Audio Engineering Society

Convention Paper

Presented at the 137th Convention
2014 October 9–12 Los Angeles, USA

An Architecture for Reverberation in High Order Ambisonics

Fernando Lopez-Lezcano[1]

[1] *CCRMA, Stanford University, USA*

Correspondence should be addressed to Fernando Lopez-Lezcano (`nando@ccrma.stanford.edu`)

ABSTRACT

This paper describes a reverberation architecture implemented within the signal chain of a periphonic HOA (High Order Ambisonics) audio stream. A HOA signal (3rd order in the example implementations) representing the dry source signal is decoded into an array of virtual sources uniformly distributed within the reverberant space being simulated. These virtual sources are convolved with independent, decorrelated impulse responses, optionally tailored to model spatial variations of the simulated reverberation. The output of each convolver is then encoded back into HOA and mixed with the original Ambisonics dry signal. The result is a convolution reverberation engine with a HOA input that outputs HOA and maintains the spatial characteristics of the input signal.

1 INTRODUCTION

In creating a virtual acoustic environment, it is desired to not only locate dry sound sources in space, but also to spatialize the reverberant response of the simulated space. Periphonic High Order Ambisonics (HOA)[1][2] is commonly used to position sound sources within a space, and a number of tools are available for manipulating the spatial nature of HOA-encoded signals. This work focuses on synthesizing and controlling the spatial character of the environment reverberation, and presents an architecture for 3D reverberation based on HOA.

The method processes a periphonic HOA input with a set of decorrelated, spatially distributed impulse responses to produce a periphonic HOA output of the same or lesser order as the input. By driving a set of statistically independent directional impulse responses with signals derived from the HOA dry signal, the directionality of the dry signal can be preserved in the reverberated soundfield.

2 ARCHITECTURE

Current practice for Ambisonics reverberation directly processes B-Format first-order incoming signals, the components of which correspond to physical directions, or converts them to A-Format, processes them in this domain and re-encodes them back to B-format[4]. HOA components have no relationship to spatial directions, so the method presented here decodes the incoming signal to a set of virtual sources that define a virtual space, creates the reverberation in that virtual space and encodes the resulting reverberated soundfield back into HOA.

2.1 The virtual space

To create a virtual space in which to process the dry signal we define an array of virtual sources that spans the whole sphere and design an Ambisonics decoder for them. Since the arrangement of sources is arbitrary we pick regular 3D arrangements that make it easier to design the decoder. For orders up to 2nd we use platonic solids. For higher orders we use spherical t-designs which also arrange the virtual sources in the sphere regularly[7], and choose a t-design that can correctly sample the spherical harmonics of the desired order. In both cases it is easy to design proper decoders for them using the ADT (Ambisonics Decoder Toolkit)[3] software package.

This decoder is only used to create the reverberation soundfield in the ideal virtual space. Once generated, the soundfield is not used for diffusion but is immediately encoded back to Ambisonics. The tradeoffs and optimizations needed for choosing, designing and tuning a decoder do not necessarily apply to this case (but they do apply later, when the reverberated soundfield mixed with the dry signal is diffused in a real space). For the examples discussed below we used simple one band max-r_E decoders because they maximize the concentration of energy in the source direction

2.2 The reverberator

To create the reverberant soundfield we start with a simple topology that convolves each one of the virtual sources with a separate, statistically independent impulse response. Each convolution describes the reverberation characteristics of the space being simulated in the direction of the virtual source. The output of all the convolvers creates a soundfield at the center of the virtual space that represents the reverberation characteristics of the whole space, as defined by the impulse responses being convolved.

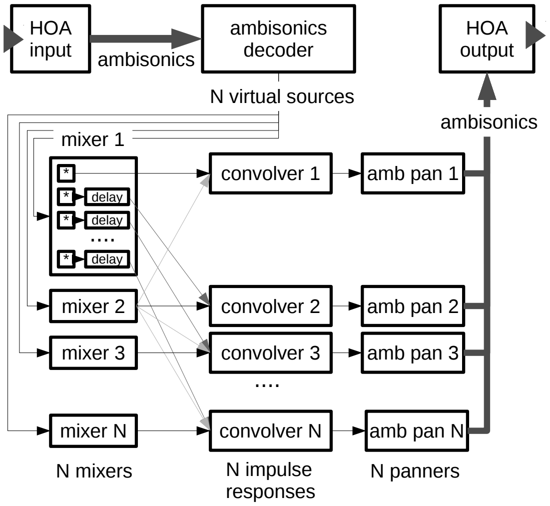

Fig. 1: signal flow for the reverberator

A signal encoded in the dry HOA soundfield with a position described by azimuth and elevation angles will be convolved with the impulse responses that correspond to the virtual sources that contribute to the decoded soundfield as defined by the Ambisonics decoder being used, and will generate a reverberation signal in the same direction as the HOA dry signal. Multiple dry signals in the input soundfield will generate multiple reverberation components in their respective directions. This simple scheme will only generate “local” reverberation, as there will be no reverberation energy coming from directions other than the one of the dry signals.

To generate global diffuse reverberation we feed each virtual source to all convolvers through a mixing vector. The gain coefficients of the mixing vector are determined by a transfer function that maps the distance between any given virtual source and all others to attenuation factors. The further away a virtual source is from the current one, the less signal will be sent to its convolver.

A power law can be used to provide good control of the spatial spread of the reverberation. It will create no attenuation for the convolver associated with a given virtual source, and increasing attenuation as the distance increases:

$$L = O + (B^{(-D)} * (1.0 - O)) \tag{1}$$

L is the level that the first source will contribute to the convolver of the second source, D is the distance between the pair of virtual sources, and B is the base of the power curve. A base B of 1.0 will generate purely global reverberation, that is, all virtual sources will contribute equally to all convolvers. Values of B greater than one will generate more localized reverberation, the higher the value of B is the more localized is the reverberation soundfield. A non-zero offset O is used to ensure that all virtual sources contributes a minimum of energy to all convolvers for a more enveloping reverberation soundfield.

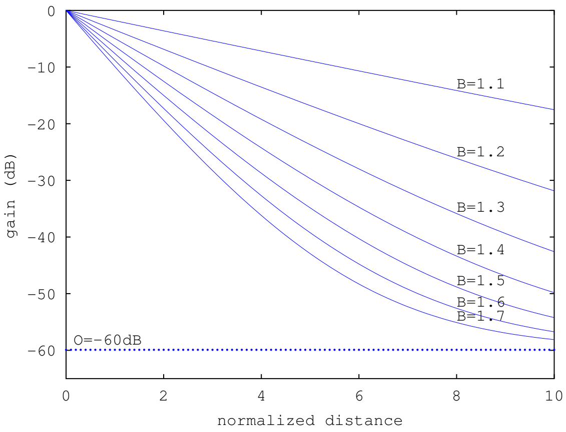

Fig. 1: gain versus normalized distance of virtual sources

The gain coefficients for all virtual sources can be normalized so that the total energy fed into the reverberant soundfield remains constant as the base B or offset O are changed. A power law with an offset provides a simple transfer function with only two coefficients that can easily change the spatial nature of the reverberation.

We add optional delay lines to each output of the mixing vector so that the contribution of each virtual source to other convolvers can be delayed in proportion to the distance between the virtual sources. This better approximates the behavior of a real reverberant space.

To create the final reverberated HOA signal we pan the output of each convolver by using an Ambisonics panner of the appropriate order, and the azimuth and elevation angles of the corresponding virtual source. The outputs of all panners are mixed together and constitute the reverberated HOA signal.

2.3 Optimizations

It is useful to find strategies that minimize the number of convolutions needed for a sufficiently convincing reverberation effect.

We do not require very precise localization of the reverberated HOA signals, so we can relax the size requirements for the array of virtual sources and use less sources than Ambisonics theory requires. For example, while for a regular dodecahedron we can create proper Ambisonics decoders only up to 2nd order, we could still use a dodecahedron to decode 3rd order signals with reduced accuracy.

It may also not be necessary to simulate the reverberation characteristics of the whole sphere. If we choose to ignore reverberated energy coming from negative elevations we can use a dome configuration of virtual sources. The current version of ADT can easily create good decoders for such configurations, using either the Allrad or Slepian methods.

3 ENHANCEMENTS

A number of enhancements are available to augment the basic approach. Each impulse response can be tailored to the reverberation characteristics of the virtual space in the direction of its associated virtual source, that is, the reverberation character and length can change according to the elevation and azimuth angles of the direction of the dry signal.

As the reverberated signal is Ambisonics, all available Ambisonics soundfield transforms[8] can be used to manipulate the resulting soundfield. These transforms can be applied to the output reverberated signal, or to the panned Ambisonics outputs of individual convolvers. The most useful are the warping and directional loudness transforms which can be used to "push" the reverberation in a certain direction.

This simple architecture can be further enhanced by splitting the processing into two parallel convolution stages to separately create the early reflections and late field reverberation. The early reflections can be created out of a larger number of shorter convolutions and the late field reverberation can be a smaller number of longer convolutions. We can use different virtual source locations for both corresponding to differences in the spatial nature of the early and late reverberation of the modeled virtual space.

3.1 Using an FDN core

While the current power of CPUs can handle a large number of convolutions as required by this method, all convolutions can be replaced with an FDN (Feedback Delay Network) based reverberation engine. This has been tested by using the FDN reverberator core corresponding to the Zita Rev1 reverberator[9]. This alternative approach trades realism of the reverberation for lower CPU load suitable for less capable processors, while maintaining the goal of having an HOA-in HOA-out reverberation.

4 IMPLEMENTATIONS

This architecture provided the reverberation for a full 3D third-order Ambisonics piece composed by the author during a short residency at ZKM, Karlsruhe, Germany (“Space, S[acred|ecular]”, May 5-12, 2014). The reverberator used was a simplified version of the scheme described here. It used 24 statistically decorrelated impulse responses for the Hagia Sophia Hall created by the Icons of Sound Project[6] and used for the realtime virtual space of the “From Constantinople to California” Bing concert[5]. Each impulse response is originally 16 seconds long, but was scaled to only about 12 seconds for the purposes of the piece.

The figure below shows a diagram of the 3rd order Ambisonics reverberation as implemented for the piece in an Ardour session[10]. The session was created with a 16 channel master bus that can directly accommodate 3rd order Ambisonics mixing.

The virtual space for the reverberation was defined by using 15 virtual sources arranged in a dome configuration. A max-r_E Ambisonics decoder was created using ADT to decode the dry 3rd order signal to the 15 virtual sources, and was inserted as a plugin in the 16 channel input bus. The 15 outputs of this bus (the dry virtual sources) were routed to stereo buses which hosted stereo IR convolver plugins. The outputs of the individual channels of the convolution buses were then routed to 15 additional buses, one for each reverberated virtual source. 3rd order Ambisonics panners in each of these buses pan the reverberated components in the virtual space, at the same azimuth and elevation angles as the original virtual source positions. The outputs of these panner buses are sent to the master bus of the session which is the reverberated 16 channel 3rd order Ambisonics output signal.

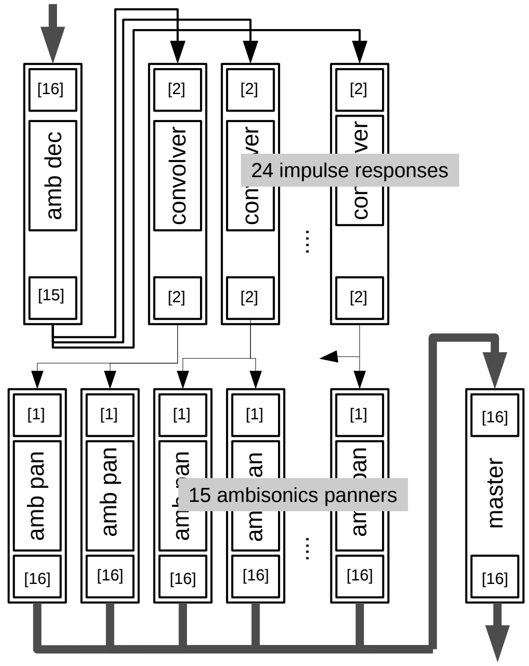

Fig. 2: a simple implementation of a 3rd order HOA-in HOA-out reverberator using Ardour and plugins

Eight additional impulse responses enveloped to remove the early reflections were used to create global reverberation. Each one was fed from 2 different virtual sources and connected to the panner buses in the opposite quadrant of the virtual space. This is a simplification of the scheme of a mixing vector that feeds all convolvers from each virtual source, but proved to be sufficient for the purposes of the piece.

The result was a very natural sounding HOA-in, HOA-out reverberation that matched the 3D Ambisonics soundfield of the piece, following the spatial characteristics of the piece yet providing a convincing global reverberated soundfield. This simplified implementation was a consequence of the limited time available to create it, and the limited CPU power available to render the effect.

All processing was done using open source free software without any custom programming. Software used included Ardour, the LV2 IR plugins, the LADSPA Ambisonics 3,3 panner and ADT for the Ambisonics decoder design.

In a subsequent test a similar design used the core of the Zita Rev1 reverberator as implemented in Faust to add 3D 3rd order full sphere reverberation to a different HOA piece with similar results. The regular controls included in the Zita Rev1 core were used to control reverberation length at two frequency bands and high frequency damping.

5 CONCLUSIONS

A simple architecture for HOA reverberation was presented and has been used to provide a very smooth and enveloping reverberation for 3rd order Ambisonics materials. Two coefficients can change the spread of the reverberation effect, and the reverberation automatically follows the spatial characteristics of the input signal. This architecture was used to provide two full 3D reverberant environments for two pieces, one based on decorrelated impulse responses for the Hagia Sophia Hall, and another based on an FDN reverberator core.

6 ACKNOWLEDGMENTS

Many thanks to Jonathan Abel who encouraged me to submit this paper and helped enormously to craft the original precis. Also thanks to Aaron Heller, Eric Benjamin and Jörn Nettingsmeier for suggestions and corrections to the original precis and this paper. I also want to thank the open source free software audio community for the many professional tools that helped make this project a reality.

7 REFERENCES

[1] Michael Gerzon, "Ambisonics in multichannel broadcasting and video", Journal Of The Audio Engineering Society, 1985, Volume 33, Number 11

[2] Jérôme Daniel, "Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia (*Acoustic field representation, application to the transmission and the reproduction of complex sound environments in a multimedia context*), PhD Thesis, Université Paris 6, 2001

[3] Aaron Heller, Eric Benjamin, "The Ambisonic Decoder Toolbox," presented at the Linux Audio Conference 2014

[4] Joseph Anderson, Sean Costello, "Adapting Artificial Reverberation Architectures for B-Format Signal Processing," Ambisonics Symposium 2009

[5] Fernando Lopez-Lezcano, Travis Skare, Michael Wilson, Jonathan Abel, "Byzantium in Bing: Live Virtual Acoustics Employing Free Software," Linux Audio Conference 2013

[6] Jonathan Abel, Nicholas Bryan, Patty Huang, Miriam Kolar, Bissera Pentcheva, "Estimating Room Impulse Responses from Recorded Balloon Pop," AES129

[7] R. H. Hardin and N. J. A. Sloane, "McLaren's Improved Snub Cube and Other New Spherical Designs in Three Dimensions," Discrete and Computational Geometry, 15 (1996), pp. 429-441

[8] Matthias Kronlachner and Franz Zotter, "Spatial transformations for the enhancement of Ambisonic recordings", International Conference on Spatial Audio, 2014.

[9] Fons Adriansen, "Zita-Rev1, a quick guide", http://kokkinizita.linuxaudio.org/linuxaudio/zita-rev1-doc/quickguide.html (last accessed: July 8 2014)

[10] Jörn Nettingsmeier, "Ardour and Ambisonics, A FLOSS approach to the next generation of sound spatialisation", http://cec.sonus.ca/econtact/11_3/nettingsmeier_ambisonics.html (last accessed: July 8 2014)