MORPHING AURAL SPACES IN 7.1 MUSIC

Matt Hitchcock
Queensland Conservatorium Griffith University
Music Technology Dept.

ABSTRACT
This paper outlines and discusses a method for improving interactivity in using space (aural and musical) as a compositional device in surround composition. The relative merits of 7.1 over 5.1 in this context are summarised. The question is asked if ‘space’ were granted equal compositional importance with the sounds themselves, what forms might this take, and how might one negotiate technical limitations that are imposed by digital audio workstations. Examples of technical and creative approaches are given.

1. INTRODUCTION

As a composer and music technologist I have an interest in creating music\(^1\) in surround for broader dissemination than just for those people fortunate enough to experience immersive works at ‘one-off’ special events with advanced technical catering. Once such events are completed, so is the opportunity to experience these works. Consequently, effective broader dissemination of music mixed for surround sound requires utilisation of consumer technologies.

Further, the level of detail required to localise sound has been the subject of psychophysical studies [1, 12, 18, 22, 36] and perceptual studies [6, 9, 24, 28, 29, 31, 33]. It has been proposed that, “at least six loudspeakers (not including LFE) are needed to reproduce the spatial impression of a diffuse sound field” [12]. My experiences in working with surround support these findings, most particularly with respect to the importance of phantom images created between pairwise speakers [8, 20, 21, 23, 25, 29, 32], commonly identified as an issue with 5.1 [6, 9, 12, 24, 28, 29, 31, 33].

There is a wide array of consumer accessible formats and configurations [14, 15, 16, 19], that comprise more than 6 speakers, however the HD 7.1 surround configuration utilised by Dolby\(^\circ\) TrueHD and DTS-HD Master Audio\(^\text{TM}\) is the default for the Blu-ray DVD format. Consequently I suggest that this configuration is most likely to be a ‘default’ configuration and have accordingly chosen this as a surround platform for my research.

The Dolby\(^\circ\) TrueHD and DTS-HD Master Audio\(^\text{TM}\) 7.1 surround formats (Blu-ray) attempt to address the recognised weaknesses of 5.1 and 3/4.1 [6, 9, 12, 24, 28, 29, 31, 33] by using 6 equidistant speakers each separated by 60° (in conjunction with the centre and LFE channels). This configuration therefore allows a consistent set of speaker pairings to occur around the entire 360° arc.

My current research focuses on utilising a series of pairwise phantom images available within the HD 7.1 configuration that exploit psychoacoustic elements of sound localisation and physical attributes (generalised) of HRTF. I am treating the surround speaker array as isotropic for compositional purposes and exploring the manipulation of multiple aural spaces that can themselves be moved around the panoramic field (360°). I employ techniques to move space in conjunction with or separately to the sounds that excite space and (partly inspired by the idea of quantum mechanics) that move sound across space (without moving sound or space). Of particular interest in my work then is the ability to morph sounds and spaces fluidly, whether in tandem or separately, and to assess levels of cognitive dissonance or consonance that occur.

2. BACKGROUND AND CONTEXT.

Working variously as a musician and composer, sound designer and engineer, I am deeply engaged by the concepts of sonic energy and narrative and the musical elements of timbre and dimension\(^2\).

I started working with surround sound in 2002. I experimented with Quadraphonics, Ambisonics and 5.1 before predominantly settling on 5.1. At this time 5.1 was gaining momentum as a publicly accessible (driven by home theatres, later supported by HDTV surround broadcasts) and, therefore, viable platform for the wide dissemination of music as an art form whose medium is sound.

\(^{1}\) While there is no single and intercultural universal concept defining what music might be, within the context of this paper I define music as an art form whose medium is sound.

\(^{2}\) Herein, dimension includes: [1] the ambient field in which the sounds sit (A piece can have multiple ambient fields); [2] the spectral height from low to high frequencies; [3] the perceived distance from the listener; and [4] the panoramic width of each sound itself.
dissemination of surround music and “auralisation” [36].
Commercial releases of musical works were on the rise. Available repertoire included: classical works which were relatively easy to produce in surround, albeit the surround channels commonly for room reflections; re-issues of older albums such as Pink Floyd’s Dark side of the moon, Eric Clapton’s Slowhand, Billy Joel’s Piano Man, Queen’s A night at the Opera, Beach Boys’ Pet Sounds, etc.; and works re-envisioned by the artist or created specifically for surround such as Sting’s Ten Summoner’s Tales, Chick Corea’s Remembering Bud Powell, Studio Voodoo’s Club Voodoo, Dave Grusin’s Two for the road, etc. At the same time, experimental and art music artists were exploring new ground musically and technically [2, 10, 11, 13, 17].

Two fundamentally different aims of rendering spatial audio were apparent: (I) to achieve physiological realism (to emulate or document self-created, biological, architectural or natural acoustic environments); and (II) to create a sonic canvas for surreal (conflicting but apparently real[7]) or completely imaginary forms of creative expression.

This paper focuses on the creation and manipulation of surreal or imaginary spaces for compositional purposes and makes no attempt to reconcile the difficulties inherent in documenting reality or physiological realism. The context being scrutinised is therefore one where the artist is either the architect of the aural space(s), or one where a work uses aural space as part of the musical structure – ergo an aural architect [3].

3. MUSICAL AND AURAL SPACE

The term ‘space’ in music is most commonly used in reference to two concepts, motion or environment. With reference to motion of sound (e.g. melodic shapes), the term space signifies action and non-action. I refer to this usage herein as musical space. The second usage of the term space refers to the sonic characteristics of the location with which the sound interacts, herein referred to as the aural space. Aural space has 4 key domains - width, depth, height and temporality (length, width, depth and height over time).

With reference to aural space, Emmerson [7] proposes the idea of a frame (a defined area of interest), applied progressively from the largest to the smallest scale. The frame comprises a landscape bounded by the acoustic horizon (distance limit from which a listener can hear sonic events) [3, 30, 34]. The landscape then comprises an arena, within which we find a stage, upon which we frame an event as depicted in figure 2 [7].

Figure 2 - Soundfield Frames , after Emmerson [7]

The aural architect, during both recording and post production stages, has control over the creation of landscapes, arenas, stages and events from the foreground out to the acoustic horizon. Further, multiple frames can co-exist, within which multiple and different landscapes, stages and arenas can be created in cognitive dissonance or consonance.

3.1. Sound as Objects in Space

An increasingly broad array of sounds—abstract, generated, permuted or extracted from other sounds, often unidentifiable or other-worldly—have become considered common constituent elements of complex textures in music from art through to contemporary commercial music, soundtracks and installations. This is evident in works from composers such as Pratella, Russolo, Varese, Schaeffer, Boulez, Cage, Schafer, Xenakis, Stockhausen, Brecht, Horner, Zimmer, DJ Spooky, Cabaret Voltaire, Scanner, Newton Howard, and Elfman to name a few.

Similarly, advances in technology and parallel changes in aesthetic sensibilities have lead to an expanded use of ambience and aural spaces—abstract, generated, permuted or extracted, sometimes unidentifiable and other-worldly—with which these sounds are heard to exist. A large range of sounds now exist that do not have a pre-established connection to a particular aural space. Not only can we place sounds within as many different types of ‘space’ as imagination allows, but as Trevor Wishart points out “we can combine real (that is, recognisable as being possible in the real world) or imaginary, objects or spaces in any combination” [36:146].

This creates a (wonderful) paradox between sound and aural space highlighting one of the paradoxes of compositional process; does a composer think more about using sounds to build a composition or using composition as a means to organize sounds? In designing a sonic work, this chicken and egg question can equally be reframed as to what comes first, the ‘space’ or the ‘objects’ in it? There may be no definitive answer, however this paper of course prioritises discussion around Emmerson’s concept of framing sounds (events) within arenas and landscapes.

---

3 It is important to note that these approaches are not mutually exclusive.

4 A fifth of course can be seen in changes (delta) to each of the 4 key domains.
4. MAKING THE CONNECTIONS

Mixing music is a performative art form. The same real-time reactive and collaborative sensibilities that are central to performing with other musicians are present when mixing. As a performer I typically engage with one instrument at a time, however as a composer or mix engineer I engage with all instruments simultaneously. This demands a heightened sense of flow, speed and dexterity in decision-making, and an ability to creatively stand back from the ‘whole’ while at the same time delving deeply into the minutiae of all the sound elements. No different to any instrumental performance, a mixing performance requires technologies that enable interactivity and control-efficiency such as a performance controller (e.g. Monome) or mix control surface rather than the limited control of a computer mouse or track pad. As will be shown, surround functionality in Digital Audio Workstations (DAWs) is currently greatly limiting in terms of accessibility and functionality, technically and creatively.

Of particular interest herein therefore are two aspects of the surround panner: accessibility and functionality.

4.1.1. Accessibility

Surround panning functionality is typically only available to mono or stereo tracks. Multichannel tracks with more than two (e.g. stereo) channels presume hardcoded outputs and as a result do not present full featured surround pan controls. Figure 3 shows two tracks in Protools for example, the left track being a 7.1 track with no pan controls and the right being a stereo track with surround panning controls. In other DAWs (e.g. Logic) a single locator is enabled for all outputs. This technical limitation brings significant creative limitations, not the least of which is a lack of control over creating phantom images.

This handicap can however be circumvented to a degree. For example, one single 7.1 event can be broken down into 5 tracks comprising 3 stereo pairs (L-R, Ls-Rs, Lss-Rss) and 2 mono outputs (C and LFE). This however increases the performance challenges substantially. For example, a work may easily have at least 4 different blanket spaces. These may include early reflection ‘ambience’ to provide extreme depth control (commonly a convolution reverb), a short room-styled space to provide a bloom or halo to a sound, a longer reverb (typically with minimal early reflections), and a delay (to simulate early reflections without smearing or clouding the mix). Each track (5) of the single 7.1 instrument requires an aux send for each space (4) resulting in 20 aux send faders (each requiring control) for each single ‘instrument’. The pieces I typically work on have between 40 and 140 tracks. For this reason, the routing limitations inherent in surround-sound DAWs are understood to be significant.

4.1.2. Functionality

A surround panner is a device for controlling the location of a sound source using a positional reference on an x-y linear grid. Various points on the grid equate to discrete outputs, with points in between these using amplitude panning to determine levels to multiple outputs. While surround panners are typically used to locate sound within a surround speaker system, the method proposed herein utilises surround panner functionality to route signal to different auxiliary busses for the purpose of applying different digital signal processing (DSP) to discrete paths.

This method requires the creation of multichannel (e.g. 7.0) auxiliaries, which are then separated into stereo or mono busses in the DAW. At the time of writing this, the only DAWs that provide this capability are Avid’s Protools (using either Protools HD or Protools LE with surround tools) or MOTU’s Digital Performer. Avid’s Protools has been chosen for the purposes of this paper.

Figure 4 shows the bus setup for a single 7.0 surround auxiliary send (Morph-1). This is separated into 3 stereo sub-pairs (Morph 1a – 1c) and a mono send (Morph 1d). The matrix view to the right of image shows which panner outputs are used for each pair. With reference to figure 5 (superimposed numerals), Morph 1a is then using outputs 1 and 2, Morph 1b is using outputs 3 and 4, Morph 1c is using outputs 5 and 6, and Morph 1d is using output 7.

In this scenario then, a single 7.0 auxiliary is acting as the morphing controller for up to four discretely configurable spaces, therefore can be considered as a morphing-panner. The stereo returns (with associated DSPs) have their inputs assigned to Morph 1a, Morph 1b and so on. The outputs of the DSP channels are assigned a position in the surround field using an output surround panner.

5 Different DAWs have been tested for features, including Apple’s Logic Pro (v6-9), Avid’s Protools (v7-10), Steinberg’s Cubase (v5), Cockos’ Reaper and MOTU’s Digital Performer (v6-7).
The positional reference point on the morphing-panner therefore no longer correlates with the perceived location of the event, instead representing send levels to multiple DSP effects. Multiple outputs operating simultaneously from a single controller results in the ability to morph multi-various spaces using simple gestures.

Given that the panner also has 3 dimmable locations (Center[7] and Sides[3,4]), the complexity of results achievable from simple movements on a single controller (x-y positional reference) can be quite profound. Adding 3D divergence (front, rear and front-rear) provides further routing possibilities with the potential to send to all (four in this example) arenas from a single controller if desired. Figure 6 shows the signal path of a single stereo channel and two of its associated stereo returns (Morph 1a and Morph 2d).

In this example (figure 5), a mono event is being sent (ignoring pan law adjustments) at 100% level to outputs 4 and 6, and 50% level to outputs 3 and 5. If 3 and 4 were disengaged completely, the controller would effectively become quad (+ centre), meaning that the same controller placement results in 100% level to output 6, and 50% level to outputs 2 and 5.

Figure 6 shows a setup in Protools (4 mono audio tracks and 4 stereo pairwise spaces) with the auxiliary surround panners (morphing-panners) configured as morphing controllers. These are nested above the output surround panners that then determine the location of each event within the surround speaker setup.

Figure 7 depicts the signal flow for 2 of the 8 available (in this example) spaces when using two surround sends.6

5. CREATIVE APPLICATION

The following example7 describes the framing of an event where the only change is to the extent and nature of the landscape. The event is a 13 second extract from a whale song recorded using a hydrophone. I have subjected the sample to a forensic cleansing to remove extraneous noises so the result also contains no depth information. Subsequently it is not possible to gauge proximity with the result that the sound appears close. The sample is repeated 3 times over a 2-minute interval.

The event is routed via a single surround panner to 5 separate stereo arenas with a combination of stereo and mono inputs with all having stereo outputs)

- L routed to input Left distant halo (stereo input ambience)
- R routed to input Right distant halo (stereo input ambience)
- Ls routed to input bandpassed ping-pong delays 1 (mono-in, stereo-out in series with the stereo halo)
- Lss routed to input bandpassed ping-pong delays 2 (mono-in, stereo-out in series with the stereo halo)
- Rs routed to input long reverberant space (stereo input reverb)
- Rss routed to input long reverberant space (stereo input reverb)
- C routed to nowhere (allows option for source signal only)

Figure 8 depicts the (approximate) location of each space (positioned using output surround panners) during the time of the event. The whole song event remains in the same position for the 2-minute timespan, located between Left and Left Surround speakers. With the morphing-panner at 0º (position 7) the overall event is framed in a strong early-reflection ‘halo’ (aural space) that places the sound into the distance. As the event

---

6 The example used throughout this paper would result in 8 configurable spaces, however up to 14 would be available across two sends if all mono inputs were used (or 16 if LFE controls were included).

7 This example is taken from an original musical work (unpublished) titled Viscosity.
continues, the morphing panner is brought back and left toward the 270° position that keeps the event in the distance, but further introduces some bandpassed (distant) long delays (aural/musical space) to support the impression of spatial distance. The delays are located toward the rear left of a front-facing listener and are therefore distinctly locatable but also appear (in series) within the halo and therefore enhance the overall sense of size. From there the morphing panner is brought toward 90°, which introduces a long reverberant (musical) space that is located in an opposing location to both the delay and the original distance halo. The soundfield starts by seeming real, but morphs into the surreal via the introduction of surreal and imaginary arenas. As a call and answer, another short whale song sample is rendered through a middle distance room sound (110º), which provides a combination of halo (aural) as well as temporal (aural/musical) space. The overall effect is still relatively subtle, stopping well short of auditory sensory saturation [4] that could be achieved with more extravagant choices.

![Figure 8 – Example multi-space placement](Image)

In contrast to placing a mono sound (the event) in front-left, this event can now be framed on multiple stages in multiple locations. This creates an expanded landscape that occupies the complete surround field without ever shifting the event from front left. Importantly, this whole scene has been made to happen with a simple movement of a single controller.

6. CONCLUSION

In contrast to the rather obvious effect of moving sound events around the listener, moving sound in associated aural spaces through pairwise planes utilising surround panning functionality produces a sophisticated fluidity and seamlessness. Space can now take on a more significant role in composition.

Being able to route sounds into multiple spaces allows a greater sophistication in defining each sound’s (and subsequently the overall) timbre, depth, width, proximity and perceived energy. The physical constraints on doing this may typically make the approach seem unfeasible, however I have found that reducing the interactions to a single controller means that creative flow, speed and dexterity in decision-making have all been enhanced.

The equidistant speaker array used in 7.1, and the isotropic nature of it, greatly enhance (in my experience, not yet empirically validated) the ability to use multivariable musical and aural spaces in pairwise planes and from any direction. Certainly, I have found the ability to achieve smooth morphing of space to be ineffective in 5.1, but very effective in 7.1. Importantly, being able to create, manipulate and mediate sophisticated spaces and cause them to interact in their own right (e.g. series vs parallel) has proven to be a rich source of compositional inspiration.

7. REFERENCES


