Measuring the Head-Related Transfer Functions of an Artificial Head with a High Directional Resolution

Bjarke P. Bovbjerg, Flemming Christensen, Pauli Minnaar, and Xiaoping Chen
Aalborg University
DK-9220 Aalborg, Denmark

Presented at
the 109th Convention
2000 September 22-25
Los Angeles, California, USA

This preprint has been reproduced from the author's advance manuscript, without editing, corrections or consideration by the Review Board. The AES takes no responsibility for the contents.

Additional preprints may be obtained by sending request and remittance to the Audio Engineering Society, 60 East 42nd St., New York, New York 10165-2520, USA.

All rights reserved. Reproduction of this preprint, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.
Measuring the head-related transfer functions of an artificial head with a high directional resolution

Bjarke P. Bovbjerg, Flemming Christensen, Pauli Minnaar, Xiaoping Chen.

Department of Acoustics, Aalborg University
Fredrik Bajers Vej 7 B4, DK-9220 Aalborg, Denmark
acoustics@acoustics.auc.dk

109th AES Convention, Los Angeles, California, USA, 22-25 September 2000

Abstract

Head-related transfer functions were measured on the artificial head VALDEMAR for nearly 12000 directions on the sphere around the head. The design and the precision of the setup are described and the results are presented. Due to the high directional resolution, contributions from the different parts of the artificial head to the head-related transfer functions are revealed.

0. Introduction

Head-related transfer functions (HRTFs) describes the transmission of sound from a point in space to the ears of a human or artificial head. The HRTFs are defined as the sound pressure measured at the blocked ear canal (P2) divided (in frequency domain) by the free-field pressure in the middle of the head, with the head absent (P1). The HRTFs changes as a function of azimuth and elevation angle, and since all directional information is captured by the HRTFs, these changes reveal the features used for sound localization. In order to allow a detailed analysis of the HRTFs, measurements were made with an unusually high directional resolution. Besides an analysis, such an HRTF database can be used in a binaural synthesis system to implement virtual sound sources. For more details about binaural synthesis see Møller [1]. Another application for a high-resolution HRTF database is to determine the minimum requirements for a binaural system. As an example the resolution needed when measuring HRTFs for binaural synthesis, was shown in Plogsties et al. [2].

The measurements were made under anechoic conditions on the artificial head VALDEMAR, developed at the Department of Acoustics, Aalborg University, Denmark. The design of VALDEMAR is described in Christensen et al. [3].

The design of the setup was done on basis of some general aims. First the setup was to introduce a minimum of reflecting surfaces. Next the setup was to be simple in design and easy to operate only because of the practical use. By this is meant that the direction definitions should be clear, and the positioning of VALDEMAR and the loudspeaker should be kept simple. Finally the whole sphere surrounding the head of VALDEMAR was to be covered by the measurement in such a way that the great arc distance, for a given elevation, between two neighboring measured directions was two degrees or less.
This preprint describes the measurement setup and its quality as well as the results that were obtained by means of the large measurement session of HRTFs.

1. Measurement setup

The measurements were done in an anechoic room and VALDEMAR was setup horizontally in the center of a half-circular arc. The principle of the setup is illustrated in Figure 1.

![Figure 1: VALDEMAR and loudspeaker in the measurement setup.](image)

VALDEMAR was placed horizontally and mounted on a wooden plate and then connected to a turning device, which made it possible to turn VALDEMAR to an arbitrary position. VALDEMAR was wearing a thin pullover to reduce the body reflection from the hard plastic. VALDEMAR has two ½” microphones from G.R.A.S. Sound and Vibration Aps. type 40AD placed 0.004 m into the ear canal. In Figure 2 the final setup for VALDEMAR is seen.

![Figure 2: VALDEMAR mounted on the turning device on the stand.](image)
An arc of 1” squares steel was put up in the anechoic room by wires and two supporting poles. The arc was centered around VALDEMAR in order to minimize variations in the measurement radius. Since the arc was not perfectly circular shaped, the resulting measurement radius varied between 2.04 m – 2.13 m. A ball loudspeaker with a 6½” 2017 unit from VIFA was mounted onto a sliding device that was connected to the arc. The sliding device offered the possibility of fastening the loudspeaker at an arbitrary position on the arc. To suppress mechanically transmitted sound, vibration dampers were used both at the mounting points of the arc and the connections between the loudspeaker and the arc.

In the setup an azimuth and elevation coordinate system was used where azimuth was defined by the turning of VALDEMAR and the elevation defined by the loudspeaker position. This is illustrated in Figure 1 by the two arrows.

The turning device had a plastic disc mounted onto it, which had 0.9 mm holes in steps of two degrees in an outer ring and in steps of three degrees in an inner ring. Laser pointers were modified and mounted to point through the holes. In this way the turning and positioning of VALDEMAR could be controlled visually. The device is seen in Figure 3 where the laser is shooting through one of the holes.

![Figure 3: The turning device that controlled the azimuth orientation of VALDEMAR.](image)

Another laser was mounted on the loudspeaker to control that the loudspeaker was pointing towards the interaural axis of VALDEMAR. Three other lasers were used to ensure the repositioning of VALDEMAR. One was marking the top of the head, one marking the middle of the microphone diaphragms in the ears and one marking the stand. The two first lasers were also used to define the point for the P1 measurements in the middle of the head with VALDEMAR absent.
In Figure 4 the complete setup including both VALDEMAR and the arc is shown. Both the arc, VALDEMAR’s stand and the stands for the lasers were wrapped with a tight fastened acoustic damping material surrounded by a loose wrapped in order to slightly increase the impedance from the arc to the air and avoid reflections.

![Figure 4: The complete setup in the anechoic room.](image)

The setup in the anechoic room was controlled from an adjacent room where the acquisition equipment was placed. This consisted of a standard PC and a multi-channel sound card from Gadget Labs with a break out box. A t.c electronic Gold Channel delivered the phantom power to the G.R.A.S. microphones and also served as a preamplifier. A Pioneer A656 amplifier was used to drive the ball loudspeaker. On the computer a newly developed Maximum Length Sequence (MLS) measuring system by Olesen et al. [4] was installed. The MLS system made it possible to operate the measurements from the anechoic room by means of a remote terminal, which prevented the operator from having to leave the anechoic room in between each measurement.

The new MLS measurement system offers the advantages of having both pre-averages to suppress correlated noise and averages with different MLS sequences to suppress non-linearities in the measured transfer functions. See [4] for further explanation. To determine the MLS order and number of pre-averages and averages to use a number of test measurements were made. A reasonable compromise between measuring time and quality was to use a 14th order sequence and test for different averages and pre-averages. In addition, the sound pressure level (SPL) at the P1 point produced by the loudspeaker was tested together with the averages to find the most optimal setting for the measurement equipment.
In Figure 5 three measured and normalized binaural impulse responses (BIR) are shown all measured with a 14th order MLS including four times of averages and pre-averages. The top BIR has the highest SPL, and here it is possible to distinguish the non-linearities from the noisefloor (see arrows) hence the signal to noise ratio (SNR) is around 60 dB. In the middle figure the level has been turned down and here the non-linearities and the noisefloor are approx. equal with at SNR of 70 dB. If then the level is reduced even further the noisefloor increases and the SNR decreases to around 55 dB. The best compromise is chosen, which is the middle BIR, and at that level the free-field SPL generated by the loudspeaker is measured to be 57 dB linear weighted. This may be a rather surprising result that indicates that one should search for the best compromise between SPL and SNR instead of pushing the loudspeaker to its limit.

![Figure 5: Normalised binaural impulse response measured at different loudspeaker sound pressure level.](image)

The equipment for a P2 measurement is shown in Figure 6.

![Figure 6: The equipment setup for P2 measurements.](image)
The setup for the P1 measurement was similar to the P2 setup except that VALDEMAR was replaced by one of the G.R.A.S. microphones. To reduce the reflections from a regular microphone stand the \( \frac{1}{2} \)" microphone, for the P1 measurement, was connected to a vertical mounted \( \frac{1}{2} \)“ pole with the length of 1.7 m.

### 2. Measurement procedure

#### Directions measured

The measurement directions were organized in order to discretize the elevation by a two degree step size, and then, in a given elevation plane, to keep the azimuth arc distance between two neighboring measurements less than or equal to two degrees. With this as a starting point the measuring resolution in the different elevation planes was calculated by:

\[
\Delta \phi = 2 \sin^{-1} \left( \frac{\sin \left( \frac{\theta}{2} \right)}{\cos \theta} \right)
\]

\( \theta \) is the elevation angle which is defined as steps of two degrees from \(-90\) degrees (bottom) to \(+90\) degrees (top). \( \Delta \phi \) is the needed angular resolution in order to assure that the great arc distance between the two neighboring directions was less than or equal to two degrees. The azimuth resolution was rounded towards zero to the nearest integer, which was dividable by the possible measuring resolutions of two or three degrees. One exception was made with a five degree resolution, which was represented on the turning device by switching between the two and three degree step size. The resulting measuring resolutions are seen in Table 1, where only the upper half sphere is represented.

<table>
<thead>
<tr>
<th>( \theta )</th>
<th>( \Delta \phi )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 0^\circ - 48^\circ )</td>
<td>( 2^\circ )</td>
</tr>
<tr>
<td>( 50^\circ - 66^\circ )</td>
<td>( 3^\circ )</td>
</tr>
<tr>
<td>( 68^\circ - 70^\circ )</td>
<td>( 5^\circ )</td>
</tr>
<tr>
<td>( 72^\circ - 76^\circ )</td>
<td>( 6^\circ )</td>
</tr>
<tr>
<td>( 78^\circ )</td>
<td>( 9^\circ )</td>
</tr>
<tr>
<td>( 80^\circ - 82^\circ )</td>
<td>( 10^\circ )</td>
</tr>
<tr>
<td>( 84^\circ - 86^\circ )</td>
<td>( 18^\circ )</td>
</tr>
<tr>
<td>( 88^\circ )</td>
<td>( 45^\circ )</td>
</tr>
<tr>
<td>( 90^\circ )</td>
<td>( 360^\circ )</td>
</tr>
</tbody>
</table>

**Table 1:** Calculated azimuth resolutions with respect to a given elevation.
One more restriction that was added to the covering of the sphere was that the three often-analyzed planes - median, horizontal and frontal – had to be measured. This restriction led to the exclusion of some resolutions (e.g. four degrees, where the azimuth directions of ±90 are not represented). With these calculated directions the cover of the sphere met the specified restrictions and the coverage of a 1/8 sphere is shown in Figure 7.

![Figure 7: Grid of measured directions on a 1/8 sphere.](image)

**Procedure**

Since the measured directions were organized in elevation planes with various azimuth resolution the measurements were organized in the same way. This means that the procedure was to measure from the bottom through all elevation planes towards the top, each time completing an azimuth rotation at the given resolution. For each elevation step the loudspeaker was repositioned.

After finishing the P2 measurements a series of P1 measurements were made. In the setup 91 different loudspeaker positions occurred and for each position a P1 measurement was made to allow a proper extraction of the HRTF. The axial direction of the microphone was perpendicular to the arc, assuring the same frequency characteristics at the microphone for all directions. These measurements included the arrival time differences arising from the variation in the arc radius, thus eliminating these in the calculated HRTFs.

The measurements were carried out by six people, and split up into time slots of two hours which was time enough to complete one two degree elevation plane. The complete measurement session lasted for three weeks.
3. Quality

Directional resolution inaccuracy

In order to align VALDEMAR to the setup and thereby ensure that the measurements on the sphere correspond to the decided directions, a test measurement was done. In this VALDEMAR was positioned by the laser marks and the frontal part of the median plane was measured in 4 degrees resolution. Here the difference in delay seen in the impulse responses between left and right ear measurements should be zero. This delay was iteratively adjusted to 1/10 of a sample at the sampling frequency of 48 kHz by moving the arc up or down. This corresponds to a 9 mm horizontal deviation with respect to the measurement radius of approx. 2.10 m.

Replacement of VALDEMAR

Previous HRTF measurements have shown to be sensitive to the placement of the artificial head or person and if it is not placed at the same position, the HRTF deviates in the way that dips in the frequency response can move and small magnitude differences are seen as well. To keep track of this, a repeatability test was done. In total, 16 random directions were chosen and repeated three times, while VALDEMAR was removed and replaced. By the use of laser marks the replacement of VALDEMAR was very easy, which also was the outcome of this test. The general observed deviation was around $\pm 2$ dB and mainly concentrated at the high frequencies.

Reflections from the setup

One of the philosophies of this setup was to have a “clean” acoustical environment. By this is meant that, in spite of the anechoic room the setup was placed in, the more equipment the setup consists of, the more reflections, resonances and diffractions will influence the measurements. Therefore, the setup was set for P1 measurements because here it was obvious that only a loudspeaker response should be measured. Normally, one single measurement reveals inaccuracies, but in order to track reflections and resonances a series of four degrees P1 measurements were made.
This gave 46 measurements in total, which are seen in Figure 8 where all minimum phase impulse responses are plotted. The figure is structured in the way that the X-axis is the loudspeakers position on the arc and the Y-axis is the time in milliseconds. Notice the zoom on the dB scale, which range from $-35$ dB to $-55$ dB.

![Figure 8: Impulse responses of the P1 test measurements. The grayscale is the level in dB. Here the reflections from surfaces are clearly seen. Arrow 1 and 2 are the supporting poles at $\pm 90$ degrees elevation and arrow 3 is one of the laser stands. Arrow 4 is the biggest reflection and is caused by the operator.](image)

Some of the important improvements are the microphone stand, the wires and poles for the arc and that all possible reflecting surfaces were wrapped with the two layers of acoustic damping material. The remaining reflections are unavoidable since two of them (marked as (1) and (2)) are the supporting poles at $\pm 90$ degrees. (3) is one of the laser stands and (4) is the operation person, which is seen to be the largest reflecting object in the setup. The P1 test measurement shows a very clean loudspeaker response within the first 3 ms, which is the time window to be used when extracting the HRFTs.

**Transducers**

Quality evaluation of microphones relies on two parameters, the calibration tones and noisefloor measurements. The equipment was calibrated using a 1 kHz tone at 94 dB SPL. The noisefloor
was estimated by replacing the loudspeaker by an 8 Ohm resistor and then make a measurement with the microphones in the free-field. The practical obtainable SNR is around 55 – 65 dB. The calibration tones were also used to check that the sensitivity of the microphones did not change at the end of the three week measuring period compared to the beginning, which was not the case since only deviations of ±0.05 dB occurred. The frequency characteristic of the loudspeaker is showed in Figure 9.

![Frequency Characteristic of Ball Loudspeaker](image)

**Figure 9:** The frequency characteristic of the ball loudspeaker.

**Temperature control**

Another important parameter, which was investigated, was temperature influences. It was seen that the temperature could fluctuate up to 6 degrees, due to the lights and the presence of a person in the anechoic room. This influence was actually revealed in the first test measurements and was seen as changes in the arrival time due to changes in the speed of sound. In order to prevent this the lights were kept on during the night and thereby the temperature was stabilized at 24 degrees Celsius.
**P2 check measurements**

Another set of control measurements was made to check that the system gain did not change during the measurements. These control measurements included all elevation angles at azimuth equal to zero, which is the frontal part of the median plane. These directions were spread over the three weeks period and could be compared to the initial control measurements in order to confirm that all equipment was working properly. For each completion of an elevation plane, the comparison was made. If no errors were found the next elevation plane was measured.

**P1 check measurements**

A P1 control measurement was included in the measurement schedule. This consisted of measurements of all the 91 loudspeaker positions on the arc before and after the three weeks measurement period. This control measurement was made to reveal any changes in the physical setup as well as changes in microphone or loudspeaker characteristics. No difference between the two measurement series was found.

**4. Results**

The results of the measurement session are 11975 HRTFs measured on the sphere around the artificial head and torso VALDEMAR. The derivations of the HRTFs are as described in the introduction where the P2 measurements are divided by the P1 measurement in the frequency domain from the same loudspeaker position. The results are plotted as HRTFs in the frequency domain for the horizontal- frontal and median plane and only for the left ear due to the accurate symmetry. This makes it possible to evaluate the structures and compare to other measurements in the same planes, made in Christensen et al. [5].

In Figure 10 the HRTFs in the horizontal plane are shown. In the figure the x-axis is the frequency and the y-axis shows the azimuth direction starting at 0 degrees and then moving to –90 degrees, which is the contralateral side and in this case the right side. After crossing the back at 180 degrees the ipsilateral side occurs at 90 degrees and finally back to 0 degrees azimuth again. The structure of an amplification from 3 – 10 kHz interfered by a somewhat complex pattern of dips, combined with a generally lower level on the contralateral side is seen. The air volume in concha causes the amplification, whereas the dip structures origins from shoulder and pinna reflections. The reflections are seen as comb filter effects where the close reflections in pinna dominates at high frequencies and the late reflections from the shoulder mostly influences frequencies below 10 kHz, but exists at higher frequencies as well.

The median plane is shown in Figure 11. Here the y-axis is the elevation and starts below VALDEMAR at –90 degrees. Then moving up in the back and crossing the top at 90 degrees, where after the direction is moving downwards in the front and ends below VALDEMAR in –90 degrees again. Here the general amplification from 3 – 10 kHz is seen again but interfered by a different dip structure. The comb filter effect from the shoulder reflection is clearer than in the horizontal plane and at higher frequencies the influences of the pinna are again revealed as the dominating dips.
Figure 10: The magnitude of HRTFs in the horizontal plane.

Figure 11: The magnitude of the HRTFs in the median plane.
The results of the measurements in the frontal plane are shown in Figure 12. The elevation on the y-axis is defined so it starts at the bottom and crosses the contralateral side first followed by the top and then the ipsilateral side to end at the bottom again. The magnitude in the frontal plane shows a similar structure as the median plane but because the frontal plane has a contralateral side the level and dip structure is more complex from –90 to 45 degrees.

![Figure 12: The magnitude on the HRTFs in the frontal plane.](image)

The HRTFs are shown here for the three planes because they are the most common ones to evaluate, but in principle it is possible to extract an arbitrary plane or curve from the measured HRTF database. This high resolution set of HRFT measurements makes it possible analyze and follow a specific feature, such as the shoulder reflection, influences on the HRTFs throughout the sphere, or it could be one single possible cue, such as a specific dip. Due to the large amount of data it is not possible to illustrate such effects in this preprint.

**Interaural time difference**

Since the measurements cover all directions on the sphere, a complete analysis of the Interaural Time Difference (ITD) around the head was done. The ITD is determined as the interaural group delay difference of the excess phase evaluated at 0 Hz. This was shown by Minnaar et al. [6] to be the correct ITD to be used with minimum phase HRTFs. The ITD is plotted as points on a sphere, grouped in intervals of 50 µs. The sphere plot was split up in different views in order to evaluate each of them separately. The different views are left, top, bottom, front and back.
ITD results are only shown for the left half sphere, since the measurements are symmetrical within 6 µs.

In Figure 13 the ITD sphere is seen from the left side, where the x-axis is the bottom and the y-axis is the front directions. Here it is seen that the ITD intervals of 50 µs almost form concentric circles except from the maximum, which is moved up and forward. Also a small bend in the curves is seen below at the back.

**Figure 13:** The ITD sphere seen from the left side.
In Figure 14 the viewpoint has changed and the figure includes both the front and the back views. Here the circles from Figure 13 are, due to the changed viewpoint, seen as vertical lines. Especially for the front view the ITD intervals are parallel until ±45 degrees elevation are reached. From here they tend to bend upwards, which corresponds to the fact that the maximum ITD is found around 10 degrees elevation. On the back view the same pattern is seen but here the parallel ITD intervals bends towards the median plane in the bottom.

**Figure 14:** The ITD sphere seen both as front and back view.

The last two viewpoints are seen in Figure 15, which are the top and bottom views. Here the x-axis is the back and the y-axis is to the left. On the top view vertical lines are seen, and when moving towards the left side they bend towards left. This shows again that the maximum is moved up whereas the curves from Figure 13 are seen more from the top. The bottom view reveals vertical lines again bending towards the bottom.
5. Summary

The HRTFs of the artificial head VALDEMAR were measured in an anechoic room. The artificial head was mounted horizontally and could be rotated to define azimuth angle. A single loudspeaker was moved on a half-circular arc to determine the elevation. The measurements were made with a directional resolution of two degrees or better resulting in 11975 HRTF measurements.

The set-up, procedures and control measurements, used to ensure the high quality of the HRTFs, was described. A marking system of 6 lasers enabled the accurate placement of the artificial head and the loudspeaker. The set-up was wrapped in acoustic damping material to suppress reflections. An MLS measurement system that suppresses external noise in the measured transfer functions was used. It was found important to maintain a constant temperature in the anechoic room during the three weeks measuring period.

Finally, some of the results were presented. The magnitudes of the HRTFs were shown in the horizontal, median and frontal planes. The changes in the magnitude as a function of both azimuth and elevation angle can be seen in exceptional detail. Especially features as the shoulder- and pinna reflections are clearly seen. It was found that the ITD of VALDEMAR is left/right symmetrical within 6 µs and has its maximum on the side of the head about 10° above the horizontal plane. Furthermore, it was seen that the ITD is generally larger in the lower half-plane to the back than in front of the head.

Figure 15: The ITD sphere seen both as top and bottom view.
6. Acknowledgements

Thanks to the laboratory technician Claus Vestergaard for his help during the design of the setup, and the support with the technical equipment. Bila A/S delivered the arc and is hereby given thanks for that. The measurement group consisted of the authors, Jan Plogsties and Søren Krarup Olesen, who put a huge effort into the measuring of the HRFTs during the three weeks.

7. References


