Proceedings of 20th International Congress on Acoustics, ICA 2010

23-27 August 2010, Sydney, Australia

Representations of HRTFs using MATLAB: 2D and 3D plots of accurate dummy-head measurements

György Wersényi

Széchenyi István University, H-9026, Győr, Egyetem tér 1, Hungary

PACS: 43.66.Qp, 43.66.Pn

ABSTRACT

Human Head-Related Transfer Functions describe the transmission from the free-field to the eardrums. HRTFs are measured on human subjects or on dummy-heads, characterized by the angle of incidence. The dummy-head measurement method allows the acquisition of data in high spatial resolution. Our setup provided HRTF data in 1 degree horizontal and 5 degrees elevational steps in different environmental settings. Spectral evaluation in spatial hearing research requires proper representation methods of detailed measurement data. Different 2D and 3D representation methods will be presented here, using different coordinate systems, color maps and additional filtering methods programmed under MATLAB. Figures are mainly helpful for HRTF analysis but MATLAB features allow other use for applications where directional characteristics, polar plots are required.

INTRODUCTION

Human spatial hearing research includes directional hearing tasks, localization performance investigations, measurement and recording techniques, playback methods in virtual reality applications, etc. Finding the location of the sound source is the most critical issue. The auditory system utilizes the interaural level differences (ILD) and the interaural time differences (ITD) measured between the sound pressures at the eardrums [1-3]. Furthermore, spectral cues introduced by the direction-dependent filtering of the outer ears (mainly the pinna, the head and the torso) help resolve localization tasks [4-6]. The latter cue is described by the transfer function of the outer ear: the Head-Related Transfer Functions (HRTFs) or its time-domain equivalent, the Head-Related Impulse Responses (HRIRs). These complex transfer functions describe the transmission from the free-field to the eardrum, and are characterized by the distance and angle of incidence [7-9]. The commonly used rectangular coordinate system is attached to the head and HRTFs for a fixed source distance are given by the azimuth angle φ (horizontal plane) and elevation angle δ (median plane). Another coordinate system sets the axis through the head and the ears and describes sound source directions using polar and lateral angles [10]. Monaural HRTFs for each ear filter the incident sound waves simultaneously, and thus, interaural spectral differences appear between the two ears. Based on these cues, the auditory system tries to determine the direction of the sound source. According to the literature, the localization blur depends on various parameters such as signal bandwidth, environmental parameters, playback method, etc. [2, 11, 12].

In case of free-field listening (anechoic chamber environment) human subjects use their own ears and individual HRTFs for localization. Signals are played back using loudspeakers. The best spatial resolution under optimal conditions can be as accurate as about 1 degree horizontally and 5-10 degrees vertically [2]. On the other hand, virtual localization tasks use headphone playback systems and additional HRTF filtering. In practice, a mono sound file is played back and it is filtered with the left and right ears' HRTFs. Furthermore, the playback chain, first of all, the headphone, has to be spectrally equalized [11, 13]. Theoretically, this results in a "perfect" spatial simulation. In general, localization performance (e.g. front-back confusion rates, in-the-head localization, etc.) is inferior to free-field listening [14-16]. During virtual localization the HRTFs play a significant role. Individually recorded or custom selected HRTFs (from a given set of HRTF sets [17]) are better than non-individual and dummy-head HRTFs [18-21]. Furthermore, the spatial resolution of the filter set can affect the results (interpolated HRTFs) [22, 23].

HRTFs can be recorded on human subjects or on dummy-heads. There are different methods using different excitation signals and spatial resolution. Dummy-head HRTFs proved to be inappropriate for accurate localization, but they are suitable for accurate recordings [24]. Whilst human subjects tolerate only short-term measurements (using impulse response techniques and low spatial resolution of HRTF data), dummy-heads allow long-term measurements including high spatial resolution, noise or sweep excitation signals, averaging of results and thus, more accurate measurements.

The role of accuracy, reproducibility and overall "quality" of HRTFs in virtual simulation and sound field rendering has been a question for a long time [22, 25, 26]. Variations in the fine structure of the magnitude of HRTFs may lead to increased localization errors in virtual audio simulation. The hearing system is more sensitive to minor changes in the HRTF structure in case of virtual sound field rendering, and this indicates the fine structure to be significant. On the other hand, under free-field listening conditions the hearing system is less sensitive to such variations. In order to investigate how the environment near the head influences the HRTFs, a dummy-head measurement system was installed in the anechoic chamber. After recording the naked dummy-head HRTF set, different environmental settings were applied. The torso was equipped with hair, clothing, caps, glasses, etc. Using high spatial resolution and spectral accuracy, a huge database of HRTFs was recorded (approx. 30,000 HRTFs for each ear).

Handling this amount of data requires appropriate presentation methods [10, 27]. This includes one-dimensional plots of spectral data, two-dimensional rectangular and polar diagrams as well as three-dimensional color plots. In addition, in order to use different filtering and averaging methods on measured data, to eliminate measurement errors, to reduce the dynamic range, to scale axes either linear or logarithmic and to have a user friendly GUI, a suitable software environment is required.

This paper first introduces the measurement procedure, the system setup and data formats we used. Then, different figures plotted by pre-defined and modified MATLAB routines, will show a powerful way to represent measured data in order to evaluate them spectrally. We will highlight problems, introduce a mathematical tool (HRTFD) to assist evaluation and show a collection of figures to demonstrate the variations of the fine structure in HRTF data.

MEASUREMENT SETUP

HRTF Differences

HRTFs are defined with (1)

$$HRTF = \frac{p_1(j\omega)}{p_2(j\omega)} \tag{1}$$

where p_1 is the sound pressure at the eardrum and p_2 is the sound pressure of the reference signal measured with an omnidirectional microphone at the origin of the head-related coordinate system.

In addition to (1), interaural differences can be defined by dividing the left ear's HRTF by the right ear's HRTF. Interaural HRTFs show spectral differences between the two ears' HRTFs from the same sound source direction.

Similarly, monaural HRTF differences (HRTFD) can be defined as the quotient of two monaural HRTFs from the same direction: a reference condition and a modified environmental condition [28, 29]. E.g. the reference condition is the naked torso and the modified environment is the dummy-head equipped with clothing. Thus, the HRTFD will give the spectral properties (the transfer function) of the clothing. This means, HRTFDs are free from any kind of individual property. Accurately measured and calculated HRTFDs are powerful tools for investigating changes and differences in the fine structure of HRTFs of about 1 dB [28, 30].

Setup

The Brüel & Kjaer Head and Torso Simulator Type 4128 was placed on a turntable in the anechoic chamber. The turntable was controlled by a computer in one degree steps. A laser targeting system was used to set the elevation of a loudspeaker from -40 degrees up to +90 degrees in 5 and 10-degree steps [31]. Excitation signal was a pseudo-random white noise signal of 81,92 ms having a frequency independent SNR [32]. HRTFs were measured for both ears simultaneously with 50 kHz sampling frequency and 16 bits resolution. The AT&T Ariel DSP card was used and programmed with its assembler and C++ routines. Responses were accumulated and averaged over time against uncorrelated measurement noise that resulted in a measurement SNR of about 89 dB [32].

Time functions were stored in .dat files using 4-byte longintegers for two-channels (left and right ear). These were accumulated and averaged over time by the DSP card. After recording, a 4096-point FFT was applied and HRTFs were stored for both ears separated in binary format. Each ear's HRTF file includes 2048 samples from 0 to 25 kHz in 12,2 Hz linear spectral resolution. Since the excitation was a pseudo-random white noise signal (having random phase information) and because our goal is to investigate the magnitude response, only the latter was saved.

PRESENTATION AND EVALUATION OF HRTF DATA

Data management requires a presentation method that is meaningful, spectacular and comfortable to use. The MATLAB package is able to read and interpret binary formatted data structures. The basic function such as *plot, plot3, mesh* and *contour* can serve a GUI. For one-dimensional and two-dimensional plotting we applied *plot* using logarithmic axes. *Plot3* is monochrome, *mesh* and *contour* use color scales and fillings during three-dimensional plotting similar to geographical topos and maps.

For the representations we used logarithmic axes as usual, both in frequency and magnitude (dB). Because the dummy-head is symmetrical and stable during measurement (in contrast to human subjects), evaluation was made here monaurally, for one ear and for the magnitude response only. Our former evaluation used grey-scaled 2D plots of unsigned absolute values for evaluating the HRTFDs in 10-degree horizontal resolution [28–30]. The improved representation methods offered here allow us to investigate HRTF data in one degree horizontal resolution and 2D and 3D color plots.

Presentation as function of frequency

In addition to the figures presented here, MATLAB allows by using plot function to create animated videos and save them as AVI files, composed of different figures in a selected order. For example, we are able to watch horizontal plane HRTFs vary as function of azimuth monaurally and binaurally as well.

Figure 1 shows HRTFs of the left and right ear if the sound source is at 90 degrees of azimuth in the horizontal plane, so the left ear is radiated directly while the right ear is in the head-shadow. Note the decreased sound pressure level of the shadowed ear.

The head-shadow area is created by the head as the ear is shadowed and the sound source is on the contralateral side. This shadowing effect creates a "noisy" spatial domain where signal level is low, high frequency evaluation of HRTFs is difficult and localization performance decreases [33]. Figure 2 shows 20 HRTFs from 250 to 270 degrees in the horizontal plane that are affected by the head-shadow the most.

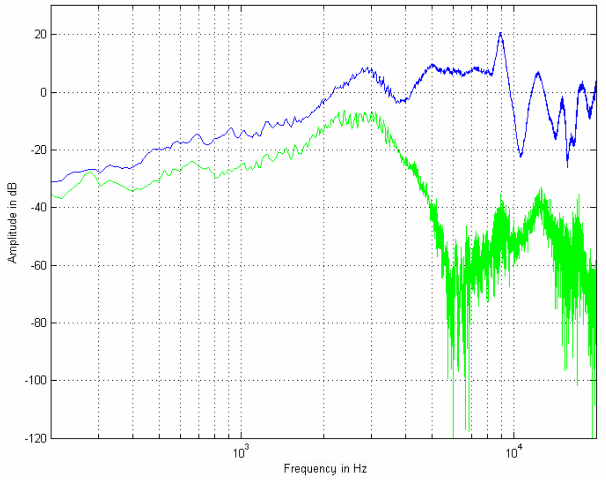

Figure 1. HRTFs of the naked dummy-head for both ears: left ear (blue), right ear (green) as sound source direction is φ=90º.

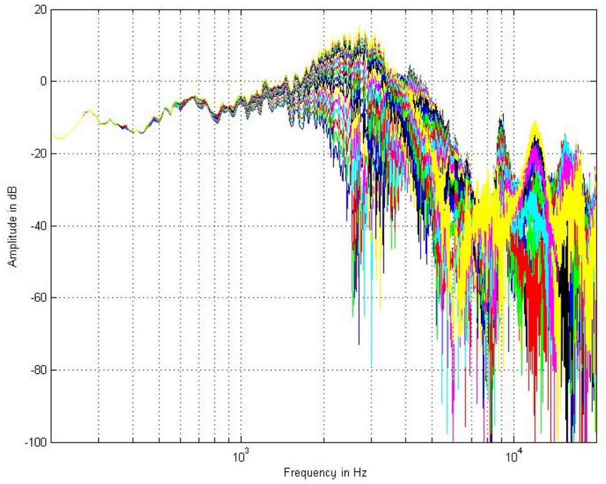

Figure 2. Monaural HRTFs from the head-shadow area, φ=250º-270º.

Figure 3 presents all 360 HRTFs from the horizontal plane for one ear. Note the enormous dynamic range of more than 120 dB due to the head-shadow area (Fig.2.). Such a dynamic range can appear by plotting HRTFDs as well, because if we divide noisy signals the result also will be noisy. Compare Fig. 3. with Fig. 4. where all 360 HRTFs are plotted at the elevation +90 degrees that is exactly over the head. In this case, turning the dummy-head is pointless, all the HRTFs should be exactly the same.

Using Fig. 4. we can test the accuracy of the measurement settings and setup. The deviations are limited to ±1 dB. The 3D plot also supports the theory: all 360 measured HRTFs look the same. The upper side of Fig.4. could be any “slice” of this 3D plot.

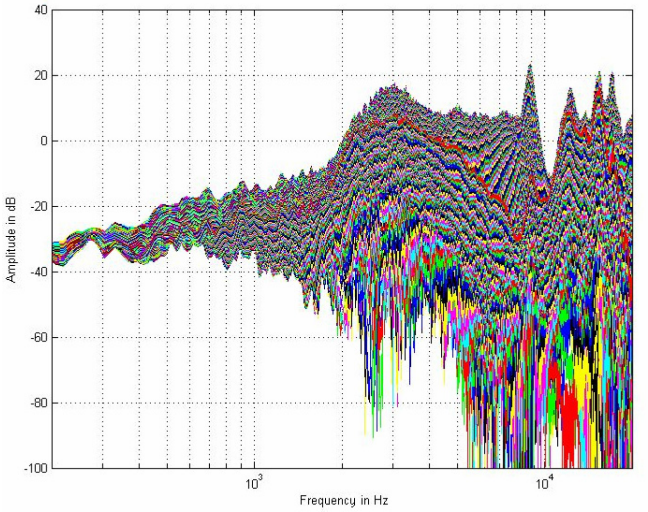

Figure 3. 360 monaural HRTFs in one-degree resolution from the horizontal plane (δ=0º). Compare with Fig 4.

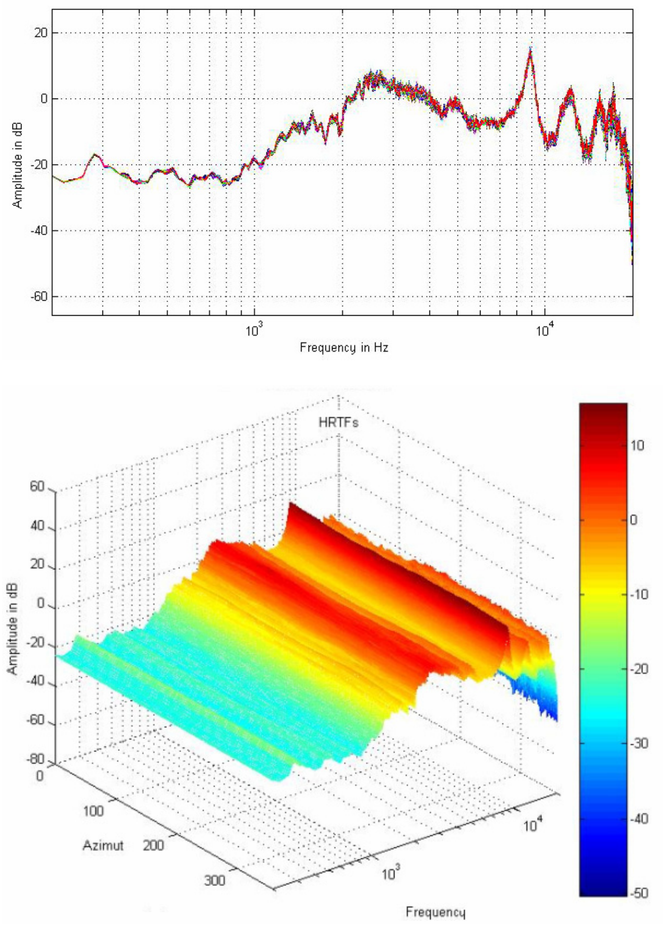

Figure 4. monaural HRTFs in one-degree resolution from “above the head” (δ=90º) plotted in 2D and in 3D. Compare with Fig 3.

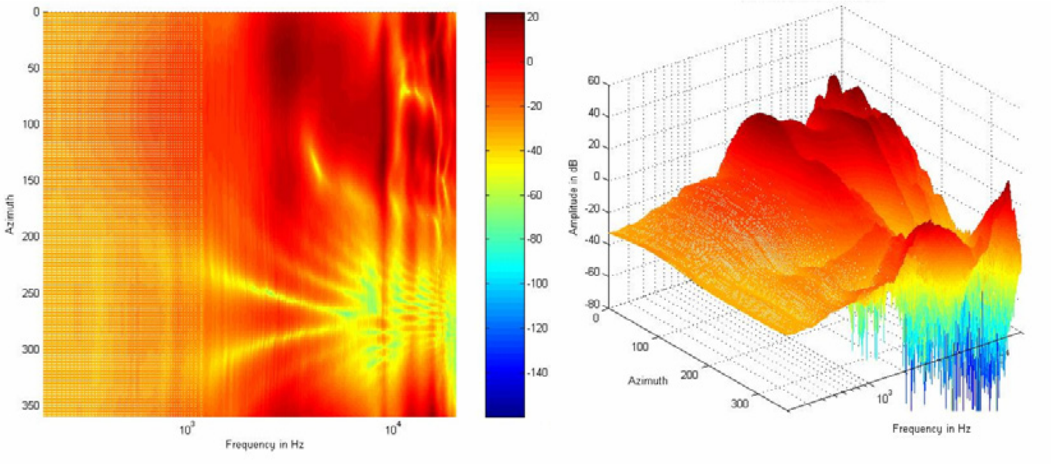

Figure 5. 360 monaural HRTFs in one-degree resolution from the horizontal plane (δ=0º) based on the same data as Fig.3. Compare with Fig 7.

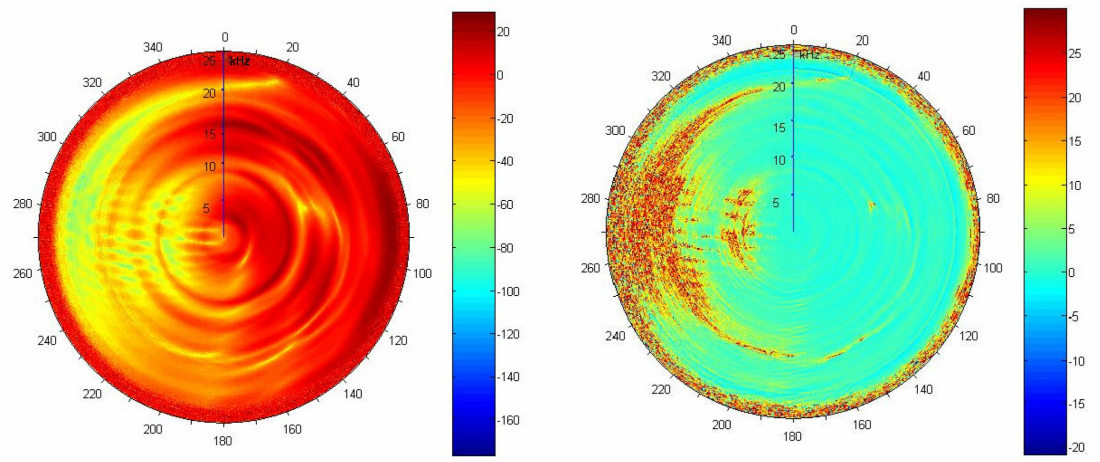

Figure 6. Polar diagrams based on Fig. 5. (left) and Fig. 7. (right).

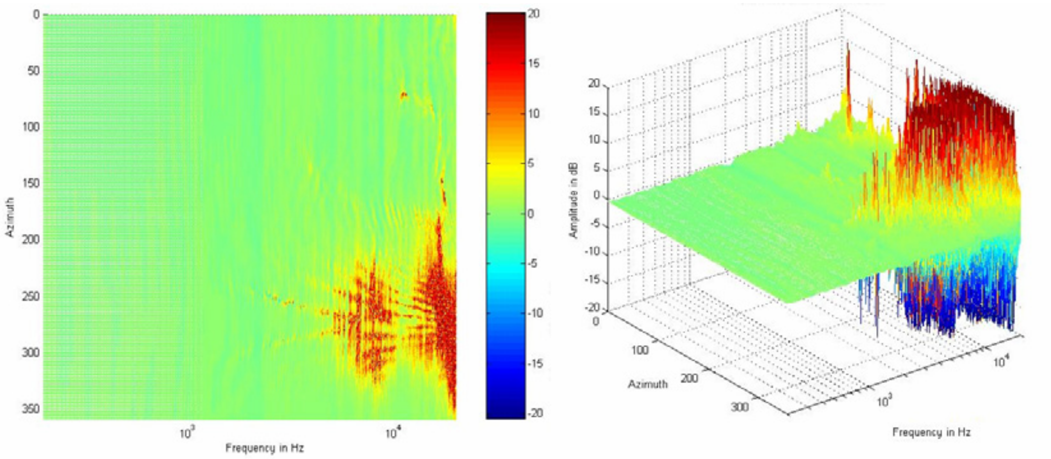

Figure 7. 360 monaural HRTFDs in one-degree resolution from the horizontal plane (δ=0º). In the head-shadow area re-measured HRTFs from the same direction show large deviations. Note the different scaling in magnitude and the notch at 11 kHz at around 60-80 degrees.

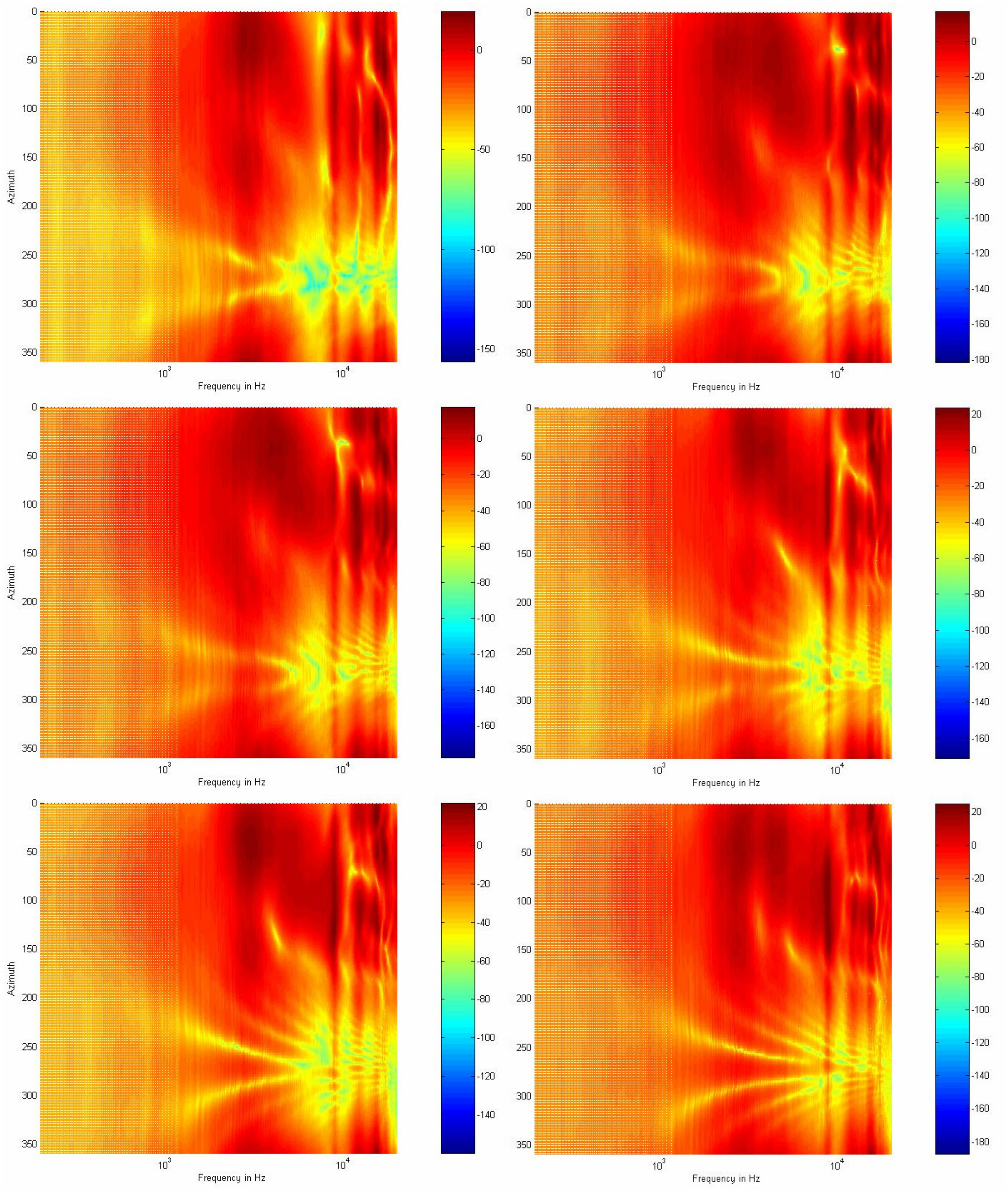

Figure 8. HRTFs as function of elevation: -40 (top left), -30 (top right), -20 (middle left), -10 (middle right), horizontal plane (bottom left), +10 (bottom right). Note the different scales automatically set by MATLAB.

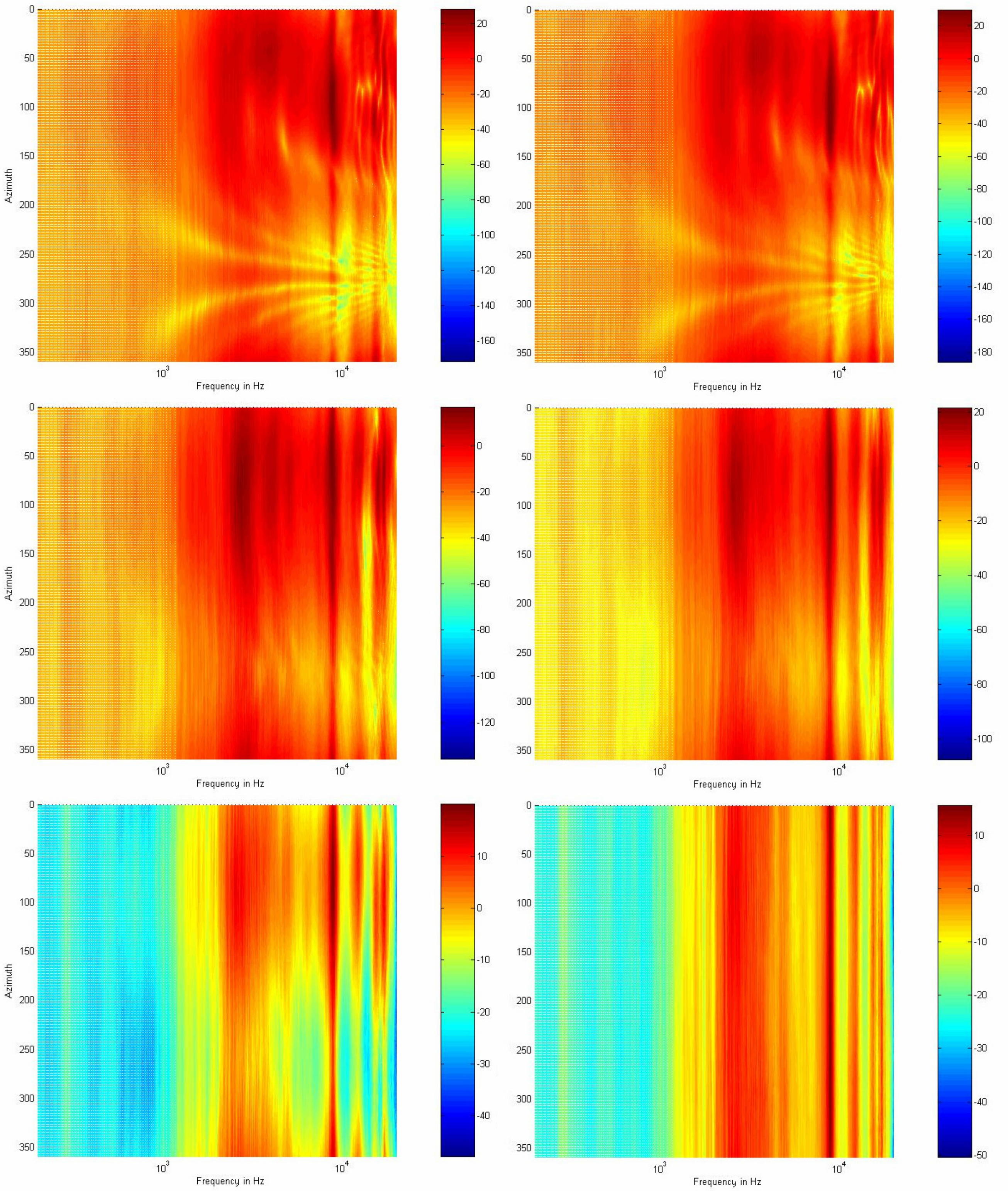

Figure 9. HRTFs as function of elevation: +20 (top left), +40 (top right), +60 (middle left), +70 (middle right), +80 (bottom left), +90 (bottom right). Note the different scales automatically set by MATLAB.

Presentation as Function of Frequency and Azimuth

This method includes 2D color plots and 3D visualization. Figure 5 shows all 360 HRTFs of one ear in 1-degree spatial resolution as function of frequency and azimuth together. The polar diagram in the left side of Figure 6 was plotted using the same data that was used for Figure 3 and 5. The polar diagram can also be used for HRTFDs. The right side of Figure 6 can be compared to the left side. All three presentation methods are well suited for deeper analysis. Nevertheless, the introduction of HRTFDs requires magnitude limitation since interesting and substantive parts can be masked if the dynamic range is too large and only few colors and tones appear on the plot. Therefore, our program has a built-in function to truncate the samples if needed: every sample that exceeds 20 dB in the HRTFD can be rounded to 20 and labeled as “difference more than 20 dB”. Figure 10 shows equal-level contours as function of frequency in the horizontal plane. We have found this kind of presentation method the less helpful.

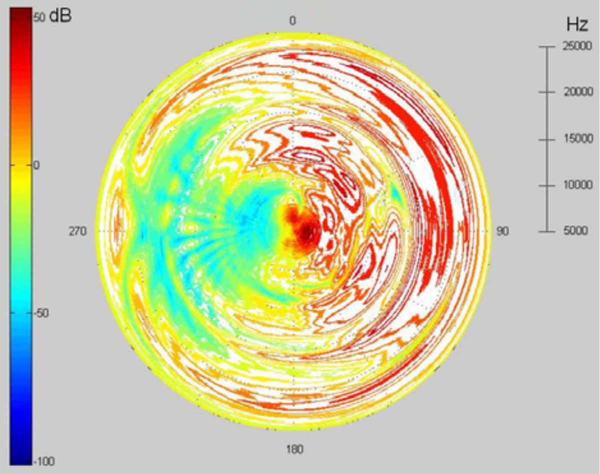

Figure 10. Equal-level contours of horizontal plane HRTFs (similarly to geographical maps).

HRTFDs were actually calculated by subtracting magnitude responses in dB. As the reference condition is the naked torso’s HRTF magnitude from a given direction, these data were subtracted sample by sample from another HRTF magnitude from the same direction. The case of no-difference would result in a flat 0 dB line over the frequency range. In addition, this method is a powerful tool to test the measurement system’s accuracy and reproducibility as well, by simply calculating this quotient with re-measured transfer functions without changing the environment. In this manner, the dummy-head has a very good reproducibility property for different directions, if the ear is radiated directly. On the other hand, head-shadowing causes large variations and thus, re-measured HRTFs from the same direction could show large deviations [28-30, 33]. Figure 7 shows HRTFDs from repeated measurements with the naked torso in the horizontal plane. In comparison with Fig.5. note the different scales in magnitude. In the head-shadow area it is not possible to re-measure the same HRTFs due to the low SNR. Note the large peak at around 60-80 degrees at 11 kHz. Responsible for this is the pinna - if we remove it from the torso, this peak disappears. The right side of Figure 6 was plotted based on Figure 7.

Figures 8-9 show the monaural HRTFs for selected different elevations from -40 degrees up to +90 degrees. Note the different scaling automatically set by MATLAB. Figures are plots using the same method as on the left side of Figure 5. We observed that pinna reflections (notches at the typical resonant frequencies) and dips in the head-shadow area (such as on Fig.2.) disappear with elevation. Above 60 degrees in elevation both effects decrease and HRTFs become more similar. Figure 4 was plotted based on the same data as the bottom right of Figure 9.

Former investigations already tried to measure high accuracy HRTFs. *Riederer* reported results in a measurement using 55 real human heads and two dummy-heads [34, 35]. This included 252 source directions, blocked ear-canal entrance measurement points by means of a miniature microphone, computer controlled turntable and pseudo-random noise excitation – conditions very similar to ours.

Although their goal was to create a high accuracy HRTF database, the spatial resolution was relatively low (7 in elevations and 10-degree steps horizontally), human subjects were not really optimal for long-term measurements and initial positioning of the subject as well as proper placement of the microphone were hard to realize. Some results are shown in Figure 11 for comparison. This resolution does not allow high accuracy evaluation. The author concluded that

- repeatability of the measurement is very sensitive to type of the ear-plug,

- middle frequencies are influenced by small movements of the subject during measurement,

- deviations up to 5-15 dB can appear in the region 5-15 kHz depending on the routine of the person who places the ear-plug,

- reflections of the knee-leg area due to the sitting position are detectable,

- and that repeatability was only investigated with one person and it is declared to be good within 1 dB deviations.

Dummy-heads are stable, they do not have to „sit” and body parts can be removed and modified easily in order to measure effects of hair and clothing. Repeatability using dummy-heads can reach 0,5 dB or less, also supported by our results. The low SNR in the head-shadow area causes the largest deviations in repeated measurements, while the best reproducibility is in frontal directions. *Riederer’s* results showed head-shadow effects up to 20 dB above 4 kHz (our results showed the same but above a clearly defined value of 3,6 kHz).

For calculating HRTFDs, averaging is also required. Four different kinds of glasses, four different but similar baseball caps and three toupees were applied on the torso and the whole measurement was repeated several times. We did not observe any significant difference among these variations, however, the results were averaged over the four glasses, four baseball caps and three haircuts. Spectral averaging was made simply by calculating the mean value sample by sample [28]. *Riederer’s* measurements can not make conclusions about the effect of hair and clothing, because only one subject participated and only few source directions were measured [35]. He observed that in sitting positions the missing trousers create larger knee-reflections as well as without a shirt. Shoulders also affect regions at 1600 and 2700 Hz. There is no clear effect of hair. Observations indicate that there is no difference between naked and dressed versions below 600 Hz. Some reflections from knee and shoulders are present between 1-4 kHz mostly on the shadowed side, and some influence around 7-9 kHz at the ipsilateral side when the ear is radiated directly by the sound source.

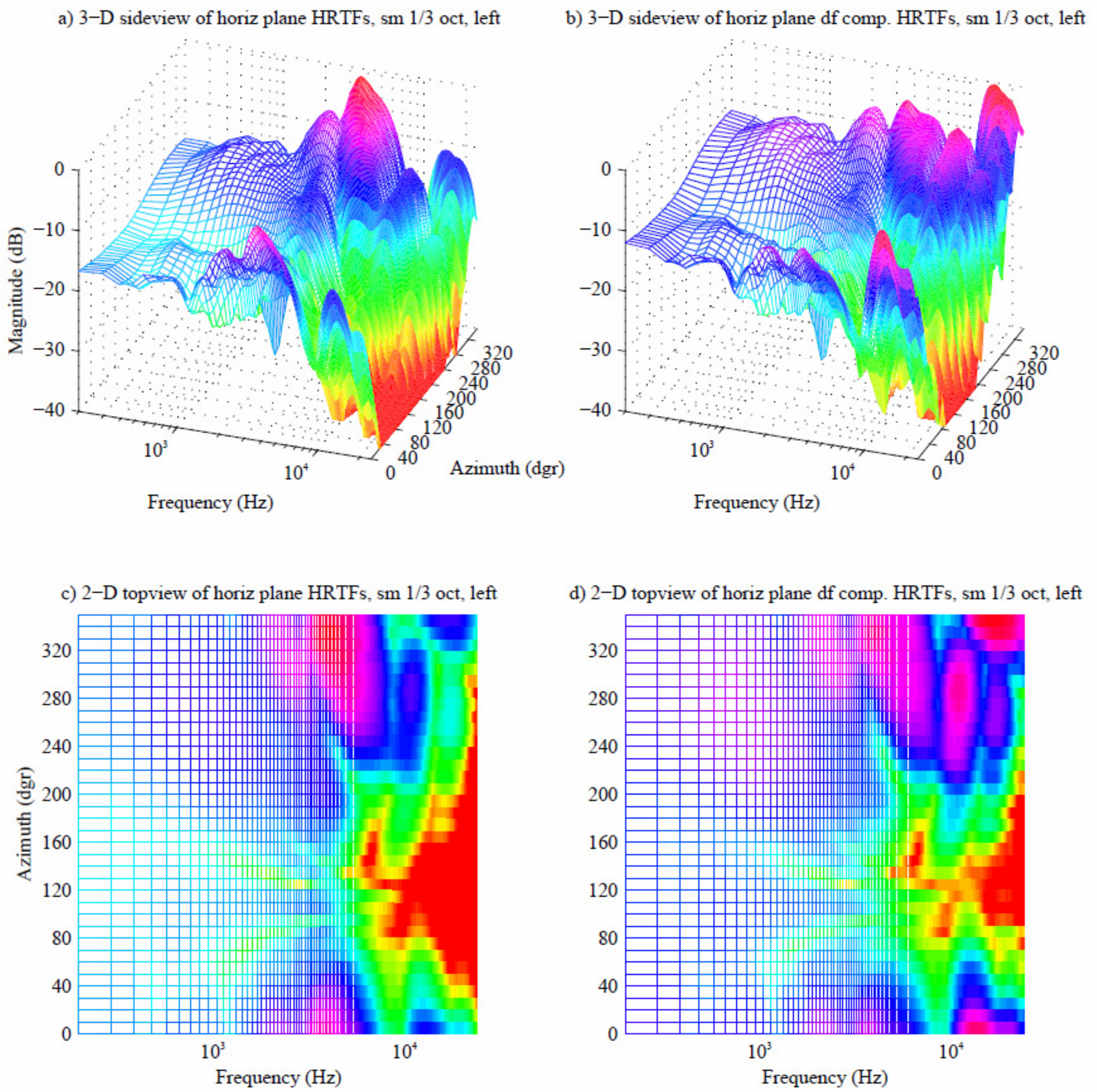

Figure 11. (a) and (b): 3D plots; (c) and (d): 2D plots of one subjects in the horizontal plane in *Riederer*'s measurements (averaged). The symmetry to the 90 degree axe is the „cone of confusion" – directions having the same ITD [34].

FUTURE WORKS

Future works include further mathematical analysis under MATLAB, first of all polar diagrams, directional characteristics plots and coordinate system transformations, mostly 3D. Furthermore, a listening test is being prepared where these HRTFs will be applied together with proper headphone equalization in order to test different environmental conditions virtually. Changes in the localization performance (whether these variations become audible or not) using naked dummy-head HRTFs versus "dressed" HRTFs will be evaluated.

SUMMARY

Different 1D, 2D and 3D visualization methods were presented for the evaluation of detailed HRTF data of a dummy-head. Monaural HRTFs and HRTF differences (HRTFDs) were calculated for different environmental conditions. The HRTFD for repeated measurements was used to test the reproducibility of the system. HRTF set of the naked dummy-head was presented in the horizontal plane and for selected elevations. The measurement data of the naked torso showed how the shadowing effect of the head and pinna reflections vary with elevation and azimuth. These presentation methods are well suited for the evaluation of HRTF and HRTFD data using different environmental settings. Figures were plotted and the GUI was programmed in MATLAB. Measured data can be filtered, modified, re-scaled and limited in magnitude if needed. This software environment together with an accurate measurement system allows the user to record high accuracy transfer functions in high spatial resolution and to create a user-friendly GUI for plotting and evaluation.

REFERENCES

1 W.M. Hartmann, "How we localize sound?" *Physics Today*, 24-29 (1999 Nov.)
2 J. Blauert, SpatialHearing (MIT Press, MA, 1983)

3 J.C. Middlebrooks and D.M. Green, “Sound localization by human listeners,” Ann. Rev. Psychol. 42, 135-159 (1991)

4 J. Blauert, “Sound Localization in Median Plane” Acustica 22, 205-213 (1969)

5 M. Morimoto and H. Aokata, “Localization cues of sound sources in the upper hemisphere” *Journal of Acous.Soc. of Japan* **E 5**, 165-173 (1984)

6 D. Hammershøi and H. Møller, “Sound transmission to and within the human ear canal” *Journal of the Acoust. Soc. America* **100(1)**, 408-427 (1996).

7 S. Mehrgart and V. Mellert, “Transformation characteristics of the external human ear” *Journal of the Acoust. Soc. America* **61(6)**, 1567-1576 (1977)

8 E.A.G. Shaw, “Transformation of sound pressure level from the free-field to the eardrum in the horizontal plane” *Journal of the Acoust. Soc. America* **56(6)**, 1848-1861 (1974)

9 H. Møller, M.F. Sorensen, D. Hammershøi and C.B. Jensen, “Head-Related Transfer Functions of human subjects” *Journal of the Audio Eng. Soc.* **43(5)**, 300-321 (1995)

10 P. Leong and S. Carlile, “Methods for spherical data analysis and visualization” *Journal of Neuroscience Methods* **80**, 191-200 (1998)

11 H. Møller, “Fundamentals of binaural technology” *Applied Acoustics* **36**, 171-218 (1992)

12 M. Kleiner, B.I. Dalenbäck and P. Svensson, “*Auralization – an overview*” *Journal of the Audio Eng. Soc.* **41(11)**, 861-875 (1993)

13 H. Møller, D. Hammershøi, C.B. Jensen and M.F. Sorensen, “Transfer Characteristics of Headphones Measured on Human Ears” *Journal of the Audio Eng. Soc*. **43(4)**, 203-216 (1995)

14 H. Møller, M.F. Sorensen, C.B. Jensen and D. Hammershøi “Binaural Technique: Do We Need Individual Recordings?” *Journal of the Audio Eng. Soc.* **44(6)**, 451-469 (1996)

15 F.E. Toole, “In-head localization of acoustic images” *Journal of the Acoust. Soc. America* **48**, 943-949 (1969)

16 N. Sakamoto, T. Gotoh and Y. Kimura, “On „out-of-head localization” in headphone listening” *Journal of the Audio Eng. Soc*. **24(9)**, 710-716 (1976)

17 IRCAM. LISTEN HRTF database http://recherche.ircam.fr/equipes/salles/listen/

18 R.L. Martin, K.I. McAnally and M.A. Senova, “Free-Field Equivalent Localization of Virtual Audio” *Journal of the Audio Eng. Soc*. **49(1-2)**, 14-22 (2001)

19 F.L. Wightman and D.J. Kistler, “Headphone Simulation of Free-Field Listening I.-II” *Journal of the Acoust. Soc. America* **85**, 858-878 (1989)

20 A.W. Bronkhorst, “Localization of real and virtual sources” *Journal of the Acoust. Soc. America* **98**, 2542-2552 (1995)

21 E.M. Wenzel, M. Arruda, D.J. Kistler and F.L. Wightman, “Localization using nonindividualized head-related transfer functions” *Journal of the Acoust. Soc. America* **94(1)**, 111-123 (1993)

22 P.F. Hoffmann and H. Møller, “Some Observations on Sensitivity to HRTF Magnitude” *Journal of the Audio Eng. Soc*. **56(11)**, 972-982 (2008)

23 E.H.A. Langendijk and A.W. Bronkhorst, “Fidelity of Three-Dimensional-Sound Reproduction Using a Virtual Auditory Display” *Journal of the Acoust. Soc. America* **107(1)**, 528-537 (2000)

24 H. Møller, D. Hammershøi, C.B. Jensen and M.F. Sorensen, “Evaluation of artificial heads in listening tests” *Journal of the Acoust. Soc. America* **47(3)**, 83-100 (1999)

25 D.J. Kistler and F.L. Wightman, “Principal Component Analysis of Head-Related Transfer Functions” *Journal of the Acoust. Soc. America* **88**, pp. 98 (1990)

26 A. Illényi and Gy. Wersényi, “Discrepancy in binaural tests and in measurements of sound field parameters” *Proc. of the International Békésy Centenary Conference on hearing and related sciences*, Budapest, 160-165 (1999)

27 C.I. Cheng and G.H. Wakefield, “Introduction to Head-Related Transfer Functions (HRTFs): Representations of HRTFs in Time, Frequency, and Space” *Journal of the Audio Eng. Soc.* **49(4)**, 231-249 (2001)

28 Gy. Wersényi and A. Illényi, “Differences in Dummy-Head HRTFs Caused by the Acoustical Environment Near the Head” *Electronic Journal of Technical Acoustics (EJTA)*, Russia, 1-15. http://ejta.org/en/wersenyi1 (2005)

29 A. Illényi and Gy. Wersényi, “Evaluation of HRTF data using the Head-Related Transfer Function Differences” *Proc. Forum Acusticum*, Budapest, 2475-2479 (2005)

30 A. Illényi and Gy. Wersényi, “Environmental Influence on the fine Structure of Dummy-head HRTFs” *Proc. Forum Acusticum*, Budapest, 2529-2534 (2005)

31 Gy. Wersényi, “Measurement system upgrading for more precise measuring of the Head-Related Transfer Functions” *Proc. Inter-Noise 2000 Vol.II.*, Nice, 1173-1176 (2000)

32 Gy. Wersényi and A. Illényi, “Test Signal Generation and Accuracy of Turntable Control in a Dummy-Head Measurement System” *Journal of the Audio Eng. Soc.* **51(3)**, 150-155 (2003)

33 Gy. Wersényi, “Spatial and spectral properties of the dummy-head during measurements in the head-shadow area based on HRTF evaluation” *Proc. Inter-Noise 2006,* Honolulu, 10 pages (2006)

34 K.A.J. Riederer, “Head-related transfer function measurements” *Master Thesis*, 1998, Helsinki University of Technology.

35 K.A.J. Riederer, “Repeatability Analysis of Head-Related Transfer Function Measurements” *105th AES convention*, preprint 4846, San Francsico, USA (1998)