a singular value decomposition. For more information on tensors and the Tensor SVD refer to [4], [24], [2] and [21]. The following algorithm was adapted from [16].

Step 1: Solve the minimization problem using the DTFs:

$$\min_{\hat{D} \in \mathbb{R}^{N_d \times N_f \times N_p}} \|\mathcal{D} - \hat{D}\|_F$$

Step 2: Decompose the rank- (r_d, r_f, r_p) tensor \hat{D} as a trilinear multiplication of a rank- (r_d, r_f, r_p) core tensor $\mathcal{C} \in \mathbb{R}^{r_d \times r_f \times r_p}$ with three full-rank matrices $X = (x_{ij}) \in \mathbb{R}^{N_d \times r_d}$, $Y = (y_{ij}) \in \mathbb{R}^{N_f \times r_f}$ and $W = (w_{ij}) \in \mathbb{R}^{N_p \times r_p}$ which is defined by $\hat{D} = (X, Y, W) \cdot \mathcal{C}$

Step 3: Use eigenvectors W to solve $B = W\tilde{A}(\tilde{A}^T\tilde{A})^{-1}$

Step 4: Multiply the new anthropometric data \tilde{a}_{new} on B to obtain $w_{new} = B\tilde{a}_{new}$

Step 5: Calculate $D_{new} = (X, Y, w_{new}^T) \cdot \mathcal{C} \in \mathbb{R}^{N_d \times N_f}$

Step 6: Reverse the DTF by adding the logarithmic mean:

$$H_{new}(i, j, k) = 10^{\frac{1}{20}(D_{new}(i,j,k) + \frac{1}{N_p} \sum_{k=1}^{N_p} 20 \log_{10} |\mathcal{H}(i,j,k)|)}$$

Step 7: Calculate ifft and add the initial delay to be able to listen to results

3 Customization by Regression

GLRAM

Short for Generalized Low Rank Approximation. It uses three lower rank matrices $X \in \mathbb{R}^{N_d \times r_d}$, $M \in \mathbb{R}^{r_d \times r_p}$ and $W \in \mathbb{R}^{N_p \times r_p}$ to approximate a matrix of higher rank. The left matrix X is then used to solve the equation $w_k = B\tilde{a}_k + \epsilon$ where in this case $w_k = R \in \mathbb{R}^{N_p \times r_p}$. After multiplication with new anthropometric data, M and W are multiplied to the result from the right side in order to reconstruct the new personalized HRIR.

For more information on GLRAM refer to [14].

The following algorithm was adapted from [16]:

Step 1: Solve the optimization problem using the DTFs:

$$\min_{X, W, \{M_i\}_{i=1}^{N_p}} \sum_{i=1}^{N_p} \| (\mathcal{D}(:, :, i) - XM_i W^T) \|_F$$

Step 2: Construct a 3D array $\mathcal{M} \in \mathbb{R}^{r_d \times N_f \times r_p}$ with $\mathcal{M}(:, i, :) = M_i$ for $i = 1, \dots, N_f$

Step 3: Use eigenvectors W to solve $B = W\tilde{A}(\tilde{A}^T \tilde{A})^{-1}$

Step 4: Multiply the new anthropometric data \tilde{a}_{new} on B to obtain $w_{new} = B\tilde{a}_{new}$

Step 5: Calculate new direct transfer functions $D_{new}(:, i, :) = X\mathcal{M}(:, i, :)w_{new}$

Step 6: Reverse the DTF by adding the logarithmic mean:

$$H_{new}(i, j, k) = 10^{\frac{1}{20}(D_{new}(i, j, k) + \frac{1}{N_p} \sum_{k=1}^{N_p} 20 \log_{10} |\mathcal{H}(i, j, k)|)}$$

Step 7: Calculate ifft and add the initial delay to be able to listen to results

PLSR

The PCA bears the problem that even if the found PCs describe the original data very well, it does not necessarily also provide a good prediction for new data. The PLSR on the other hand uses the new anthropometric data and tries to find the influence of the subject's anthropometric data on the HRTF data base. The PLSR was implemented using the Matlab Help while referring to [22], [18] and [11].

Step 1: Calculate the DTFs \mathcal{D}_X and load the anthropometric data A

Step 2: Decompose $A = TP^T$ and $\mathcal{D}(i, :, :) = UC^T$ with T, U being the feautres and P, C being the weights.

Step 3: Use $\mathcal{D}(i, :, :) = TC^T$ with $T = AW^T$

Step 4: Calculate the regression coefficients $B = W^T C^T$

Step 5: Calculate the new DTFs using the new subject's anthropometric data: $D_{new}(i, :) = B\tilde{a}_{new}$

Step 6: Reverse the DTF by adding the logarithmic mean:

$$H_{new}(i, j, k) = 10^{\frac{1}{20}(D_{new}(i, j, k) + \frac{1}{N_p} \sum_{k=1}^{N_p} 20 \log_{10} |\mathcal{H}(i, j, k)|)}$$

Step 7: Calculate ifft and add the initial delay to be able to listen to results

For more information on PLSR the reader is referred to [18], [11] or [22].

4 Experimental Set Up

The proposed algorithms are subjected to a mathematical evaluation based on cross validation. The used error measurement is the spectral distortion.

The cross validation has the advantage of using as much data as possible as well as being very common.

4.1 Used Data

Two data bases were used in the course of this paper. On the one hand, the well known and well established CIPIC data base - on the other hand, the newly created LDV data base. The CIPIC data base was recorded with a sampling frequency of 44100 Hz, the HRIRs are 200 samples long and there are a total of 1250 directions for each subject. It comprises 35 human HRIRs and two KEMAR HRIRs.

The LDV data base was recorded with a sampling frequency of 48000Hz, the HRIRs are 1024 samples long. It was measured at 6 different elevations with 360 degrees of azimuth per elevation totaling at 2160 directions per subject. It consists of the HRIRs of 35 people, including both male and female subjects. To establish a more direct comparison between the two data bases, the KEMARs were excluded from the CIPIC data base, so it only contained human HRIRs.

Results of the CIPIC data base were almost always slightly worse for the right ear which is based on the fact that only one ear was measured for the anthropometric data. To avoid this problem in future calculations, the LDV data base includes the sizes of both ears.

Please refer to [Martin Stimpfl] for more information.

Cross Validation

Since Cross Validation is a very common way of evaluation, it served as a ground truth in this paper. It holds the advantage of including as many subjects into the training data as possible. HRTF, which can then be compared to the measured one from the data base. Since this paper is supposed to further investigate the findings of [16], there is also a set of results, in which 30 people out of the 35 of each data base are used for training.

Spectral Distortion

In order to evaluate the results, spectral distortion was used. It is calculated with

$$SD = \sqrt{\frac{1}{N_f} \sum_{i=1}^{N_f} \left(20 \log_{10} \frac{|\mathcal{H}_i|}{|\mathcal{H}_{new_i}|} \right)^2}, \quad (4.1)$$

4 Experimental Set Up

where $\mathcal{H}_i \in \mathbb{R}^{N_d \times N_f \times N_p}$ is an HRTF at its i -th frequency from a data base and \mathcal{H}_{new_i} is a HRTF calculated with a regression algorithm at its i -th frequency. N_f denotes the number of frequencies. This method was already used by [10] and [16].

5 Results

All of the beforehand mentioned algorithms can be varied in numerous ways. The set of best parameters was found in [15] through extensive tests. For further information about what parameters can be altered and how they should be selected, the reader is referred to [15].

All the plots and tables are the results of cross validation. Comparing plots of the GLRAM's results can be found in the following. The remaining plots and tables can be found in the appendix.

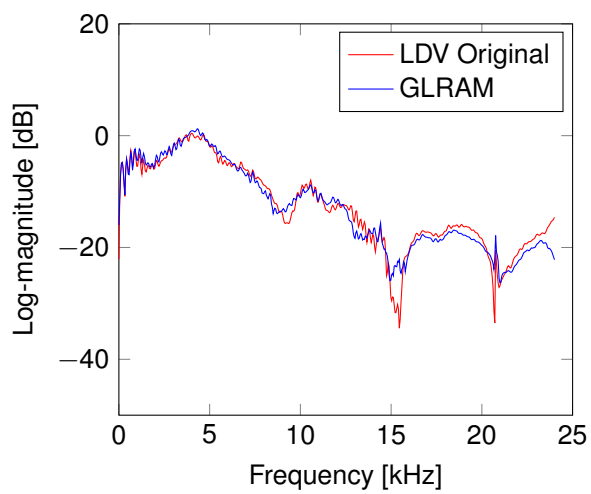


Figure 5.1: XValid, GLRAM, LDV, Subject 20, Azimuth 180°, Elevation 0°

5 Results

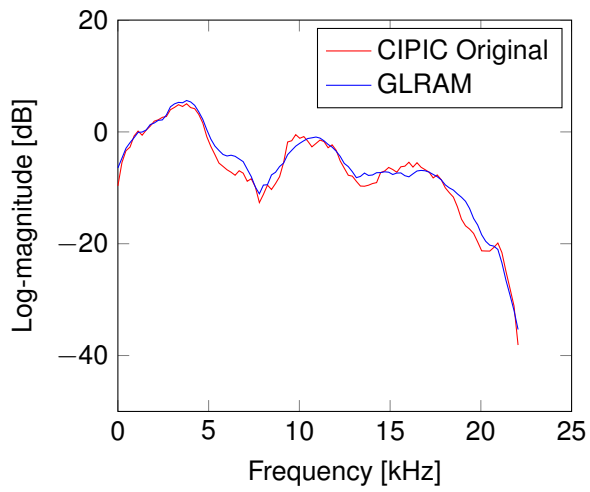


Figure 5.2: XValid, GLRAM, CIPIC, Subject 30, Azimuth -10° , Elevation 208.125°

With the GLRAM being the best approach, the subject and angle of the GLRAM's best performance were selected for the plots. These parameters are: LDV data base:

Left: elevation 0° , azimuth 180° , Subject 20 — 540

Right: elevation 40° , azimuth 307° , Subject 8 — 2107

CIPIC data base:

Left: elevation 208.125 , azimuth -10 , Subject 30 — 546

Right: elevation -16.875 , azimuth 65 , Subject 30 — 1156 (without Kemars)

6 Discussion

In this paper, five regression methods were presented. They were evaluated using spectral distortion which describes the difference between the regression HRTF and the measured one. Since the SD values were fairly small, customization by regression seems to provide a good alternative to the very time consuming approach of measuring HRIRs individually.

The very common PCA provided the highest SD values, whereas the GLRAM provided the best ones.

7 Figures and Tables (Appendix)

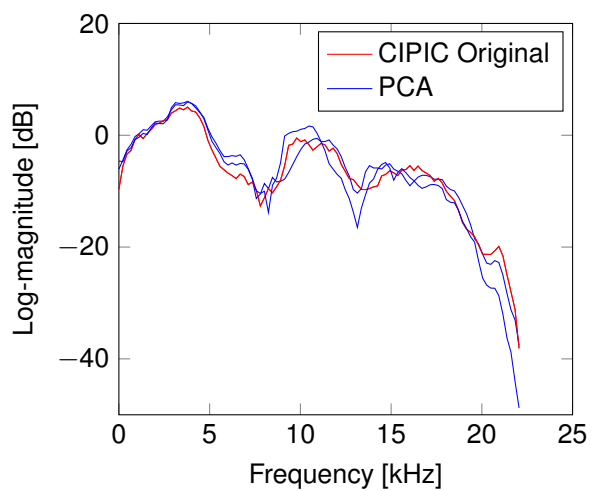


Figure 7.1: XValid, PCA, CIPIC, Subject 30, Azimuth -10° , Elevation 208.125°

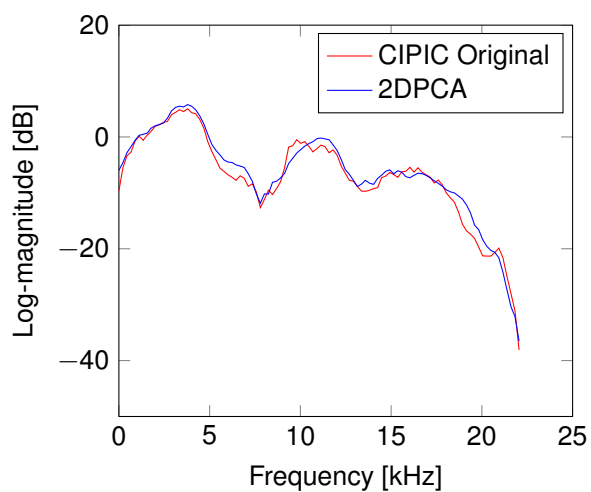


Figure 7.2: XValid, 2DPCA, CIPIC, Subject 30, Azimuth -10° , Elevation 208.125°

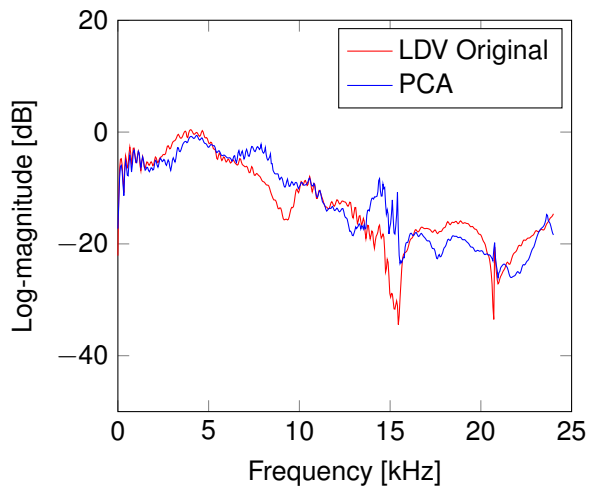


Figure 7.3: XValid, PCA, LDV, Subject 20, Azimuth 180°, Elevation 0°

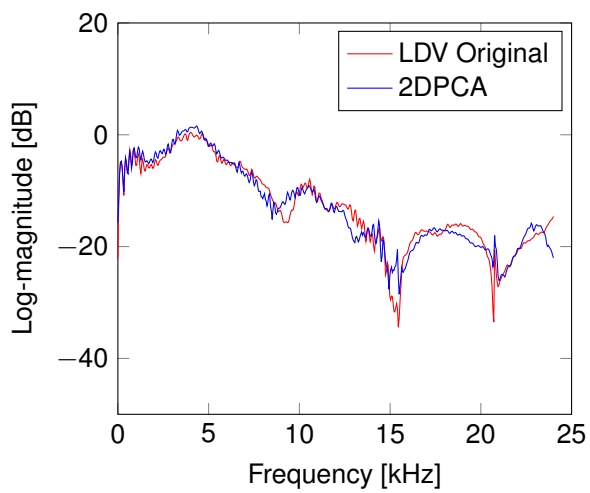


Figure 7.4: XValid, 2DPCA, LDV, Subject 20, Azimuth 180°, Elevation 0°

7 Figures and Tables (Appendix)

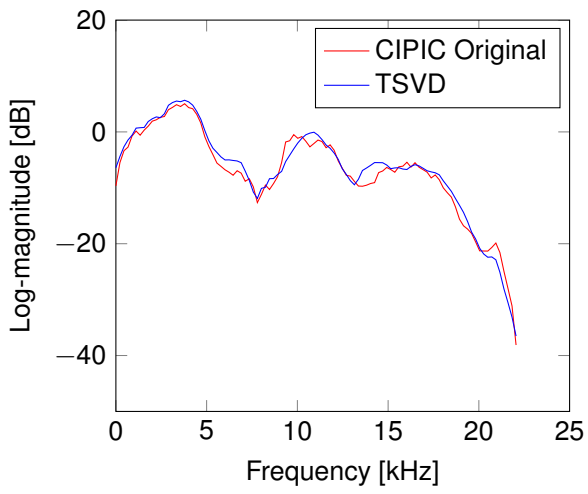


Figure 7.5: XValid, TSVD, CIPIC, Subject 30, Azimuth -10° , Elevation 208.125°

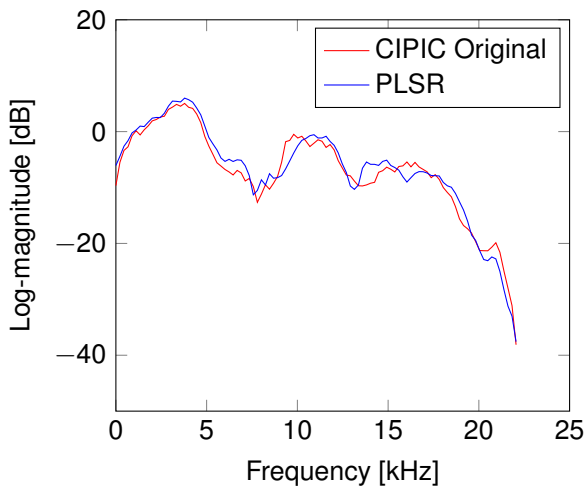


Figure 7.6: XValid, PLSR, CIPIC, Subject 30, Azimuth -10° , Elevation 208.125°

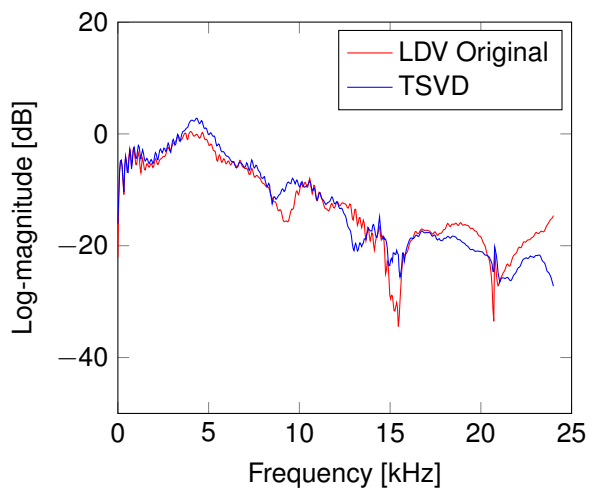


Figure 7.7: XValid, TSVD, LDV, Subject 20, Azimuth 180°, Elevation 0°

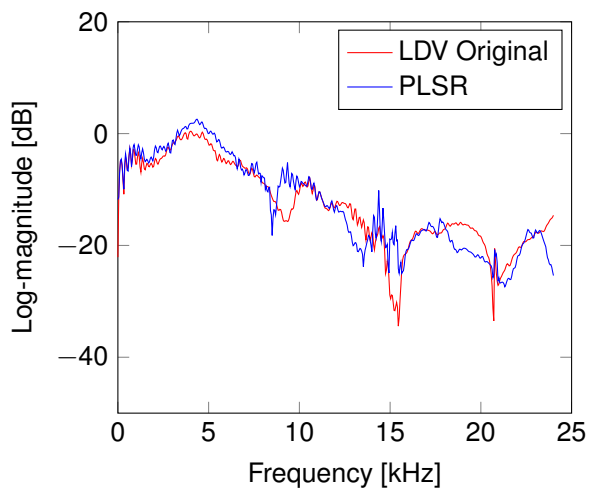


Figure 7.8: XValid, PLSR, LDV, Subject 20, Azimuth 180°, Elevation 0°

7 Figures and Tables (Appendix)

Subject	PCA	2DPCA	GLRAM	TSVD	PLSR
1	6.86	5.68	5.34	5.72	6.27
2	6.81	5.50	5.37	5.96	6.36
3	7.95	6.77	6.54	6.70	7.01
4	6.88	5.66	5.55	5.86	6.26
5	6.96	5.48	5.37	5.62	5.86
6	7.29	6.17	5.95	6.40	6.85
7	7.81	6.54	6.22	6.57	6.96
8	6.60	5.61	5.39	5.70	5.96
9	7.00	5.93	5.85	5.85	5.92
10	6.89	5.60	5.52	5.80	6.14
11	8.74	6.72	6.49	7.08	7.73
12	6.77	5.65	5.56	5.83	6.04
13	6.81	5.36	5.31	5.36	5.47
14	7.00	5.94	5.89	6.23	6.53
15	7.27	6.13	6.09	6.18	6.29
16	7.79	6.74	6.66	6.70	6.85
17	6.65	5.75	5.60	5.69	5.95
18	6.61	5.42	5.31	5.44	5.68
19	6.17	4.92	4.75	4.92	5.21
20	7.12	5.75	5.60	5.81	6.12
21	7.48	7.14	7.12	7.28	7.49
22	7.26	5.73	5.65	5.76	6.07
23	7.08	5.90	5.81	6.12	6.45
24	6.33	4.98	4.96	5.12	5.34
25	6.52	5.74	5.69	5.85	6.11
26	9.10	7.44	7.41	7.71	7.98
27	6.54	5.73	5.61	5.71	5.85
28	7.41	6.01	5.99	6.04	6.14
29	7.54	7.18	7.18	7.20	7.24
30	5.71	4.39	4.36	4.48	4.59
31	7.06	5.68	5.41	5.87	6.41
32	6.22	5.25	5.15	5.19	5.41
33	6.17	5.04	4.89	5.20	5.45
34	7.47	6.13	6.01	6.38	6.66
35	6.49	5.33	5.18	5.58	5.91
mean	7.04	5.86	5.74	5.97	6.25

Figure 7.9: CIPIC - Left

Subject	PCA	2DPCA	GLRAM	TSVD	PLSR
1	6.08	5.13	5.07	5.08	5.36
2	6.47	5.32	5.11	5.49	5.84
3	6.32	5.59	5.37	5.51	5.72
4	5.98	5.02	4.98	5.10	5.22
5	5.72	4.82	4.72	4.71	4.92
6	6.17	5.04	4.82	5.24	5.83
7	7.41	6.72	6.47	6.62	6.87
8	6.34	5.41	5.33	5.51	5.75
9	6.18	5.19	5.05	5.39	5.64
10	6.16	5.24	5.18	5.33	5.49
11	7.30	6.35	6.25	6.52	6.72
12	6.82	6.17	5.97	6.09	6.33
13	6.74	5.83	5.67	5.87	6.10
14	6.43	5.49	5.36	5.49	5.86
15	6.23	5.19	5.16	5.33	5.44
16	6.59	5.61	5.55	5.76	6.07
17	6.10	5.30	5.19	5.19	5.42
18	6.31	5.35	5.21	5.45	5.62
19	5.83	4.83	4.74	4.88	5.07
20	5.56	4.54	4.36	4.59	4.86
21	6.02	5.15	5.03	5.10	5.36
22	7.01	6.15	6.09	6.20	6.38
23	6.26	5.45	5.44	5.55	5.67
24	6.27	5.33	5.16	5.39	5.71
25	5.71	4.59	4.47	4.64	4.83
26	5.62	4.56	4.46	4.77	4.91
27	6.06	4.76	4.69	5.07	5.45
28	5.72	4.87	4.84	4.96	5.13
29	6.30	5.29	5.19	5.45	5.77
30	6.47	5.60	5.55	5.63	5.74
31	5.69	4.66	4.56	4.64	4.95
32	6.19	5.45	5.38	5.47	5.57
33	5.96	4.99	5.03	5.14	5.21
34	7.15	6.13	5.96	6.19	6.42
35	5.80	5.00	4.85	5.00	5.25
mean	6.26	5.32	5.21	5.38	5.61

Figure 7.10: LDV - Left

7 Figures and Tables (Appendix)

Subject	PCA	2DPCA	GLRAM	TSVD	PLSR
1	6.46	5.16	4.89	5.34	5.93
2	7.89	6.48	6.34	6.65	7.01
3	7.65	6.48	6.32	6.41	6.70
4	8.00	6.72	6.61	6.83	7.18
5	6.38	5.32	5.19	5.30	5.53
6	7.04	5.90	5.60	5.86	6.31
7	8.07	6.76	6.39	6.65	6.99
8	7.00	5.69	5.47	5.69	5.97
9	7.20	6.30	6.29	6.30	6.37
10	8.19	7.44	7.31	7.46	7.73
11	7.96	6.04	5.89	6.92	7.59
12	7.06	5.92	5.89	6.02	6.24
13	6.86	5.51	5.44	5.60	5.71
14	7.69	6.54	6.47	6.69	6.95
15	7.58	7.03	7.03	7.09	7.18
16	7.76	6.67	6.65	6.83	7.00
17	6.80	5.88	5.74	5.93	6.16
18	6.74	5.48	5.46	5.50	5.72
19	6.38	4.78	4.72	5.02	5.29
20	7.01	5.89	5.76	5.96	6.29
21	7.26	6.06	5.97	6.17	6.42
22	6.94	5.70	5.52	5.74	6.08
23	7.38	5.98	5.84	6.07	6.37
24	6.03	4.69	4.68	4.82	5.07
25	7.14	5.92	5.76	5.88	6.13
26	7.85	6.67	6.60	6.96	7.29
27	7.31	6.21	6.15	6.19	6.34
28	6.34	5.08	5.02	5.15	5.26
29	7.66	6.54	6.53	6.59	6.64
30	6.22	5.09	5.02	5.20	5.30
31	7.23	5.73	5.52	6.04	6.62
32	6.20	5.16	5.08	5.13	5.33
33	6.59	5.47	5.40	5.69	5.91
34	6.87	5.82	5.55	5.72	5.99
35	6.39	5.34	5.12	5.38	5.76
mean	7.12	5.93	5.81	6.02	6.30

Figure 7.11: CIPIC - Right

Subject	PCA	2DPCA	GLRAM	TSVD	PLSR
1	5.77	4.69	4.60	4.83	5.10
2	5.97	5.08	4.79	5.07	5.40
3	5.94	4.98	4.87	4.99	5.21
4	5.94	4.84	4.81	5.06	5.24
5	6.10	5.29	5.12	5.30	5.44
6	6.08	4.49	4.46	5.04	5.64
7	7.32	6.75	6.62	6.62	6.88
8	5.69	4.47	4.44	4.80	5.10
9	6.41	5.76	5.31	5.59	5.94
10	5.45	4.60	4.63	4.62	4.67
11	6.53	6.23	5.67	5.79	6.02
12	6.86	6.35	5.90	6.08	6.29
13	5.92	5.20	4.88	5.01	5.28
14	6.29	4.85	4.97	5.42	5.75
15	6.42	5.61	5.49	5.67	5.87
16	6.17	5.26	5.16	5.33	5.69
17	6.85	6.30	5.97	6.08	6.44
18	5.98	4.90	4.89	5.07	5.29
19	6.15	5.54	5.26	5.30	5.50
20	6.17	4.96	5.01	5.35	5.60
21	6.62	5.99	5.63	5.82	6.07
22	6.04	5.24	5.09	5.18	5.38
23	5.91	5.00	4.94	5.04	5.22
24	6.60	5.48	5.74	5.82	6.07
25	5.72	4.86	4.76	4.83	4.93
26	5.60	4.78	4.54	4.74	4.92
27	5.93	4.95	4.88	5.13	5.34
28	5.83	4.96	4.76	4.96	5.15
29	6.25	4.85	5.12	5.49	5.72
30	5.77	4.73	4.71	4.83	5.05
31	6.56	5.72	5.41	5.73	6.00
32	5.90	5.14	5.03	5.13	5.21
33	5.70	4.78	4.67	4.84	4.99
34	6.33	5.26	5.25	5.48	5.73
35	6.07	4.96	5.04	5.20	5.45
mean	6.14	5.22	5.10	5.29	5.53

Figure 7.12: LDV - Right

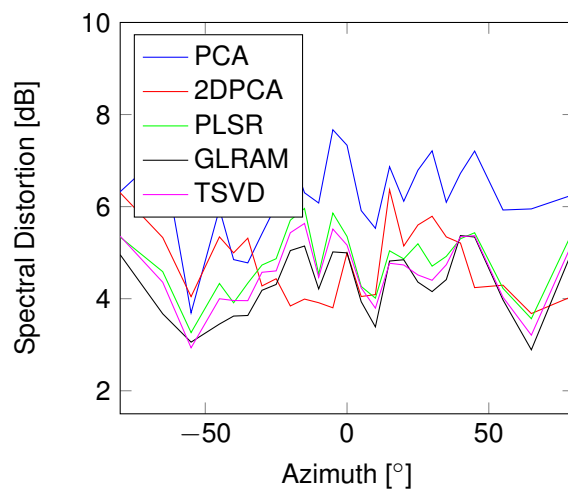


Figure 7.13: SD values for fixed elevation, CIPIC, Subject 30, Elevation 0° - Left

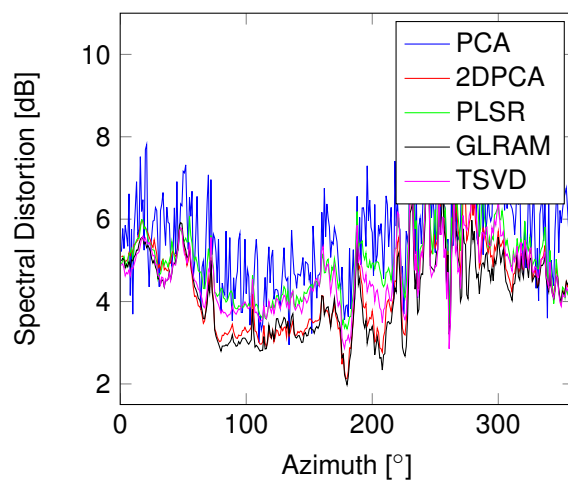


Figure 7.14: SD values for fixed elevation, LDV, Subject 20, Elevation 0° - Left

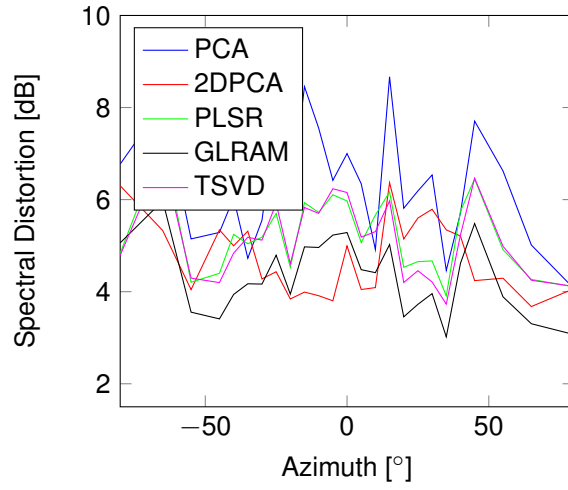


Figure 7.15: SD values for fixed elevation, CIPIC, Subject 30, Elevation 0° - Right

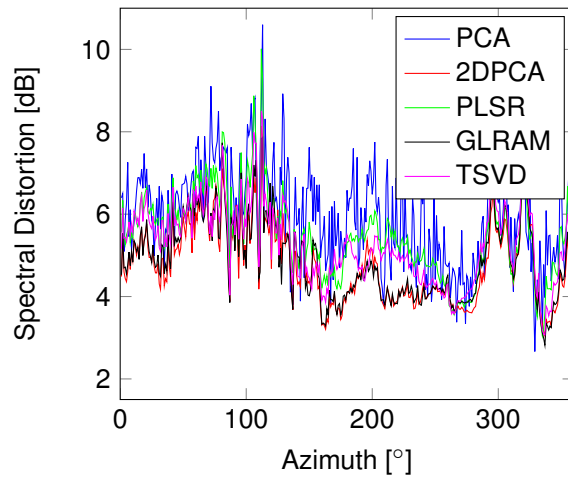


Figure 7.16: SD values for fixed elevation, LDV, Subject 20, Elevation 0° - Right

Bibliography

- [1] V.R. Algazi, R.O.Duda, and D.M.Thompson. The use of head-and-torso models for improved spatial sound synthesis. In 113th *Audio Engineering Society Convention*, pages 1–18, 2002.
- [2] B.Savas and L.-H.Lim. Best multilinear rank approximation of tensors with quasi-newton methods on grassmannians. Technical Report LITH-MAT-R-2008-01-SE, Department of Mathematics, Linköpings Universitet, 2008.
- [3] B.Seeber and H.Fastl. Subjective selection of nonindividual head-related transfer functions. In *In Proceedings of the International Conference on Auditory Display*, pages 1–4, 2003.
- [4] C.Andersson and R.Bro. The N-way Toolbox for MATLAB. *Chemometrics and Intelligent Laboratory Systems*, 52(1):1–4, 2000.
- [5] D.J.Kistler and F.L.Wightman. A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction. *The Journal of the Acoustical Society of America*, 91(3):1637–1647, 1992.
- [6] D.N.Zotkin, J.Hwang, R.Duraiswami, and L.S.Davis. Hrtf personalization using anthropometric measurements. In *Workshop on Applications of Signal Processing to Audio and Acoustics*, pages 157–160, 2003.
- [7] Nail A. Gumerov, Ramani Duraiswami, and Dmitry N. Zotkin. Fast multipole accelerated boundary elements for numerical computation of the head related transfer function. In *International Conference on Acoustics, Speech and Signal Processing*, pages 165–168, 2007.
- [8] Zoltan Haraszy, David-George Cristea, Virgil Tiponut, and Titus Slavici. Improved head related transfer function generation and testing for acoustic virtual reality development. In *Proceedings of the WSEAS international conference on Systems*, pages 411–416, 2010.
- [9] Zoltan Haraszy, Daniel Ianchis, and Virgil Tiponut. Generation of the head related transfer functions using artificial neural networks. In *Proceedings of the WSEAS international conference on Circuits*, pages 114–118, 2009.
- [10] H.Hu, L.Zhou, J.Zhang, H.Ma, and Z.Wu. Head related transfer function personalization based on multiple regression analysis. In *International Conference on Computational Intelligence and Security*, pages 1829 – 1832, 2006.
- [11] H.Martens and T.Naes. *Multivariate Calibration*. John Wiley & Sons Ltd., New Jersey, 1989.
- [12] I.T.Jolliffe. *Principal Component Analysis*. Springer-Verlag, New York, second edition, 2002.

- [13] J.Yang, D.Zhang, A.F.Frangi, and J.Yang. Two-dimensional PCA: a new approach to appearance-based face representation and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(1):131–137, 2004.
- [14] J.Ye. Generalized low rank approximations of matrices. *Machine Learning Journal*, 61:167–191, 2005.
- [15] A. Kuhn and M. Rothbucher. HRTF Customization by Regression. Technical report, 2013.
- [16] M.Rothbucher, M.Durkovic, H.Shen, and K.Diepold. HRTF customization using multiway array analysis. In *Proceedings of the European Signal Processing Conference*, pages 229–233, 2010.
- [17] Makoto Otani, Tatsuya Hirahara, and Shiro Ise. Numerical study on source-distance dependency of head-related transfer functions. *The Journal of the Acoustical Society of America*, 125:3253, 2009.
- [18] S.de Jong. SIMPLS: An alternative approach to partial least squares regression. *Chemometrics and Intelligent Laboratory Systems*, 18(3):251–263, 1993.
- [19] S.Dong, R.Hu, W.Tu, X.Zheng, J.Jiang, and S.Wang. Enhanced principal component using polar coordinate pca for stereo audio coding. In *International Conference on Multimedia and Expo*, pages 628–633, 2012.
- [20] S.Hwang, Y.Park, and Y.-S.Park. Modeling and customization of head-related transfer functions using principal component analysis. In *International Conference on Control, Automation and Systems*, pages 227–231, 2008.
- [21] S.Weiland and F.van Belzen. Singular value decompositions and low rank approximations of tensors. *IEEE Transactions on Signal Processing*, 58(3):1171–1182, 2010.
- [22] S.Wold, M.Sjöström, and L.Eriksson. PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, 58:109–130, 2001.
- [23] S.Yairi, Y.Iwaya, and Y.Suzuki. Individualization feature of head-related transfer functions based on subjective evaluation. In *Proceedings of the International Conference on Auditory Display*, pages 340–343, 2008.
- [24] T.G.Kolda and B.W.Bader. Tensor decompositions and applications. *SIAM Review*, 51(3):455–500, 2009.
- [25] W.Hugeng and D.Gunawan. Improved method for individualization of head-related transfer functions on horizontal plane using reduced number of anthropometric measurements. *Journal of Telecommunications*, 2(2):31–41, 2010.
- [26] W.W.Hugeng and D.Gunawan. Effective preprocessing in modeling head-related impulse responses based on principal components analysis. *Signal Processing: An International Journal*, 4(4):201–212, 2010.