3D sound field rendering under non-idealized loudspeaker arrangements

Acoustics'08 Paris
June 29 - July 4, 2008
euronoise

Alois Sontacchi, Franz Zotter and Robert Höldrich

contact [sontacchi@iem.at]

Institute of Electronic Music and Acoustics, Inffeldgasse 10/3, 8010 Graz, Austria
University of Music and Dramatic Arts, Graz, Austria
http://iem.at

Abstract

The approach to realise periphonic sound field reproduction based on spherical base solutions of the wave equation has already been well-known as Ambisonics and High Order Ambisonics (HOA), respectively (see [1]). By the aid of an N-dimensional orthogonal set of vectors any arbitrary source free sound field can be described. Reproduction is realized by projection of the encoded sound field on a regular loudspeaker distribution over a spherical surface. The used set of vectors exhibits a defined hierarchic with interesting symmetries. In the original scheme sound sources represented by plane waves (sources in far distance) can be encoded independent of the decoding process on the regular loudspeaker layout. Usually, in practice - in contrast to theory, 3D loudspeaker layouts are requested for the upper hemisphere.

This restriction is caused by the geometry of many rooms. First of all that restriction bounds the reproduction of sound sources to the upper area. With a complete set of N-th order spherical harmonics in such discussed cases a regular design of a HOA decoder is impossible.

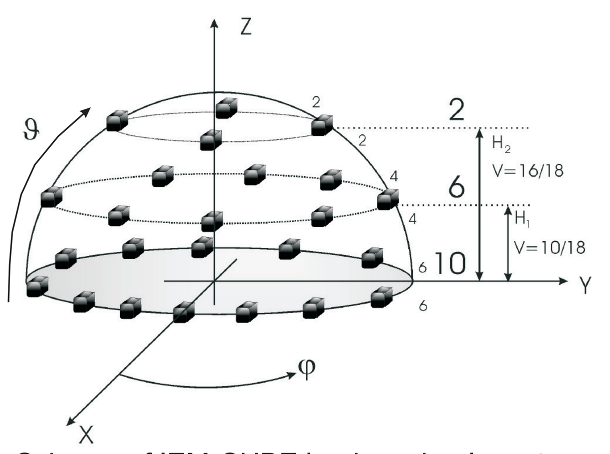

Figure 1: Schema of IEM-CUBE loudspeaker layout.

Goal

Within this case study we will show how the symmetries of the spherical harmonics can be used to obtain optimized decoding rules and to overcome insufficient irregular loudspeaker arrangements in regard of incompletely sampled spheres. Based on an existing loudspeaker arrangement (IEM-CUBE, see Fig.1) for the upper hemisphere we will highlight that a reduced set of the spherical harmonics can cope with these conditions.

Experimental Arrangements

Basically, the loudspeaker configuration could be determined by using the vertices of Platonic solids. However, this yields a strongly restricted number of possible configurations. Other layouts can be found empirically and by numerical optimization (cf. [3]).
To provide numerical analysis, and comparison by listening the given real loudspeaker layout in the upper hemisphere (IEM-CUBE, see Fig.1) was expanded by various virtual arrangements in the lower hemisphere.

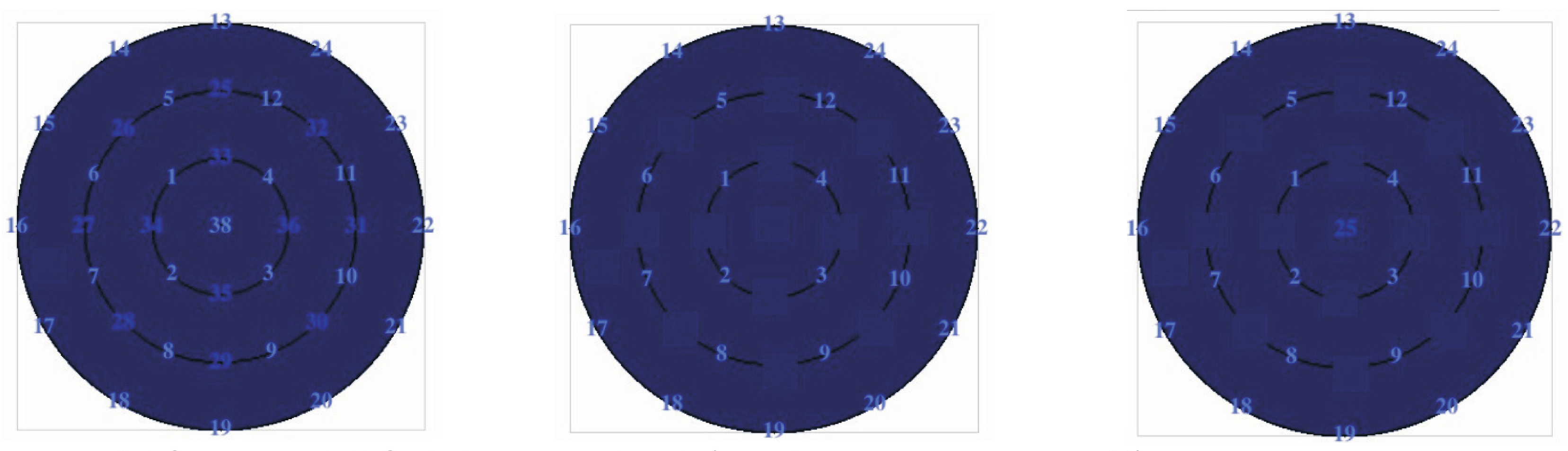

Figure 3: left: expanded IEM-CUBE loudspeaker layouts (dark blue numbers, and number 38)
middle: original layout used for reduced SH set. And right: additiv loudspeaker to improve reduced SH-Set decoding.

The optimized layout regarding objective evaluation criteria (see below) and informal listening tests is compared to the direct application of the reduced set of spherical harmonics. 3D audio rendering will be restricted in both cases for virtual sound sources within the upper hemisphere. Furthermore it should be shown that even the amount of real loudspeakers *N* is much lower than the required number defined by L_{SH} the system provides reliable and satisfactory results. The theoretical limit of required real loudspeakers is now bounded by a much lower number $L_{red,SH}$.

Background Information

HOA is a 3D audio rendering technique, which is based on spherical harmonics. Beside the compact notation it provides the great advantage to separate encoding and decoding. Therefore the same encoding can be used for many different speaker layouts. In principle a sound source is assumed to be located far away from the ideal listing position and hence it can be modelled by a plane wave.
The representation ***B*** of a plane wave emitting the source signal *S* form a direction given by the source angle ϕ (azimuth) and θ (elevation) can be written as the signal S times a vector $\mathbf{Y}_{n,m}(\phi,\theta)$ consisting of the spherical harmonics evaluated at the specified angles (Eq.1).

$$\mathbf{B} = S \cdot \mathbf{Y}_{n,m}(\phi,\theta) \qquad (1)$$

Within the 3D audio rendering procedure N loudspeakers arranged uniformly over the surface of a sphere will try to reproduce the corresponding sound field. Each loudspeaker is regarded to be placed far away from the ideal listing position, too. Therefore the resulting representation (Eq.2) can be written as a weighted sum of the Ambisonics representation of each loudspeaker *i* at the position specified by the angles (ϕ_i,θ_i)

$$\tilde{\mathbf{B}} = \sum_{i=1}^{N} p_i \cdot \mathbf{Y}_{n,m}(\phi_i,\theta_i) = [\mathbf{Y}_{n,m}(\phi_1,\theta_1), \mathbf{Y}_{n,m}(\phi_2,\theta_2), \cdots \mathbf{Y}_{n,m}(\phi_N,\theta_N)] \cdot \mathbf{p} \qquad (2)$$

To provide perfect reconstruction the individual loudspeaker feeds p_i have to be adjusted to meet Eq. 3.

$$\tilde{\mathbf{B}} \cong \mathbf{B} \qquad (3)$$

This can be achieved (Eq. 4) by an inversion of the loudspeaker matrix in Eq. 2. If a perfect uniform symmetric loudspeaker arrangement exists the inversion can be realized by simply transpose the matrix[1]. In general cases where $N \geq L_{SH}$ (whereby $L_{SH} = n^2 + 2n + 1$ is the number of used spherical harmonics, related to used order n) the inversion is obtained by calculating the pseudo inverse.

$$\mathbf{p} = S \cdot inverse\{[\mathbf{Y}_{n,m}(\phi_1,\theta_1), \mathbf{Y}_{n,m}(\phi_2,\theta_2), \cdots \mathbf{Y}_{n,m}(\phi_N,\theta_N)]\} \cdot \mathbf{Y}_{n,m}(\phi,\theta) \qquad (4)$$

If the loudspeaker arrangement deviate severe form symmetrical layouts the (pseudo) inversion might lead to imperfect solution, where perfect reconstruction cannot be guaranteed over the whole rendering surface.

[1] *in 2D: equal spacing angles, in 3D: platonic solids* ($N \geq L_{SH}$)

Results

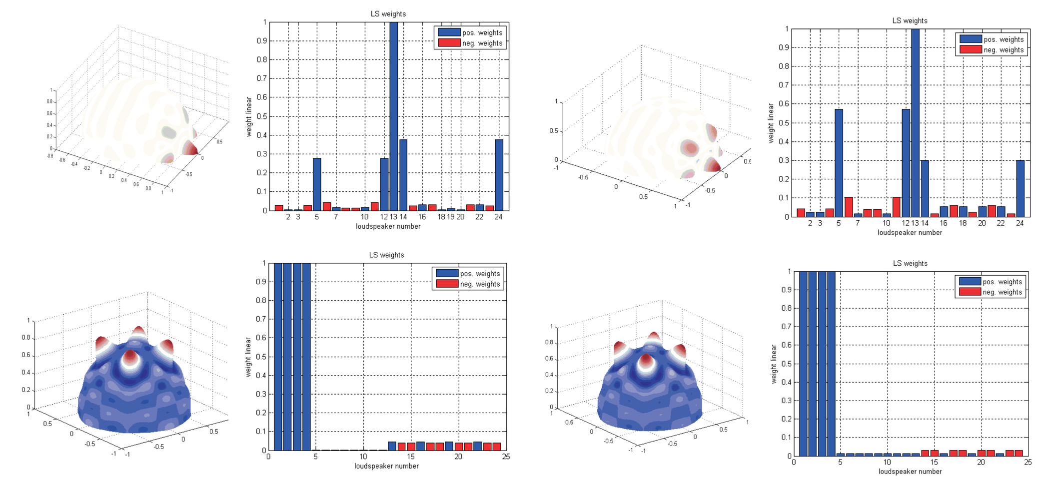

Figure 4: loudspeaker weights, on the left: IEM-CUBE expanded (LSH) & on the right: orig. Layout (both top performing, see Table 1)
top: ideal source direction, bottom: non-ideal source direction

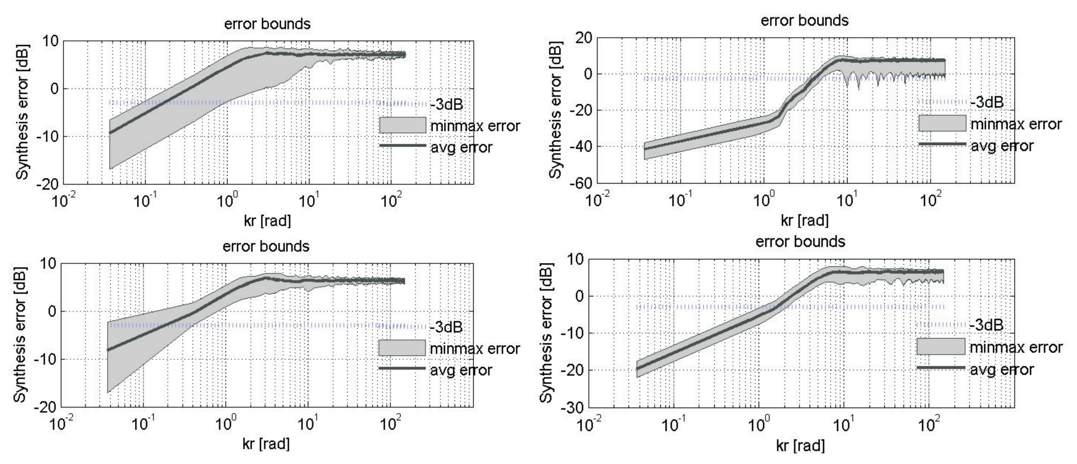

Figure 5: D2D-Error, on the left: IEM-CUBE expanded (LSH) & on the right: orig. Layout top: without weights, bottom: with weights (right plus virt. loudspeaker)

Global overall quality description for low and high frequencies:

The quantity LF_{var} (see Eq. (5)) measures the amplitude variability of various sampled virtual source directions at the ideal listing point. Therefore for each rendered source direction the loudspeaker feeds are summed and the ratio of the maxima and the minima is given. This measure is a quality criterion for low frequencies.
In case of the HF_{var} (see Eq. (6)) the squared loudspeaker feeds are summed. This measure is related to the reproduced energy and is a quality criterion for high frequencies.

$$LF_{var} = 20 * \log 10 \left(\frac{\max(\sum_{i=1}^{N} p_i)}{\min(\sum_{i=1}^{N} p_i)} \right) \qquad (5)$$

$$HF_{var} = 10 * \log 10 \left(\frac{\max(\sum_{i=1}^{N} p_i^2)}{\min(\sum_{i=1}^{N} p_i^2)} \right) \qquad (6)$$

Table 1

	weights	LF in [dB]	HF in [dB]
L_{SH}	[1 1 1 1 1]	0.3	5.7
L_{SH}	[1 .94 .78 .56 .32 .126]	0.8	1.8
$L_{red,SH}$	[1 1 1 1 1]	0	4.9
$L_{red,SH}$	[1 .94 .78 .56 .32 .126]	0	3.2
$L_{red,SH}$	[1 .94 .78 .56 .32 .126] + virt. LS	0	2.1

Assumptions

Based on the investigations of Zhiyun Li in [2] concerning semi-spherical microphone arrays it is evident, that the contribution of spherical harmonics showing asymmetric relations concerning the z-axis cannot be controlled within (upper) hemisphere system arrangements. Looking to the table of spherical harmonics approx. half of the components are affected (see Fig.2).

Up to order n, $\frac{(n-1)\cdot n}{2}$ components can be discarded, therefore only $L_{red,SH} = \frac{n^2}{2} + n + 1$ components remain.

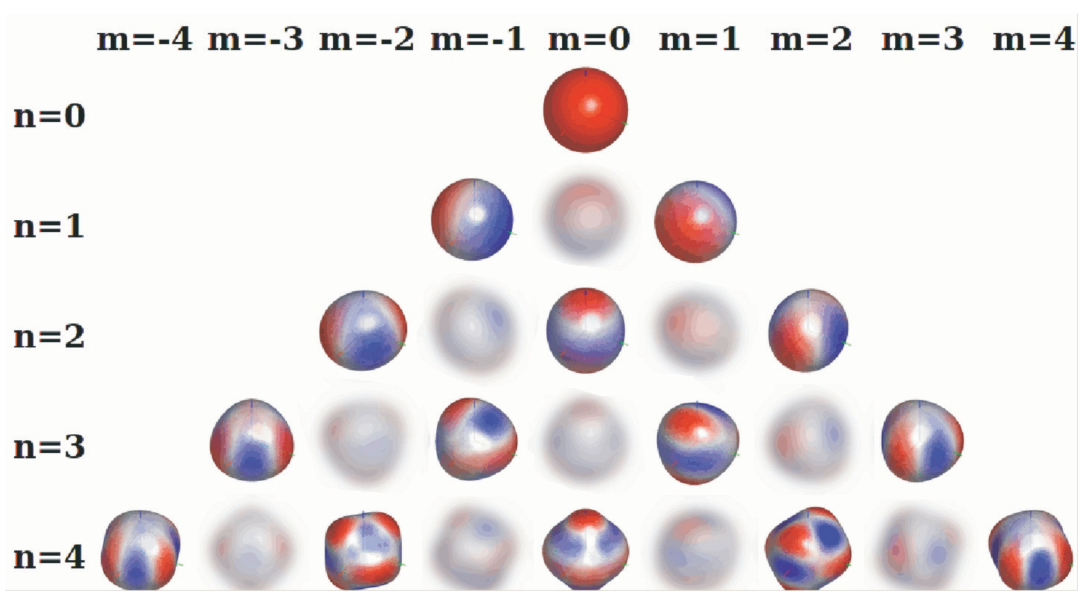

Figure 2: Set of z-symmetrical spherical harmonics ($LS_{red,SH}$)

As a direct consequence of the reduced set of spherical harmonics the inversion of the loudspeaker matrix in Eq. 4 for arrangements in the upper hemisphere is much more robust against singularities, due to improperly chosen loudspeaker layouts.

Summary & Outlook

Within this case study concerning loudspeaker arrangements in the upper hemisphere it has been shown that the usage of a reduced spherical harmonics set will lead to minimize the ambiguities and to provide linear independence, respectively. Therefore the complicated process of establishing 3D layouts, the higher computational load and requested double amount of intermediate representation channels can be circumvent.

Furthermore this case study motivates to discover further symmetry relations within the spherical harmonics to provide specific reduced representations for loudspeaker arrangements, which are reduced to a specific rendering area e.g. in 3D the segment of a sphere (c.f. [8]) or in 2D the ITU arrangement for surround sound (c.f. [9]).

Acknowledgement
We would like to thank Thomas Musil for the evaluation of various empirical loudspeaker layouts and the implementation and informal listening sessions of promising decoders in our laboratories at the IEM in Graz.

References

[1] M. A. Gerzon, "Periphony: With-Height Sound Reproduction", presented at the 2nd AES Conv. Munich,Germany, 1972, J. Audio Eng. Soc.1973.
[2] Z. Li and R.Duraiswami, "Hemispherical Microphone Arrays for Sound Capture and Beamforming", WASPAA, 2005.
[3] F. Hollerweger, "Periphonic Sound Spatialization in Multi-User Virtual Environments", Diploma thesis, Institute of Electronic Music and Acoustics, 2006.
[4] M.A. Gerzon, "General Metatheory of Auditory Localisation",92nd AES Convention, Preprint No. 3306, March, Vienna, 1992.
[5] J.S. Bamford, „An Analysis of Ambisonic Sound Systems of First and Second Order", Thesis presented to the University of Waterloo, Waterloo, Ontario, Canada; 1995.
[6] J. Daniel, J.B. Rault, and J.D.Polack, „Acoustic Properties and Perceptive Implications of Stereophonic Phenomena", in Proc. AES 16th Int. Conf. On Spatial Sound Reproduction, pp. 91-102, 1999.
[7] M.A. Poletti, „A unified theory of Horizontal Holographic Sound Systems", J. Audio Eng. Soc., Vol. 48, No. 12, December, 2000.
[8] A. Farina, et al., "Employing WFS (Wave Field Synthesis) and HOA (High Order Ambisonics), ICSV14, 9-12, July, 2007, Cairns, Australia.
[9] M. Poletti, "Robust Two-Dimensional Surround Sound Reproduction for Nonuniform Loudspeaker Layouts" J. Audio Eng. Soc., Vol. 55, No. 7/8, 2007 July/August.