A CONCEPT-BASED MODEL FOR THE LIVE DIFFUSION OF SOUND VIA MULTIPLE LOUDSPEAKERS

Dr James Mooney
Newcastle University
Culture Lab

Dr David Moore
University of Sheffield
Department of Music

ABSTRACT

This paper presents a conceptual framework for sound diffusion: the process of presenting multiple channels of audio to an audience in a live performance context, via loudspeakers. Terminology that allows us to concisely describe the task of sound diffusion is defined. The conceptual model is described using this terminology. The model allows audio channels (sources) and loudspeakers (destinations) to be grouped logically, which, in turn, allows for sophisticated abstract methods of control that supercede the restrictive 'one-fader-one-loudspeaker' approach. The Resound project – an open source software initiative conceived to implement and further develop the conceptual model – is introduced. The aim is, through further theoretical and practice-led research into the conceptual model and software respectively, to address the technical, logistical and aesthetic issues inherent in the process of sound diffusion. *Keywords* – Sound Diffusion, Spatialisation, Software, Resound.

1. INTRODUCTION

Sound diffusion involves the reproduction of Audio Streams, via loudspeakers, to an audience. A common application is in the performance of electroacoustic music from CD, but the principles are relevant in a diverse range of activities including theatre sound, cinema, virtual reality, and indeed any scenario where sophisticated real time control over the auditory environment is required. Where references are made to the performance of electroacoustic music, it is to be understood that the same principles could apply in a variety of sound diffusion scenarios.

Audio Streams can originate from any combination of a wide variety of sources, including recording and playback devices (e.g. CD players, audio sequencers), microphones and pickups, synthesis technologies and audio processing technologies, any of which may be implemented in hardware or software [1, p17-20]. The Audio Streams are reproduced via a loudspeaker array comprising a potentially large number of loudspeakers. A mixing device (often an audio mixing desk but, increasingly, a software implementation) acts as an intermediary between audio sources and loudspeakers, along with an interface that allows the diffuser (performer) to control the auditory results.

This research focuses mainly on the issues inherent in the *real-time* diffusion of multiple Audio Streams. The ergonomics of the interface used to control the diffusion are therefore of key importance. Furthermore, because the specific technical demands of electroacoustic works – in terms of audio sources, the number of channels used, and so on – can vary considerably from composition to composition within the context of a single concert programme, a dynamic and configurable audio routing architecture is required.

The authors have previously developed the M2 Sound Diffusion System, whose design and implementation has been thoroughly documented elsewhere [1, 2, 3]. The prototype system has regularly facilitated live performances since its inaugural performance in March 2004 and, at the time of writing, continues to serve as the principal sound diffusion system at the University of Sheffield Sound Studios. The M2 system demonstrated some significant advantages: it is portable quick to set up; signal routings can be re-configured comparatively quickly and easily; the interface supplements familiar fader-based diffusion paradigms rather than replacing them with unfamiliar methods. Three years of practical experience with the system – as well as highlighting the benefits of the system – has revealed some unforeseen issues [1, p247-265]. It is clear that significant improvements are still to be made in the technology of sound diffusion, and that future systems will benefit from a thorough conceptual reappraisal of what, exactly, sound diffusion involves.

2. AUDIO STREAMS

For the purposes of this article, an Audio Stream is defined as any non-symbolic representation of audio, analogue or digital, including that which is recorded, synthesized or transduced via microphones or pickups. The continuously varying voltage generated by a microphone is an example of an Audio Stream. It is a directly analogous representation, obtained via a process of transduction, of the fluctuations in air pressure that constituted the original auditory event. It is not symbolic, unlike, for example, western classical notation or Morse code: these are abstract codified systems that need to be deciphered according to predetermined schemes. Importantly, the Audio Stream is one, and only one, single channel of audio: a stereophonic recording therefore comprises two Audio Streams. (Clearly the two Streams are related; we will return to this later.) We identify the Audio Stream as an irreducible unit in the context of sound diffusion.

3. COHERENT AUDIO STREAM SETS

Multiple Audio Streams are very often used concurrently to encode spatial information. Two-channel stereophony [4] involves the concurrent use of two Audio Streams to encode spatial information with respect to a single (nominally horizontal) axis. In principle, this concept can be extended to any number of Audio Streams. A Coherent Audio Stream Set (CASS) is defined as a group comprising an arbitrary number of Audio Streams that are co-dependent on account of having been used collectively to encode spatial information in this manner. Some examples

Email: james.mooney@newcastle.ac.uk, d.moore@shef.ac.uk

are given in Figure 1. Collectively, the constituent Streams of the CASS represent a single, homogeneous spatio-auditory image. Given an appropriate loudspeaker configuration it is possible to reproduce the spatio-auditory image embedded within a CASS by simply mapping the constituent Audio Streams directly to loudspeakers. This distinguishes the CASS from higher order encodings such as Ambisonics and Dolby's AC3 encoding, which require preliminary decoding in order to obtain the constituent Audio Streams before reproduction over loudspeakers is possible.

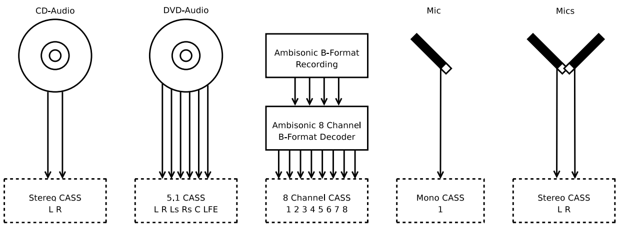


Figure 1. Coherent Audio Stream Sets (CASS)

4. COHERENT LOUDSPEAKER SETS

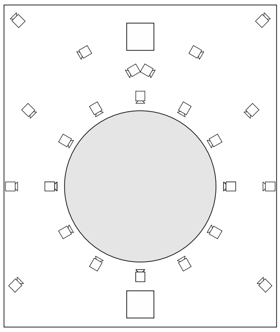
A Coherent Loudspeaker Set (CLS) is a group of loudspeakers arranged in such a way that it can be used to correctly broadcast a Coherent Audio Stream Set. For example, two loudspeakers arranged according to the normal stereophonic convention can be used to correctly reproduce a stereophonic CASS. (What constitutes "correct reproduction" in this context is, clearly, open to interpretation.) Similarly, five loudspeakers and a sub-woofer deployed according to the accepted convention for 5.1, can be used to correctly play back a 5.1 channel recording. A CLS can comprise any combination of loudspeakers from the array (a loudspeaker "array" connotes all of the loudspeakers in the system) that are deemed appropriate for the reproduction of a given CASS. Thus, an array can contain many more CLSes than there are loudspeakers, and any given loudspeaker could conceivably be a member of several different CLSes. Some arbitrary examples are given in Figure 2.

5. REDEFINING SOUND DIFFUSION

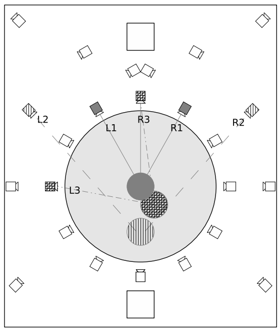
Sound diffusion can now be defined as a practice that involves the real time broadcast of one or more Coherent Audio Stream Sets via one or more Coherent Loudspeaker Sets. Nowadays it is not at all uncommon for electroacoustic works to comprise multiple CASSes comprising variable numbers of Audio Streams. Consider Jonty Harrison's eight-channel work *Rock 'n' Roll*, described by the composer as follows:

> I used a six-channel hexagonal [loudspeaker] array as a surround,' and I had two solo speakers – the mains if you like [...] – for close up stuff. What I would like to do [...] is be able to diffuse the twos and the sixes completely independently.[5]

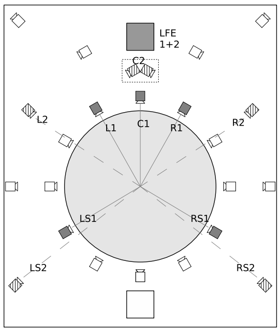
What Harrison describes is an eight channel work comprising two Coherent Audio Stream Sets: one six-channel Set and one two-channel Set. To be able to diffuse the two Sets independently will clearly depend on a flexible mix architecture and an equally flexible user interface.



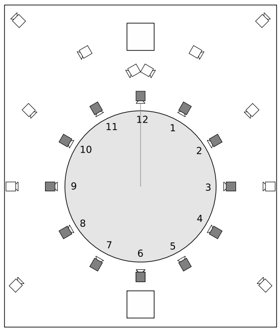
(a) Example loudspeaker array.



(b) Three potential stereo CLS mappings: L1-R1, L2-R2 and L3-R3.



(c) Two potential 5.1 CLS mappings.



(d) 12 Channel CLS circular mapping with sequencial "clock face" ordering.

Figure 2. Coherent Loudspeaker Sets (CLS). A selection of CLS groupings are mapped onto the loudspeaker array.

6. MATRIX MIXING

Ultimately, sound diffusion concerns the dynamic distribution of Audio Streams from input channels to output channels. A flexible way of achieving this in practice is with a mix matrix, an example of which is illustrated in Figure 3. Input busses are shown as horizontal lines, output busses as vertical lines. The points at which these lines intersect are referred to as input-to-output nodes (or cross-points); these represent potential input-to-output buss routings. Each input-to-output node has an attenuator, as opposed to simply being switched. In addition to the input-to-output nodes we also have the usual input nodes output nodes for each channel.

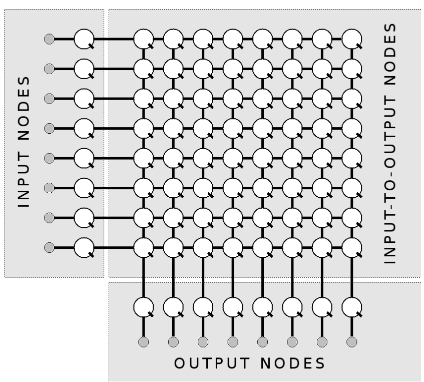


Figure 3. A simple 8-in-8-out Matrix mixer schematic

Switched-buss mix architectures (as found in most conventional mixing desks) necessitate fixed routing of inputs to outputs, and allow us simple control over buss input and output levels only. In sound diffusion this can be limiting, especially where multiple and/or greater-than-stereo CASSes are involved. The mix matrix architecture, on the other hand, allows us to dynamically mix any input signal to any output channel in any proportion. Clearly, this is very useful in the context of sound diffusion. Accordingly, the mix matrix is increasingly frequently implemented in sound diffusion systems, and forms the underlying architecture of, amongst others, the DM-8 system [6], Richmond's SoundMan system [7], and the M2 Sound Diffusion System, which we will return to later. Recent developments at the University of Birmingham have also explored the matrix architecture [8]. It has additionally been noted that matrix nodes can be used to accommodate extended signal processing, resulting in further increased flexibility [2, p36-38].

In summary, the mix matrix provides us with a multitude of parameters through which we can diffuse Audio Streams. The complexity involved in the individual control of such large numbers of parameters, however, becomes a concern.

7. PHYSICAL INTERFACE DEVICES AND ABSTRACTED VERSUS DIRECT CONTROL

The expression Physical Interface Device will be used to describe any hardware means of user input into a system. Single axis slide potentiometers – faders – are the most common Physical Interface Devices in sound diffusion. A fader affords us direct control over the value of a single matrix parameter only. A simple eight-in-eight-out attenuation matrix (as shown in Figure 3) has eighty nodes – eight for the inputs, eight for the outputs, and sixty-four for the input-to-output nodes – and would therefore require eighty faders. Clearly, then, a direct one-to-one mapping of faders to parameters is problematic. If the matrix architecture is to be useful in the context of sound diffusion it is necessary to have abstracted control over the matrix parameters, whereby the fader manipulates multiple parameters indirectly via some predefined algorithm. If a fader is to afford us abstracted control over multiple parameters in this way, then we need to be able to specify – firstly – which parameters we want to control and – secondly – precisely what kind of control we want to have.

8. COLLECTIVES

A Collective is used to specify which mix matrix parameters (nodes) a Physical Interface Device will control. A Collective is, essentially, a group of parameters in a particular order. More specifically, it is a set of elements in a specific order, where each element can contain one or more matrix parameters. This is illustrated in Figure 4. In programming terminology a Collective is simply a two-dimensional array of parameter addresses. A Coherent Audio Stream Set can, therefore, be expressed as a Collective whose elements can only be matrix input nodes. A Coherent Loudspeaker Set – at least as far as the routing of inputs to outputs is concerned – can be expressed as a Collective whose elements consist only of matrix output nodes. In terms of how Parameters are addressed, a useful system – whereby addresses are given in URL form – is utilised in the OSC protocol [9]. We might, for example, describe an input-to-output node's attenuation Parameter address as follows: /matrix/in1/out1/att.

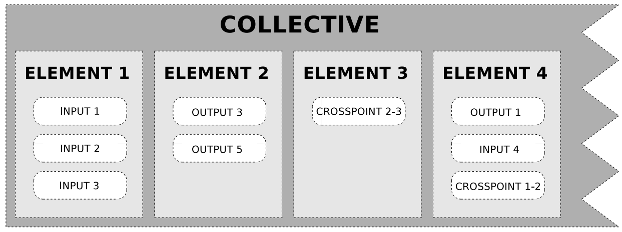


Figure 4. Example showing the structure of a Collective

9. BEHAVIOURS

A Behaviour determines precisely how the parameters of a Collective will be acted upon. More accurately, a Behaviour has one or more of its own input parameters and uses these to manipulate the parameters contained within a designated Collective in a specific way. A Behaviour's own parameters can be assigned to Physical Interface Devices, thus allowing interactive control.

As a very simple example, consider a Behaviour called 'Group.' The 'Group' Behaviour has one parameter of its own – Level – direct control of which has been assigned to Fader *A*. The Behaviour also has a Collective assigned to it, which contains a number of mix matrix parameters. When the value of Fader *A* changes, its value is mapped on to all of the parameters of the Collective, and thus the fader behaves like a 'group.'

In a more sophisticated example, a Behaviour may be designed to make use of the ordered nature of the Collective in some logical manner. A 'Chase' or 'Sequence' Behaviour could be used to iterate through the elements of a Collective in turn, resulting in a semi-automated movement of sound between loudspeakers. In addition to a 'Level' parameter it might also have a 'Frequency' parameter, to determine the rate at which the chase sequence would iterate. In principle, any number of different Behaviours could be defined that interact algorithmically with matrix parameters in various ways.

10. PARAMETER SUMMING PRINCIPLE

Consider the possibility of individual matrix Parameters being under the simultaneous control of multiple Behaviours: which Behaviour has final control over the Parameter? Of course we could decide that multiplicitous control is not allowed, but this is limiting. Another solution would be to allow the last accessing Behaviour to set absolute value, however, this is problematic when Behaviours all act at the same time. What if we want one Interface Device to act on the full range of a Parameter and another to provide fine control over a small range? This example requires the Parameter to have two Behaviours operating on it: 'coarse' and 'fine,' if you like. This kind of action can be achieved by adopting a scheme whereby the influences of all of the Behaviours affecting a Parameter are summed together to produce the final absolute value. A similar parameter summing concept was employed in the M2 Diffusion System and is explained in greater detail elsewhere [2].

11. CLASSIFICATION OF CASS AND CLS ELEMENTS

An advantage of software-based systems is that input-to-output routing configurations can be saved to disk and re-loaded at a later time. In theory, this should save time when setting up performances. However, loudspeaker arrays can vary considerably from performance to performance in terms of the number and relative positioning of loudspeakers. The routings for one setup may not map directly on to those of another, meaning that the system will have to be configured from scratch.

Consider the following scenario. As part of our diffusion configuration we have a fader to control the level of a stereo CASS sent equally to two stereo CLSes: 'rear' and 'distant rear.' When we raise the fader fully the stereo source emanates from both loudspeaker pairs. At a later date, we perform the same work in a different venue, with a slightly different loudspeaker array, which only has one pair of loudspeakers at the rear of the hall. We now want the fader to diffuse to one pair of 'rear' loudspeakers rather than two. We are clearly dealing with closely related scenarios yet, frustratingly, there is no way to map between the old configuration and the new, as the former contains references to audio outputs that are either absent, or routed to different loudspeakers, in the new configuration.

In order to make the mapping automatically, two things are necessary. Firstly, we need a system whereby loudspeakers can be identified according to their role rather than by which audio outputs they are connected to. This way loudspeaker pairs could be labelled 'rear,' 'distant rear,' and so on. Secondly, we need to be able to identify 'rear' and 'distant rear' as functionally related. More specifically, 'distant rear' could be defined as a hierarchical subdivision of 'rear.' This process has been described more fully elsewhere [1, p278-80].

Such labelling could also be applied to CASS elements. (Recall that in our conceptual model CASSes and CLSes are both simply special instances of parameter Collectives.) This means that mappings between CASSes and CLSes could take place semi- or fully automatically if required. The left channel of a stereophonic CASS – let's call this Audio Stream *A* – might be associated with the labels 'Mono' and 'Left.' Two loudspeakers within a CLS – Loudspeakers *x* and *y* – might be classified as 'Mono,' 'Left' and 'Mono,' 'Right,' respectively. If asked to identify a suitable automatic mapping for Audio Stream *A*, we would choose Loudspeaker *x*, because its identifiers match more closely the labels of the Audio Stream *A* than do those of Loudspeaker *y*.

In effect, what we are describing is a hierarchical structure of abstract metadata identifiers.

12. RESOUND

In order to realise the conceptual model and develop it further, the authors have founded the Resound project [10]. This is an open-source software project licensed under the GNU Public License (GPL). The software comprises a server component, which amongst other things implements the mix architecture and parameter addressing paradigms described, and a graphical client, which deals with user input as well as facilitating the manipulation of Collectives and Behaviours.

13. CONCLUSION

In summary, the conceptual model of sound diffusion is as follows. The 'raw materials' are Audio Streams; these are grouped in to Coherent Audio Stream Sets (CASSes). The CASSes are then diffused, via a process of abstracted control over a mix architecture comprising a matrix of Parameters, to an array of loudspeakers comprising multiple Coherent Loudspeaker Sets (CLSes). Abstracted control is achieved by defining two things: a two-dimensional array of target matrix parameters (a Collective), and the way in which these parameters behave (a Behaviour). A Collective-Behaviour combination can then be bound to a Physical Interface Device (e.g. a fader) for user control, and the process of mapping Collective-Behaviours can be repeated for as many faders as there are available. Conflict between multiple Behaviours simultaneously acting on the same Parameters is resolved via parameter summing within a fixed range. The model is further refined by way of a hierarchical structure of abstract metadata identifiers for parameters, which can be used for semi-automated CASS-to-CLS routings and to increase interoperability between differing loudspeaker arrays.

14. REFERENCES

[1] J. Mooney. *Sound Diffusion Systems for the Live Performance of Electroacoustic Music: An Inclusive Approach led by Technological and Aesthetical Consideration of the Electroacoustic Idiom and an Evaluation of Existing Systems.* PhD thesis, University of Sheffield, 2005. Online at http://www.james-mooney.co.uk/research/sound_diffusion.

[2] D. Moore. *Real-time Sound Spatialization, Software Design and Implementation.* PhD thesis, University of Sheffield, 2004. Online at http://music.dept.shef.ac.uk/phd_thesis/rtss.pdf.

[3] A. Moore, D. Moore, and J. Mooney. M2 Diffusion The Live Diffusion of Sound in Space. In *Procedings of the 2004 ICMC International Computer Music Conference*, pages 317–320., Miami FL., November 2004.

[4] A. Blumlein. Improvements in and relating to Sound-transmission, Sound-recording and Sound-reproducing Systems. International Pat. No. 394325, 1931. Online at http://www.doramusic.com/patents/394325.htm (17th June 2007).

[5] J. Mooney, editor. *Sound Diffusion Systems for the Live Performance of Electroacoustic Music: An Inclusive Approach led by Technological and Aesthetical Consideration of the Electroacoustic Idiom and an Evaluation of Existing Systems*, chapter Interview with Jonty Harrison. 2005.

[6] C. Rolfe. A practical guide to diffusion. 1999. Online at http://cec.concordia.ca/econtact/Diffusion/pracdiff.htm (20th June 2007).

[7] Richmond Sound Design. Soundman server theatre sound design / audio sound control software. 2007. Online at http://www.richmondsounddesign.com/sm-sspecs.html (20th June 2007).

[8] S. Wilson and J. Harrison. Introduction to beast's diffusion research project. 2007. Presentation given at CBSO Centre, Birmingham, 27th May 2007. Reference online at http://www.ea-studios.bham.ac.uk/BEAST/concerts/2007.html.

[9] M. Wright and A. Freed. Open sound control: A new protocol for communicating with sound synthesizers. In *Procedings of the 1997 ICMC International Computer Music Conference*, pages 317–320., Thessaloniki, November 1997.

[10] J. Mooney and D. Moore. Resound project. Online at http://resound.sourceforge.net (19th June 2007).