

ADAPTING ARTIFICIAL REVERBERATION ARCHITECTURES FOR B-FORMAT SIGNAL PROCESSING

Joseph Anderson¹, Sean Costello²

¹ School of Arts & New Media, University of Hull, UK (j.anderson@hull.ac.uk)

² Valhalla DSP, USA (sean@valhalladsp.com)

Abstract: *Auralisation ray-traced volume modeling reverberation methods presently appear to be a preferred approach for generating surround sound reverberation. However, a wide literature describing various architectures for artificial reverberation filters is extant. Many of these alternative delay-line methods give distinct performance advantages, in that they are not reliant upon convolution or volume modeling, and allow a variety of tonal and spatial parameters to be modified independently. The authors describe a method to expand or adapt these known architectures to an Ambisonic B-format context, and provide an illustrative example native B-format reverberator. Central to this approach is the adaptation of scattering junctions to a B-format context and the use of B-format spatial image transformation techniques.*

Key words: Ambisonic, Reverberation

1 INTRODUCTION

Our present discussion aims to describe the methods the authors have used to develop native B-format reverberator networks. One of the authors, Anderson, is a composer who works with B-format sound recordings. With *Kyai Pranaja* [1] Anderson began to experiment with artificial reverberation of B-format soundfields, which was initially approached by applying separate mono reverberators to each channel (W, X, W, Z) of the B-format signal. While in some cases this may generate a ‘compositionally suitable’ soundfield, the results are not necessarily wholly satisfactory.

It is with work on *Pacific Slope* and *Mpingo* [1] that the task of developing a satisfactory reverberation for B-format was revisited, this time involving collaboration with Costello. As B-format acquisition is not a necessarily a simple task, requiring a Soundfield microphone and associated 4-channel recorder, maintaining appropriate ‘Ambisonic-ness’ of the acquired sound recording was an important goal in this effort.

Firstly, this implies a suitable network should both input and output B-format. That is, the network should not simply be a mono in reverberator acting on the mono part (W) of the input soundfield. Secondly, the network should somehow ‘make sense’ in terms of how reverberant signal is distributed across the B-format soundfield. That is, energy should spread throughout the soundfield and be diffused wholly through the space. Thirdly, advantages of B-format soundfield representation should be observed. Being ‘B-format native’ should imply certain opportunities—from the development of the network itself (scattering matrices) to the ergonomics B-format

imaging matrices offer (control of early reflections and late reverberation).

This paper is primarily concerned with networks suitable for 1st order Ambisonic soundfields, but it is expected 2nd order networks may be developed in a similar way.

2 ARCHITECTURE

Our principle approach to adapting reverberation architectures involves a translation from B-format to the A-format domain, the ‘directional representation’ of the soundfield. In A-format, the outputs for each channel of a reverberation network should be mutually decorrelated.

2.1. Decorrelation by differing delay lengths

Ideally, the decorrelation of the A-format reverb outputs is created by different delay times, not just phase differences. In many published reverb examples [2], [3], [4] the outputs of the reverberation network are created by summing together the delay network outputs using an orthogonal matrix, such that the outputs have equal energy from all delay lines, and differ in phase. Using such an output structure in the A-format reverb may often result in unintended cancellations in the A-B format matrix.

2.2. Delay output options

The outputs can be taken from the ends of different delay lines, such as the parallel delay lines in a feedback delay network (FDN), or parallel combs. Alternately, the outputs can be obtained by weighted delay taps from the recursive reverberation network as shown by Dattorro [5].

2.3. Early Reflections via tapped delay lines

Early reflections can be generated by tapped and weighted delay lines, as long as each output channel has mutually decorrelated delay lengths. In practice, and for the Ambisonic periphonic case (full, with-height 3D), it has proven effective to stagger the delay lengths of each channel, such that the output tap delay lengths are positioned in the soundfield vertices of the A-format tetrahedron, located in space as follows: front-left-up, front-right-down, back-left-down, back-right-up. This is repeated for a convincing set of reflections.

2.4. Early Reflections via cascaded unitary blocks

It is also possible to generate early reflections from cascaded unitary stages [6], [7]. In such a system, parallel unitary operators, such as delay lines, allpass filters, or a passthrough signal, are combined via unitary matrices. By cascading several stages of the parallel operators and matrices, a high echo density can be obtained. The outputs can be taken from the outputs of the final unitary stage, or can be taken from weighted taps within the network (such as after each block's delay lines, or from the individual unitary matrices). Such a network can also be used to feed the inputs of a recursive late reverb network, as the outputs of the cascaded unitary stages preserve the total input energy.

2.5. Adapting existing structures to A-format

As an example of adapting an existing reverberation algorithm to A-format, we can start with the 4-channel FDN reverb published by Puckette [7], as similar algorithms can be found in the Pd distribution, as well as in algorithms dating back to the 1980's that were in use at IRCAM, and later distributed under the Jimmies title for Max/MSP. The Puckette algorithm takes a stereo input, and uses cascaded unitary stages as described above to quickly build the early echo density. The outputs of the cascaded unitary stages are fed into the inputs of four parallel delay lines, which are scaled and fed through lowpass filters,. The outputs of the four scaled lowpassed delay lines are combined via a unitary matrix, the outputs of which are fed back into the four delay line inputs.

2.6. Expanding 2-channel reverb to 4-channels

The cascaded unitary stages found in Puckette's algorithm are made up of a delay line, in parallel with a passthrough path, that are combined via a 2×2 rotation matrix with its rotation angle set to $\pi/4$ (the gain normalization is left out of Puckette's algorithm for simplicity). To work with A-format input signals, the cascaded unitary stages can be expanded to 4 parallel branches. The parallel branches can be 4 parallel delay lines, 3 delay lines and a passthrough path, series combinations of delay lines and allpass delays, etc. The 3 parallel delays, 1 pass-through path structure has the advantage of being able to cascade multiple stages while still having a zero-delay path through the structure, which can be advantageous when feeding the late reverb (to avoid an excessive delay of the onset of late reverberation).

3 ADAPTING TO A-FORMAT

Internally, the adapted network operates in A-format. We'll need to change domain from B-format to A-format to enter the network, and then from A-format to B-format on exiting, as early reflections and late reverberation.

In viewing the problem as how best to adapt Puckette's structure to an A-format space, we'll repeat the observation that for an ideally diffuse reverberation, all channels of A-format should be mutually decorrelated. Additionally, we'll need to observe that the canonical scaling on the W channel of B-format is $1/\sqrt{2}$, which is not ideal in this instance. This canonical scaling is usually referred to as an engineering consideration. For horizontal only soundfields (pantophonic), the scaling is a suitable choice, and when one considers B-format was initially intended as a successor of the 'quad' formats of the 1970s [8], this choice is credible. However, for a full periphonic soundfield (3D), a normalised scaling on W of $1/\sqrt{3}$ is more appropriate. We'll refer to signals using this scaling as 'W-normalised'.

3.1. Normalised A-format

For our purposes, we'll define two matrices for changing the signal domain from B-format to W-normalised A-format (1) and back again (2).

$$\begin{bmatrix} 1/\sqrt{6} & 1/2 & 1/2 & 1/2 \\ 1/\sqrt{6} & 1/2 & -1/2 & -1/2 \\ 1/\sqrt{6} & -1/2 & 1/2 & -1/2 \\ 1/\sqrt{6} & -1/2 & -1/2 & 1/2 \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} \sqrt{3/8} & \sqrt{3/8} & \sqrt{3/8} & \sqrt{3/8} \\ 1/2 & 1/2 & -1/2 & -1/2 \\ 1/2 & -1/2 & 1/2 & -1/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \end{bmatrix} \quad (2)$$

A useful feature to note of the W-normalised A-B matrix is that each point of the A-format tetrahedron encodes on the surface of the Ambisonic sphere, a characteristic appropriate for the encoding of early reflections.

3.2. Scattering

In making a choice for a suitable scattering matrix, the question remains as to what matrix may be regarded as equivalent in 4-channel A-format to a 2-channel rotation by $\pi/4$. Puckette [9] has given a hint when he discusses power conservation in complex delay networks in showing a cascade of 2-channel rotations to scatter more than two channels. In reviewing the 2-channel early reflection unitary cascade of Puckett's algorithm, each rotation of $\pi/4$ in the early reflection stages may be regarded as a 'maximally diffusive' [10] scattering network of the form shown below (3).

$$\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \quad (3)$$

This matrix may be regarded as a 2×2 Householder reflection matrix and is of the kind suggested by Jot and others [4], [10], [11]. As 1st order A-format consists of four channels, a 4×4 Householder matrix (4) is an appropriate choice for scattering in our adapted network.

$$\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \quad (4)$$

3.3. Geometric interpretation of scattering

If we consider the effect of (4) on the signal in the B-format domain, we'll see that it is equivalent to inverting the sign of W. Interpreting this action spatially, on each pass through the scattering matrix, the signal is reflected through the origin to the opposite side of the sphere. Visualising the A-format tetrahedron, one sees a pass through the scattering matrix reorients the tetrahedron so that each vertex is reflected, to be distributed between the three opposite vertices—resulting in a maximally diffusive scattering. That is, after a pass through, each vertex is re-oriented so that it is then strongly split across the delay lines positioned at the opposite three vertices.

3.4. Scattering in higher orders

The authors suspect that a similar procedure may be followed for adapting networks for high order Ambisonic (HOA) reverberation. (However, at this point an investigation has not been made.) In particular, it is suspected that the geometric interpretation of reflecting the soundfield about the origin will similarly result in maximally diffusive scattering.

4 CONTROLLING THE DECAY CURVE AND SPATIALISATION OF EARLY REFLECTIONS

One drawback of the cascaded unitary matrix approach to early reflections is that, as the number of cascaded blocks is increased, the amplitude response of the output grows closer to a Gaussian bell curve, with its characteristic fade in and fade out. This can be partially alleviated by taking the outputs as taps from within the cascaded blocks, either as taps from the individual delay lines, from the outputs of the branches pre-scattering matrices, or from the outputs of the scattering matrices themselves. The outputs can be weighted so as to produce a variety of amplitude responses, within limits, as the blocks further down the cascade will have a more Gaussian amplitude response. If a specific early reflections amplitude response is required, the cascaded unitary blocks approach should be replaced with tapped delay lines, with one delay line for each input channel, and with each delay line having weighted taps that are sent to each of the 4 output channels as required.

4.1. Bringing early reflections into B-format

Early reflections may be tapped off from each early reflection stage and panned into the resulting B-format output at azimuth θ and elevation ϕ using the familiar Ambisonic encoding matrix. (5)

$$\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \cos\phi\cos\theta \\ \cos\phi\sin\theta \\ \sin\phi \end{bmatrix} \quad (5)$$

While such a procedure allows detailed control of the position of each early reflection in the soundfield, the authors regard this to be 'over specified' in many cases.

A more convenient approach is to use the W-normalised A-B matrix (2), and variants of this matrix, to position early reflections in the soundfield. This approach places reflections at the vertices of a tetrahedron and may be regarded as maximally diffuse in a geometric sense. The resulting reflections may then be steered using rotation (6) (shown across the Z-axis) and the focus transform (7), or a combination of both. The early reflection cascade is illustrated in figure 1, as described, with the 'early reflection positioning network' including gain scaling and imaging.

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (6)$$

$$\begin{bmatrix} \frac{1}{1+|\sin\theta|} & \frac{1}{\sqrt{2}}\left(\frac{\sin\theta}{1+|\sin\theta|}\right) & 0 & 0 \\ \sqrt{2}\left(\frac{\sin\theta}{1+|\sin\theta|}\right) & \frac{1}{1+|\sin\theta|} & 0 & 0 \\ 0 & 0 & \sqrt{\frac{1-|\sin\theta|}{1+|\sin\theta|}} & 0 \\ 0 & 0 & 0 & \sqrt{\frac{1-|\sin\theta|}{1+|\sin\theta|}} \end{bmatrix} \quad (7)$$

The focus transform is a dominance [12] variant developed by one of the authors (Anderson). For this current application it brings particular advantages, in that reflections can be directed or 'focused' along the X-axis. For focus, θ refers to the image distortion half-angle. A value of 0 leaves the image unchanged, where $\pi/4$ focuses (or compresses) the image to front-centre. If one is interested in modeling concert hall reverberation, choosing a value of θ towards $\pi/4$ for the first reflection stage will give a compressed image. Then, with each successive early reflection tap-off, decreasing θ towards 0 successively opens up each reflection group.

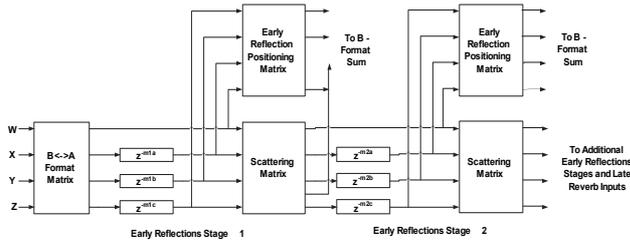


Figure 1: Early reflection cascade

5 LATE REVERBERATION

The late reverberation network [7] used in this example suffers from a low initial echo density, as well as a fairly low modal density, due to the length of the delay lines. By replacing each of the delay lines with several allpass delays in series, with a straight delay line pre or post allpass, the echo density of the network can build at a much higher rate, and the total delay of each branch can be made high enough to obtain the required modal density. The perceived modal density can be improved by modulated one or more of the delays in each branch [3], [5]. The outputs of the late reverb can be taken from the end of each branch, pre-scattering matrix, and pre-scaling and lowpass filter if desired. Alternately, the outputs can be taken from taps within each branch, or after each allpass in the branch. The outputs from each branch can be sent to a single A-format channel, or scattered between the channels as desired.

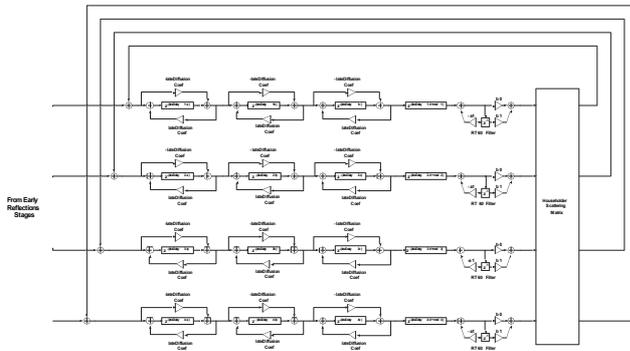


Figure 2: Late reverberation network

5.1. Bringing late reverb into B-format

Tapping out late reverb as described, the resulting A-format should be as decorrelated as the network adapted is able to provide. We can choose the A-B matrix (2) previously discussed, or other tetrahedral orientations are possible. The authors prefer the variant below (8) which places the vertices of the tetrahedron at front-left, front-right, back-up and back-down.

$$\begin{bmatrix} \sqrt{6}/4 & \sqrt{6}/4 & \sqrt{6}/4 & \sqrt{6}/4 \\ 1/2 & 1/2 & -1/2 & -1/2 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 & 0 \\ 0 & 0 & 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \quad (8)$$

As one would expect, it is possible to further steer the late reverberation as one chooses with the dominance transform (9) [12]. The forward scaling is λ , which can be represented as the forward gain of the soundfield, $g_{forward}$, in dB. (10)

$$\begin{bmatrix} \frac{1}{2}\left(\lambda + \frac{1}{\lambda}\right) & \frac{1}{\sqrt{8}}\left(\lambda - \frac{1}{\lambda}\right) & 0 & 0 \\ \frac{1}{\sqrt{2}}\left(\lambda - \frac{1}{\lambda}\right) & \frac{1}{2}\left(\lambda + \frac{1}{\lambda}\right) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (9)$$

$$\lambda = 10^{g_{forward}/20} \quad (10)$$

In application, if a concert hall reverb is desired, it may be suitable to steer the late reverberation the rear of the soundfield. To do so, one would choose a negative value for $g_{forward}$: -3 dB reduces the gain at the front of the late reverberant soundfield by 3 dB and increases the gain at the back by 3 dB, giving a 6 dB difference between front and back.

Combining this approach with the late reverberation and the steering of the early reflections discussed previously, with care, can result in a convincing hall.

5.2. Higher order approaches for late reverberation

An alternate approach for the late reverberation is to use a higher order feedback delay network, where each A-format channel output is taken as a sum of the outputs of some of the delays in the network. Puckette includes a 16-channel FDN in the Pd distribution. This network can be adapted to A-format by dividing the delays into groups of four, summing the outputs from those delays, and using the summed outputs as the late reverberation outputs for the respective A-format channels. The feedback paths can remain the same, and the outputs taken pre-scattering matrix. An example of this approach can be found in [13].

6 CONCLUSION

This paper presented an approach to adapting known artificial reverberation networks to a B-format context, creating native Ambisonic reverberators. Resulting reverberators both input and output B-format, distribute signal energy appropriately through the resulting reverberant soundfield, and take advantage of the ergonomics Ambisonic imaging techniques offer for shaping the spatial impression of a soundfield.

A stereo network was adapted as an example. Similarly a variety of known stereo and multi-channel reverberators may be adapted to create a number of native B-format reverberators with varying characteristics.

7 ACKNOWLEDGEMENTS

The authors would like to thank Dave Malham for initial introductions, discussions and inspirations.

REFERENCES

- [1] J. Anderson, Epiphanie Sequence, Sargasso SCD28056, 2008. Audio CD.
- [2] M.R. Schroeder, "Natural-sounding artificial reverberation," *J. Audio Eng. Soc.*, 10(3), 1962, 219-233.
- [3] J. Stautner and M. Puckette, "Designing multichannel reverberators," *Comput. Music J.*, 6(2), 1982, 52-65.
- [4] J.-M. Jot and A. Chaigne, "Digital delay networks for designing artificial reverberators," *Proc. 90th Conv. Audio Eng. Soc.*, 1991, preprint 3030.
- [5] J. Dattorro, "Effect Design, Part 1: Reverberator and Other Filters", *J. Audio Eng. Soc.*, 45(9), 1997, 660-684.
- [6] M. Gerzon, "Synthetic Stereo Reverberation, Parts I and II", Part 1: *Studio Sound*, 13, Dec. 1971, 632-635, Part 2: *Studio Sound*, vol. 14, Jan. 1972, 24-28.
- [7] M. Puckette, "Artificial Reverberation", *The Theory and Technique of Electronic Music*, 2006. Online. Available: <http://crca.ucsd.edu/~msp/techniques/latest/book-html/node111.html> [Accessed: Mar. 2009]
- [8] R. Elen, "Whatever happened to Ambisonics?", *Audio Media*, Nov. 1991, 50-54.
- [9] M. Puckette, "Power conservation and complex delay networks", *The Theory and Technique of Electronic Music*, 2006. Online. Available: <http://crca.ucsd.edu/~msp/techniques/latest/book-html/node110.html> [Accessed: Mar. 2009]
- [10] D. Rocchesso, "Maximally Diffusive Yet Efficient Feedback Delay Networks for Artificial Reverberation", *IEEE Signal Processing Letters*, 4(9), 1997, 252-255.
- [11] D. Rocchesso and J.O. Smith, "Circulant and Elliptic Feedback Delay Networks for Artificial Reverberation", *IEEE Transactions on Speech and Audio Processing*, 5(1), 1997, 51-63.
- [12] P.S. Cottrell, "On the Theory of the Second-Order Soundfield Microphone", Ph.D. Thesis, University of Reading, 2002, 105-107, 170-174.
- [13] S. Costello, "B-Format Reverb", *DXARTS 567 / Sound in Space*, Online. Available: <http://www.dxarts.washington.edu/courses/567/08WIN/BFRev.rtf>. [Accessed: Jun. 2009]