HOA Introduction

HOA is a technology provided by Technicolor that enables a new generation of sound capture, processing, and reproduction systems for 3D immersive sound. Instead of introducing more and more audio channels in formats like 5.1, 7.1, 11.1 up to 22.2, HOA is based on a sound field description using Higher Order Ambisonics and therefore is a future proof, loudspeaker layout independent system. HOA is based on linear matrix operations and fits seamlessly into today's existing infrastructure. Every single audio source can be transformed into the HOA representation using a linear encoder matrix and loudspeaker signals can be derived from the HOA representation using a linear decoder matrix. In the following, an overview of existing audio content representations is given to explain the different paradigms in audio production.

Channel-Based - Today most sound systems are loudspeaker channel based. That means that sound mixtures are produced for pre-defined loudspeaker layouts. The compliance of the loudspeaker positions to the addressed layout is mandatory for a successful sound field reproduction. To address other layouts, automatic down- and up-mixing methods exist, but today still suffer from well-known problems like comb filtering and volume differences between sound elements.

Object-Based – During the production phase sound designers are mixing single audio sources using panning laws and effects into the said loudspeaker channel-based formats. Alternatively, the content can be transmitted as separated audio-objects together with panning instructions and other meta data. In this case the reproduction system will perform the mixing process for the locally installed loudspeaker layout. Hence, sound designers are not restricted to loudspeaker layouts anymore, but limited by the properties provided by the objects.

HOA-Based – The HOA-based approach describes the entire sound field and therefore circumvents the problem of modeling a sound scene by objects. The sound field is described by a mathematical transform describing the spatial audio content. This transform can be applied to individual audio objects as well as to ambient sound fields recorded using microphone arrays.

Recordings carried out with a spherical microphone array [1] show how easy it is to do live recordings, transmissions and reproductions of an ambient sound from a sport event or live concert [2].

HOA with Objects - A combination of HOA together with audio-objects is still loudspeaker setup independent. In addition it would allow interactive modifications and adjustments of the mixture at the reproduction location [2]. If interactive modifications are not required, audio-objects may be rendered into the HOA format and added to the existing HOA content by linear mixing.

HOA Distribution Format

Linear-matrixed – HOA based on Higher Order Ambisonics is basically a linear matrix system. The required matrix elements are based on a mathematical description of a sound-field using spherical harmonics [3].

HOA coefficient channels - HOA content consists of a certain number of audio signals, where each audio signal contains the time-domain HOA real-valued coefficients representation. A physically exact sound field description is given by an infinite number of HOA coefficient signals. Due to the limited spatial resolution required for reproduction, this series can be truncated. The maximal number of HOA coefficient signals is given by the HOA order 'N' and equals to '$(N + 1)^2$'. The minimal required order depends on the minimal angular spacing of loudspeakers and the listening area. Higher spatial resolutions support large listening areas whereas smaller resolutions are sufficient for smaller listening areas e.g. for Home Theatres.

As a rule of thumb, the HOA order should be greater-equal to 4 for diffuse audio objects and should be between 6 and 8 for high and very high spatial resolutions.

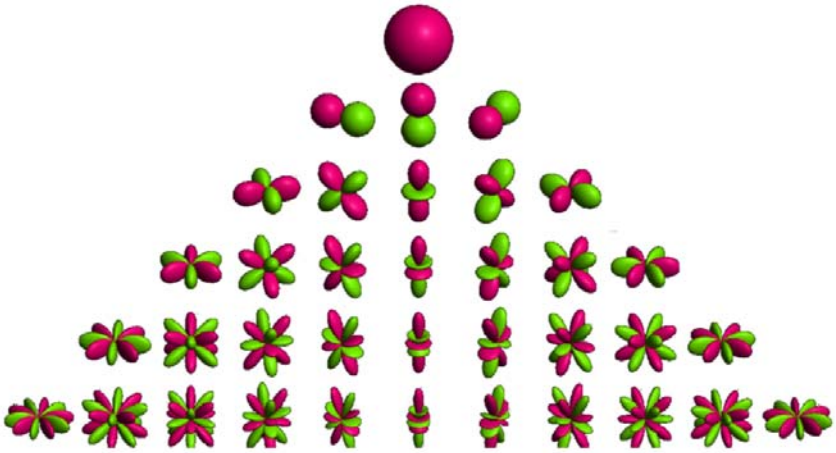

Figure 1: Spherical Harmonics. Figure shows 1st row: order 0 with 1 coefficient up to 6th row: order 5 with in total 36 real coefficients. As higher the order, the higher the spatial resolution. Red portions represent regions where the function is positive, and green portions represent regions where the function is negative.

Recording - Spherical microphone arrays, like in [1], allow direct capturing of spatial live sound. The 32 capsules mounted on a rigid sphere deliver an maximum HOA order of 4 [5].

Scalable - HOA is scalable in terms of the spatial resolution. If the transmission channel bandwidth is restricted and the given reproduction system angular spacing is wide, like for Home Theatres, higher orders can simple be neglected to obtain a lower order HOA format.

Compression - Further bandwidth reduction is possible using a HOA compression format, as it is currently under development by the ISO/IEC MPEG audio group. 1.2 to 1.5 Mbit/s are expected for a transparent transmission.

Rendering Engine and Loudspeaker Layout

Rendering engine - The HOA renderer may combine incoming HOA, audio object and channel-based signals into a common HOA format before rendering the content to the existing loudspeaker setup. Alternatively, the HOA renderer can render HOA formats only and add audio objects using e.g. VBAP-based rendering techniques [6]. In addition channel based content could simply drive the predefined loudspeaker directly, if they exist.

Loudspeaker Layout - HOA content is loudspeaker layout independent, therefore a HOA renderer can fully take advantage of any loudspeaker layout [4]. Of course, some setups are more ideal than others. As a rule of thumb, it is preferred to have spatially equal distributed loudspeakers over the area the sound is expecting to arrive from.

Technicolor recommends adding to the existing cinema loudspeaker row behind the screen a lower and an upper row, to control every existing loudspeaker in the room as individual speaker and to add some speakers in the front side areas of the cinema where speakers are missing. On ceiling level it is recommended to add a regular distributed sparse loudspeaker array.

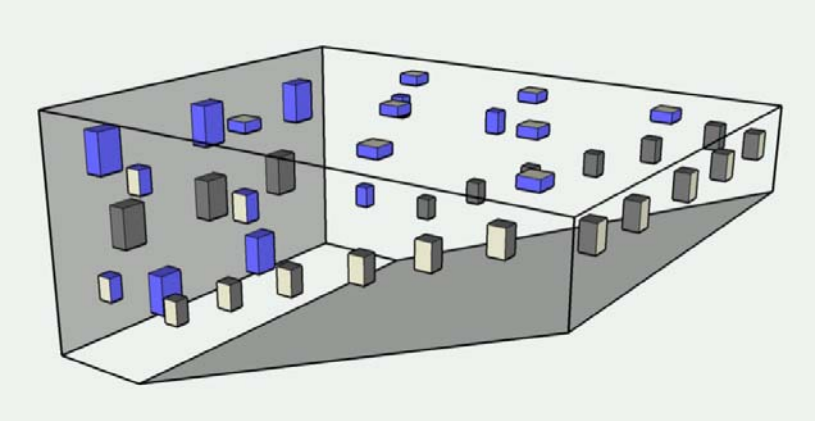

Figure 2: HOA example loudspeaker layout for cinemas – grey: existing loudspeakers – blue: new proposed loudspeakers

Compatibility - Fully backward compatibility to other existing reproduction systems like 5.1, 7.1, 22.2 and higher is guaranteed due to the HOA format description.

The HOA format, provided by Technicolor has been tested and evaluated in various movie theater sites. In particular the Technicolor Interoperability Test Center (ITC) in Burbank has a movie theater where the format was tested. We did comparisons using a conventional 7.1 surround sound system and experimental loudspeaker installations for 3D sound, having up to 36 loudspeakers. At other movie theater sites already installed Auro-3D layouts where used for playback.

Personal systems - In addition to the room-based systems, personal systems will exist. Headphones with appropriate spatialisation would also be a means to deliver an immersive 3D audio signal. Head movements, measured by a head tracker, can easily be compensated with an inverse sound field rotation in the HOA domain, which leads to a better externalization.

Conclusion

HOA breaks with the existing trend to simply add more and more audio channel formats. Instead it describes the desired sound field and allows for one production format that is suitable to all loudspeaker setups. It is a new paradigm.

Live recordings, using a spherical microphone array, as well as encoding of conventional microphone channels can easily be done using a linear-matrixes process.

The HOA format is scalable in terms of spatial resolutions and can easily be truncated to a lower resolution if required.

The rendering enables the reproduction of an immersive 3D sound field onto any arbitrary loudspeaker setups including irregular loudspeaker layouts, without up- and down-mixing restrictions. Sound designers are no longer restricted to loudspeaker layouts.

References

[1] Eigenmike, http://www.mhacoustics.com/mh_acoustics/Eigenmike_microphone_array.html

[2] "3D Audio in FascinatE", http://www.fascinate-project.eu/index.php/tech-section/audio/

[3] Spherical Harmonics, https://en.wikipedia.org/wiki/Spherical_harmonics

[4] "Investigation of Robust Panning Functions for 3D Loudspeaker Setups", Johann-Markus Batke, Florian Keiler, 128th Convention of the Audio Eng. Soc.2010

[5] "Optimization of Spherical Microphone Array Recordings", Sven Kordon, Alexander Krüger, Johann-Markus Batke, Holger Kropp, ICAS 2011

[6] "Spatial Sound Generation and Perception by Amplitude Panning Techniques" Pulkki, V. (2001), PhD thesis, Helsinki University of Technology