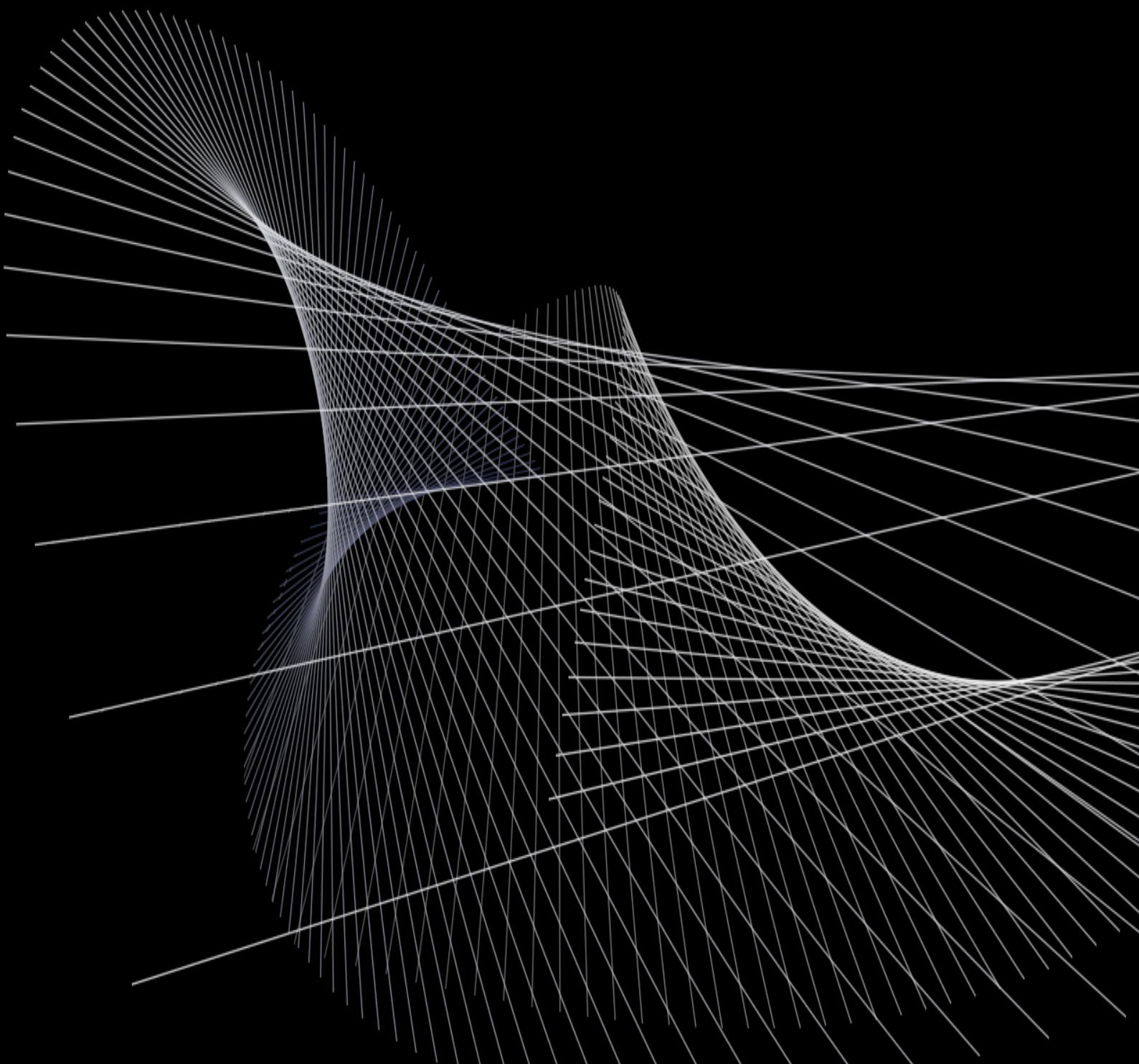


SOUND : SPACE



Conference Proceedings

Australasian Computer Music Conference

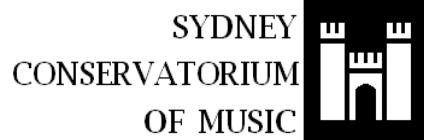
10 - 12 July 2008

SOUND : SPACE

Proceedings of the Australasian Computer Music Conference 2008

The Sydney Conservatorium of Music
The University of Sydney
Sydney, Australia

10th -12th July 2008



The University of Sydney

Editors

Sonia Wilkie

Anthony Hood

Published by the Australasian Computer Music Association

Organising Committee

Anthony Hood

Robert Sazdov

Ivan Zavada

Sonia Wilkie

P.O. Box 284
Fitzroy
VIC 3056
Australia

July 2008-06-19

ISSN 1448-7780

Keynote Speakers

Robert Normandieu

Roger Dean

All Copyright remains with the authors.
© Copyright 2008

Paper Jury

Lydia Ayers

Ros Bandt

Andrew Brown

Warren Burt

Amanda Cole

Jon Drummond

Christian Haines

Donna Hewitt

Gordon Monro

Tim Opie

Garth Paine

Rene Wooller

The paper refereeing process is conducted according to the specifications of the Australian Government for the collection of higher Education research data, and fully-refereed papers therefore meet Australian Government requirements for fully-refereed research papers.

* Denotes papers selected by peer review of full paper.

Special Thankyou to our Sponsors Liquid Architecture 9 and Sound Devices.

Thankyou to Peter Mcilwain and the ACMA Executive Committee

Thanks also to the many volunteers that have fulfilled tasks in preparation for the conference.

Keynote Speakers

Professor Roger Dean	1
Professor Robert Normandieu	2

Papers

* Lydia Ayers and Andrew Horner Synthesizing Some Australian Birds	3
* Warren Burt Do You Suppose He Didn't Know What He Was Doing? On "Not Knowing" and Computer Music	10
* Steven Campbell Flow, Riverscape, & Come Home: Three Installation Designs for the PLaY+SPaCE Ultrasonic Sensing Environment	13
* Toby Gifford and Andrew R. Brown Stochastic Onset Detection: An approach to detecting percussive attacks in complex audio	19
* Callum Hedemann, Andrew Sorensen, and Andrew R. Brown Metascore: User Interface Design for Generative Film Scoring	25
* Sarah Keith Universes of belief: a context-focused framework for performance design	31
* Julian Knowles Liminal Electronic Musics: Post-Punk Experimentation in Australia in the 1970s-1980s	37
* Tim Opie Rendering Soundscapes	46
* Garth Paine Interfacing for dynamic morphology in computer music performance	50
* Mark Pedersen User Configurable Intention Maps for Gestural Control and Live Spatialisation in Computer Music performance	55
* Jean Penny Amplified breath - (dis)embodied habitat: exploring sonic interfaces of performance, electronics, space and flautist identity	60
* Christian Ritz, Greg Schiemer, Ian Burnett, Eva Cheng, Damien Lock, Terumi Narushima, Stephen Ingham and Diana Wood Conroy An Anechoic Configurable Hemispheric Environment for Spatialised Sound	65
* Graeme Smith and Andrew Johnston Interactive Software for Guitar Learning	69
* Andrew Sorensen and Andrew R. Brown A Computational Model For The Generation Of Orchestral Music In The Germanic Symphonic Tradition: A progress report	78

* Wendy Suiter and Eva Cheng Listening Chairs: Personal Acoustic Space in Public Places	85
* Rene Wooller and Andrew R. Brown A framework for discussing tonality in electronic dance music	91

Artist Talks

Ros Bandt Finding the Isobue: From the Hydrophone to the Radio studio to the Gallery; spatial sonic designs of an endangered sound	96
Brigid Burke Interfacing for dynamic morphology in computer music performance	100
Chi-Hsia Lai Hands On Stage: An Interface for Sound and Image Improvisation	103
Stephen Stanfield Lost in Space: A Composer's Journey	107

Posters

Robert Bell PAD – the Percussive Audio Database: Initial Concept and Design	111
Ben Swift Thinking Through Musical Parameter Space	112

Workshops

Ajay Kapur 21st Century Raga: Digitizing North Indian Music	113
--	-----

Keynote Speaker

Professor Roger Dean

The University of Western Sydney
MARCS Auditory Laboratories
AustrALYSIS
Sydney, Australia
roger.dean@uws.edu.au
www.australysis.com

Biography

Roger Dean is an Australian composer/ improviser, sound and multi-media artist. He has performed in more than 30 countries, as bassist, keyboardist, and lap-top computer artist. His compositions include computer and chamber music, to commissions from the Australian Chamber Orchestra, Sydney Alpha Ensemble, Wallace Collection and Chaconne Brass (UK) and others. His works are published by Open University (UK/USA), Red House, La Trobe University, and Sounds Australian, on cd-rom with the International Computer Music Association, and on the web (multimedia commissions from the Australian Film Commission, How2, and others).

His work is also available on more than 30 commercial recordings, including LPs/CDs on Audio Research Editions, Discus, Mosaic, Soma, Future Music Records (FMR) (UK); Jade, Rufus and Tall Poppies (Australia); and Crayon, Cuneiform, and Frog Peak (USA). He is developing computer-interactive networked improvisation, sound and intermedia work. He has published five books and many articles on improvisation, particularly in music, and on music cognition. With Hazel Smith he wrote *Improvisation, Hypermedia and the Arts Since 1945* (Harwood, 1997). A subsequent book was on computer-interactive sound improvisation: *Hyperimprovisation*, published by A-R Editions (USA; with cd-rom, 2003). In 2005 he published *Sounds from the Corner: Australian Jazz on CD* (Australian Music Centre). His current research focuses on music cognition and computation.

He is the founder and director of austrALYSIS, the international sound and intermedia arts group (commenced as LYSIS, UK in 1970); and also founded the Sonic Communications Research Group at the University of Canberra. Roger has the unusual distinction of being a subject in both the new Grove Dictionary of Music and that of Jazz. He appears in an international group led by British jazz composer Graham Collier, and first recorded with him in 1975; a 2007 release is on Rune (USA). austrALYSIS regularly presents new work in Sydney and elsewhere, and its most recent CD is *Sonic Stones* (Tall Poppies, 2006). Dean was until early 2002 the founda-

tion director of the Heart Research Institute, Sydney, and has more than 300 substantive scientific publications. From 2002-2007 he was the Vice-Chancellor and President of the University of Canberra. He is now a full time research professor in Sonic Communication at the MARCS Auditory Laboratories, University of Western Sydney. See www.australysis.com for information about his artistic work.

Keynote Speaker

Professor Robert Normandeau

Université de Montréal

Montréal, Canada

robert.normandeau@umontreal.ca

Biography

After a BMus in Composition (Electroacoustics) from the Université Laval (Québec City, 1984) he moved to Montréal and completed an MMus in Composition (1988) and the first Ph.D.Mus in Electroacoustic Composition (1992), under Marcelle Deschênes and Francis Dhomont. He is a founding member of the Canadian Electroacoustic Community (CEC, 1987). From 1986 to 1993, he was an active member of the Association pour la création et la recherche électroacoustiques du Québec (ACREQ), where he produced the Clair de terre concert series at the Montréal Planetarium. In 1991, he co-founded Réseaux, an organization for the production of media arts events, notably the acousmatic concert series Rien à voir.

After a certain interest in instrumental and mixed works, his current endeavours are focused on acousmatic music. More specifically, his compositions employ esthetical criteria whereby he creates a 'cinema for the ear' in which 'meaning' as well as 'sound' become the elements that elaborate his works. Along with concert music he now writes incidental music, especially for the theatre.

Robert Normandeau is an award winner of numerous international competitions, including Ars Electronica, Linz (Austria, 1993, Golden Nica in 1996), Bourges (France, 1986, 1988, 1993), Fribourg (Switzerland, 2002), Luigi-Russolo, Varese (Italy, 1989, 1990), Métamorphoses, Bruxelles (Belgium, 2002, 2004), Musica Nova, Prague (Czech Republic, 1994, 1995, 1998), Noroit-Léonce Petitot, Arras (France, 1991, 1994), Phonurgia-Nova, Arles (France, 1988, 1987), and Stockholm (Sweden, 1992).

He is Professor in electroacoustic music composition at Université de Montréal since 1999. He received two Opus Awards from the Conseil québécois de la musique in 1999: "Composer of the Year" and "Record of the Year-Contemporary Music" for "Figures". The Académie québécoise du théâtre (AQT) has given him a Masque Award: "Best Music for Theatre" for "Malina" in 2002.

He received commissions from The Banff Centre for the Arts, CKUT-FM, Codes d'Accès/Musiques et Recherches, Groupe de recherches musicales de Paris, Groupe de musique expérimentale de Marseille, Jacques Drouin, événements du neuf, Claire Marchand, Arturo Parra, Musée d'art contemporain de Montréal,

Open Space Gallery, Radio-Canada, Réseaux, Sonorities Festival, Vancouver New Music and ZKM (Zentrum für Kunst und Medientechnologie). He was composer in residence at the studios of Banff (Canada, 1989, 1992, 1993), Belfast (Northern Ireland, 1997), Bourges (France, 1988, 1999, 2005), Mons (Belgium, 1996), the GRM in Paris (France, 1990, 1994), Ohain (Belgium, 1987) and ZKM (Karlsruhe, Germany, 2004)..

Lydia Ayers

Andrew Horner

The Hong Kong University of Science
and Technology

Clear Water Bay, Kowloon

Hong Kong

layers@cs.ust.hk, horner@cs.ust.hk

Abstract

This paper describes an additive synthesis design for synthesizing bird calls which models their dynamic spectra and many of their expressive characteristics. The Csound design produces tones that are very similar to several Australian bird calls. The design also works for other animal sounds, such as crickets. Finally, excerpts from a musical composition illustrate the musical uses of the design and the flexibility of varying the calls.

Introduction

When hiking around Australia in 2007, we saw some beautiful birds (see Figure 1) and listened to some attractive bird calls. Although it is a mystery whether the silent birds we saw made any of the calls we actually heard when we weren't looking at them, we've synthesized some of the calls using additive synthesis in Csound (Vercoe 1992).



Figure 1. Birds at Uluru

This paper first gives some background on sound synthesis research and then describes our source material. As the additive synthesis design is very similar to a previous design, we focus mainly on the spectral data used for synthesizing

Synthesizing Some Australian Birds

bird calls. We then discuss several examples of the bird call transcriptions. Finally, we conclude with excerpts from a composition illustrating the musical uses of the design with the flexibility of varying the calls, and an extension to synthesize a cricket.

Background

Ayers wanted to use some sampled bird recordings in a composition, but they contained a lot of background noise which filtering did not sufficiently remove. Could our previous synthesis design for musical instrument tones model the bird calls? She tried it without initially considering the research implications of the project. The goal was simply to replace the recordings, and the method worked. On further reflection, when we realized that most of the bird calls used in computer music compositions that we've heard have been sampled, and that few have been synthesized, we decided to share our results with the community.

We previously designed group additive synthesis models of many wind instrument tones (Horner, Ayers and Law 1999; Horner and Ayers 1998a, 2002) and synthesized timbre tremolos and flutter tonguing on wind instruments (Ayers 2003, 2004; Ayers and Horner 2004a and b). We have synthesized the Chinese flutes and the *suona* using additive synthesis (Ayers 2005; Ayers and Horner 2006), and we have also synthesized the *didgeridoo* using timbre morphing (Ayers and Horner 2007).

Other researchers have focused their attention on the acoustics (Greenewalt 1968; Fletcher and Tarnopolsky 1999) or physical modelling of bird calls (Kahrs and Avanzini 2001). Some of these models are quite complex and their goal seems to be pure research rather than creating a bird instrument to use in compositions. Our research is basically working in the opposite direction: while most of the other researchers are trying to synthetically reconstruct the bird and consider how its physical characteristics affect the sound, we're trying to reconstruct the sound without considering the shape of the bird that made it. We found one additive synthesis application in Matlab, "The Additive Synthesis Forest," which doesn't seem intended as a composition tool (Sturm 2003)..

The Source Material

We recorded some attractive bird calls when hiking around Watarrka, Uluru, Kata Tjuta and Alice

Springs in central Australia, Namadji National Park near Canberra, and around Cairns in 2007. We recorded the sounds using the only recording device we had brought with us, our digital video camera. With the environmental background noise and camcorder quality, it was less than optimal, but we got very interesting synthesis results in spite of the limitations. Our motivation for synthesizing the sounds was to clean up the background noise which made them difficult to use in a sampled instrument design.

In order to make the recording usable, we cut them up into the cleanest segments, and then filtered them to minimize the background noises, such as wind, camera clicks, other birds, etc. Although such a process is necessarily “lossy,” the soundfiles we used produced a pretty fair representation of the bird sounds. Figure 2 shows a spectrogram (Horne 2001) analysis of a short segment used for the synthesis, with four calls from one bird, and two calls from another bird in the background. The pitch contour is fairly flat for the fundamental of the bird call we want, at around 2000 Hertz, and the second bird is lower, around 750 Hertz.

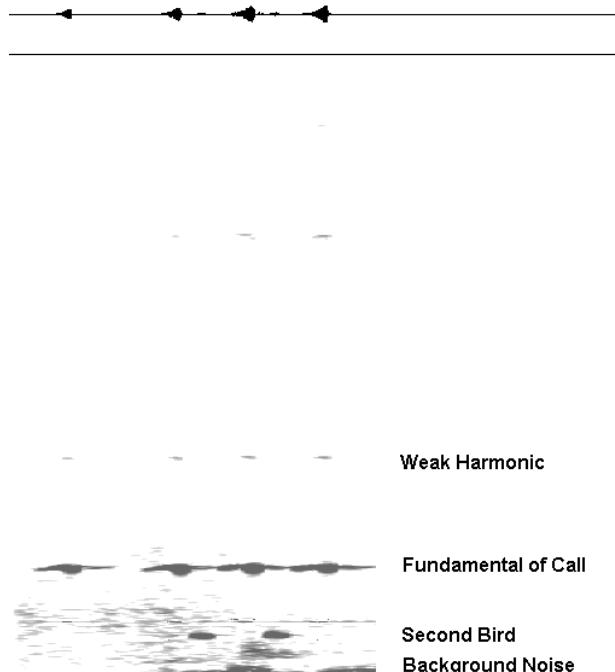


Figure 2. Spectrogram and Waveform of Original Call for Bird 1

In order to remove the hissy background and the two calls from the second bird, we filtered the bird calls. This produced the cleaner sound shown in Figure 3, with a further decrease in amplitude of the already weak higher harmonics.

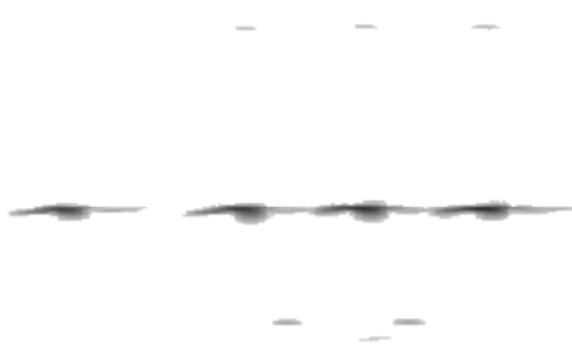


Figure 3. Spectrogram of a Filtered Call for Bird 1

The Additive Synthesis Design

The bird calls modelled in this paper are clearly pitched, but they don't divide up into 12-tone equal temperament scales. After examining the recordings, we found that we could get the musical results we wanted from synthesizing the average spectra of just three bird calls. We grouped two calls together for Bird 1, and kept one call separate for Bird 2.

Our previous additive synthesis designs use up to 63 sine waves to model the rich spectra of various timbres such as the *suona* (Ayers and Horner 2006). Most bird calls are higher than *suona* tones, their spectra aren't as rich in harmonics, and 10-20 harmonic sine waves were enough to model their timbres. Since the bird spectra worked well in the additive synthesis design, we won't review its code here. Instead, we will focus on how we customized the data to produce bird calls. An important feature of our design is that repeated performances of the bird calls will have varied spectra in order to make the performance sound more natural, which is useful for producing a more expressive composition.

Special Amplitude Envelope for “Bird 1”

As shown in Figure 2, the higher call for Bird 1 has an unusual amplitude envelope that rises for most of the sound and then decays very quickly. In order to handle this special case, we added a function table that produces the envelope with a slight randomized variation (see Figure 4). The minimum amplitude is .1% of the total amplitude and the maximum amplitude is 100% at about 85% of the duration of the tone. If this special envelope is not needed, the design uses a simple rise, sustain and decay.

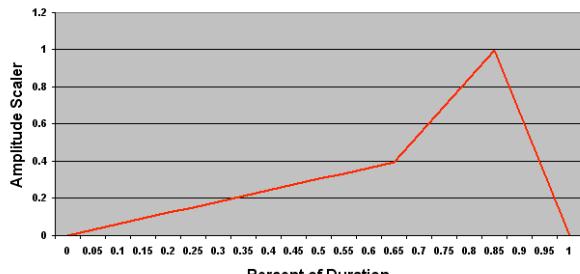


Figure 4. Amplitude Function for Bird 1

The Bird Spectra

This section describes the spectra of two groups of calls with similar spectra which we combined as Bird 1 and the separate call for Bird 2.

The Spectral Components for “Bird 1”

Figure 5 shows an 11-harmonic time-varying phase vocoder analysis (Dolson 1986; Wun 2000) of the group of four calls from one bird. The pitch contour is fairly flat, at around 2000 Hertz, and the segment contains mainly one loud low harmonic and much instability among the higher harmonics.

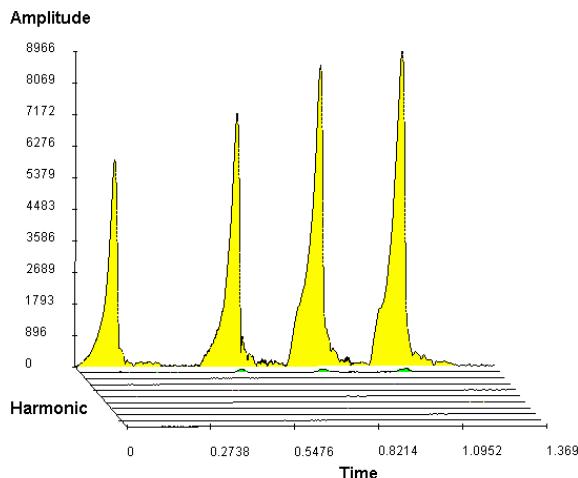


Figure 5. Time-Varying Phase Vocoder Analysis of Filtered 2000 Hertz Call for Bird 1.

We use the average harmonic amplitudes to model this bird call based on our previous additive synthesis wind instrument designs (see Figure 6). The first harmonic is quite strong and the others very weak. Based on this spectrum alone, we might consider just using a sine wave to synthesize the calls, as others have done (Assman). However, we found that the low amplitude harmonics above the fundamental do contribute to the bird call’s timbre, particularly in the lower frequency calls described in the following discussion.

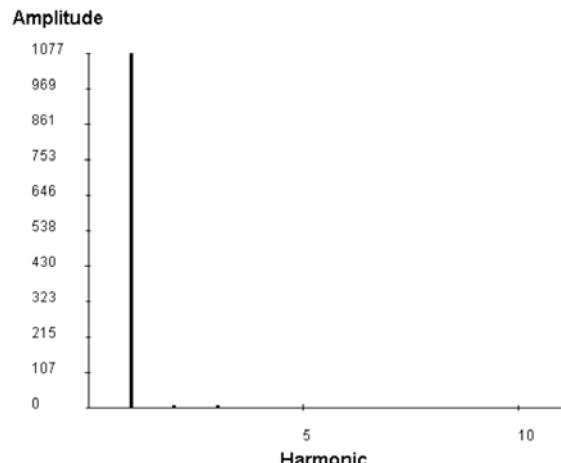


Figure 6. Average Harmonic Amplitudes of Filtered 2000 Hertz Call for Bird 1.

Figure 7 shows a 20-harmonic time-varying phase vocoder analysis of a lower bird call at 1100 Hertz. As with the higher call, the lower call contains mainly one loud low harmonic and much instability among the higher harmonics.

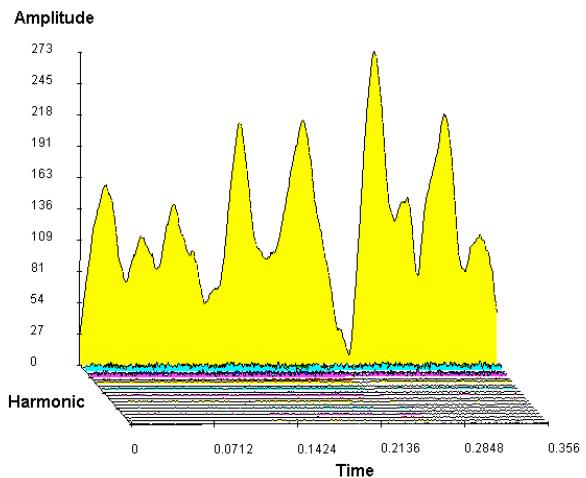


Figure 7. Time-Varying Phase Vocoder Analysis of Filtered 1100 Hertz Call for Bird 1.

Figure 8 shows the average harmonic amplitudes used to model this bird call. The first harmonic is still stronger than the others, but proportionally the amplitude of the other harmonics is stronger, with more noise in the higher harmonics. As the amount of variation in the higher harmonics is within the range of random variation incorporated in the design, we did not include additional function tables to cover bird calls between these two frequencies.

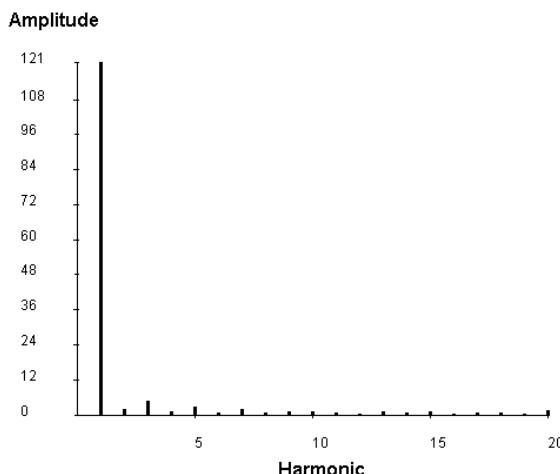


Figure 8. Average Harmonic Amplitudes of Filtered 1100 Hertz Call for Bird 1.

We stored the amplitude data from the graphs in two Csound score function tables (see Figure 9).

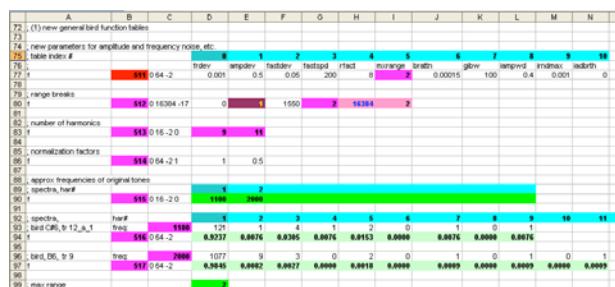


Figure 9. Csound Function Tables for Bird 1

The Spectral Components for “Bird 2”

Although most of the calls have spectra similar to the average amplitudes used for Bird 1, a few calls have a much stronger second harmonic and a very weak fundamental. Did we make an error in finding the fundamental frequency of these calls? As the synthesis for these calls really does work better with the weak fundamental, strong second harmonic spectrum than it works with lowering the fundamental by an octave and using either of the other spectra, we separated it from the other tones as “Bird 2.”

Figure 10 shows a spectrogram analysis of the call used for Bird 2, with six similar calls, each preceded with a portamento grace note. The fundamental frequency of this call is around A6 at 1820 Hertz, with the grace note beginning at around 1650 Hertz (about G#6).



Figure 10. Spectrogram of Filtered Call for Bird 2

Figure 11 shows a 12-harmonic time-varying phase vocoder analysis of the segment. The segment contains mainly one loud second harmonic and much instability among the higher harmonics.

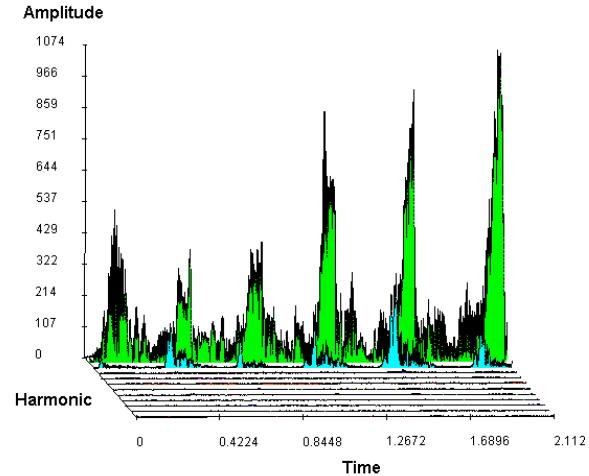


Figure 11. Time-Varying Phase Vocoder Analysis of Six Filtered 1820 Hertz Bird Calls.

Figure 12 shows the average harmonic amplitudes used to model this group of calls. The second harmonic is stronger than the others, with harmonics 1 and 3 also having significant amplitudes and little amplitude in the higher harmonics.

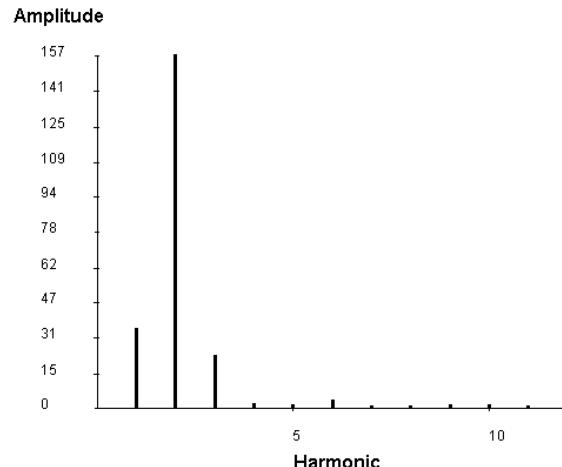


Figure 12. Average Harmonic Amplitudes of Filtered 1820 Hertz Call for Bird 2

After using the phase vocoder analysis to find the three average spectra and converting them to the Csound function tables, we concentrated on transcribing the wide variety of the other bird calls.

Transcribing the Bird Calls

This section describes transcribing the bird calls. As already shown, the spectra of these birds are fairly close to sine waves, so the most distinctive feature of each call is its expressive pitch contour. As shown in Figure 3, the four-note call used for Bird 1 has a flat frequency contour. The actual frequency, around 2000 Hertz, is quite close to B6, and it has a clear rhythm (see Figure 13). The four

sounds in this call use the amplitude envelope with a continuous rise and almost no decay shown in Figure 4. Although this motive seems quite simple, this little bird turned out to be very useful as a rhythmic marker in the composition described below.

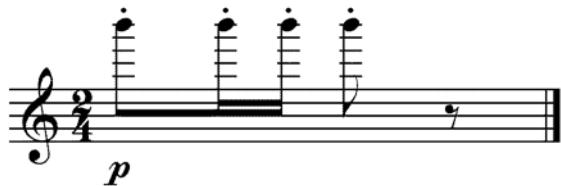


Figure 13. Transcription of Bird Call 1, Example 1.

We added a lower frequency for Bird Call 1, quite close to C#6 (around 1100 Hertz), using the first tone of a different bird's call (see Figure 14). That complete call has a rising motif, followed by an answering melody (from the same bird). The bird's rhythm is slightly different from Ayers' transcription, which adapts it to her musical purposes. Some of the bird tones have an amplitude modulation which she modelled using the previous flutter tongue design. (Ayers 2004)



Figure 14. Transcription of Bird Call 1, Example 2.

The six-call motif used for Bird 2 (see Figure 15) has an overall flat contour, with preceding portamento grace notes providing expressive interest. The fundamental frequency of this call is around A6 at 1820 Hertz, with the grace note beginning at around 1650 Hertz (about G#6). Changing the grace notes to G natural made them fit better with the pentatonic E minor scale on the native American flute.

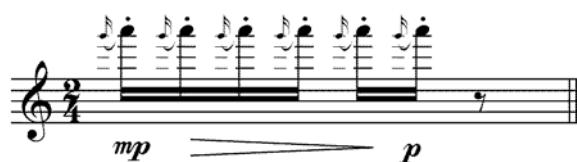


Figure 15. Transcription of Bird Call 2, Example 1.

A number of other interesting calls proved useful for the piece; Figure 16 illustrates a particularly striking pair of calls. Many times, a bird repeated its call with noticeable microtonal variations on the pitches, and it was quite fun to exploit this characteristic in the composition.



Figure 16. Bird Call Example 3.

Musical Use of the Design

From the "Dreamways" of the Iroquois to the "Dreamtime" or "Dreaming" of the indigenous Australians, ancient peoples have left their stories carved or painted on stone. In *Rock Art in the Dream World*, Ayers has created a musical story about these people, who travelled through their natural world and their dreams to create rock art. The live native American flute in E and *didgeridoo* in E represent two human characters. Accompanied by their synthesized dream selves, they stroll through a natural environment including a native American medicine rattle and recorded and synthesized Australian birds, crickets, kangaroos, rock taps and tap sticks.

The composition is divided into the following sections:

- Journey to the Art Place**
- Creating the Art**
- Reflecting on the Art**
- Finishing the Art**
- Into the Dream World**
- Return**

The composition uses a number of other bird call transcriptions using the three spectra previously described to model all the birds. For example, both of the bird calls illustrated in Bird Call 1, Example 2 (see Figure 14) are transformed in both the real and synthesized dream flute parts. In the slower section of "Into the Dream World," some of the motives, including this one, are stretched into "slow-motion" time (see Figure 17).



Figure 17. "Slow Motion" of Bird Call 1, Example 2.

The Flute Part

The native American flute part grew out of the motivic material transcribed from the bird calls. However, each flute typically has a range of slightly larger than an octave, with a basic pentatonic minor scale. Using two flutes an octave apart expanded the range, but the birds remain out of range of both flutes. Figure 18 shows parallel use of motive material from bird calls in the flute part. At times, the flute and birds play together, at other times they chase each other, with the birds keeping out of reach. The dream flute shadows ahead, behind or in harmony with the live performer, too.



Figure 18. Parallel Motives in Flute and Bird Calls

The Didgeridoo Part

The live *didgeridoo* performer is free to improvise his part as he wishes. A demo version gives a sense of how his live performance may interact with his dream self. The synthesized *didgeridoo* used for the dream self and the demo live part is based on the design described in Ayers and Horner 2007.

Cricket Chirps

We were also able to synthesize a cricket chirp using the same design. The cricket makes longer single chirps and groups of shorter chirps. Each chirp is made up of a high frequency sound with amplitude modulation (see Figure 19).



Figure 19. A Single Cricket Chirp.

The spectrum of the cricket chirp is somewhat similar to the spectrum of Bird Call 2, though for the cricket the second harmonic is the only one with significant amplitude (see Figure 20).

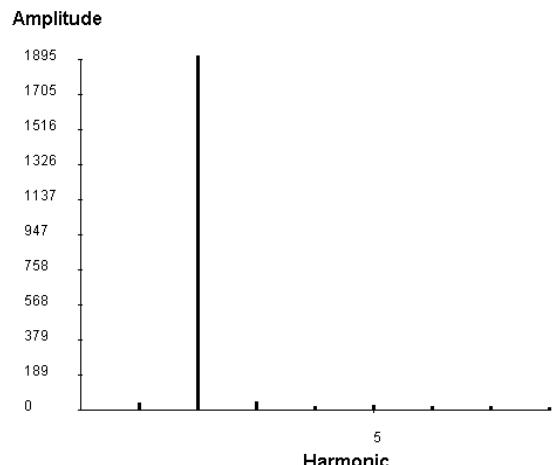


Figure 20. Peak Amplitude of Cricket Chirp.

Figure 21 shows several cricket motives, sounding an octave higher, that function as rhythmic markers in the composition.



Figure 21. Cricket Chirps

Conclusion

We used additive synthesis to model the dynamic spectra and musical expression of some Australian birds. The Csound design captures the subtle characteristics of the bird calls, and also works for cricket chirps. In addition, Ayers transcribed the calls and used the design in the soundscape of a composition, *Rock Art in the Dream World*, for flute and *didgeridoo*, which demonstrates the expressiveness of the design.

Acknowledgements

This work was supported by the Hong Kong Research Grants Council Competitive Earmarked Research Grant 613806.

References

- Aßmann, P. www.utdallas.edu/~assmann/trackdraw/bird.html
- Ayers, L. 2003. "Synthesizing Trills for the Chinese Dizi," *Proceedings of the International Computer Music Conference*, Singapore.
- Ayers, L. 2004. "Synthesizing Timbre Tremolos and Flutter Tonguing on Wind Instruments." *Proceedings of the International Computer Music Conference*, Miami, Florida, USA.
- Ayers, L. 2005. "Synthesizing Chinese Flutes Using Csound," *Organized Sound*. 10:1.

- Ayers, L. and Horner, A. 2004(a). "Synthesis of Chinese Dizi Ornaments," *Ghost in the Machine: Performance Practice in Electronic Music*, Australasian Computer Music Conference. Wellington, New Zealand.
- Ayers, L. and Horner, A. 2004(b). "Expressive modeling of Chinese folk wind instruments," *Journal of the Acoustical Society of America*, San Diego, California, USA, 116:4, part 2.
- Ayers, L. and Horner, A. 2006. "Synthesizing the Chinese Suona," *Medi(t)ations, Australasian Computer Music Conference 2006*, Australasian Computer Music Conference. Adelaide, Australia, pp 15-22.
- Ayers, L. and Horner, A. 2007. "Didgeridoo Synthesis Using Timbre Morphing," *Trans: Boundaries / Permeability / Reification, Australasian Computer Music Conference 2007*. Canberra, Australia, pp 18-25.
- Dolson, M. 1986. "The Phase Vocoder: A Tutorial," *Computer Music Journal*. 10:4.
- Fletcher, N. and Tarnopolsky, A. 1999. "Acoustics of the avian vocal tract," *Journal of the Acoustical Society of America*, 105(1), pp. 35-49.
- Greenewalt, C. 1968. *Bird Song: Acoustics and Physiology*. Smithsonian Institution Press, Washington, D.C.
- Hibbert, B. 2003. *Bell Waveform Analysis Program*. Great Brookham, Surrey, UK. www.hibberts.co.uk.
- Horne, R. 2001. *Spectrogram*. Visualization Software. www.visualizationsoftware.com/gram.html
- Horner, A., Ayers, L. and Law, D. 1999. "Synthesis Modeling of the Chinese Dizi, Bawu, and Sheng," *Journal of the Audio Engineering Society*. 47:12.
- Horner, A. and Ayers, L. 1998(a). "Modeling Acoustic Wind Instruments with Contiguous Group Synthesis," *Journal of the Audio Engineering Society*. 46:10.
- Horner, A. and Ayers, L. 1998(b). "Modeling Chinese Musical Instruments," *Proceedings of the 135th Meeting of the Acoustical Society of America*. Seattle, WA, 4.
- Horner, A. and Ayers, L. 2002. *Cooking with CSound, Part 1: Woodwind and Brass Recipes*. AR Editions.
- Kahrs, M., and Avanzini, F. 2001. "Computer Synthesis of Bird Songs and Calls," *Proceedings of the COST G-6 Conference on Digital Audio Effects (DAFX-01)*, Limerick, Ireland, December 6-8, 2001.
- Sturm, B. 2003. "The Additive Synthesis Forest." www.mat.ucsb.edu/~b.sturm/SSUM/docs/birds.htm.
- Vercoe, B. 1992. *Csound: A Manual for the Audio Processing System and Supporting Programs with Tutorials*. MIT Media Lab, Cambridge, MA.
- Wun, S. 2000. *PVan*. www.cs.ust.hk/~simon/pvan

Warren Burt

Sonic Arts Research Network, University of Wollongong; and Illawarra Institute of TAFE
PO Box U-27
Wollongong, NSW
Australia
waburt@melbourne.dialix.com.au

Do You Suppose He Didn't Know What He Was Doing? On “Not Knowing” and Computer Music**Abstract**

Is there a place in computer music for not knowing? Is there a place in computer music for suspension, or transcendence of the ego? Is there a place in computer music for ecstatic expression? Is there a place in computer music for non-mediated creation? Is there a way in which creating computer music can be a spiritual practice? This short essay asks these questions in a non-linear manner, not so much as a means of proposing answers, but as a means of suggesting problems to be dealt with.

First Sound Excerpt: duration 1:30

I recall, a number of years ago, a composer colleague saying, "I can account for every note in this piece!" His pride was that not only was his piece constructed totally rationally, but that each decision along the way had a justification as well. Contrast this with the familiar story of Morton Feldman bringing his early student work to John Cage – after looking at the piece, Cage asked Feldman how he had made it. Feldman replied weakly, "I don't know how I made it." Cage exploded with enthusiasm: "Isn't that marvelous! It's so beautiful and he doesn't know how he made it!"

The quote which opened this paper, and from which it takes its title, is from a series of conversations between Morton Feldman and John Cage, recorded by WBAI, New York in 1966 and 1967. These are now available from <http://www.archive.org/details/CageFeldman5>. The quote comes from the very beginning of the fifth dialog. At the end of the fourth dialogue, they had been talking about, among other things, music education, and Edgar Varese. Then the tape ran out. The 5th dialog simply opens in mid sentence. Whatever transpired while the tapes were being changed has been lost. Here is the complete moment:

Cage: Do you suppose he didn't know what he was doing....or, knew what he was doing and didn't want anyone to know.

(long pause)

Feldman: I think that he knew what he was doing, but he didn't want to know what he was doing.

(long pause)

Cage: Well, in a very real sense, that's what we're all doing, because even though we might think

we knew, the thing will only come to life for someone else when he knows something that we don't know. (Cage / Feldman 1966-67)

This quest for not knowing, for the inspired unconscious act, for the making of music in other than a state of determined rationality, has been around for much of the 20th and 21st centuries. It can be traced back even farther, of course, back to the 14th century Christian mystical text, *The Cloud of Unknowing*, the writings of Islamic mystics such as Rumi and Hafiz, and innumerable Buddhist and Zen texts. And it turns up in other places. Arnold Schoenberg, mis-represented by many as the height of compositional rationality, in a 1914 letter to Wassily Kandinsky states that since 1909 he has been searching for 'complete freedom from all forms', and later says 'We search on and on (as you yourself say) with our feelings. Let us endeavour *never* to lose these feelings to a theory.' (Craft 2006:29-30.)

Given that in computer music, so much of what we do is done in programming languages and environments, which often involve a lot of rational, step-by-step analytical thinking, it's important to ask about the irrational, the unknown, the ecstatic state (three very different things) as a basis for musical creativity. This essay is posed as a series of questions. I ask these questions as a secular composer in a technological medium, not pushing or endorsing any kind of religious or spiritual belief. However, since our art is concerned with the study of attention (how do we listen?), and since the meditative and spiritual disciplines have been concerned with this for centuries, we should feel free to listen to them.

Second Sound Excerpt: duration 1:34**Encounters with Remarkable People**

In 2003, through a mutual friend, I was invited to meet Sheikh Abdul Aziz, the British-born leader of the Mevlana Sufis of Melbourne. I found him to be an absolutely inspiring teacher, one who used irrationality and paradox in his talks in a most delightful way. I was also impressed with how the Melbourne Sufis made their spiritual practice the centre of their lives, organizing work, play, and trips away around their Sufi practice. While observing this, the thought occurred to me: "This is all very beatiful, but I already have a spiritual practice – it's called music

composition!" I realized that what I was doing was a secular equivalent to a spiritual discipline – a path of self-alteration and self-growth involving the exploration of sound and the processes behind it. My readings in neo-Pythagoreanism, especially Iamblichus, Porphyry and Plotinus, had shown me that mathematical contemplation was indeed a meditative, if not overtly spiritual, practice and that my explorations into various mathematical-musical structures and tuning systems were a part of that meditative, spiritual path.

Thoughts such as this, and Cage's early 1950s ideas of using chance as a means of avoiding what I would now call "first level" (note-to-note) ego-driven choices in composition, lead to considering the ego. One of the best writings about the ego I've read recently is a 2006 interview with Robert Frager, head of the Institute of Transpersonal Psychology, who is also Sheikh Rangip, leader of the Redwood City, California Helveti-Jerrahi Sufis. In this interview, he talks at length about the ego and its seemingly contradictory place in modern psychology and traditional spiritual practice. I will now quote from it at length, interspersing quotes from it with questions of my own as to how the issues he raises might relate to computer music.

One way of putting the problem is that in using the term "psychology" in an academic setting, in an institution that offers a Ph.D. Degree, we're taking on the whole Western academic tradition with its emphasis on the head alone – certainly not heart, much less soul. If you break apart the very term "psychology," "psyche" means spirit or soul in Greek; and therefore, psychology or psychoanalysis is literally the scientific analysis, the logical cutting up, or parsing, of the soul, which in itself is pretty crazy. How in hell do you parse the soul? How can you be analytic when it comes to the soul?

I don't think it's too much of a stretch to relate this to our field, computer music, which has mostly lived in the realm of the academy, and on the edges of the sciences. And yet, our field is music, which totally lacks that ability to be proven which is the essence of scientific method. I mean, how do you prove music? There's nothing to prove. Sound simply is. (And we here note that economic success or failure proves nothing. Or at least nothing that is relevant to the present arguments.)

When you even use the term "psychology," you're buying into something that says logic will do it. But logic is a very limited tool. Certainly, logic has caused me to make a lot of wrong decisions in my life. And in

Sufism, as soon as you get to the higher stages, forget logic. It doesn't figure anymore because you have a paradox; what is that soul in you that's transcendent? What is before the before? And after the after? These are not questions logic is ever going to handle.

How does paradox enter into our composing computer music? What is the place for the transcendent in computer music? Can we use the logical construct of the "patch" as a substrate for ecstatically based artistic activity?

I think certain practices frankly don't have any power unless you've been given them by a teacher. They won't work. So I think this business about being your own teacher ignores the importance of transmission, of lineage, of initiation. The spiritual path is not merely logical or mechanical. It's not psychological or spiritual bodybuilding. It's something much more subtle. I think there's an energetic connection with the teacher. We talk in Sufism about the *rabita al kalb*, the connection of the heart.

We might ask – In Australian computer music, do we have this idea of transmission, lineage, initiation? Should we? If we don't have it, why not? If we need it, is there some way to establish it?

One very important component of the struggle to develop oneself spiritually is service – service to humanity but also service to the world, to all of creation. One of the great tools to do that is the personality structure, including the ego, the sense of self. Now even as you're working to divest yourself of that separate sense of self, which is the last stage, in order to get there, paradoxically, you need to use that self well. It is the beast on which the Buddha rides.

Is the ego - the desires - the beast on which our ecstatic sounding can ride? Is technology the beast upon which our Buddha can ride? Is programming logic the substrate on which we can build our sonic and spiritual explorations? In Sufism, the aspects of the personality which hold us back are called the *nafs*. Is there a way to make computer music free of the *nafs*?

I had a wonderful teacher, Moshe Feldenkrais, who is an incredible teacher of movement and bodily functioning. He could work directly with anyone – from those with the most severe physical handicap limitations from accidents and birth defects, all the way up to great athletes

and musicians – to improve their functioning. And he said, “When I’m working with someone, I don’t even think in sentences. Because the structure of grammar would get between my nervous system experiencing the nervous system of the person I’m touching.” (Frager/Hamilton 2006)

Is there a way to incorporate that kind of non-grammatical immediate nervous system to nervous system communication in computer music? Is there some way that a structure built on grammar (a program) can be used for that kind of non-verbal expression? Even farther, and perhaps most mystical of all: Is there a way our energy and the energy of our machines – nascent intelligences they very well may be – can cooperate in some kind of higher spiritual or energetic song? Are we already transformed by our machines, even as we are transforming them?

Third Sound Excerpt: duration 1:13

Questions and Music

Kenneth Gaburo’s “scatter” method of composition, that he developed in the 1980s, was put forward as one way of transcending logic, rationality, “the lick.” It used the physicality of his body, often after extended sensory deprivation experiences, to make an output – a diagram, a tracing of some kind. These tracings, the product of unmediated physical energy, were then analyzed and transcribed into various other media. In dealing with machines, Gaburo often worked with them directly after a sensory deprivation experience. In one case, “*Rerun*” he worked with the loudspeakers off, using his sense of physicality to direct his working with the machine. These processes are described in his essays “ISIT” and “LA.” (Gaburo, 1986, 1987)

So can I not know what I’m doing? How can I work with deterministic technology in not just a non-deterministic way, but also a way where I literally don’t know what I’m doing, or what my results will be? And how can this be reconciled with the desire for both very carefully controlled sonic results, and for exact attention to the details of unexpected sonic results. In other words, how much “not knowing,” and in which contexts, are we talking about? How can I get to such a state of familiarity with my equipment that I can compose with it ecstatically, intuitively, spontaneously? Can I use my medium in an unmediated manner?

The three musical excerpts heard here were attempts at that not knowing, and of becoming, as is said by Tibetan Buddhist writer Lobsang Phuntsok Lhalungpa, “listening itself.” In the first, the quote from John Cage is treated with the Composers’ Desktop Project “Distort-Cycles-Repeat” and “Extend-Scramble” functions to produce an extended and fragmented result. These functions are such that predicting their

results is extremely difficult. John Dunn’s ArtWonk was used to generate lists of random parameters for the CDP program to produce time-varying results. The second excerpt takes 10 quotes from Chris Mann’s interactive website “The Use” <http://theuse.info/> (last accessed April 6, 2008) and convolves each of them against all the others. No sounds were listened to during the gathering of the samples, or the editing, or the convolution process. All work was done visually, and kinesthetically. Only once the samples were placed into the VSampler sampler, and a random selection process in ArtWonk had been devised, were the results listened to, and the ArtWonk program modified slightly. The final excerpt is made with some electro-acoustic percussion boards I built. I improvised on the instruments, recorded the result, then used the CDP function “Edit-Random Chunks” to cut each of four improvisations into 61 random fragments each. This produced fragments I would not have otherwise chosen. These fragments were assigned to individual keys in the Wusikstation sampler, and this assignment was then further unpredictably altered by placing the Wusik into a 64 note microtonal tuning. This was then controlled by another random selection program made with ArtWonk.

Fourth Sound Excerpt: duration 0:03

References

- Cage, J. & Feldman, M. 1966-67. “In Conversation: Radio Happening V of V” recorded at WBAI, New York, <http://www.archive.org/details/CageFeldman5> (all websites last accessed April 6, 2008).
- Craft, R. 2006. “Down a Path of Wonder”, Naxos Books, Norfolk, UK.
- Frager, R. & Hamilton, C. “The Man With Two Heads” *What is Enlightenment?* Issue 34, September-December 2006. Article available online at <http://www.wie.org/j17/frager.asp?page=1> (last accessed April 6, 2008).
- Gaburo, K. 1986 “ISIT” *IS Journal* No. 2, pp. 39-51.
- Gaburo, K. 1987 “LA” *Perspectives of New Music*, Vol. 25, No 1 / 2, pp. 496-510.
- Guthrie, K. 1987 *The Pythagorean Sourcebook and Library*, Phanes, Grand Rapids.
- Lhalungpa, L. P. 1975 “Tibetan Buddhism: The Way of Inward Discovery” in Needleman, J. ed. *Spiritual Tradition and Present Need*, Viking, New York.
- Johnston, W., 1967 *The Mysiticism of the Cloud of Unknowing*, Anthony Clarke, Wheathampstead.
- Mann, C. “The Use” interactive website <http://theuse.info/> (last accessed April 6, 2008).
- Plotinus 1964 *The Essential Plotinus* tr. O’Brien, E., Hackett, Indianapolis.

Steven Campbell
 School of Creative Arts
 James Cook University
 Townsville, 4811
 Australia
 steven.campbell@jcu.edu.au

Flow, Riverscape, & Come Home: Three Installation Designs for the PLaY+SPaCE Ultrasonic Sensing Environment

Abstract

Sound installation design presents numerous compositional challenges, not least of these the requirement that an installation is to be 'played' by the public. In presenting three sound installations undertaken with an ultrasonic sensing system, possibilities for sound installation design using non-tactile sensing systems within the public space are demonstrated. Each of the three installations has required a different design in sensor placement, audio assignment and spatialisation, the various designs indicative of the functionality of non-tactile ultrasonic systems in installation environments.

Installation Design

A diversity of approaches to sound installation design, particularly in Australia, is well documented in Bandt (2005, 2006). Bandt presents representative examples of installations that "include permanent public and ephemeral sculptures, time-dense computerised sound installations, museum designs, exhibits in airports, art galleries, car parks, digital and interactive media exhibitions, and real-time virtual habitats on and off the web" (Bandt 2005).

The three installations discussed here (*Flow*, *Riverscape* and *Come Home*) are all indoor temporary works, the first developed as a stand-alone sound installation within a festival, the second and third accompanying and augmenting community-based visual art works in exhibition settings. In their design, the three installations necessarily vary as a result of the size and acoustics of the installation spaces, and the nature of the audio outputs required in each setting (music, texts, environmental sounds etc). From one installation design to the next, these factors alter a wide range of design parameters including the number and placement of sensors within the installation space, the number of divisions or triggering points within a sensor's range, the number of speakers to be used, their configuration (mono, stereo or multi-channel) and placement in the space, the density of audio textures, selection of pitch and rhythmic materials, degrees of predictability and change, and levels of interactivity. In the following overviews, these parameters are discussed in relation to the design of each installation and the installation setting.

Sensing Environment

The three installation designs discussed are based on the use of the PLaY+SPaCE ultrasonic sensing environment. The PLaY+SPaCE system uses up to eight ultrasonic sensors to detect positions of people within each sensor's detection range, each sensor individually detecting human positions over a range of 10 meters, with a minimum resolution of 15cm per position. The sensors use a proprietary hardware interface to map sensor data to MIDI. A software interface developed in MAX/MSP is then used to map incoming MIDI data. The mapped data is subsequently used in further MAX patches to control audio/music outputs designed for each individual installation. Further technical attributes and the design of the overall system have been detailed previously in Campbell (2003, 2005).

As a non-tactile system, the PLaY+SPaCE environment allows installation visitors/participants to move within the space unencumbered by any physically attached sensing devices or wiring. The eight sensors of the system are generally located to form a 4 x 4 grid within a 100sqm sensing space, the sensors placed at (adult) waist height to detect movements of people moving to and away from each sensor across the space. This grid configuration was used in the first installation (*Flow*). Sensor location is not however limited to grid placements as each sensor may be placed in any location within an installation space, this allowing a considerable level of flexibility in installation design. The second installation discussed (*Riverscape*) used such an alternative approach to present three separate audio environments within a single installation space. The third installation (*Come Home*) used only two sensors placed on a single wall to create a sensing space within a small room.

Audio output configurations necessarily alter from one installation to the next, the three installations discussed here using a stereo system (*Flow*), a quad system (*Come Home*) and a stereo combined with two mono output configurations (*Riverscape*).

Whilst the discussion here is focused on the PLaY+SPaCE system, it is possible to apply the presented installation designs to alternative non-tactile systems and achieve similar results. Such systems could be based on, for example, pressure sensitive mats or video-based sensing.

Reactive Sound

From an artistic and compositional perspective, sound installations provide challenging pursuits for the audio artist in that they must be 'played' by the general public. The three installations discussed here are all 'interactive' in the most elemental sense of the term, though as discourse regarding interactivity in new media and musical contexts suggests (Bongers 1999, Manovich 2001, Winkler 1998), each of the three installations discussed relies on 'reactivity' as opposed to interactivity.

In essence, 'reactive' systems do not rely or expand on collected or predicted data (e.g. past or possible future states) to influence the current state of the system. More simply, reactive systems singularly output data in response to single data inputs. Such systems do however provide sufficiently entertaining, engaging and rewarding outputs for the public, particularly when the public is unfamiliar with the technology and the manner in which to engage with it. In many sound installations presented for the use of the public, a simple reactive design is indeed a necessity, required in order to actively engage the public with the themes and contexts of an installation.

Installation 1 – 'Flow'

Flow was designed as a work for inclusion in an annual regional festival (Riverfest) and was the inaugural public presentation of the PLaY+SPaCE system in 2002. The installation was presented in a bare space (c.80sqm) on the festival site, as shown in Figure 1. This was a purely audio installation, and with the exception of written directions for the public to explore the sounds presented through their movement through the space, no further media was provided.



Figure 1 – 'Flow' Installation Space

The work utilised a mode developed for the system in which the audio output is based on the measurement of activity levels within the space over two second periods (called 'Activity Mode'). Ten activity levels are recognized in this work, each relating to a musical environment based on differing tonal harmonic structures, rhythms and textural densities.

The work has a programmatic basis of a river and water flow. It has a very simple premise in that as activity levels increase in the physical space, so too do the levels of musical activity, these associated with the increase in the flow of a river from a calm and quiet start through to white-water rapids. The program uses various samples of water, marimba, vibraphone, percussion and a plucked piano bass, the timbral combinations, dynamic levels and harmonic tension increasing with the levels of physical activity, as shown in Figure 2.

Activity	Texture	Dynamic	Harmony
1	Vibes/Water	pp	I
2	Perc/W	pp	n/a
3	V/P/W	p	II
4	Marimba/P	p	IV
5	All	mf	V
6	All	mf	V
7	All	f	I
8	All	f	VI
9	All	ff	II
10	All	ff	I

Figure 2 – Flow Activity Levels and Musical Settings

In keeping with the programmatic basis of an ever-changing river, the differing timbral and rhythmic textures are generated in the Flow patch using a variety of controls to present sample sets at each Activity Level that appear to be constantly changing. In the case of rhythm, different combinations of (quasi) random numbers and altered metronome triggering speeds are applied at each activity level to generate rhythmic variations.

As an example, the Percussion sample set at Activity Level 2 is comprised of two conga samples, assigned with index numbers 1 and 2. A random generator for the samples provides numbers from 1 to 10 at this Activity Level, so that only the numbers 1 and 2 provide audible percussion outputs while the numbers 3 through 10 provide silence. With such a configuration the texture tends to be sparse. At the high Activity Levels of 9 and 10 the sample set is comprised of nine percussion samples. Random number generation is again set to provide numbers from 1 to 10, resulting in only one of the ten random numbers providing silence.

Additionally, at the higher Activity Levels, random numbers are used to provide changes in note triggering speeds. At the lower Activity Levels the periodicity of random number generation is set at 250ms. An underlying and ever-present ostinato in the work is set at 1000ms, set-

ting a tempo of $\frac{1}{4} = 60\text{bpm}$. The note length at the lower Activity Levels is thus $1/16\text{th}$. At the higher Activity Levels the note length is randomly changed from 250ms to 125ms ($1/16\text{th}$ to $1/32\text{nd}$ notes) providing constantly changing patterns in both rhythm and timbre.

The linear programmatic basis of this work, moving from calm waters to rapids, provides a musical architecture that will inevitably take an arch shaped form as participants strive to attain the 'rapids' then relax their movement. With the relaxation of the participants after reaching 'the rapids', the system responds with less dense textures and rhythms, and a tonal harmonic setting that eventually returns to tonic harmony.

Whilst 'Flow' was originally designed for a river festival, the design of the work results in it being applicable in further installation settings, so long as the visitor is able to relate the sounds and activity of the audio output to the river-based programme.

Installation 2 – 'Riverscape'

Riverscape was part of a 2003 exhibition commissioned by the Townsville City Council, with support from the Australia Council and Delphin. The project was multi-disciplinary, with professional artists working in conjunction with amateur artists resident in the Riverside Gardens suburb of Townsville.

The culmination of the project was the exhibition of various art works developed by nine local participants working with professional artists. A series of poems, photographic works, nine 'Concept Maps' (mixed media visual art works), banners (displaying original poetry by the participants), a 'Ribbon Map' (an aerial photo adorned with participants' names and ribbons illustrating their daily movements in and out of the locality) and soundscapes (derived from interviews with the participants and recordings made within the local environment) were all included.

The relatively small area of the exhibition space, the live acoustics of the room (polished wood floors and numerous windows) and the quantity of exhibition materials limited the number of individual audio areas available in the space, and the size of each area. The resulting assignment of audio materials was to three audio areas, relative to one of the poem banners, two of the 'Concept Maps', and the 'Ribbon Map' as shown in Figure 3.

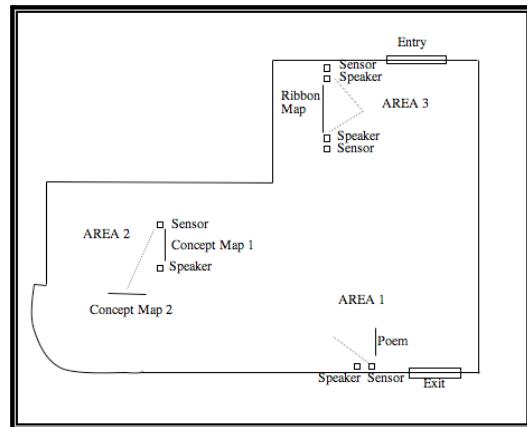


Figure 3 – Riverscape Sound Areas

In Area 1 (the poem banner) the sensor's range is divided into two triggering points. As a viewer approaches the banner, a range of environmental sounds is triggered, relative to the content of the poem. Moving closer to the poem, the viewer activates a second triggering point resulting in the playing of a recorded reading of the poem by its author. Should the viewer remain stationary, the poem will be heard in its entirety. If the viewer moves out of the sensing area for less than five seconds the reading will pause at the end of a stanza and resume on re-entry at the beginning of the next stanza. Inactivity lasting any longer than five seconds resets the recording to the start. The purpose of this was to allow visitors a short time to reflect on each stanza of the poem.

In Area 2 the 'Concept Maps' are accompanied by written texts by the artists explaining the artworks. Concept Map 1 has a central theme of transition from a life/career on the sea to life on land. Concept Map 2 is by a Kenyan immigrant and relates the journey to Australia via China.

The proximity of the two artworks to one another allowed the use of a single sensor in the area, the range of the sensor divided into six triggering points as shown in Figure 4. The trigger points nearest the artworks (Figure 4 points 2 and 3 for Map 1, and points 5 and 6 for Map 2) were assigned for audio output.

Audio outputs relative to Concept Map 1 included thirteen samples from a recorded interview with the artist, along with water-based environmental sounds relative to the artist's theme. Trigger points 2 and 3 are assigned to trigger both the interview and environmental sounds. Similarly, Concept Map 2 sounds are based on interview excerpts and further water-based environmental sounds to represent the artist's ocean journeys.

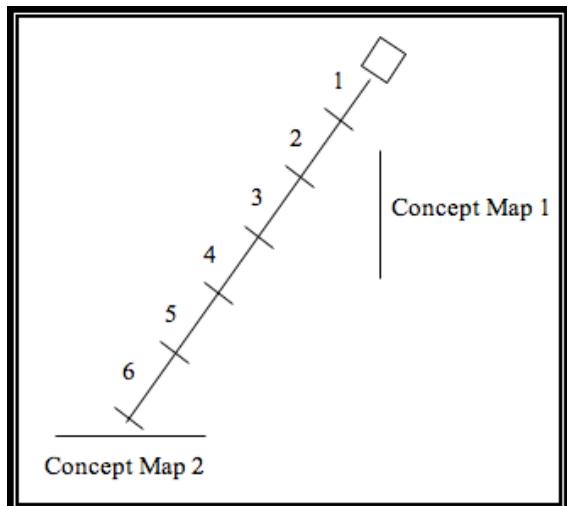


Figure 4 – Riverscape Area 2, Sensor Triggering Points

Figure 5 shows the ‘Ribbon Map’ placed in Area 3 of the installation. The large size of the map, in comparison to the artworks in the other areas, facilitated the use of two sensors, these placed so that the sensor beams crossed, as shown in Figure 6.



Figure 5 – Riverscape Area 3, Ribbon Map

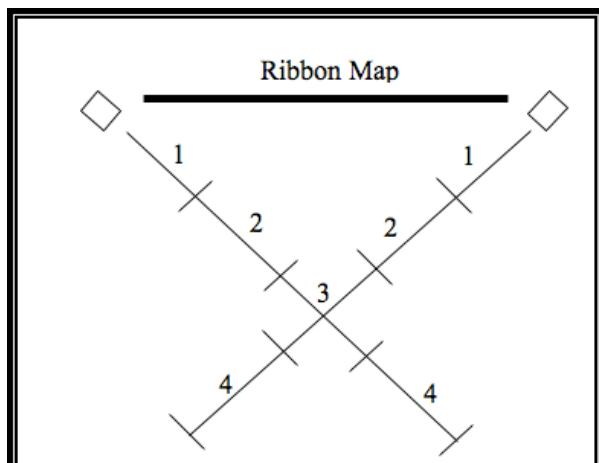


Figure 6 – Riverscape Area 3, Sensor Triggering Points

Due to the proximity of the sensors to the peripheries of the map, point 1 of the range is not used to trigger samples. Point 2 of the range, of both left and right sensors, is assigned to a series of 31 brief samples of the nine participants’ voices. The text content is primarily focussed on life in the participants’ locality, and relative to the content of the Ribbon Map.

Part 3 of the sensor range is assigned to children’s voices either reciting or singing, in alphabetical order, the names of streets in the locality. Due to the central placement of the part in the Ribbon Map area, these samples were assigned a probability of 25% in occurrence, i.e. one in four triggers would result in the samples being heard. Additionally, the samples were faded in and out over a five second period in order to avoid abrupt intrusions to the text samples.

Part 4 of the sensor range triggers environmental or ambient sounds, of which there are 44 individual samples. Content of these sound files included sounds of bicycles, commentators on boat rowing competitions, sounds of tennis matches, and so forth. These samples were triggered as visitors entered through the front door, along with the occasional children’s voice samples. Moving closer towards the Ribbon Map resulted in the triggering of the participants’ voice samples.

The samples triggered by both parts 2 and 4 were also panned in the stereo field. When a viewer triggered a sample with the left sensor, the sample would be heard from the left speaker, and similarly from the right speaker when the right sensor was triggered. As most of the samples were a number of seconds in length, the sample would be heard in its entirety before another trigger could be heard. However, if the viewer, having triggered a sample on one side and hearing it through the corresponding speaker, then moved to the other side of the map, the currently playing sample would pan to the opposite speaker to ‘follow’ the movement of the viewer.

Due to the proximity of the three areas within the exhibition and the live acoustic of the space, problems were encountered with audio spill from one area to another. These were overcome however through speakers being placed/angled in such a way as to minimise interference from one area to the next. Whilst audio from one area could certainly be heard in another, the clarity of the sound samples and the placement of the speakers provided a clear audio experience relative to each individual area, the audio from the other areas generally being indistinct and heard as background sound.

The incorporation of sound into the exhibition heightened the experience for the visitor by providing audio that the visitor could control. Further, the recordings of the artists’ interviews and the reading of the poem allowed visitors to make a greater connection with the artists and their work.

Installation 3 ‘Come Home’

Come Home was a collaborative work (2007) involving a group of professional artists working with children from three Townsville primary schools to develop artworks for an exhibition in a small empty house. The children were asked to identify an elder and under the guidance of the artists, to develop artworks based on their chosen elder and their relationship with that elder.

Stories from the children were recorded with the intention that the living room of the house would become a sound installation as part of the overall exhibition. Further recordings were made of environmental sounds from each school's locale.

The living room of the house (Figure 7) was small (approximately 3m x 4m) with a passageway marked by the entrance (the front door of the house) and an exit to a hallway as shown in Figure 8. The two sensors used for the installation were placed to make the majority of the room an active sensing space, with the exception of the passageway.

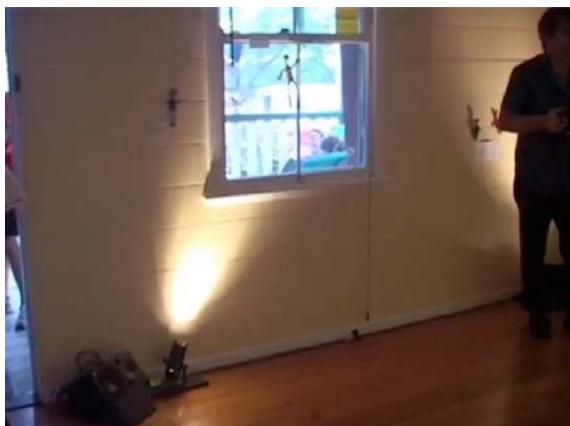


Figure 7 – Come Home Installation Space

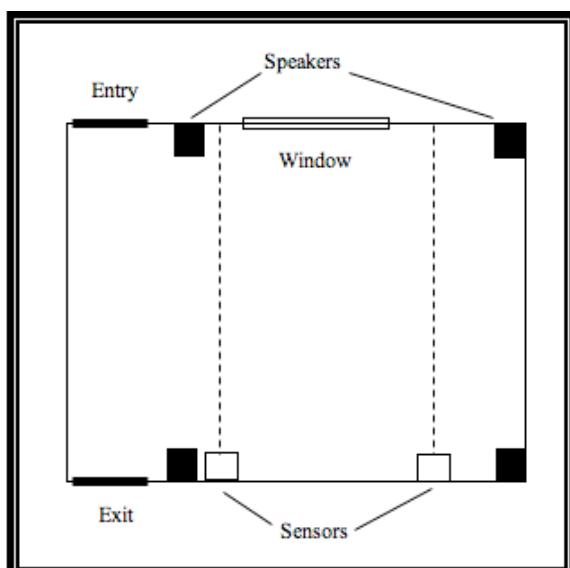


Figure 8 – Come Home Installation Space Floor Plan

The range of the two sensors was divided into two triggering points, providing a total of four triggering points for the entire space, as labelled 1 – 4 in Figure 9.

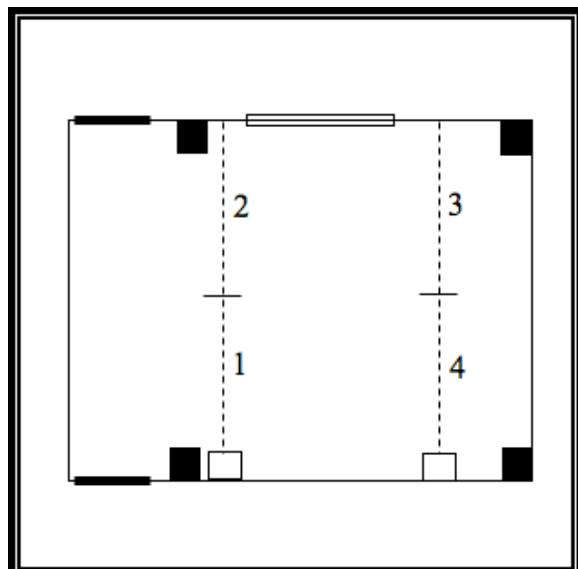


Figure 9 – Come Home Sensor Trigger Points

Even with this minimal number of triggering points, the design of the installation needed to allow for a variety of audio outcomes. The sounds were divided into three categories of environmental sounds, children's texts and short sound effects, each category triggered independently from the others. This resulted in a constantly changing sound environment, as lengths of the recordings in each category were highly varied.

The environmental recordings contained a wide variety of sounds including birds, traffic, planes and a gramophone playing. Ranging in length from 30" to over 3'00, these formed a sonic background that was triggered at all four triggering points, though limited to one recording being played at any one time. Incoming triggers were also monitored so that should there be a period of any more than 10" without a trigger from the space, any environmental sound still playing would fade out. These sounds were distributed equally to all four speakers in the quad setup.

Text recordings were typically around 30", though could extend to 2'00 in length. These were organised in the same way as the environmental sounds, i.e. triggered at all four triggering points, one text at a time, and faded out when there was no activity in the space. The texts were subject to a simple level of spatialisation wherein the audio would be heard from the nearest speaker to the triggering point. Should a visitor move through another trigger point, the audio would pan to the speaker nearest the triggering point. Using this panning, a text could be triggered and as the visitor moves around the space the text 'follows' the movements.

The sound effects were again triggered at all triggering points, and being short (under 30") did not require any fade out. Similar to the text spatialisation, the sound effects were distributed to the four speakers, however for clarity the sounds were distributed to the speaker opposite the triggering point as opposed to the speaker nearest the triggering point.

The inclusion of this sound installation within the overall exhibition augmented the experience for visitors as they were able to hear the voices of the children and to hear the children speaking about their chosen elders and their relationships with their elders. As the majority of the exhibition was based on the children's presentation of visual art works, the inclusion of this sound installation provided variety and another dimension to the exhibition.

Conclusion

The three installations discussed illustrate the diversity of approaches that may be employed in the design of installations using non-tactile ultrasonic sensing and the variations possible in design according to the installation setting. In Flow, a requirement that the work be suited to its river festival setting resulted in the PLaY+SPaCE system's Activity Mode being utilised to facilitate a programmatic depiction of the course of a river. As a non-tactile system, PLaY+SPaCE allowed the public to move freely throughout the space, using their own levels of activity to realise the programmatic and musical aspects of the work.

In *Riverscape*, the three areas of the exhibition, and the varying art works presented in each, required three different approaches to audio assignment and triggering. The basic premise of Reactivity was used in this work to provide simple audio triggering that could be easily understood and controlled by visitors. The audio materials presented within each area served to heighten visitors' experiences of the art works in each area of the exhibition.

Similarly, the *Come Home* exhibition used simple triggering of audio samples. The varying lengths of the samples used for text, environmental sounds and sound effects ensured that the audio texture created with all three categories of sounds was constantly changing. The multi-channel quad setup was beneficial in enhancing the experience of the children's texts through panning, and allowed the distribution of texts and sound effects to different speakers for clarity.

References

- Bandt, R. 2005. "Designing Sound in Public Space in Australia: a comparative study based on the Australian Sound Design Project's online gallery and database." In *Organised Sound* 10:2, 129-140.
- Bandt, R. 2006. "Sound Installation: Blurring the Boundaries of the Eye, the Ear, Space and

- Time." In *Contemporary Music Review*, 25:4, 353-365.
- Bongers, B. 2000. Physical Interfaces in the Electronic Arts. Interaction Theory and Interfacing Techniques for Real-time Performance. In Wanderly, M. & Battier M. (Eds.) *Trends in Gestural Control of Music*. Paris, IRCAM - Centre Pompidou.
- Campbell, S. 2003. "Bats, Max, Boids and Music." *Proceedings of the 2003 Apple University Consortium Conference*. Available from auc.uow.edu.au/conf/conf03/papers/AUC_DV2003_Campbell.pdf [Accessed 18th March, 2005].
- Campbell, S. 2005. "PLaY+SPaCE: An Ultrasonic Gestural MIDI Controller." *Proceedings of the 2005 Australasian Computer Music Conference: Australasian Computer Music Association*: 43-49.
- Manning, P. 2004. *Electronic and Computer Music*. Oxford: Oxford University Press.
- Manovich, L. 2001. *The Language of New Media*. Cambridge, Mass. MIT Press.
- Winkler, T. 1998. *Composing Interactive Music, Techniques and Ideas Using Max*. Massachusetts. The MIT Press.

Toby Gifford

Queensland University of Technology
 Victoria Park Road
 Kelvin Grove, 4059
 Australia
 toby.gifford@qut.edu.au

Andrew R. Brown

Queensland University of Technology
 Victoria Park Road
 Kelvin Grove, 4059
 Australia
 a.brown@qut.edu.au

Stochastic Onset Detection: An approach to detecting percussive attacks in complex audio**Abstract**

This paper introduces a new onset detection algorithm for the extraction of percussive attack times from a musical audio signal. The crux of the technique is to search for patterns of increasing noise in the signal. We therefore refer to it as the Stochastic Onset Detection (SOD) technique. This technique is designed for use with complex audio signals consisting of both pitched and percussive instrumental sounds together, and aims to report solely on the timing of percussive attacks. In contrast to most onset detection algorithms it operates in the time domain and is very efficient; suiting our requirements for real-time detection. In this paper we describe our approach to onset detection, compare this with other approaches, outline our detection algorithm and provide preliminary results from musical trials to validate the algorithm's effectiveness.

Introduction

The extraction of onset time information from musical signals is an important process for a number of applications. These include music information retrieval (MIR) systems seeking note and pulse identification, beat tracking systems for rhythmic segmentation, and real-time interactive music systems where music analysis and synchronisation are the goals. A number of publications have surveyed the techniques for onset detection (Bello et. al., 2004, Collins 2005). These surveys reveal that these techniques generally operate in the frequency domain and perform best on a particular class of onsets, with the most important class distinction being between pitched sounds and non-pitched sounds. The algorithm we present in this paper is targeted at onset detection in non-pitched sounds and operates in the time domain.

We have designed this algorithm for use in an interactive music system that performs real-time percussive accompaniment to a complex music signal. For example, a system that adds musical parts against an audio input, or where the system acts as an 'automatic' DJ. Existing techniques for

onset detection are confounded by the presence of pitched material to varying degrees. The aim of this algorithm is to perform better than existing techniques on complex audio signals, such as recordings of multi-part performances. The results of the algorithm have been evaluated by the use of mimicry – by having the algorithm play along with the audio track triggering a percussive sound when it detects an onset. The algorithm's success was assessed aesthetically by the musicality of the output, that is to the extent it detects musically significant percussive onsets and ignores insignificant ones. Audio examples that accompany this paper can be found online at http://runtime.ci.qut.edu.au/ListeningForNoise_Examples.zip

Existing techniques

In the survey of techniques for onset detection by Bello et. al. (2004) they describe an approach shared by many techniques; the input signal is distilled into a reduced form called the detection function; the detection function is then searched for recognisable features, often peak values; and these features are filtered, and then reported as onsets.

The simplest method for detecting onsets is to look for growth in the amplitude envelope. However, in the presence of complex audio signals containing multiple musical parts this technique is not viable.

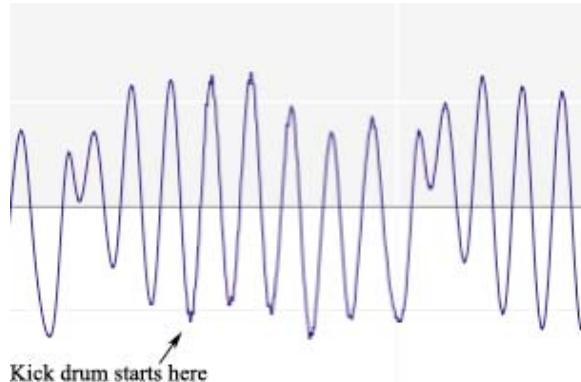


Figure 1. Attacks can be masked in multi-part signals.

For example, figure 1 shows the waveform for a sustained synthesizer note with a kick drum sound in the middle (corresponding to the example audio file *kick_and_synth.mp3*). The kick drum is clearly audible but its onset does not correspond to a peak in the amplitude envelope.

To overcome situations like this, where timbre is more significant than amplitude, a number of onset detection algorithms first split the signal into frequency bands using a Fourier transform. Onsets are then associated with growth in the energy in any band. One algorithm using this technique is Miller Puckette's (1998) bounded-Q onset detector available as the **bonk~** external for Max/MSP. Another example is the High Frequency Content (HFC) detection function of Masri and Bateman (1996) which aggregates energy across all bins but preferentially weights higher frequencies.

However, for complex audio signals in which there is power throughout the spectrum, the growth of energy in a frequency sub-band due to a percussive attack may still be masked by the ambient power of the signal in that band. The SOD technique described in this paper is designed to address this problem by seeking time domain artefacts of percussive attacks that are absent in periodic signals.

The Rapidly Changing Component

In this paper we adopt the Deterministic Plus Stochastic model of Serra (1997) for modelling musical signals. In this model a musical signal is considered to consist of a deterministic component, which may be described as a combination of sinusoids, and a stochastic component, which is described by a random noise variable. The crux of our onset algorithm relies on the assumption that a percussive onset will be characterised by an increase in the noise component of the signal.

In Serra's model noise is equated with randomness. For example a totally random (digital) signal would be one where the amplitude of the signal at each sample point is drawn from a probability distribution and is independent of the amplitude at any other sample point.

Informally speaking our algorithm operates by separating the stochastic component from the deterministic component of the signal, and then making two queries:

- (i) how loud is the stochastic component?
- (ii) how random is the stochastic component?

It may seem superfluous to measure the randomness of the stochastic component; presumably if we have done a good job of the separation then it will be totally random. The reason for making this measurement is that a perfect separation is, perhaps, impossible and certainly time consuming. Instead of seeking a perfect split our algorithm operates in two steps; first by separating out a portion of the signal that contains the stochastic component plus some small amount of the deterministic component, and then estimating the frac-

tion of this separated portion that is due to the stochastic component of the signal.

To this end we define a new notion called the Rapidly Changing Component (RCC). The RCC can be thought of as the zig-zags in the signal. For example, referring to figure 1, the signal is smooth when the synthesiser is playing on its own. The time at which the kick-drum starts is visually discernible because the signal becomes rougher (more zig-zags) at that point.

The RCC consists both of high frequency sounds and noisy sounds. Our algorithm operates by separating out the RCC from slower moving components, and then measuring the loudness of the RCC and estimating what fraction of the RCC is due to noise.

People often think of white noise as having a 'flat' Fourier spectrum, in other words equal power at all frequencies (within some band). However this picture is somewhat misleading, at least for noise as we are talking about it in this paper. In fact, if a digital signal is completely random then its spectrum is also completely random. There is a sense in which the spectrum can be described as flat - namely that the power of the spectrum in any given frequency bin will be a random number drawn from the same distribution as every other bin. But for any particular window of signal (here we are talking about the Short Time Fourier Transform with a rectangular window) the actual spectrum will not have equal power in all bins – it will be totally random. And so from one analysis window to the next there will be no relationship between the spectra, save that the total energy will be approximately the same. Consequently onset detection algorithms that operate by looking for patterns in the spectra of successive windows are ill-suited to detecting noise.

Transient detection

A straightforward onset detection technique is to look for growth in the energy of the signal. However, for complex audio signals where the amplitude of pitched material exceeds that of the percussive onsets as shown in figure 1, simple amplitude tracking will not suffice.

As discussed above, many transient detection schemes look for growth within frequency bands. From the preceding discussion we would expect that random noise would appear in various different frequency bins inconsistently from one window to the next. This is often informally referred to as smearing¹. One method devised to deal with this situation is the High Frequency Content (HFC) technique (Masri & Bateman, 1996). As the name suggests this approach aggregates all of the energy in high frequency bands (to be precise it aggregates all bands but linearly weights by fre-

¹ Strictly speaking smearing is where a particular frequency shows up in several frequency bins due to the quantised nature of the Short Time Fourier Transform. The use of this term in the above context is inappropriate in the same way that the use of the STFT to detect noise is inappropriate.

quency). Doing so avoids the problems of smearing to a large extent.

So why is the HFC suited to finding noise? We suggest one aspect of this can be understood with reference to Serra's Deterministic Plus Stochastic model. The stochastic component should be random at all time scales. In particular it should be random from one sample to the next, so that a burst of noise should create an increased 'jaggedness' of the signal at very short timescales, or high frequencies.

What are the drawbacks of the HFC approach? It suffers from the same basic problem as a direct amplitude approach but in more limited circumstances; if the periodic part of the signal has a lot of energy in high frequencies, then the growth in the HFC due to the percussive onset may be small in comparison to the ambient level of HFC, degrading the signal/noise ratio for the detection function.

What can be done about this? Our approach is to look at the short-timescale activity and measure how random it is. Then we can look at the growth in that randomness. This way the presence of background high frequency periodic content will not affect our detection function and the beat detection will be more robust and reliable.

Description of the SOD Algorithm

Our Stochastic Onset Detection (SOD) algorithm is designed for real-time use with minimal latency. The input signal is processed in short windows of 128 samples. Each window we measure the level of noise in the signal. This measurement consists of four steps:

1. Separate out the RCC
2. Measure the size of the RCC
3. Measure the randomness of the RCC
4. Estimate the loudness of the stochastic component – this is our detection function.

Having obtained the most recent value for the detection function, we then employ an adaptive peak-picking algorithm (described below) to look for significant growth in the noise. Points of significant growth that exceed an absolute noise threshold are marked as percussive onsets.

Splitting out the RCC

The first step in the construction of our noise measure is separating out the Rapidly Changing Component (RCC) from the rest of the signal. To do this we use a little rocket science - drawing inspiration from a technique developed at NASA called Empirical Mode Decomposition (Huang et. al., 1998). This is a technique for extracting 'modes' from a non-linear signal, where a mode may have a varying frequency through time. The basic idea is that to get the RCC we look at adjacent turning points of the signal (i.e., the local maxima and minima) and consider these to be short timescale activity around a carrier signal which is taken to be halfway between the turning points. It's a bit like creating a smoothed carrier

wave by using a moving average with a varying order, and taking the RCC to be the residual of the signal from the carrier wave. The process is illustrated in Figure 2.

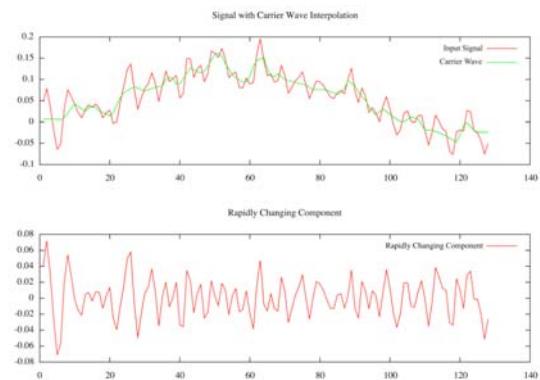


Figure 2: Splitting out the RCC.

Stochastic components of the RCC

Generally the RCC will be comprised of both the stochastic component of the signal and high frequency parts of the deterministic component. So as to get a sense of the relative sizes of these contributions to the RCC, we make a measurement of the level of randomness in the RCC.

The statistic that we use to measure the level of the stochastic component is the first order autocorrelation, which measures how related the signal is to itself from one sample to the next. The stochastic component of the signal should have each sample statistically independent, and so will have an autocorrelation of zero. The deterministic component, on the other hand will be strongly related to itself from one sample to the next, and so should have autocorrelation close to one. The autocorrelation of the RCC will then reflect the relative amplitudes of these two components of the RCC; an autocorrelation of close to zero means that the RCC is mostly stochastic, whilst an autocorrelation close to one means that the RCC is mostly deterministic.

Another measure of the randomness that could be considered is the signal entropy (Shannon, 1948). The use of entropy in searching for changes in the signal noise was explored by Bercher & Vignat (2000), who give an adaptive procedure for estimating the entropy. However, their procedure is not intended for real-time use, indeed the calculation of entropy is computationally expensive (Hall & Morton, 2004). Furthermore, the autocorrelation measure has the advantage that it has a direct interpretation as approximating the percentage of the RCC that is deterministic. Conversely, if we take our measurement of randomness to be $1 - c$ where c is the autocorrelation, then this will be an approximate measure of the percentage of the RCC attributable to noise. For these reasons we prefer the autocorrelation measure to entropy.

Description of Noise Measure

Having extracted the RCC we can report on how loud it is. Then, having also estimated the stochastic component of the RCC, and hence the approximate percentage of the RCC attributable to noise, we can make an estimate of the loudness of the noise in the signal by multiplying the amplitude of the RCC by its stochastic component. In more detail, our noise measure is constructed as follows:

1. Split the signal into rectangular analysis window (we have used a window size of 128 samples).
2. Calculate the Rapidly Changing Component
 - (i) Find the turning points of the signal
 - (ii) The carrier wave is assumed to be halfway between adjacent turning points of the signal, so construct the carrier wave by linearly interpolating between these midpoints
 - (iii) The Rapidly Changing Component is the difference between the signal and the carrier wave.
3. Calculate the size of the RCC:
 $\text{Size}_{\text{RCC}} = \text{Std. Dev. of the derivative of the RCC}$
4. Calculate the randomness of the RCC
 $\text{Randomness}_{\text{RCC}} = 1 - \text{autocorrelation of the RCC}$
5. Calculate the noise
 $\text{Noise}_{\text{RCC}} = \text{Size}_{\text{RCC}} * \text{Randomness}_{\text{RCC}}$

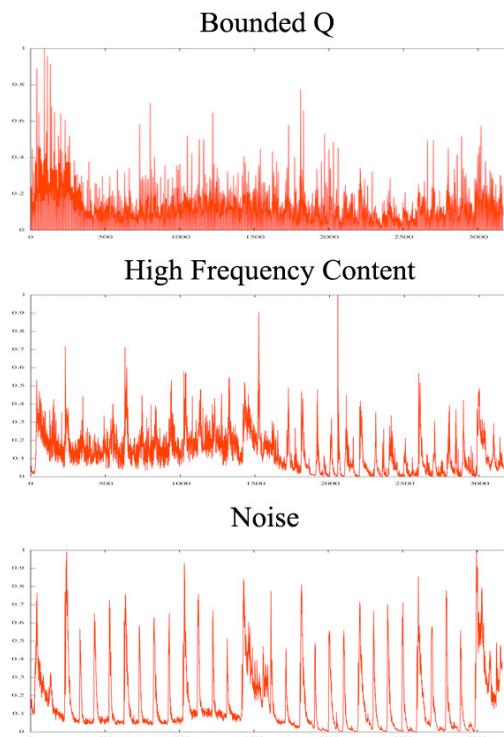


Figure 3: Comparison of Detection Functions.

An example of the Noise detection function for a short audio sample (corresponding to the file JungleBoogie.mp3 in the examples) is shown in figure 3. Also shown for comparison is the HFC detection function, and a Bounded Q detection

function similar to that used by **bonk~**. The horizontal axis is time, and onsets are identified by looking for peaks in these detection functions. The noise detection measure has much more clearly discriminated peaks than the other two.

Adaptive Thresholding

Having calculated the noise function we then want to identify peaks, which we will interpret as percussive attacks. In fact, what we are really looking for is sudden growth in the noise, followed by a peak, and then a decay. To do this we look for ‘significant jumps’ in the noise function. Different pieces of music may have markedly different noise characteristics; the size of a jump which is significant will depend on the ratio of the ambient noisiness of the pitched instruments compared to percussive instruments. To deal with this variation between musical signals we have used an adaptive thresholding technique.

We maintain a measure of the mean and standard deviation of noise in the recent past using an Exponentially Weighted Moving Average. For each new window we update these measures by accumulating a weighted value of the preceding window (we currently use a weighting of 8%). So for each new window the measures of mean and standard deviation of recent history will be 92% of what they were before + 8% of the values for the immediately preceding window. This process allows us to identify a significant jump in the noise level: where the noise level is some number of standard deviations above the mean of the recent past.

Once an onset is detected using this technique, it is not necessary to report any more onsets until the current attack is completed. A common strategy for measuring attack completion is to maintain a high and low threshold; where, for an onset to be reported the detection function must exceed the high threshold, and then no further onsets will be reported until the detection function has dropped below the low threshold. We have utilised an adaptive version of this technique for reasons mentioned previously. Once a significant jump is detected, an ongoing measure of the peak value of the detection function is maintained, and the attack is considered to be ongoing until the detection function has dropped sufficiently that recent past is significantly lower than the peak (using the same exponentially weighted moving average scheme as for detecting the onset).

The detected onsets are then further filtered by an absolute noise threshold. To be considered as an attack, a significant jump must have a peak value higher than this threshold. To allow for real-time responsiveness to the signal with minimum latency, the onset is allowed through the filter as soon as the ongoing measure of its peak value exceeds the noise threshold. For example, an open high-hat onset will have a rapid increase in the noise level but not a quick decay – so that if the algorithm were to wait until the noise had peaked

before reporting the onset it would have significant latency.

Computational Efficiency

The stochastic onset detection (SOD) algorithm presented in this paper is quite efficient. No FFT is required because it works in the time domain. In our real-time implementation, a 128 point sample window took approximately 8 samples to process. It is also quite responsive because the RCC measurements can be calculated on small sample buffers, typically as small as 32 samples providing a latency of less than 1 millisecond.

Experimental Results

We applied this algorithm to a selection of audio snippets containing complex audio with percussion. The snippets can be found in the online examples accompanying this paper. In addition to our few hand-selected tracks, we tested the algorithm against the MIREX Audio Tempo Extraction training data set. Training snippets in this set that did not have any percussion parts were omitted.

The Noise detection function generally seems to have a superior signal to noise ratio than the HFC or Bounded-Q detection functions. For example referring back to Figure 3, of these three detection functions the Noise detection function has the most clearly defined peaks.

We evaluated the algorithm by having it ‘jam’ along with the audio track (in real-time) mimicking what it hears by triggering a MIDI percussion sound when it detects an onset. The Noise measure also gives an estimate of the amplitude of the onset, and so this information is used to determine the velocity of the MIDI imitation. As a contrast, we performed the same trials using the **bonk~** external for Max/MSP.

These results are preliminary in that we have tested the algorithm with a limited range of musical examples and only performed aural analysis of the results. However, they clearly show that our approach is generally more robust than the algorithm in **bonk~** but is still not entirely consistent. In particular, our algorithm makes few mistakes in detecting onsets but does not detect all onsets. The onsets it does predict do not always correlate with those that seem most significant to human judgments, but this is not surprising given that our algorithm does not build expectations about pulse as humans do.

The algorithm appears to be particularly attuned to high-hat and cymbal onsets. For example, referring once again to Figure 3, in the snippet from Jungle Boogie, the Noise Detection algorithm follows the high-hats solidly, whilst the HFC algorithm appears more drawn to the guitar rhythm (and the Bounded-Q algorithm is totally at sea). The evaluations of these three algorithms may be heard online in the examples as JungleBoogie_nd.mp3, JungleBoogie_hf.mp3, and JungleBoogie_bq.mp3.

The other examples are of the form name_nd.mp3 for the Noise Detection sample and name_bk.mp3 for the Bonk sample.

Conclusions

In this paper we have presented a new approach to onset detection of percussive sounds in audio signals we call Stochastic Onset Detection. This approach works with complex audio signals that have a polyphonic mixture of pitched and unpitched parts. Our approach analyses signals in the time domain and detects percussive onsets by measuring significant changes in the noise component of the signal that is typically associated with percussive attack transients. We have developed an algorithm based on this approach and provided preliminary test results that indicate that it is efficient and effective. The algorithm seems to be particularly good at detecting high pitched percussive sounds such as high-hats, which could be useful for tempo tracking of dance / rock tracks as the high-hat is often used to keep the pulse.

We hope to pursue further comparative testing with existing onset detection methods using the same hand marked test database as a benchmark for comparison used by Bello et al. (2005) and Collins (2005). We have plans to undertake future developments of this approach that include the addition of predictive assistance based on regularities and psychoacoustic models of expectation that we anticipate will particularly allow for variations in transient attack rates and allow the algorithm to have more sense of syncopated or irregular rhythms.

References

- Bello, J. P., Daudet, L., Abdallah, S. A., Duxbury, C., Davies, M., & Sandler, M. (2005). A Tutorial on Onset Detection in Music Signals. *IEEE Transactions on Speech and Audio Processing*, 13(5), 1035-1047.
- Bercher, J., & Vignat, C. (2000). Estimating the Entropy of a Signal with Applications. *IEEE Transactions on Signal Processing*, 48(6), 1687-1694.
- Collins, N. (2005). A Comparison of Sound Onset Detection Algorithms with Emphasis on Psychoacoustically Motivated Detection Functions. *Proceedings of the AES 118th Convention*, Barcelon, Spain.
- Hall, P., & Morton, S. (2004). On the Estimation of Entropy. *Annals of the Institute of Statistical Mathematics*, 45(1).
- Huang, N., Shen, Z., Long, S., Wu, M., Shih, H., Zheng, Q., et al. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis. *Proceedings of the Royal Society of London A*, 454, 903-995.
- Masri, P., & Bateman, A. (1996). Improved Modelling of Attack Transients in Music Analysis-Resynthesis. *Proceedings of the International Computer Music Conference*, Hong Kong.

- Puckette, M., Apel, T., & Zicarelli, D. (1998). Real-time audio analysis tools for Pd and MSP. *Proceedings of the International Computer Music Conference*.
- Serra, X. (1997). Musical Sound Modeling with Sinusoids Plus Noise. In C. Roads, A. Piccialli & G. De Poli (Eds.), *Musical Signal Processing*: Swets & Zeitlinger.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell Systems Tech Journal*, 27, 379-423; 623-656.

Callum Hedemann

Queensland University of Technology &
Australasian CRC for Interaction Design
Kelvin Grove, 4059
Australia
callum.hedemann@qut.edu.au

Andrew Sorensen

Queensland University of Technology &
Australasian CRC for Interaction Design
Kelvin Grove, 4059
Australia
andrew@moso.com.au

Andrew R. Brown

Queensland University of Technology &
Australasian CRC for Interaction Design
Kelvin Grove, 4059
Australia
a.brown@qut.edu.au

Abstract

This paper outlines the development of the Metascore application which is the result of investigations into visual interfaces for the control and coordination of parameters for generative film music scores. Metascore is part of a project under development at the Australasian CRC for Interaction Design (ACID), the aim of which is to create a real-time automatic film scoring system. Metascore allows the user to control generative music algorithms using high-level descriptions of compositional intent and in synchronisation with video. It provides control of constant parameters within sections that demarcate film cues, and control of continually varying parameters through free-form break point envelopes. Parameter values in Metascore can themselves be expressions or functions, enabling extensive flexibility for adaptive control of the algorithmic score. This paper will describe the features which we have found to be most useful in the composition of generative film music the design of the Metascore interface which exposes these features to the user.

Introduction

Film music provides both structural and emotive support to the narrative of a film. As a consequence, film music is constrained by the pace and structure of the film's narrative. Metascore is a film scoring and composition environment designed to make the setting and adjustment of video synchronised musical parameters quick and easy to manage. In combination with generative processes, Metascore is designed to make the composition process flexible and efficient, whilst

Metascore: User Interface Design for Generative Film Scoring

providing the user with sufficient control over the development of the music to meet the expressive demands of the film narrative. Metascore's controls allow a score to be defined using descriptions of the video's structure, such as the timing of film cues and scene changes, and abstracted music descriptions, such as the harmonic tension or rhythmic density. These descriptions are then processed by Metascore to generate the film score.

When developing a film scoring tool, important considerations include the emotional content of the music and the time-based issues around synchronisation of music to video and visual events. This paper is primarily concerned with issues surrounding synchronicity and the design of a user interface to accommodate them; the compositional algorithms controlled by Metascore can vary from style to style and is discussed in related papers.

Design Considerations

When composing for film the integration of music and video needs be addressed. A general pattern in the film industry is that music must be adapted and shaped to fit the film. In the design of Metascore and its interface, we have considered two ways to integrate music with film.

Firstly, it is common practice for the score to emphasize and articulate the actions in the film, so that musical events coincide with visual events or actions. This may involve placing strong beats in the music to land on hit points, matching sectional changes in the music with scene changes, using crescendos to build up to a plot event, or other musical articulation of visual events. In order to accommodate these

cases, the tempo and time signatures used in the score must be carefully managed to ensure that the music is synchronised with the film. Given that film cues can fall at any interval, synchronisation attempts can lead to rather irregular musical structures. It can be a challenge for composers to comfortably fit the music around them to maintain the pace of the music.

A second scoring consideration is matching the emotional content of the music with the emotions in the film. The mood of the music must follow the film in an appropriate manner, usually done to achieve empathy or contrast with the intent to create a heightened collective experience. Adding to the challenge, the mood of a film can change either progressively over time or abruptly from scene to scene. The composer must have sufficient control over the music to allow both subtle and agile emotional change.

These are issues that we have had to consider while creating Metascore. The design of Metascore's user interface accommodates both structural synchronisation and composition control through the implementation of a number of specialised features. In contrast to the majority of film scoring applications currently available (Sony Creative Software Inc. 2008, Smartsound Software Inc. 2008), Metascore uses compositional algorithms to generate its musical material. This allows the user to set the parameters of the music, such as key structural points and compositional elements, in order to control the timing and mood of the score over time. Metascore then takes care of calculating how to fit the music to the score, a task usually undertaken by the composer.

The interface has been developed to facilitate our research of the underlying generative algorithms employed. This means that it is not intended to be end-user ready, which is apparent, for instance, in the fact that some parameters require Lisp expressions. Although this would be unsuitable as a final release, it allows us as developers to powerfully express the input for these algorithms. It helps us to discover what types of user input and functionality are effective and allows us to explore various possibilities for end-user interface design. The Metascore interface also allows us to rapidly test our ongoing work on the underlying compositional algorithms.

Similar Projects

There are a number of other projects which are related to our work on Metascore. Some of these include:

Both Cinescore (Sony Creative Software Inc. 2008) and Sonicfire Pro (Smartsound Software Inc. 2008) are commercial scoring tools which take the approach of adapting pre-recorded fragments or themes to video, achieved primarily by stretching or compressing passages to fit the length of a scene or section. The user outlines the structure of

the scenes in the film with a series of markers and selects a recorded fragment which suits the mood they desire. While the features and implementation vary between the two programs, both are constrained to the library of available recordings and limit the user's creative input, yet provide a quick and easy way to create accompaniment for a film.

Scorebot (Pierce 2000) concerns the theory and design of an automated film scoring tool and, although it was not implemented, identifies many issues surrounding generative film scoring. Its design explores the manipulation of pre-composed music to be used as a film score by processing a given theme to align with a set of scene information and cues, such as the timing, mood and events associated with the scene. It is proposed that this be achieved through altering the compositional aspects of the theme, including the dynamics, modality, rhythms, tempo and dissonance of the themes.

Qsketcher (Abrams, Bellofatto, Fuhrer, Oppenheim, Wright, Boulanger, Leonard, Mash, Rendish, Smith 2001) is a composition environment that considers the creative workflow of composition (in the context of film music) and ways in which to enhance the role of the computer in this process. Its design addresses the representation of musical ideas, their context and how they relate to one another, in such a way as to support the cognitive association with composition and creative workflow.

Hyperscore (Farboud, Kaufman, Jennings 2007) is a graphical composition environment in which the user composes music by drawing various lines, curves, shapes and textures. The graphical information is then interpreted by Hyperscore's compositional algorithms (in accordance with musical motifs) to produce the composition. This gives non-musically inclined users the ability to express compositional intent in an abstract manner and to explore musical creativity. While not directly applicable to film composition Hyperscore is another significant commercial application in the generative music space and is one of the only tools which also investigates the generation of orchestral scores.

These applications all identify pertinent issues concerning the creation of music synchronised to film and the user interface design required to facilitate this. Each deals with these issues using various approaches, providing different advantages through their implementation.

Metascore overview

Metascore is being developed in the Impromptu environment (Sorensen 2005) and makes use of the Vienna Symphonic Library Special Edition for sample playback. It facilitates the experimentation of algorithmic processes for genera-

tive film music and allows the user to control parameters of the underlying generative algorithms that focus on high level artistic issues such as intensity, complexity, tension and the like. These

higher level considerations are often referred to as meta structure (Xenakis 1992) and control of them as meta-level composition.

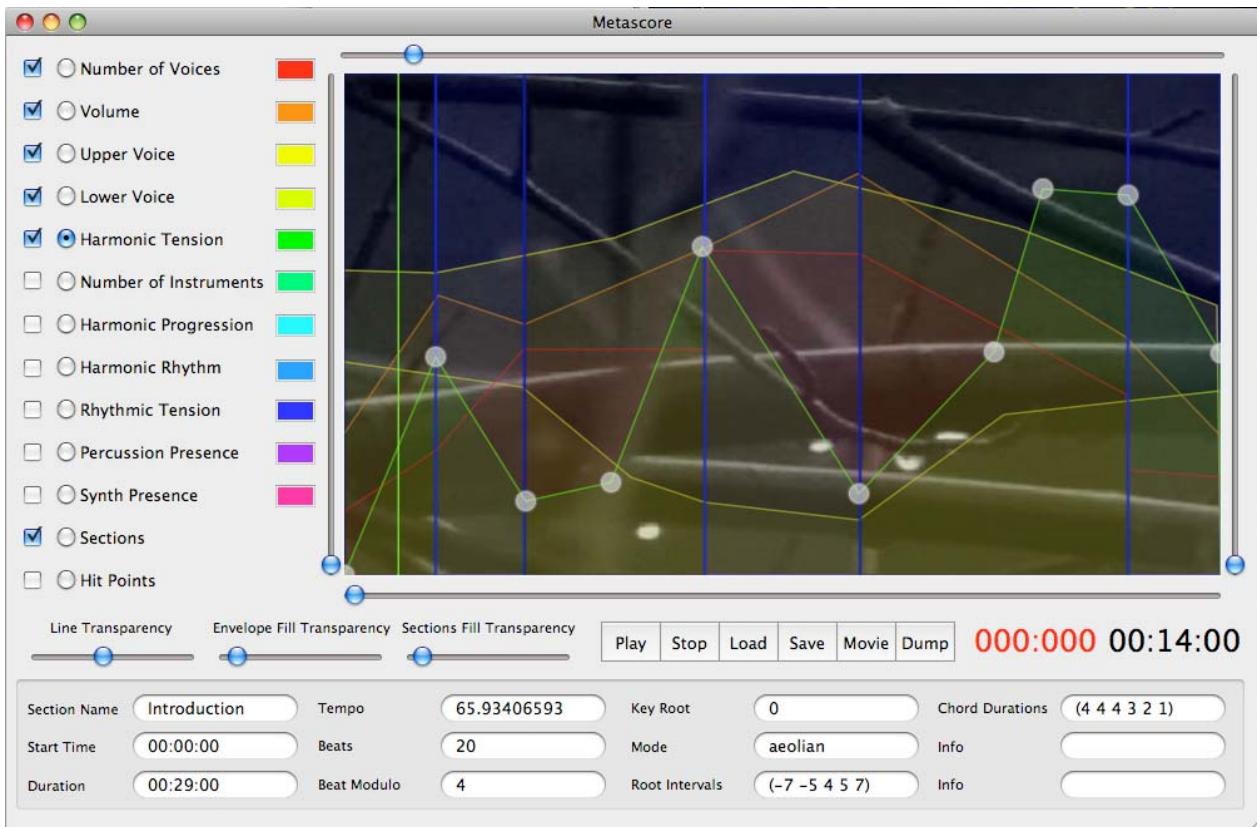


Figure 1. Screen shot.

Sections

Metascore enables the division of the score into sections which group together global parameters for specific parts of the film. Each section has its own set of values that are used to provide local state for the duration of the section. When combined, these sections outline the development of the entire composition over time. The main function of sections is to provide high-level musical structure with a cohesive relationship to the underlying film. Section markers are used to segment the film into cues. Metascore calculates bars, beats and tempi to ensure that natural musical boundaries always match cue points. Various parameters, including tempo, key and chordal information, are sectional and their settings remain constant for the duration of a given section.

In addition, Metascore provides a layer of parameter envelopes that allow for continual variation in parameters over time. Parameter values are interpolated between break points in the envelopes, which are used as input into the compositional algorithms that generate the music.

In figure 1, section markers are indicated by vertical lines on the video display area, which acts as a time line. The various parameter envelopes are shown overlaid on top of these vertical section markers. The position within, and resolution of, the time line is controlled by the sliders beside and above the display. The left side of the interface shows the musical parameters than can be controlled by envelopes and allows for them to be toggled on and off. The lower part of the display includes transport controls and fields for parameter values.

Tempo Snapping

The use of sections relates to events in the film on a macro scale. One important structural aspect of music, as distinct from film, is the grouping of regular repetitions of pulses into bars and beats. Film duration is usually measured in minutes, seconds and frames. In order to make the translation from the film world, based in irregular durations in frames, into the music world of more regular temporal patterns in beats, it is important to determine a musical tempo that will accommodate the requirements of both domains. Metascore allows the user to specify a tempo for

each section, but in such a way as to ensure that it contains a whole number of beats or bars. The specified tempo is adjusted to the nearest value which would accommodate a whole number of beats over the duration of the section. This calculation will usually push the tempo up or down by up to five beats per minute. If the new tempo is too far from the composer's intent, a change to the number of bars and beats can be made and the process repeated. When applied to all sections, this effectively synchronises the beats of the music with the structure of the film. These Tempo Snapping calculations occur whenever a section is created or modified.

Beat Modulo

A feature related to tempo snapping is Beat Modulo. This allows for the grouping of beats, so that a whole number of these groups fit within the duration of the section. Again, this is achieved by further modification of the tempo. This feature is useful for specifying the bar length used in the section (related to specifying a time signature), which prevents the use of an irregular number of beats and allows for further constraints over the musical length of a section while maintaining tight links to the structure of the film.

Sectional Harmonic Parameters

The remaining variables for a section are the key, mode, chord intervals (steps between sequential chords) and chord durations. These are important parameters that greatly influence the musical character of each section. For instance, the pitch scale used in a section can be easily defined by specifying the root note of the key and name of the desired mode.

Another useful feature is the ability to supply a list of possible root progression intervals and chord durations which are used as input to the underlying compositional algorithm. This allows the output of each section to be specified in more detail. These features can be used to support the linear structure of a film composition. When these parameters are used effectively, and in coordination with the other sections, it is possible to produce a well controlled harmonic basis for the overall composition.

Hit Points

Even with the use of sections to align the composition and film structures, there still may be cases where particular off-beat actions within a section of the film require musical accentuation. This is a situation commonly encountered in film composition and is known as a hit point. Metascore deals with hit points by subdividing a section at the point where this accentuation is needed and forcing the first beat of the nearest bar to lie on

the hit point. Subdivisions created before and after the hit point are tempo snapped. Both subdivisions exist as a part of the same section and still share all of the other section characteristics.

We also plan to provide independent hit points, which are unrelated to any beat association or synchronisation, although at present our interest is in seeing how far we can push these tempo relationships.

Envelope Control

Metascore provides control over many compositional elements via continuously varying parameters. These are most frequently used to express compositional intent through features such as volume, orchestration, melodic continuity, texture, pitch compass and the like. The details and workings behind these elements and algorithms are discussed at greater length in related papers about generative music algorithms that Metascore can control (Sorensen & Brown 2008).

The control of these elements is given to the user as a series of break-point envelopes that consist of a custom number of editable nodes as shown in figure 2. These envelopes represent the value or state of a parameter over the duration of the entire film. Their values are used as input into the compositional algorithms which, in turn, generate the score on a note-by-note level in real-time. The various combinations and arrangement of these envelopes, in conjunction with the various section parameters discussed earlier, are the building blocks of any given composition.



Figure 2. Break point envelope in detail.

The envelopes themselves can be easily modified, allowing the user to create new nodes, drag around existing ones, or delete unwanted nodes from the break-point envelope. A key editing feature is the fact that the envelopes are overlaid on to the video. This makes for easy editing, as the user can drag the time slider to specific visual events and align envelope nodes with precision, ensuring that key compositional points match up with corresponding events in the film. The user can zoom in on an envelope for even finer control. Additional editing capabilities include the option to view multiple envelopes simultaneously, as well as to control the transpar-

ency of the overlayed envelopes. Envelope display options can be used to enhance the editing process by putting related envelopes into simultaneous view and editing them in concert. Envelopes can also be used to depict a meta “score”. By increasing the transparency of the envelopes it is possible to view, and hear, the interaction between the compositional parameters in real-time.

With effective arrangement of the various envelopes over the duration of the film, the user has essentially composed the film’s score using the contours of the envelopes. This method of composition emphasises a balance between control and efficiency. Control of the composition on a note-by-note level is delegated to the generative algorithms and the user engages with envelope control of compositional elements over time.

The intention behind these compositional envelopes is that they can influence the emotional content of the generated music. For instance, the envelopes could be arranged to indicate action by using a high rhythmic tension value, high harmonic tension for suspense, and a large number of instruments for a more general crescendo. But these examples are overly simplistic. A more accurate description of their use would include the careful arrangement of all the parameters working together to create a specific mood. Potentially, the more parameters that were available, the greater the possibilities for more complex emotions. The integration of parameters to achieve specific moods is an area we are still investigating through the development of Metascore. However, even with the current editing capabilities, it is quite possible to arrange the compositional elements to accurately coincide with the film’s emotional content and to create some basic moods in the music to match. The governing idea is that with the interface features described above, the user will be able to sculpt the mood of the music around the structure of the film, using the compositional envelopes in any combinations they see fit.

Using Metascore

The method of composition provided by Metascore is reliant on both the ability to compose accurately with the interface and on the quality of the interpretation by the algorithms. The accurate realisation of user compositional intent has been one of our main concerns in the development of Metascore’s interface. This has lead us to provide the user with meta-level control over the composition which allows us, as researchers, to explore the full range of output that Metascore is capable of generating.

The inherent limitation of this method is the fact that it takes away the absolute control of composition usually available to the composer. However, this limitation is counter-balanced by

the efficiency of synchronisation between music and film, the ability to compose on a structural level, the ability for music to follow edits in the film, and the ability to rapidly and interactively experiment with compositional options.

Conclusion

This paper has introduced a new interface for film composers designed to support the structural and emotional support for the actions and events which occur visually in the film. In order to fulfil these roles successfully, synchronisation between the structure of the music and the structure of the film is vital. As film structures are irregular in comparison to common musical structures, the musical structures for film require special consideration. The design of the Metascore interface addresses these issues for the generative music composer and provides a platform for efficient experimentation with a variety of parametric settings and thus a quick exploration of the compositional space occupied by any particular compositional algorithm.

We have discussed how the use of sections can support the kind of linear development which is effective in the composition of film music, particularly for harmonic and metric structure. We have described how Metascore makes use of section breaks and parameters envelopes to align musical features with the structure of a film and to potentially craft the emotional content of the music in relation to the visuals.

The future development of Metascore lies in other areas inherent to film composition, such as the use of pre-composed themes and thematic development. As this development continues, the interface design of Metascore will evolve also to allow for the effective experimentation of generative music for film.

Acknowledgments

The work reported in this paper has been supported by the Australasian CRC for Interaction (ACID) through the Cooperative Research Centre Programme of the Australian Government’s Department of Education, Science and Training.

References

- Abrams, S., Bellofatto, R., Fuhrer, R., Oppenheim, D., Wright, J., Boulanger, R., Leonard, N., Mash, D., Rendish, M. and Smith, J. (2001), ‘Qsketcher: An Environment for Composing Music for Film’, Proceedings of the International Computer Music Conference. ICMA, Havana, Cuba.
- Farboud, M., Kaufman, H., Jennings, K. (2007). *Composing with Hyperscore: An intuitive interface for visualising musical structure*. Proceedings of the International Computer Music Conference.

- Conference. ICMA, Holmen Island, Copenhagen
- Pierce, M. (2000), 'Scorebot: Theory and Design of an Automated Film Scoring Application', Master's thesis, University of Leeds, Leeds.
- Smartsound Software Inc. (2008), *Sonicfire Pro: The Future of Library Music*, 10 April 2008, <<http://www.smartsound.com/sonicfire/>>
- Sony Creative Software Inc. (2008), *Cinescore: Professional Soundtrack Creation*, 10 April 2008, <<http://www.sonycreativesoftware.com/cinescore>>
- Sorensen, A. (2005). Impromptu: An interactive programming environment for composition and performance. *Proceedings of the Australasian Computer Music Conference 2005*. Brisbane: ACMA, pp. 149-153.
- Sorensen, A. and Brown, A. (2008). A computational model for the generation of orchestral music in the Germanic symphonic tradition: A progress report. *Proceedings of the Australasian Computer Music Conference 2008*. Sydney: ACMA.
- Spring comes to the mountaintop* (2007), video recording, open source, produced by C. E. Price. Creative Commons license <<http://creativecommons.org/licenses/by-nc-sa/3.0/>>
- Xenakis, I. (1992). *Formalized Music: Thought and Mathematics in Music*. Stuyvesant NY, Pendragon Press.

Sarah Keith

Macquarie University
 Dept. of Contemporary Music Studies
 North Ryde, 2109
 Australia
 sarah.keith@humn.mq.edu.au

Abstract

The incorporation of new and unfamiliar technologies to musical performance is a recurring practice, both within experimental and avant-garde practices and within more commercial settings. The use of the laptop in real-time music creation represents one example of this incorporation, which occurs across a range of musical and performance settings. A framework has been composed which introduces a range of features to be considered in performance design incorporating new technologies. A central assumption of this framework is the nature of performance as a culturally mediated event dependent on contextual factors. This paper describes the theoretical foundations for situating performance as a culturally defined procedure, how this supposition relates to the framework, and discusses how the framework may be practically applied.

Introduction

This paper explores a context-focused approach to designing technologised musical performance. It will examine the theoretical basis for this stance, and introduce a four-step framework building on this principle. This framework is intended for practical implementation in the construction of novel performance systems.

Current Research

The formulation of this framework constitutes a part of the author's current objective of designing a laptop-based performance environment for contemporary electronic music. The term contemporary electronic music refers to the popular music form commonly referred to as "techno" or "electronica" which originated in the mid-1980s and exists in diverse manifestations today. Performance approaches to music of this broad type are historically derived from the turntable DJ paradigm. Throughout the last decade, however, advances in computing power have enabled the live playback and manipulation of digital audio. This development has compelling applications for musical performance, particularly with regard to techniques explored within the computer music community. To date, my research has consisted of exploring the features and commonality of these two diverse musical fields. This examination has led to the framework I will be discussing, which will be used to structure the creative

Universes of belief: a context-focused framework for performance design

component of the author's doctoral thesis, due for completion in late 2009.

As this undertaking involves both an analytical and a creative component, two roles are required; that of the researcher, and that of the artist or performer. Incorporating these two roles involves formulating a methodology to relate the technical to the creative aspects of this project.

Theoretical foundations

Much computer music research corresponds to classical Western musicology, where musical text and intra-musical elements are prioritised over context and extra-musical details. Some critiques have been made of this strategy (Landy [1999], Myatt [2008]), questioning the lack of attention to social, cultural, and musical context in current research. This type of research approach is particularly unsuitable for analyses of performance for two main reasons. Firstly, musical performance is representative of a transaction between an individual (a performer) and a given cultural field, and so requires a consideration of context. Secondly, although the areas of contemporary electronic music and computer music possess a very broad technological similarity, their aesthetics are different. No shared cultural framework exists, and both areas must be re-appraised in order to determine how they may be related. These two assertions will now be explained in more detail.

The theoretical link between creative works and contextual factors has been explored by several researchers. Csikszentmihalyi, in his study of the process of creativity, concludes that creativity is invariably the result of an interaction between an individual, social institutions, and the cultural domain (1988: 325).

Regarding performance, Carlson identifies that all performance (theatrical, musical, or other) exists according to at least one of three principal concepts (1996: 4). The first is defined simply as the display of skills (*ibid*). The second is the display of recognised and culturally coded patterns of behaviour (*ibid*). The third is the demonstration of success in an activity in light of existing standards of achievement (*ibid*: 5).

Phrases such as 'display of skills', 'culturally coded patterns of behaviour', 'demonstration of success', and 'standards of achievement' underline that performance exists as a qualitative function of a given social group or culture. Its methods, and the judgement of its success, are contin-

gent upon culturally determined principles, such as skill and established behaviours.

As stated earlier, this research, as a practical undertaking, involves the roles of researcher and artist. As a researcher, defining what constitutes notions such as 'skill' necessarily involves examining the practices of a given musical performance culture. The sociological work of Bourdieu is of use here, particularly the notions of cultural field and habitus.

Bourdieu's metaphor of the cultural field describes the institutions and conventions that enable the production of works of cultural value. The cultural field also determines the discourses and activities that are acceptable within that field. Bourdieu further defines the cultural field as "a separate social universe having its own laws of functioning independent of those of politics and the economy" (Bourdieu, 1993: 162). Knowing which laws, discourses and activities are permitted within a performance culture is 'natural' to those who are ensconced within an established group. Attempting to formulate new approaches to performance within a given cultural field, however, compels an active attention to these rules, conventions, and categories. The inclusion of new technologies and techniques directly challenges existing performance practices. Existing practices of the cultural field in which performance is to occur must be assessed, in order to determine whether this new type of performance is congruent with pre-existing codes. Attending to these systems is critical for creating value and legitimacy, as Bourdieu states that the artistic cultural field produces "not only the object in its materiality, but also the value of the object, that is, the recognition of artistic legitimacy" (ibid: 164). Specific aspects of the field of musical performance that may be considered include type of musical output, conventions of venue, and the role of the performer.

An idea linked to the cultural field is Bourdieu's notion of habitus. Habitus can be understood as the residual effects of experience on the individual, which shape how the individual develops subjectivity. The habitus of the subject affects the way that the subject engages in practices. This concept relates to the creative component of this research. Individuals seeking to participate in a cultural field inevitably emerge from a range of different backgrounds, which affect the way in which they connect with the field. As Bourdieu states, "All agents, writers, artists or intellectuals construct their own creative project according, first of all, to their perception of the available possibilities afforded by the categories of perception and appreciation inscribed in their habitus" (1993: 184). By actively examining the habitus of an agent in relation to the cultural field, the subjectivity, specificity and non-universality of the nature of engagement is addressed. In this research, this involves ascertaining the subject's musical preferences and performance capabilities. The intended result is a

performance environment that addresses the habitus of the artist.

A final notion proposed by Bourdieu that relates to cultural field and habitus is agency. Bourdieu writes, "To understand the practices of writers and artists, and not least their products, entails understanding that they are the result of the meeting of two histories: the history of the positions they occupy and the history of their dispositions" (1993: 61). Agency is thus used to describe the potential for interaction between an individual and a given context, or, to use the terms above, an individual's negotiation between habitus and cultural field. Performance provides a moment for agency to occur and to be evaluated. A successful performance can be understood as an effective demonstration of agency between habitus and the constraints of the cultural field.

The requirement of negotiating the cultural field in the production of artistic value is directly elucidated by Bourdieu:

The artistic field is a *universe of belief*. Cultural production distinguishes itself from the production of the most common objects in that it must produce not only the object in its materiality, but also the value of this object, that is, the recognition of artistic legitimacy. This is inseparable from the production of the artist or the writer as artist or writer, in other words, as a creator of value. (Bourdieu, 1993: 164, italics in original)

Understanding performance within this model accentuates its specificity to a given cultural context. In designing new approaches to performance, an attention to the cultural field is required in order to comprehend existing structures and codes. In actually engaging in performance, confronting the habitus of the performer draws attention to subjectivity. Linking these concepts is agency, representing the interaction between cultural field and habitus, and illustrating performance itself.

Framework

This research asserts that the laptop, as a novel and unfamiliar object, challenges conventional conceptions of performance, which are based in the acoustic tradition. These conceptions of performance are established within the cultural field, and affect the subjectivity of both audience and performer. Effective performance entails an efficient agency between cultural field and habitus. New approaches therefore need to consider cultural field and habitus on a symbolic and semantic level, in order to develop a relationship between them.

The designed framework, which is presented in the following section, aims to achieve this by considering contextual factors in the design of

technologised music performance environments. Although it was designed in regard to the specific instance of performing electronic music on a laptop, it is hoped that the framework is broad enough to be used for a range of applications. The framework identifies points to consider in negotiating approaches to performance that seek to engage with pre-defined contexts and constraints.

Each stage in the framework comprises three aspects. The concept defines a desired characteristic that is derived from traditional or acoustic performance archetypes. The object represents the issue to be resolved in that particular stage. The action involves the practical steps that are necessary to attain the object, and therefore the concept.

1: Instrument, Structure, Identification

Stage 1	
Concept: Instrument	An instrument-focused function for the laptop performance environment
Object: Structure	Definition of fundamental actions, interactions, and output
Action: Identification	Identifying and prioritising musical aspects to be incorporated to performance

Figure 1. Framework for context-focused laptop performance environment design: Stage 1

The first step in the framework aims to situate the laptop as an instrumentally-derived means for music production, rather than a functional tool for controlling output. Brown (2000) identifies five modes of compositional engagement between the computer and performer. In this arrangement, the computer may fill the role of artifact, tool, instrument, model, or creator (*ibid*: np). As an instrument, the computer “becomes more than an efficient tool, rather it is considered a partner in the compositional process,” and furthermore, permits improvisation (*ibid*). Central to this venture is the definition of interactions. These interactions proceed from the desired musical output. For instance, certain musics may require precise tonal or timbral modulation, whereas others may involve altering complex rhythmic structures. These attributes may be derived from the cultural field, as conventions of musical genre, or from habitus, as the composer or performer’s own preference. Based on these characteristics, musical behaviours need to be identified so that interactions can be determined accordingly. As an example, the formative research of Meyer (1956) derives emotive structures from musical works. This type of analytical

approach may be used to determine key musical functions. Once such behaviours have been established, they may be categorised according to their importance to form a usable map of interactions. Expressive modelling of this type has been explored by Camurri et al. (2005), Madison (2000), and De Poli (2004).

2: Expression, Behaviour, Representation

Stage 2
Concept: Expression
An instrument capable of performability and expression
Object: Behaviour
Specification of musical behaviours to be carried out in performance
Action: Representation
Representing desirable musical performance actions as objectives executable by the computer

Figure 2. Framework for context-focused laptop performance environment design: Stage 2

The second step aims to implement these behaviours within a computer-based environment. This involves translating identified behaviours to a format intelligible to the computer. In other words, these variables need to be rendered as semantically precise structures. As an example, generative processes may be employed in order to allow the performer to assume higher-level control over musical processes. Tailoring these processes and their contingent variables is necessarily a detailed and technical undertaking. The object of this stage is to ensure that the computer-mediated musical output is representative of the performer/composer’s own preference, as well as the boundaries of the musical style in which the creative work is situated. Significant consideration is required in order to ensure that the generated output is musically satisfactory. Collins (2002), Birchfield (2003), and Wessel and Wright (2002) have investigated the formulation of generative behaviours for the creation of coherent musical structures. The overall purpose of this undertaking is to define the artistic habitus of the performer, thereby constructing a potential for agency.

3: Virtuosity, Interface, Execution

Stage 3	
Concept: Virtuosity	A potential for learnable and skilled interactions
Object: Interface	Mapping of appropriate physical input to desired musical output
Action: Execution	Effectively implementing identified structural and behavioural characteristics in the performance environment

Figure 3. Framework for context-focused laptop performance environment design: Stage 3

The third stage attends to the transaction between performer and computer through the hardware or software interface. The central concern in this step is the effective mapping of performer input to sonic output, in order that processes can be practically controlled in a performance context. This may involve a simple one-to-one mapping of gesture to acoustic result, or more complex relationships. Incorporating complex mapping harnesses the processing capabilities of the computer, and simultaneously addresses interactive limitations in the production of complex output; see, for example, Hunt et al. (2003), Hunt and Wanderley (2002) and Goudeseune (2002). A vital concern at this point is the cognitive and physical limitations of the performer with regard to the number and complexity of control options, discussed extensively by Pressing (1990, 1998) and Jordà (2004). An associated issue is the design or choice of control interface, determined subsequently to the establishment of interactions, mapping strategies, and scope for control. A final point that may also be addressed at this stage is the intelligibility of interactions to an audience, historically a contentious topic in acousmatic and electroacoustic performance. This issue is explored by Cascone (2002), with particular reference to the use of the laptop in performance, while D'Escrivan (2006) and Croft (2007) discuss issues such as the importance of perceived effort and intelligible interactions. In a performance context, the audience can be understood as an aggregate of individuals constituting a manifestation of the cultural field. The target of this stage is thus to ensure that performance agency can be fulfilled.

4: Completeness, Evaluation, Usability

Stage 4	
Concept: Completeness	An environment which is conceptually sound and sustainable across a range of performance instances
Object: Evaluation	Critical evaluation of the success of the system constructed, with reference to contextual factors
Action: Usability	Refining the practicality of the system, including comprehensibility for both performer and audience

Figure 4. Framework for context-focused laptop performance environment design: Stage 4

This stage aims to critically appraise the usefulness of the designed system and identify avenues for improvement. The principal notion here is conceptual completeness, a term coined by performer and researcher Bob Ostertag (2002: 13). This phrase is used to define an instrument that is sustainable across a range of performance instances and resistant to obsolescence. The central determinant of success of a developed performance environment is its flexibility and responsiveness to the musical needs of the performer, allowing for agency to be carried out. Wanderley and Orio (2002) propose a method for quantifying and evaluating the expressivity of musical interfaces, and identify four criteria for success. These include learnability, explorability, feature controllability, and timing controllability (*ibid*: 71), features which must be present in any interface capable of satisfying basic musical and performative tasks. Furthermore, considering the potential of the environment for improvisation may be used in assessing its conceptual completeness. Allowing for expansion and modular additions to the system, in terms of embedding a potential for new interactions, is likewise an important aspect of the environment's flexibility. Finally, it is important to consider subjective as well as technical judgments of success and failure, including aesthetically and culturally informed factors such as quality of musical output and visual and communicative aspects of performance.

Application of the framework

This framework fosters the implementation of instrumental skill, virtuosity, and expressivity to computer-based music performance, based on an attention to the cultural field in which creativity occurs. This will ideally promote the construction of environments capable of less restricted, nonlinear and improvisational music creation according to performer-specified musical goals. The framework is being practically applied in the

development of the creative component of the author's thesis, as mentioned at the outset of this paper. Currently, Stage 1 (Instrument, Structure, Identification) is the main focus, and musical behaviours and interactions are being prioritised. Aspects of other stages have also been considered, such as the construction and control hierarchies of generative procedures which will be used to produce output. Interfaces have been considered, with the commercially available Novation ReMOTE ZeRO SL MIDI controller (Novation DMS Ltd, 2008: np) and Wacom Cintiq graphics tablet (Wacom, 2008: np) presenting the most promising possibilities at this point. A functioning model is scheduled for completion by end 2008.

Conclusion

This paper has proposed that performance is contingent upon an understanding of contextual factors derived from the cultural field. The development of a performance environment incorporating relatively new technology, such as the laptop computer, therefore needs to consider the existing cultural field in order to create a coherent means for performance. Achieving effective performance, understandable as an instance of agency between cultural field and habitus, is both a technical and a musical process. This requires an analysis of how desired features of performance and appropriate musical output may be practically implemented. To this end, the framework presents these concerns in a cogent structure that can be used for the practical construction of a performance environment.

References

- Birchfield, D. 2003. "Generative Model for the Creation of Musical Emotion, Meaning, and Form." *Proceedings of the 2003 ACM SIGMM Workshop on Experiential Telepresence*, Berkeley, California.
- Bourdieu, P. 1993. *The Field of Cultural Production: Essays on Art and Literature*. Polity Press, Cambridge.
- Brown, A. 2000. Brown, A. 2000. "Modes of Compositional Engagement." *Proceedings of the Australasian Computer Music Conference*, Queensland University of Technology.
- Camurri, A., De Poli, G., Friberg, A., Leman, M., Volpe, G. 2005. *The MEGA Project: Analysis and Synthesis of Multisensory Expressive Gesture in Performing Art Application*. Journal of New Music Research, 34:1, pp. 5-21.
- Carlson, M. 1996. *Performance: a critical introduction*. Routledge, London.
- Cascone, K. 2002. *Laptop Music - counterfeiting aura in the age of infinite reproduction*. <http://interdisciplina.dmethods.com/microsound-wiki/index.php/Texts> - Accessed 20/04/07.
- Collins, N. 2002. "Interactive Evolution of Breakbeat Cut Sequences." *Proceedings of Cyber-sonica, Institute of Contemporary Arts*, London, UK.
- Croft, J. 2007. *Theses on liveness*. Organised Sound, 12:1, pp. 59-66.
- Csikszentmihalyi, M. 1988. *Society, Culture and Person: A Systems View of Creativity*. In Sternberg, R. J. (ed), 1988. *The Nature of Creativity: Contemporary Psychological Perspectives*. Cambridge University Press, New York.
- d'Escriván, J. 2006. *To Sing the Body Electric: Instruments and Effort in the Performance of Electronic Music*. Contemporary Music Review, 25:1, pp. 183-191.
- De Poli, G. 2004. *Methodologies for Expressiveness Modelling of and for Music Performance*. Journal of New Music Research, 33:3, pp. 189-202.
- Goudeseune, C. 2002. *Interpolated mappings for musical instruments*. Organised Sound, 7:2, pp. 85-96.
- Hunt, A., and Wanderley, M. M. 2002. *Mapping performer parameters to synthesis engines*. Organised Sound, 7:2, pp. 97-108.
- Hunt, A., Wanderley, M. M., Paradis, M. *The Importance of Parameter Mapping in Electronic Instrument Design*. Journal of New Music Research, 32:4, 00. 429-440.
- Jordà, S. 2004. *Instruments and Players: Some Thoughts on Digital Lutherie*. Journal of New Music Research, 33:3, pp. 321-341.
- Landy, L. 1999. *Reviewing the musicology of electroacoustic music: a plea for greater triangulation*. Organised Sound, 4:1, pp. 61-70.
- Madison, G. 2000. *Properties of Expressive Variability Patterns in Music Performances*. Journal of New Music Research, 29:4, pp. 335-356.
- Meyer, L. B. 1956. *Emotion and Meaning in Music*. University of Chicago Press, Chicago 60637
- Myatt, T. 2008. *New aesthetics and practice in experimental electronic music*. Organised Sound, 13:1, pp. 1-3.
- Novation DMS Ltd, 2008. *Novation Music | Global Products | MIDI Control | ReMOTE ZeRO SL*. http://www.novationmusic.com/products/midi_control/remote_zero_sl/ - Accessed 21/03/08.
- Ostertag, B. 2002. *Human Bodies, Computer Music*. Leonardo Music Journal, Vol. 12, pp. 11-14
- Pressing, J. 1990. *Cybernetic Issues in Interactive Performance Systems*. Computer Music Journal, 14:1, pp. 12-25.
- Pressing, J. 1998. *Psychological Constraints on Improvisational Expertise and Communication*. In Nettl, B. and Russell, M. (eds), 1998. *In the Course of Performance: Studies in the World of Musical Improvisation*. University of Chicago Press, Chicago.
- Wacom, 2008. *Cintiq Product Information*. <http://www.wacom.com/cintiq/> - Accessed 21/03/08.
- Wanderley, M. M. and Orio, N. 2002. *Evaluation of Input Devices for Musical Expression: Bor-*

- rowing Tools from HCI.* Computer Music Journal, 26:3, pp. 62-76.
- Wessel, D. and Wright, M. 2002. *Problems and Prospects for Intimate Musical Control of Computers.* Computer Music Journal, 26:3, pp. 11-22.

Julian Knowles

Queensland University of Technology
Musk Ave,
Kelvin Grove, QLD, 4059
Australia
julian.knowles@qut.edu.au

Abstract

This paper presents a survey of some key experimental electronic music scenes in Australia in the 1970s and 1980s that were situated outside the frame of music institutions and the electro-acoustic music tradition. Specifically it focuses on experimentalism in the popular music field in this period, outlining the 'little bands' scene in Melbourne and the scene around the Terse Tapes and M Squared labels in Sydney. It locates a set of international reference points in No Wave, industrial music and krautrock and discusses the work of some individuals who traversed the popular music and experimental music scenes, acting as a nexus for musical ideas. The paper proposes that experimental practices in popular music exert a significant force on experimental electronic music practices in a wider sense and the links to more academic forms of music experimentation need to be better understood. In doing so the paper calls for a more expanded and complete view of the historical development of electronic music as a way of better understanding current experimental electronic music scenes in Australia.

Introduction

When one considers the contemporary electronic music scenes that have coalesced around artist-run spaces and warehouses in Australia since the late 1990s, one is struck by the differences between these scenes and the electro-acoustic music scene associated with conservatoires and tertiary music schools. At the same time, one is struck by the similarities between these contemporary scenes and the 'DIY' post-punk scenes of the late 1970s and early 1980s in Australia. Despite these strong resonances, there has been little attempt to connect these scenes historically, or to see contemporary electronic music scenes as having roots in these scenes, or indeed in any tradition other than the electro-acoustic or western art music traditions¹. I would suggest that an expanded view of music history, which accounts for sonic experimentalism in popular music scenes alongside developments in the electro-acoustic music scene, provides the necessary context to understand contemporary electronic music scenes. This paper constitutes a small step towards drawing some of these links.

Liminal Electronic Musics: Post-Punk Experimentation in Australia in the 1970s-1980s

The rise and subsequent canonisation of serialism, *musique concrète* and *elektronische musik* through the 1950s and 60s fuelled the proliferation of electronic music studios in music institutions, radio stations and research institutes throughout Europe and the USA. Such facilities became critical sites for musical experimentation. Due to the somewhat conservative institutional climate in Australia during this period, it took some time for comparable studios to be established in institutions, a process which commenced in the late 1960s. Whilst there is evidence of some small private electronic music facilities outside music institutions, almost all were contained within institutions. Such a concentration was largely due to the fact that, prior to the mid 1980s, music and studio technologies were extremely expensive and mostly out of the financial reach of independent artists.

From the late 1970s the hegemony of institutions was disrupted, as tools for electronic music production became affordable for independent musicians outside institutions. Accordingly the sites for musical experimentation became more diffuse and much exploratory electronic music was made outside an institutional context. In a popular music context, experimental approaches to electronics and production were thriving in Australia in the late 1970s, as they were in other parts of the world. This diffusion of experimentation yielded significant outcomes, rarely accounted for in mainstream accounts of the development of electronic music in Australia.

It would be fair to say that the continuing diffusion of powerful tools for electronic music de-centred the long-held view of innovation in electronic music practice to the point that the electro-acoustic music scene became viewed as somewhat 'traditional' in comparison to the innovative practice happening in experimental electronic music scenes from the mid 1990s onwards. These tensions and debates are well documented and came to a head when electronic dance music artist Richard James (aka Aphex Twin) and video artist Chris Cunningham were awarded the Golden Nica in the Digital Musics category at Ars Electronica in 1999. The trend toward critiquing the dominant paradigm of electro-acoustic music continued. The statement from the 2001 Ars Electronica Digital Musics Jury was direct.

It's a fact: on the cusp of the twenty-first century, the most innovative,

¹ The former is actually a sub-branch of the latter.

com-pelling and startling work being produced in the impossibly broad area of Digital Musics comes from musicians whose backgrounds have largely by-passed academic study and customary career paths. Instead their work speaks of an intense, autodidactic engagement with the hyperlinked worlds of post-industrial cultures: conceptual and performance art, installation and video work, improvised music, post-industrial cultures, eco-activism, post-colonialism, as well as the post-techno/hip hop/dub grass-roots diaspora of blunted beatnuts and bedroom boffins. (Digital Musics Jury 2001)

Such a statement proposes that electronic music practitioners find relevant histories outside the western art music tradition. One might ask how the above might be understood in an Australian context, specifically what local reference points might exist outside the western art music tradition to support such a notion?

It has been suggested on many occasions that musicological engagements with new music in Australia have been scant. I might add to this that most musicological engagements with music scenes in Australia tacitly and uncritically accept a split between so-called 'art music' and 'popular music'. The same shortcoming can be observed in the majority of writing around both popular music and western art music in Australia. I would suggest that uncritically accepting this split prevents us from fully understanding contemporary electronic music practices. Rather than suggesting that we should ignore musical differences arising from different musical contexts and traditions, I would argue that different contexts and traditions have significant fields of intersection and these are significant drivers of musical innovation. To fully understand developments in musical practice we must acknowledge this. Whilst the dialogue across different musical scenes is ever present, I wish to focus on flow between the experimental electronic music and post-punk music scenes in the 1970s and 1980s, as this marked a particularly active time, driven by the diffusion of technology into the wider musical community for the first time. The visual arts scene can also be seen to have exerted a significant influence on both music scenes, as art schools were an important site for experiment in a broad range of time based media, including sound.

Interactions with Post-Punk Music

Whilst it is the musicological norm to draw sharp distinctions between popular music practices and art or experimental music practices, the experimental electronic music scene in the late 1970s and early 1980s had a demonstrably robust set of engagements with the punk and post-punk

music scenes in Australia. This is evidenced by the significant quantity of experimental music distributed through post-punk's extensive DIY infra-structure network consisting of vinyl and cassette labels, fanzines, venues, specialist publications and radio programs. Rather than simply serving as distribution networks, the post-punk scene was an important site for musical experimentation in its own right and referenced a different set of histories from the contemporary classical tradition.

Beyond post-punk music as a site for experimentation in itself, an interesting and fertile set of interactions took place between musicians working in the more adventurous end of the post-punk scene and musicians in other areas who were interested in exploratory music. The principal sites for these activities and interactions were Sydney and Melbourne, although there was significant interaction between the experimental art and post-punk music scenes in Brisbane at this time. Art schools, galleries and visual artists² with an interest in music were part of the network of interactions. The post-punk and art school scenes were critical to the development of experimental music in Australia and again, posed challenges to the classical music establishment. These scenes could be interpreted as antithetical to the conservatoire scene due to their close association with popular culture and 'non-musicians' respectively. Ironically, their exclusion from the classical music mainstream made them powerful sites for unbridled musical experimentation and much important work arose from this engagement.

The Clifton Community Music Centre as a Liminal Space

The Clifton Hill Community Music Centre (CHCMC) is one of the better-documented scenes in Australian experimental music history. Whilst it is understood to have occupied an important place in the history of new music in Australia, it operated in a liminal space between the institutional (read 'academic') new music scene and non-institutional scenes. When David Chesworth took over the directorship of the CHCMC in 1978, he was young, still a third-year student at La Trobe University Music Department and had broad musical interests which included popular music. Importantly he was a friend and collaborator of Philip Brophy and the involvement of both in the CHCMC and in the Melbourne experimental music scene would be pivotal in terms of generating a dialogue across popular music and experimental genres. It became a liminal space in terms of genre. Embracing the post-punk DIY aesthetic, Chesworth and Brophy established a label, Innocent Records,

² See John Nixon's *Anti-Music* project 1980-83.

and a magazine New Music, with the latter fulfilling the similar function of a post-punk fanzine by serving as a tool to connect a community.

Both individuals had avant-pop projects with hybrid electronic/amplified resources that they started in this period. Brophy had formed the band tsk tsk tsk (the group's name was technically three arrows – pointing to the right, up, and to the right respectively) in 1977. Chesworth formed the band Essendon Airport in 1978 with guitarist Robert Goodge and a drum machine. Both Essendon Airport and the various incarnations of tsk tsk tsk were regular performers at the CHCMC during Chesworth's five-year tenure as director. Due to the cross-overs created by Chesworth and Brophy, CHCMC audiences started to build in the following years, augmented by the audience for post-punk music, which was thriving in Melbourne at the time. At the time, Chesworth and Brophy's bands were performing in post-punk music such as the Crystal Ballroom, considered the home of the scene at the time. Chesworth (1980), cited in Fox (2002) says,

The musical ideology of punk/new wave is in many ways similar to that of early, new and experimental music. Both basically involve a rejection of accepted musical values and formats in favour of re-asserting and re-defining the fundamental processes involved in music making the "any-one can do it" attitude figures prominently in both areas.

Chesworth's music at this time reflected his eclectic influences. Listening to his solo album *Layer on Layer* (Innocent Records, 1981) one cannot help but hear his influences from the American repetition-based minimalism of Steve Reich (probably encountered through his academic studies at La Trobe), fused with the funk-crossover of key post-punk albums such as Talking Heads *Remain in Light* (Sire Records 1980), and the processed guitar, vocal and synth gestures of UK post-punk bands from this period. Both Chesworth and Brophy were not afraid to create exploratory music with groove, two concepts which were not usually connected in the Australian scene at the time.

This blending of the experimental music scene with the post-punk scene in Melbourne was important in understanding how other genres may have influenced the experimental music tradition. Certainly the enthusiasm for the work of these two individuals was not shared by all, and there are some accounts that the CHCMC scene around Chesworth and Brophy was too 'intellectual' in the context of popular music. The ties to the university were regarded with some suspicion. As Clinton Walker (1996), a well-known popular music writer and authority on post-punk, noted,

The scene centred around the Clifton Hill Community Music Centre was... cerebral, abstract. Although the CHCMC also embraced older guard musicians like David Tolley, its stars were the likes of tch-tch-tch [tsk tsk tsk] and David Chesworth, who had taken to the mini-Korg like it was the key to a whole new kingdom.

Walker's comments, coming from someone with no connections to the experimental or academic music scenes, are illuminating, in that they give an insight into how the CHCMC might have been seen from the outside at the time. They indicate that, from a post-punk perspective, Chesworth and Brophy were the most notable figures associated with the centre (the 'stars'). This provides a different perspective to that represented in the much of the musicological and artist-centred writings about the CHCMC.

Tsk Tsk Tsk and Philip Brophy

Whilst not a purely electronic band, but one of mixed instrumentation - drums, synthesizer, guitar and saxophone - tsk tsk tsk proved to be influential in serving as a nexus for ideas to flow between the experimental music scene and the popular music scene in the late 1970s/early 1980s.

Tsk tsk tsk was at the same time a band and a series of eclectic art projects. It was strongly informed and influenced by Brophy's reaction against contemporary classical music pedagogy, which he encountered briefly as a student at La Trobe University music department, and his emerging body of theoretical critique around high and popular culture dynamics which followed. Brophy often stated that he saw no problems in being interested in Stockhausen and Gary Glitter and "that there seemed no obvious way of relating the two seemed to offer a cultural lesson" (Jenkins 1988). Tsk tsk tsk became a vehicle to explore this liminal space between art and pop.

The work of tsk tsk tsk is at the same time an analysis of the concept of genre, and a form of 'writing' (to use Brophy's term) to communicate ideas that emerge from this process of analysis. This level of abstraction, a music about music, a *meta-music*, seeks to create a space where ideas from popular culture and experimental art can freely mingle.

In the following passage from his 1987 article *Avant-Garde Rock: History In The Making?*, Brophy attempts to unpack the position of avant garde in popular music in relation to the histories of the avant garde in art music, in an attempt to define a space for avant garde music to exist outside the unreasonably tight constraints of western art music.

There are two possible ways of doing this. The first is to see the avant garde of rock as a pitiful bastardisation of

the original thrust of twentieth century avant garde art, a cooption of the polemic intensity that motivated the radical nature of its ideas and pursuits. The second is to acknowledge its nature as mutation, as an artistic activity born of visions that arise more from a developed social environment than from a studied historical lineage. The first approach is idealistic. The second is realistic...[A]vant garde purists in the academic realm of the contemporary, experimental and new music fields, rarely display the flexibility to accommodate the often sloppy approaches to rock experimentation, while the bulk of avant garde rock can at times be frustratingly uninteresting and uninspiring. (Brophy 1987)

Brophy's criticism that experimental and new music fields cannot accommodate sloppy and amateurish playing proposes that this sonic quality is worthy of musical investigation itself and can be considered an essential aspect of the whole work.

[I]n the reality of a pub gig, we generally appear to be incompetent, uno-
riginal, impersonal and anony-mous.
In short - the essential rock band.
(Brophy 1983)

Brophy embraced and aestheticised this aspect of a band. He called for an understanding of the musical material as a total sonic experience rather than a compartmentalised understanding which seeks to separate notions of musical content from its execution in performance. Interestingly, this could be seen to be advocating a form of total, or abstracted listening akin to the listening modes identified by Pierre Schaeffer and refined by subsequent artists and theorists.

Both Brophy and Chesworth would become well known in the Melbourne post-punk scene at the time, which of which two prominent strands were developing. The first was around Nick Cave and band The Boys Next Door (later called The Birthday Party), a strikingly original band with a dark rock sound and angular dissonant guitars from Rowland S. Howard. This represented the experimental end of the post-punk scene in respect of standard rock instrumentation. The other prominent scene was not so much around a band, but a collection of bands, many of whom used electronic resources. This was called the 'little bands' scene³.

Melbourne: The Little Bands Scene

The little bands scene, which flourished from 1978-1981, grew out of North Fitzroy, where two bands at the centre of the scene lived in adjacent houses. These were Whirlywirl (core members Ollie Olsen – vocals / synthesizer / guitar and John Murphy – drums, along with various other members, including Arne Hanna - guitar) and The Primitive Calculators (Stuart Grant – guitar / vocals, Denise Rosenberg - keyboards, David Light – bass / keyboards, and Frank Lovece - drum machine / vocals). These bands had been influenced by early electronic bands such as the Silver Apples, the UK industrial act Cabaret Voltaire and the New York No Wave bands including Suicide, DNA, Teenage Jesus and the Jerks and James Chance and the Contortions. Most of the bands used some form of electronic instrumentation and the music ranged from abrasive drum machine driven sounds of bands such as the Primitive Calculators, to angular electronic pop exemplified by Whirlywirl among others. Like many music scenes of this type, there were blurred boundaries between the bands and their audiences⁴ and people were encouraged to form bands, hence the proliferation of 'little bands' that formed the essence of the scene. The emphasis was on being part of the scene rather than aiming for commercial success. The scene was given its name when a local record store owner paid for and released a compilation EP entitled *Little Bands* in 1980 which featured tracks from several bands in the scene (Griffin 2006).

The comparison to the New York No Wave movement, emerging at around the same time, was a significant one. Centred on downtown venues such as CBGBs, the Mudd Club and Max's Kansas City, the No Wave scene lasted for only a short period but proved to be very influential. It was a reaction against the new wave music coming out of the UK and what was considered to be the conformity and commercialism of punk rock which was still seen as conventional rock and roll by many in this scene. The No Wave bands were angular, abrasive, atonal, used dissonant tunings, were sometimes primitive in their musicianship and drew on eclectic musical influences. The most cited musical reference was Alan Vega and Martin Rev's electronic duo Suicide, who abandoned guitars in favour of a synthesizer, drum machine and vocals. Other common influences were the Velvet Underground, Lou Reed's *Metal Machine Music* and the 'krautrock' bands Can and Faust.

Central to the No Wave movement was a conscious desire to push artistic boundaries and to approach music from a more experimental, anti-commercial perspective. Martin Rev, keyboard player for Suicide said "I think No Wave

³ A fictionalised account of this scene can be seen in the 1986 Richard Lowenstein film 'Dogs in Space'.

⁴ The same could be said of contemporary experimental music scenes.

was a valid avant-garde extension of rock, and it incorporated free and often atonal improvisation. I think... it was comparable to the avant-garde in any other form, such as jazz or European classical" (Nobakht 2004). This was perhaps due to the fact that No Wave was more than just a music scene. The musical activities served to connect individuals from a wide variety of disciplinary backgrounds including visual arts, film and music, many of who had relocated to New York to study or to further their careers. Influenced by the punk aesthetic, they formed bands, even if they had limited musical abilities in a conventional sense.

Significantly, both Glenn Branca and Rhys Chatham would go on to gain a level of acceptance in the contemporary classical music scene, perhaps due to the fact that they possessed traditional musical literacies and were versed in the contemporary classical tradition. In Australia, a comparison might be drawn to David Chesworth, or to a lesser degree Philip Brophy, who although accepted to some extent, rejected the contemporary classical tradition early in his career.

The Australian connections to the No Wave musical sensibilities cannot be underestimated and many musicians active in the more adventurous ends of popular music at this time cite similar reference points. The connections were more than just musical - they extended to the overall cultural milieu of the two scenes. John Murphy of Whirlywirl, in an interview, draws similar links in his description of the little bands, outlining its connections to the visual arts scene and how experimental practices were blended with a punk aesthetic.

The little bands thing was meant to be wild and chaotic and punk... art, experimental stuff, and not just electronic. A lot of the original participants were actually artists who applied the dada... approach of their painting. It was the attitude and idealism of punk, but applied to a post-punk art type thing. (Walker 1996)

Such a description is remarkably similar to those of the No Wave scene. Given the ages of the participants at the time, many of these visual artists would have been at art school, or recently graduated, which underlines the role that art schools played in providing a platform for young artists to experiment across a range of media, which often included music. Murphy, in the course of this interview, emphasises the differences between the little band scene and Philip Brophy's work at the time, which sought to avoid the 'emotion' associated with punk and post-punk. From Murphy's perspective it was more academic and detached.

In Australia, the reactionary drive of the little bands (and the similar scenes in other cities) was generated more in response to the mainstream

dominance of rock and roll in Australia at the time as to the failings of punk rock to deliver a radical alternative. The reaction was manifested not just in attitude and musical influences, but in sonic terms. The prominence of synthesizers and primitive drum machines opened up quite abstract sound worlds and these bands would prove to be influential to later bands in electronic scenes. The sonically adventurous attitude of the Primitive Calculators motivated them to push sonic boundaries, which made for some extreme sonic outcomes, "In a time when synth based 'new wave' was losing its edge the Primitive Calculators made clangour sounding like the destruction of a collapsing building. (Lunsford 2004). Alan Bam-ford, a musician (Alan Bamford Experience, Hugo Klang), sound technician and radio presenter on alternative station 3RRR, provided this account of mastering the Primitive Calculators LP in 1979.

When I mastered the Primitive Calculators LP...I didn't have a clue what I was doing and Stuart [Grant] said 'maximum bass and treble' so we did that but then we had to reduce the bass when we cut the acetate, as it cut right through to the plate. The Primitive Calculators always cut right through. (Bamford 2004)

This evidences the band's desire to explore extreme sonic terrain in their production processes.

Despite the strong punk element and obvious differences between this scene and the more academic ends of the new music scene, connections prevailed. Two of the main figures in the little bands scene, Ollie Olsen and Arne Hanna from Whirlywirl (who were flatmates in the 'Whirlywirl house' in Fitzroy), both studied composition and electronic music with Felix Werder (a German/Australian composer who had himself studied with Stockhausen). They would, therefore, both have been well aware of the European avant garde. This link to the contemporary classical music scene is not mentioned in the popular accounts of the little band scene, but underlines that the experimentalism of the little bands scene had roots that extended further than common accounts might suggest.

Sydney: Alternative Venues, Terse Tapes and M Squared

Sydney was also a site for crossover between the experimental music scene and the post-punk scene. Like in Melbourne, a sense of community was formed around venues, bands a social networks. Interactions across post-punk and experimental electronic music scenes occurred through shared venues, labels, record stores and social networks.

A venue which played an important role in Sydney in the late 1970s was the Paris Theatre, which had been vacated by the Hoyts cinema company in 1977 and became used as a venue for live music, cabaret and theatre performance frequented by punks, hippies and alternative types. It rose to particular prominence as a post-punk music venue and its bills for all the main acts of the day, including the Laughing Clowns, the Birthday Party and the Go-Betweens. In 1977 the venue has become host to an emerging punk scene with local and touring bands such as X, JAB, Rocks, the Survivors, News, and the Babeez all playing gigs there (Coomber 2000; Morris 2000). The experimental music scene also intersected with the space. Rik Rue, a prominent Sydney tape collagist and improviser was organising gigs there in the late 1970s (Jenkins 1988) and 'Towards a Relative Music', an early Fringe Benefit Records release from Jon Rose and the Relative Band, credits one of the tracks as being recorded live at the Paris Theatre on May 25, 1978 (Rosenstein and Roussel).

Art Galleries also hosted a number of gigs for both experimental music and post-punk groups. These included the Sculpture Centre, Art Unit and Central Street Gallery (Rosenstein and Roussel; Jenkins 1988; Blades 2003). Squat spaces were also important, one of the most significant being SideFX, an artist squat in a former Catholic school building in Darlinghurst, which hosted music gigs and experimental film events⁵. Consulting the gig lists of post-punk bands and experimental musicians from the late 1970s and early 1980s indicates a strong overlap in the venues used. These venues were associated with a particular cultural milieu and ideas were exchanged across different media. Ideas around contemporary art processes mixed with ideas from popular music and contemporary western art music.

Sydney was home to a strong culture of tape cut-up and collage, and many artists in both the post-punk and experimental scenes experimented with tape manipulations. A key early example in Sydney was Ian Hartley, a musician, music venue promoter and publisher of the *Spurt* punk fanzine. Hartley also owned a shop called *Skin Deep* in inner-city Sydney which sold post-punk and mod clothes and fanzines into the Sydney subculture scenes. A late 1979 issue of *Spurt* featured a cassette of Hartley's tape collage and he would go on to release more tapes of collage under the name Mice Against God (Blades and Andrews 2007). Other exponents of tape collage at this time included Rik Rue, visual artist and musician John Gillies, The Horse He's Sick (Ian Andrews), Sw Sw Thrght (John Jacobs) and

Severed Heads. Many of these tape collagists and manipulators were involved in radio at the time. They were members of the 'contemporary collective', a group of individuals who presented experimental or adventurous radio programs from midnight to dawn on 2MBS-FM, a community radio station. This scene eventually spawned The Loop Orchestra in 1983 (core members - John Blades and Richard Fielding), a well-known Sydney group that makes music exclusively from tape loops. The Loop Orchestra, through its shifting guest membership, would act as a hub connecting a range of individuals from the post-punk and experimental scenes in Sydney.

There were two main experimentally influenced music labels that emerged from the post-punk scene in Sydney. The first was Terse Tapes, established in 1979 by Tom Ellard of Severed Heads. Terse Tapes was named after the TRS-80 home computer (which Ellard was using to make music) and the logo consisted of a clipping from an advert for the computer with the label name written over it in white-out fluid (Ellard 2008). As the label name suggests, the output of Terse Tapes was mostly in cassette format, but some vinyl was produced.

The first vinyl release on Terse was the collaborative album *Earbitten 12"* in 1980, a shared release featuring Severed Heads material (featuring extensive use of tape loops and electronic synthesizer noise) on one side and tracks from the group Rhythmyx Chympx on the other. Severed Heads had already been receiving radio play, after sending Peter Doyle (2-JJJ announcer) some of their earlier cassette releases (Ellard 1990). Tom Ellard, the principal musical force behind both Severed Heads and Terse Tapes, had a reputation as a technology futurist, adopting and employing new technologies before their more generalised uptake into the market⁶. For the track *Dance* on the *Earbitten* EP, Ellard wrote a computer program in BASIC to generate interference tones on a nearby portable radio. The range of sounds made by the radio as the tuning was adjusted was then employed as part of the track (Blades and Andrews 2007).

Ellard's dry, humorous account of early Severed Heads gigs provides an insight into the climate in the more adventurous ends of the post-punk scene at the time.

[In 1980] we were offered a 'gig' at a venue called I.C.E. [Institute for Cont-emporary Events] run by Ian [Hartley] and Michael (who now runs DOME, Kinselas etc.). I think I

⁵ A number post-punk groups and experimental music practitioners were experimenting with experimental Super 8 film at the time, fuelled by its widespread use in art schools. Examples include Philip Brophy, David Chesworth, SPK, Debra Petrovich.

⁶ Ellard cites his early musical influences as David Bowie, Brian Eno, Eno, Roxy Music and Kraftwerk, and the industrial band Throbbing Gristle as his first encounter with experimental music. He cites seeing the posters for Philip Brophy's 'Punk Gunk' gigs on New Years Eve 1977 as his first encounter with the Melbourne 'underground' music scene Ellard, T. (2007). Email correspondence with Julian Knowles. October 9, 2007. J. Knowles. Brisbane.

bumped into Ian at a record shop or something and he thought we were the sort of people who frowned a lot and wore camouflage. Anyway 33 people turned up to a pretty messy tape loop extravaganza, including Stephen Jones [who would later join Severed Heads as its video artist] (Ellard 1990).

Various elements within Severed Heads consciously strove to innovate and react against what was seen as the boring Australian rock and roll mainstream at the time. Ellard describes Richard Fielding being particularly sensitive to this issue and in his dry style describes how Fielding left the band for a period to pursue more experimental work "Richard started to react against the way the band was becoming even the tiniest bit 'rock and roll' - gigs, records - and developed his own group The Nobodies - a collection of tape recorders for which he was the roadie" (Ellard 1990).

Like many bands working with electronic resources, experimentation was critical to their process and their approach to constructing tracks was adventurous. In addition to employing emerging technologies, they developed novel processes with existing technologies - modifying tape loops, physically manipulating vinyl through the addition of holes, cutting, warping and modifying the surface. Notably they developed some interesting techniques for controlling analogue sequencers. All of these insights they happily shared with their fans. This quote provides an insight in the creative process behind the track *Epilepsy 82* from the 1983 album *Since the Accident*.

This one is tricky so you'll have to watch closely. Now, we fold a tape loop so, and we record drum sounds thus. Here we place a pulse signal with a hi-hat sound, run a line from the desk to John Blades' Pro One, load a sequence here and behold we have a sequence clocked by a tape loop. So add two sections of Gamelan orchestra, a lecture regarding membrane depolarization, baying wolf and yapping cat, mad gorilla and serve. When we remixed this we added a wolf baying every now and then stuck in a digital delay. (Ellard 1990)

Not content to confine their experimentation to the music, they also experimented with visual aspects of their live shows via "slide shows, film loops, bubble machines, blow up backing singers and computer graphics". In 1982 video artist Stephen Jones organised a collaboration with Severed Heads for which he employed a video synthesiser. This collaboration resulted live shows and the album *Blubberknife*, released as a C90 cassette, shrink wrapped with electronic

television components. Jones would subsequently join the band as a resident video artist (Ellard 1990).

Arguably one of the most important contributions from Terse in this period was the release *One Stop Shopping*, a three-cassette survey of Australian underground music⁷, released in March 1981. This was a national survey (not just of Terse artists). It included contributions from the Melbourne scenes - Clifton Hill artists David Chesworth (solo and with Essendon Airport), Ernie Althoff, Graeme Davis and Philip Brophy (tsk tsk tsk/tch tch tch); and the little bands - Alan Bamford Experience, Use No Hooks, Swinging Hogs and the J.P. Sartre Band. Alongside this there were a range of tracks from Sydney bands from Terse and from the other important Sydney label, M Squared including the Sydney Quads and Systematics. Such a release evidences an awareness of a national scene around exploratory 'under-ground' music, both from the post=punk and experimental music spheres, and a desire to present it as a coherent and sustained movement. Interestingly, according to Ellard, this network came into being through live performances. In his experience, most of the bands which underpinned the Sydney scene were independently formed by individuals who had been accessing music and publications entering Australia from overseas and that they only met each other when their bands started playing live (Ellard 2007).

M Squared, the other important Sydney label from this time, was established in 1979 and operated as a collective of electronic bands who shared a recording studio facility in Wilshire St. in Surry Hills. Key personnel in M Squared were Mitch Jones (who was a live sound engineer for both Ed Kuepper and The Numbers), Patrick Gibson, and Michael Tee. The most well known acts associated with M Squared were Scattered Order, the Systematics and the Makers of the Dead Travel Fast. All these bands employed extensive electronic instrumentation including analogue synthesizers and drum machines alongside the more traditional instruments of electric guitar and voice. None had drummers. The music had a strong machine-like quality with primitive and somewhat bare electronic rhythm programming which was highly distinctive to both the label and the period. Many consider M Squared to be similar to the scene around the little bands in Melbourne and while the Primitive Calculators and Whirlywirl were seen to drive the Melbourne scene, Scattered Order, the Systematics and the Makers of the Dead Travel Fast were considered the principal bands of the Sydney scene.

This description from Michael Tee of M Squared paints not only a picture of their musical

⁷ 'Underground music' was the term used by Terse Tapes.

reference points, but their curious (and by all accounts real) reactions to their drum machines

In 1980...PiL had just done *Metal Box*; Ian Curtis was still warmish; The Residents hadn't stretched their muse; the ghosts of Can + Neu + Faust still lived on; synths were still monophonic (not poly) and analogue; Eno's voice and lyrics were still potent; Bowie hadn't done *Let's Dance*; Cabaret Voltaire... and Throbbing Gristle... made noise... people would shout out "where is your drummer!"; the Australian Musician's Union would hassle us about our drum machines putting drummers out of work. (Tee 2003)

The reference points, as they were for so many of the 'underground bands' in Australia at the time, were the more adventurous ends of English punk and post-punk (The Pop Group, PiL, Wire, Joy Division), early industrial bands (Throbbing Gristle, Cabaret Voltaire), glam and art rock (Bowie, Eno) and the German 'krautrock' bands from the 1970s (Can, Neu, Faust). This list was often extended to include the prog rock/art rock bands Henry Cow (featuring Fred Frith), Slapp Happy, This Heat and Pere Ubu along with New York No Wave.

Such musical reference points stood in stark opposition to the mainstream Australian culture of pub rock and many ironic references were made to distinguish the scene from the mainstream. M Squared dubbed themselves the "builder's labourers of the avant garde" as a form of ironic self-deprecation (Tee 2003). The label's releases promoted strong and diverse reactions from writers in the independent music press as the following quotes indicate (Tee 2003)

Some of the music has bordered on the un-listenable, being extreme, erratic, naive and experimental. But at its best has been amongst the most inspired and original music recorded in recent years (Stuart Coupe)

M Squared specialises in non rock n' roll. Describing these records is impossible, they cannot be related to other music, because they aren't music. (Peter Botrell)

When one considers the output of M Squared and the bands with whom the label was associated, it is clear that they were considered to be pushing musical boundaries and re-defining to a certain degree what might be considered popular music. A spirit of experimentalism and a reaction to mainstream rock was mixed with an electronic pop aesthetic. This engagement was fuelled by musical reference points in innovative popular music and electronic music scenes from other cities and from the more adventurous ends

of 70s electronic music and conceptual rock, as opposed to the western art music tradition.

Conclusions

The cross-overs between post-punk and experimental electronic music practice in this period mark the beginning of a sustained period of cross-fertilisation through the 1990s to the time of writing. One could trace similar connections and intersections between experimental electronic music and the various offshoots of industrial music, hip-hop, turntablism, minimal techno, downtempo, ambient music, post-rock and new electronica hybrids. This is a complex set of interactions that has scarcely been documented. By way of contrast, the electro-acoustic music tradition has remained comparatively 'pure', maintaining close links to the conservatoire and the contemporary classical music tradition. Fairly minimal impact is seen when one examines the programs of international events such as the International Computer Music Conference. As such, the experimental electronic music scene continues to drift further away from the electro-acoustic music mainstream, although in recent years some experimental music festivals included work from electro-acoustic music composers who are seen to be 'influential' to practitioners in this scene. Evidence would suggest that electro-acoustic music festivals and events have not been so porous to other electronic music practices.

References

- Bamford, A. (2004). "Primitive Calculators: Quotes." Retrieved November 8, 2007, from <http://www.project.net.au/primitivecalculators/quotes.html>.
- Blades, J. (2003). "My Personal Journey though Post-punk, Art, Music and Radio." No Night Sweats: Sydney Post-Punk Memoirs Retrieved July 3, 2007, from http://www.users.bigpond.com/pturnbul/nns_memoirs_jb01.htm.
- Blades, J. and I. Andrews (2007). Post-Punk Electronic Music (unpublished manuscript for forthcoming book chapter). G. Priest. Sydney, UNSW Press.
- Brophy, P. (1983). Made by Tsk tsk tsk. Melbourne, Tsk Tsk Tsk.
- Brophy, P. (1987). Avant-Garde Rock: History In The Making? Missing in Action - Australian Popular Music in Perspective. Melbourne, Graphics Publications.
- Chesworth, D. (1980). "The Clifton Hill Community Music Centre." New Music (1978-1979): 3-4.
- Coomber, M. (2000). "News: The Original 1970's Punk Band." Retrieved January 15, 2008, from http://www.geocities.com/babeeez_news/
- Digital Musics Jury (2001). Digital Musics Diaspora: Digital Musics Jury Statement. Cyber-

- arts 2001: International Compendium Prix Ars Electronica. H. Leopoldseder and C. Schöpf. New York, Springer Verlag; 168-169.
- Ellard, T. (1990). "History: A short guide to Severed Heads 1979 - 1990." Sevcom Booklet: The Journal of Severed Communications 5.
- Ellard, T. (2007). Email correspondence with Julian Knowles. October 9, 2007. J. Knowles. Brisbane.
- Ellard, T. (2008). "Terse Records." Severed Heads Archive Retrieved February 24, 2008, from <http://www.severed-heads.co.uk/sevcomdocuments/Terse/>.
- Fox, R. (2002). MA Thesis. Experimental Music in Melbourne: 1975-1979. School of Music. Melbourne, Monash University.
- Griffin, R. (2006). "Discography." Australian Post-Punk 1976-1981 Retrieved October 12, 2007, from <http://members.ol.com.au/rgriffin/postpunk/Various.html>.
- Jenkins, J. (1988). 22 Contemporary Australian Composers. Melbourne, NMA Publications.
- Lunsford, K. (2004). "Primitive Calculators: Quotes." Retrieved November 8, 2007, from <http://www.project.net.au/primitivecalculators/quotes.html>.
- Morris, G. (2000). "Rocks." Retrieved January 15, 2008, from <http://www.breakmyface.com/bands/rocks.html>.
- Nobakht, D. (2004). Suicide: No Compromise. London, SAF Publishing.
- Rosenstein, M. and P. Roussel. "Discography of Jon Rose." Retrieved January 15, 2008, from <http://www.wnur.org/jazz/artists/rose.jon/discog.html>.
- Tee, M. (2003). "The M Squared Secret Squirrel Society." M Squared Memoirs Retrieved December 6, 2007, from <http://www.spin.net.au/~mifilito/msquared.html>.
- Tee, M. (2003). "What they said about M Squared." M Squared Memoirs Retrieved December 3, 2007, from <http://www.spin.net.au/~mifilito/M2press.html>.
- Walker, C. (1996). Stranded: The Secret History of Australian Independent Music 1977-1991. Sydney, Pan Macmillan.

Timothy Opie

Box Hill Institute
1000 Whitehorse Road
Box Hill, 3124
Australia
timopie@fastmail.fm

Rendering Soundscapes

Abstract

The recording of an aural soundscape can be used as a structural framework in which the composer can reconceptualize and reshape the entire space. Using an eco-structuralist process the recording is reduced to a framework that affords an entirely new sonic perception, whilst retaining a sense of the original space. This paper discusses the framework and rendering of a seaside sound event.

Introduction

Eco-structuralism contains a set of rules and techniques that can be employed when working with recorded sounds. The techniques are designed to extract and maintain the structure of a recorded sound so that it may be used in a number of different compositional forms. In this paper the techniques will be used on a common seaside sound event to reshape the soundscape. The process and results will be analysed with a focus on the continuity and discontinuity of the sonic space.

Extracting the Structure

Eco-structuralism is a compositional technique that employs very strict rules for extraction and processing of sound event structures (Opie & Brown 2006). The rules are designed with persistence of the structures as a primary outcome. The persistent nature of the structure affords an understanding of the original encapsulating sound event, whilst also allowing the new schizophrenic¹ event to generate its own generational identity (Gibson 1966, Schafer 1994). Eco-structuralism attempts to interact with the sound event in a non-trivial way (Dunn 1984).

By zeroing in on the structure of the sound event itself, it becomes possible to extend the network of meanings along most time scales and into the realm of perceptual cues. (Keller 2006)

The structure used in this paper is a seaside sound event containing a common ocean wave and swell keynote sound of many beaches without any discernible soundmark. Because the pro-

cess requires sonic audition to make an informed decision on the process, this paper is attempting to represent some of the audible ideas graphically by applying similar techniques to a related image. Represented first in Figures 1 & 2 are the original source files. The image is just a quick snapshot of the space, representing only a few milliseconds of the event, whereas the recording represents 25 seconds of the event as it took place.



Figure 1. Image of the Ninety Mile Beach.

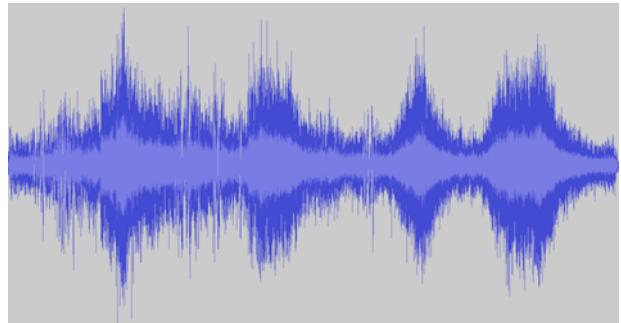


Figure 2. Waveform of beach sound event

The simplest structure to extract is the amplitude envelope. *ExtractAmplitude.java* performs an amplitude attribute study that analyses and extracts the peak amplitudes (Opie 2005). This program generates an XML document containing information on the audio file, including a description, file formatting, analysis length, samples taken, minimum and maximum values, links to the audio file and extra comments, and most importantly it lists the peak amplitudes. Being in an XML format the information is easily interoperable with other programs, and easily accessible. Figure 3 shows the beach image with the outlines highlighted, and the rest of the image information falling into obscurity. This is essentially what the composer is left with after performing an amplitude extraction process. All that is left is an envelope, which contains information on at-

¹ It is acknowledged that the sound is already schizophrenic, having been removed from its originating source and place in time, however this new process creates a much deeper level of schizophrenia.

tack peaks, sustains, decays and releases. Comparing Figure 2 to Figure 4, it is clear to see the correlation between the waveform of Figure 2, and the peak amplitudes of the graph in Figure 4. Audio tests with these amplitude structures have shown just how important the amplitude envelope is in determining the origin of the sound event (Opie 2005).



Figure 3. Image of the Ninety Mile Beach showing just outlines – comparable to amplitude structure.

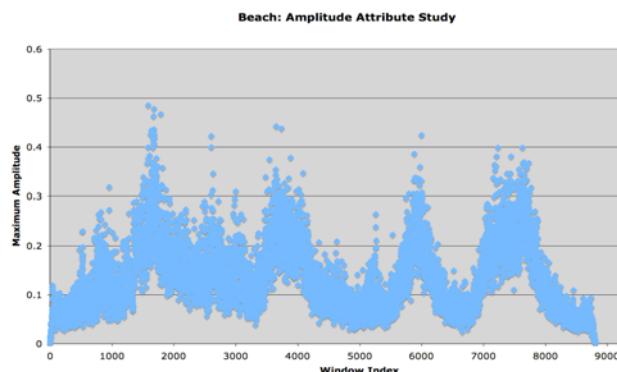


Figure 4. Beach sound event amplitude structure results graphed.

The second structure to be extracted is the timbre. Following the timbre attribute study *ExtractTimbre.java* performs an FFT analysis and determines ratios that define the timbral structure of the beach event. This information is collected in an XML document that is structured in the same way as the amplitude XML document. The timbre structure contains a coarse structure of relations between layered sounds within the event. The best graphical representation of the structure would be that of Figure 5 in which the outlines are becoming blurred, and textures are looking more like coarse grains. The timbre attribute study seeks to set up ratios by which a similar timbred sound can be created by applying the ratios to its fundamental frequency.

The timbre attribute study still needs some refinement, which may reduce the coarseness of the structure, although this may come at the cost of a ridiculously large data file. There are some avenues of investigation that need to be followed in order to improve this process, but these will not be discussed here. Figure 6 shows a spectrograph of the beach event, from which the timbral data is gathered. As can be seen from this figure, the timbre becomes very rich and noisy as the waves crash on the shore. There is a high correlation between the amplitude and timbral richness

that can be utilised and exploited with these structures. For example a sound model of a beach sound event could be emulated effectively by increasing the timbral richness proportionally to the amplitude increase. Such a model would operate entirely by manipulation of the amplitude, similar to the crackling fire model demonstrated by Opie (2005).



Figure 5. Image of the Ninety Mile Beach with the texture amplified – comparable to timbre structure.

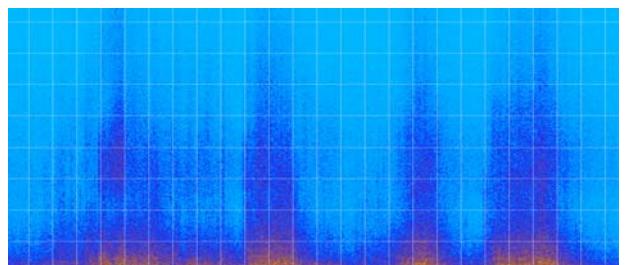


Figure 6. Beach sound event spectrograph.

The third structure to be extracted from the seaside event is the fundamental frequency. *ExtractFrequency.java* calculates and extracts the fundamental frequency, using a fundamental frequency attribute study (Opie 2006). The results are again recorded in an XML document that is structured in the same way as the amplitude document. The fundamental frequency study seeks to find the base frequency upon which all other harmonics rest. This is achieved by finding the significant-zero-crossing (SZC) and calculating the frequency based on the period. The SZC is determined by the peak amplitude value, of the waveform. The peak must fall within the mean range threshold of previous peaks. This process reduces multiple crossings that occur due to minor fluctuations that occur around the crossing point. Once the fundamental frequency is known, it creates a place with which to employ the timbre structure. Applying the amplitude structure to the combined frequency and timbre structure gives the resulting sound a sharper contour.

The fundamental frequency extraction process itself would be best visualised by slightly pixelating the image, as in Figure 7. There is a significant loss of timbral richness from the sound event, the base tone being the only feature retained. The picture can still be rebuilt with that knowledge, but it would look more like a paint-by-number exercise than a complete artwork. Figure 8 displays a graph of the fundamental

frequency. Most frequencies stay under 3000Hz, although there are a few anomalies reaching as high as 8000Hz during the crashing of the wave, which indicates some fluctuation of the signal. The lowest portion of the graph is most telling, as there is a definite rise in frequency that coincides with each crash of the waves.



Figure 7. Image of the Ninety Mile Beach showing average colour segregated into small windows – comparable to fundamental frequency structure.

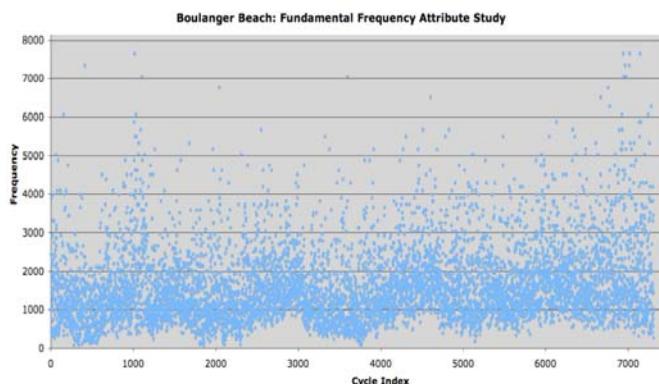


Figure 8. Beach sound event fundamental frequency structure results graphed.

A spatial attribute study could also be made, but as this is a mono recording such a test would reveal nothing. A durational attribute study has also been considered as the next step, but it has not been developed yet, although an approximation of the results reveal that the information is also captured in the timbre structure, although not yet in a singular accessible format.

Processing the Structure

With these three structures a new soundscape can now be rendered. The structures are like any building structure. They constitute the framework. Eco-structuralism allows for the rotation, stretching and compression of these structures (Opie & Brown 2006), although for this exercise they will be left unchanged, acting entirely as the framework from which they originated. The first rendering will impose a bell ringing upon the beach sound event. The bell ring contains a full rich tone, which can easily fill out the structures, in terms of amplitude and spectral detail. The fundamental frequency structure will cause a varying transposition throughout the bell ring, causing a warping of the sound. The timbre structure will also transpose the timbre of the bell

ring to suit that of the beach, and the amplitude structure will give the bell a new dynamic shape. Figure 9 depicts this process as adding a chrome finish to the beach image structure. There is definitely some transposition and rippling of the image, as well as new content added to the image. The beach form can still be made out, but it has an entirely new rendered façade. The sonic space is the same, although the subject has been imbued with a new entity. Because the beach structure is still intact, the perception of the beach should also remain somewhat intact, or it should at least afford the notion of a beach, as well as that of a bell.

Visually this has been captured. Figure 10 is the waveform of the new soundscape. As can be seen it is very similar to the original waveform. On listening to the new soundscape the perception of the beach is still present, the bell just creating a new perspective on the beach.



Figure 9. Image of the Ninety Mile Beach structure with a chrome finish – comparable to a using a rich tone as the content within which to fill the structure.

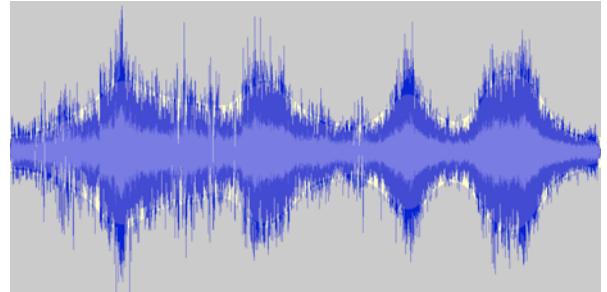


Figure 10. Waveform of beach sound event structure infused with a rich tone.

The Bell demonstrates this simple experiment effectively, but what if a much more contrasting sound is used as the new structure filler? What happens when you mix water and fire? As with the previous example the structure itself will not be transformed, it will act only as a direct framework in order to create a balanced juxtaposition. Visually the beach structure in Figure 11 has been filled with a dried earth image, and a stone wall for the sky, as this also seemed an appropriate visual contrast. The dry land has water washing over it, with a brick wall the horizon. The perception of the beach still exists, but the contrast between wet and dry causes the viewer to re-evaluate the entire image.

For an audio example a fire sound event works much better than the recording of dry

land. A campfire sample has been placed within the confines of the beach structure. The fire does not fill the space as readily as the bell. It leaves gaps in the rendered walls of the structure. The spectrum of the fire easily fills the beach structure, but the amplitude structure is much larger than that of the fire.

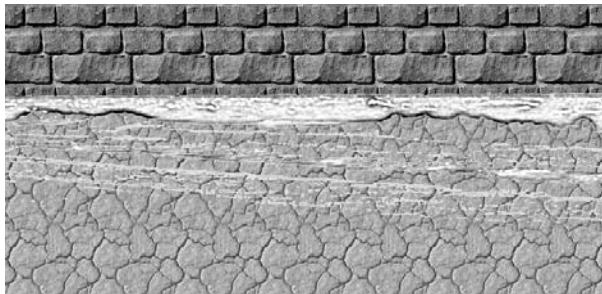


Figure 11. Image of the Ninety Mile Beach with entirely new images built within the structure - comparable to juxtaposing sound events created within the confines of the sound structure.

The recording of the fire is clearly discernable within the confines of the beach structure, but the beach structure can still be perceived. Water and fire can coexist peacefully, although not without causing the listener to re-evaluate what they think they are hearing. The affordances of water and fire would not have many overlapping features, they would instead cancel each other out more often than not.

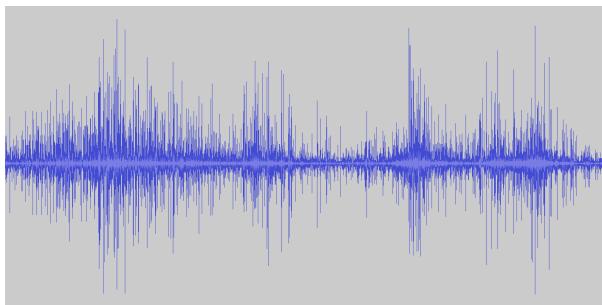


Figure 12. Waveform of beach sound event structure infused with a crackling fire sound event.

Future Structural Devices

A couple of future structural directions have already been mentioned in this paper, namely creating a duration attribute study, and also a beach-modelling project in which the beach sound event is physically modelled and emulated. Also the notion of structural transformation has also been raised. For this paper the space remained the same, whilst the contents changed. It would be interesting to examine how persistent the beach was after a number of structural transformations, especially if those transformations were quite radical. Eco-structuralism allows for quite a number of transformations to take place so long as the data remains in serial unity.

Other avenues that need investigating are the psychological ramifications of the extraction of structures, transformation and juxtaposition pro-

cesses. Whilst the structures can be perceived, are they still strong enough to afford the same response as would be received from the original recording (Dunn 2001)? Also if the sounds are used in the described way, would they promote better listeners? If the audience needs to re-evaluate all the sounds they are hearing would they listen more closely? If they are listening more closely to the newly rendered soundscapes, does that mean they would listen more closely to the real sounds in their environment (Schafer 1994)?

Conclusion

This paper has explored the extraction of a number of structures from a beach sound event, discussing some of the potential outcomes of these structures. It has then used the structures as a framework within which other sounds events were placed. The paper has shown that the structures can still be perceived despite a radical change of audio content. The perception of the two juxtaposing sound events causes re-evaluation and perhaps more meaningful listening of the newly rendered soundscape.

References

- Dunn, D. 2001. *Santa Fe Institute Public Lecture*, <http://www.daviddunn.com/~david/writings/sfilecture.PDF> Accessed 16/09/2007.
- Dunn, D. 1984. *Music, Language, And Environment*. <http://www.daviddunn.com/~david/writings/mle.pdf> Accessed: 16/09/2007.
- Gibson, J.J. 1966. *The Senses Considered as Perceptual Systems*. Boston, MA: Houghton Mifflin.
- Keller, D. and Capasso, A. 2006. "New Concepts and Techniques in Eco-Composition", *Organized Sound*. 11(1). Cambridge, UK. University Press.
- Keller, D. 2004. *Paititi: A multimodal journey to El Dorado*. Doctor in Musical Arts Thesis. Department of Music, Stanford University. CA, USA.
- Kivy, P. 2002. *Introduction to a Philosophy of Music*. Clarendon Press.
- Opie, T. and Brown, A. R. 2006. "An Introduction to Eco-Structuralism", *Proceedings of the International Computer Music Conference*. New Orleans, USA. ICMA, pp. 9-12.
- Opie, T. 2006. "Fundamental Frequency as an Attribute of Eco-Structualism", *Proceedings of the Australasian Computer Music Conference*, Adelaide, Australia.
- Opie, T. 2005. "Amplitude Analysis as Musical Composition", *Proceedings of the Australasian Computer Music Conference*, Brisbane, Australia.
- Schafer, R. M. 1994. *The Soundscape: Our Sonic Environment and the Tuning of the World*. Destiny Books.
- Xenakis, I. 1971. *Formalized Music*. Indiana University Press, Bloomington, USA.

Dr. Garth Paine

University of Western Sydney
 School of Communication Arts
 Coordinator VIPRE Research Lab
 BD 1.14 Werrington South
 Penrith South DC 1797
 Australia
 ga.paine@uws.edu.au

Abstract

Laptop performance of computer music has become wide spread in the electronic music community. It brings with it many issues pertaining to the communication of musical intent. Critics argue that performances of this nature fail to engage audiences as many performers use the mouse and keyboard to control their musical works, leaving no visual cues to guide the audience as to the correlation between performance gestures and musical outcomes. Interfaces need to communicate something of their task. The author will argue that cognitive affordances associated with the performance interface become paramount if the musical outcomes are to be perceived as clearly tied to realtime performance gestures, ie. That the audience is witnessing the creation of the music in that moment as distinct to the manipulation of pre-recorded or pre-sequenced events.

Introduction

Acoustic Instrument Model

In traditional musical performance, the interface has always been part of the instrument - part of the excitation-sonification system. For instance, all instruments in the string family share a mechanism for retaining the string which sees the string attached to the tail piece, run over the bridge, which sits on the sounding board, (dispersing and amplifying the vibrations of the string) and runs to the peg, held in the scroll allowing for variations in tuning. This mechanism provides for multiple independent sound sources (four or five strings) to be held on a single instrument body. Whilst this point may seem trivial, it allows a single musician to execute multiple parallel musical ideas on a single instrument, a techniques well illustrated in the solo cello suits by J.S Bach or the solo violin works of Paganini, although in reality, the design of the instrument only allows two notes to be sounded simultaneously, with three or four note patterns sounding in quick succession (appearing almost simultaneous).

Recent work carried out by the author (Paine, Stevenson, & Pearce, 2007) at the University of Western Sydney examined the fundamental control parameters utilised by expert musicians on traditional instruments. The resultant model (see Figure 1) presents the musical parameters; Dynamics, Pitch, Vibrato, Articulation and Attack/Release,

Interfacing for dynamic morphology in computer music performance

tack/release as the focus of the physical instrument control, and of primary focus in achieving a well developed instrumental tone, the principle concern for all interviewed musicians. The notion of musical expression, or musicianship plays a role in aesthetic decisions, and acts as an overall metric for the musical sensibility that underlies all instrumental training and musical decision-making. The above musical parameters were then considered in terms of the physical control musicians employ in order to achieve musical outcomes (see Figure 2). For instance one subject commented that "pitch is controlled by the position of the bow, as well as the movement of the bow between the bridge and the fingerboard", a flute player commented that a "fast air stream produces a dark tone, whilst a relaxed, slow air stream produces a lighter, softer tone. The higher pitch registers are achieved by using a faster air stream, angled slightly upwards,

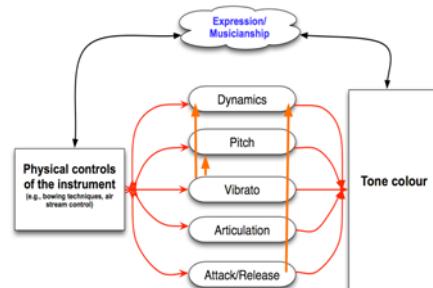


Figure 1. Musical parameter overview

whilst the lower registers require a slower downward air stream". An analysis of such statements lead to the identification of four primary physical controls; pressure, speed, angle and position (Figure 2).

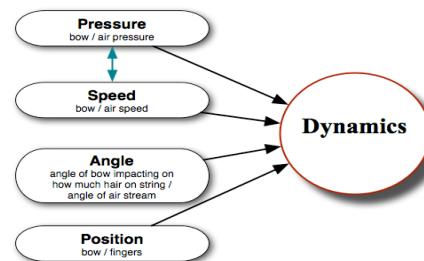


Figure 2 Control parameter model

Applying the Model in Computer Music

In computer music, the excitation-sonification relationship is broken into interface and synthesis algorithm.

The selection of an interface that provides the above defined physical control parameters; pressure, speed, angle and position, and the careful mapping of the parameter space can encourage a link between generation (gestural quality – intense, fast, light etc) and musical outcome. The author has been experimenting with the application of the above physical control parameter model for electronic music performance using experimental interfaces such as the Wacom Graphics Tablet and the Nintendo Wii Remote (WiiMote). This discussion will focus on the Wiimote.

Nintendo Wii Remote

The Nintendo Wii Remote (Wiimote) has been the focus of an explosion of music performance, DJ and VJ experimentation. Software frameworks quickly appeared for the Macintosh computer and the Windows and Linux operating systems, and were compiled into objects or plugins for programming languages such as Max/MSP, Quartz Composer (QCWii), Isadora, OSCulator (providing an OSC bridge) and were embedded in applications such as WiiToMIDI, Wiinstrument and Wii Loop Machine (actually developed in Max/MSP)

Wiimote: X-axis accel, X-axis accel, X-axis accel, Pitch, Roll, Yaw, 2 or 3 Buttons, WiibarX, WiibarY

Nunchuck: X-axis accel, X-axis accel, X-axis accel, Pitch, Roll, Yaw, X/Y of joystick, 2 Buttons
Table 1 Wiimote and Nunchuck data set

Part of the attraction will have been that the interface is wireless (Bluetooth), but a major factor will also have been the range of control afforded by the Wiimote. The Wiimote contains eleven buttons (momentary) and a three-dimensional accelerometer (6 continuous streams of data – see Table 1 below). In addition to this already substantial data set, the Wiimote accepts an accessory called a Nunchuck, which is held in the remaining free hand and contains a second three-dimensional accelerometer, a traditional two-dimensional joystick and two trigger style buttons (C and Z).

One of the characteristics of such an interface is that the buttons and or the joystick can be used in parallel with, and independent of the three-dimensional accelerometer. This means that the

following data streams can be performed semi-independently and simultaneously.

An infrared Sensor Bar (Wiibar) can be added to this setup to sense the absolute position of the Wiimote on the X and Y-axis, providing an additional two continuous data element. Table 1 indicates that the Wiimote can produce up to nine simultaneous but partly inter-related data streams (6 continuous and 3 momentary) and the Nunchuck can produce up to ten simultaneous but partly inter-related data streams (8 continuous and 2 momentary). When combined with the Wiibar a data set of 16 continuous and 5 momentary elements can be produced.

This represents too many pieces of data for an individual to use constructively at any one time. In order to consider the Wiimote/ Nunchuck combination against the acoustic instrument model outlined above, an analysis of the parameter space was examined against the characteristics of pressure, speed, angle and position.

Pressure: XAccel, YAccel, ZAccel, XJoy, YJoy. Pitch, Roll and Yaw may all be interpreted in this way as they provide a continuous data stream.

Speed: can be calculated from all continuous data streams

Angle: Pitch, Roll and Yaw

Position: Pitch, Roll and Yaw, NunJoyX/Y, WiibarX/Y and all buttons (11 with up to 5 simultaneously)

As the continuous data streams are simultaneous but partly inter-related, some decisions need to be made regarding the use of control parameters in synthesis where by their musical outcomes could conflict. For instance the X/Y/Z Accel data streams cannot be independent – a movement in any one axis will cause some variation in the other axis. Pitch, Roll and Yaw may also be affected by accelerative movements in the X/Y/Z planes, however within themselves they present more independence than the acceleration data, which in all planes is subject to gravity. The buttons are of course independent although only momentary. The fundamental consideration in constraining/choosing the data set is an analysis of the independence or otherwise of the generation of the control data, guided by an understanding of the gestural qualities best suited to the timbral/pitch/temporal qualities of the composition, and the way in which these musical potentials could be navigated in performance.

It should also be noted at that this point that the more advanced Wacom tablets such as the Intuos3 provides at least thirteen pieces of continuous data and two buttons from a single pen (multiple pens may be independently mapped, but not used simultaneously) as is illustrated in the Wacom pen assignment page in the Kyma software (see Figure 3).

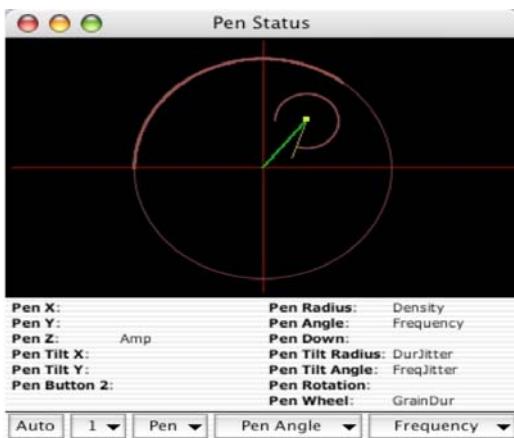


Figure 3 Kyma software Wacom interface status screen

Instrument Design

Once the controller parameters are understood, the instrument design begins. In addition to the above discussion, electronic music performance differs from acoustic in that the composer designs and builds the instruments to suit the demands of the composition. The gestural quality of interface parameter sets plays an important part in the associated gesture of the musical outcomes – slow or fast transitions, fast chaotic timbral changes, accuracy in pitch or amplitude and repeatability are just some of the considerations at hand. It should be noted that the composition, instrument design, interface design and mapping strategies apply constraints that define the musical work and the performance options. Instrument design requires some serious consideration of musical aesthetic and the relationship between performer and audience. The interface and its implementation therefore serve two primary goals: 1) to increase performability, allowing the musician to nuance musical outcomes in a way not possible with existing interfaces or using the mouse/keyboard computer interface; 2) To increase communication with the audience, displaying something of the energy and intent of the performer, providing a conduit for engagement in the realtime qualities of the performance – ie. The ritual of performance.

Any implementation of a new musical interface must therefore consider the ecology of this environment. Gurevich and Treviño (2007) discussed the development of a framework for an ecology of musical action:

An ecological framework without the assumption of a commodity or a singular creator makes it admittedly difficult to unify or relate the experiences of the individual actors in the system. Donald Norman's (Norman, 2004) formulation of three levels of processing in the human brain and associated modes of experience facilitates a meaningfully descriptive but inclusive consideration of the musical experience from a variety of points of view. The three levels of processing are

visceral, automatic and pre-wired reactions to sensory stimuli; behavioral, involved in the subconscious control of learned everyday actions (driving a car, taping, playing a violin); and reflective, the highest-level conscious thought in which we form opinions, plans, and abstractions. ... While Norman argues that good design requires a balanced appeal on all three levels, it is also clear that all three levels are engaged in creating music. ... Norman describes the skilled performer's ability to play a piece unconsciously (Behavioral) while simultaneously considering matters of the large-scale form (reflective). The listener reacts viscerally to the sound and may also contemplate meaning. (pg. 109)

Gurevich and Treviño point to the difficulty of identifying a unified source for musical creation in a complex system where the interface and the sonification mechanism are separated. Their adoption of the three levels of processing outlined by Donald Norman (Norman, 2004) illustrate some of the cognitive associations brought into play when engaging both as a performer and an audience member during a musical performance. The author supports Gurevich and Treviño statement that musicians are working at all three levels identified by Norman when improvising and performing dynamic scores. The visceral and behavioral levels are enshrined in the kinetic gesturing that brings about the musical outcomes, representing a sonification of the performative gesture. The reflective layer is brought to bear by the musician who is actively, but probably subconsciously planning form, structure and harmonic progression. The momentary and the future abstract are always co-existent and active. Schwartz and Godfrey (1993) define seven principle concepts in contemporary composition: Pitch Logic, Time, Sound Color, Texture, Process, Performance Ritual and Parody (Historicism). Each of these can be considered in terms of both instrument design and mapping strategies for musical interfaces and extend the Norman proposal towards a more detailed musical focus..

Ortiz Pérez, Knapp and Alcorn (2007) also discuss this issue when developing a composition-driven approach to using novel interfaces in Díamair for choir and integral music controller. The diagrams in Figure 4 illustrate the way in which the compositional decisions influence the interface and in turn how the interface design influences the instrument design (software synthesis) and compositional decisions.

A combination of these considerations is brought to bear when developing a musical interface that also addresses composition constraints. One example is the author's work with Michael Atherton in their ensemble SynC. The compo-

sition titled *Encounter*¹, for Hurdy-Gurdy and live electronic processing from the Parallel Lines CD (Paine & Atherton, 2006), utilised the Capybara/Kyma system² for live electronic processing of the Hurdigurdi sound, and implemented the WiiMote as the control interface.

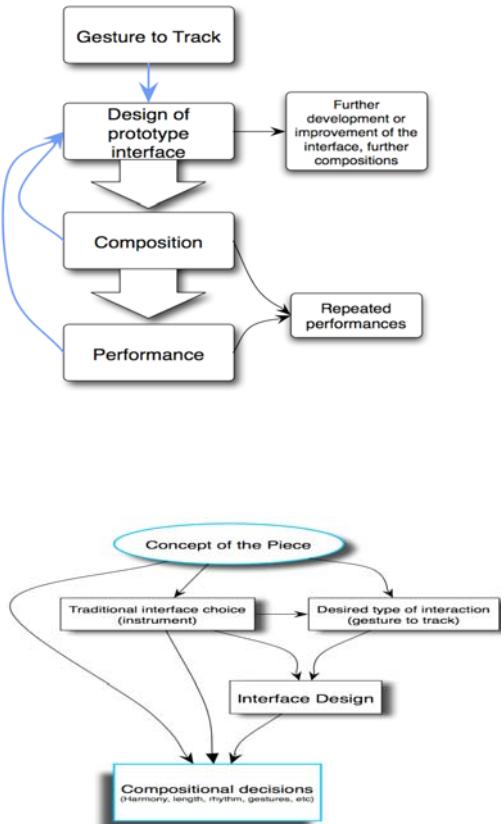


Figure 4 Composition driven approach to using novel interfaces (above and below)

Encounter – an example of the Wii-

Mote as musical controller

In developing the synthesis mapping for the Wiimote in this musical work, a number of compositional determinates were considered. The composition contains a large number of synthesis variables. It was not possible to constrain the variable in such a manner that they could all be controlled in realtime. The author chose to automate some variables and leave the following under realtime control; filter center frequency, the triggering of buffer recordings (four buffers), the rate/density of granulation, the triggering of a sample playback (one pre-recorded sample is

used in the work), the control of two frequency/pitch variables, the sample-and-hold rate of these oscillators and the delay feedback time also associated with this oscillator instrument which is the solo electronic instrument in the central section of the work.

WiiMote mapping for the musical work *Encounter* is outlined in Table 2.

Synthesis Variable	Control Source
WispCrossSynth	→ Nun.Joy.Y
GrainDesnsity	→ Nun.Joy.X
PlayOn	→ Nun.Accel
RecordBuffer1	→ Nun.Trigger2
RecordBuffer2	→ Nun.Trigger.C
LiveRandomBuffer	→ WiiTrigger
Record3	→ WiiButton A
SampleTrigger	→ WiiAccel
Freeze	→ Nun.Joy.Y (acts on Buffer 1, 2, 3 in the Track&Hold Time Freezer)
ChopA	→ WiiPitch (acts on Buffer 3)
RCGain	→ WiiRoll
RCHoldTime01	→ Nun.Yaw
RCFreqMod	→ Nun.Pitch

Table 2 data mapping for *Encounter*

Conclusion

The author has illustrated an application of the Nintendo WiiMote and NunChuk interface for realtime electroacoustic music performance. The composition *Encounter* was used to discuss the gestural qualities of the interface, compositional constraints and engagement with an audience in addition to the current mapping strategy. The work is still in progress - no formal conclusion has been reached, however, it is clear that the proliferation of wireless gaming interfaces, notably the Nintendo WiiMote, make these questions more relevant and approachable. This makes the development of a shared theoretical basis for interface design and comparison increasingly important. Gurevich & Treviño's (2007) discussion of a framework for an ecology of musical action concludes that Norman's three levels of processing offer a new currency for describing the experience of music creation [placing] the electronic music interface in context. This framework has three distinct advantages: 1) it admits a broader range of aesthetic concerns; 2) it provides a more meaningful way to 'evaluate' an interface; and 3) it expands the scope for the consideration of novel interfaces (pg.109). It is hoped that the musical instrument control model outlined in Figures 1 & 2 will also contribute to this discourse.

References

- Cacone, K. (2002). Laptop Music: Counterfeiting Aura in the Age of Reproduction. *Parachute*, 107, 56.

¹ From Sync, Parallel Lines, Celestial Harmonies, 2006. An excerpt can be heard at http://www.syncsonics.com/blog/?page_id=39 viewed 18/09/07

² See <http://www.symbolicsound.com> viewed 18/09/07

- Gurevich, M. & Treviño, J. (2007). Expression and Its Discontents: Toward an Ecology of Musical Creation. Paper presented at the International Conference on New Interfaces for Musical Expression (NIME07), NY.
- Hunt, A. & Kirk, R. (2000). Mapping Strategies for Musical Performance. In M. Wanderley & M. Battier (Eds.), Trends in Gestural Control of Music. Paris: IRCAM - Centre Pompidou.
- Lazetta, F. (2000). Meaning in Musical Gesture. In M. Wanderley & M. Battier (Eds.), Trends in Gestural Control of Music. Paris: IRCAM - Centre Pompidou.
- Miranda, E. R. & Wanderley, M. M. (2006). New digital musical instruments : control and interaction beyond the keyboard, Wis: A-R Editions.
- Mulder, A. (1994). Virtual musical instruments: Accessing the sound synthesis universe as a performer. Paper presented at the First Brazilian Symposium on Computer Music.
- Mulder, A. (1996). Getting a grip on alternate controllers: Addressing the variability of gestural expression in musical instrument design.
- Mulder, A. & Fels, S. (1998). Sound sculpting: Performing with virtual musical instruments. Paper presented at the 5th Brazilian symposium on computer music, Brazil.
- Norman, D. A. (2004). Emotional Design. New York: Basic Books.
- Ortiz Pérez, M. A., Knapp, B., & Alcorn, M. (2007). Díamair: Composing for Choir and Integral Music Controller. Paper presented at the International Conference on New Interfaces for Musical Expression (NIME-07), NY.
- Paine, G. (2002). Interactivity, where to from here? *Organised Sound*, 7(3), 295-304.
- Paine, G. & Atherton, M. (2006). Parallel Lines [Musical Composition]. Tucson, Arizona: Celestial Harmonies.
- Paine, G., Stevenson, I., & Pearce, A. (2007). The Thummer Mapping Project (ThuMP). Paper presented at the International Conference on New Interfaces for Musical Expression (NIME07), New York City, NY.
- Schwartz, E. & Godfrey, D. (1993). Music since 1945: issues, materials, and literature. New York Toronto New York: Schirmer Books Maxwell Macmillan Canada Maxwell Macmillan International.
- Wanderley, M. (2001). Gestural Control of Music. Paper presented at the International Workshop Human Supervision and Control in engineering and Music, Kassel Germany.
- Wanderley, M. (2002). Quantitative Analysis of Non-obvious Performer Gestures. In I. Wachsmuth & T. Sowa (Eds.), Gesture and Sign Language in Human-Computer Interactions. Verlag: Springer.
- Wanderley, M. M. & Orio, N. (2002). Evaluation of Input Devices for Musical Expression: Borrowing Tools from HCI. *Computer Music Journal*, 26(3), 62-76.
- Zicarelli, D. (2007). Max/MSP. Retrieved 18/09, 2007, from <http://www.cycling74.com>

Internet References

- i See <http://www.wacom.com.au> viewed 13.09.07
- ii See <http://www.nintendo.com.au> viewed 13.09.07
- iii See http://www.wiili.org/index.php/Wii_mote_driver viewed 13.09.07
- iv See http://www.iamas.ac.jp/%7Eaka/max/#aka_wiiremote viewed 13.09.07
- v See <http://www.daisyrust.com/2007/02/12/> viewed 18/09/07
- vi See <http://www.troikatronix.com/isadora.html> viewed 18/09/07
- vii See <http://www.osculator.net/> viewed 18/09/07
- viii See <http://mike.verdone.ca/wiitomidi/> viewed 18/09/07
- ix See http://screenfashion.org/releases/the_wiinstrument/ viewed 18/09/07
- x See <http://www.theamazingrolo.net/wii.html> viewed 18/09/07
- xi From SynC, Parallel Lines, Celestial Harmonies, 2006. An excerpt can be heard at http://www.syncsonics.com/blog/?page_id=39 viewed 18/09/07
- xii See <http://www.symbolicsound.com> viewed 18/09/07

Mark Pedersen

Chailight Productions
 PO Box 26
 Moreland, 3058
 Australia
 markp@chailight.com

Abstract

We describe an implementation of a gestural interface for manipulation of computer mediated musical material based on a user-configurable intention map, in order to support more direct expression of the performer's expressive intention than is typically achievable with direct manipulation of parameters. The particular application of intentional maps and gestural control to live spatialisation in surround sound performance is explored.

Introduction

Interfaces for computer music performance typically suffer from being directly oriented toward the control of sound production (e.g. directly controlling pitch, duration, tempo). In addition, more subtle controls relating to timbre of synthesised sounds are frequently oriented toward a single parameter among many, making physical control of complex timbres difficult. The overall result is that computer music performance interfaces lack expressivity, inasmuch as they hinder the performer's effort to express complex performance intentions (Rovan 1997). Furthermore, the typical lack of obvious connection between a performer's action and the music being performed is a well-documented issue with respect to audience engagement during computer music performance (Schloss 2003).

Related Work

Canazza et. al. (2003) describe an abstract control space for the expression of intentions in music performance which incorporates the results of studies on the emotional response of listeners to differing performances of a piece of music. The listeners' responses are correlated with changes in the performance characteristics in order to identify the broad dimensions of a control space oriented around the expression of emotional intentions and a mapping to performance parameters.

The development of an interface which incorporates knowledge of a mapping from emotional intention to performance parameters holds promise for overcoming the limitations typically associated with computer music interfaces. The challenge presented by Canazza's proposed abstract control space is how to implement a phys-

User Configurable Intention Maps for Gestural Control and Live Spatialisation in Computer Music performance

ical interface which retains transparency with respect to intention and performance in such a way that it is engaging for both the performer and the audience.

Designing such a physical interface has drawn upon prior work in gestural control and parameter mapping and interpolation. Both topics are too broad to fully review here, however in the area of parameter mapping particularly relevant work has been done by Hunt & Wanderly (2002) and Verfaillie et. al. (2006). Bencina's (2005) work on parameter interpolation is the inspiration for the system used here. In the specific area of gestural control for spatialisation, work by Schacher (2007) was a significant starting point.

User Configurable Intention Maps

We propose a gestural interface for manipulation of computer mediated musical material, utilising the concept of a user-configurable intention map, in order to support more direct expression of the performer's communicative intention than is typically achievable with computer music interfaces oriented around direct control of sound production parameters. We augment the control space proposed by Canazza et. al. (2003) with a gesture recognition system that allows for nuanced expression of emotional intention by physical means.

Description

The architecture of the system consists of a gestural recognition layer, an intention control layer, and an audio control layer, with mappings between each layer. The user is able to configure the gestures to be recognised, the mapping from gestures to intentions, and the mapping from intention to performance parameters. The user can interact with the intention control layer directly (i.e. the mapping to the gesture space is one-to-one) or via complex gestures. Gesture recognition can be used to create an expressive 'shorthand' for complex control positions within the intention space.

The system can be used to control performance at varying levels of detail, for example the tempo of sequence playback as well as the articulation of notes and the timbre of synthesised material. Multiple sound objects (or parts in a polyphonic performance) can be controlled individually. An implementation of the system in

Max/MSP has been developed, providing a variety of physical input methods including video recognition, Wacom tablet, and wireless motion sensors. The particular interface explored in this paper is based on the monome 40h (Crabtree and Cain 2007): an 8x8 button grid device augmented with a two dimensional accelerometer and an infra-red distance sensor. The infra-red distance sensor is mounted underneath the device, measuring distance to the surface over which the device is being held.

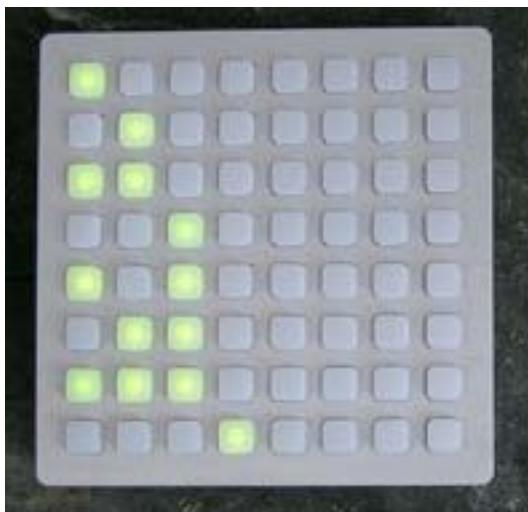


Figure 1. The monome 40h provides an 8x8 button grid with independent visual feedback via LEDs within the buttons.

Implementation

The implementation in Max/MSP incorporates a number of pre-existing modules.

The gesture recognition layer makes use of Frederic Bevilacqua et al's Gesture Follower, which uses a Hidden Markov Model to recognise patterns in multi-dimensional data based on a supplied set of examples (Bevilacqua 2007).

The gesture recognition layer outputs a label for each recognized pattern, corresponding to a position in the intention map. The intention map is based up Oliver Larkin's INT.LIB (Larkin 2007): an interpolated preset interpolation control for Max/MSP. Positions in the intention map are arranged according to the two dimensional distribution of performance intentions proposed by Canazza (2003), namely: Bright, Dark, Hard, Light, Soft, Natural, Heavy. Movement between these positions in the intentional map results in smooth interpolation of control parameters associated with each position, as shown in Figure 2.

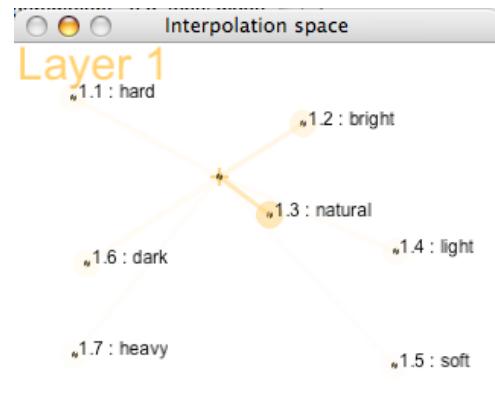


Figure 2. Example layout of interpolation space following Canazza 2003.

User configuration involves two phases:

- training the gesture recognition layer to recognise input patterns corresponding to one of the available intention labels
- configuring/adjusting performance parameters according to the target performance device.

The gesture recognition system takes 3 inputs from the monome 40h: the X and Y axis of the accelerometer input, and distance from infrared sensor, providing the user with an approximation of 3-dimensional motion sensing. Gesture recognition is not continuous: the user triggers the start and end of an intended gesture by selecting a dedicated button on the monome 40h.

In addition to the mapping of symbolic gestures to intentions, the system also affords direct manipulation of position in the parameter space. Direct manipulation maps the 2 dimensional accelerometer input from the monome 40h into motion in the 2 dimensional intention mapping space, allowing the user to tilt the device in the desired direction. Direct manipulation can be handled either in absolute positioning or as interpolation from the current position, allowing the user to 'nudge' the interpolation cursor in the direction of tilt. Switching between gesture recognition and direct manipulation modes is selected using dedicated control buttons on the monome 40h.

Gesture Selection

The author's experience arising from both development of and performance with the intention mapping and gesture recognition system suggests that care must be taken to select gestures which are sufficiently orthogonal so as to give robust results from the recognition engine. Simple flicks to the left and the right, forward and back gave strong results. Likewise, vertical movement is easily recognised. Combinations of the above resulted in some ambiguous results in terms of recognition. For instance, rapid side-to-side motion was not always distinguished from rapid front back motion, presumably due to the

difficulty of isolating rapid movements to be only along one axis.

This issue highlights another challenge with the gesture recognition, in that the user must train themselves to repeat gestures reliably enough when training the system that the gestures are recognisable. Limited variation in the gestures helps the robustness of the recognition, as minor variations are mapped to the same intention, however gross variations will reduce the effectiveness of the recognition algorithm.

A further limitation is that gestures are mapped to intention categories without any scalar factor indicating the "strength" of the gesture. In comparison, typical human interaction, for instance with an orchestra conductor, would afford recognition of symbolic gestures as well as the degree or strength of the gesture through factors such as speed or expansiveness of the gesture. Alternative inputs, such as pressure sensitivity, which remain orthogonal from the recognition system, are being considered as means of overcoming this problem.

Mapping to Parameter Space

In the context of computer music performance, the available parameters for intention mapping are vast. However it is sensible to restrict the possibilities to those which are typical for musical performance.

Working within the framework of using traditional instruments such a piano to render sequenced note information, Canazza et. al. focus on mappings from intention states to energetic and kinetic parameters of performance such as tempo and dynamics.

In terms of traditional instrument performance, Paine (2007) groups performance parameters broadly under the headings of Dynamics, Pitch, Vibrato, Articulation, Attack and Release as the mediating parameters between the physical instrument and the desired tone colour. For specific instruments, additional parameters (such as bow pressure, position and speed for violin or cello) also contribute to the desired tone colour.

We adopt a similar approach, replacing vibrato with a more general and direct set of parameters relating to timbral brightness and noise. Given the possibility of using the intention mapping technique to manipulate sequenced material, it is also possible to influence the tonal mode of the playback, in much the same way that a composer may choose to work in a specific key, or modulate to a new key in order to express an emotional state or intention.

Mapping of these abstract qualities needs to be made specific for any particular synthesis engine. For example, brightness may be expressed in terms of filter cutoff frequency and resonance, EQ or even reverb amount. The system allows for parameters to be communicated via MIDI to external synthesis engines or communicated directly to embedded effects that have been developed within the system itself.

Application to Live Spatialisation

Spatialisation of sound for surround-sound performance can be handled as simply another parameter in the intention space, however a number of challenges emerge:

- spatialisation performances typically manipulate multiple sound objects, while the existing intention mapping architecture is designed for the manipulation of only a single sound object
- merely associating a single spatial position with an expressive intention is overly restrictive
- directly manipulating a parameter which effects the rendering of a sound object on a 2 dimensional plane is potentially confusing for a performer who is navigating an abstract 2 dimensional plane.

In terms of the first issue, control of multiple sound objects is afforded by the use of layers in Larkin's (2007) interpolation space. Each layer in the interpolation space is configured with an identical intention map layout, with each layer connecting to the control parameters for the associated sound object. Thus separate sound objects, with separate synthesis engines, can be controlled from a common set of intentional gestures. Selection of the active sound object is achieved through dedicated buttons on the monome. Simultaneous control is limited currently to affecting all layers simultaneously.

A significant disadvantage of the layering approach is the issue of additional cognitive load generated by using a single interface device to control multiple virtual objects. Context shifting between the objects tends to break the continuity of parameter settings and the control surface, requiring the user to keep track of which object they are controlling and beware of making unintended changes to the active object.

This issue is somewhat alleviated by having a common intention map across all objects, and common gestures for each intention. Interpolation from the current position in the intention map to the desired goal helps to reduce the incidence of discontinuities arising from context switching, and for this reason gestural control or 'nudging' is preferred over direct manipulation of the intention mapping space when working with multiple objects and layers.

In addressing the issue of decoupling spatial position and intention, a number of approaches have been taken. While it may suit the compositional/performance requirements of some pieces to associate a particular performance intention or tonal quality (dark, bright, heavy, etc) with a particular position in the sound field, being limited to always positioning sound objects in the "bright corner" or the "dark corner" is not desirable. Hence a more sophisticated approach to mapping intention to spatialisation parameters is required.

In the first instance, we associated complex trajectories with each intention. Thus allowing, for instance, a shift into the “light” intention space to result in the sound object bouncing swiftly around the sound field, whereas a move into the “heavy” intention space slows the movement down to settle in the centre. Care must be taken to manage switching between trajectories to avoid unintended discrete jumps from one spatial position to another.

The current implementation associates fixed trajectories with each intentional position and applies changes in trajectory once the sound object passes through a common point. Trajectories can be recorded by the user directly and stored for later use. This allows the user to develop a library of spatialisation trajectories that can be linked with performance intentions when configuring the system. Other scalar factors such as translation in the XY plane, and scaling of trajectories can likewise be applied.

Trajectory coordinates are not the only parameter available for control. Speed, and direction along the trajectory (forward/reverse), are controllable and, as scalar values, are easily interpolated using the current architecture.

Apart from automated control of spatial position via fixed trajectories, direct manipulation of the spatial position of a sound using the monome 40h is also possible. In this instance, a separate spatialisation layer is added to the parameter interpolation control, in which the preset reference points are positioned to mirror the speaker layout of the performance space.

Movement in this layer directly maps on to movement in the sound field. Again, the use of additional layers of control adds additional cognitive load, subjectively speaking, however the context switching is not as problematic as when switching between multiple objects with an overlaid control space. While possible, maintaining control of a separate intention map and spatial position map for multiple objects has proven challenging in performance, particularly as context switching between layers becomes more problematic.

Visual Feedback

The advantage of the monome 40h as a control device is that it affords a degree of visual feedback within the device itself, which is not available in other devices such as Wacom tablets or handheld accelerometers such as the Wiimote, although these devices may provide more detailed feedback via computer displays.

Selection of sound objects is achieved by selecting a button along the bottom row of the monome. Buttons representing available objects (up to eight) are lit, and the selected object flashes.

Likewise, toggling between the timbral control layer and the spatial control layer can be done by selecting a button on the top right of of

the monome, with a flashing status indicating spatial control.

Position within the control space (for the purpose of direct manipulation) can be indicated by a “cursor” on the monome, which moves according to the tilt of the device. However, this form of feedback was found to be inadequate for performance purposes compared to the display of the interpolation space on computer screen, which provided details of the various significant positions within the space. In addition, the spatial control position could also be monitored concurrently on the computer screen, whereas the monome 40h display required context switching.

Future Work

A number of areas of future work remain. Foremost is a more formal evaluation of the intention mapping system with a range of performers to assess usability in a performance context, particularly looking at issues of cognitive load in surround surround performance, and robustness of gesture recognition.

Switching between fixed trajectories is somewhat of a limitation in terms of choice of trajectories and achievement of a similar level of interpolation in the spatial parameter space as in the timbral parameter space. Interpolation between trajectories requires calculation of intermediate trajectories which can be applied progressively without noticeably interrupting the current motion of the object. A solution to this problem, based on approaches to non-intersecting graph morphing, is currently being investigated.

Refinement of the interface, including the addition of pressure inputs or other means of indicating the intended strength of a gesture is highly desirable.

Given the limitation on the number of directly controllable sound objects, hierarchical grouping and control along the lines suggested by Schacher (2007) has been identified as a potentially fruitful direction for further research into the management of multiple objects within the intention mapping architecture. Linking gestures to complex transitions within the parameter space, either for individual sound objects or multiple groups would also extend the expressive potential of the system.

Conclusion

The system developed addresses issues with expressivity and transparency in computer music performance interfaces by providing the performer with a spatial representation of expressive options that can be navigated either directly or by means of configurable gestures, allowing complex control of performance parameters in a manner more closely linked to the performer's own expressive intention.

In particular, the application of intentional maps to control of live spatialisation has been

explored and a number of issues identified and workable solutions developed.

References

- Bevilacqua, F., Guédy, F., Schnell, N., Fléty, E., Leroy, N., 2007, "Wireless sensor interface and gesture-follower for music pedagogy", *Proceedings of the International Conference of New Interfaces for Musical Expression (NIME 07)*, New York, NY, USA, pp 124-12
- Canazza, S., et al., 2003, "An Abstract Control Space for Communication of Sensory Expressive Intentions in Music Performance," *J. New Music Research*, vol. 32, no. 3, 2003, pp. 281-294.
- Crabtree, B. and Cain, K. 2007, monome, Philadelphia, USA, <http://www.monome.org>
- Hunt, A. and Wanderley. M. 2002, "Mapping Performance Parameters to Synthesis Engines". *Organised Sound*, 7(2):97–108, 2002.
- Larkin, O., 2007, "Int.Lib – A Graphical Preset Interpolator For Max MSP", *Proc. International Computer Music Conference 2007*, Copenhagen
- Paine, G., 2007, "Interfacing For Dynamic Morphology In Computer Music Performance", *International Conference on Music Communication Science*, Sydney, Australia
- Schacher, J., 2007, "Gesture Control of Sounds in 3D Space" *Proceedings of the 2007 New Interfaces for Musical Expression Conference (NIME 07)*, New York, NY, USA.
- Schloss, W. A., 2003, "Using Contemporary Technology in Live Performance: The Dilemma of the Performer", *Journal of New Music Research*, Volume 32, Issue 3 September 2003 , pages 239 – 242
- Rovan, J., Wanderly, M., Dubnov, S., Depalle, P., 1997, "Instrumental gestural mapping strategies as expressivity determinants in computer music performance" *Proceedings of the Kansei - The Technology of Emotion Workshop*, Genova - Italy, Oct. 1997.
- Verfaille, V., Wanderley, M. and Depalle, P., 2006 "Mapping Strategies for Gestural and Adaptive Control of Digital Audio Effects", *Journal of New Music Research*, 2006, Vol. 35, No. 1, pp. 71 – 93

Jean Penny

Queensland Conservatorium,
Griffith University
mjpenny@bigpond.net.au

Amplified breath - (dis)embodied habitat: exploring sonic interfaces of performance, electronics, space and flautist identity

Abstract

The electronic spatialization of flute sound represents an intersection of acoustic resonance and technology, ancient sonic imagining and present day cultural imprints. The performative impacts of technological interventions and the new responses thus stimulated in performance practice, invite rigorous exploration and reflection. In this paper I aim to articulate my personal responses to spatialization technologies and live interactive electronics, to explore the expanded capacities provided for the flautist, and the new performative elements introduced by this genre. Focusing on breath tone and gesture, transformations of sonority, performer identity, space, physicality and interactivity will be investigated from the flautist's point of view. Warren Burt's new work for flute and live interactive electronics, *Mantræ* (2007), is presented as an illumination of performance exploration. In this work, the movements of the flautist trigger sound events that radically alter the flute timbre, using camera signals to activate computer sound modification software, *Plogue Bidule* and *Cycling '74's Hipno*.

Introduction

Solo flute works characteristically exploit spatial tonal qualities. The flute's resonance has motivated (and perhaps seduced) composers and performers to explore ways of capturing and expanding its translucent sonic qualities into evocative expression. This sonic ethos stretches across the centuries, from the architectural structures of Baroque works, such as J. S. Bach's *Partita in A minor* or Jacob van Eyck's *Der Fluyten Lust-hof*, to the more recent works of composers such as Takemitsu and Saariaho, in which the resonance and malleability of the tone and breath are a critical element of composition. Edgard Varese's seminal composition for solo flute, *Density 21.5* (1936), initiated new tonal directions through the use of key clicks, dynamics as sound manipulation and extremes of tessitura. This work generated immense change in the perception of flute sound and capabilities, and encouraged practitioners to explore new sonic paths. New texts began to appear examining such techniques as microtones, whistletones, breath tones, tongue clicks, key slaps, multiphonics, harmonic, glissandi, new embou-

chure techniques, vocalization and varieties of articulation by such authors as Bruno Bartoluzzi (Bartoluzzi, 1982), Pierre-Yves Artaud (Artaud, 1995) and Robert Dick (Dick, 1989). These techniques became the source material of composers, scrutinized for sonic properties and expanded out into the musical milieu through amplification and spatial projection.

Contemporary flute players have embraced these techniques, especially in conjunction with electronic interventions, revelling in the expanded sonic palette, a fresh abundance of ideas and new discoveries in expression and performance. The overlapping layers of instrumental performance and extended practice through electronic techniques have evolved into a complex search for interpretative cohesion and confluence, an engrossing search full of chance meetings with thoughts and new responses to movement, emphasis and projection.

My focus in this paper is on transformations of sonority, physicality, identity and space as they occur in the performance of music for flute and electronics. My vehicles for exploration are the breath tone, gesture and interactivated space.

Transformed Sonority

Amplification is the core electronic process that provides the technological means by which sonic transformations can be projected. With the convergence of extended flute techniques with amplification, the sonic landscape of flute music has been transformed, the inaudible has become audible, the interior has become exterior, and performer explorations of identity and expression aroused.

Breath tone begins with an unfocussed flute sound, which develops a character and strength of its own, an individual set of colours that can be manipulated acoustically and electronically to produce intense musical effects. Varied embouchure shape, internal mouth and throat shape, air stream direction, air speed, tongue positions, articulations, differing focus, use of vibrato, pitching, closed or open embouchure and body positions are used by the flautist to create the effects. These air sounds can be expanded into enormous wind noise, melded with

normal flute tone, voice, percussive techniques, greater or lesser amounts of hissing, or create tentative wisps or dots of micro sound. Percussive sounds on flute are usually key slaps or tongue clicks, and they are effectively used in combination with breath tones. A most astonishing soundscape results from extensive use of amplified percussive sounds, as seen in the third movement of Marco Stroppa's *little i* for flutes and electronics: *Moderato Percusive - Molto Minuzioso*. (Figure 1)



Figure 1, percussive articulations and breath tones in Stroppa's *little i*, third movement, bars 20 – 26.

In this music, various combinations of pizzicato, flutter tonguing, tongue ram, key slaps and breath tone are sent through speakers away from the player, generating an illusion of invisible presences and a sense of mystery and uncertainty. Diffusion techniques such as these, manipulations of timbre and sound placement, capturing and conversion techniques, expansion and contraction of the sound, all transform the breath tone into a viable and highly expressive musical device.

Hyper-magnification also generates an increased capacity to draw the listener into the intimate sound world of the performer. These expansions and contractions invoke changed expectations, impose new meanings and alter perceptions of the listener and performer. The breath tone sound can be stunningly large, creating a broad spatial implication, or coloured by minute tongue or percussion sounds, expanded out into the room, creating an impression of intimacy. Added vocalizations can add the sense of other presences, and the electronic processes such as delay or harmonization collect layers or mirroring reverberations to increase density or texture. These sounds can traverse across the room, or dart from speaker to speaker, create an immersion, or disconnection, or further illusions of intimacy or distance through panning or delay. The proximity of the breath tone can give an illusion to the audience of moving into the performer's sound space, an access to a private space, which can approach the confronting. The indistinct, grainy sound of breath tones can also produce a sense of uncertainty, of muffled undercurrent in the music, and can create a breathless, tense and dissonant sensation in the listener.

Working up the technique for breath tones is a remarkable journey, full of self-scrutinization and surprise. The very intimacy of the use of breath as a direct sound source places this technique in a different plane from normal resonant tone. Hearing this sound projected out into the room, shaping it and colouring it through minute mouth and blowing movements encourages much play and experimentation, generating a different physical

approach, a re-positioning of the soloist in the space and new sense of identity.

Identity

The performer's identity is a combination of multiple interior and exterior factors. The outward identity (style, sound, interface) and inward identity (self awareness, interpretative imaginings) blend to form the performing persona. This persona is a projection of the self, intermingled with compositions and artistic goals. The basis of the instrumentalist's life is the commitment to music as a core activity. One's identity is strongly bound up with being a particular type of musician, and making specific musical sounds. As an extension of the self, the instrument acts as an amplifier. Inward and outward identity is displayed through performance, revealing aspects of inner processes and thoughts, responses and emotional connections, through interpretational signals. This self-projection creates a perception in both player and listener, imprinting both visual and aural impressions on a listener, and a sense of positioning and potential to the player.

A new sense of oneness in the sonic world can occur through electronic processes. These include the melding of techniques to give enhanced freedom of expression, sonic empowerment and flexibility. There is an expanded sense of the energy source – the gestures, breath, interpretations; and increased sense of sonic choice. There is a new sense of ensemble, and (at times) a spatialization of solo playing into layers and illusions of other presences. There is a magnification of the micro, intimate elements of playing, and the possibility to mask and confuse.

Timbre manipulations influence sonic identity, as new sound objects become familiar in one's sound world. Through memory, sounds become recognizable, "admissible" (Jonathan Harvey, 1999) and hence usable – and audiences also become significantly more comfortable with sounds they know or can relate to the known. The new sound worlds become elements you hold as part of yourself, part of your own "vocabulary", a representation of personality in sound. The map below (Figure 2) shows exploratory pathways as they pertain to breath tones in flute playing. The musical example included in the map is from Stroppa's *little i*, first movement, indicating a breath tone multiphonic changing to normal tone. Techniques are defined in the upper section of the map, and effects in the lower.

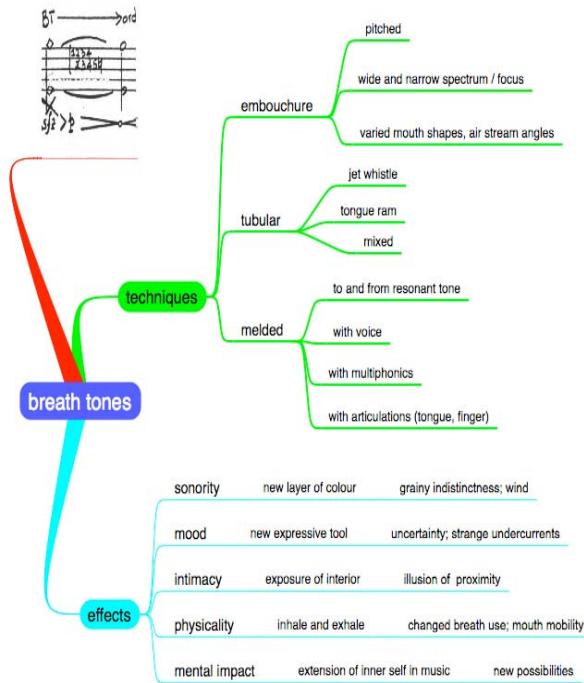


Figure 2, breath tone explorations

The following maps (Figures 3 and 4) indicate interior and exterior identity elements I have found in my own work, a representation of explorations and the pathways to new self-perception and awareness, projection and capabilities invoked by electronics.

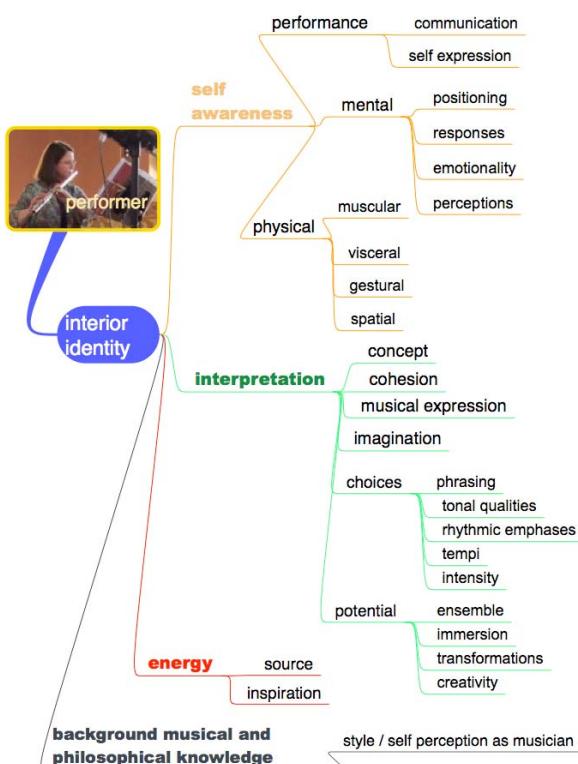


Figure 3, Performer identity explorations: interior

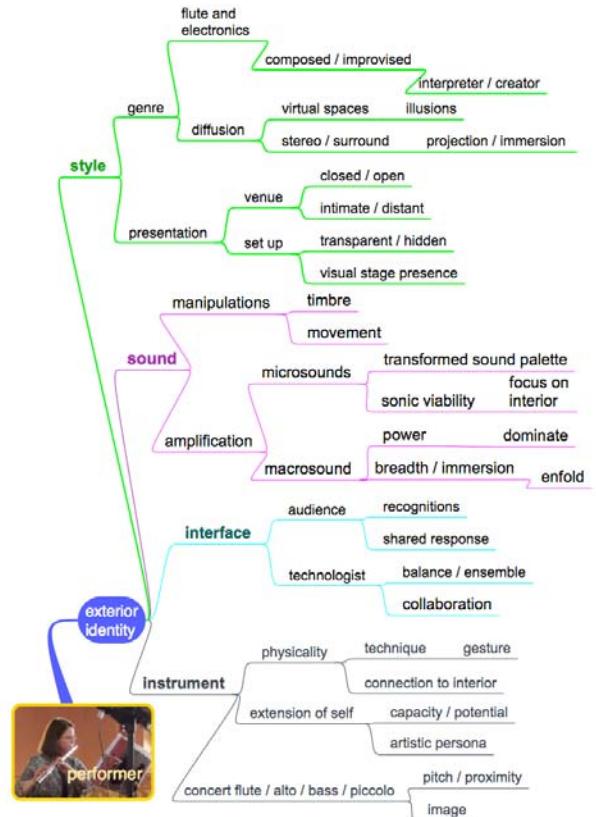


Figure 4, Performer identity explorations: exterior

Spaces: performer's habitat

The play-as-text can be performed in a space, but the play-as-event belongs to the space, and makes the space perform as much as the actors perform. (Wiles, 2003 p.1)

The transformation of space in a performance for flute and electronics occurs with the physical set up and visible dynamics, the real space, and the virtual space of sonic illusion and invisible interconnections and processes. I alluded earlier to the sense of proximity created by hyper-magnification of intimate breath and tongue sounds. This expansion of internal bodily actions out into the room has a two-fold impact. The performer experiences an exposure of self, an open sharing of the internal in pursuit of sonic and communicative pathways. The self at the same time is empowered, enlarged, and given a new context for existence. The listener, on the other hand, experiences the extreme "close up", illusions and sensations of proximity and distance, unexpected sound movement and virtual dialogues with invisible sources.

The creation of a performance space is influenced by many elements of historical context or tradition and physical practicalities. Emergent modes of spatial representations, produced via technological, architectural and performative means have developed. The spaces of music (hall, church, outdoors), the forms of presenta-

tion (formal, intimate, invisible), the position of the audience (opposite, or around) all impact on the performance. The natural acoustic within the playing space has a major role in projection and synthesis, influencing listeners' and performers' perception of the music.

The influence of the sound in the space on the self in the space manifests in the way the player tunes in and 'plays' the space. Performing in a virtual space is entirely different to performing in an acoustic space. Similarities exist between surround sound and extremely reverberant cathedrals, for example, but the virtual spaces created through electronics do not necessarily follow natural rules of sonic behavior. The virtual space becomes a player in the performance, controlled in the main by the technologist. It is an element of expression and communication, and this changes the relationship of performer and performance. Sound sources can become entirely mixed up through live and pre-recorded sound sources, creating ambiguity, the sense of invisible presence and unexpected interconnections. The player can treat the space as a collaborator or an opposition. A sense of enclosure, or disorientation, uncertain proximity and passing of time can result, (for example, in Jean Claude Risset's *Passages*), or a dialogue with spatialized voices (as in Stroppa's *little i*). The performer can be magnified or obscured, empowered or reduced.

The dominant partnership creating the interconnections in this space is the instrumentalist – technologist relationship. This is an incontrovertibly vital, central element of electro-acoustic performance, but one that is rarely discussed from a performative point of view. The balance, the working methodology, the construction of a functional zone that activates and recreates the habitat for the piece are central elements to any performance. Putting together the hardware and the software, working together for confluence and functionality, frequently dominate preparation – the mechanics often demanding energy and attention throughout the rehearsal period. Re-interpretation of composed music implies a search for the composer's ideal, a pursuit of perfection. Premier performances are very likely experimental, an exploration of presentation ideas and a step towards this state of perfection. The relationship of instrumentalist and technologist is consequently one built on great trust, acceptance and a willingness to tread risky paths together. Most importantly, it also thrives when a commonality of artistic and aesthetic goals exist, when the myriad shades of balance, tonal softening or distortion, response and imagination create together a truly interactive gestural environment and ensemble.

Physicality and Gesture

Any sense of causality with visible instrumental gesture is secondary.....There is a strong tendency for the instrument to reestablish its presence separated from the electroacoustic soundworld. This is sometimes through body gestures....breath contour....but

also...idiosyncrasies of amplification. (Emmerson (2007) p. 105)

Kinaesthetic awareness is heightened in performance with electronics through the intensification of focus on micro techniques, and the sense of physical freedom that comes from amplification. Flute playing is a mix of body and sound awareness. The emergence of sonic identity through gesture and shaping of musical material is founded on physical actions (posture, breathing, throat and mouth shaping, tonguing, arm and hand movement, emotionally reflective body language) learnt over many years of practice. Electronic processes impact very powerfully, highlighting variations of tone and intensity, and techniques that demand refined embouchure adjustments, new finger movements and microphone techniques. Amplification liberates the flautist's sense of energy: the sensation of projection is overwhelmingly overtaken by electronics, the audibility is no longer a matter for concern. The impact on the player is an increased relaxation of musculature, as the sound dimension is taken over, and subtle alterations can be made through minute movements. The focus is on this micro-sound world, on the link between physicality and sound texture and timbre, the projection of meaning in the interaction of breath, instrument and electronics.

The electronic separation of the sound from the body in the room gives the sound image an enlarged capacity, and invokes new listening and playing modes in the performer. Breath sounds, key movements and tonguing, are magnified and removed from the source, heard from a distance, giving these sonic units an entity and presence separate from the physical domain of the body. The breath tone is a fragile gesture, of blurred, diffuse and indefinite character. It is the link between interior and exterior elements, and represents a direct extension and magnification of the body inciting heightened perceptual awareness, focus on micro elements, and exploration of physical and expressive meanings and contexts. The immateriality of much of this sound world can also create a sense of the other, of dreams, and the separation of the sound from the body via electronics adds to this effect.

Articulation of breath tones employs physical gestures that take on meaning from electronic manipulations, musical context and perceptions of player and listener. For example, an explosive attack, or a soft pizzicato tongue sound will transmit opposite sensations; the emergence of breath tone from normal flute tone, or vice versa, will move the musical gesture towards a grainy, uncertain mood; the use of voice and breath tone will create impressions of multiple characters or shadows; percussive key slaps will add startling thwacks or pattering patterns; pitch movements will create mini skiffs of air, or sweeping gestures from one end of the flute to the other; jet whistles engage a more forceful muscular movement, and create in-

tensity and extroversion in the sound; tongue rams create a thud effect below the usual range of the flute. Ambiguity reigns in this environment, and a cognitive dissonance with reality can occur through the removal of visual gestural cues, the blurring of sound sources, and conjunction of virtual and real elements.

Interactivity

The first major work for flute and interactive electronics was Phillippe Manoury's *Jupiter* of 1987. This work developed and established score following with Max technologies in the field. Many works have followed, employing such triggering mechanisms as MIDI pedals, pitch sensors, movement sensors, and buttons with cables attached to the flute. The success of these processes can vary, depending at times on the characteristics of the performance venue. For example, Russell Pinkston's *Lizamander* for flute and Max/MSP (2003) uses pitch tracking. Two microphones are used, one for tracking and one for amplification and audio processing. The computer program detects specific pitches played by the flute at precise moments in the score, at times searching for notes above or below a particular threshold, in other instances, an exact pitch. This identification allows movement to the next section, triggering such actions as capturing and granulating sounds, harmonization, or the playing of computer-generated material, but does not always work effectively. If the hall is overly resonant, for example, the computer will have difficulty detecting pitches, and the work can only move forward with manual activation by the technologist.

An exploration of the use of breath as sonic material, and gesture as interactive trigger, introduces Warren Burt's new work, *Mantrae* (2007) for flute and live interactive electronics. Sharply articulated breath and flute tone provide the sonic base for this work, which captures the flautist's movements with a camera, and uses these to activate radical sound effects. The flautist moves between the three *Mantrae* at will, and the three music stands also, whilst maintaining a chant or speech-like style. The flute sound is played through a computer using Cycling '74's Hipno sound processing modules, and a random mixing routine is set up in Plogue Bidule, the host program for the processing. The soundscape is constantly changing, and the flautist is freely changing position and Mantrae whilst maintaining a sense of focus and inner calm. The minimalist style and the conjunction of spiritual expression of the *Mantrae* further evoke ambiguity – a meditative calm in the midst of dramatic sound treatments, representing a constantly changing and unpredictable world. Kinaesthetic awareness is acute in the performance of this work, the broad gestural movements controlling sound events, serving as a challenge to the flautist to adopt a resolute, intense focus on the distilled yet powerful contour of the *Mantrae*, surrounded by complexity and multifariousness.

Conclusion

Through explorations of the interconnections of flute performance and technology, renewed sonic imaginings and responses build significant artistic discourse and illumination. Highlighting the multiple features of flute performance with technology, the transformation of the inaudible to the audible, the new approaches and perceptions of self and expression, new elements and new identities, new relationships and reinventions of space and physical responses, the performer's world attains an expanded meaning and refreshed understanding. The instrument and performer become one, the microphone nourishes this relationship, and new sounds and identities are built.

References

- Artaud, P-Y. 1995. *Present Day Flutes*. Billaudot, Paris.
- Bartoluzzi, B. 1982. *New Sounds for Woodwind*. Second edition. Oxford University Press, London.
- Burt, Warren. 2007. *Mantrae* for flute and live interactive electronics. Self published. (Commissioned by Jean Penny with assistance from the Australia Council for the Arts)
- Dick, R. 1989. *The Other Flute*. Second edition. Multiple Breath Music Company, Missouri, USA.
- Emmerson,S. 2007. *Living Electronic Music*. Ashgate Publishing Limited, Aldershot, UK.
- Harvey, J. 1999. "The Metaphysics of Live Electronics.' *Contemporary Music Review*, 18:3, 79-82.
- Stroppa, M 1996. *little i* for flutes and electronics. Unpublished edition.
- Wiles, D. 2003. *A Short History of Western Performance Space*. Cambridge University Press, Cambridge.

**Christian Ritz^{1,2}, Greg Schiemer^{1,2}, Ian Burnett^{1,2}, Eva Cheng²,
Damien Lock¹, Terumi Narushima¹, Stephen Ingham¹,
Diana Wood Conroy¹**

¹Sonic Arts Research Network, Faculty of Creative Arts / Faculty of Informatics, University of Wollongong

²Whisper Labs / Telecommunications and Information Technology Research Institute (TITR), Faculty of Informatics, University of Wollongong

{critz, schiemer, ianb, ecc04, dnl463, tn649, singham, diana_wood_conroy }@uow.edu.au

Abstract

This paper reports on the recently completed and significant upgrade of the University of Wollongong's Configurable Hemispheric Environment for Spatialised Sound (CHESS). The CHESS studio, which housed a 16 speaker hemisphere for creating spatial sound, has been converted into an anechoic chamber and a new 3D speaker system has been designed. The recent work is a continuation of a successful cross-disciplinary research activity between the Faculty of Informatics and the Faculty of Creative Arts. Also reported are new research initiatives that will be taking place in the facility.

Introduction

The original Configurable Hemispheric Environment for Spatialised Sound (CHESS) was a 3D sound environment that allowed for playback of up to 16 channels of audio using loudspeakers attached to a hemispheric framework (Schiemer, et al., 2004). It was created in 2001 as a collaborative project between the 3D audio team in Informatics, led by Ian Burnett, and staff in the sound, composition, music and production program from Creative Arts, led by Stephen Ingham.

CHESS has been the principal infrastructure used for research activities in spatial sound (Potard and Ingham, 2004). A photo of the original CHESS system, which includes a 3D loudspeaker array, is shown in Figure 1. Technology resulting from this research was used in an international symposium on sonification (ICAD 2004), and for the production of a 16-channel electroacoustic composition created collaboratively between the Faculties of Informatics and Creative Arts (Schiemer & Potard, 2004).

An Anechoic Configurable Hemispheric Environment for Spatialised Sound



Figure 1. The Original CHESS facility (Schiemer, et al., 2004)

The recently upgraded facility has created an anechoic environment for spatial sound. This will enable a wealth of new research activities previously not possible with the original CHESS facility. This report will outline the design and construction of the chamber, the design of a new 3D loudspeaker array, and summarise existing and future research activities to be supported by the facility.

Creating an Anechoic Environment for CHESS

In 2007, a team from the Faculties of Informatics, Creative Arts, and Science, led by Christian Ritz, was awarded a Research Infrastructure Block Grant (RIBG) to build the University of Wollongong's first anechoic chamber: a new anechoic home for CHESS.



Figure 2. Anechoic CHESS

A traditional anechoic chamber requires sound insulation of all 4 walls, the ceiling, and the floor. Such insulation is generally expensive, and requires installation of large wedge shaped soundproofing. The solution adopted for the new facility at the University of Wollongong is the flat-walled multi-layer technique invented by Jingfeng Xu (Xu, *et al.*, 2005). The approach uses standard building insulation material installed in multiple layers on the walls, ceiling and floor to absorb sound within the room (Xu, *et al.*, 2005).

The anechoic chamber was designed to provide 99% acoustic absorption at any one-third octave band centre frequency between 250 Hz and 10 kHz. A photo of the interior of the chamber is shown in Figure 2.

Compared to traditional wedge-based anechoic chambers, employing the design of (Xu, *et al.*, 2005) resulted in significant economic savings whilst achieving similar acoustic characteristics. Similar in size to the chamber reported in (Xu, *et al.*, 2005), the anechoic facility has approximate dimensions of 4.8m in length, 3.3m in width, and 2.4m in height. An additional advantage when compared to conventional wedge based designs is the reduced sound insulation thicknesses, making the technique suitable for smaller spaces (Xu, *et al.*, 2005).

As well as the installation of sound insulation, a number of other building modifications were required. These included the design and installation of a raised floor, and modification of the air-conditioning system to eliminate noise created in the chamber from air-flow through the ducting and outlets. Whilst a number of alternative solutions were considered, the most cost-effective solution was to install a damper system. This system significantly reduced noise generation and allowed for user control to toggle the air-conditioning on/off when total silence is required (e.g. when conducting sound recording experiments).

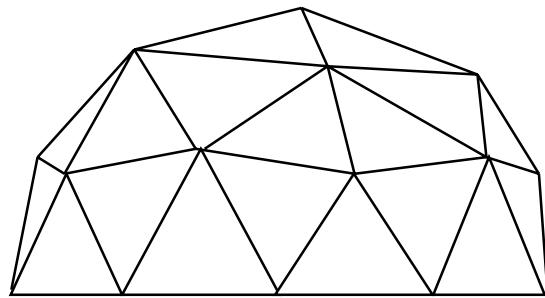


Figure 3. A 2V Geodesic dome (Landry, 2002)

New Geodesic Speaker Array

The original CHESS system consisted of a customised structure formed from six steel frames arranged to form a ‘globe’ (Schiemer, *et al.*, 2004) (see Figure 1). Hinged brackets allowed for attachment of loudspeakers and placement at desired azimuths and elevations. Rebuilding CHESS within the anechoic chamber required a new design due to the reduced space resulting from the sound insulation.

The new design will consist of a geodesic array built from metal tubing. The new array will use lighter (and lower cost) active studio monitors (Genelec 8020A) and will be constructed using lightweight material to ensure that it is more configurable and easier to accurately manipulate speaker positions than the original CHESS structure.

The dome design will be the upper hemisphere of a 2V Geodesic dome as illustrated in Figure 3 (Landry, 2002). The use of this dome will allow placement of loudspeakers at regular spatial positions and is particularly suited to 3D audio reproduction (Vennonen, 1995). The new CHESS speaker array will be formed from a regular decagon shaped base constructed from 10 metal tubes. The remainder of the dome will then be built from metal tubes of two different lengths and customised hinged brackets will be constructed to allow attachment of the loudspeakers at arbitrary positions on the framework.

Audio Recording and Playback System

The complete CHESS system included a user interface hosted on a Pentium 4 PC that communicated via a LAN with the 3D audio Digital Signal Processing (DSP) software and commercial audio hardware hosted on a MAC G3. The 16 channel speaker array was interfaced to the MAC G3 with audio conversion hardware (Alesis AI-3) via a Digidesign Digi-001 controlled by Pro-Tools LE. Custom 3D audio software was also developed for the original CHESS system using Max/MSP on the G3. Spatial audio reproduction was achieved using the Ambisonics approach (Malham, 2005), where the appropriate loud-

speaker signals were created using custom software implementing 4th order Ambisonics encoding (Schiemer *et al.*, 2004).

For the new facility, experiments will be conducted using both the existing control system as well as an alternative system: the MAC G3 is replaced by a dual-core Pentium D PC interfacing through a RME HDSP9652 Hammerfall DSP system to three Behringer ADA8000 pre-amps, controlled by Adobe Audition (Windows) or Ardour (Linux). The alternative system provides the advantage of increased flexibility for simultaneous multichannel recording and playback through the same system, which has been used in recent research published by the authors (Cheng, *et al.*, 2006).

Key Research Projects to be Supported by the New Facility

Location-Based Audio Object Annotation

tion

The new anechoic chamber will allow for new experiments into the recording and analysis of spatial sound. Authors Ritz, Burnett and Cheng have researched computationally efficient algorithms that provide event-based metadata extraction from spatial positioning information derived from multichannel speech recordings in real acoustic environments, where problems caused by room reverberation and background noise are generally hard to predict and account for (Cheng, *et al.*, 2006). The new anechoic environment will allow for controlled experiments into the recording and annotation of spatial audio objects which are not affected by room reflections and noise signals. This will enable new scientific experiments comparing results between anechoic and reverberant recordings.

New Techniques in Recording and

Playing 3D Audio

The research, design, and development of new microphone array techniques for recording 3D (spatial) sound require an anechoic environment for accurate characterisation of the array acoustic response and performance evaluations. The authors are currently developing new techniques for the recording of spatial sound using non-obtrusive array designs, distributed arrays, and arrays of directional vector sensors. The anechoic facility will also be used to study frequency response in thin film speaker material used in the Orbophone (Lock and Schiemer, 2006), a new interface for sound radiation.

Spatial (3D) Audio Coding

Authors Ritz and Burnett and their PhD student Bin Cheng have developed a new spatial audio

coding technique called Spatially Squeezed Surround Audio Coding (S3AC) (Cheng, *et al.*, 2007; Cheng, *et al.*, 2008). This technique has been applied to the compression of multi-channel surround audio as well as Ambisonic B-format recordings, with results from perceptual listening experiments conducted in real environments having shown significant advantages over existing approaches (Cheng, *et al.* 2007; Cheng, *et al.* 2008). The anechoic chamber and new 3D speaker array will be used for new experiments investigating the human perception of spatial sound without interference from unwanted reflected signals. The results of these experiments will be used for developing new coding techniques that further reduce the bit rate for compressing 3D audio encoded in various signal formats.

Auditory Perception of Microtonal

Music

An anechoic facility will be useful for conducting perceptual studies involving auditory perception of microtonal music (Narushima, *et al.*, 2008) by composers, performers and untrained listeners as part of a new ARC Discovery Project ARC. Two international collaborators taking part in this project will investigate harmonic structures and psychological processes involved in microtonal composition. The Sonic Arts Research Network, a research collective formed in the University of Wollongong to foster interdisciplinary research between the Faculty of Informatics and the Faculty of Creative Arts, provided funding to support these joint initiatives.

Acoustics of Ancient Architecture

The way acoustic sound is enhanced through specific material surfaces and textures is of great interest in the understanding of performance in ancient architecture, such as the Hellenistic theatre in Paphos, Cyprus, where members of the faculty have worked. The Sonic Architectures exhibition in October 2006 in the Faculty of Creative Arts Gallery was a start in exploring these issues (Wood Conroy, *et al.*, 2006). The anechoic chamber will allow the effects of different material surfaces on the acoustic properties of sound to be studied in a controlled environment. It will give researchers new insights on sound projection based on understanding how such a theatre might have sounded to audience members in ancient times.

Conclusion

The University of Wollongong's CHESS facility has been a successful collaboration between two faculties that has led to a number of research outcomes. The upgrading of CHESS to provide a new anechoic environment has produced a unique facility that will ensure many new re-

search and creative projects involving sound in all dimensions.

Acknowledgements

The upgrade of the CHESS facility was supported by a University of Wollongong 2007 Research Infrastructure Block Grant (RIBG) Pool 2 as well as funding from the University of Wollongong's Telecommunications and Information Technology Research Institute (TITR) and the University of Wollongong's Sonic Arts Research Network (SARN). We would also like to thank Fred Zylstra from the University of Wollongong, who was the building project manager for the new facility, Colin Schoknecht and colleagues from the University of Western Sydney who provided advice on certain aspects of the chamber construction and Jingfeng Xu, who was the acoustic consultant providing the sound insulation specifications.

References

- Cheng, E., Burnett, I., Ritz, C. 2006. "Using Spatial Audio Cues from Speech Excitation for Meeting Speech Segmentation", in *Proc. 8th International Conference on Signal Processing (ICSP '06)*, 16-20 Nov. 2006, Guilin, China.
- Cheng, E., Burnett, I., Ritz, C. 2006. "Varying Microphone Patterns for Meeting Speech Segmentation using Spatial Audio Cues", *Advances in Multimedia Information Processing – PCM2006, Proceedings of the 7th Pacific-Rim Conference on Multimedia (PCM2006)*, Hangzhou, China, Nov. 2-4, 2006, *Lecture Notes in Computer Science* 4261, pp. 221 - 228, Springer-Verlag, 2006.
- Cheng, B., Ritz, C., Burnett, I. 2007. "Principles and Analysis of the Squeezing Approach to Low Bit Rate Spatial Audio Coding", *Proc. IEEE 2007 International Conference on Acoustics, Speech and Signal Processing (ICASSP'2007)*, Honolulu, USA, 15-20 April 2007.
- Cheng, B., Ritz, C., Burnett, I. 2008. "A spatial squeezing approach to Ambisonic Audio Compression", to be presented at *IEEE 2008 International Conference on Acoustics, Speech and Signal Processing (ICASSP'2008)*, Las Vegas, USA, 30 Mar-4 April, 2008.
- Landry, T. 2002. Desert Domes, [online], 2002. Available : www.desertdomes.com/index.html
- Lock, D. Ong, L. and Schiemer, G. 2006. "Orbophone: a new interface for radiating sound and image" *Proceedings of the 12th International Conference on Auditory Display*, London, UK, June 20-23, 2006 pp. 206-209.
- Malham, D. G. (ed.). 2005. "Sound in Space", University of York Music Technology Group, 2005. www.york.ac.uk/inst/mustech/3d_audio/
- Narushima, T., Schiemer, G., Schubert, E. and Parncutt, R. 2008. "Recognition of microtonal musical intervals by performers and composers" submitted to *The 10th International Conference on Music Perception and Cognition (ICMPC10) & The 3rd Triennial Conference on Asia-Pacific Society for the Cognitive Sciences of Music (APSCOM3)* 25th-29th August 2008, Sapporo, Japan.
- Potard, G and Ingham, S. 2003. "Encoding 3D sound scenes and music in XML", *Proc. of 2003 International Computer Music Conference*, ISBN 0-9713192-1-9 pp. 65-74, National University of Singapore, October 2003.
- Schiemer, G. M., Ingham, S. F., Balez, D. V., Scott, J. P., Lock, D., Hull, A., Jenkins, G. & Burnett, I. S. 2004, "Configurable Hemisphere Environment for Spatialised Sound", *Ghost in the Machine: Australasian Computer Music Conference 2004. Proceedings* (pp. 53-56). Victoria, Australia: Australasian Computer Music Association.
- Schiemer, G and Potard, G 2004, "Sonification of the Coherence Matrix and Power Spectrum of EEG Signals: Listening to the Mind Listening", *Proc. International Conference of Auditory Display*, July 6-8 2004, Sydney, Australia.
- Schiemer, G and Potard, G 2004, "Neural Dialogues: an electro-acoustic composition for 16-channel space array", *Listening to the Mind Listening*, curated by S. Barrass, The Studio, Sydney Opera House, July 8, in *Proc. International Conference of Auditory Display*, July 6-8 2004, Sydney, Australia
- Vennonen, K. 1995. "A practical system for three-dimensional sound projection", [online], Copyright Kimmo Vennonen, 1995, Available: www.kvp.net.au/3S.94.abstract.html
- Wood Conroy, D., Bunt, B. S., Epoff, D. & Ingham, S. F. 2006. *Sonic Architecture: Mapping the ancient theatre through sound and image*, May 10, part of *Illawarra Innovation Showcase & Sonic Connections*, May 9 - 11, 2006, SARN, University of Wollongong.
- Xu, J., Bucholz, J.M., Fricke, F.R. 2005. "Flat-walled multilayered anechoic linings: Optimization and application", *J. Acoust. Soc. Am.*, pp. 3104-3109, 118(5), November 2005.

Graeme Smith

Andrew Johnston

Creativity and Cognition Studios,
University of Technology, Sydney
PO Box 123
Broadway, NSW, 2007
Australia
grasmith@gmail.com
aj@it.uts.edu.au

Abstract

In this paper we present software designed to help address problems encountered by beginning guitarists, using interactive software to find effective solutions to enhance the learning process. Software can be utilised to improve a player's ability to hear mistakes in their performance, as well as to create a fun and entertaining learning environment to motivate the player to practice. A software prototype has been developed, which served as a basis for usability testing, to highlight the usefulness of various methods of feedback and provide a way forward in developing valuable software for guitar tuition.

Introduction

In the early stages of learning a musical instrument, such as guitar, a player may lack awareness of mistakes in their performance, due to the high level of concentration required to play the instrument. The player may be making mistakes without realising, resulting in the continued practice of incorrect technique. This can slow the learning process, and may discourage the player if they do not improve and do not understand why.

Listening skills are important to develop, which allow a player to hear mistakes whilst performing. Playing along with a song recording is a common approach to practising, however this can wash out the sound of the player's performance, with mistakes going unnoticed. Playing without a recording can sound bare, and it can be difficult to maintain a good sense of timing.

This project aims to provide insight into ways of overcoming the difficulties encountered when trying to listen to a performance and play at the same time. An interactive software system has been developed and evaluated, aiming to help develop listening skills and make practicing a more enjoyable experience, by providing intuitive feedback to players.

Instrumental Music Education Software

For decades research has been conducted to find effective ways in which software can assist in the learning of musical instruments. Percival, Wang

Interactive Software for Guitar Learning

and Tzanetakis [6] have surveyed recent work in computer-assisted musical education, and explain that a complex system for music education should have clearly defined goals, the most important being:

1. Enhancing the lessons with a teacher
2. Enhancing the player's practice
3. Motivating the player

The choice of these goals dictates the purpose and intended users of the software, and presents several important factors that need to be considered, including whether to give feedback during or after the performance, whether to present the information qualitatively or quantitatively, and how to motivate a player.

Feedback During or After a Performance

ance

Giving real-time feedback during a performance allows a player to realise and correct errors whilst playing, whereas giving feedback after a performance allows for critical sections of the performance to be highlighted and examined in more detail than is possible whilst playing.

Iwami and Miura [3] state that a player can be more aware of their weak points and tendencies if they recognise mistakes whilst performing. It is important to give careful consideration to the player's cognitive load when providing feedback during a performance. Playing an instrument can require a high level of concentration, particularly for beginners. As such, real-time feedback needs to be simple and easy to understand. It should not distract the student from listening to their performance [6] and forming mental models of their desired performance. Music-oriented video games, such as SingStar (<http://www.singstargame.com>), Guitar Hero (<http://www.guitarhero.com>) and Rock Band (<http://www.rockband.com>), accommodate a wide range of users, with differing musical ability, and provide very simple and intuitive real-time feedback.

On the other hand, Percival, Wang and Tzanetakis [6] argue that software interaction should occur after a performance, as means of confirming and correcting a student's judgement. A player can listen to their performance whilst playing, and then see if the software agrees with their understanding of the performance. A re-

view can provide a complete view of a performance, which can be useful in identifying mistakes and problem sections, as well as evaluating the accuracy of specific performance characteristics. However, a performance review should not be so time consuming that it detracts from actually playing the instrument.

Both approaches provide significant benefits if used appropriately, and we found a combination of the two was most useful. Careful consideration needs to be given to ensure that a player does not become dependant on the feedback. The focus must remain on teaching the player to listen to their performance, so that when the software is taken away they will be able to identify their own mistakes.

Quantitative and Qualitative Feedback

While traditional feedback provided from a music teacher during a lesson is qualitative in nature, data captured via computer interfaces enables a detailed quantitative look at a player's performance.

Qualitative feedback can be used to provide advice on mistakes in a performance, or to simply give the player an overall impression of how well they are playing. In either case, it generally results in a more human-like style of feedback. This requires a level of 'intelligence' in the software to interpret the quantitative data, as seen in the Piano Tutor [1]. While software can be used to simulate human feedback, it is important not to forget about the important role that human teachers play. A human teacher can make intuitive decisions on a student's weak points, and read into the subtle psychological aspects of learning, such as providing encouragement to a disheartened student.

Quantitative feedback aims to take a more objective look at a performance, and is useful for measuring the accuracy of a player's performance. Feedback may be presented in the form of a graph, showing both the expected and actual performances. Interpreting graphs may require some musical knowledge, to identify sections that were within an acceptable accuracy range and sections that were incorrect. Therefore, this style of feedback is more suited to assisting a teacher in a lesson, and not as relevant to a student practising alone. Alternatively, quantitative feedback can be given in the form of an accuracy score, which can be useful for motivation and self-improvement. The software must be robust and accurate, so as not to disadvantage, confuse or aggravate players. However, with an art form like music, there is no strict right or wrong, and a mathematical dissection of someone's performance is sometimes not helpful. A computer may not know what is actually right or wrong, and may penalise creative expression.

Motivation

Percival, Wang and Tzanetakis [6] comment that motivation is the single most useful factor when

using software for music education. They claim that the benefits of keeping a player motivated and interested in learning outweigh the benefits of effective multimedia feedback. Motivation can be achieved by instilling positive feelings in a player, such as a sense of achievement after skill improvement, and by providing a player with a fun and entertaining environment with which to perform.

Positive feedback is an important aspect of the learning process, providing a player with encouragement to stay motivated. Positive feedback is an aspect of music education software that is often overlooked, with many systems focussing solely on identifying mistakes in a performance. Whilst the mistakes identified may be accurate and relevant, it may not be very motivating for a player.

Juul [4] notes that learning to play an instrument, such as guitar or piano, can be very tedious, whereas instrument simulation games, such as Guitar Hero and Rock Band, empower the player with the feeling that they are superbly skilled. Williams [4] adds that part of the appeal of instrument simulation games comes from the compressed learning curve, which gives the player the satisfaction of having learnt a skill.

If a player has fun playing their instrument, they will be keen to practice. An effective way to achieve an entertaining performance environment is to provide a social element. Music is a form of creative expression, and should be shared with other people. A player may lack motivation to perform music without anybody listening. The Family Ensemble [5] system aims to make practising a shared experience, to increase motivation and encourage a deeper appreciation of music. The StarPlay (www.starplaymusic.com) system instils motivation by simulating that the player is part of an orchestra. Facilitating competition, among peers or against oneself, can be another powerful motivating factor.

Designing Software for Guitar Tuition

This section outlines the approach taken to design software for guitar tuition, with the aim of finding effective techniques to:

- Motivate a player to practice
- Develop a player's listening skills

A software prototype has been developed, which has served as a basis for usability testing, to provide insight into methods of effective feedback, as well as key areas for further development in interactive guitar tuition. The prototype is designed to be a framework for a more advanced and feature-rich system.

Usability testing was conducted in two stages: during design and development of the prototype, and after the final iteration of the prototype. Tests were conducted with beginner and intermediate players in the first stage of usability testing, to shape the design towards the needs of the target users. The second stage of

testing involved professional players and music teachers, to gain ideas for improvement. Seven participants were tested, with a diverse range of skills in guitar, music and computing. The testing sessions were filmed, to enable careful review of their experiences with the software and suggestions.

Prototype Design

This section details the final iteration of the prototype design, including some outcomes from usability testing which impacted the design.

Features

The prototype was designed for use in a typical practice session. The user selects a song to play, along with some settings, including what speed to play the song at, whether to play the whole song or just part of it and whether to loop the song. Using code from the open source Tux-Guitar Java program (<http://www.tuxguitar.com.ar/>), Guitar Pro (<http://www.guitar-pro.com>) and Power Tab (<http://www.power-tab.net>) song files may be loaded by the software.

The software is divided into two main stages: the performance and the performance review. During the performance, the player is accompanied by a MIDI version of the song, reading from the scrolling guitar tablature and receiving simple accuracy feedback (Figure). After the performance, the user is given an accuracy score and is able to review their performance accuracy (Figures 2 and 3). They may listen to a recording of their performance, which may be played along with the MIDI song tracks.

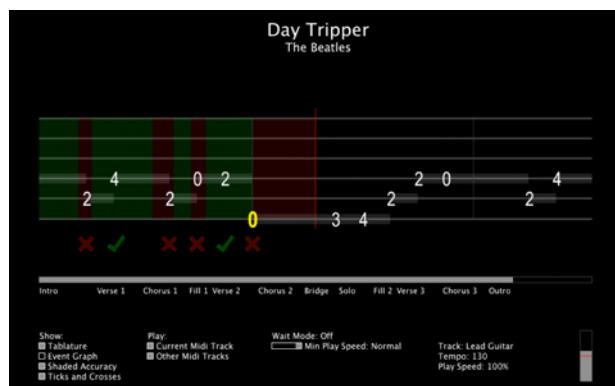


Figure 1. Performance stage, with real-time feedback given on the scrolling tablature.

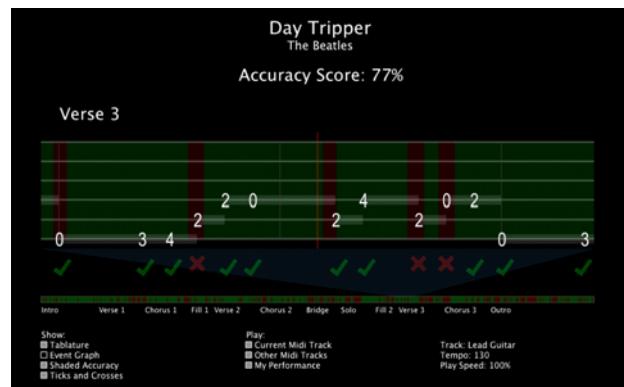


Figure 2. Performance review, feedback given on the tablature.

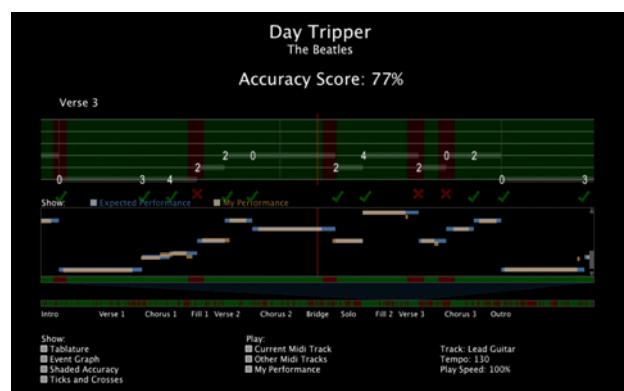


Figure 3. Performance review, showing both the tablature and piano roll displays.

One of the key considerations in designing the software was for it to suit the needs and expectations of a variety of players with differing skill levels. As such, the user is given control over what is heard and what is displayed. Through use of checkboxes (at the bottom of the user interface), the user may toggle between the audio tracks played, as well as the notation and accuracy feedback style displayed.

MIDI Playback

Using TuxGuitar code, a MIDI song is generated from the loaded song file. The player has the option of hearing or muting the current track they are performing and the backing tracks. It was found that hearing the current MIDI track whilst performing can be useful to hear differences between what is being played and what should have been played. On the other hand, muting the current MIDI track allows a player to hear their own performance more clearly. To enhance the experience, one user suggested acquiring the actual song recordings, however there are licensing issues and may be difficulties synchronising the song recording with the song tablature file.

Notation Display

Using an approach similar to that used in the Digital Violin Tutor [8], simple notation styles were chosen, to assist in the rapid development of skills. The user may toggle between two display formats: guitar tablature and piano roll notation. Tablature presents a simple and intuitive notation that can be read whilst playing, whereas piano roll offers more detail and is useful when reviewing a performance. Both displays can be viewed together, as seen in Figure, allowing the player to experience the benefits of both notation styles simultaneously. We recognise that musical score would be a useful addition, as noted by several music teachers, and this is an area for future development.

Tablature

Guitar tablature is a style of guitar notation often used by beginner and intermediate players, due to its simplicity and resemblance to the physical appearance of a guitar. It consists of six lines, which represent the strings of the guitar. Fret numbers are placed on the lines to represent notes to be played.

An important consideration is the notation of time in an animated notation display. There are two key approaches to this: karaoke style and scrolling displays. The karaoke style involves using a bouncing ball or changing the colour of the notes or lyrics, as seen in SingStar. On the other hand, a scrolling display, implemented in both Guitar Hero and Rock Band, uses a time bar to indicate when to perform each note. A scrolling display allows a player to maintain a good sense of timing, as it moves at a consistent rate. For this reason we have implemented a scrolling tablature. The tablature scrolls horizontally across the screen, and a vertical time bar indicates when each note is to be played. Notes are highlighted when they need to be played, by increasing the size of the font and changing the colour.

All usability testing participants found the scrolling tablature easy to follow in terms of note onset timing, and the accuracy of striking notes at the correct time was observed to be quite good. However, difficulties were encountered with recognising which string to play, and the strings have been made thicker and further apart to overcome this.

One of the disadvantages of tablature is that it has no rhythmic notation, which makes accurate sight-reading near impossible. For this reason, guitar tablature is often used when the rhythm of a melody is already known, or is accompanied with a musical score. An attempt to overcome this was made in the prototype by displaying the duration of each note as a semi-transparent bar, as shown in Figure. However, the usability testing participants did not find this entirely useful, and this is an area for further de-

velopment. Several music teachers suggested a hybrid notation, using music score notation to indicate rhythm on a guitar tablature.

Several music teachers noted that it would be good to see notation for fingering on the tablature. We experimented with displaying the finger number in small text above or next to each fret number on the tablature. This looked acceptable for simple riffs, however, the tablature became cluttered for music containing chords, significantly decreasing its readability. Notation aside, the player still needs to translate the finger notation into how to position their hand, which would be a difficult task whilst performing a song. It is likely the player would make a mistake the first time, then look at the display to understand the fingering, and then attempt the section again. As a result, virtual demonstration, such as video and 3D computer animation, may be a more effective way of illustrating fingering, as it simulates a teacher's demonstration in a lesson.

To help establish a more inviting and comfortable environment, one teacher suggested making the tablature look more like a guitar, to make it feel more like music and less like a graph. A picture of the guitar headstock could be added to the left edge of the tablature, and the tablature could be made to look more like a fretboard.

Piano Roll

Piano roll is a bar graph style notation, based on the physical appearance of a piano, whereby the vertical axis contains the notes of a piano and the horizontal axis represents time. The piano roll display (Figure) allows the user to compare their performance to the expected performance, in terms of pitch and timing. The visibility of the expected and actual performances can be toggled using checkboxes. The user is able to zoom into the graph using the mouse scroll wheel, and can navigate using the arrow keys.



Figure 4. Piano roll display.

By presenting a detailed view of their performance, a player may be able to understand why sections of their performance were inaccurate; for example, piano roll effectively illustrates the timing accuracy of a performance, allowing for easy recognition of early, late, missed and incomplete notes. While incorrect notes are evi-

dent, it can be difficult to determine the reason for a wrong note, such as fingering the correct fret but the wrong string.

Unfortunately most users found the piano roll to have a very scientific look, which was not well received by beginners. SingStar was noted as having a very effective piano roll notation, as seen in Figure. This presents a more symbolic style of feedback, with the expected note bars being ‘filled up’ when the correct note is performed and incorrect notes being represented with smaller bars. While this offers less precision, it is a more intuitive notation. One teacher commented that this simplistic, non-scientific look makes a player feel more comfortable and not be quite as afraid to make mistakes.



Figure 5. SingStar screenshot [2].

Performance Accuracy Feedback

Feedback given on the accuracy of a performance aims to assist a player in understanding if, how and why their performance differs from the expected performance. This can provide insight into aspects of a performance that a player may have otherwise been unaware of. Ultimately, the feedback given to the player should be designed to enhance their listening skills.

Accuracy feedback is given to the player both during and after their performance. The tablature is intended to be viewed during the performance, and is annotated with accuracy information. The piano roll is more useful after the performance, offering a more detailed look at the performance.

Real-Time Feedback

Whilst performing, the user may receive accuracy feedback in the form of green and red shading on the tablature, and/or ticks and crosses under the tablature. The user is able to toggle the visibility of each feedback style via the checkboxes, allowing them to use the feedback style that suits them best, or even turn off real-time feedback if it is distracting.

The shading on the tablature indicates green for correct sections and red for incorrect sections. This allows the player to recognise if part of a note was performed accurately or not. Beginner users found this style of feedback to be too detailed.

A number of teachers mentioned that a big problem for beginners is not playing each note for its full duration. If a player thinks they have played mostly correct, by playing notes at the right time, they may find it demoralising to see red sections on the tablature and receive a low score. From this, the tick and cross style of feedback was suggested and implemented. This simplifies the feedback to a Boolean response for each note: the note was either played mostly correct or mostly wrong. One teacher noted that this tick and cross style of feedback provided more incentive to play accurately and try to generate a tick under each note. To increase motivation, a brighter tick could be used to indicate that a note was very accurate. The tick could be replaced with icons or animations, such as smiley faces or thumbs up signs, to target specific users, such as children and adults.

Performance Review

A performance review allows the student to receive more detailed feedback than they can interpret whilst playing. After a performance, the player is given the opportunity to review their performance accuracy and hear their performance again. Using performance reviews may slow the learning process, as the player spends less time actually playing the instrument. We have minimised the time needed to review a performance by providing a complete view of the performance accuracy under the tablature or piano roll display (Figures 2 and 3), allowing the player to easily find and skip to critical sections of their performance.

The player is given an accuracy score, which simply indicates the percentage of time they played correctly. While this is not the most ideal method of calculating a player’s performance accuracy, most users found that their score improved the more they played, prompting motivation to improve their score. Providing a more detailed and reliable accuracy score is an area for further development.

Performance Evaluation

Performance evaluation can be divided into two stages. The first stage is transcribing the performance using real-time pitch tracking, to translate the audio signal into a more meaningful representation. The second stage is to compare the transcribed performance to the expected performance, using a matching algorithm, to assess the accuracy of the performance.

Pitch Tracking

The pitch tracking functionality was implemented using Pure Data, an open source graphical programming environment that can be used to process real-time audio. Pure Data was chosen for its speed and robustness with audio signal processing. The fiddle~ object estimates the pitch and amplitude of an incoming audio signal, both

continuously and as a series of MIDI note events [7]. A window size of 2048 was required to detect all of the notes on a guitar with standard tuning. Unfortunately, this reduces the precision of onset detection. The attack detection provided by fiddle~ does not offer high enough precision to be of benefit, and the bonk~ object is suited to detecting percussive attacks, rather than guitar attacks. We implemented attack detection by detecting spikes in the amplitude output from fiddle~ (output approximately every 23 milliseconds). An independent signal power analysis is required to achieve accurate attack detection, which is a possible area for future development, however user testing by experienced players indicated that accuracy of the prototype was adequate.

While monophonic pitch tracking has reached high levels of accuracy, polyphonic pitch tracking is still an area of avid research, with a reliable real-time solution yet to be developed. This project does not aim to make any improvements to polyphonic pitch tracking, but rather utilise existing technologies to design software with the focus of providing effective feedback to a player.

Matching Algorithm

Percival, Wang and Tzanetakis [6] comment that there are no clear-cut ‘correct’ or ‘incorrect’ answers when assessing a musical performance. Decisions need to be made to determine if a note is sufficiently in tune and played at an acceptable time. We have developed a simple matching algorithm, which compares the pitch of the expected and actual performances, to inform the player if they played the right notes, and whether they played them at the right time. The matching algorithm determines the percentage of time that the pitch of the actual performance (rounded to the nearest semitone) matched the expected performance. In Figure, the green sections illustrate a match between the actual and expected performances and the red sections represent the performances not matching. This reflects the shaded accuracy feedback presented on the tablature display.

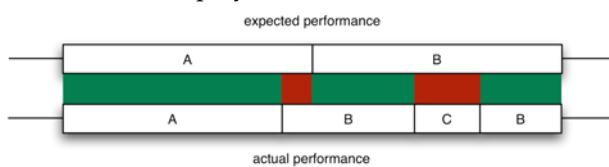


Figure 6. Performance comparison using the matching algorithm.

The ticks and crosses are generated by determining the accuracy of the note onset. If the player strikes the correct note within a short period of time either side of when the note should be struck (we found 50 milliseconds to be fair), a tick is generated. Otherwise, a cross is generated

(Figure). The performance of the rest of the note is not considered, as many beginners feel that as long as they strike the note at the right time they have played correctly.

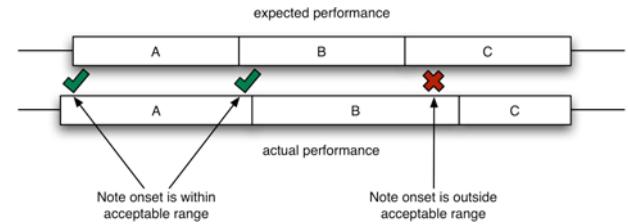


Figure 7. Calculation of ticks and crosses.

Wait Mode

Wait Mode is a mode of play whereby the speed of the song is adjusted to wait for the correct notes to be played. The user can select the minimum speed for the song to slow down to, as well as the normal speed to play the song. The speed of the song slows down when the expected note is not played, and the speed increases up to the normal speed when the correct notes are played (Figure).

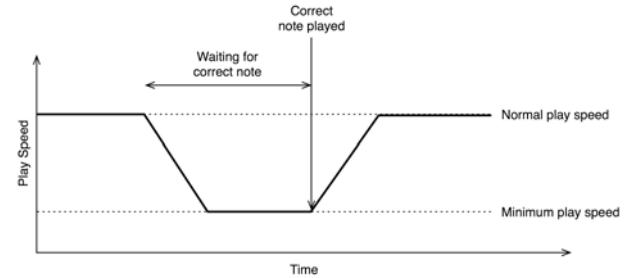


Figure 8. Timeline for Wait Mode.

This mode allows the user to ensure they are playing the correct notes in a performance, something which can get lost when the speed of the song is increased. Beginners often need to look at the guitar whilst they are playing, which makes it difficult to look at the computer screen to see what they need to play. Wait Mode aims to give the player time to look at the screen and not feel as rushed in their performance.

The most notable outcomes of usability testing of Wait Mode were the difficulty to maintain a good sense of timing, and the conflicting opinions that emerged as to when the play speed should start slowing down and whether the song should come to a complete stop.

Minimum Play Speed

Some participants wanted the song to come to a complete stop, whereas others wanted the song to slow down to a specified minimum speed. The main reason for wanting the song to stop is to allow the user to spend time thinking about what they played and how it was wrong, without feeling rushed. If the song is not stopped, the prob-

lem note may be passed over before the player is able to play it correctly.

Slowing Down

There were three ideas for deciding when to start slowing the song down: as soon as a note is missed, half way through a note if it has not been played, or after a certain number of consecutive mistakes. For the first two approaches, it was agreed that the song should reach the minimum play speed by the end of the note (which may involve stopping). By stopping after a certain number of consecutive mistakes, the player is able to keep a better sense of timing, as the play speed is not continually adjusted for each note.

Sense of Timing

Most participants found it hard to maintain a good sense of timing in Wait Mode, as the play speed changed too abruptly. It was generally found that the play speed increased too quickly once a correct note is played. While the song may slow down quickly, the return to the normal play speed needs to be more gradual. If the player has been looking at their guitar to find the correct note to play, they will need time to look back at the computer screen to see what they need to play next.

Implementation

Due to the varied opinions on the behaviour of the play speed during Wait Mode, the user needs to be given appropriate controls to use the software in a manner that suits them. This allows the user to find a balance between maintaining accurate timing and making sure every note is played. Sliders have been implemented on the user interface, to define the minimum play speed, when to start slowing down and how quickly the play speed is changed. For example, the slider for the minimum play speed has one end indicating that the song will come to a complete stop, and the other end indicating that the song will not slow down at all (turning Wait Mode off).

Possible Future Work

Usability testing provided a wealth of ideas for further improvement to harness the capabilities of software to motivate a player and help develop their listening skills. The key areas for improvement are noted as enhancing the practising experience, providing more intelligent feedback and keeping up to date with pitch tracking improvements.

Enhancing Practice

Enhancing the experience of practising is one of the key goals of the software developed in this project. Providing players with reminders, integrating video game elements and virtual demon-

strations have been identified as notable areas for improvement.

Practising Reminders

Ideally, a player should not make mistakes that have been discussed in a previous lesson if they practice often enough, by maintaining a regular practice routine. Unfortunately, many students do not follow this, and may forget important information. The software practising tool could assist with this problem by allowing the teacher to enter reminders to be given to a player whilst practising. The software could even be used as an organiser for students, to plan when they will practice, remind them when they need to practice and log the hours of practice. Some players may find this unnecessary, although the teacher may be interested to keep track of the amount of practice a student has done, which may motivate the player to practice.

Integration of Video Game Elements

Some people will spend hours per day playing video games, trying to advance through the levels of the game and trying beat other players. Music-oriented video games, such as Guitar Hero, SingStar and Rock Band, are no exception. By integrating video game elements into a music practising tool, the player is provided with a more motivating and entertaining environment. One such way to achieve this is to facilitate progressive learning of a song with a video game level structure, back story and characters.

Virtual Demonstration

A player needs to be aware of what they should be doing, rather than just what they did. Virtual demonstrations, such as video and 3D computer animation, could be utilised to illustrate certain performance aspects, such as ideal fingering and strumming patterns. Rather than making a player watch the entire performance, they should be able to select a particular note, chord or small section of the song to see demonstrated. This gives more immediate feedback, allowing the player to quickly understand what they have done wrong and go back to performing. Unfortunately there are several limitations that make effective demonstration difficult. Videos offer no interactivity, in terms of the viewing angle and zoom, and 3D computer animation lacks realism of fine performance details.

More Intelligent Feedback

The ‘intelligence’ of the software can be improved to provide more meaningful and intuitive information to the user, increasing the quality of the interaction. Providing encouragement, detecting repeated mistakes and adjusting the level of detail in feedback have been recognised as the most notable areas for improvement.

Providing Encouragement

Several music teachers stressed the need for more positive feedback, such as encouraging messages after sections of correctly performed notes, or animations around notes played correctly. One teacher commented that the presence of red feedback outweighing green could be quite intimidating to a beginner, especially if they are insecure. This could be overcome by only showing the green accuracy feedback, or only showing the ticks and no crosses.

Detecting Repeated Mistakes

During a lesson, a music teacher will notice if a student is continually making the same mistake over successive performances, and may suggest technical exercises for the student to undertake to correct the mistake. This capability could be added to the software practising tool. While the software could be given a set of rules to rank mistakes, several teachers noted that they would like to be able to configure the priority of mistakes.

Level of Detail in Feedback

To make the feedback more relevant, the level of detail should be adjusted to match the ability level of the player, starting simple and then intuitively building up as the player practises more and develops their skills. There is a logical transition from the tick and cross feedback to the more detailed shading on the tablature. The accuracy feedback would also need to be adjusted to highlight more detailed issues as the player improves.

Pitch Tracking Improvements

Pitch tracking presents a significant limitation to guitar practising tools. Current polyphonic pitch tracking does not provide results accurate enough to present the user with a reliable representation of their performance.

Certain instruments are widely available with a MIDI interface, such as keyboard and drums, due to the nature of the instrument. This can provide a very accurate software representation of a performance. While string instruments are not as well suited to a MIDI implementation, MIDI devices are available, such as divided pickups, which contain a short-range microphone for each string of the instrument. This allows for monophonic pitch tracking to be conducted on each string, resulting in an accurate transcription of a polyphonic performance. Presenting the player with a representation of their actual performance could provide a more meaningful look at mistakes, such as notes being played on the wrong string. Several music teachers commented that some students may think they have played something correctly, despite being told they have not. By giving a visual representation of their performance in tablature no-

tation, these students may realise and understand what they have actually played.

Conclusion

We have provided insight into how software can be effectively utilised to overcome the difficulties encountered by a learning guitarist. By providing intuitive feedback to a player, software can provide valuable assistance to a learning guitarist, most notably by:

- Motivating a player to practice, by providing accompaniment and giving positive feedback
- Development of listening skills, by making the player more aware of aspects of their performances

A software prototype has been designed and developed, which provides intuitive feedback to a player during and after their performance. This prototype served as the basis for usability testing, conducted with several users of a diverse range of guitar ability. The usability testing prompted a valuable discussion of the effectiveness of various feedback approaches and highlighted the ways in which software can be successfully harnessed to enhance the learning process. The prototype provides a framework from which a more advanced and feature-rich system can be developed.

References

- [1] Dannenberg, R.B., Sanchez, M., Joseph, A., Capell, P., Joseph, R. & Saul, R. 1989, 'A Computer-Based Multi-Media Tutor for Beginning Piano Students', *Interface - Journal of New Music Research*, vol. 19, no. 2-3, pp. 155-173.
- [2] ign.com 2008, IGN: SingStar Image, ign.com, viewed 8 February, 2008 <http://au.media.ps3.ign.com/media/819/819473/img_357836_6.html>.
- [3] Iwami, N. & Miura, M. 2007, 'A Support System for Basic Practice of Playing the Drums', *Proceedings of the International Computer Music Conference*, International Computer Music Conference, Copenhagen, Denmark, pp. 364-367.
- [4] Juul, J. 2007, 'Guitar Hero II: Playing vs. Performing a Tune', viewed 16 November 2007, <<http://www.jesperjuul.net/ludologist/?p=312>>.
- [5] Oshima, C., Nishimoto, K. & Suzuki, M. 2004, 'Family Ensemble: A Collaborative Musical Edutainment System for Children and Parents', *Proceedings of the 12th annual ACM international conference on Multimedia*, ACM, New York, New York, USA, pp. 556-563.

- [6] Percival, G., Wang, Y. & Tzanetakis, G. 2007, 'Effective Use of Multimedia for Computer-Assisted Musical Instrument Tutoring', Proceedings of the international workshop on Educational multimedia and multimedia education, ACM, Augsburg, Bavaria, Germany, pp. 67-76.
- [7] Puckette, M. 1996, 'Pure Data: another integrated computer music environment', Proceedings, Second Intercollege Computer Music Concerts, Tachikawa, Japan, pp. 37-41.
- [8] Yin, J., Wang, Y. and Hsu, D. 2005, 'Digital violin tutor: an integrated system for beginning violin learners', Proceedings of the 13th annual ACM international conference on Multimedia MULTIMEDIA '05, ACM Press, Hilton, Singapore, pp. 976-985.

Andrew Sorensen

Queensland University of Technology &
Australasian CRC for Interaction Design
Kelvin Grove, 4059
Australia
a.sorensen@qut.edu.au

Andrew R. Brown

Queensland University of Technology &
Australasian CRC for Interaction Design
Kelvin Grove, 4059
Australia
a.brown@qut.edu.au

Abstract

In this article we report on progress at the Australian CRC for Interaction Design investigating the computational generation of orchestral music based in the Germanic Symphonic tradition. We present an introduction to the project including a brief overview of our intended methods and some guiding principles for the project. We then outline the current state of the project and introduce our initial algorithmic system with a special emphasis on an implementation of Paul Hindemith's harmonic system. We conclude with some initial findings and future goals. We provide an extensive selection of audio examples online that accompany, verify and enhance information provided in this report.

Introduction

Starting in 2008 the Australian CRC for Interaction Design began an exploration into the computational generation of orchestral music based loosely around a mid to late romantic aesthetic. There are several reasons for our current interest in computational approaches to the generation of romantic orchestral music. Firstly, there is a commercial aspect to this investigation. A significant amount of the music currently used in mainstream media is orchestral; feature films, documentaries, television dramas/sitcoms and computer games all make significant use of orchestral soundtracks. Surprisingly, perhaps, the computer plays a disproportionately small role in the creation of orchestral music beyond its role as a notation and recording device. Arguably this is due to the structural complexity of orchestral music of the mid-late romantic period and a cultural gap between electroacoustic and orchestral musical aesthetics.

Secondly, the generative agenda of the project builds on a long tradition in computer music research, but still seems to be a contentious area of investigation in the broader music community, and with some in the computer music community. While we acknowledge the significant efforts of projects such as Cope's EMI software (Cope 1992, 2001), there remain many problems with the application of such research within a real-world con-

A Computational Model For The Generation Of Orchestral Music In The Germanic Symphonic Tradition: A progress report

text. A cursory survey of the lack of generative music processes in current professional music software is all that is required to see that this is the case.

Thirdly, we have an interest in music of the Germanic Symphonic tradition that has had an influence on music for film and, more recently, computer games and the rich musicological tradition that surrounds it. Although our own musical practice extends well beyond this genre, there is a fundamental musical vocabulary here that we perceive as integral to our growth as computational musicians and we believe is generally applicable to many musical styles. This paper outlines some of our recent attempts to apply musical theories from the orchestral tradition to generative computational techniques that can be applicable in practical contemporary contexts such as cinema and computer games.

Background

We acknowledge that this is an ambitious project and our hopes for success are largely based on the hypothesis that many of the problems inherent in algorithmic composition are implementation issues. In other words, we do not believe that a more detailed specification of the problem is required, but instead a more practical solution for its implementation. Our hope is that we have reached a turning point where the technology is capable of adequately supporting the vast wealth of musical and computational theory available to us. In particular we hope that our use of Impromptu (Sorensen 2005), a real-time dynamic system, will facilitate our exploration of this highly temporal medium.

After many years of building computational music systems the authors are now guided by two beliefs. Firstly, the seemingly obvious belief that algorithms derived from musical analysis are far more likely to provide effective musical solutions than those from formalisms outside the music domain. Having investigated Neural Nets, Genetic Algorithms, Genetic Programming, Cellular Automata and the like (Towsey et al. 2000; Brown 2005, 2005a; Wooller et al. 2005; Gerber and Brown

2006, Brown et al. 2008) the authors have returned to a simple set of principles based around probability, linearity, periodicity, set theory and recursion (Sorensen and Brown 2007). This is not to say that there is not or will not be a role for the other processes, only that we believe sonifications resulting from them tend to produce music outside the bounds of the Romantic orchestral style that is our focus here. Our goals are first and foremost to achieve a musical outcome and although we may stumble across issues related to machine or human creativity these are not primary goals of the project.

Secondly, and very much in line with the first statement, we are interested in holistic solutions to our problem. We question the usefulness of a reductionist approach that focuses on one musical element, e.g., pitch or rhythm, in isolation when dealing with real world compositional problems. We have found that it is only by experimenting with musical results that attempt to simultaneously address all dimensions of the musical puzzle, even if to differing degrees, that generative musical process can succeed. The musical whole is always more than the sum of the parts in our experience. We anticipate that we will be better able to address sub problems while experimenting with the whole and in this paper focus on one sub-problem, harmonic tension, by producing works that also include generative melody, rhythms, form, dynamics and so on.

Research design

At this stage of the project we have been proceeding using a pseudo-scientific method whereby various hypothesis, regarding music theories and computational implication, are tested through experiment for observable results. Indeed the datum of our project, are the musical compositions generated by our system. These data, as with all experimental research, are the core outcome of our work and the final and authoritative guide to the success of our project. We therefore provide numerous examples online to accompany this paper. Where doubt arises about information in this paper the musical examples should be considered the canonical source.

Hindemith's Harmonic System

We began by searching for a suitable harmonic system that accommodated the diatonic and chromatic content of the Germanic Symphonic music of the later half of the nineteenth century. Not only would the chosen harmonic system need to be musically applicable but also able to be implemented within the constraints of a real time system. There are many harmonic systems that could provide suitable stepping off points but based on the success of early experimentation with Hindemith's system we did not consider an exhaustive survey of alternative harmonic systems and techniques necessary before moving forward.

Paul Hindemith's "The Craft of Musical Composition - Book 1" appeared to fit our purposes (Hindemith 1970). It outlines a harmonic system based around intervallic relationships and the tensions inherent in these relationships. While certainly not the only comprehensive theory based around intervallic relationships, Hindemith's is simple, elegant and suitable for fairly direct computational implementation. Because Hindemith's tension relationships are subjective some have argued about the theoretical soundness of his choices. Also some suggest that its reliance on harmonic relationships may only formally be applicable to music using just intonation. They argue that an equally tempered system blurs functional distinctions between major and minor triads (Clark 2007). Despite these reservations Hindemith's system presented a good starting point for our investigations because it afforded an opportunity for rapid implementation and subsequent aural experiment. Another strong validation of Hindemith's harmonic system is Hindemith himself who made extensive use of the system for his own compositions, going so far as to rewrite many of his early works so that they would conform to this system which he developed later in life.

It is important to clarify that our intentions were never to test Paul Hindemith's system as originally conceived but to use it as a starting point to further our own harmonic enquiry for generative purposes. We therefore took the liberty to add, modify and remove aspects of Hindemith's system as our test results dictated. Usually these modifications were made to support an autonomous version of the system never intended by Hindemith himself.

Hindemith's System Of Chord Qualities

Hindemith devised a system of chord qualities that grouped all possible chords into one of six chord-groups, numbered I-VI. Hindemith's system segregated chords based upon their intervallic relationships between chord tones. The set of internal intervals identify a chord-group to which each chord is assigned. Chords in a particular group are assumed to share a similar pureness, stability or harmonic tension. The six groups are defined as follows:

- i. No tritone, no seconds, no sevenths
- ii. No minor seconds, no major 7ths, with tritone
- iii. No tritones
- iv. Contains one or more tritones
- v. Indeterminate
- vi. Indeterminate (with tritone predominating)

An example may help to clarify how this works in practice. Consider the dominant 7th chord C7 [C E G Bb] containing the following set of intervals [M3 P5 M6 m6 m7 m3 M2 tritone]. This interval set places C7 in chord-group II.

Cmaj7 [C E G B] containing the intervals [M3 M6 m6 P5 m7 m3 m2] would be placed in chord-group III. Hindemith makes additional sub-segregations based on root position but we ignore this addition in our current work.

Hindemith also provided a guide for the creation of root progressions also based on intervallic relationships. Hindemith defined two interval series (series I and II) that he uses for the creation of root progressions and other melodic materials. We do not provide further explanation here as our present system makes only indirect use of these series for the creation of root progressions.

In a simple sense, Hindemith's theory progresses by assigning chord quality to root notes where the choice of chord quality is based upon the amount of harmonic tension required at any given point in the composition. All changes in tension or resolution are relative to the current chord quality, rather than measured against an absolute scale. It should be pointed out that Hindemith never intended for the system to be autonomous and as such it provides a general, but not absolute, guide for us to follow.

Investigation and Modification

Our initial investigation of Hindemith's system indicated that it would provide musical benefits including; chromatic flexibility, relatively direct control of musical tension, the ability to cleanly separate chord quality from root progression and the trivial variation of existing root progressions. Some of the technical advantages of the system are its suitability for real time work because it works within limited temporal restraints (i.e., There is no need to look ahead or behind) and because of the efficiency of the algorithm.

During the development cycle there have been a number of technical additions that provide constraints over Hindemith's system. Firstly, there is a Pitch Class Set (PCS) parameter that offers the opportunity to help constrain chordal choice. For example, Hindemith's system places no constraints over the choice of major or minor chord qualities, leaving this instead as a subjective choice for the composer. For example, both Cmaj [C E G] and Cmin [C Eb G] are contained within chord-group I. Our implementation uses the PCS parameter as a means for weighting these types of choices as a way of maintaining "obvious" constraints that Hindemith left to composers. Additionally, the current implementation of the algorithm only makes use of Hindemith's first three chord groups (I-III) leaving out the indeterminate choices (V-VI) and the multiple tritones subordinate group (IV). This was an early choice to impose limits on the range of harmonic variety to a practical, as opposed to theoretically complete, range. To date, our limited subset has proven to provide adequate variation. We may choose to add the further three groups at a later stage or for use in different musical contexts. Additionally, we do not currently implement inversions as distinct chordal subgroups, as Hindemith does. This allows us to more

simply manage the construction of musically meaningful root movement and voice leading.

Before beginning a detailed investigation of the algorithmic implementation we will describe in brief the overall design of the music production system.

System Overview

The generative process begins with harmonic selection. Even though there is no requirement to begin this way, we feel that this is reasonable given the dominance of harmonic considerations in late romantic music and the emphasis on thematic rather than more traditionally melodic movement. Our harmonic selection process begins by generating a root progression including both pitch and rhythm dimensions. The root progression is generated to fit a given time period; eight bars for example. The algorithm then progresses linearly through the root progression. For each root note several activities take place. Firstly, a chord quality is selected for the given root. The harmonic tension of this chord quality is substantially influenced by the Hindemith chord group value (I, II, or III) passed to the chord quality generator. Once the chord quality is selected a score is constructed based on current settings for orchestration, pitch range, harmonic accompaniment and voice leading. Finally, a bass line is added and thematic material is selected and manipulated to fit the current harmonic context.

Generative processes for aspects other than harmonic considerations are relatively rudimentary placeholders. We have already started to reap benefits from the approach of implementing many aspects of the system in parallel, even if not with equal sophistication. One interesting result of this experience is just how much is possible with a very superficial implementation provided that coverage is broad. We will now discuss each aspect of the overall system in more detail.

Rhythm Generator

As the rhythm generator is discreet and is used by many parts of the system we will begin by describing its implementation. The rhythm generator is a simple stochastic function providing control over duration, tactus, level of syncopation and rhythmic value list (herein called the RVL parameter). The generator initially selects a rhythm value at random from a user provided RVL. The generator will continue to select rhythm values at random from the RVL until either (a) the maximum duration is reached or (b) the maximum duration is surpassed at which point the algorithm will backtrack to a point at which it can successfully continue forward.

Implementing gestalt laws of proximity and similarity that evaluate the results of randomly chosen rhythm patterns provides the musical effectiveness of this algorithm. The weighting of this gestalt selection is made through a combination of the tactus and syncopation parameters. At a tactus

point a stochastic selection will always be made, at all other times a percentage choice, the syncopation value, will determine a gestalt or stochastic choice. Random reselection is forced until patterns pass evaluation. This simple algorithm has proven very effective at handling all of our existing rhythmic generation requirements.

Root Selection

The root selection algorithm currently operates using a relatively simple point-to-point style approach. There is no explicit cadential knowledge built into the root selection algorithm. The first part of the root selection process defines a harmonic rhythm selected using the stochastic rhythm generator described above. The point-to-point process begins with parameters for a starting Pitch Class (PC), ending PC and interval value list (IVL). The algorithm selects N-2 (i.e., minus starting PC and ending PC) values from the IVL that when combined move from starting PC to ending PC.

Cadential knowledge can be added by setting an appropriate PC ending sequence. A half cadence for example requires no modification while a perfect cadence is trivially implemented by appending a resolving I chord. This simple process provides the basis for implementing all major cadential forms. The critical role that bass movement plays in cadential movement is acknowledged and is discussed later.

The harmonic complexity of the root progression is controlled by the contents of the IVL, however two further Pitch Class Set (PCS) quantization methods are available. Firstly, the output of the root progression generation can be quantized against a PCS. Secondly, the IVL can be implemented as a step value list (SVL) that is applied against a PCS.

In conclusion, our control over the root selection process is fairly fine grained. We can choose to quantize to a PCS or not, we can choose to work within a traditional system of cadences or not and we can control the degree of harmonic complexity in our root progression by adjusting the values in the IVL.

Chord Quality

A chord quality is defined for each root in the root progression as it is played. The chord quality is selected using Hindemith's chord quality system. Our initial implementation of Hindemith's algorithm used a simple iterative stochastic process. A random selection of PC's is made and the complete set of intervals joining these PC's is calculated. A random selection of the PCS [C E G] would produce the interval list [M3 m3 P5 M6 m6]. This set of intervals is then tested against a user specified Hindemith chord group (i.e. I, II or III) for validation. Successful validation of the random selection results in it's being returned to the user as appropriate. At this stage it is important to realize that the chord is a set of interval rela-

tionships free of any definite pitch classes. As a final step in the process the user provides a root for the chord and an appropriate PCS is generated. The PCS does not designate an inversion nor octave displacement; rather, it is an unordered set of pitch classes.

As an example let us consider a group I chord that can be either a major or minor triad. If the user asks for a group I chord with root 2 then the algorithm will probabilistically return either a Dmaj or Dmin chord with no extensions. Chord groups II and III provide a far greater range of chord qualities and the distribution of type I, II and III chord qualities is, along with the root progression itself, the greatest source of control over harmonic tension.

Bass Note

After assigning a chord quality the next stage of the process assigns one or more bass tones for the current chord duration (the bass may move). For cadential purposes the generator attempts to accentuate the current musical key while avoiding strong cadential reference. Our initial implementation achieves this by using a simple weighting to reject root and bass correlation (i.e., we use inversions instead of root positioning) away from phrase boundaries. At this stage we freely invert using any option available within the context of the current chord quality. The generator also uses root position chords where possible, and attempts to maintain minimal linear voice movement of the bass.

Scoring

The next phase in our process involves the voicing and orchestration of the chord. We now have a fairly complete harmonic picture; we know the root, quality and bass of the chord. Using this information we proceed to orchestrate the chord using a further range of system parameters including the current instrumentation, number of active voices, and lower and upper pitch bounds. Additionally the scoring algorithm uses the previously scored harmonic block as a reference point for smooth linear part movement.

Our linear part movement algorithm follows a simple point-to-point approach searching for the shortest path between the two chords while maintaining complete chord coverage (i.e., making sure each chord tone is represented). Finally, the algorithm makes decisions about the style of accompaniment to apply. At present we have only implemented two accompaniment styles, an arpeggio style and a homophonic style. A choice about whether or not to perform the bass part is also made at this stage and this decision depends upon the instrumentation, range and accompaniment style chosen.

Thematic Material

The last stage in the process involves generation of a thematic fragment to fit the current harmonic context and duration. First we generate a rhythm equal to the duration of the current chord, and then select intervals at random from a weighted list to determine pitch contour. The generated theme can then be transposed to commence on any degree of the currently active chord. The theme is quantized to a musical scale that agrees with the current chord quality, chord root and current key. We choose a scale by correlating the pitches from the chord with the standard church modes, rooted against the current key. There is an additional option to select a scale simply from the chord root; this is primarily applied against chordal roots outside of the current key.

With a series of simple extensions this trivial thematic generator can be tuned to provide passable melodic material. In part this is due to the dominance of thematic rather than melodic invention in romantic music.

We provide the ability to choose not to play a melodic fragment and provide the option to accompany the chosen thematic fragment with a delayed copy played on a second instrument and possibly in a separate register. Thirdly, we make extensive use of themes previously generated during the course of performing the piece using random repetition and recapitulation to provide some global coherence. The final stage of thematic generation is to assign an instrument for the performance of the theme.

Orchestral Performance

Following the generation of the thematic material our musical data is ready to be performed. We use the Vienna Symphonic Library for sample playback.

We have designed a quasi agent-based approach for instrumental playback to take advantage of the low level control we have over sample playback. Instruments act semi-independently in their responses to musical note information choosing a sample patch based on a notes volume, duration, articulation style, and so on. For example, a trumpet knows which sample bank it should play from given a heavy attack, loud volume and short duration, as well as providing some simple range checks and instrument specific performance options (muted options for brass, pizzicato for strings etc.). We also apply a variety of gestural control mechanisms at this level, such as multi layered oscillators for dynamic modulation, control of legato performance parameters, fine grained volume control and alike as detailed previously (Sorensen and Brown 2007). These performance details greatly enliven performance.

Additionally the playback system supports multiple independent metronomes allowing us to modulate tempo for each individual part if required. This allows us to fairly trivially add rubato

playing where instruments are linked to a single metronome or move independently to their own individual metronome.

The process described above is rapid enough to be calculated and performed in real-time for a relatively complete orchestration of flutes, oboes, clarinets, bassoons, trumpets, horns, trombones, tuba, violins, violas, cellos, basses and percussion. A detailed discussion of all aspects of our performance system would require another full paper of equal length, so we will defer that report for another time.

Shortcomings

Before we begin a discussion about the results of our current research we would like to point out some of the known shortcomings of our current system so that we can take account of them when discussing our results.

We make very little allowance for good voice leading beyond our simple shortest path solution. Voice leading is an area in which computational study has arguably had its greatest successes (Ponsford 1999, Huron 2001, Hömel 2004). Given that we pay only superficial regard to voice leading we can expect to suffer from many of the classic part writing concerns, parallel movement, consecutive fifths and octaves etc. Another concern of our current shortest path algorithm is our lack of control over the distribution of voices over the complete pitch range. In other words there is little to stop bunching of parts at the top, centre or bottom of the pitch range. Given previous research in this area we feel confident that voice leading can be easily improved in the future.

There is no explicit shaping of the point-to-point root progression that should result in non-directed root progressions and therefore non-directed harmonic movement. Additionally, our current cadential support is superficial and we would expect this to manifest in contrived and stilted phrase and section boundaries.

We use no melodic shaping whatsoever. In fact, aside from simple pitch constraints and basic repetition, we provide no melodic support at all. One might expect this to result in generally worthless melodic material. A related shortcoming is the absence of any counter-melodic material aside from trivial thematic cannoning that is occasionally employed.

The lack of higher-level structure is a common weakness in generative music systems. Our system is currently limited to the manual enveloped control of some global parameters in the shaping of overall form and structure. We expected that this would provide reasonable control over musical intensity as dictated by harmonic tension, dynamic and orchestration for example, but provide no value for other structural features, such as repetitions and variations that require more than local memory.

We currently used structured accompaniments. This is actually not as unusual as many people would expect for either manual or genera-

tive compositional systems. There are a range of accompaniment patterns that are found time and time again in orchestral music and programs such as Band-In-A-Box rely almost entirely on them. We currently provide no mechanism for escaping beyond the bounds of the preconceived accompaniment patterns. We anticipated that this would severely limit the range of stylistic output that the system was capable of, but this limit is not difficult to extend and will certainly be a focus of further developments.

Results and Future Work

Here we outline our subjective reactions to the musical output of our system (the data) and sincerely hope that interested readers will also review the musical examples made available here: http://impromptu.moso.com.au/acmc08_examples.html

Listening back to the music generated by our system included a happy surprise for us that local cohesion operated surprisingly well. With only a few simple operations, such as thematic repetition and re-use of root progressions we feel the music has reasonable local structure without being overly restrictive. Most importantly we feel that our harmonic implementation is currently operating well enough to provide a balance between harmonic novelty and structural cohesion. Overall we are happy with our current implementation of Hindemith's chord quality system. We do however need to make improvements to the sophistication of our cadential treatment in order to help drive the music at section boundaries.

When listening to the music generated by our system there are many obvious deficiencies in the music currently produced. Most pronounced is the lack of meta-structure. This is hardly surprising as it is a common complaint about generative music systems and we foreshadowed this in our own shortcomings section above. We are not overly concerned about this at this stage, however, as we hope that introducing a "memory" into our system will help to introduce opportunities for global cohesion. We are also working with various user interface tools to allow higher-level human control over a parameter automation process that will assist to provide meta-structure. The development of these interfaces is directly in line with our desire to produce usable tools for working professionals in film, TV and computer game environments, and early prototypes of these interfaces are discussed in an accompanying publication.

Our voice leading, as we speculated earlier, does indeed cause some concern, although only partially. Our concerns over common voice leading issues such as parallel movement seem to be largely misplaced. In general we have no major complaints about the shortest path approach, which provides relatively smooth voice movement. Of greater concern however is the arbitrary wandering of this movement, which can lead to voice bunching and unnecessary voice crossover. We will need to address this by forcing some form

of inner voice boundary checking and possible use of medium scale melodic contour control.

Our minimal set of accompaniment styles is a severe limitation on the stylistic variety possible with the current system. Indeed, our requirement that accompaniments be somewhat pre-defined is an inherent weakness in the system more generally. In future work we would like to investigate extending the instrument agent definition to provide some form of environmental listening, enabling each instrument "agent" to have the ability to perceive its relationship to surrounding material. In this way we may be able to provide automatic accompaniment based around various gestalt principles of grouping, symmetry and self-similarity.

One of the great surprises for the work so far is how well completely random thematic material can work. This is not to say that it is reliable, indeed thematic materials can be as bad as they can be good. However, given that our thematic implementation is a placeholder only, we find its operation surprisingly serviceable. Our intention is to extend this current system to provide greater local scope. For example, we currently only support short thematic fragments, we would like instead to support melodic material that could be sub-divided into smaller thematic components. This would maintain the interesting expositional features that we currently enjoy in the system while enabling longer, more cohesive, melodic passages. We hope this might also provide us with more cohesive phrase level operation, a current deficiency linked to the aforementioned cadential weakness.

This brings us to our most significant insight. Overall we are convinced that the success of the system to date is very largely due to our holistic approach to implementation. In other words, we feel that the output of our present system is far greater than the sum of its often greatly limited parts. This is in line with our initial expectations, and on the surface would appear to support our approach. However, we are increasingly concerned that a by-product of this success is a severe limitation to the range of stylistic output the system is capable of. In other words, we are concerned that our approach may be producing output closer to a single "work" when we would hope for a system capable of producing significantly varied "works". At this early stage it is difficult to know if the system's aesthetic limitations are due to its partial implementation or whether there is a more serious methodological problem with our approach. We will need to keep this strongly in mind in the future.

Conclusion

This has been an introduction to the ongoing work at the Australian CRC for Interaction Design into the generation of orchestral music. This is a report on early outcomes and we have much work still to do ourselves and in partnership with others.

Acknowledgments

The work reported in this paper has been supported by the Australasian CRC for Interaction (ACID) through the Cooperative Research Centre Programme of the Australian Government's Department of Education, Science and Training.

References

- Brown, A. R. (2004). An aesthetic comparison of rule-based and genetic algorithms for generating melodies. *Organised Sound*, 9(2): 191-198.
- Brown, A. R. (2005). Exploring Rhythmic Automata. Proceedings of the *Applications of Evolutionary Computing: EvoWorkshops 2005*, Lausanne, Switzerland. Springer, pp. 551-556.
- Brown, A. R. (2005a). Generative Music in Live Performance. Proceedings of the *Australasian Computer Music Conference*, Brisbane, Australia. ACMA, pp. 23-26.
- Brown, A. R., Wooller, R. and Miranda, E. R. (2008). Interactive Evolutionary Morphing as a Music Composition Strategy. In, *Music As It Could Be: New Musical Worlds from Artificial Life*. E. R. Miranda, Ed. Madison, Winconsin, A&R Editions.
- Clark, J. W. (2007). *Hindemith's Analyses in The Craft of Musical Composition*. Retrieved 7 April, 2008, from <http://jasonwclark.com/hindemith.aspx>.
- Cope, D. (1992). Computer Modeling of Musical Intelligence in EMI. *Computer Music Journal*, 16(2): 69-83.
- Cope, D. and Hofstadter, D. R. (2001). *Virtual music : computer synthesis of musical style*. Cambridge, Mass., MIT Press.
- Gerber, A. and Brown, A. R. (2006). Visualising Music with Impromptu. Proceedings of *Medi(t)ations: The Australasian Computer Music Conference*, Adelaide. ACMA, pp. 38-43.
- Hindemith, P. (1970). *The Craft of Musical Composition: Theoretical Part 1*. Mainz, Germany, Schott.
- Hömel, D. (2004). ChordNet: Learning and Producing Voice Leading with Neural Networks and Dynamic Programming. *The Journal of New Music Research*, 33(4): 387-397.
- Huron, D. (2001). Tone and Voice: A Derivation of the Rules of Voice-Leading from Perceptual Principles. *Music Perception*, 19(1): 1-64.
- Ponsford, D., Wiggins, G. and Mellish, C. (1999). Statistical learning of harmonic movement. *The Journal of New Music Research*, 28(2): 150-177.
- Sorensen, A. (2005). Impromptu: An interactive programming environment for composition and performance. Proceedings of the *Australasian Computer Music Conference 2005*, Brisbane. ACMA, pp. 149-153.
- Sorensen, A. and Brown, A. R. (2007). aa-cell in practice: an approach to musical live coding. Proceedings of the *International Computer Music Conference*, Copenhagen. ICMA, pp. 292-299.
- Towsey, M., Brown, A. R., Wright, S. and Diederich, J. (2000). Towards Melodic Extension Using Genetic Algorithms. Proceedings of *Interfaces: The Australasian Computer Music Conference*, Brisbane. Australasian Computer Music Association, pp. 85-91.
- Wooller, R., Brown, A. R., Miranda, E. R., Berry, R. and Diederich, J. (2005). A framework for comparison of processes in algorithmic music systems. Proceedings of the *Generative Arts Practice*, Sydney, Australia. Creativity and Cognition Studios Press, pp. 109-124.

Wendy Suiter¹, Eva Cheng²

¹Faculty of Creative Arts

²School of Electrical, Computer, and Telecommunications Engineering

University of Wollongong

Wollongong NSW 2500

Australia

ws999@uow.edu.au, ecc04@uow.edu.au.

Listening Chairs: Personal Acoustic Space in Public Places

Abstract

Creation of a personal acoustic space for listeners of audio works in a public venue, which encourages the listener to focus and engage with the musical work, is the primary goal of this collaboration. Considering the context of this artificially created acoustic space various subsidiary goals are also important. These include visual aesthetics, as well as the audio capacity of the employed technology, while considering the portability of the equipment and the overall cost.

Introduction

The primary goal of this collaboration between engineer and composer is to create an environment in which listeners of audio works are encouraged to focus and engage with the musical work. The type of environment we are seeking to create is a personal acoustic space within a larger public venue. Despite the apparent simplicity of the goal there are underlying philosophical questions about the relationship of listener to sound works and the environment in which these works are presented, as well as the numerous technical questions and challenges we have met along the way.

This paper is an interim report on a work in progress. It will outline the context of this collaboration, make explicit the underlying philosophical assumptions for the project, and describe the process of designing and creating a new piece of equipment, with regard to optimising both the quality of the audio reproduction and the capacity to duplicate the equipment with regard to the overall cost.

While we describe the final goal of the project as an installation, we use the word in a rather abstracted way. Usually artists will use the word installation to refer to a specific genre of visual art, which may also include some sound or musical elements to create an alternate environment to the one in which it is physically located. Usually this new environment is an attempt to communicate the artist's conceptual ideas on a particular theme. We take the notion of creating a smaller environment for a specific purpose, and so use the word 'installation' to describe our project. However, it should be noted that we are using the concept quite abstracted from any specific music or

sound that may be reproduced through the system.

As we come from different but overlapping backgrounds this collaboration has also been conducted in a spirit of enquiry into how other disciplines function with regard to their research and approach to technology.

Context of This Collaboration

Wendy is a composer of both acoustic and audio works, currently undertaking doctoral studies at the University of Wollongong. One of her interests is the presentation of her work in an accessible environment, which enhances understanding and focus on the work, while maintaining the integrity of the work itself.

It is not possible to work in the audio domain without having to learn something of the underlying technology and underlying theoretical concepts, even though the primary focus of my learning in this field is for creative purposes.

My compositional background in both contemporary classical music and experimental music has led me to explore a number of different compositional techniques, including sonification. This lead to my meeting of Eva at the ICAD conference in Sydney in 2004.

Eva is a telecommunications engineer specialising in digital speech and audio processing, with an underlying passion for music. Her doctoral research investigates the extraction of spatial information from multi-microphone speech recordings in reverberant environments. In parallel with researching in and developing new audio engineering technology and processing techniques, Eva questions and experiments with the importance of spatial sound in human experiences and how sounds correlate to certain environments. Eva's interest in 3D audio recording and reproduction techniques has uncovered a curiosity about acoustic space, 'surround sound' contemporary compositions, and how composers and engineers view and treat the roles of acoustics, space, and their listeners.

Philosophy and Design Requirements

Listeners in concert halls are automatically given the opportunity to be focussed on the music by fact of being in the concert hall environment. The characteristics of the concert hall which encourage

listener focus are: 1. a defined space into which an audience member enters with the intention of listening to music; 2. the audibility of the work, including lack of background and distracting other sound; 3. seating for listener comfort; and 4. low levels of visual stimulation or distraction. Indeed often the only visual stimulation are the performers.

Some other composers have worked on solutions to the problem of providing a particular acoustic space for listening to audio works. For example Francoise Bayle created the Acousmonium, a loudspeaker orchestra, in 1974 at the Groupe de Recherches Musicales (GRM) in Paris. The Acousmonium contained eighty speakers of different sizes placed across a stage at varying heights and distances. Their placement was based on their range, their power, their quality, and their directional characteristics. This essentially replicates the concert hall situation replacing the performers with speakers.

Our philosophy of making music more available to a wider audience led to the concept of bringing audio works to a public space in a comfortable yet personal listening environment. Such a listening environment would not then be an experience limited to those who can afford to attend or reach museums, concerts, or similar venues generally located in metropolitan areas.

Previous works exploring seated personal acoustic space in public places include Bernhard Leitner's 'Firmament' (1996), which provided a six speaker 'symmetric' listening environment for a seated listener. Although the speakers are focused to provide an exploration of zenith, the installation is not flexible for different listeners and the leakage sound is likely be noticeable due to the open arrangement of the speaker array. Ros Bandt's outdoor work, Listening Place (2003), utilises underground speakers to present spoken stories to listeners seated on a park bench. Creating a diffuse soundfield, the installation is site-specific and is thus not easily portable. Iain Mott's 'Talking Chair' (1994) interactive installation places the listener in an immersive six-speaker personal soundfield. Although the shape of the 3D soundfield is wirelessly controlled by the listener, the sound is quite diffuse (and needs to be for 3D spatialisation) and the physical structure not flexible or easily portable.

To fulfil our philosophy of providing personal acoustic space in public places, the installation must be as flexible as possible, in order to be able to accommodate a variety of venues, and acoustic environments (including the outdoors), and listeners of varying physical dimensions and abilities. In addition, the installation can be used to present any stereophonic audio work so the design is independent of the audio composition for maximum flexibility in this dimension. Thus, our work may be considered as presenting a flexible and comfortable stereophonic playback system

This distinguishes our work from these earlier chairs which are specially designed sound art installation encompassing a fixed physical structure which integrates with the specific audio work being presented.

We have chosen to focus on stereophonic audio output for the creation of an intimate, personal soundfield for the listener. This is still the most commonly available audio format for general use, enabling use of off-the-shelf components in the construction of the installation. However, there is no philosophical, design, or practical reason why more speakers cannot be deployed around the listener for surround sound. In fact, this is a future extension of the installation. We chose to use speakers over headphones due to listener comfort, freedom of movement, and most importantly, engaging the listener in a personal acoustic space that is not isolated from the greater public space. In this way, the listener is still part of the public space, rather than isolated in an individualised virtual acoustic space created by wearing headphones.

The installation also aims to address issues of presentation of audio works and how they are consequently consumed by listeners. Often, audio works are presented in museums, galleries or other public spaces where the listeners are limited to enjoying the works in passing, as there is no seating available to offer a comfortable means to engage with the work. Our philosophy considers the listener and their comfort as an intrinsic part of the installation. However, one assumption is made about the listener: that a seated position is more comfortable and conducive to engaging with the audio work than a standing position. As concerts traditionally offer seating (before moshpits and dancefloors became fashionable!) and people generally listen to music seated, we do not consider this assumption to limit the installation.

Considering the listener and listener comfort as central to our installation also raises questions about how the listener is perceived and accounted for when acoustic spaces are discussed in general. Are listeners considered part of the acoustic space? Conversely, without listeners, what does an acoustic space mean? Often, engineers overlook the listener's experience in favour of technical specifications for acoustics, such as room reverberation characteristics. Alternatively, how often do (or should) composers decide upon and compose for a specific acoustic space or simply leave such considerations to a museum curator or concert manager for maximum flexibility in presentation of the work? Or are characteristics of an acoustic space part of creative expression and thus part of the composition?

Installation Technical Design

Our philosophy considers the listener to be central to acoustic space. Thus, the listener's comfort and

experience is one of the primary motivations behind the installation concept and design. A second, equally important motivation, is to bridge the ‘musical divide’ by assisting in making art music more available to the general public by creating a personal listening environment located within a public space.

Current research in, and commercially available products for, directional, focused sound reproduction address the issue of creating a personal acoustic place in a public place with variable levels of success depending on the technique: beam-steering, reflector-based, or ultrasonic.

Beam-steering, adopted from antenna theory and sonar, is a large-scale approach, that steers sound in a particular direction: manipulation of the magnitude and phase of each loudspeaker arranged in an array steer the sound in a particular direction by controlling where the sounds sum or cancel. The most common loudspeaker arrangement is linear, which allows for beam-steering in the plane perpendicular to the array. Two-dimensional loudspeaker arrays extend the capabilities to focusing sound in both planes, potentially converging to a central focal point. The advantage of the beam-steering systems (Duran Audio, EAW, Meyer Sound) is that the frequency response is adopted from the loudspeakers in the array (so high fidelity is possible); the disadvantages are variation in frequency response with the radiation pattern and the cost involved with the number of loudspeakers and multichannel hardware required.

Reflector-based technologies apply the theory of satellite dishes to use curved reflective surfaces to amplify or direct sound in acoustic bandwidths (Wahlström, 1985). However, the frequency response of the system is variable and highly dependent on the curve size and shape e.g., parabolic, hemispheric, or a hybrid of various shapes; the amplification is greatest with high frequencies due to the shorter wavelengths. Parabolic reflectors (MuseumTools, Meyer Sound), although achieving a greater range than other shapes, cannot create stereo images, whereas hemispheric and hybrid sound domes can create virtual stereo imaging through an audio ‘hologram’ at the listener’s ears (Brown Innovations, SoundTube Entertainment). The main advantage of these systems is the low cost (relative to other focused sound approaches) and ease of use.

Adopting techniques from AM radio propagation, using ultrasonic frequencies to modulate audible sounds relies on the nonlinear interaction of sound waves in air: two frequencies emitted in close proximity will generate secondary signals at their sum and difference frequencies (Yoneyama et al., 1983; Pompei, 1998). Thus, when ultrasound is modulated by an audio signal and transmitted, the nonlinear wave interactions through propagation in air demodulates the ultrasound signal to recover the original audio. To achieve focused



Figure 2. Prototype of one component for one listener.

sound, wave propagation theory states that the source directivity depends on the size of the source compared to the emitted wavelengths. Thus, due the short wavelengths of ultrasonic signals, the sound propagates in a narrow beam – ideal for focused sound. The disadvantages of ultrasonic systems (Holophonic Research Labs, American Technology) are that, due to the direct sound beam, the apparent sound source is inside the listener’s head; and, stereo images cannot be produced as coherently combining two beams is non-trivial.

The technique most comparable and relevant to the goals of our project are hemispheric or hybrid reflectors (e.g., the Localizer from Brown Innovations or Focus Point Speakers from SoundTube). Hemispheric or hybrid reflectors are the only small-scale, focused sound technology to support stereo imaging – spatiality, in addition to directivity in a personal acoustic space. While these objectives are central to our goals, the use of such commercially available products in our installation is currently limited by budget and, quite significantly, the frequency and polar response characteristics. Although the Localizer and FPS bandwidth encompass the audio range (approx. 150Hz - 20kHz, +/-3dB), the low-frequency attenuation is significant, and the frequency response is not as flat as conventional speakers, with resonant peaks clearly exhibited. In addition, polar radiation patterns can vary greatly with different frequencies. However, adopting the technical concepts of hemispheric reflectors for a more sophisticated playback system is a potential future avenue for this work.

The challenge is then to map our philosophical questions and subsequent design requirements into a physical installation, using equipment of optimal reproductive capacity. Ease of use (for the

listener and curator), ease of reproduction of components, portability, visual aesthetics, acoustic quality, and overall cost are the key design drivers; how each of these physical, visual, and acoustic design constraints were addressed will be discussed in more detail.

Physical Design

The installation conceptually consists of four armchairs placed in a square facing outwards, with each chair sporting a pair of small stereo speakers adjustable to head height, enclosed in an adjustable floor based housing located behind the chair. Figure 1 conceptually illustrates the arrangement of the chairs.

Figure 2 illustrates one prototype component,

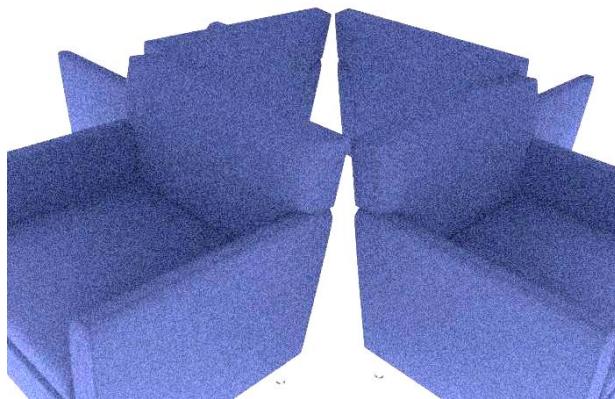


Figure 1. Arrangement of the chairs.

that is the arrangement of elements for one listener. Each chair has an independent audio source: the speakers are connected to an out-of-sight portable CD player.

The audio flow path and interaction between installation components is shown in Figure 3. It will be necessary to provide a power source for the audio hardware (speakers and CD player). To accommodate for a range of venues, this power source can be a 240V wall outlet for indoor installations, or we can also work with portable power sources for the outdoors (e.g., solar power or batteries) with minimal alterations to the technical design.

For aesthetics, ease of use, and ease of producing a number of sets of the components (by using 'off-the-shelf' equipment), a floor lamp was chosen for the speaker housing. The decision to use floor lamps to hold speakers is a result of many attempts to custom-design speaker housing that could be mounted onto the back of any armchair. However, having customised mounts limited the ease of duplicating the system, since each would need to be individually made by a technician. In addition, designing a mount flexible enough to accommodate a variety of chairs proved quite dif-

ficult. Thus, a floor-mounted option was selected as this has only one requirement: that floor space is available behind the armchair.

The particular floor lamp model used (Austrabeam Glee 32336 in black), shown close-up in Figure 4, was selected based on cost, having a pair of separable lampshades, lampshade shape and size (to hold the speakers), ease of re-cabling, flexibility (to accommodate different chairs and listeners), and portability (the lamp can be dismantled and flat-packed).

Flexibility is a major concern as chairs and listeners can vary in width and height, and listeners will vary in seating position e.g., slouching. Thus, flexible lamp components, such as the long 'goosenecks' used to connect the lampshade to the lamp body (see Figure 4), are ideal for accommodating these requirements. We are currently sourcing goosenecks to use in the lamp body to add further flexibility to the physical design.

In summary this physical design meets the requirements of having the:

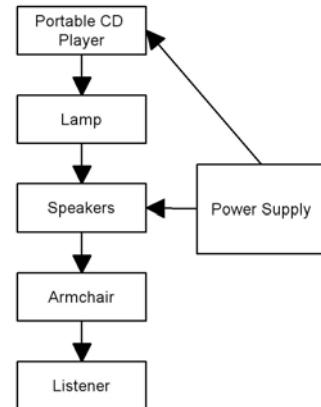


Figure 3. Block diagram of audio signal flow and installation components

- Listener central to the acoustic space;
- Listener comfort intrinsic to each component;
- Creation of a personal acoustic space in a public space. Note that we can only control the personal acoustic (sub) space, and that we assume nothing of the ambient, greater public space.
- Flexible enough to accommodate a variety of listeners' physical requirements.
- Is relatively easy to duplicate the component parts.
- Is portable and can be used in a variety of environments.

Acoustic Design

Acoustic quality is obviously of utmost importance in an audio work. However, cost, ease of duplication, matched acoustic responses, and size can severely affect technical characteristics such as frequency response (esp. at low frequencies) and



Figure 4. Speaker housed in lampshade with acoustic insulation and gooseneck connection

audio clarity, tonality, and distortion (esp. at loud volumes).

To continue using ‘off-the-shelf’ equipment for ease of duplicating each component, several pairs of (cheap) computer speakers were tested for their trade-offs in technical characteristics, as listed above, and evaluated subjectively by the authors with a range of audio sources. The Altec Lansing 120i model, being a pair of 3 watt, 4 ohm speakers, was chosen due to its low cost, audio quality, availability of power, volume, and tone controls on the circuit board, and easy dismantling from its casing to re-cable for fitting in the floor lamp. The speakers have a frequency response of approx. 85Hz to 18kHz, a range that we would consider to be suitable for most audio works.

The creation of a personal acoustic space (or small ‘sweet spot’) requires the soundfield to be concentrated at the listener’s head i.e., between the two speakers. This is the main acoustic design reasons we chose to use the floor lamp in Figure 4. The diameter of the lampshade snugly fit our speakers with a shape and length ideally suited to direct the audio towards the listener’s ears and minimise sound diffusion.

To further reduce sound diffusion, beyond the directionality provided by the speaker mounting, the rear and inside of the lampshades were lined with flexible but dense polyester building insulation. Adding insulation to the inside of the lampshade served two purposes: it reduced acoustic reflections (i.e., reverberation inside the metallic lampshade) to give a clearer sound; and, minimised acoustic leakage from the patterned holes at the edge of the lampshade, presumably present for visual aesthetics and light diffusion.

Small movements made by the listener’s head out of the ‘sweet spot’ results in a significantly different soundfield, as verified by our pilot listening tests, discussed in the next section. Such a personal soundfield in a comfortable environment allows the listener to focus on the musical work as their physical needs are accounted for. Although the acoustic design maximally contains the optimal soundfield to the space between the speakers, some sound leakage is allowed such that passers-by can hear enough to rouse their aural curiosity.

While some small leakage may assist in attracting their attention, in order to make the installation as flexible as possible, each listener will automatically hear the audio work from the beginning when they are seated. We are still investigating how build in an automatic reset function for the sound when the listener is seated. We would like the reset to occur as the listener’s head is located between the speakers.

Although we have devised a concept for this part of the project we have not yet begun to build the required electronics. The technical concept is to use infrared (IR) sensors in the two speakers, and as soon as the listener’s head enters the soundfield (i.e., breaks the IR beam between the speakers), this triggers the reset function on the audio source (CD player). Minimal electronics are required for this concept, and with some tweaking of electronics we hope to incorporate the extra components as add-ons to the CD player circuit boards.

Visual Design

The use of these floor lamps is thought to be aesthetically pleasing, while also playing with the expectations of the audience who may be intrigued by the notion of sound emanating from a lamp shade. Although the work focuses on the aural senses, the importance of visuals cannot be overlooked. Due to the philosophy of the work requiring a personal soundfield, only softly audible, intentional ‘leakage’ audio can be heard by passers-by. Thus, potential listeners passing by the installation are more likely to be drawn in by the visual appeal, only hearing the ‘leakage’ audio at close range. Hence, a visually aesthetic installation is paramount in attracting potential listeners to engage with the audio work. This is one of the key motivations behind using ‘off-the-shelf’ equipment in non-traditional ways: this naturally draws in people’s curiosities.

Far-field visual aesthetics are essential in this work to bring in listeners; however, there must be minimal visual distractions for the listener once they are seated. As visuals can easily distract a person from their aural senses led to the physical arrangement of chairs shown in Figure 1. With the rest of the installation and other listeners out-of-sight from the listener, each listener is minimally distracted from engaging and experimenting with the aural stimulus presented by the installation.

Listening Tests

We conducted a pilot study on four listeners who simply happened to walk past our installation. A number of physical requirements, especially in regards to flexibility for accommodating different chairs and listeners, were brought to our attention from these informal listening tests using untrained listeners.

Physical design issues highlighted during listener tests include: 1. Listener sitting position: although this can be seen as an unexpected acoustic discovery! Slouching in the armchair and focusing the speakers downwards towards the ears results in a completely different listening experience (and more comfortable for some)! 2. Listener height variation even when seated is significant. This requires wide flexibility in the positioning of the speaker mounting.

The main acoustic design issue highlighted by listening tests was surprise at the concentrated soundfield, which indicates that the overall concept of encouraging potential listeners to be seated to engage with a personal acoustic space is worth pursuing.

This pilot study was really useful in highlighting some features of which we were ignorant. And we intend to conduct more tests iteratively for maximal listener feedback as we proceed with the project. In particular, we will seek comments regarding our assumptions about listener comfort, portability, flexibility and the general acoustic experience of the listener.

Extensions and Future Work

1. Auto reset as the listener engages with the installation.
2. Listener control of audio (start/stop/reset/volume).
3. Testing multiple sets of components out in field.
4. Investigating alternate power sources such as solar power.
5. Testing in anechoic chamber (an unnatural public space)
6. Possible extension of work: surround sound with lamps in a circle around the seated listener.

References

- American Technology [Online], <http://www.atcsd.com/site/content/view/34/47/>
- Brown Innovations [Online], <http://www.browninnovtions.com/sounddomes.html>
- Duran Audio [Online], <http://www.duran-audio.com/>
- EAW, [Online] <http://www.eaw.com/proutcs/DsaSeries/>
- Electronic Music Foundation, [Online]
<http://emfinstitute.emf.org/exhibits/acousmonium.html>
- Holosonics [Online], <http://www.holosonics.com>
- Meyer Sound, [Online] <http://www.meyersound.com/products/industrialseries>
- Museum Tools [Online], <http://www.museutools.com/secretsound.html>
- Pompei, F (1998), "The use of airborne ultrasonics for generating audible sound beams", *Proc.*

- 105th AES Convention, Preprint 4853, vol. 47, no. 9, 1998, pp.726-731.*
- SoundTube [Online],
<http://www.soundtube.com/cgi-bin/main.cgi?Speakers=start&series=6>
- Van Beuningen, G. W. J., Start, E. W. 2000. "Optimizing Directivity Properties of DSP Controlled Loudspeaker Arrays", *Proc. Reproduced Sound 16 Conference*, Stratford upon Avon, UK, 17-19 Nov. 2000.
- Wahlström, Sten (1985), "The Parabolic Reflector as an Acoustical Amplifier", *Journal of AES*, Volume 33 Issue 6, June 1985, pp. 418-429.
- Yoneyama, M., et al. (1983), "The audio spotlight: An application of nonlinear interaction of sound waves to a new type of loudspeaker design", *Journal of Acoust. Soc. Am.*, Vol. 73, No. 5, May 1983, pp. 1532-1536.

Rene Wooller &

Andrew R. Brown

Queensland University of Technology

Kelvin Grove, 4059

Queensland, Australia

renwooller@gmail.com

a.brown@qut.edu.au

A framework for discussing tonality in electronic dance music

Abstract

We present a framework for discussing tonality in Electronic Dance Music (EDM) which highlights how tonal techniques are employed creatively throughout EDM. While most musicological analysis of EDM focuses on rhythm, we contend that composer/producers of EDM play with a sense of tonality that often defies traditional analytic techniques. The rich tonality of EDM tracks may be illuminated by using our framework, which is broader and more contextually relevant to EDM than standard western classical music approaches to analysis. The framework consists of four attributes that are used to describe the nature of tonality over the course of the music: rate of tonal change, tonal stability, pitch/noise ratio and number of independent pitch streams. We will define and illustrate these attributes using numerous audio examples. In a practical sense, our framework may be useful for generating more detailed discussions and analysis of tonality in EDM and may lay the basis for formal systems of tonal analysis of EDM in the future.

Introduction

It is often observed that rhythm in EDM genres is of greater significance than in many other styles of music (Butler 2006: p 4-5; Neill 2002; Shapiro 2000) and, as a result, the growing body of musicological analysis of EDM appears to have overlooked the significance of tonality to some extent. This paper highlights how tonality, if conceived of more broadly than in western classical music, is employed creatively throughout EDM. This may assist in debunking the notion that the tonal elements of EDM and EDM are merely a simplification of the more mature and complex systems of tonality from classical music. EDM may appear simple at best, or incoherent at worst, through traditional tonal frameworks. However, our framework shows how EDM composer/producers can play with the same sense of tonality but in ways that often defy traditional analytic techniques. As a broad framework, no particular notational system is provided; moreover the concepts are expressed purely through sonic examples.

Defining Electronic Dance Music

We adopt a broad notion of EDM which includes any form of metric electronic music based on

loops and layers of different synthesised or sampled instruments and sounds. Overlapping styles include house, techno, electronica, downbeat, instrumental hip-hop, break-beat, and many others. While ‘pop music’ is not necessarily excluded from EDM, a number of the styles within EDM would clearly be outside of pop music, for example, minimal techno. Musically, the focus of pop music is more on vocals and lyrics, while EDM is defined here as being more concerned with rhythm and instrumental elements. Following the tradition of EDM (Shapiro 2000: p 77-78), a piece of music will be referred to as a ‘track’ rather than a ‘song’ partly due to the lack of traditional, lyrical style ‘singing’.

Tonality in a modern context

Tonality is defined here as the way a sense of tonic pitch is, or is not, suggested in the music and the way other pitches, if present, are organised with respect to it and each other to achieve particular musical effects. This is a slight broadening of the term originally coined by Féétis (in Reti 1958:7) due to the surprisingly diverse nature of tonality in EDM which does not always conform to the narrower definitions typically applied to classical music that deal primarily with chords and scales (Huron 2006:175; Reti 1958; Shenker 1935). While it would be possible for examples of EDM to be described purely in these terms, in a significant number of cases, if this were the *only* focus, there would be very little to say, despite the obvious widespread appreciation of the music. In a similar way to the low information classification of EDM as simply “4/4”, it would be a vast oversight to declare that “minimal techno is drone-based” and leave it at that.

Despite the overarching influence of a western harmonic heritage, EDM producers, simply through using new technology and often without formal music education, are less occupied with traditional notions of tonality. In such examples it is more revealing instead to explore atonality (the apparent lack of tonic), tonal ambiguity, the subtle introduction of tonality through intuitively non-tonal voices, or the harmonic properties of overtones and their manipulation within a single note. At the other extreme, one might imagine that unnecessarily complex tonal analysis may be prompted during passages where a synthesizer chord is played as though it were a lead

(musical example 1)^{*} – it would be simpler to treat the part as it appears to have been produced and perceived.

This would obviously involve a subjective judgement call, however the aim of this paper is to describe enough of the tonality of EDM so as to convey an idea of the range of techniques and possibilities within the genre. It should not, however, be construed as a theoretical attempt to authoritatively encapsulate all that is possible and denote precise generative likelihoods.

The current musicological literature on EDM (Butler 2006; Keller 2003) tends to revolve primarily around the rhythm, due to its importance; as discussed above. Literature on pop musicology does not often deal with “musical analysis” (Beard and Gloag 2005:11) so much as cultural theory (Hawkins, et al. 2007) and when analysis does occur, tonality does not appear to be discussed in detail (Tagg 1982). As well as this, the musical interest in pop music is primarily in the vocals and this does not adequately relate to the more instrumental (non-vocal), repetitive and drone oriented styles of EDM. The analytical musicology of acousmatic and electro-acoustic music also encounters problems of sonic analysis but to a much larger extent than EDM. While various approaches (Battier 2003; Hirst 2003; Windsor 1995) are somewhat relevant, there currently does not appear to be a framework for tonality that is suitable for EDM and so the following descriptive continuums were conceived, drawing from a variety of other music theories: rate of tonal change over time, the amount of recognisably pitched sounds, the level of harmonic coherence within the audible pitch set and the degree of polyphony.

We will explain these attributes and use them to define EDM. The objective here is to express the tonality of the EDM paradigm and should not be construed as an attempt at defining music for any other purpose. Throughout the explanation of terms, we will present key examples from EDM and apply the descriptive tonal attributes to them in order to build evidence for their suitability in defining EDM. As well as this, some sense of the genre may be expressed through the audible examples.

Rate of Tonal Change (Horizontal)

The rate of Tonal Change (TC) attribute relates to the level of activity within tonal parts – at one extreme, the entire track consists of a constant drone of tonic and/or pitch-class set without changing over time (musical example 2). At a level above this we might observe drones that shift pitch only once in a whole track or at the end of a lengthy cycle (musical example 3). A higher level of TC might involve typical chord

progressions in the bass line, such as the very common two-chord (musical example 4) or four-chord (musical example 5) varieties. Such progressions tend to gravitate to an underlying tonic (Bukofzer 1947 in Thomson). At a higher rate still, the bass line could form a riff that dances around an implied fundamental bass (Grant 1978) or “Urlinie” (Schenker 1980), an imagined bass line that can be reduced from notes over a span of time (musical example 6; musical example 7). Above this level, we might consider lead riffs which are changing in such a way and at such a rate as to contribute to ambiguity of the underlying tonic. This is typified by the “solo” (musical example 8). It should be noted that TC is derived from the sum of activity in the various pitched parts. For example in (musical example 9) three voices can be heard: the bass that doubles the kick, the mid-high register synthesizer fulfilling the role of bass, and the higher-register lead vibes, all different but adding up roughly to a mid level of TC overall – that is, the tonic and related pitches are not constant, but also are not so wildly variable as to confuse the tonality. Over the entire track the TC does not change dramatically. Having defined TC and provided an example of how it might be roughly gauged, it is now possible to examine how EDM overall can be described in terms of TC.

Overall, EDM is skewed more towards the “drone” end of the spectrum than the “solo”, with most tracks consisting of two, three or four primary chords in a progression and many consisting of a drone. The solo is, on the whole, a rare occurrence in EDM, although it occurs more commonly in related genres that are similar to pop music in terms of structure and emphasis on the lead-part for interest, for example, the excerpt above (musical example 8). These observations apply to whole pieces of music, whereas if the time span is narrowed onto a particular section, the level of TC may deviate drastically. For example, during a fill section, there is generally an increase in TC either through transformation of a pitched part (musical example 10) or addition of a pitched cue (musical example 11), while during a breakdown the opposite is often true, due to the introduction of sustained pads (musical example 12). In other instances ambiguity in the breakdown is partially conveyed through higher levels of TC in a kind of solo (musical example 13). The tendency for EDM to have low to mid levels of TC, to be more “drone” oriented than “solo” oriented, can be contrasted with classical music which, with continual variation and key modulation, has a relatively high level of TC. Some might argue that EDM in particular should be listened to at the macroscopic level of the DJ’s set and that at this timescale significant TC would occur. However, if one considers an orchestral work of the same length, it seems natural that the differences in TC between the two genres would remain. The broad genre of Pop music sits mostly in the middle, with complex lead ele-

* The musical examples referred to throughout this paper are available on the ACMC08 CD and online at www.lemu.org

ments and clichéd chromatic key shifts representing the upper boundary of TC and the more popular elements of EDM representing the lower boundary.

Tonal Stability (Vertical)

Tonal Stability (TS) is an estimate of how strong the sense of tonality (as tonicity) is, with primary reference to the tonic, but also to the idealised pitch schemas that the listener carries with them, for example, the minor and major scales. EDM has a mid-level of TS, but varies quite substantially. At the least stable end of the TC continuum, we could envisage pitches that do not suggest a particular tonic and do not relate to any of the scale intervals ever experienced by the listener. Above this, there may be recognisable intervals, but still no strong sense of tonic, as is often the case with whole tone scales. At the mid level, a tonic would be identifiable, but many of the other pitches may be accidentals or extraneous scale degrees that are less fundamental or less "similar" to the tonic. Above this, the tonic may be forcefully emphasised, featuring fundamental intervals such as the fifth, fourth and octave more strongly. The extreme of TS would feature only the octave.

One might ask: why are the intervals of fifth, fourth and octave given such a fundamental role in establishing tonality? Empirical qualitative research supports the claim that they are judged as "stable" and "strong" in musical terms (Huron 2006:145). The special fundamental role of these intervals is also apparent in the musicology of most other civilisations (Thomson 1999). These intervals are readily perceived as being similar on neurological (Weinberger 1999) and cognitive (Krumhansl 1979) levels. This also extends to the chords I, IV and V (Krumhansl 1983). On a physical level, the ratio of 2:3 (the fifth), produces a shorter macro-cycle ($2^*3=6$) between the two frequencies than any other ratio below 1:2 thus could be considered as something of a "best fit" on a physical level – while perfect ratios only exist in metaphysical realms, small inconsistencies in tuning, 2.03:3 for example, are typically overlooked during tonal perception, an argument made earlier by Theodore Lipps in 1900 (Thomson 1999:89). Due to all of these musical, neurological, cognitive and physical reasons we consider music which features the octave, fifth and fourth intervals to have a higher TS than other intervals.

In addition to the perfect intervals, pitch-class sets which are familiar to listeners will appear to have higher TS than those that are not. There is a lack of evidence for any particular scale with non-perfect intervals being more or less intrinsically viable from a musical perspective – that is, scales appear to be learnt (Thomson 1999), which is not to deny the evidence of certain affordances of the human mind to guide this learning, for example our propensity for five to nine discrete categories (Baddeley 1994). Because

of this, and from experience of extended exposure to unfamiliar pitch-class sets, it is reasonable to assume an increase in TS for any pitch-class set over time.

Within scales themselves, pitches have particular functions and can add to the TS by reinforcing a familiar pitch schema. Correlating surface pitches with **key profiles** is one way to assess the TS of a musical sequence and this formal approach is explained by Temperly (2007:53). Accidentals and pitches that are outside the dominant tonal schematic will also reduce the TS if they occur more often than is typical. This is supported by empirical music psychology studies which found that people have a notion that certain pitches fit a tonal context much better than others (Huron 2006:148), the individual judgements being averaged into a key profile.

As mentioned, EDM is considered to mostly have mid-level TS. This is justified because, in the vast majority of cases, there is a clear tonic, regular scales are used; most commonly pentatonic minor, followed by minor and Mixolydian (major with flattened seventh). Tonal movement in chord progressions is often between I and V if binary (musical example 14), or progressions that include V or IV if ternary or quaternary (musical example 15). Sequences with lower TS have less "perfect" intervals in their bass lines or a copiousity of unfamiliar accidentals (musical example 16). Tracks with higher TS are pure monotonic drones which span multiple octaves (musical example 17). As with the other attributes mentioned, TS is dynamic, often changing during fills and breakdowns. As argued by Huron (2006:160-161) and others, the sequence of pitches also contributes to stability, however the details would be a distraction to this current discussion. Nonetheless, the principle is exemplified here (musical example 18) where a random-walk and arpeggio are played together, outlining a pitch-class set, but not assisting in the definition of tonic.

In comparison, classical music can be considered to have mid-level TS for similar reasons, but deviating towards less TS rather than more, particularly when considering the more recent periods of tonal complexity. By way of contrast, pop music has mid-high TS due to the prevalence of standard scales, chords, I-IV-V and fifth based progressions.

Pitch/Noise Ratio

The clarity of pitches has a direct effect on the ability of the listener to develop a sense of tonality – for example, in the case of total noise where there are no discernible tones, it is impossible to conceive of the TC (rate of tonal change) and TS (tonal stability). As a result, the Pitch/Noise Ratio (PNR) is considered here to be a relevant attribute of tonality, particularly for electronic music, which has always involved a significant amount of sonic expression. The con-

tinuum can be envisaged with purely untuned and/or distorted percussive sounds and noises at the lower end (musical example 19). The highest PNR is music made from pure tones.

EDM overall has a mid range of PNR, but varying substantially between sub-genres and individual tracks. In particularly minimal instances, the percussive sounds are usually tuned in some way so as to suggest a basic tonality, or there is a very subtle application of tones, for example, the high-hat and kick (musical example 20). In other cases, sound effects such as ring modulation are used to introduce tones (musical example 21). In contrast, down-tempo artists such as Board of Canada are known for their rich tones (musical example 22), although in the main sections these tones are usually accompanied by unpitched drums. Boards of Canada often “detune” their synthesizers, which provides a distinctive sound and does not obstruct the identification of tones. However, some other forms of pitch shifting can disturb pitch clarity and thus would have to be considered as having lower PNR (musical example 23; musical example 24). Despite this, it should be noted that foreign and abnormal tuning systems are sometimes used, and these are not considered as having any less PNR due to the tones being quite perceivable (musical example 25; musical example 26). A temporary decrease in PNR is often observed during fills, breakdowns and transitions, the dissolution of tonality being associated with increased tension or intensity. For example, DJ Shadow reduces the PNR through a record slowdown (musical example 27).

The mid PNR of EDM can be contrasted with the high PNR of classical music; mid-to-high level of PNR in pop music; and the low level of PNR in acousmatic and electro-acoustic music. This is justified as most orchestral voices have a distinct pitch, including some of the percussive parts such as timpanis and triangles. In pop music, there is a heavy emphasis on tonality and pitch clarity and more conventional use of sounds than in EDM, mainly due to more conventional instrumentation and less emphasis on electronic media to assist expression. The sound-objects used to compose acousmatic and electro-acoustic music are often not easily recognisable as clear pitches and so have a low PNR. The PNR describes the clarity of tones for a given piece of music, while the TS and TC describe how these tones are organised to effect the tonality.

Number of Independent Pitched Streams (IPS)

The number of Independent Pitched Streams (IPS) relates to the number of pitched voices that can be identified as operating independently and simultaneously. At the lowest end of the continuum is a single pitched voice/part, at the highest end is a dense texture built from numer-

ous voices and in the centre is the typical three to five part tonal voicing of EDM and pop music. Usually there is one bass, one or two leads, and one or two accompaniments. Classical orchestral music can be distinguished by a high number of IPS. It should be noted that while a particular number of IPSs might be clearly identifiable through careful analysis, it is more common for listeners to concentrate on a single stream or texture at a time (Bregman, 1990).

While mid-level IPS is typical in EDM, there is often deviation from this, sometimes with extended periods of none (musical example 19), one (musical example 20), two (musical example 28) and three (musical example 29) or more pitched voices.

A subjective judgement call is sometimes needed to determine whether a part contains multiple streams or not. As shown by Bregman (1990), a single sequence of tones, if played with alternating pitches that are related beyond a certain interval, it will be more likely to be perceived as two separate streams, albeit with the focus generally on one or the other. Alternatively, a synthesizer chord that always consists of the same intervals in parallel might more easily be classified as a single stream of an interesting ‘chord-like’ timbre (musical example 1).

Conclusion

In summary, a framework for discussing tonality in EDM has been presented and applied to a brief synoptic analysis of tonality in EDM, using a number of audible examples. Four terms have been introduced: TC, TS, PNR and IPS. TC relates to the rate of tonal change, from drone through to solo. TS relates to the stability or strength of the perceived tonality at any point in time, from the strong sense of fundamental tone brought about by octaves and perfect intervals through to the vaguer sense of tonality implied by accidentals and atonal pitches. PNR relates to the clarity of the pitches, from pure tones to noise. IPS relates to the number of independent streams that it is possible to discern.

Each of these descriptive dimensions are independent of one another and all of them relevant to the unique sense of tonality that may or may not be instilled in the individual listener of EDM. While some concepts, such as stream segregation, have been studied previously, the framework we present here appears unique in its combined application of concepts to EDM, particularly in the context of tonality.

It should be noted that exploration of this framework for tonality in EDM is at an early stage – future research would involve more precise definitions, ultimately to the point where the framework is formalised into a generative theory that could be executed on a computer.

References

- Baddeley, A. 1994. "The Magical Number Seven: Still Magic After All These Years?" *Psychological Review*. **101**: 353-356.
- Battier, M. 2003. "A Constructivist approach to the analysis of electronic music and audio art - between instruments and faktura". *Organised Sound*. **8**: 249-255.
- Beard, D. and Gloag, K. 2005. "Musicology: the key concepts". Abingdon, Oxon, Routledge.
- Bregman, A. S. 1990. "Auditory scene analysis: the perceptual organization of sound". Cambridge MA, MIT Press.
- Butler, M. 2006. "Unlocking the groove". Bloomington, IN, Indiana University Press.
- Hawkins, S., Scott, D. and Sweeney-Turner, S. 2007. *Pop musicology online*. accessed on 29th of May, 2007. site: <http://www.popular-musicology-online.com/>
- Hirst, D. 2003. "Developing a cognitive framework for the interpretation of acousmatic music". *Converging Technologies: Australasian Computer Music Association Conference 2003*, P. D. Andrew Brown, David Hirst, Tim Kreger, Gordon Monro, Robert Sazdov, Lindsay Vickery. Western Australian Academy of Performing Arts at Edith Cowan University, Perth, Australia, Australasian Computer Music Association.
- Huron, D. 2006. "Sweet Anticipation". Cambridge, Mass., MIT.
- Keller, R. 2003. "Mapping the soundscape: rhythm and formal structure in electronic dance music". *School of Music*. Tallahassee, Florida State University.
- Krumhansl, C. 1983. "Perceptual structures for tonal music". *Music Perception*. **1**: 28-62.
- Krumhansl, C. 1979. "The psychological representation of musical pitch in a tonal context". *Cognitive Psychology*. **11**: 346-374.
- Neill, B. 2002. "Pleasure beats: rhythm and the aesthetics of current electronic music". *Leonardo Music Journal*. **12**: 3-6.
- Reti, R. 1958. "Tonality, Atonality, Pantonality: A Study of Some Trends in Twentieth Century Music". Westport, CT, Greenwood Press.
- Schenker, H. 1935. "Der Freie Satz (Free Composition)". New York, Longman, 1979.
- Shapiro, P. 2000. "Modulations: A History of Electronic Music - Throbbing Words on Sound". New York, Distributed Art Publishers.
- Tagg, P. 1982. "Analysing popular music: theory, method and practice". *Popular Music*. **2**: 37-67.
- Temperly, D. 2007. "Music and probability". Cambridge, Mass., MIT.
- Thomson, W. 1999. "Tonality in music". San Marino, Everett Books.
- Weinberger, N. M. 1999. "Music and the Auditory System". *The Psychology of Music*, D. Deutsch. San Diego, California, Academic Press: 47-83.
- Windsor, W. L. 1995. "A perceptual approach to the description and analysis of acousmatic music". *Music Department*. London, City University.

Dr Ros Bandt

Honorary Senior Research Fellow
 Director of the Australian Sound Design Project
 online data base and gallery
www.sounddesign.unimelb.edu.au
 The Australian Centre,
 School of Historical Studies
 The University of Melbourne
 137 Barry St
 Carlton, Vic 3053
 Australia
r.bandt@unimelb.edu.au
www.rosbandt.com

Abstract

Finding the Isobue is the journey I made with the Japanese Environmental activist, Dr Kumi Kato to the remote island of Sugashima in Japan in search of the endangered sound the isobue, the Japanese sea whistle. This sound is made by the Ama women who free dive for abalone and sea creatures maintaining a sustainable fishing practice they have preserved since the time of The Pillow Book. This paper articulates how this significant sound has been interpreted into a number of spatial sound works from original field recordings, radiophonic art, bi-lingual sonic haiku, electro-acoustic music for piano, koto, flute, sound sculptures with soundscape and the electronic art 5.1 surround sound installation , a quiet space for contemplation of a disappearing world. The double CD Isobue celebrates the significance of this sound and is dedicated to the Ama.

Introduction

I met Kumi Kato in Canberra while working at the National Museum on sounddesign of public spaces. She mentioned finding a hydrophone to capture the water environment of the Ama's sea whistle sound. In two weeks we were there together in search of the sound and loaded with 40 kilos of sound equipment, 2 DATs, hydrophone, stereo air microphones and backup minidisk. Living in the village of Sugashima, we shared the lives of the fisherfolk women and my main aim was to help Kumi, record the information she needed for research. Listening deeply to the rural Japanese and spending time in the coastal regions was an immersive window of listening. I wrote haiku everyday, like a diary. Together we recorded some 10 CDs of site specific recordings, festivals, Japanese stories, songs, dances and rituals which surround the act of monitored fishing, a system of ritual connection they have maintained for centuries.

When I returned home I realised the cultural value and creative potential of the material we had gathered. An ABC Composers residency with ABC Radio National Sydney, made possible some creative spatial sound designs, a radio-

Artist Talk:

Finding the Isobue: From the Hydrophone to the Radio studio to the Gallery; spatial sonic designs of an endangered sound

phonics story for Radio Eye, *Waiting for the Tide*, a suite of traditional music with soundscape, *Sea-Folk Voices*, the 8 electro-acoustic sonic haiku, *Shima*, the short electroacoustic Vignette, *Ama no Isobue* and the substantial 5.1 installation *Isonageki, Sea Lament*.

Trapping the Sound

How to trap the Isobue when it is made under very strict practices deep in the sea? Many of the younger women can not make it? What about the imperialism of the microphone? Kumi and I spent long hours seeking permission and sharing the stories with the blessing of the women at each site before we began. Kumi helped lug the gear on boats, up cliffs, over rocks, into the sea after her careful negotiations. Taking the weight of the metres of hydrophone cable in a deep moving sea to a depth of 10 metres is a job for two. She also swam with the hydrophone in the cold waters.

In a recent article, Recording for Experiential Flexible Sound Installation, 1977-2007 published in the ASRA Journal 2008, no 34, I have noted nine different types of recordings made to effect these different celebrations of an intangible cultural heritage.

On location in Japan these included

- * hydrophone underwater and air recordings,
- *180 degrees stereo mike into DAT for soundscape environmental recordings, the village night, early morning at the fisheries auction,
- *Japanese interviews with key women divers and fishermen,
- *90 degrees directional recordings for detailed crabs and sea creatures on the rocks, insects,
- *omni-directional dummyhead surround sound stereo recordings on the rock during the women's diving and coming up for air and emitting the whistles in a community sea location.

Creating the Sound Palette

In the ABC studios, I used techniques which would serve my creative imagination for the new works to be made. Anything was possible with Andrei Shabunov's help. I amassed a huge array

of recorded sounds and modules which would aggregate as a sonic palette to design the simulated underwater environment, the world of the women to situate their sound.

*the anechoic chamber was used to carry out many experimental recordings of situational water, whistle simulations and instrumental samples of psaltery, Satsuki's Koto, and recorders and flutes.

*the interior and keys of the slightly out of tune Bosendorfer grand piano in studio 256 was a world of underwater danger as well as a wonderful instrument for live performance of *Aominesun*.

*Studio Vocal recordings were made of my haiku in Japanese and English as well as translations of Japanese texts to be used for poetic reflection in the Radio story.

*Original samples of my glass and clay sculptures, the Flagong, Lorraines' ceramic bells, wooden and metal percussion, shells, windchimes, wooden buckets, Japanese paper, and strings were also recorded in the Melbourne studios with Richard Girvan and taken up to Sydney as files. Elaborate submixes were made weaving computer effects on ProTools FX and Wavelab. These were overlayed and dovetailed into complex modules with thematic titles, abalone shell, whistle, bucket, rack, deepsea ready for diverse combinations in the haiku, the installation and radiophonic text underlay.

Design Morphologies and Processes

Different sonic architectures and processes are required for different outcomes. Each one was made with a specific aim in mind, a radio story, a poetic bilingual talisman, an instrumental suite, an Immersive forty-three minute sound installation in 5.1. a short summary electro-acoustic trip. For each the sound files and submixes were collated and roughed out in Wavelab to experiment further. In the Radio Eye piece, a text narrative structure in two languages dominated the flow as readings and translations on many themes had to be fitted together. Kumi and Sharon worked this through in consultation with the rest of us while Andrei and I concentrated on the sonic tapestry to hold the words, underpinning them, obscuring them, enlarging them, drowning them. In the Haiku, the text had a different role, the English defining the audio content and Kumi threading the Japanese response through the tiny sonic vignettes made especially for each one. The *Sea Folk Voices* was intended as a musical suite so the soundscape recordings were subservient to the music written and recorded in two pieces, *Sea Folk Voices* and *Husa*, and improvised in two. The piece *To the Sun Goddess* was a straight recorded improvisation between Satsuki Koto and myself on medieval psaltery while *Aominesan* was a solo piano piece with soundscape providing the atmosphere of the sacred mountain the Ama pray to before they dive.

By far the densest work was *Isonageki*, Sea Lament a 5.1 surround sound mix where all the electroacoustic mixes were designed in a constantly changing polyphonic spatial array resembling an underwater environment. *Isobue* CD2 track14. This was mixed in Pro Tools in Melbourne with Richard Girvan. Elaborate scores of the interrelationship of parts to the whole were drawn to communicate the movement and morphology of continuity for all the pieces. Detailed naming of all the files and mixes allowed us to float many materials in different configurations and densities, like the ocean, ebbing and flowing over multi-layered on site recordings of Sugashima, Shirongo, Mendonohama and Champagne beaches. Detailed hydrophone recordings of elephant sharks eating pippis recorded in the Melbourne aquarium were melded with polyps from Pope's Eye in Port Phillip Bay and the Japanese waters. *Isobue* CD1, Track 13.

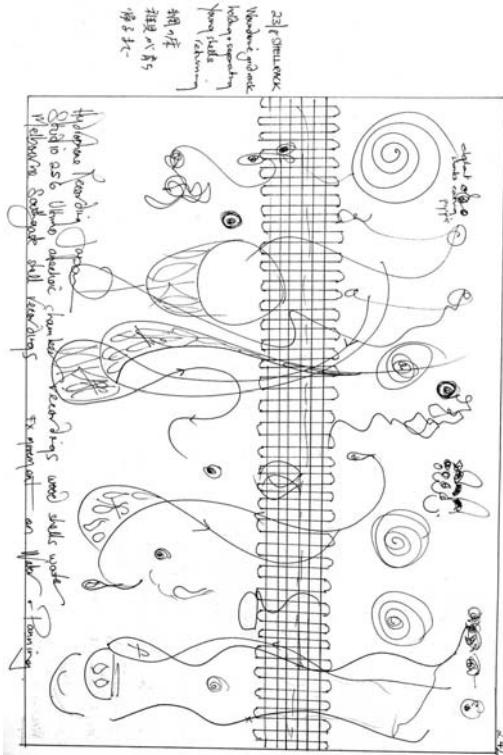


Figure 1. Score for Haiku Shell Rack

The electroacoustic soundscape was a global mix to recreate and evoke the Japanese underwater environment where the women risk their lives. The reconstruction of a pseudo natural environment full of beauty, chance and danger is one of the most fulfilling designs I have worked on to date. The spatialisation took days after the initial composition layout and Richard and I worked tirelessly on this intuitively together, carving time as shared experienced human moments and gestures.

Organising the massive shifts and dramaturgy was made possible in a dedicated space especially set up with the speaker array so everything was audited at the moment of creation.

Some of the files and scores for this work appear below.

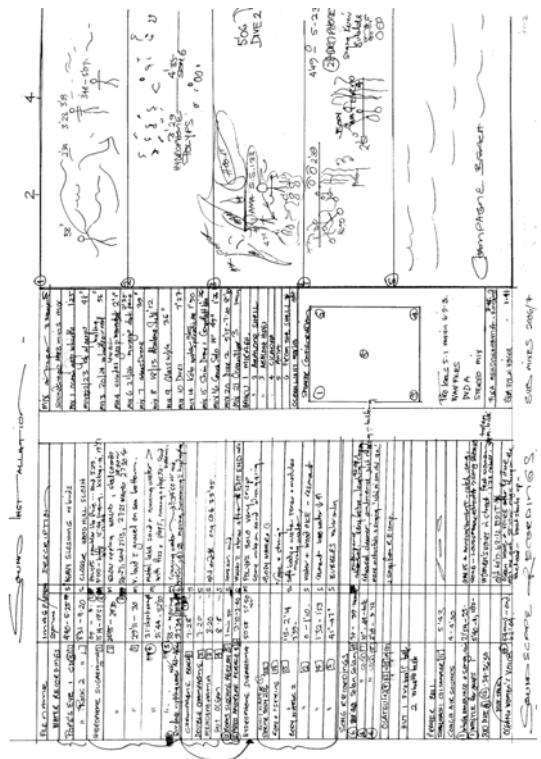


Figure 2. Installation Score *Isonageki*, files and map.

Back in Sydney at the end of the project, Andrei and I made *Ama no Isobue*, a short summary trip through this long journey, made comprehensible in a few minutes while retaining essential elements of the women and their sea whistle, the sound that orients them in the ocean to the location and each other, communicates whether their dive was successful in obtaining large abalone or just a sea cucumber, whether they are well, tired, eager for more. It is a sound which is tribally shared and individually uttered containing their life's breath in the face of danger. When there is no text, the meaning is conveyed through the electro-acoustic sound, its density, timbral amelioration and rhythmic unfolding alone.

Spatial Design for Exhibitions

Going into the Gallery with this audio requires visual design and spatial understanding of the acoustic space itself, the flow line of the spectators, the requirements of speakers to wall, floor ceiling surfaces and the general visual and architectural characteristics of the building's size, shape, volume, and surfaces. Every space is an acoustic space and the sensitive positioning of sounds in each discreet space is a skill required of the installation artist.

Sound installation is continuous and open ended. The audience get to define the length of the stay and involvement with the work within the limits of the gallery opening hours. If the en-

vironment is friendly with seating or floor coverings encouraging longer stays, it is more likely the work will be better audited. Frequent visits are also an option when works are installed for longer periods such as days and weeks.

The *Shima, 8 Haiku for Kumi* (Isobue CD1 8-13) have been exhibited in a group show at the Counihan gallery, Brunswick focusing on embodied energy and sustainable environmental practices curated by Edwina Bartlem and Penny Algar. Headphones were chosen to isolate the listeners' attention from other gallery sounds and framed parallel texts of the eight Haiku, were hung either side of the blue silk falling from the ceiling representing the womens dive. Illuminated shells resting on the rice cooker, reminiscent of the divers wooden bucket, the oke, bring the future domestic allusion to the piece, that the abalone will be sold to eat in the end. Four sets of headphones were chosen at four listening ports under the text, each one containing two haiku on loops. The longest length for each box was between two and three minutes, a good length for standing room only and the brevity of the texts.



Figure 3. *Shima 8 Haiku for Kumi*, Counihan Gallery Brunswick, June 6-29, 2008.

The response has been heartening with four people at once often listening side by side and sharing their thoughts of the listening world.

Designing the 5.1 *Isonageki, Sea lament* is a much different experience. As the theme for the conference is sounds in space it was decided that a simulated underwater environment as an immersive and contemplative surround sound installation would be an appropriate experience for the electroacoustic community. As seasoned listeners they could appreciate the evanescent 43 minute work spatial composition.. The space needs to be acoustically discreet so the contemplative qualities and sensitive perception of the moving sound masses can be fully sensed. Lighting, 5.1 speaker array with sub woofer and suspension opportunities have been requested. The fifth speaker for the 5.1 should be centrally overhead to articulate the vertical dimension of the sound as movement in a deep sea envi-

ronment. The women's spatial path diving is a mainly vertical one.

Later this year, *Isobue* will be adapted to create a quiet space setting at the University of Queensland for Kumi Kato's Peaceful Space conference in September. The sonic creations will also be taken back to the Ama women in Japan as handwrapped CD gifts. A cassette will be made for their tiny new Ama museum in the small village where technology is not really part of their lives and a simple portable CD player gifted for the repeat play scenario for their artefact display. Headphones or /and speakers will be decided in situ. They are so proud NOT to have technology. It must be kept simple.

In the double CD *Isobue*, the sea whistle as intangible cultural heritage has been celebrated through elaborate digital means and processes, available through large corporations like the ABC, in association with artists, the environmental group ecco and the sonic gallery of the Australia Asia foundation. Part of the proceeds of the disc made on recycled paper will be gifted back to ecco for creative conservation through art projects of this nature. In January the two notated pieces *Sea Folk Voices* and *Husa* will be published in handstitched composer autographed notations and taught at a summer school encouraging recorder players to interpret space notation with electroacoustic soundscape. A new piece *Free Diving* for recorder orchestra, electroacoustic soundscape and graphic notation was premiered in Melbourne in March using extended recorder techniques and the breath in all parts of the disassembled instrument. Through contact with the Isobue, many people are reclaiming the sanctity of life's breath.

In the end it is the quality of the communication between people, the reasons for their lives and the presence of life's breath which is infinitely more important than technology of any kind. Listening and breathing are our first and last experiences in being alive. They validate our presence on earth. The Ama are in tune with the natural environment through listening, and their special sound the Isobue, a rare experience these days. They are immersed in ancient sea culture and are responsive to its every moment. They are *Waiting for the tide*.

Shima: 8 haiku for Kumi is dedicated to Kumi Kato
The double CD *Isobue* is dedicated to the Ama

Thanks to:

Kumi Kato for inviting me to share the search for the *Isobue*, translator, narrator, recording assistant, Haiku responses, project manager, environmental researcher, Japanese minder.

Robyn Ravlich for the ABC Radio National Sound Artist in residence and the New Media Arts Board, of the Australia Council, Australia's national arts funding body.

Thanks to the collaborative *Radio Eye Team*, Sharon Davis, Kumi Kato, Ros Bandt and Andrei Shabunov, Readers Tony Barrell, Satsuki Odamura, reader and koto, Anne Norman, Shakuhachi.

The engineers Andrei Shabunov, Richard Girvan, Alex Stinson, Angie Grant.

Art direction and design of CD: Joseph Griffiths

C Ros Bandt, 2008. Sound Artist, Japanese soundscape and field recordings, composer, poetry, performer (psaltery, flutes, recorders, Bosendorfer piano, percussion, original glass sculptures, conch, air whistle, shells, glass, clay wood, and brass percussion, voice), Surround sound immersive installation artist.

References

- Bandt, Ros. 1985. *Sounds In Space*, Melbourne: Victorian Arts Council.
- Bandt, Ros. 2001. *Sound Sculpture: Intersections in Sound and sculpture in Australian Artworks*. Sydney:Craftsman House.
- Bandt, Ros. 2008 "Recording for Experiential Flexible Sound Installation 1977-2007", *The Australasian Sound Archive*, No 34 ASRA, edited Bradley and Jatania, 18-33.
- Kato, K. 2007. "Waiting for the Tide, Tuning In the World-Ama no Isobue: Soundscape of Abalone Diving Women", *Hearing Places*, edited Bandt, Duffy and MacKinnon, Newcastle UK: Cambridge Scholars Publishing, 213-233
- Maraini,F. 1962. *Hekura: The Diving Girls' Island*, London: Hamish Hamilton.
- Martinez,D. 2004. *Identity and Ritual in a Japanese Diving Village*, Honolulu: The University of Hawaii Press.
- Morris,I. 1967. *The Pillow Book of Sei Shonagon*. Middlesex:Penguin.
- Tanabe,S 1993. *Ama*,Tokyo: Mitsui Shobo..

Brigid Burke

Clarinetist, composer and visual artist
 brigid@alphalink.com.au
<http://www.brigid.com.au>

Abstract

A series of performance works will be presented that explore the processes of

integrating sound and visual counterparts. 'Frenetic Illusion', 'Strings' and 'Memories of a Shadow' for clarinet, live interactive audio and visuals (DVD).

The works explore the sonic vocabularies of extended clarinet acoustic micro-tonality techniques, interactive mapping audio devices and visual components. The performance works are an interaction between the performer, computer and real-time digital audio and visual devices. The interactive audio techniques used are pitch-shifters, frequency changes, room placements and granulation. All these filters and parameter modulations can be controlled in live performance using a mapping software device. The visuals go through similar processes as the audio samples but are pre recorded and include video footage of dancers, photos and drawings.

Introduction

The aim in 'Frenetic Illusion', 'Strings' and 'Memories of a Shadow' is to make sound and image structurally integrated. To achieve this integration in performance of these works, the audio is analyzed and used directly to control the manipulation of specific aspects of the audio guided by the fixed visuals. The real energy of this idea comes from combining the strength of the interaction, real-time processing and sound/image linking and mapping into the singular work that explores all forms of expression. When linking the music and visuals I question whether there are any real correspondences between sound and vision. The computer has helped this interaction especially in my recent works. As Kapuscinski quotes, "Even an unlikely collision of sound and image can cause both of them to be evaluated with equal attention. An idea that is very relevant to my work. It may even combat the usual dominance of sight and hearing. He also notes that it is not the equity between the media forms but how different the media interacts. In mapping audio through various programs it has given me endless possibilities that I can vary from piece to piece. I am using the live interactive audio program AudioMulch and software developer that allows me to facilitate this flexibility and allows me to address mapping in a modular way that is easy to reconfigure throughout performance.

Artist Talk: Integrating Sound and Visuals in Performance

Process

When creating these works I examine the media elements I am going to use in the composition/improvisational elements. I think about basic audio elements: sonic realm, amplitude (volume), pitch, timbre (tone quality), duration, tempo, rhythm and density. I then take these forms and add extended clarinet techniques such as (micro-tonality, voice, key clicks, multiphonics, monophonics, quarter tones, over-blowing and interrupted tones) and filters such as (pitch shifters, reverbs, flangers, room placements, harmonics, sine waves, ring modulators, delays, phases, granulation and EQ.) The process continues with manipulation of files into different layers and multi channels, concentrating on microtonal interaction between the samples. A similar process is applied to the visual materials including analysis of brightness, colour, contrast, duration, speed and complexity. The images have two categories: graphic based images and film/still images. The sound and image influences the shape and analysis of each of the works. The audio in the compositions uses a real-time environment of acoustic sound and generative structures.

The other added facet is to combine live acoustic clarinet. Audio Mulch controls the modulating parameters (for example pitch shifters, granulators, phases, loops, switches) controlling the amount of dynamics, on and off switches and loops during performance on the clarinet. A pressure pad controls the computer, which is situated under the thumb set of the clarinet, which is attached to a Pedal Midi Controller Box. The Pedal Midi Controller Box is a device that controls the selected Midi parameters in the computer in real time. At the moment I am controlling each sample manually reacting off the visuals.

Compositions

Memories of a Shadow

A visual and sound composition, using shadows of a figure (Omar Rigo) moving through confined spaces and language that will depict memories and snap shots of individual's dreams. Sigmund Freud's, book of 'Interpretation of Dreams', inspires the text quoted throughout with the famous dream of the Guillotine by Maury. It explores inner thoughts of identity, acoustic and manipulated language to create

sound, movement, performance, line and colour. These ghost-like shadows appear and recede in a dark space depicting the wanderings and shapes of figures. The outcome is an entwining of audio murmurings, drawn shadows, limited edition prints (silk screened), and mixed-media art works on canvas and video footage. The audio utilises spacialization and morphing of Maury Freud's Dream as it comes to life using acoustic and modern filter techniques including ring-modulation, vocoding, flanging, multi-channelling, granulation and equalization. The opening is a mixture of female and male voices telling us of the dream. The speech is interspersed with speech rhythms and ring modulators that create hollow and bell like sounds with slight reverbs and a small delay added to make the whispers inaudible. One hears the air sounds at ends of phrases that are granulated. Single words in the middle section are fragmented – slight delays are added and microtones and multiphonics are played on the clarinet. These clarinet sounds are treated with pitch shifters, spacialization and reverberant harmonics. Throughout the work the granulation of the text occurs and as a whole is stretched, compressed and multi-channelled. The textual fragments are reshaped, spiralling in and out of understanding coherency and audibility. This fragment of text can never be viewed as a fixed object as the content is hundreds of years old and has been handed down through generations by oral poets.

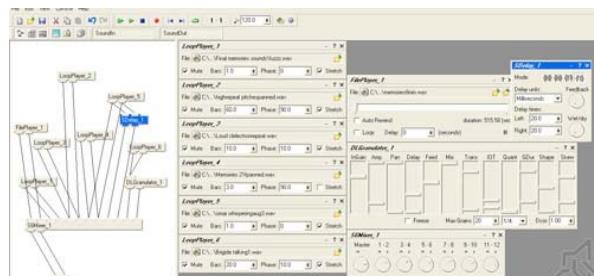


Figure 1. Example of AudioMulch Patch for 'Memories of a Shadow'



Figure 2. Memories of a Shadow stills from DVD

Strings

The concept of the composition is to transform the acoustic prepared piano to another timbral plane of textural colors. The sounds and transformations came from images of strings, wiring systems that connect urban cities and how strings and manifestations of these communication systems connect our lives. The visuals are heavily manipulated in various ways to create surreal imagery creating density of the wiring systems. The source of these images is from a

series of silkscreen prints I created in 2006 based on the wiring systems of Inchigaya in Tokyo Japan. The video footage is boiling pots of string, spaghetti and food coloring. The transformation of the sounds aim to match the sounds is to match timbres to reflect a rich canvas of sonorities around similar pitches and rhythms. These are all interwoven with live electronic and effected clarinet textures.

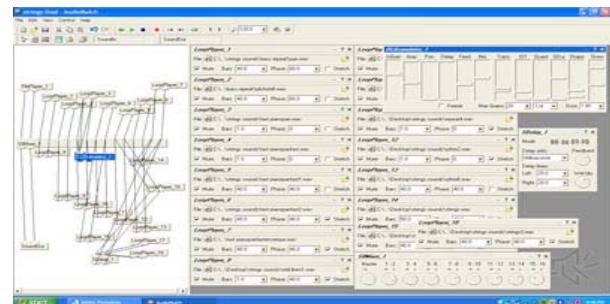


Figure 3.Example of AudioMulch Patch for Strings

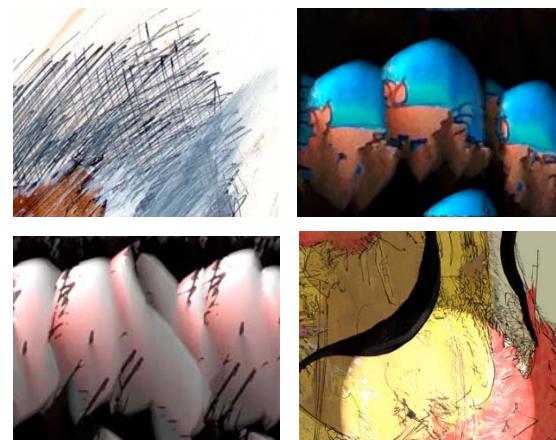


Figure 4. Strings stills from DVD

Frenetic Illusion

This work is an integrated sound and video work, which fuses dance, layered digital images, film footage and live acoustic and digital sound. The dancer (Sela Kiek) merges with my line and ink drawings and paintings. These multi-layered projections ripple and distort as the dancer moves in and through hanging cloth, reflecting the activities of insects or animals, constructing, burying, fixing and struggling to control its fragile environment. Frenetic Illusion explores the ephemeral, fragile, natural world and its ability to create complex structures that both sustain and ensnare its inhabitants. The piece echoes the natural cycles of growth and decay as the music, dancer and layers build and pulls apart sculptural forms with the cloth. The sound-scape echoes the frenetic activities of construction and degeneration found in nature at the microscopic level. Frenetic Illusion was developed into a live dance, sound and video work to be performed in July 2007. The visuals were projected on a huge web-like cloth structure suspended by ropes, as

though emerging organically from one corner of the space itself with images of the projected dancer. Working with the visual imagery and with the fabric used for projection, Sela created a series of phrases and structured tasks to produce movement, and the character of the duet began to emerge. The digitally manipulated sound created in rehearsal also influenced the development of themovement vocabulary. This project has provoked both of our interests in the connection between music and dance in live situations and experimentation of both artists moving together, sharing the same space.

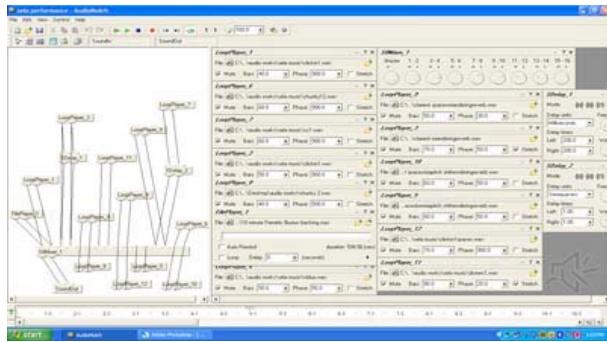


Figure 5. Example of AudioMulch Patch for Frenetic Illusion

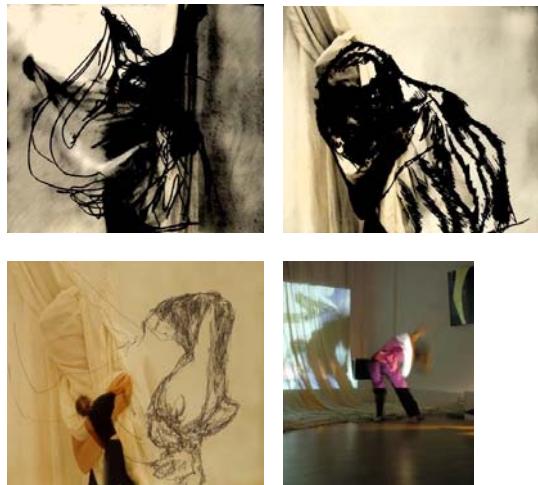


Figure 6. Frenetic Illusion stills from DVD

Conclusion

Interaction with the visuals and sound through analysis and mapping offers new opportunities and challenges that deserve original and creative application. The conclusion I can only draw from the performance works I will present, is currently the interactive performance applications available cannot compare with the practice of years of training. Immediate reactions and vocabularies are hard for the computer artist to compete with as technology is forever changing giving more scope for artists to work with, but also creating more problematic situations to cope with in performance.

Acknowledgment

"This project has been assisted by the Australian Government through the Australia Council for the Arts, its funding and advisory body."



References

- Kapuscinski, Jaroslav Compositions with sounds and Images, http://www-crca.ucsd.edu/njareh/ideas_top.html
- Rush, Michael. Video Art Thames & Hudson
- Eldridge, Alice. Fond Functions: Generative processes in live improvised performance.
- Woolman, Matt. *Sonic Graphics Seeing Sound* Thames & Hudson

(Lisa) Chi-Hsia Lai
 Australian National University
 School of Art
 121 Childers Street
 Canberra, ACT 0200
 Australia
 chi-hsia.lai@anu.edu.au

Abstract

Hands On Stage is a performance interface designed from a percussionist's perspective. It explores the inter-relationship between the performer's hand movements, real-time video tracking techniques, and the recording and processing of sound in an integrated configuration. This paper describes the concept, interface design, technical development, and the creative approach to the visual and sound elements. The paper also discusses the project from a personal performance perspective.

Introduction

The intention of the project was to seek new performance approaches that would sit between two practices, percussion performance and new media art in a concert performance context. The initial research for the project has made me realise that many current experimental and new media concerts involve the integration of audio and visual domains, while the world of percussion encompasses a wide variety of performance practices and an extensive range of instruments. I decided to combine the two practices and build my own instrument, which would take percussion performance practices into another realm where the performer could have control over audio and visual domains with hand movements.

Hands On Stage was designed to be used as a tool to explore the relationship between the performer's movements, sounds produced by the performer and projected images. This interface has enabled the communication and interaction between three domains: hand movement, sound, and visual elements. In the aspect of sound design in relation to hand movement, the objective was to move beyond deterministic results and explore ways of controlling and mixing multiple sound elements in a real-time performance context. For the visual aspect, the projected images are simply captured and processed with video effects and projected in real-time.

Hands On Stage is a table-like interface. In creating the interface a number of areas, including human computer interface design, motion tracking, mapping, and contemporary performance practice were explored. Recent projects of interfaces designed for controlling and processing sound and/or image have included the *re-*

Artist Talk: **Hands On Stage: An Interface for Sound and Image Improvisation**

actable and the *Morphing Table*, both designed for collaborative multi-user music making. There is also an interactive audiovisual installation, *Sound Rose*, which consists of a touch sensitive table and a projector. While these projects may use purpose-built table-like interface, the technology and the objectives are different.

This paper begins with an overview of the conceptual and technical development of *Hands On Stage*, and followed by a description of the configuration of sound and visual elements for an improvisation titled, *Untitled #1*. Finally this paper ends with a discussion of the project's future directions.

Hands On Stage

This section discusses the creation and the development of the interface, *Hands On Stage*.

Concept

The initial purpose of the project was to explore and enable the interaction between performer's gesture, and real-time sound and image processing to form an audiovisual performance approach from a percussionist's perspective.

Early experimental stages included looking for suitable materials and technologies in order to construct a table interface. Although sensor technologies could probably offer a more precise response to the performer's actions, it was my intention to simplify this part of progress. More time was spent in exploring the performer's hand movements as a control to sonic and visual outcomes. As a result, motion-tracking technology was applied instead of sensor technology.

The framework of the interface design was inspired by Thierry de Mey's composition *Musique de Table* for three percussionists who produce sounds by striking a table in many ways with their hands. Thus, *Hands On Stage* is constructed as a table-like interface enables the performer to have control over sets of configured parameters of the designed synthesis, patterns manipulation, and visual effects using hand movements. The input data for these parameters are analysed from the performer's hand movements, the percussive strikes of the table surface, and the activation of the buttons on the Wiimote controller.

Interface Construction

The construction of the *Hands On Stage* interface is rather simple. It is made of plywood pieces, a piece of Perspex covered with a sheet of tracing paper, two contact microphones, and a web camera.

Five pieces of plywood are used as the body and the bottom part of the table. Perspex is placed at the tabletop surface, and it is considered as the centre performing stage for the hands. The purpose of using a sheet of tracing paper is to diffuse the light. Two contact microphones are placed at two corners of the tabletop and these are connected to a M-Audio FireWire recording interface. A web camera is placed underneath the tabletop surface facing at tabletop for motion tracking purposes. However, one issue occurred while using an *I-Sight* camera as the choice of the web camera. Because *I-Sight* has a built-in microphone and it tends to reset the computer's audio input setting. As a result, the M-Audio recording interface crashes when there is an audio input conflict. This problem was resolved by replacing the camera to an *iMage* camera, which does not have a built-in microphone.

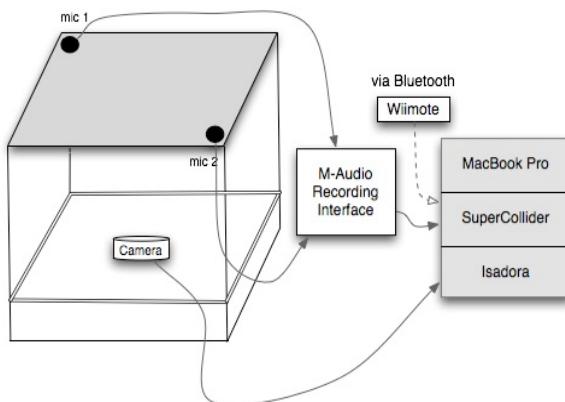


Figure 1. The table interface construction diagram.

After constructing the physical interface, I soon realised there is one major issue in this interface. That is the lack of control over two microphones and a camera. Therefore, a Wiimote controller is added as the extension of the interface and its main function is to act as gates that allow or block the motion tracking data mapped to the sound parameters.

Technical Development

The interface technical development is categorised into three stages, generating input data from performer's action, analysis of the data, mapping of the parameters. In creating a personal approach towards audiovisual performance practices, the project eventually led to an integrated configuration through the use of two software programs, and *Isadora* running on one MacBook Pro. Also, as mentioned earlier, a Wiimote controller was included as a part of the interface.

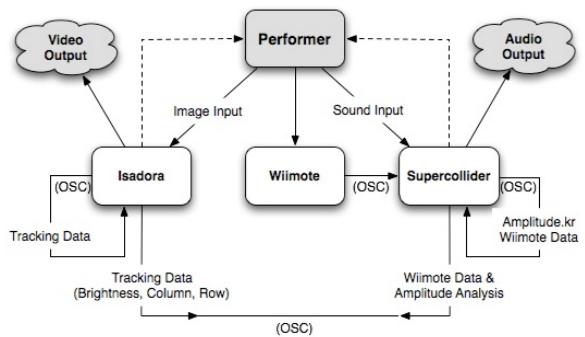


Figure 2. The diagram of the abstract relationship involves in utilising the *Hands On Stage* interface

The sound elements, such as synth definitions, pattern-loops, and the amplitude analysis, are created and processed in *SuperCollider*. The amplitude analysis is used to trigger event when the amplitude value detected is above a certain threshold. It requires using two Ugens, a SynthDef and an OSCpathResponder.

```

SynthDef(\mic1, {|trigID=1|
//pathID=1
SendTrig.kr(Impulse.kr(40), trigID,
Lin-
Lin.kr(Amplitude.kr(AudioIn.ar(1)),
0.001, 0.5, 1, 20) );
}).play;

OSCpathResponder(s.addr,['/tr',nil,1], { |t,r,
msg| var threshold=20;
m = msg[3].round(1).postln;
if ( m > threshold, {
s.sendMsg("/s_new", \recBuf,
s.nextNodeID, 0, 1, \bufnum,
~recBufnum.next);
}
).add;
}

```

The image effects and video tracking analysis are done in *Isadora*. For video tracking analysis, a few Actors including Eyes and Threshold were used in *Isadora* in order to track the position of the shadow produced by the performer's hand. Also, an OSC Listener and an OSC Transmit is used here for receiving and sending OSC messages from and to *SuperCollider*. Here is an example patch:

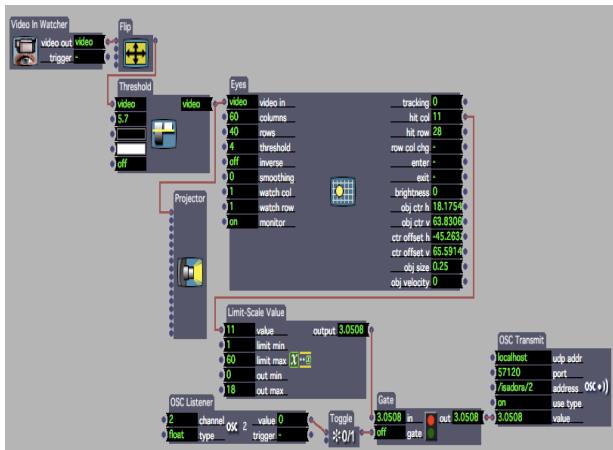


Figure 3. An example of the video tracking analysis and OSC data transmit in a graphic programming environment, Isadora.

Because both programs support OSC (Open Sound Control) protocol, data communication is enabled between the two programs. For example, the performer's hand position is analysed in *Isadora*, and the values of the hand positions are passed onto the sound parameters in *SuperCollider* via OSC, when the gate function is called by the activation of the Wiimote buttons.

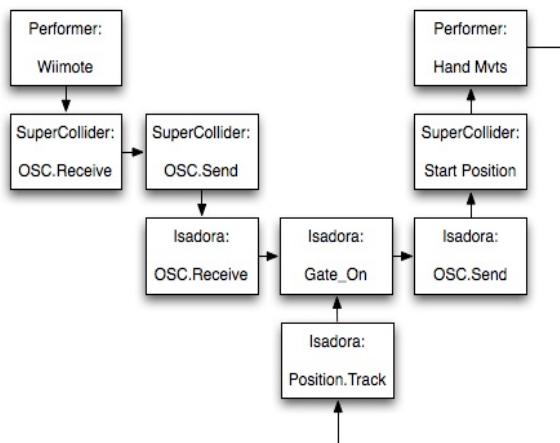


Figure 4. An example of the process of the communication between the performer, SuperCollider, and Isadora.

Improvisation: Untitled # 1

The configuration for this sound and image improvisation, *Untitled # 1*, was first created and performed in 2007 at the CNMA's annual event, DUST. A revised version of this configuration was done in 2008. *Untitled # 1* is a structured improvisation, not a composition. Due to the nature of the mapping decisions and parameters, the result of the performance sounded quite different every time; thus the idea was to create an abstract plan for the improvisation. In fact, this idea suited the objective of the *Hands On Stage* interface, which aimed to include generating sounds that are beyond obvious results.



Figure 5. An abstract idea plotted for the improvisation: Untitled # 1.

Untitled # 1 consists of three groups of sounds: percussive sounds, bell-like sounds, and white noises. The improvisation began real-time recorded percussive sounds. The percussive sounds are produced through striking the tabletop by the performer's hands. These sounds are recorded as sound files and loaded into the buffer, which execute a pattern as the code below, when called. Parameters in the *Pbindef* class, such as the start position (*\startpos*) and the buffer rate (*\bufrate*) are mapped to the video tracking data of the hand's position. The pattern used in the improvisation is a loop of a four-voiced pattern, which is produced by utilising the same sound file playing with different buffer rates and starting positions.

```

~loop1= Pbindef(\loop1,
\instrument, \playBuf,
\bufnum, nil,
\startpos, 0,
\out, 0,
\dur, Pseq([1, 1, 0.5, 0.5,
0.25],inf),
\amp, 0,
\bufrate, 1
);

~loop2= Pbindef(\loop2,
\instrument, \playBuf,
\bufnum, nil,
\startpos, 4, //different startpos
\out, 1,
\dur, Pseq([0.5, 0.5, 1, 1,
0.75],inf),
\amp, 0,
\bufrate, 2 //different bufrate
);
  
```

A graphic user interface (GUI) window is also created for this improvisation performance. The purpose of the creation of the GUI is to provide feedbacks of the status of the Wiimote button. Because they are used as on and off buttons, the colour of the GUI buttons changes in according to the Wiimote status.

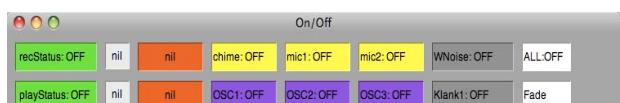


Figure 6. The GUI (Graphic User Interface) for the Improvisation: Untitled # 1.

As for the visual aspects, three effects are designed for this improvisation. The web camera not only does the video tracking analysis, but it

also captures the images of the shadow movements on the tabletop, and processes them in real-time in *Isadora*.



Figure 7. An image of projected visuals from the rehearsal of Untitled # 1.



Figure 8. An image of the Hands On Stage interface from the rehearsal of Untitled # 1.

Conclusion

Utilising the *Hands On Stage* interface for a sound and image performance is a challenging task. First, there is the technical knowledge required, in order to produce and have control over the sounds and image effects. Second, a major distinction between performing this experimental instrument and the conventional instruments is the constant involvement in listening, watching the processed visuals, and reading input data is always required while using such interface. It is perhaps not an expressive musical instrument, because much more attention was needed in order to gain more control over sound and image processing.

Future directions include exploring various configurations and approaches toward a theatrical performance, as well as implementing the technologies, designed for the *Hands On Stage*, onto a different surface, such as a drum skin, to resolve the limitation of the sounds produced from the striking of the tabletop.

References

- Brown, A., Wooller, R. and Thomas, K. 2007. "The Morph Table: A collaborative interface for musical interaction", *Australasian Computer Music Conference*, Canberra, Australia.
- Crevoisier, A., Bornand, C., Guichard, A., and Matsumura, S. 2006. "Sound Rose: Creating Music and Images with a Touch Table", *Proceedings of the International Conference on New Interfaces for Musical Expression*, Paris, France.
- Jordà, S., Kaltenbrunner, M., Geiger, G., and Benica, R. 2006. "The reacTable", *Proceedings of the International Computer Music Conference*, Barcelona, Spain.
- McCartney, J. The SuperCollider Home Page. <http://www.audiosynth.com> (accessed on 24/03/08)
- Riddell, A. 2006. "Sound Performance within an Augmented Reality Theatre Context", *Australasian Computer Music Conference*, Adelaide, Australia.
- Riddell, A. 2007. "Towards Interactive Gesture Control of Sound Patterns Using the Wii-mote", *Australasian Computer Music Conference*, Canberra, Australia.

Stephen Stanfield

17 Arkwright Street
Tarragindi
Brisbane 4121
Australia
stephenstanfield@aapt.net.au

Abstract

As a composer trained in the acoustic Western art-music tradition, the development of new technologies has given rise to new compositional possibilities in regards to the expressive qualities of sound, the transformation of sound through time, and the movement of sound in space. The following presentation is an overview of the artist's creative practice pertaining to personal classifications of space, the application of space in his electroacoustic composition output, and compositional processes involved.

Who am I and where have I come from?

I am currently enrolled in a creative practice Doctor of Musical Arts at Queensland Conservatorium Griffith University. The submission will be a folio of original surround-sound electroacoustic works exploring the concepts of gesture and space in music composition. My initial composition training was through the Queensland Conservatorium and my influences were fairly atypical of many composition students of the mid to late-80's being primarily the works of post-WWII European Modernists (Donatoni, Ligeti, Berio, Nono etc) and working solely with acoustic media (with not much encouragement to move outside of this sphere). I was fortunate to also be exposed to the American avant garde music scene (Cage, Feldman, Crumb, Partch) and further exploration uncovered stereophonic electronic/electroacoustic works by the likes of Oliveros, Davidovsky, and Subotnik. Concurrent with my composition apprenticeship I was indulging in the enjoyment of popular music through playing bass guitar in various original and cover bands. During this time I discovered the Fairlight CMI generated works of British sampling outfit Art of Noise. This was the inspiration I required to dabble in sample-based compositions when the Conservatorium acquired a Fairlight CMI, in addition to composing the occasional work for acoustic instruments with live electronics (delay & reverb) or tape accompaniment.

After graduating I continued to compose mostly for acoustic instruments but maintained my curiosity in digital music creation through the purchasing of a hardware sequencer, synthesiser and drum machine. Much of my output was abstract 'dance' pieces for a choreographer col-

**Artist Talk:
Lost in Space: A Composer's Journey**

league. This MIDI equipment was later superseded by the Atari ST running Notator, followed by a Power Mac 7200 running Emagic's Logic (initially MIDI only) and Finale 2.5. It was this latter system that influenced a progressive move away from composing for acoustic media towards electronic/digital media and using the technology as a performance/presentation medium. This shift generated a number of exploratory works involving lighting effects, projections, and choreography.

Composition Techniques

My personal compositional language and technique in electroacoustic music is influenced by the traditional undergraduate composition training I received where the preconceived organisation of pitch and rhythmic structures (thematic/motivic based ideas) hold a significant importance and governing influence at the micro and macro levels of the composition. Through analysing a selection of my acoustic compositions I have identified my personal usage of thematic/motivic structures as that of 'musical gesture'. Wilson Coker explains in *Music and Meaning* that a musical gesture "comprises a recognizable formal unit and consists of a selection and organization of sonic and rhythmic properties in sonorous motion, which signifies other purely musical objects or non-musical objects, events, and actions" (Coker 1972:18). From this statement I have defined a musical gesture for my own purpose as being: a characteristic and identifiable element in the composition embodying pitch and rhythmic organisation, dynamic and expressive design, timbral and textural qualities, and structural implications.

Fernando Lazzetta states in his article *Meaning in Musical Gesture* that "gesture does not mean only movement, but a movement that can express something. Therefore, it is a movement that embodies a special meaning. It is more than a change in space, or a body action, or a mechanic activity: gesture is an expressive movement that becomes actual through temporal and spatial changes" (Lazzetta 2000:260). This last sentence holds particular importance for my interpretation of a musical gesture and the implementation of this concept in my creative practice. Composing in a surround sound electroacoustic domain means no single sound, or combination of sounds, is confined to a static locational placement but is able to be moved freely around and through different areas of the performance/listening space. This

effectively generates a different expressive meaning and quality of the musical gesture through spatial movement by comparison to the musical gesture being experienced in a static or fixed locational space.

The notions of 'space' for my creative practice have evolved from the development of an analytical framework for my own works and those of other composers. I categorise space into three distinct types, however in his lecture *Music/theatre as a theatre of ideas* (1990) Richard Vella argues that there are only two types of space: physical space and temporal space. Vella's physical space includes all spaces that may contain a performance (real, virtual, projected etc) and in musical terms includes the movement of sound in a stereo field, the registral space of an instrument or group of instruments (high, middle or low), and harmonic space (the intervallic distance between pitches). It is the last two elements contained in this description of physical space, with the inclusion of texture/density, that I believe warrant a separation from the physical space and categorised as 'registral space'. From a compositional perspective the notion of registral space (the vertical organisation of sonic materials, the intervallic distance between pitches, and texture/density) is a separate and independent entity from the physical or locational space. In the context of an electroacoustic composition presented in a three dimension multi-speaker performance environment these two notions of space (registral and physical) could in fact be in contradiction with one another when the normal acoustic perceptions are altered. For example a low register sound could emanate from a loudspeaker positioned at the ceiling height, and a high register sound could emanate from a loudspeaker positioned on the floor. This contradictory effect could also be achieved in a binaural processed recording. This locational and registral space contradiction confers to me the separation of these events is warranted.

From these classifications, the three notions of 'space' employed in my creative practice are:

- *Temporal Space*: the structural design, linear organisation, and motion of musical events (gestures) through time
- *Registral Space*: the registral placement of sonic material, the intervallic distance between pitched material (harmonic structures), and the textural arrangement (density) of timbres and musical gestures
- *Locational Space*: the placement and movement (including distance and perspective) of musical events within the stereo or surround sound performance/listening environment

Through self analysis, and consequently the development of an analytical framework, a number of distinct compositional techniques and processes have come to the fore that typify what may be categorised as 'my style' (whatever that is!). These compositional pre-occupations can be categorised under the banners of musical gesture, temporal space, registral space, and locational

space. It is worth noting that the following list of techniques constitutes my 'compositional toolbox' and not every technique or concept is employed in each composition.

Musical Gesture

For a large proportion of my output the musical gesture is the underlying component. The character and composition of the musical gesture influences much of the work's development and design in regards to pitch centres, register contour, rhythmic motion, dynamic shape, timbral and textural design, and structure. This is very much a notion of self-similarity where there is a high degree of unity between the micro and macro elements of the work. The expressive character of the musical gesture is then heightened and transformed when moved around and through the surround sound performance space. This phenomena is not dissimilar to the effect of experience that Trevor Wishart presents regarding spatial motion that "a motion is characterised not only by its path in space but also by its behaviour in time" (Wishart 1996:222). The speed of the musical gesture's movement and the type of movement that is undertaken (direct trajectories, cyclic motion, dispersion) contributes significantly to the experienced character of the material.

Temporal Space

Being the structural design, linear organisation, and motion of musical events (gestures) through time, the development and transformation the musical gesture is achieved through formulaic and intuitive processes of repetition and variation. Compositional processes and techniques include:

Rhythm

- Cellular motives
- Cyclic/isorhythmic structures
- Additive and subtractive structures
- Mechanical and fluid

Linear Pitch

- Cellular motives/pitch-set
- Cyclic phrases/isomelos
- Harmonic derivative
- Cyclic progressions

Form & Structure

- Influenced by the design and architecture of the musical gesture

Registral Space

Is the registral placement of sonic material, the intervallic distance between pitched material (harmonic structures), and the textural arrangement (density) of timbres and musical gestures. Compositional techniques and materials include:

Harmonic Structures

- Polyinterval chords
- Hierarchical pitch organistaion/Overtones

- Linear derivative

Texture

- Mechanical and fluid layers
- Circle and line (continuum and cantus firmus)

Locational Space

Composing in a stereo, quad, or 5.1 surround sound environment the placement and movement (including distance and perspective) of musical events is paramount to the compositional framework and is often designed for the specific performance/listening experience. The positioning and movement of sonic materials is influenced by the character of the musical gesture and the desired perspective of the material (foreground, middle ground, background). The following spatial motion descriptors are taken from Wishart's *On Sonic Art*:

- Fixed location
- Direct motion
- Cyclic motion
- Oscillatory motion
- Counterpoint of spatial motions
 - Multi-directional
 - Multi-temporal

Composition Process

My compositional process can be categorised into two components or periods of creative time. The gestation period is a significant process in the early stages of composition where the experimentation and exploration of compositional materials occurs. It is here that decisions are made concerning what type of material will constitute the piece (pitch, duration, dynamic, timbre and texture), what techniques will be employed to manipulate the chosen material, what material will be retained and discarded after the manipulation processes, and how will the remaining material be organised in time and in some instances, locational space. During the early stages of this process the compositional intention and overall concept of the work emerges, and in most cases this will be reflected in the title of the work. The time-line of these processes might be as follows:

- concept development - title of the work.
- the development of a musical gesture through improvisation and exploration.
- analysis of the gesture in terms of pitch, rhythm, expression, timbre, texture, and registral contour.
- exploration of the gesture's spatial placement and movement
- exploration of gestural characteristics at micro and macro levels.
- development of additional material through transformational processes.
- selection and de-selection of material transformation outcomes.
- the organisation of the musical structure and material transformations

The application period is the construction of the composition through the organisation and implementation of the materials and concepts developed during the gestation period. During this period the material will be further refined and decisions made regarding the use and development of the chosen material, the structure and timing of musical events, the spatial placement and movement of gestural material, and the solving of notation and/or production challenges related to the musical elements and their application within the composition. The time-line of these processes might be as follows:

- the realisation of the concept and musical materials - writing/producing the music.
- solving issues that arise concerning the notation and/or production of the musical ideas.
- maintaining a balance between processed (pre-composed) and intuitive composition techniques.
- maintaining a sense of development , timing, and dramatic tension through reflection and adjustment of composed material.
- reviewing of recorded repertoire appropriate to the concept of the work in progress.

Production Techniques and Processes

I have been fortunate during the evolution of my compositional interest in the digital music domain that technology has been running along side, if not influencing me with its developments and advancements in MIDI control, audio capturing, and signal processing. From the acquisition of my first computer, commercial software products have served my purpose for a number of musical intentions - Coda Technology's Finale for notation based works, and Emagic's (now Apple) Logic for MIDI sequencing and audio processing. This is not to say that other pieces of software, particularly shareware applications, have not been used in some part of the composition process.

As my requirements and expectations grew, so too did the computer and software capabilities. I can classify my computer-based composition journey into three phases brought about by marked increases in computer processing power, and significant advancements in the software I was using. The first phase was MIDI based only and used external MIDI sound modules. Signal processing, filter and equalisation editing were all controlled from within the host software using a custom designed sysex editor for each piece of hardware. Phase two saw the addition of audio recording, editing, and audio signal processing to Logic's capabilities. This phase also included the exciting development and implementation of VST software instruments, real-time audio processing, and the automation of controller movements (volume, panning, filter and equalisation modifications, auxiliary send levels, etc). The third (and current) phase was the discarding of external hardware sound sources in favour of a sonically self sufficient software-based composition environment. This included the emulation of 'classic'

analogue synthesisers, the increased emergence of sample-based software instruments, and more sophisticated signal processing plugins. In addition to the computer processing power required, a significant development in this phase was the ability to easily and inexpensively create surround sound DVDs.

The desktop computer has now become my composing toolbox, my timbral encyclopaedia, and my performance ensemble.

Conclusion

The transitional journey from an acoustic compositional world to an electronic compositional world is far from over for me. Every time I sit at the computer and compose for my 'performer' I am confronted with new challenges and experiences. The shift into a technology based composition process has enabled me to explore sound at a deeper level through the creation of sound worlds incorporating sampled acoustic sounds, signal processed sounds and synthetic sounds. The greatest revelation however has been the ability to mobilise sound and project those sounds into different listening spaces. The ability to move sound through space adds an entirely new dimension to the process of composing, and consequently generates an enhanced expressive quality and meaning to the musical material.

References

- Coker, W. 1972. *Music & Meaning: A Theoretical Introduction to Musical Aesthetics*, The Free Press, New York.
- Lazzetta, F. 2000. "Meaning in Musical Gesture", *Trends in Gestural Control of Music*, M.M. Wanderley and M. Battier, eds. Ircam, Paris.
- Vella, R. 1990. "Music/theatre as a theatre of ideas", *New Music Articles 8*, Fran Dyson ed. Frog Peak Music, New Haven.
- Wishart, T. 1996. *On Sonic Art*, Simon Emmerson ed. Harwood, Amsterdam.

Robert Bell

Swinburne University of Technology
Faculty of Life and Social Sciences
Melbourne, Australia
robert.bell@vu.edu.au

Poster:**PAD – the Percussive
Audio Database: Initial
Concept and Design****Abstract**

The organization and identification of sound files in sample libraries is a topic of concern in the fields of Music Information Retrieval and Timbre Classification. Of interest is the notion of labelling and searching for files using sonic and timbral criteria.

The author proposes classifying percussive sounds using the 'Percussive Audio Database' (PAD).

PAD is a multi-dimensional taxonomy specifically for drum and percussion libraries, employing an Object-Relational model. It describes sounds by a set of 34 classes, each one defining a specific spectral, temporal or timbral attribute. Examples include Spectral Density, Decay Rate and Attack 'Hardness'. Variables such as beaters, excitation methods and morphology are also classified as potential modifiers of a given sound. PAD's object-relational model allows both hierarchical and nonhierarchical relationships between data to be represented. Classes are labelled with an alpha-graphical notation and visually displayed as a three dimensional matrix/flow chart, using hierarchical layers and distance measures to connect them.

Software is being designed that automatically analyses samples for spectro-temporal data, and permits user-input of various perceptual responses. Samples are then automatically classified with all 34 attributes. The user is able to search, filter and compare any classes between samples in a library.

Artificial Intelligence models are also explored, with a view to allowing PAD to 'learn' patterns of similarity and difference between different samples in a collection. This in turn would further automate the classification process, and improve efficiencies in updating libraries with new samples – especially shared libraries housed online and on local networks.

Ben Swift

Australian National University
Department of Computer Science
121 Childers Street
Canberra, ACT 0200
Australia
benjamin.j.swift@gmail.com

Poster:**Thinking Through Musical Parameter Space****Abstract**

The aesthetics of the musical output of a music generation system are often determined by a small number of high-level parameters. Choosing a trajectory through this ‘parameter space’ is one of the main artistic decisions facing the designer of such automated music systems. Control of these parameters is often achieved by a musician by directly manipulating various knobs and sliders. The Mind-Modulated Music Interface (MMMI) is a system where control of these high-level parameters is provided in real-time by measuring the listener’s brain with an EEG.

By monitoring a neural synchrony parameter which has been linked to ‘musical processing’, the MMMI provides the music generation engine with a time-varying indication of the listeners attention to the stimulus. This information can then be used in various ways to adapt and modulate the stimulus by mapping it to the appropriate aesthetic parameters. in this way, the listener is no longer a passive observer of the musical system, but actively involved in its evolution and journey through the musical parameter space it occupies.

A working prototype of the MMMI has been constructed, and a detailed user study is in the planning stage. The authors goal is both to build a system for his own musical expression, and also to suggest a paradigm by which other artists can harness the power of the brain to experience and control their work in a new way.

Ajay Kapur

California Institute of the Arts
Valencia, CA,
USA
akapur@calarts.edu

Abstract

This workshop describes methods for digitizing, analyzing, preserving and extending North Indian Classical Music. Custom built controllers, influenced by the Human Computer Interaction (HCI) Community, serve as new interfaces to gather musical gestures from a performing artist. Modified tabla, dholak, and sitar will be described. Experiments using Wearable sensors to capture ancillary gestures of a human performer will also be discussed. A brief history through the world of Musical Robotics will be followed by an introduction to the MahaDeviBot, a 12-armed solenoid-based drummer used to accompany a live sitar player. Presentation is full of video examples showing evolution of the body of work in the laboratory to the live performance stage. Live demonstrations will also be included in workshop..

Links to work:

<http://www.digitalraja.com> (video on bottom from Knowledge Network)

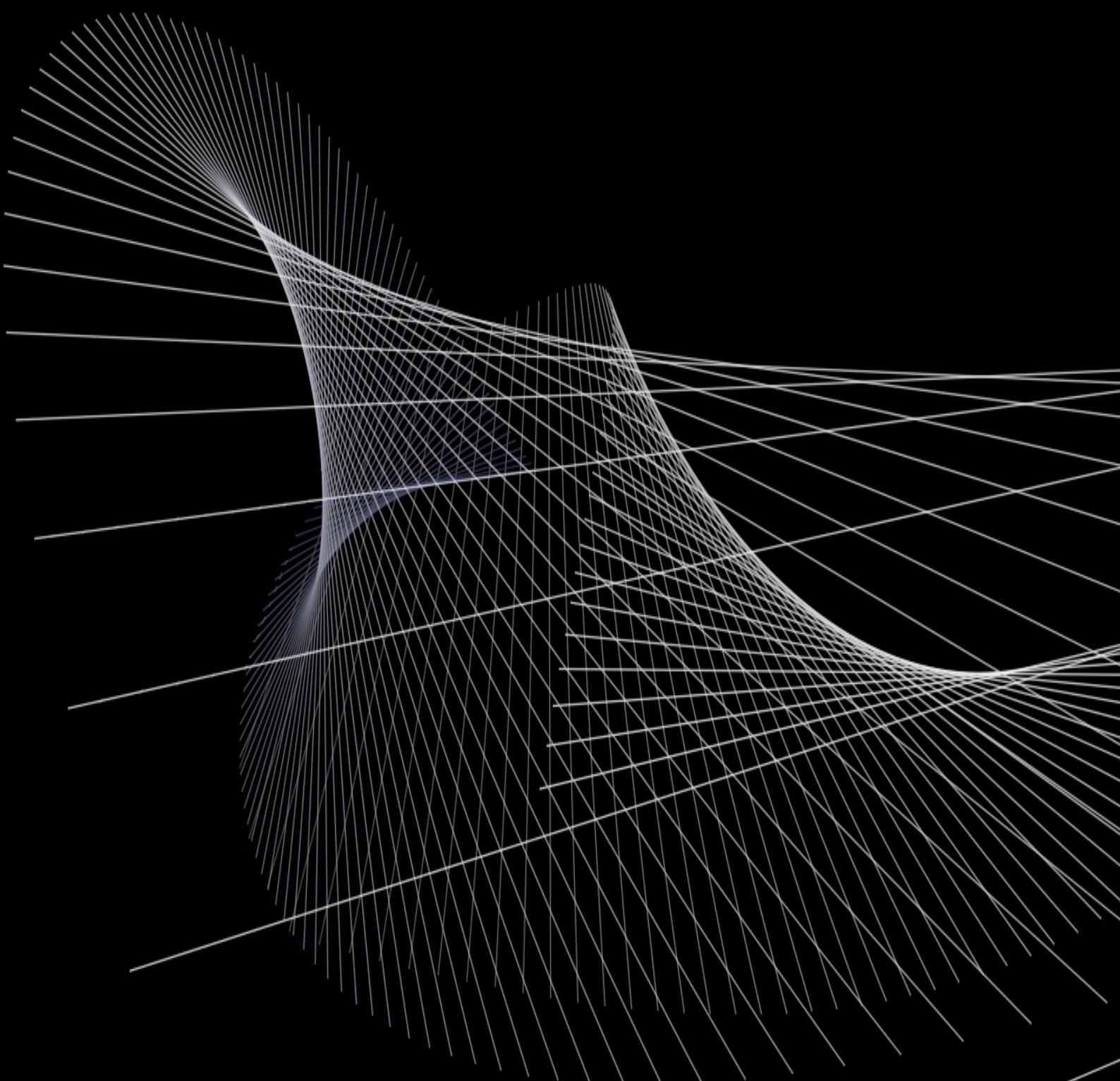
<http://www.youtube.com/watch?v=I6hkLQVnCoQ>



Workshop: 21st Century Raga: Digitizing North Indian Music

Biographical information

Ajay Kapur is the Director of the Interactive Music Technology, Intelligence and Design at California Institute of the Arts. He received an Interdisciplinary Ph.D. in 2007 from University of Victoria combining Computer Science, Electrical Engineering, Mechanical Engineering, Music and Psychology with a focus on Intelligence Music and Media Technology. Ajay graduated with a Bachelor in Science and Engineering Computer Science degree from Princeton University in 2002. He has been educated by music technology leaders including Dr. Perry R. Cook, Dr. George Tzanetakis, and Dr. Andrew Schloss, combined with mentorship from robotic musical instrument sculptors Eric Singer and the world famous Trimpin. A musician at heart, trained on Drumset, Tabla, Sitar and other percussion instruments from around the world, Ajay strives to push the technological barrier in order to make new music



a c m a
000000101100000
australasian computer music association

Sound Devices

liquid architecture 9

SYDNEY
CONSERVATORIUM
OF MUSIC



The University of Sydney