

Design of a Fast Broadband Individual Head-Related Transfer Function Measurement System

Bruno Masiero, Martin Pollow, Janina Fels
Institute of Technical Acoustics, RWTH Aachen University, Germany.

Summary

In order to improve the quality of binaural based virtual reality systems, the use of individual head-related transfer functions (HRTF) plays a major role. This paper presents a system for individual HRTF measurements developed at the Institute of Technical Acoustics in Aachen. The main characteristic of this system is the close distance between head and sound sources. This allows to make use of audible near-field effects as well as to extrapolate the measurement result to different ranges with a small error.

In this set-up the common far-field assumption of a plane wave incidence on the head is no longer valid and thus great care is taken to design a loudspeaker which radiates as similar as possible to an ideal point source in a wide frequency band (200 Hz to 16 kHz). The supporting arc is constructed with thin metal sticks in a trellis structure in order to minimize disturbing reflections. Up to 40 loudspeakers can be positioned along the arc, allowing a flexible choice of the employed measurement grid. A neck support is used to assist the person under measurement to stay at rest at the same position for the entire duration of the measurement, which should take no more than ten minutes to complete.

PACS no. 43.20.Ye, 43.66.Pn

1. Introduction

For the research on binaural hearing and spatial audio reproduction the head-related transfer function (HRTF) is of fundamental importance since it describes the individual perception of sound for different angles of incidence. When hearing a binaural signal synthesized with a generic HRTF (i.e. not the listener's own HRTF) localization is degraded, especially the perception of elevation since it is strongly related to the individual geometry of the pinna. To improve even further the quality of the virtual acoustic system at the Institute of Technical Acoustics in Aachen the use of individualized HRTFs is paramount.

The main concern when measuring individual HRTFs is that the subject under measurement should not move during the whole measurement, otherwise results would be inaccurate, especially at higher frequencies. To stay still for a long period of time is a very uncomfortable (and probably impossible) task. The measurement time thus must be as short as possible.

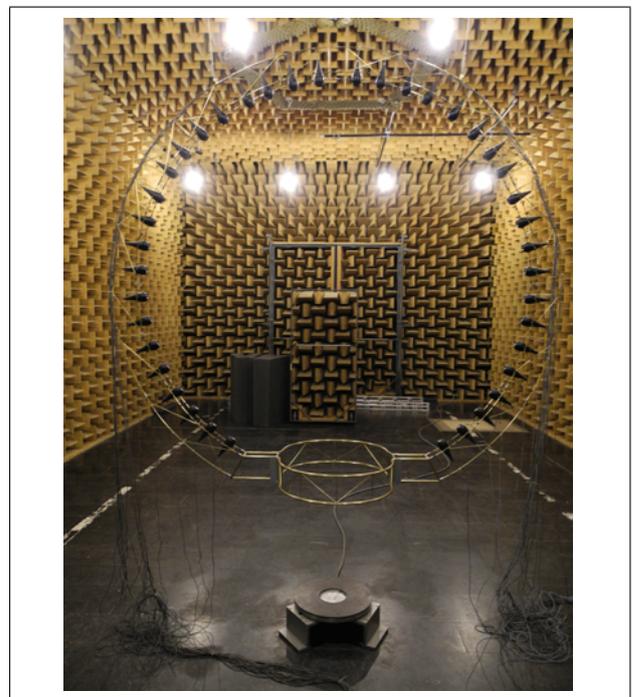


Figure 1. The new HRTF measurement system composed of supporting arc, 40 loudspeakers and a turn-table.

A method commonly used in numerical acoustics to reduce computation time is the use of reciprocity, where source and receivers are exchanged. Reciprocal HRTF measurement was proposed by Zotkin et. al. [1], where they use a miniature sound source and 32 microphones distributed in a spherical structure of 0.7m radius. This method is very fast, as the excitation signal has to be played only once for each ear, on the other hand the use of a miniature source delivers a considerably small signal-to-noise ratio (SNR) and restricts the measurement frequency range to frequencies above 1 kHz. Zotkin used a model based extension of the HRTF for frequencies below this limit, losing the individual character of the HRTF measurement. If a high spatial resolution is desired, then many microphones are needed, increasing hardware costs. For this reasons a reciprocal measurement system has been discarded.

Direct measurement systems need two miniature microphones, placed ideally in the entrance of the blocked ear canal [2], and at least one loudspeaker. The different systems can be roughly divided in three categories regarding the number of required sound sources:

- Dense array: with as many loudspeakers as directions to be measured.
- Hybrid array: with a group of loudspeakers placed at an arc, where either the arc of the subject is turned.
- Sparse array: with only one loudspeakers that is moved for every direction to be measured.

Just as with the reciprocal measurement, if high spatial resolution is desired, a dense set-up will require a large number of loudspeakers, amplifiers and digital-to-analog converters, dramatically increasing the costs of the system. To the best of our knowledge, no such system exists for HRTF measurements.

On the opposite side, a sparse measurement system uses only one loudspeaker that has to be moved to each measurement position. Even using fast positioning equipments, such a system will always be the slowest of all methods as no parallelization in the measurement procedure is possible. For example the system built at TNO in the Netherlands takes 2.5 h to measure 976 directions [3]. Subjects have to wear a head tracking device to verify that their heads are always in the correct position.

The hybrid arrays allow a good compromise between speed and hardware complexity/cost and are therefore the most commonly found. The well known CIPIC HRTF database was measured with a hybrid system, using a moving arc around the subject. They used MLS or Golay codes as excitation signal to allow parallel measurement of distinct directions, but nevertheless the measurement of 1250 points still takes around 1.5 h to complete [4]. The arc's size also varies considerably; e.g. the CIPIC's system has an arc with

1 m radius while the system build in Aalborg University has a radius of 1.95 m [5], allowing the subjects to be measured in standing position (a camera helps the subjects to stay in place).

Until recently all HRTF measurement systems used pseudo-random sequences as excitation signals to allow measurement parallelization and thus achieve a considerable speed-up. However, current state-of-the-art acoustic measurements use sine sweeps for excitation [6]. Thus Majdak et al. proposed the interleaved sweep technique to speed-up HRTF measurements with sine sweeps in their 22 loudspeaker hybrid set-up [7]. Recently, Enzner proposed a new technique for the spatially continuous measurement of HRTFs in a circle. This technique allows a very high azimuth resolution in a very fast manner; 20 s per elevation angle [8].

A very important contribution in the field of HRTF measurement is the range extrapolation technique, that describes "a way to obtain the range dependence of the HRTF from existing measurements conducted at a single range!" [9]. Because this is a holography problem, the closer to the source the measurement is made, the better the extrapolation results are expected to be.

The HRTF measurement system presented in this paper was designed with the goal of allowing fast measurement and flexible source positioning. Since a broad-band measurements system is desired, the reciprocal method was discarded. As the best compromise between cost and speed, a hybrid set-up has been chosen. The arc has a radius of 1 m and was constructed with thin metal rods in a trellis structure to minimize disturbing reflections. Near-field measurements require the sound source to radiate as similar as possible to an ideal point source and should work in a wide frequency band. To meet these criteria, a nearly omni-directional loudspeaker was designed using a small, high quality broadband driver mounted in a drop-shaped enclosure to minimize diffraction. 40 units of the drop-like loudspeaker were built. The system is completed with a turntable onto which the person under test is placed and fixed in a static position.

The following sections describe the design and construction of the loudspeaker and the supporting arc and as well as the first measurement results with the system.

2. Source

In general, to guarantee a good SNR, many other HRTF measurement set-ups use relatively large loudspeakers, sometimes even two-band systems. At close ranges these speakers will have near-field effects in their directivity pattern and two-band systems also possess an unstable acoustic center. For measurements in close ranges it is desirable that the loud-

speakers behave as a point source and no influence of neighboring loudspeakers on the radiation pattern should occur. The loudspeaker should ideally irradiate in the complete audible frequency range.

2.1. Driver

An ideal loudspeaker for a HRTF measurement set-up should meet the following design criteria:

1. Broad-band reproduction.
2. Low non-linearity.
3. Omni-directional directivity.

A loudspeaker fulfilling all these requirements is practically impossible to find. So the desired frequency range was reduced from the complete audible range to the range between 200 Hz and 16 kHz, since below 200 Hz HRTFs show very little individual variation. Still, criterion number 1 is only fulfilled by specific broad-band drivers. An initial group of seven drivers with membrane sizes varying from 20 mm to 50 mm was selected. The frequency response of each of these drivers was measured and it was verified that only three of these drivers met the frequency range specification.

These three drivers were further tested regarding their maximal sound pressure level (SPL), non-linear distortion and directivity. These measurements assisted on the choice of the best suitable driver. The driver with the highest maximal SPL and consequently lowest non-linear distortion has a bundled directivity and relatively large dimensions, being discarded for these reasons. The other two drivers show similar characteristics, with lower maximal SPL, higher non-linear distortion and directivity pattern closer to omni-directional (consequence of the smaller membrane diameter). Since the loudspeakers on the arc will be relatively close to the microphones, maximal SPL is not critical and so the smaller driver with 32 mm diameter was chosen, as it also allowed an easier fixation at the enclosure.

2.2. Enclosure

To radiate in the low frequency region the chosen driver needs an enclosure. According to the measured Thiele-Small parameters, a volume of 100 ml is enough. But even such a small enclosure can influence the radiated sound field in consequence of border diffraction effects. An optimization of the enclosure was carried out to minimize these effects.

First the driver's membrane velocity was measured with a laser vibrometer at 154 points and these values were used as input data for the loudspeaker simulation. The vibrometry results showed the presence of resonance modes on the membrane; however the phase difference on the membrane at these frequencies was relatively small.

Three forms of enclosure were simulated: a cylinder with rounded front edge, a cylinder with both front



Figure 2. Prototype of the drop-like loudspeaker.

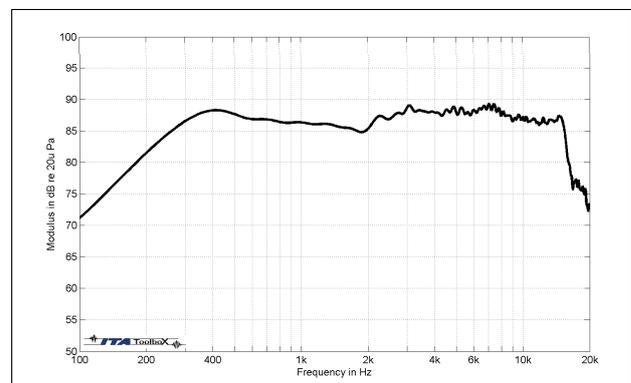


Figure 3. Frequency response of the drop-like loudspeaker measured at 1 m and 1 V.

and back edges rounded and a drop-like enclosure. All forms avoid, in varying degrees, sharp edges responsible for diffraction. Simulation results showed that the drop-like enclosure has the least influence in the loudspeakers frequency response and directivity.

Also the influence of neighboring loudspeakers should be as little as possible. To verify which form delivers the best results, a simulation was made with three identical loudspeakers placed at an arc of 1 m radius placed 10° apart. The central loudspeaker was set as the sound source while the other two loudspeakers were left inactive, as mere diffraction bodies. Again, the drop-like enclosure shows a slightly lower influence in the radiated sound field and was therefore chosen as the form for the loudspeaker's enclosure.

3. Supporting Arc

The design of the arc to hold the loudspeakers is also focused on minimizing its influence on the radiated sound field, i.e. to avoid reflection and diffraction effects as much as possible. Although easier to manufacture, bulky structures have a very high influence on the sound field and should thus be avoided. On

the other hand, a thin metal rod can be considered as acoustically transparent if its diameter is much smaller than the wavelength. The supporting arc was therefore designed with thin metal rods in a trellis structure seeking to minimize disturbing scattering effects.

For virtual reality applications the use of near-field HRTFs is highly interesting, so range extrapolation of measured HRTFs becomes of importance. From a theoretic point of view, outward range extrapolation (i.e. from near-field measurements to the far-field) should deliver a more accurate extrapolation than the inward range extrapolation. Therefore, the radius of the supporting arc was defined to be 1 m. As a person has to stay in the middle of the loudspeaker array, the use of a complete circle is not practicable, hence an arc of 300° was chosen, allowing measurements of elevation angle from -60° to 90° .

A first prototype consisting of a 30° arc section was built to verify its influence on the sound field radiated by the drop-like loudspeaker. The frequency response of the drop-like loudspeaker was measured at an anechoic chamber with and without the supporting arc. It turns out that the measured arc section produces no noticeable effects on the frequency response and directivity of the loudspeaker. Variation of less than 1 dB were observed in the frequency response and no obvious difference, that could be traced back to the supporting arc, was verified in the impulse response.

Due to its lightly built structure, the supporting arc has an under-damped oscillatory behavior. If the supporting arc were thus to be rotated, a long pause would be necessary until the arc reaches again its rest position. For this reason it was decided to keep the arc stationary and to rotate the subjects inside the arc with the help of the turn-table (already available at the institute).

4. Measurement

A first measurement was undertaken to verify how well the system works. A relay switch matrix was used to route 8 input channels into 40 output channels that were later amplified and reproduced by the drop-like loudspeakers. For that reason only 8 loudspeakers could be simultaneously used. As interleaved sweeps were used as excitation signal, this restriction increased fivefold the measurement time. Nevertheless, a measurement grid with 20 elevation angles and 48 azimuth angles (totaling 960 points) was concluded in under six minutes using log-sweeps of FFT degree 14 (0.341 s at 48 kHz sampling rate) interleaved at a 0.1 s interval.

A Gaussian sampling grid of spherical harmonics order 23 was used, with the four lowest elevation angles left missing. It is important to highlight that not all loudspeakers could be positioned at the specified elevation angle due to structural restrictions of the

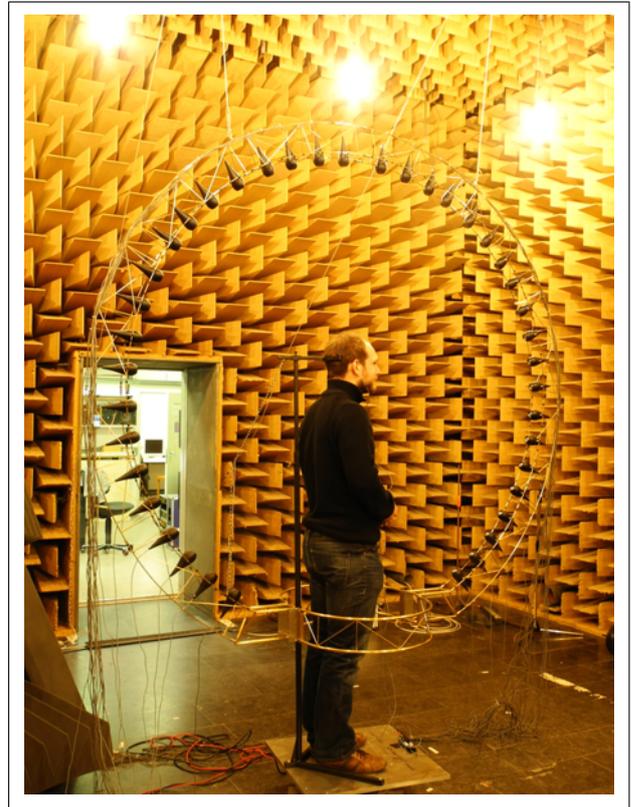


Figure 4. A subject having his HRTFs measured. Please note the headrest used to reduce movement of the subject under test.



Figure 5. Photo of the HRTF measurement set-up while measuring the HRTFs of a dummy head.

supporting arc. The real loudspeaker positions were used for processing the measured data.

5. Results

As a proof of concept the HRTFs of a dummy head were initially measured. A measurement with a dummy head offers the obvious advantage that the subject under test does not move itself and can be precisely positioned. On top of that, a comprehensive

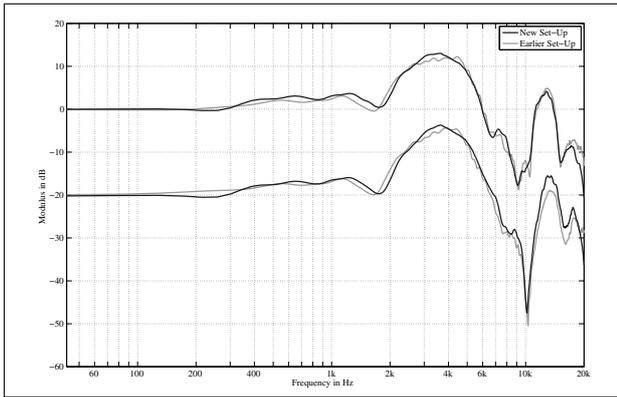


Figure 6. HRTF measured at 86° elevation and 0° azimuth. Right channel is shifted -20 dB for clarity.

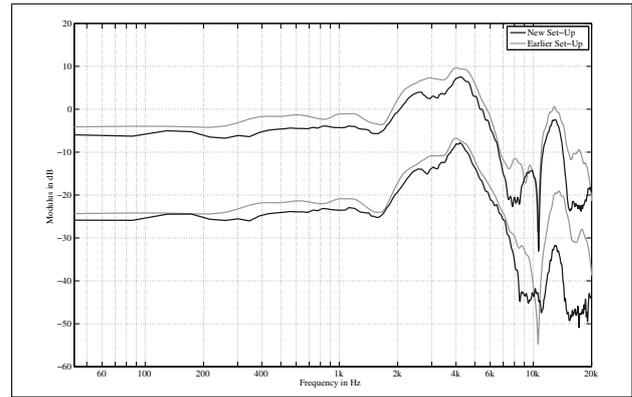


Figure 7. HRTF of frontal direction ($90^\circ, 0^\circ$). Result with the new system was interpolated and the range was extrapolated from 1 m to 2 m. Right channel is shifted -20 dB for clarity.

HRTF database of the same dummy head had already been acquired with another measurement system¹.

First two HRTFs measured at the same direction with both systems are compared. Despite the fact that the earlier HRTFs were measured at 2 m distance, the two measurements agrees very well, as can be seen in Fig. 6 for the frontal direction. Only little deviation in the higher frequencies is clearly noticeable. This shows that the measurement task, per se, and the post processing stages are adequate.

To achieve lower measurement times the new set-up also make a trade-off between speed and number of measured points. In post-processing the fewer number of measurement points should be able to be compensated by an adequate interpolation algorithm. It is therefore important to verify how well the interpolation and also the range extrapolation work.

Fig. 7-8 compare the HRTFs measured at two positions using the earlier system with the HRTFs obtained through interpolation and range extrapolation of the results obtained with the new system. Once again the curves agree well, especially for low and medium frequency. At the moment no explanation could be found why the left ear shows a better agreement then the right ear.

For higher frequencies the agreement is compromised. This occurs because the interpolation was truncated at spherical harmonics order 23 and higher frequencies usually require higher order components to be correctly described. This effect is intensified in the HRTF measurement as the source (or reciprocally the receiver) is displaced in relation to the center of the spherical coordinate system, thus requiring an even higher order for the correct description of the HRTFs at higher frequency.

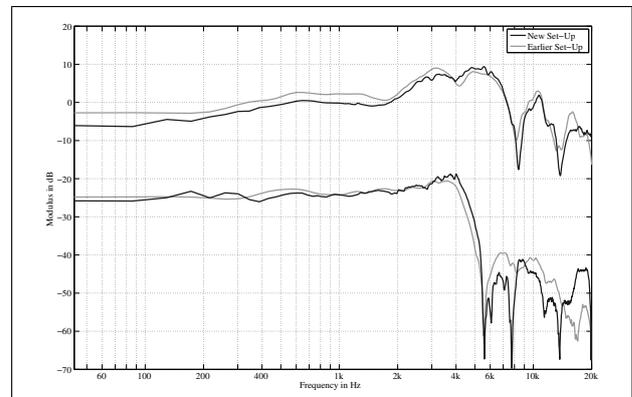


Figure 8. HRTF of left direction ($90^\circ, 90^\circ$). Result with the new system was interpolated and the range was extrapolated from 1 m to 2 m. Right channel is shifted -20 dB for clarity.

6. Conclusions

The development of a new HRTF measurement set-up was described. The new system was designed to allow a fast and correct acquisition of individual HRTFs with minimal influence of the measurement set-up itself on the measurement result. The system is composed of 40 loudspeakers, a supporting arc and a turntable with a head fixation. HRTFs are measured at 1 m range, i.e. the supporting arc has 1 m radius. For later range extrapolation it is important to have a sound source as similar as possible to an ideal point source. A loudspeaker driver with a small membrane and broad frequency range was used. The loudspeaker's enclosure was designed with the format of a drop to minimize diffraction. To minimize unwanted reflections, the supporting arc was constructed of thin metal rods in a trellis structure. The arc is kept fixed and the subject under measurement is turn inside the arc by a turn-table.

The first measurement done with the new system showed good agreement with other measurements of

¹ The earlier system was of the sparse array type and took 8 h to measure a grid with 2° by 2° resolution.

the same subject. The complete measurement was concluded in less than six minutes. If instead of eight forty audio channels were available, the measurement time could drop to around two minutes. Interpolation and range extrapolation works well for low and medium frequencies. Higher frequencies require a higher truncation limit. For future measurements a Gaussian grid of higher order shall be used intercalating the elevation angles on both halves of the supporting arc and turning the subject 360° instead of only 180°. This will obviously increase the measurement time, so new hardware to provide 40 audio output channels will be required.

Acknowledgement

We would like to thank Malte Sartor for the design of the loudspeaker and the mechanical and electrical workshop of the Institute of Technical Acoustics in Aachen for their superb work at building the 40 loudspeakers, the supporting arc and the relay switch matrix.

Bruno Masiero is a scholarship holder from the National Council of Scientific and Technological Development; CNPq - Brazil.

References

- [1] D. N. Zotkin, R. Duraiswami, E. Grassi, and N. A. Gumerov, "Fast head-related transfer function measurement via reciprocity," *The Journal of the Acoustical Society of America*, vol. 120, no. 4, p. 2202, 2006.
- [2] H. Møller, D. Hammershøi, C. B. Jensen, and M. F. Sørensen, "Transfer Characteristics of Headphones Measured on Human Ears," *Journal of the Audio Engineering Society*, vol. 43, no. 4, pp. 203–217, 1995.
- [3] A. Bronkhorst, "Localization of real and virtual sound sources," *The Journal of the Acoustical Society of America*, vol. 98, no. 5, p. 2542, 1995.
- [4] V. Algazi, R. Duda, D. Thompson, and C. Avendano, "The cipc hrtf database," in *Applications of Signal Processing to Audio and Acoustics, 2001 IEEE Workshop on the*. IEEE, 2001, pp. 99–102.
- [5] H. Møller, M. F. Sørensen, D. Hammershøi, and C. B. Jensen, "Head-Related Transfer Functions of Human Subjects," *Journal of the Audio Engineering Society*, vol. 43, no. 5, pp. 300–310, 1995.
- [6] S. Müller and P. Massarani, "Transfer-function measurement with sweeps," *Journal of the Audio Engineering Society*, vol. 49, no. 6, pp. 443–471, 2001.
- [7] P. Majdak, P. Balazs, and B. Laback, "Multiple exponential sweep method for fast measurement of head-related transfer functions," *Journal of the Audio Engineering Society*, vol. 55, no. 7/8, p. 623, 2007.
- [8] G. Enzner, "Analysis and optimal control of LMS-type adaptive filtering for continuous-azimuth acquisition of head related impulse responses," in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, vol. 2. IEEE, 2008, pp. 393–396.
- [9] R. Duraiswami, D. N. Zotkin, and N. A. Gumerov, "Interpolation and range extrapolation of HRTFs," in *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 1000. IEEE, 2004, pp. 45–48.