Binaural Sky - Examination of Different Array Topologies

K. Laumann[1,2], G. Theile, H. Fastl[2]
[1] *Institut für Rundfunktechnik GmbH, Germany, eMail:* laumann@irt.de
[2] *AG Technische Akustik, MMK, Technische Universität München, Germany*

Introduction

Binaural technology, a widely used body of methods for capturing and creating spatial acoustical scenes, requires an exact reproduction of ear input signals. Naturally, the first choice for playback are headphones, but particularly in the context of virtual or augmented reality a reproduction free of headphones is often requested. Loudspeaker systems designed to equal transmission characteristics of headphones are generally termed *virtual headphones*[1].
Following [1], the fundamentals will be explained on a simplified example: We assume a common stereo setup with two loudspeakers in free field conditions. In this case four transfer functions can be specified, from each source to the ipsi- and the contra-lateral ear.

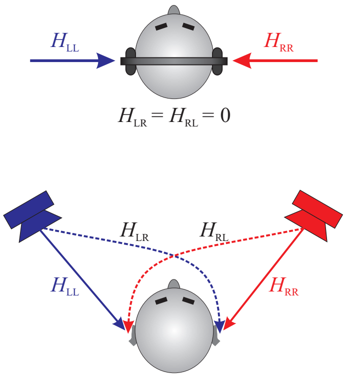

Figure 1: Transfer functions in case of headphone (upper part) and loudspeaker (lower part) reproduction.

Headphones on the other side are characterized by a high channel separation: idealized the direct transfer functions H_{LL}, H_{RR} are equal to one without crosstalk to the opposite ear, thus H_{LR}, H_{RL} equaling zero (see Fig. 1).
If the input channels are processed by an appropriate network, the correct reproduction of ear input signals on loudspeakers is possible. In this simplified case the speaker signals y_{left}, y_{right} can be calculated from the two input channels x_{left}, x_{right} as follows — x,y being specified in the frequency domain:

$$\begin{pmatrix} y_{\text{left}} \\ y_{\text{right}} \end{pmatrix} = H^{-1} \begin{pmatrix} x_{\text{left}} \\ x_{\text{right}} \end{pmatrix}$$

The matrix H^{-1} is the inverse of matrix H containing all four transfer functions as specified in Fig. 1:

$$H = \begin{bmatrix} H_{LL} & H_{RL} \\ H_{LR} & H_{RR} \end{bmatrix}$$

[1]Note: In consumer electronics headphones simulating surround sound speakers are also called virtual headphones.

One major restriction is identified in praxis: The transfer functions in H change if the listener turns the head or moves. The network H^{-1} is no longer valid, and therefore an adaptive solution is necessary.
A concept for virtual headphones, named *Binaural Sky*, was proposed, proofed [2, 3], and further developed [4], which — instead of recalculating the network H^{-1} — relies on a circular array of speakers mounted above the listener rendering focussed sources based on the principles of Wave Field Synthesis. These focussed sources are moved simultaneously with the listener so that they keep a constant position in the head-related coordinate system. The matrix H is then also constant (in the ideal case), and so the realization of an adaptive network H^{-1} is avoided. Fig. 2 visualizes this idea.

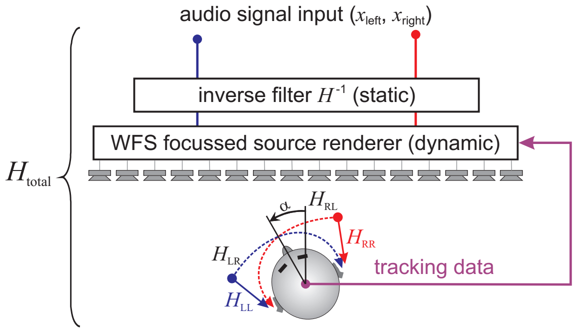

Figure 2: Schematic of a concept for virtual headphones; a dynamic Wave Field Synthesis renderer places focussed sources at a constant position in the head-related coordinate system. The input signals are pre-filtered by a static network.

Experiment

The prototype based on the concept mentioned above (see Fig. 2) was designed with a precisely defined layout [2, 3, 4]. A circular configuration was chosen for the reason of symmetry to allow a robust compensation of head rotations for a single listener.
It must be considered that the theory of Wave Field Synthesis is derived for a half space defined by an infinite plane, or an enclosed volume, bounded by a continuous distribution of an infinite number of secondary sources. All implementations can be approximations only and thus will produce artifacts. Furthermore, focussed sources are not covered by the fundamental theory, but an extension based on reciprocity (time reversal principle), and require special attention to ensure causality [5].
These, very basic, considerations lead to the assumption that the reproduction quality and stability of virtual headphones based on the proposed concept is influenced by several design parameters due to possible artifacts.

The experiment outlined in the following was planned to estimate geometrical restrictions for the loudspeaker array. Bearing possible applications, e. g. combination with displays, in mind, three **configurations** with 24 speakers (ELAC 301) in each case, and two listener positions at 0.6 m and 1.2 m distance, were tested:

1. one "regular" WFS line array with 24 speakers
2. two *vertical* lines with 12 speakers each, in front of the listener to the left and right
3. two lines with 12 speakers each, on either side *above* the listener

According to the proposed concept four focussed sources were placed close to the head (15 cm to the center) and symmetrical with respect to the median plane. This configuration results from a further experiment, for details see [6].
The matrix of transfer functions H between focussed sources and and the two ears was measured with a dummy head (Neumann KU 100). H is not a square matrix in this case, so the pseudo-inverse H^+ must be calculated. For stability reasons, the transfer functions in H were smoothed with a sliding third-octave band window, and the resulting network H^+ was bandpassed between 120 Hz and 14 kHz.
To check for the compensation of head rotations, four different **scenarios** were evaluated: *First*, the head orientation was straight forward and the virtual headphones initialized; *second*, the head was turned 30° to the right without any reaction of the system; *third*, focussed sources were repositioned according to the proposed concept; *fourth*, H^+ was recalculated for the new head orientation.

Results

An objective measure for the quality of virtual headphones are the resulting ear input signals, or more precisely the four transfer functions between the two input channels x_{left}, x_{right} and the two ear input signals of the listener. As stated in the introduction, the transmission path of the two direct channels should be frequency independent and cross-talk should be eliminated.
Measurements of all three loudspeaker configurations were taken with the KU 100 dummy head in an anechoic chamber, at both listener positions and covering all four scenarios. Fig. 3 exemplarily shows data of the measurement on the regular line array at 0.6 m distance, and the following general results can be stated:

- for *scenario 1* (head orientation straight forward) the results are similarly good, independent on the configuration of and the distance to the array: the amplitude response of the direct path is nearly constant between 200 Hz and 12 kHz and crosstalk is attenuated by 20 dB and more
- for *scenario 2* (head orientation to the right), the quality degrades — as expected — if the head is turned to the right and the system does not react

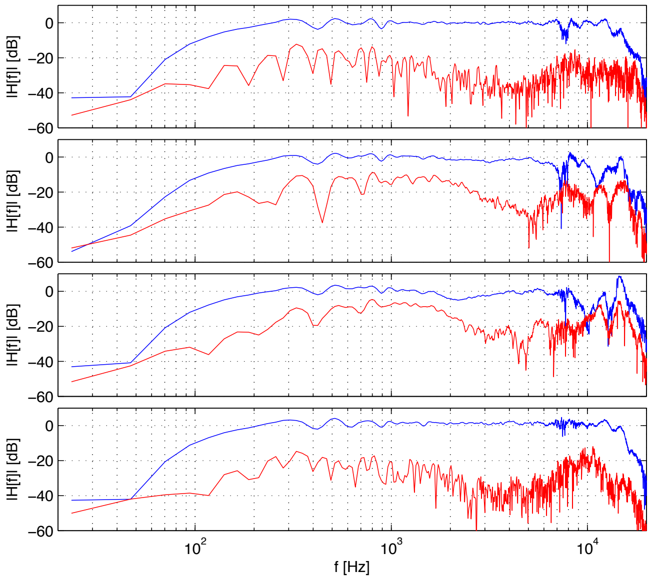

Figure 3: Amplitude responses H_{total} of virtual headphones between left input channel, and left (blue) and right (red) ear; measurements were taken of a 24 channel WFS line array at 0.6 m distance.
From top to bottom: experiment scenarios 1-4 as outlined in the section *Experiment*.

- for *scenario 3* (head orientation to the right), — surprisingly — no improvement was found if the focussed sources are repositioned; the results are in some cases even worse
- for *scenario 4*, if the head is orientated to the right and the network H^+ is recalculated, the results are nearly as good as for *scenario 1*

Conclusion

Two main statements can be deducted from the results of the experiment: First, the quality of virtual headphones based on Wave Field Synthesis according to the concept proposed in [2, 3] is independent on the configuration of and the distance to the loudspeaker array. However, second, the system is less stable with respect to head rotations as expected. Repositioning of the focussed sources in a way that they remain at a fixed position in the head-related coordinate system lead to no improvement of the measurement results in the studied cases.

Acknowledgments

H. Petritsch (Institut für elektronische Musik und Akustik, Graz, Austria) and *M. Meier* (Institut für Rundfunktechnik, Munich, Germany) have contributed major parts of work. In addition to the co-authors *G. Theile* and *H. Fastl* the following persons are acknowledged for helpful discussions and comments: *F. Völk*, *D. Menzel*, *H. Wittek*, *S. Goossens*, *M. Link*

References

[1] J. Bauck, D. H. Cooper, "Generalized Transaural Stereo and Applications", *J. Audio Eng. Soc.*, vol. 44,

no. 9, pp. 638-705 (1996)

[2] D. Menzel, H. Wittek, G. Theile, H. Fastl, "The Binaural Sky: A Virtual Headphone for Binaural Room Synthesis", *International Tonmeister Symposium*, Hohenkammer, Germany (2005)

[3] D. Menzel, H. Wittek, H. Fastl, G. Theile, "Binaurale Raumsynthese mittels Wellenfeldsynthese - Realisierung und Evaluierung", *Tagungsband Fortschritte der Akustik - DAGA 2006*, Braunschweig, Germany, pp. 255–256 (2006)

[4] R. Franke, G. Theile, K. Laumann, D. Leckschat, "Berücksichtigung der Kopfposition bei der Erzeugung eines virtuellen Kopfhörers mittels Wellenfeldsynthese", *Tagungsband Fortschritte der Akustik - DAGA 2008*, Dresden, Germany, pp. 895–896 (2008)

[5] J. Ahrens, S. Spors, "Notes on Rendering Focused Directional Virtual Sound Sources in Wave Field Synthesis", *Tagungsband Fortschritte der Akustik - DAGA 2008*, Dresden, Germany (2008)

[6] T. Weissgerber, K. Laumann, G. Theile, H. Fastl, "Headphone Reproduction via Loudspeakers using Inverse HRTF Filters", *Proceedings NAG/DAGA 2009*, Rotterdam, The Netherlands (2009)