INVESTIGATIONS ON DISTANCE CODING IN 3D SOUND FIELDS

Alois Sontacchi, Robert Höldrich

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

alois.sontacchi@kug.ac.at, robert.hoeldrich@kug.ac.at

ABSTRACT

This investigation proposes a possibility to synthesise a true 3D sound field over loudspeakers. A new approach concerning the distance coding is presented. We tried to combine the benefits both using the Wave Field Synthesis (WFS) approach and Higher Order Ambisonics (HOA). Therefore the proposed system can be divided into two main parts. Firstly the determination of the driving functions of the sound sources using a derivative of the WFS approach. Secondly the coding for transmission and/or storage whereby the scheme is based on the Ambisonics approach using higher orders.

The paper is organised in four sections. The first section gives a brief introduction about the WFS and the HOA approaches. In the second section the derivation of the driving functions is presented and the coding scheme of the derived source signals is explained. Results are given in section three. Finally the paper is concluded and further possible research directions are identified.

1. INTRODUCTION

In the following a brief introduction about Wave Field Synthesis and Higher Order Ambisonics is given. Further extended information can be found concerning WFS in [1], [2], [3] and in the case of HOA in [3], [4], [5].

1.1. Wave Field Synthesis

The WFS approach is based on the Huygens Principle that is mathematically described by the Kirchhoff-Helmholtz-Integral (see Eq. 1). The Kirchhoff-Helmholtz-Integral implies that the wave field of a source free volume V can be described by the knowledge of the pressure along the enclosure surface S and the gradient of the pressure normal to the surface S.

$$P(\vec{r}_R) = \frac{1}{4\pi}\oint_S \left[P(\vec{r}_S)\cdot\nabla_S G(\vec{r}_R|\vec{r}_S) - G(\vec{r}_R|\vec{r}_S)\cdot\nabla_S P(\vec{r}_S)\right]\cdot\vec{n}dS \qquad (1)$$

whereby $G(\vec{r}_R|\vec{r}_S)=\frac{e^{-jk|\vec{r}_R-\vec{r}_S|}}{|\vec{r}_R-\vec{r}_S|}$ is known as the Green's function.

Therefore each arbitrary sound field inside a source free volume can be reproduced with distributed monopole and/or dipole sources along the surrounding surface. This leads to a technique called „holographic audio“ [1] or also known as „holophone systems“. In the case of our investigation we use a related approach. The feeds of distributed loudspeakers are filtered in order to obtain the specified sound field in a defined area. The calculation procedure is presented in section 2.

1.2. Higher Order Ambisonics

In general the Ambisonic approach is presented with the well-known B-format with the signals W, X, Y and Z. The system approach is based on the sound field's spatial decomposition in spherical harmonics of 0^{th} and 1^{st} order. This approach can be extended to higher order systems (HOA) [5], [6], resulting in better localisation properties and a wider listening area. However increasing the system order will also increase the required transmission channels and also the amount of necessary loudspeakers. In the 3D case, the mathematical description of the arriving planar wave is extended in a series leading to an infinite sum of cosine and sine weighted Legendre-functions, leading to the real spherical harmonics. Using higher order signals will reduce the reconstruction errors and also energy spread over the loudspeakers. There are compromises necessary to overcome the problem of finite system orders, and even new complex microphone characteristics are required in case of recording real sound fields.

1.3. Motivation and System Design

The basic idea is to build up a system based on the advantages of HOA (e.g. channel coding, independence of the sound space and the location of the real loudspeakers) and WFS (possibility to mime a realistic transmission path). Using the curvature of the direct arriving sound field we try to mime the distance perception of a virtual source So the system design is motivated by coding the curvature for single sources with a derivative of the WFS approach and rendering the whole sound field using the HOA approach. The resulting system can therefore be divided in two parts (see figure 1). Each single source $S_i(\vec{p})$ is defined by a signal or sound file, and a time dependent position $\vec{p}(t) = p(\varphi(t), \vartheta(t), r(t))$ (azimuth, elevation and distance) related to the system origin, which is identical to the ideal listening position.

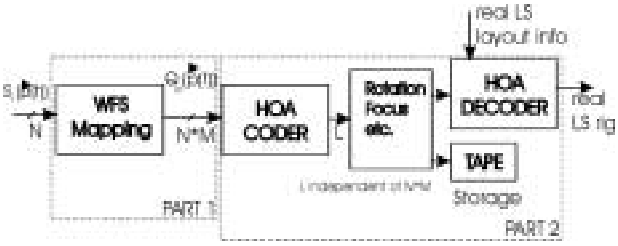


Figure 1: *System model.*

In the first part each source signal is mapped under defined conditions to virtual loudspeaker feeds $Q_l(\tilde{\vec{p}})$. These loudspeakers are located at a defined fixed distance r_0 which is equal to the average distance of the real loudspeaker positions to the origin (calibrated system).

In the second part these loudspeaker feeds are coded to the HOA domain according to their position $\tilde{\vec{p}}$ and rotated to their defined virtual position $\vec{p}$. In the HOA domain all those sound fields are summed up and different sound field manipulations can be performed easily (e.g. acoustic focusing). Afterwards the Ambisonics signals are decoded to the real loudspeaker rig by a decoding matrix (see e.g. [5] or [6]).

2. DISTANCE CODING SCHEME

2.1. Description

To recreate the sound field curvature of a virtual point source with a unique loudspeaker the two position (neglecting the radiation properties of a real loudspeaker) must be identical. Therefore if the positions differ, the curvatures of those two fields won't match in any proper defined area.

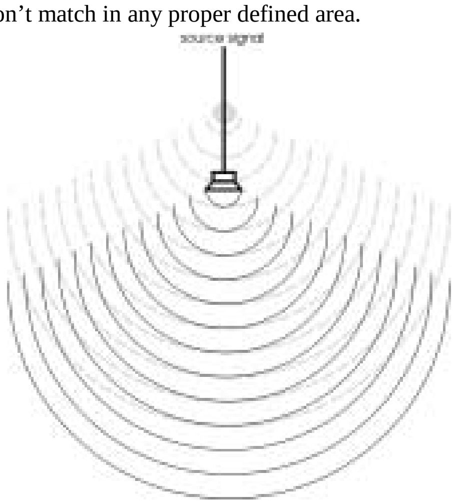


Figure 2: *Different curvature caused by different positions.*

Using an arbitrary finite amount of discrete distributed loudspeakers (according to the Huygens' principle) we will be able to synthesise a specified sound field within a defined area. Within the defined area the reference field P_{REF} (the field caused by the virtual source) and the system field P_{SYS} (the superposed field produced by the loudspeakers) are compared. In the following we consider the 2 dimensional case, reducing additional mathematically complexity. We assume that our virtual loudspeakers are positioned along a sector around the ideal listening position at a defined fixed distance r_0. The apex angle between the virtual loudspeaker and the number of used loudspeakers can be chosen arbitrary, but should be related to the amount of real speakers.

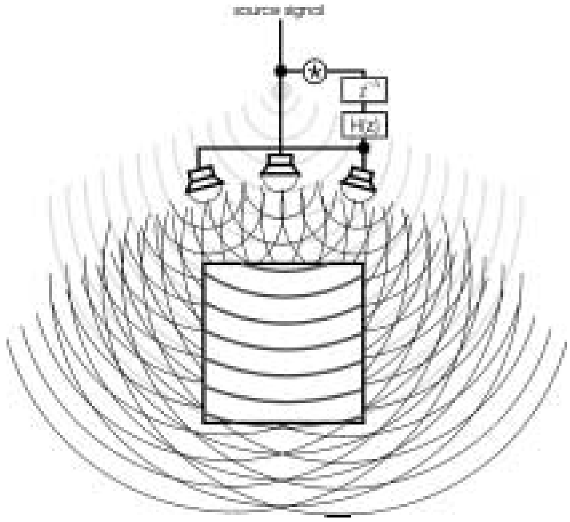


Figure 3: *Scheme of virtual loudspeaker placement and restricted reproduction area (highlighted rectangle).*

2.2. Calculation

In order to minimize the overall error concerning the sound field curvatures the loudspeaker feeds must be adjusted properly. The solution is found by solving the following over determined equation system (see Eq. 2), whereby the vector $\vec{r}$ represents a set of discrete positions within the defined area, where the two fields are compared and ω is a set of discrete frequency values.

$$P_{REF}(\vec{r},\omega) \equiv P_{SYS}(\vec{r},\omega) = \sum_i G_i(\omega) \cdot P_{LS\,i}(\vec{r},\omega) \qquad (2)$$

$P_{REF}(\vec{r},\omega)$ refers to thereference sound field and $P_{LS_i}(\vec{r},\omega)$ refers to the sound fields caused by the loudspeakers. $G_l(\omega)$ describes the frequency dependent weights (filters) for the different loudspeaker feeds. By rewriting equation 2 in matrix form (see Eq. 3) the required filter sets $\mathbf{G}$ is obtain by calculating the pseudo inverse of matrix $\mathbf{P}_{SYS}$.

$$\mathbf{P}_{REF} = \mathbf{G} \cdot \mathbf{P}_{SYS} \qquad (3)$$

$$\mathbf{G} = (\mathbf{P}_{SYS}^T \cdot \mathbf{P}_{SYS})^{-1} \cdot \mathbf{P}_{SYS}^T \cdot \mathbf{P}_{REF} \qquad (4)$$

According to the distance of the reproduced virtual source the filter set will change. In order to control the distance of sound sources in real time the filters $G_l(\omega)$ are divided (depicted in figure 3) into a fixed filter H(z) and variable gains (*) and delays ($z^{-\Delta}$) depending on the coded distance. Therefore the distance of a source can be easily adjusted like the panning with a "pan-pot". The range of possible source distances is bounded by the distance and the number of the real loudspeakers. The number of the adjacent virtual loudspeakers and their apex angle are important design parameters and have to be further investigated.

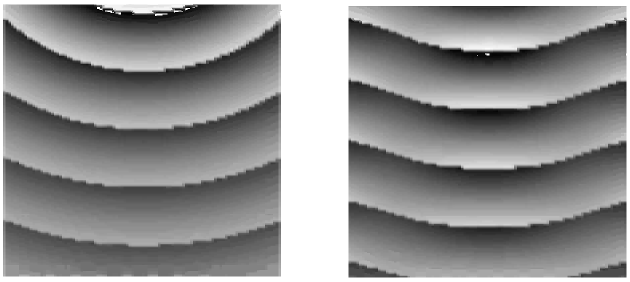
2.3. Channel coding and audio rendering

In the following coding section each "virtual loudspeaker feed" is coded around the real position of each single source into the Ambisonic domain. Subsequently these Ambisonics signals produced by the virtual loudspeakers are summed up. Therefore the advantage of this strategy is that the amount of transmission channels is bounded independent of the introduced number of single sources. Furthermore in the Ambisonic domain the sound field transformations features (e.g. rotation, acoustical focusing) can be realised easily. Afterwards the Ambisonics signals are decoded to the existing real loudspeaker rig or stored on recording tapes.

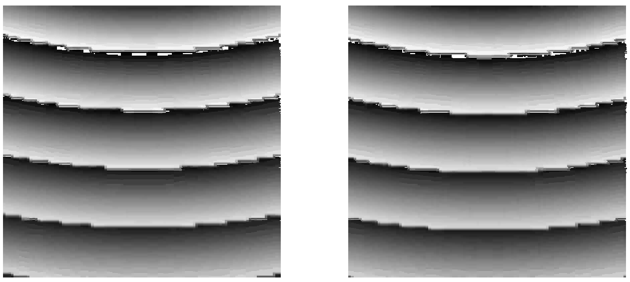
3. RESULTS

Below some results of the sound fields within the restricted reproduction area for different source distances are given. On the left hand side the reference field (target) is depicted and on the right hand side the reproduced system field is given. The size of the restricted area is 2 times 2 meters. Three virtual sources (loudspeakers) are used at r_0 = 5m and the apex angle is set to 10 degree. The coded source radiates a monochrome wave with 800 Hz. The distance of the coded source to ideal listening position, placed in the middle of the restricted area, is respectively given.

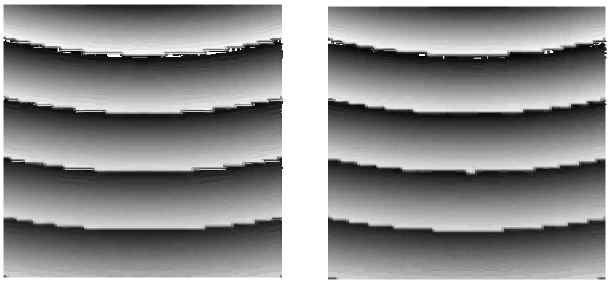
Coded source at 2 m distance.



Coded source at 4.6 m distance.



Coded source at 11 m distance.



4. CONCLUSIONS

The proposed system design enables the control of sound sources around a large auditorium. The reproduced directions of the sources are almost independent of the listening position inside the auditorium. This is achieved by dividing the reproduction of synthesised sound fields into two parts: the describing (virtual) and the rendering part.

As a future work reliable objective and subjective measurements on the proposed scheme have to be done. We will investigate the behaviour of the reconstruction error concerning the ideal number of virtual and real loudspeakers. An objective measure for the reconstruction error regarding the curvature might give an optimization cue. The number of the adjacent virtual loudspeakers, their apex angle, and the frequency distribution of the source are important design parameters and have to be further investigated.

5. REFERENCES

[1] Berkhout, A.J., "A Holographic Approach to Acoustic Control", J. Audio Eng. Soc., 36(12), pp. 977-995, 1988.

[2] Verheijen, E., "Sound Reproduction by Wave Field Synthesis", Thesis, TU Delft, 1998.

[3] Start, E.W., "Direct sound enhancement by wave field synthesis", Thesis, TU Delft, 1997.

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

[5] Daniel, Jérôme, "Acoustic field representation, application to the transmission and the reproduction of complex sound environments in the multimedia context" (Engl. translation), PhD Thesis, 2000.

[6] Malham, D.G. "Higher order Ambisonic systems for the spatialisation of sound" Proceedings, ICMC99, Beijing, October 1999.

[7] Sontacchi, A: Höldrich R., "Enhanced 3d sound field synthesis and reproduction system by compensating interfering reflections", DAFX'00 Verona, Italy, December.

[8] Sontacchi, A: Höldrich R., "Further investigations on 3D sound fields using distance coding", DAFX'01 Limerick, Ireland, December 2001.