Concrete Abstractions: Reflections on Sound Recording for Immersive Acousmatic Composition

David Rylands
d.rylands@auckland.ac.nz

Abstract

This paper documents sound recording approaches for immersive multiple loudspeaker acousmatic composition. The notion that spectrum and spatiality are inseparable is supported by perspective-based recording approaches (*foreground, middleground* and *background*) that serve to simplify real-world sound environments. A *hybrid* approach is proposed as findings suggest that it is not possible to limit recording, and in turn re-presentation, to one optimal method.

It is commonly agreed that the phonographic inventions of the late 19th century greatly accelerated developments in sound-based art forms. The inscription of sound, first associated with the acoustic etchings of the phonograph, would preserve and objectify a sonic event that for the first time could be recreated 'without being uttered' [1]. However, the disproportionate sound image and poor quality of early recording and transmission devices lacked the ability to capture and re-present the sonic content with sufficient detail. These fatal technical flaws (amongst others) were criticised by Theodor Adorno who believed the loss of 'symphonic space' through radio transmission to be detrimental to music reception, as it led to a degradation of the dynamic, timbral and spatial nuance crucial to the compositional process [2]. With drastic improvements in audio fidelity, it is apparent that many of the technical flaws inherent to the earlier technologies may have been surpassed. It is now a priority for certain artists to re-establish listening approaches that concentrate on the structural unfolding of sound within an enveloping loudspeaker space.

Sound recording plays a central role in acousmatic art making. The act *dislocates* sound from its physical origin in space and time, allowing one to contemplate the meaning of the data carried and the aesthetic elements of the fixed and unresponsive recorded event. When related to sound, a dictionary definition of *disembodiment* may act as an adequate departure point for an understanding of the acousmatic listening condition. That is, 1. separated from or existing without the body, 2. (of a sound) lacking any obvious physical source [3]. Through the act of dislocation, the once ephemeral nature of a sound is separated from its visible source-cause and established as a repeatable and manipulable sound-object in itself. When used in art practice, preserved sound may take on an independent life as a poetic image, given to one like that from words of a book. It may be that the desire to preserve sound emerges from a recordings capacity to etch for us a reminder of life's transient nature, a *memento mori* that, as Antonin Artuad alludes to, 'embodies and intensifies the underlying brutalities of life to recreate the thrill of experience'[4].

Sound recordings are documental of experience, where the microphone functions as either an inquisitive instrument, or a passive detached observer, informed by a spatial continuum from intimacy to immensity [5], [6]. This continuum suggests that close sounds embody a different psychological trace to that of distant sounds, and visa-versa. The functionality of this continuum also implies that spectral information and spatial cues are perceived as inseparable. For instance, high frequency *texture* is a primary cue for the individuation and localisation of a sound as it provides information about the place in which a sounding object finds itself, and the proximity of its position. Further to this, Denis Smalley [6] states that the spectral content and shape of sounds in themselves suggest spatiality, thus the manipulation of spectro-morphological change can assist to reanimate spatial morphology with significance. In other words, spatial articulation, to a large degree, is determined by a sounds spectro-morphology, and it is by traversing to-and-fro along the spatial continuum that much of the structural interest and creative expression is to be enjoyed when working with recorded sound.

A wide variety of sources and events have been recorded for the creative project that informs this paper. Through extensive experimentation, recordings were culled to a select few, upon which the majority of the findings are based. The main sound sources include windmills, wind blown corrugated sheds, printing press machinery, hand written gestures on varying surfaces and in differing spaces, the white sound of touch, and spoken text. Non-specific spatial settings were also recorded for use as peripheral backdrops. The broad artistic intent for the project is to explore both, the extensions of inner impulses into expressive physical form, and the impressions made by physical form on ones inner being. The theme for the accompanying compositional work is based on the problematic nature of the materialisation of thought into word (as below), and the dual power words can carry as objects to wound and debilitate, or to enable healing through compassion and personal catharsis.

> *"That was when I learned words are no good; that words don't ever fit even what they are trying to say at."* [7]

The majority of recording for this project was done out in the field. This places certain constraints on the sound recordist and tends to dictate the types of approaches possible. The limitations often mean that equipment needs to be suitable for efficient set up and ease in transportation, as too much equipment can lead to overwhelming frustration in application. This has led to the use of more compact recording devices and microphone configurations. In practice, the most versatile equipment has proven to be two and four channel portable hard-disc recorders, as they are quick to set up

and can interface with a number of different microphone input combinations (e.g. contact, minature omni, dynamic or condenser). For more complex configurations a computer with external audio hardware can be sought.

It is sensible to state that recording and re-presentation for immersive composition cannot be summed into one optimal, universal method. Instead, suitable *hybrid* recording approaches (and in turn, a *hybrid* re-presentation systems) using mono, stereo and ambisonic formats are necessary to cope with diverse sound entities and event locales. For example, combinations for this project were devised as follows:

Windmill – multiple contact, multiple stereo microphone perspectives (stereo and 4 channel)
Windy shed – multiple contact, spaced microphones, ambisonic B-format (stereo and 4 channel)
Print press – multiple contact, stereo and spaced microphones (computer interface)
Writing box – multiple contact, stereo and mono microphones (computer interface)
Gestural writing – mono contact, mono microphone (stereo)
Spatial backdrop – mono, stereo, and ambisonic microphone configurations (stereo and 4 channel)
Voice – mono microphone

Whilst recording in the field, successful application requires considerable foresight as monitoring is usually limited to headphones, or sometimes nothing at all. Continuous reflection on the recording approaches used is necessary to review both the realisation of recording intent, and the suitability of techniques and equipment. That said, certain recording approaches will generally embody sonic information in predictably ways:

Monophonic (point source)
    Is not spatially diffuse
    Spectral texture carries all spatial cues
Stereophonic (standard or multiple-stereo)
    Gives a precise and stable frontal image
    Accurate localisation
    Can be perceived as cluttered
    Lacks immersion and envelopment
Ambisonic (A and B-Format)
    Very immersive and enveloping
    Gives good rendering of ambience
    Poor localisation
    Poor image accuracy
    From a central perspective listening outwards

A *hybrid* approach towards the use of varying recording techniques has been recommended, as it helps to efficiently simplify and separate an environment into differing perspectives consisting of *foreground things* (sources)*, middleground features* (spatial relationships)*,* and *background place* (ambience character) [8]. The application of this reductive approach can capture much of a scenes plausibility, and by paying equal attention to the rendering of these three perspectives, it is possible to convey greater depth without overloading the perceptual foreground. The *Foreground* perspective concerns proximate and dynamic sources that are demanding of attention, and could be recorded with contact or intimate close microphones. This approach rejects much of the place ambience and sound bleeding from other close source material. Recording of the *middleground* functions to give a readable interpretation of the relationship between sources within a scene. An obvious example is in music ensemble recording where a relatively close stereo configuration is used. Place ambience is partially rejected. Lastly, an ambient microphone array can be used to capture details about the place where sounds find themselves. The listener is usually not attentive to *background* information, however, without a suitable backdrop, the more proximate sounds are missing a context essential for the synthesis of a coherent and believable sound field. This peripheral perspective can be successfully attained through ambisonic or spaced recording techniques. Overall, what is extracted is a simplified sonic image containing a variety of sounding perspectives that naturally exist along the intimacy to immensity spatial continuum.

Extant research on the assimilation of recorded material into immersive sound fields suggests that absolute spatial accuracy is both technologically unfeasible and perceptually unnecessary for plausible immersive results. This is because in everyday life we often deal with ambiguous and incomplete sound environments, and make sense by focusing our attention more fully on the proximate cues that seek to call us into action. This indicates that it is possible to efficiently simplify complex environments into unambiguous sound fields by synthesising the most perceptually significant information [9]. The main aim of reducing information is so one can signify intuitively comprehensive sound fields. For this project, *hybrid* sound recording approaches have been proposed as the initial step in simplifying the massive informational bandwidth of real-world sound environments. This reduction of information may been likened to the functionality of a cartoon sketch, where unnecessary information is stripped away in order to accurately deliver, and often exaggerate, relevant information. As well as being used to inform sound recording, *cartoonification* [10] directs audio processing approaches in re-presentation. For instance, the recording and re-presentation of a distant sound that requires little aural attention should not need to be defined by accurate location cues, proximity information, or signal quality.

Once playback is experienced, it is reasonable to suggest that recording and re-presentation directly inform one another. Reflective listening will show what recording techniques give the desired results, and the retained details in a recording

will carry cues suggestive of certain processes for re-presentation. If one is to work effectively with a diverse range of recorded information types, a *hybrid* re-presentation system is necessary. The spatialisation techniques used in this project were reviewed within an immersive 8-loudspeaker setting, and based on applications developed for max/MSP. These are Vector Based Amplitude Panning (VBAP) [11], Multiple-Direction Amplitude Panning (MDAP) [12], Distance Based Amplitude Panning (DBAP) [13], Ambisonic Equivalent Panning (AEP) [14], Ambisonics [15], Ircam Spat Library [16], and 8space Audio Spatialisation Matrix designed by Timothy Place. However, as suggested throughout this paper, more than angular separation is needed to create plausible virtual sound fields, as the control of proximity and ambience cues contributes most effectively to the appreciation and interpretation of location, movement and place. To satisfy this requirement, the spatialisation applications mentioned have been used in tandem with various amplitude, spectral and ambience processing techniques.

The scope of this paper allows for only a brief description of how recordings may be nested within an immersive loudspeaker setting. By using the *hybrid* recording and re-presentation techniques proposed, a way to compile material might be to create a 'sonic landscape' [17] that combines various perspectives so to present a diverse display featuring information ranging across the intimacy to immensity continuum. For example, a *foreground* recording of the gesture of writing captured with contact microphones could have subtly temporal manipulations applied to amplitude, pitch and frequency response, as well as the mapping of tangible virtual movement within the loudspeaker setting. A *middleground* perspective could be given by setting a scene in which the act of writing is taking place. This may be done by overlaying, surrounding and 'pushing back' (with spectral manipulation and subtle reverberation) a recording taken inside the small wind blown shed where close spatial details (which could continually twist from side to side) are provided by the shaking movements heard from inside the enclosed space. The *background* place could be portrayed by providing information perceived beyond the walls of the space in which the event occurs. Processing the spectrum and ambience of any combination of source or spatial recordings could create a distant setting. This may include the distant clanking of the windmill mixed with an ambisonic night landscape recording.

In conclusion, it is understood that perception logically requires *information reduction*, which is achieved by concentrating on particular elements at the expense of others. In this project, sound recording is devised as the initial step in the reduction process, extracting (and often exaggerating) important information from the *foreground, middleground* and *background* of a sounding environment. The extracted information contained in a sound recording directly informs further spectral and spatial processing so it can be plausibly nested within an artificial immersive sound field. For flexibility in recording and immersive re-presentation, a *hybrid* approach has been proposed.

The knowledge gained from this study has had a significant impact on a number of related projects. Firstly, extensive testing and development of various immersive spectral and spatial studio composition approaches have been conducted alongside this study, contributing greatly to the design and installation of a 24-loudspeaker large-scale immersive diffusion system for the New Zealand Electroacoustic Music Symposium in September 2009. This project was realised in collaboration with colleagues John Cousins and John Coulter. A related subject deserving greater depth in study is whether recorded sound can be considered as a transparent medium where sounds exist in themselves, or whether dislocated sound (no matter how abstract) inevitably embodies the representational remains of its past life. And finally, the completion of the 8-channel work that has given all the necessary experience for this paper is due for February 2010.

References

[1] Kahn, D., Whitehead, G. 1992. *Wireless Imagination: Sound, Radio, and the Avant-Garde*. Cambridge: MIT Press.
[2] Adorno, T.W. 2002. "The Radio Symphony", pp. 251–69 in Leppert, R (ed.) *Essays on Music*. Berkeley, CA: University of California Press.
[3] Hobson, A (ed.). 2002. *The Oxford Dictionary of Difficult Words*. New York: Oxford University Press.
[4] Jamieson, L. 2007. *Antonin Artaud: From Theory to Practice*. London: Greenwich Exchange. pp.21
[5] Bachelard, G.1994. *The Poetics of Space*, tr. M. Jolas. Boston: Beacon Press.
[6] Smalley, D. 1991. "Spatial experience in electro-acoustic music", pp. 121-124 in Dhomont, F. (ed.), *L'Espace du Son II*. Ohain: Musique et Recherches.
[7] Faulkner, W. 1970. *As I Lay Dying*. London: Chatto and Windus Press. Pp. 171
[8] Lennox, P.P. 2004. *Spatial Music and Spatial Perception in Artificial Environments*. D.Phil Thesis. JB Morrell Library, University of York, UK, pp. 246.
[9] D. G. Malham, "Approaches to Spatialisation", in *Organized Sound*. Vol 3, issue 2.
[10] Lennox, P., Myatt, A., Vaughan, J. 2001. "3d Audio As An Information-Environment: Manipulating Perceptual Significance For Differentiation And Pre-Selection", ICAD *Proceedings of the 7th International Conference on Auditory Display*. Espoo, Finland.
[11] V. Pulkki. 1997. "Virtual sound source positioning using vector base amplitude panning", in *Journal of the Audio Engineering Society*. Issue 45(6).
[12] V. Pulkki. 1999. "Uniform spreading of amplitude panned virtual sources", in *Proceedings of the 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. Mohonk Mountain House, New Paltz, New York.
[13] Lossius,T., Baltazar, P., de la Hogue, T. 2009. "DBAP - Distance Based Amplitude Panning", in *Proceedings of 2009 International Computer Music Conference*, Montreal, Canada.
[14] Neukom M., Schacher J. 2008. "Ambisonics equivalent panning," in *Proceedings of the 2008 International Computer Music Conference*. Belfast, UK.

[15] Schacher, J. C. Kocher, P. 2006. "Ambisonics Spatialisation Tools for Max/MSP", in *Proceedings of the 2006 International Computer Music Conference*, New Orleans, US.

[16] Jot, J.M., Caulkins, T. 2006. *Spat Reference Manual*. http://support.ircam.fr/forum-ol-doc/spat/3.0/spat-3-ref/co/spat-3.html (visited 9, December, 2009)

[17] Wishart, T. 1996. *On Sonic Art*. Amsterdam: Hardwood Academic Publishers. pp. 147