MULTICHANNEL SOUND REPRODUCTION SYSTEM FOR BINAURAL SIGNALS – THE AMBISONIC APPROACH

Stefan Leitner, Alois Sontacchi, Robert Höldrich

Institute of Electronic Music and Acoustics
University of Music and Dramatic Arts in Graz
`www.kug.ac.at/iem`

ABSTRACT

Convincing sound reproduction via headphones requires filtering of virtual sound sources with head related transfer functions (HRTF). HRTFs describe signal differences at the two ear drums in level, time and frequency weighting [1,2]. Another psychoacoustic phenomenon of the human auditory system concerning hearing in natural sound fields is the improvement of source localization if small head movements are possible [3]. In this paper a computational efficient implementation to get dynamic binaural signals (in dependence on the head position) is proposed. This realtime modell is based on Ambisonic sound reproduction [4,5].

1. INTRODUCTION

To obtain auditory localization cues from head movements, binaural headphone signals have to be updated in realtime according to head the position (figure 1). HRTFs are only available in form of impulse respones at certain discrete directions, so the problem of impulse response interpolation arises which is difficult to handle in realtime.
The proposed approach has the advantage of constant filters. Instead, the filter input signals are controlled by head movements. As a result, the interpolation of impulse responses can be avoided.

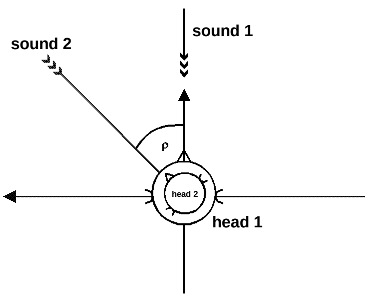

Figure 1. *The subject's head movement results in different HRTF filters for a static sound field (sound 1, sound 2). This figure demonstrates that two static sound sources demand different filtering for head position one and two after rotation.*

2. AMBISONIC

2.1. Definitions and Equations

The aim of Ambisonic [7-10] is to reproduce the soundfield of a plane wave, (1) and figure 2, with an array of loudspeakers.

$$S_\psi = P_\psi e^{ikr\cos(\phi-\psi)} \tag{1}$$

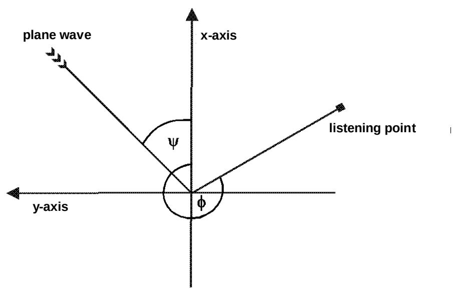

Figure 2. Ambisonic axis and angle notation.

The original plane wave (= reference wave) according to figure 2 arrives from direction ψ, the listening point is situated at (r,ϕ).The number of transmission channels for a 2D-system is (2m+1) and $(m+1)^2$ for a 3D-system where m specifies the system order. This mathematical model compares a reference plane wave with the wave produced by the ambisonic system.
Regarding to equation (1) a soundfield can be written as a series expansion of cosinus and bessel funcions.

$$S_\psi = P_\psi\left(J_0(kr) + \sum_{m=1}^{\infty} 2i^m J_m(kr)[\cos(m\psi)\cos(m\phi) + \sin(m\psi)\sin(m\phi)]\right) \tag{2}$$

This soundfield has to be approximated by a loudspeaker array. The output of the n-th speaker (plane wave radiation presumed) is

$$S_n = P_n\left(J_0(kr) + \sum_{m=1}^{\infty} 2i^m J_m(kr)[\cos(m\phi)\cos(m\phi_n) + \sin(m\phi)\sin(m\phi_n)]\right) \tag{3}$$

Superposing all loudspeaker signals, the array field results in

$$S = \sum_{n=1}^{N} P_n J_0(kr) + \sum_{m=1}^{\infty} 2i^m J_m(kr)\left(\sum_{n=1}^{N} P_n \cos(m\phi_n)\cos(m\phi) + \sum_{m=1}^{N} P_n \sin(m\phi_n)\sin(m\phi)\right) \tag{4}$$

A comparison of equation (2) and (4) supplies matching conditions (5,6,7) for N loudspeaker feeds.

$$P_\psi = \sum_{n=1}^{N} P_n \tag{5}$$

$$P_\psi \cos(m\psi) = \sum_{n=1}^{N} P_n \cos(m\phi_n) \tag{6}$$

$$P_\psi \sin(m\psi) = \sum_{n=1}^{N} P_n \sin(m\phi_n) \tag{7}$$

A limited number of transmission channels and loudspeakers means that the series expansion ends at a certain order m. Equation system (5,6,7) provides the so-called Ambisonic signals (spherical harmonics) for the encoding of the original soundfield. A first order 2-dimensional system consists of the Ambisonic signals W, X and Y.

$$W = \frac{P_\psi}{\sqrt{2}} \tag{8}$$

$$\begin{aligned} X &= P_\psi \cos(\psi) \\ Y &= P_\psi \sin(\psi) \end{aligned} \tag{9}$$

Solving Equation (5,6,7) for the loudspeaker feeds P_n with the angle ϕ_n of the n[th] speaker, one obtains Equation (10) for a symmetric array like in figure (3). This results in an approximation of the original sound field near the origin.

$$P_n = \frac{1}{N}\left(W + 2X\cos\phi_n + 2Y\sin\phi_n\right) \tag{10}$$

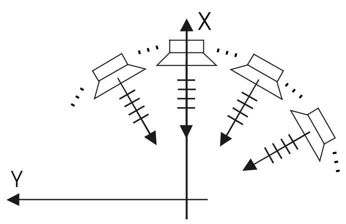

Figure 3. *Reproduction with a symmetric loudspeaker layout.*

For a 2D - second order system, we need two additional ambisonic signals U,V:

$$\begin{aligned} U &= P_\psi \cos(2\psi) \\ V &= P_\psi \sin(2\psi) \end{aligned} \tag{11}$$

The loudspeaker feeds are

$$P_n = \frac{1}{N}\left(W + 2X\cos\phi_n + 2Y\sin\phi_n + 2U\cos 2\phi_n + 2V\sin 2\phi_n\right) \tag{12}$$

2.2. Computation of a dynamic binaural stereo signal

The loudspeaker signals (10) or (12) are used as virtual mono sources for spatialisation via headphones. Each loudspeaker signal has to be convolved with a HRTF corresponding to its direction ϕ_n:

$$\begin{pmatrix} L \\ R \end{pmatrix} = \begin{pmatrix} H_{1,L} & H_{2.L} & \cdots & H_{N,L} \\ H_{1,R} & H_{2,R} & \cdots & H_{N,R} \end{pmatrix} * \begin{pmatrix} P_1 \\ P_2 \\ \vdots \\ P_N \end{pmatrix} \tag{13}$$

For a second order 2D system we obtain the Ambisonic signals W,X,Y,U and V and as a result the loudspeaker feeds (14).

$$\begin{pmatrix} P_1 \\ P_2 \\ \vdots \\ P_N \end{pmatrix} = \frac{1}{N}\begin{pmatrix} \alpha_1 & 2\beta\cos\phi_1 & 2\gamma\sin\phi_1 & 2\delta\cos 2\phi_1 & 2\varepsilon\sin 2\phi_1 \\ \alpha_2 & 2\beta\cos\phi_2 & 2\gamma\sin\phi_2 & 2\delta\cos 2\phi_2 & 2\varepsilon\sin 2\phi_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \alpha_N & 2\beta\cos\phi_N & 2\gamma\sin\phi_N & 2\delta\cos 2\phi_N & 2\varepsilon\sin 2\phi_N \end{pmatrix} \cdot \begin{pmatrix} W \\ X \\ Y \\ U \\ V \end{pmatrix} \tag{14}$$

$\alpha = \beta = \chi = \delta = \varepsilon = 1$ for symmetric loudspeaker layouts.

2.3. Soundfield Rotation

The great benefit of our approach is that soundfield rotation with respect to head movements can be performed in a very easy way using a rotation matrix (15). Consequently spatial interpolation is less complex at the ambisonic stage than at HRTF stage.

$$\begin{pmatrix} W' \\ X' \\ Y' \\ U' \\ V' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \cos\rho & -\sin\rho & 0 & 0 \\ 0 & \sin\rho & \cos\rho & 0 & 0 \\ 0 & 0 & 0 & \cos 2\rho & -\sin 2\rho \\ 0 & 0 & 0 & \sin 2\rho & \cos 2\rho \end{pmatrix} \cdot \begin{pmatrix} W \\ X \\ Y \\ U \\ V \end{pmatrix} \tag{15}$$

To summarise the above points, we calculate a binaural headphone signal as follows:

$$\begin{pmatrix} L_{Bin} \\ R_{Bin} \end{pmatrix} = \begin{pmatrix} \text{constant} \\ \text{filters} \end{pmatrix} * \begin{pmatrix} \text{rotation} \\ \text{matrix} \end{pmatrix} \cdot \begin{pmatrix} \text{ambisonic} \\ \text{signals} \end{pmatrix} \tag{16}$$

3. IMPLEMENTATION

The rotated Ambisonic signals are the input signals for 256 point FIR filters representing the KEMAR HRTFs [6].Computation of the constant filter coefficients was performed with MATLAB 5.1. A third order realtime system was implemented in MAX/FTS on a SGI/O2. The system block diagram is shown in figure 4.

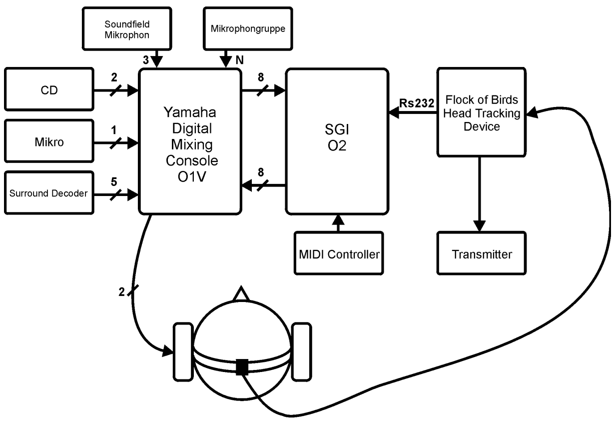

Figure 4. *Block diagram of a system for generating dynamic binaural signals. Interactivity is provided via head position.*

4. REFERENCES

[1] Gilkey, Robert H., Anderson Timothy R.: "Binaural and Spatial Hearing in Real and Virtual Environments", Lawrence Erlbaum Associates, Publishers; Mahwah, New Jersey, 1997.

[2] Blauert, Jens: “Spatial Hearing, The Psychophysics of Human Sound Localization”, The MIT Press, Revised Edition, 1997.

[3] Mackensen P., Reichenauer K., Theile G.: “Einfluß der spontanen Kopfdrehungen auf die Lokalisation beim binauralen Hören”, Tagungsband, 20.Tonmeistertagung, 1998.

[4] Bamford, J.S.: “An Analysis of Ambisonic Sound Systems of First and Second Order”, Thesis presented to the University of Waterloo, Waterloo, Ontario, Canada, 1995.

[5] Malham David G., Myatt Anthony: “3-D Sound Spatialization using Ambisonic Techniques”, Computer Music Journal, 19:4, pp.58-70, 1995 Massachusetts Institute of Technology, Winter 1995.

[6] Gardner B., Martin K.: “HRTF Measurements of a KEMAR Dummy-Head_Microphone”, http://sound.media.mit.edu/KEMAR.html, 1994.

[7] Nicol, R., Emerit M.: “3D-Sound Reproduction Over An Extensive Listening Area, A Hybrid Method Derived From Holophony And Ambisonic”, AES 16th International Conference, pp.436-453, 1999.

[8] Nicol, R., Emerit M.: “Reproducing 3D-Sound for Videoconferencing: a Comparison between Holophony and Ambisonic”, DAFX98 Proceedings pp.17-20, 1998.

[9] Gerzon Michael A.: “The Design of Precisley Coincident Microphone Arrays for Stereo and Surround Sound”, Preprint No.20, 50th Convention of the Audio Engineering Society, London, March 1975.

[10] Farrah K.: “The Soundfield Microphone”; Wireless World, pp.99-103, November 1979.