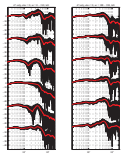
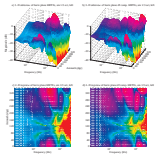


Computational Quality Assessment of HRTFs



Klaus A J Riederer

Helsinki University of Technology (HUT), Laboratory of Computational Engineering & Laboratory of Acoustics and Audio Signal Processing
Miestentie 3, P.O. Box 9400, FIN-02015 HUT, Finland. Email: Klaus.Riederer@hut.fi URL:http://www.lce.hut.fi/~kar

ABSTRACT

This paper presents an approach to calculate an objective deviant measure of HRTFs from the common trend of a HRTF database. The perceptually justified deviance index is applied for the evaluation of the HRTF measurement system and database. The index shows the relative behavior of HRTFs but it does not show their explicit "goodness".

HEAD-RELATED TRANSFER FUNCTION, HRTF

Measured HRTFs include all the basic binaural cues received by the human auditory system, i.e., interaural level and time differences (ILDs & ITDs).

- Effective HRTF length is usually considered as ca. 100-200 sample long impulse responses (HRIRs, head-related impulse responses), which are synthesized by FIRs or IIRs to produce 3-D sound.
- HRTFs depend highly on the person, direction, and frequency; hence *there are no "correct" HRTF shapes.*
- Measuring HRTFs is not straightforward, there are many kinds of disrupting factors:
  - acoustic (background noise, unwanted reflections etc.)
  - physical (head shadowing, movement of the test subject etc.)
  - electrical (A/D conversion errors, i.e., DC-offset etc.).

⇒ Degraded HRTFs are difficult to separate from the uncorrupted ones.

PCA MODELING of HRTFs

Too precise HRIRs introduce vivid person-dependent HRTF fine structures which make it very difficult to find imperfections in such kind of responses. Too high precision may be also excessive from the human perception point of view; the auditory resolution decreases at high frequencies, and the localization accuracy depends on the direction.

⇒ Modeling of HRTFs is needed in order to able to estimate a common HRTF structure.

Method

Based on good results presented by Wightman and Kistler (1992), their method based on principal component analysis (PCA) was applied as a starting point:

- Database consisted of both ear responses HRTF magnitude responses from 55 subjects and 259 measurement angles (7 elevations, 37 azimuth angles).
- An optimal BARK warping (Smith and Abel 1995) was used to create a perceptually correct frequency resolution.
- The amount of HRIR sample points was studied in order to obtain proper amount of spectral resolution; 128 samples were applied.
- The HRTF database was PCA modeled per one measurement angle and per ear , i.e., input data was 128 *55 (rows * columns; p*n).
- Before the singular value decomposition, the input data was first centered and scaled to diminish subject-dependency and direction-indepency from the measured HRTFs.

⇒ The first 2-7 principal components (PCs) explained ca. 95% of the total variance depending on the direction; PC-1 explained ca. 80-90% of the total variance.

Logical concept

The principal component reconstruction equals to a weighted summation over the input data, similar as in Fourier synthesis.

- Because PC-1 contains the minimum amount of individual fine structures, PC-1 holds the information of the "common trend" from all the 55 different HRTFs (per measurement angle and ear).
- A reference (modeled) HRTF is reconstructed by multiplying only the PC-1 (the first basis vector) and the corresponding individual weights (& adding the subject-mean).

⇒ The PC-1 modeled HRTF (PCA_1) denotes to a "clean" HRTF, approximating the common HRTF shape and neglecting the individual fine structures, thus minimizing the individual faults in the original raw measurements.

CALCULATION OF DEVIANCE

An *error-count* (*ERR*) was obtained by:

$$ERR_{person,angle,ear} = \frac{\sum_n (PCA_{1,person,ear,n} - HRTF_{person,ear,n})^2}{\sum_n PCA_{1,person,ear,n}^{\ 2}}$$

ERR is scaled to get a suitable range for *deviance* (*DEV*):

$$DEV_{person,angle,ear} = 1 - \tanh(\alpha_{angle,ear} * ERR_{person,angle,ear}),$$

where the scaling factor has been chosen factually as

$$\alpha_{angle,ear} = \frac{arctanh(0.5)}{mean(ERR_{person,angle,ear})}$$

⇒ The value of the deviance *DEV* is [0,1] limited.

- A measured HRTF that differs strongly from the common trend of all the HRTFs has a low DEV, close to 0.

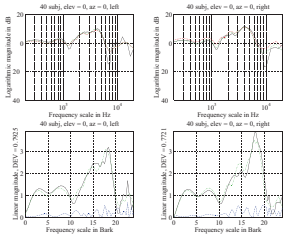


Figure 1. Only a slight deviation from the common trend is noted, except at the antiresonances at high frequencies.

PC-1 modeled HRTF = green curve

Mean of all HRTFs = red curve

Measured HRTF = black curve

Green - black = blue curve, area of which ~ DEV

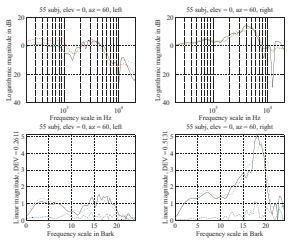


Figure 2. In the left ear rather high deviance (also at low frequencies), for the right ear only at the antiresonancies at high frequencies.

STATISTICAL (QUANTITATIVE) ANALYSIS of HRTF DATABASE

Investigating the distribution of the deviance indexes reveals qualitative and quantitative quality aspects of the HRTF database:

- the overall quality of the HRTF measurement system
- the atypical measurements, and the atypical test subjects.

Depending on the view point, the calculated deviance indexes show different aspects from the HRTF database.

1) Deviating HRTFs with DEVs under a threshold

- Crosses forming vertical lines reveal the most deviant test subjects – threshold = 0.2 in Fig. 3 indicates the deviating dummy heads (subject numbers 2 and 54) with wide and flat ear canals.
- there are no "faulty" azimuths in the measurement system clearly – those would be indicated by crosses forming horizontal lines.

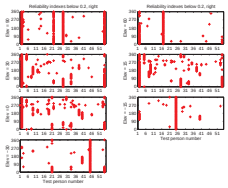


Figure 3. Right ear HRTFs, DEV < 0.2.

2) Median plots per elevation (37 azimuths, averaged per person)

- Only minor variation between the subjects, medians varying < 0.2, except at the 90° elevation (~ 0.3); see Fig. 4.
  - 90° elevation seems to be more susceptible to asymmetrical measurement positions in the median plane, i.e., pivoting (sideward tilting) of the head, than in the other directions [Riederer (1998a) has shown that the pivoting is minor].

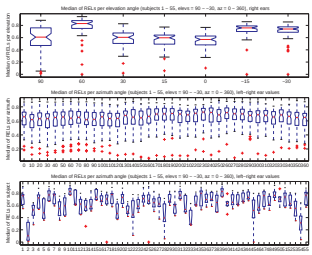


Figure 4. Median values for right ear HRTF measurements; a) per elevation angle b) per azimuth angle, c) per test subject.

3) Median values per azimuth angle (subjects averaged per subject over all 7 elevations)

- Much less variation (0.7..0.8) is seen than in the prior elevation case – now averaging is done only over 7 elevations (notice also Fig. 3.)
- Minor variation between different azimuth angles, slightly higher DEVs are obtained at ipsilateral directions (the same is observed for the left ear data).
  - HRTF shape is smoother in the head shadow directions, which causes slightly smaller ERR values.
- Minor variation between different azimuth angles, slightly higher DEVs are obtained at ipsilateral directions (the same is observed for the left ear data).

4) Median values per test subject (259 HRTFs per one subject)

- Large variation between the test subjects; the maximum deviance is noted with subject numbers 2 and 54 (medians ca. 0.1 & 0.3) and minimum with subjects 10 and 39 (median ca. 0.8)
- The inter-individual results give an indication for a HRTF clustering analysis.

The deviance index is an implicit measure, showing relative behavior of HRTFs; however, it does not show the explicit "goodness" of HRTFs.

DISCUSSION

HRTFs pose a complicated problem field due to their intricate acoustic nature; e.g., the highly person-dependency at high frequencies (above ca. 7 kHz). E.g. ,vivid head movements cause strong frequency deviations that might be undistinguishable from very atypical head shapes.

The presented method proved the high quality of the applied HRTF measurement system but did not show an explicit quality measure for particular HRTFs. For the latter purpose, more sophisticated method analyzing also the time-domain and band-limited frequency responses is necessary.

The most important is the perceptual effect: when HRTFs are applied in 3-D sound reproduction, listening tests are indispensable in assessing the quality of HRTFs.

ACKNOWLEDGEMENTS

The author wishes to thank Dr.Tech. Jyri Huopaniemi (Nokia Research Center), M.Sc. Matti Hämäläinen (Nokia Research Center), Dr.Tech. Vesa Välimäki (HUT), Prof. Matti Karjalainen (HUT), M.Sc. Timo Kostiainen (HUT) and M.Sc. Harri Lappalainen (HUT)