Audio Engineering Society

Conference Paper

Presented at the 21st Conference
2002 June 1–3 St. Petersburg, Russia

Distance Coding in 3D Sound Fields

Alois Sontacchi[1] and Robert Höldrich

Institute of Electronic Music and Acoustics, University of Music and Dramatic Arts, A-8010 Graz, AUSTRIA (http://iem.at)
[1]Correspondence should be addressed to alois.sontacchi@kug.ac.at

ABSTRACT

This investigation proposes a possibility to synthesise 3D sound fields over loudspeakers taking distance coding into account. The system can be divided into two parts combining the benefits both using the Wave Field Synthesis (WFS) and the Ambisonic approach. In order to code the virtual source distances the driving functions using a derivative of the WFS approach are primarily calculated. In the second step the apparent solid angle of the sources are coded using the extended Ambisonic approach.

1. INTRODUCTION

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 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. 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]. In the second section the derivation of the distance coding 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.

1.1 Wave Field Synthesis

The WFS approach is based on the Huygens` Principle which is mathematically described by the Kirchhoff Helmholtz Integral (see eq. 1). This 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.

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

Whereas $G(\vec{r}_R|\vec{r}_S)=\frac{e^{-jk|\vec{r}_R-\vec{r}_S|}}{|\vec{r}_R-\vec{r}_S|}$ 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 our approach the derivation of the virtual loudspeaker feeds is closely related to the WFS approach. The feeds of the distributed loudspeakers are filtered in order to obtain the specified sound field in a defined area. In [2] and [3] the calculation of this filters for linear and planar loudspeaker arrays are presented. Caused by the finite and discrete arrangement artefacts are introduced which are addressed there, too. In our case a special arrangement of virtual loudspeakers along a segment of a circle is used. The calculation procedure for these loudspeaker weights (filters) is given in section 2.

1.2 Higher Order Ambisonics

In general the Ambisonics approach is presented with the well-known B-format with the signals W, X, Y and Z. This system approach is based on the sound fields spatial decomposition in spherical harmonics of 0th and 1st order. In common 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 three dimensional case the sound field produced by a plane wave arriving from direction $\vec{k}$, can be expanded into a series of spherical harmonics in a point of interest $\vec{r}$ as follows

$$e^{i\vec{k}\cdot\vec{r}} = \sum_{l=0}^{\infty}\sum_{\nu=1}^{2l+1}(2l+1)\cdot i^l\cdot\eta_{l,\nu}\cdot Y_{l,\nu}(\vec{r})\cdot Y_{l,\nu}(\vec{k})\,j_l(|\vec{k}||\vec{r}|) \qquad (2)$$

whereby j_l denotes the spherical Bessel functions related to the order *l,* further $\eta_{l,\nu}$ denotes the Neumann functions, and $Y_{l,\nu}(*)$ the spherical harmonics. If the wave field is inspected in the origin where $|\vec{r}|\equiv 0$ the mathematical description of equation 2 can be reduced to a series of an infinite sum of cosine and sine weighted Legendre-functions. This leads to the common coding and decoding equations which can be found elsewhere in [5] and [6]. 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. We suggest the approach of applying spatial filters, which we call window applied decoding. Unwanted artefacts are reduced by weighting higher order signals less than lower orders. Therefore the localisation blur is slightly increased but the perceived direction is much more stable [7], [8]. In the case of recording real sound fields even new complex microphone characteristics are required. A possible solution approach is given in [9].

1.3 Motivation and System Design

The basic idea is to build up a system based on the advantages of HOA (e.g. sound field rendering) and WFS (e.g. distance coding). We use the sound field curvature of the direct sound to mime the distance perception of a virtual source. Therefore the system design is based on coding the curvature with a derivative of the WFS approach and rendering the sound field with the HOA approach. Thus, the resulting system can be divided into two parts (see fig. 1). Each 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 systems origin, which is identical to the ideal listening position. In the following the two dimensional case is treated.
In the first part each source signal is mapped to virtual loudspeaker feeds $Q_i(\vec{p})$. These loudspeakers are located on a segment of a circle with radius r_0 which corresponds with the average distance of the real loudspeaker positions. The spacing of the virtual loudspeakers should correspond with the spacing of the real ones, too. The mapping procedure neglects the direction of the sound source, only the distance from the origin is taken into account. Therefore the required complex weights (filters) for the virtual loudspeakers have to be adjusted in order to the source distance.
In the second part these loudspeaker feeds are coded to the HOA domain around the (before neglected) according direction of each sound source. All those sound field are summed up in the HOA domain. Different sound field manipulations can be performed easily (e.g. rotation, acoustic focusing etc.). Afterwards the Ambisonic signals are decoded to the real loudspeaker rig by a simple decoding matrix given in [5], [6]. According to the Ambisonic assumptions the loudspeakers should be placed symmetrically (as possible) on a circuit in the 2D case, and arranged symmetrically (as possible) on the surface of a sphere in the 3D case [7]. Therefore the real loudspeaker array has to be calibrated.

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 (see fig. 2).
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.

2.2 Calculation

In order to minimise the overall error concerning the sound field curvatures the loudspeaker feeds must be adjusted properly. We use the LMS approach to calculate the appropriate loudspeaker weights which produces more or less the same results using the WFS approach with additional optimisation methods to reduce the artefacts caused by the discrete and finite arrangement. In [10] a similar investigation was introduced. The results can be compared even in our case a solution is found by splitting the dependency of the source distance and source angle into two steps (see fig.1). The solution is found by solving the following over determined equation system (see eq. 3), 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.

$$P_{REF}(\vec{r},\omega) \equiv P_{SYS}(\vec{r},\omega) = \sum_i G_i(\omega)\cdot P_{LS\,i}(\vec{r},\omega) \qquad (3)$$

$P_{REF}(\vec{r},\omega)$ refers to reference 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 3 in matrix form (see eq. 4) the required filter sets $\mathbf{G}$ is obtain by calculating the pseudo inverse of matrix $\mathbf{P}_{SYS}$.

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

$$\mathbf{G} = (\mathbf{P}_{SYS}^{T} \cdot \mathbf{P}_{SYS})^{-1} \cdot \mathbf{P}_{SYS}^{T} \cdot \mathbf{P}_{REF} \quad (5)$$

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 fig. 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 position of the single sound sources into the Ambisonic domain. Subsequently the corresponding Ambisonics signals are summed up. The advantage of this strategy is that the number of transmission channels is bounded and independent of the introduced number of sources. Furthermore in the Ambisonic domain the sound field transformation 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 in figure 4 to 6 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 r0 = 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.

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. 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., Vol. 36 No. 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, J., "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. Noisternig, M. Majdak, P., Höldrich, R., "An Objective Model of Localisation in Binaural Sound Reproduction Systems", 21st AES Conference, St. Petersburg, 2002.

[8] Sontacchi, A. Noisternig, M. Majdak, P., Höldrich, R., "Subjective Validation of an Objective Localisation Model", 21st AES Conference, St. Petersburg, 2002.

[9] Poletti, M. A.,"A Unified Theory of Horizontal Holographic Sound Systems", J. Audio Eng. Soc., Vol. 48, No. 12, 2000.

[10] Kirkeby, O., Nelson, P. A., "Reproduction of plane wave sound fields", J. Acoust. Soc. Am. Vol. 94, No. 5, 1993.

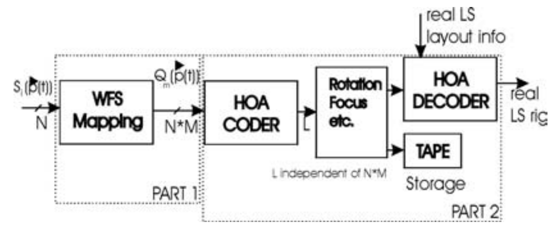

Figure 1: *System model.*

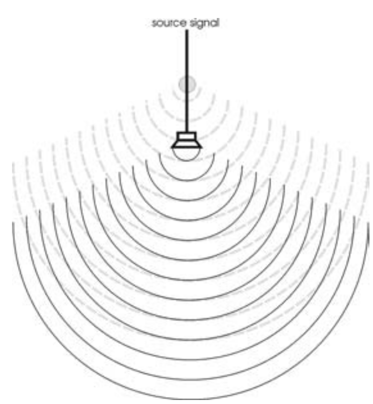

Figure 2: *Different curvature caused by different positions.*

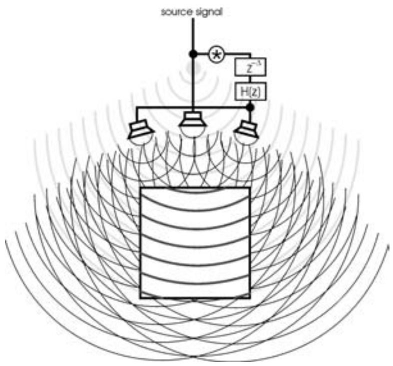

Figure 3: *Scheme of virtual loudspeaker placement and restricted reproduction area (highlighted rectangle).*

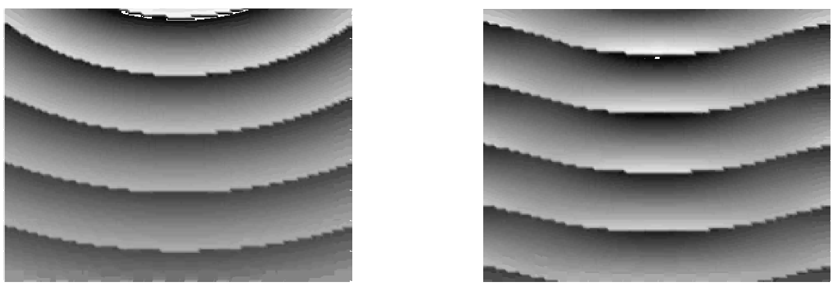

Figure 4: *Coded source at 2 m distance.*

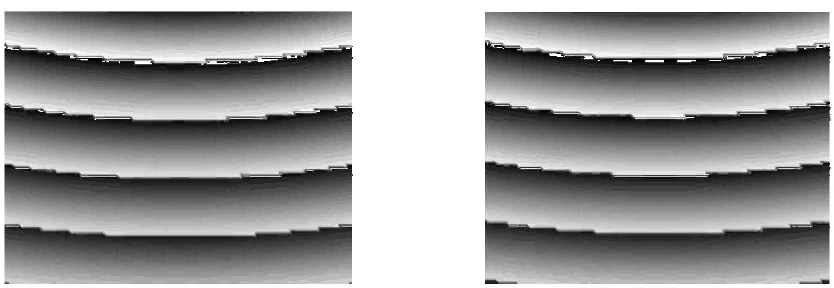

Figure 5: *Coded source at 4.6 m distance*

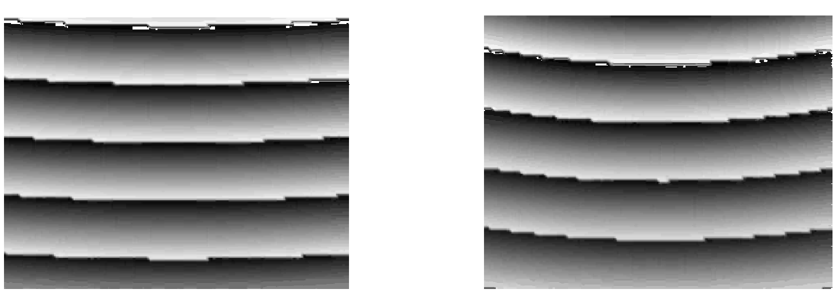

Figure 6: *Coded source at 11 m distance*