Australasian Computer Music Conference 2004

Proceedings

Ghost in the Machine
Performance Practice in Electronic Music

Victoria University of Wellington
1-3 July 2004

Ghost in the Machine
Performance Practice in Electronic Music
Victoria University of Wellington
1-3 July 2004

Conference Chair: Lissa Meridan
Keynote Speaker: John Cousins
Conference Administrator: Haidi Ankers

Papers Jury
Paul Doornbusch
Dugal McKinnon
Michael Norris
Philip Brownlee

Music Jury
Lissa Meridan
Philip Brownlee
Michel Norris
Jack Body

Technicians
Stacey Pilcher
Andrew Traveller
James Dunlop
Jeremy Brick
Roy Carr
Marcus Wilson

With special thanks to:
Victoria University School of Music
Creative New Zealand
Australasian Computer Music Association
Universidad Nacional de Tres de Febrero
Gaura Yoga
Protel International Technologies Ltd.
Sounz Centre for NZ Music
Anwar_pro@lycos.com
John Downie
and all the School of Music students who gave their time and energy.

Published by
The Australasian Computer Music Association
P.O. Box 284
Fitzroy, Victoria
VIC 3056
Australia

July 2004

ISSN 1448-7780

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.

Contents

ACMC PAPER Sessions

Session 1: Thursday 1 July 11am, Gamelan Room

Warren Burt	Picking up the threads of a history more extensive than previously known: Percy Grainger's work with music technology.

Robin Maconie	Performance Theory and Practice in Stockhausen's *Kontakte*

Miriama Young	Artist Talk: Titlipur

Ryan Cockburn	Artist Talk

Session 2: Friday 2 July 9am, Gamelan Room

Andrew R. Brown & Greg Jenkins	The Interactive Dynamic Stochastic Synthesizer

Tim Opie	Granular Synthesis: Conception and Continuity

Dave Burraston & Ernest Edmonds	Global Dynamics Approach to Generative Music Experiments with One Dimensional Cellular Automata

Lissa Meridan	Studio Report: The ArMiDillo

Session 3: Friday 2 July 11am, Gamelan Room

Lindsay Vickery	Interactive control of higher order musical structures

David Hirst	Fission Or Fusion: Analysing The Acousmatic Reaction

Greg Schiemer, Stephen Ingham, John Scott, Aaron Hull, Damien Lock, Didier Balez, Gareth Jenkins, Ian Burnett, Guillaume Potard & Mark O'Dwyer	Configurable Hemisphere Environment for Spatialised Sound

Jim Barbour	Exploration of the Height Dimension in Audio Reproduction

Session 4: Friday 2 July 3.30pm, Gamelan Room

Greg Schiemer & Mark Havryliv	Wearable Firmware: the singing jacket

Ivan Zavada	Quadraloop: real-time audio looping software

Rene Wooller	Developing a software meta-instrument for live electronic dance music

Colin Hemmingsen & Leigh Jackson	Artist Talk: The use of MIDI instruments as performance tools in Jazz and Fusion improvisation

Session 5: Saturday 3 July 11am, Gamelan Room

Caleb Stuart	The Object of Performance: Performativity in Contemporary Japanese 'Onkyo' Music.

David Sanders	Artist Talk: Song of the Kokako

Alejandro Iglesias Rossi	An experience of synergy of Latin American indigenous instruments and electronic live media

Lydia Ayers & Andrew Horner	Synthesis of Chinese *Dizi* Ornaments

ACMC Concert Schedule

GHOST IN THE MACHINE: PERFORMANCE PRACTICE IN ELECTRONIC MUSIC
1–3 July 2004, Adam Concert Room, Victoria University of Wellington.

COMPOSER/PERFORMER	Title	Format
Concert 1 : Thu 1 July 1:30pm		
Paul Doornbusch	Continuity 4	Bassoon + Tape
Philip Brownlee	Ex Machina	Stereo Diffusion
Stephen Gard	Mbu	DVD
Andrew Brown, Greg Jenkins	IDSS02	Interactive Dynamic Stochastic Synthesizer
Fabio Cifariello Ciardi	Limens Limine	Stereo Diffusion
Concert 2 Thu 1 July 5:30pm	**Douglas Lilburn Complete Electronic Works CD Launch**	
Concert 3 Thu 1 July 8pm	**UNTREF Fronteras del Silencio (Argentina)**	
Concert 4 Fri 2 July 1:30pm		
Jeremy Yuille	Human beings are animals too	5.1 Surround Sound
Brigid Burke	Gesturing on the Move	DVD
David Hirst	La Vie Naturelle	Stereo Diffusion
Panayiotis Kokoras	Response	Stereo Diffusion
Robert Sazdov	Mesecina	5.1 Surround Sound
Michael Parsons	Skitter	5.1 Surround Sound
Concert 5 Fri 2 July 5:30pm		
Gordon Monro	What are you really thinking?	5.1 Surround Sound
Warren Burt	Poems of Rewi Alley	Laptop
Robin Maconie	Measures	UHJ Ambisonic
Catherine Schieve	Aviary	Cross-Grainger Electric Eye Tone Tool
Greg Schiemer	A Dekany in Memoriam	Stereo Diffusion
Concert 6 Fri 2 July 8pm		
John Rimmer & Richard Nunns	Cosmic Winds	Traditional Maori instruments + Tape
Ivan Zavada	Mirage	Violin & live electronics
Robin Fox	Backscatter	8-CH electrical audio signal, single light photon, phosphorous screen
Donna Hewitt	Dystonia	Vocal, eMic & Laptop
Miriama Young (Kristian Larsen: Dancer)	Titlipur	Dancer, Accelerometer, DVD and Laptop
Concert 7 Sat 3 July 1:30pm		
David Downes	Noise (theme and variations)	VHS/Stereo
Jim Barbour	Outback	5.1 Surround Sound
Lydia Ayers	Nostalgia Strata	Stereo Diffusion
Kyung Mee Choi (Rachel Jefferies: Marimba)	Sublimation	Marimba and Tape
Mike Norris (Arnold Marinissen Saw)	In:flection	Saw and Tape
Guiseppe Rapisada	Artico	Flute and Tape
Concert 8 Sat 3 July 8pm	**While You Were Sleeping**	
Ryan Cockburn	Live	Turntables
Emile De La Rey	Live	Live electronics
Simulus	Noodle Spider	Virtual Reality Glove
Joseph Waters & Joel Bluestone	AirEarth	Percussion & electronics
Signer	Live	Guitar and Live Electronics
& VJs of the WYWS Collective	Live	Live Video

Welcome: Conference Chair

Lissa Meridan

School of Music
Victoria University of Wellington
PO Box 600, Wellington
Email: lissa.meridan@vuw.ac.nz

Welcome to Ghost in the Machine - the 2004 Australasian Computer Music Association Conference.

In a field so defined by technology as ours, it is easy to become bogged down in the technicalities of the art and its tools, and to assume that we as artists and composers have the ultimate control or ownership over the sounds that we are producing. It is with this in mind that we have chosen to celebrate the Ghost in the Machine, that intangible something that is beyond our control, that gives life to sounds, enables them to transcend the purely physical and psychoacoustic, and become capable of communication beyond our immediate realities. Here our sounds take on a life and purpose of their own as they transport our experience beyond the mundane.

This conference is as much about building bridges between individuals, practices, institutions and ideologies as it is about coming together to celebrate electronic music, in all its richly diverse guises. Please take this opportunity to listen and share, to enhance the understanding and cross-fertilisation of ideas, and to inspire and be inspired.

We are honoured to host the CD launch of the complete electronic works of Douglas Lilburn, the "father of New Zealand composition", and indeed the man who began and facilitated the development of electronic music here in New Zealand. Lilburn established the first facility dedicated to the research and composition of electronic music here at Victoria University in 1966, and it is with deep respect and gratitude that we can now celebrate this auspicious occasion in his memory.

I would like to thank Creative New Zealand for enabling us to present Cinema for the Ears, an octaphonic installation work by our keynote speaker, John Cousins, and to thank John for his valuable participation, and support. I am thrilled that we have this special opportunity to feature his work during Ghost in the Machine, and hope that you take this opportunity to experience Cinema for the Ears.

I would also like to welcome our special guests from the Universidad Nacional de Tres de Febrero, Argentina, and thank them for coming such a long way to share their musical culture and heritage with us.

Be inspired!

"Don't worry about saving these songs!
and if one of our instruments breaks,
it doesn't matter.

We have fallen into the place
where everything is music.

The strumming and the flute notes
rise into the atmosphere,
and even if the whole world's harp
should burn up, there will still be
hidden instruments playing.

So the candle flickers and goes out.
we have a piece of flint, and a spark.

This singing art is sea foam.
The graceful movements come from a pearl
somewhere on the ocean floor.

Poems reach up like spindrift and the edge
of driftwood along the beach, wanting!

They derive
from a slow and powerful root
that we can't see.

Stop the words now.
Open the window in the centre of your chest,
and let the spirits fly in and out.

-Rumi

Translated by C. Banks, Penguin Books, England, 1995.

Keynote Address: The Ghost in the Machine

'We are the bees of the invisible.
We distractedly plunder
The honey of the visible in order to
Accumulate it within
The golden hive of the invisible'
Rainer Maria Rilke

John Cousins

E-mail: jec20@xtra.co.nz

I have to declare at the outset that I presume the central concern uniting everyone at this conference is the production of art. No matter which genre we use, or direction we may be individually taking, and how seemingly incompatible our attitudes and methods, we are all, I take it, committed to getting something inside, out! Although the materials and tools we use may vary, the process of making good this commitment is always complex and fraught with difficulties. It is not only the fact that we all use digital technology in a variety of ways which places us on common ground, it is also our sharing of the experience of having to deal to the less tangible phantoms of our imaginations.

It is with this in mind, that I proffer the following comments.

The conference title is:

'The Ghost in the Machine' Performance practice in Electronic Music.

It is the 'ghost' of the title upon which I would like to dwell.

Some initial insight into this phrase can be gained from the person who coined it. viz.: the Wayneflete Professor of metaphysical philosophy at Oxford and the editor of the journal *Mind* for nearly twenty five years, one, Gilbert Ryle. A synopsis of his writings states:

'In 'Systematically Misleading Expressions' (1932) he proposed a philosophical method of dissolving problems by correctly analysing the derivation of abstract inferences from uses of language. Applying this method more generally in 'Categories'(1938), Ryle showed how the misapplication of an ordinary term can result in a category mistake by which philosophers may be seriously misled. Dealing with the traditional mind-body problem in 'The Concept of Mind'(1949), Ryle sharply criticised Cartesian dualism, arguing that adequate descriptions of human behaviour need never refer to anything but the operations of human bodies. This form of logical behaviourism became a standard view among ordinary-language philosophers for several decades.'

Cartesian dualism held that the spiritual dimension in human beings, commonly referred to as the 'soul', is a separate category from the physical corporeal aspect, existing independently of it.

Ryle's theory however, relegated the soul to a ghostly category, it being merely a mirage produced by the machine of the body. No body – no soul. Although I tend to agree with him (that our spiritual experiences are essentially the results of body chemistry). I empathise with those who regard the spiritual dimension as an entity in its own right. It can often seem that way. Perhaps it is a little like our everyday experience of the flatness of the earth, despite the contrary data provided by our satellites.

The scale of our physiology in relation to the planet makes it impossible for us, at least in our everyday existence, to perceive its curvature.

Of more interest perhaps, is a recent paper by Bruce Mangan, of the University of California, Berkley, entitled 'Sensation's Ghost' The Non-Sensory "Fringe" of Consciousness', the abstract for which reads:

' Non-sensory experiences represent almost all context information in consciousness. They condition most aspects of conscious cognition including retrieval, perception, monitoring, problem solving, emotion, evaluation and meaning recognition. Many peculiar aspects of non-sensory qualia (e.g. they resist being 'grasped' by an act of attention) are explained as adaptions shaped by the cognitive functions they serve.

The most important non-sensory experience is coherence or 'rightness'. Rightness represents degrees of context fit among contents in consciousness, and between conscious and non-conscious processes. Rightness (not familiarity) is the feeling-of-knowing in implicit cognition.'

A 'feeling-of-knowing' which cannot be grasped by an act of attention.

Sound familiar?

A close analogy is our experience of peripheral vision. We usually look directly at anything we wish to examine, taking little or no notice of the objects away from the centre of our gaze. Attempting a close examination using only peripheral sight will result in us giving up, and turning to look fully at the object. It feels unnatural *not* to do this.

Using the corner of one's eye to notice what is fully displayed centre stage, rather than what is lurking dimly in the wings, is an acquired skill.

This 'feeling-of-knowing' is of course what we all as art makers constantly rely upon. It is our compass and chronometer in the uncharted universe of the soul (or as Ryle would have it, our own impenetrable cerebral illusion). Of course artists don't externalise it as a 'feeling'. We just accept it as 'knowing'. If we chose to be concerned by it for its own sake, we would become philosophers and cease to be artists.

The evolving preference for objective scientific thought over understandings achieved through essentially subjective means (via art, or religion for example) is still very much on the move in our culture which heralds the facts of science and is suspicious of the feelings of art. It is not entirely surprising that this is so, for it is true that the sense of rightness referred to by Mangan, is a notoriously slippery animal. It is the greasiest of pigs, being fiendishly difficult to pin down. But just because you can't catch it, does not mean it is not there! Hopefully, gatherings such as this may validate the power, accuracy and enormous regenerative energy contained in that 'feeling of knowing'.

I am personally convinced that the true value of conferences like this, resides primarily in the collective human contact which coming together in one place facilitates. The enabling solidarity of being in the company of other maniacs, partaking of the collective store of passion inherent in this sort of forum, is, in my opinion, the hit we most need. It is also of course useful to experience one another's work, and to present and respond to papers on various topics. But if I were forced to remove every activity except the one most valuable, I would choose a massive, all in, 72 hour rave with a prize for the last person left standing.

The repercussions on each of us of attending a conference like this are unique and complex. Being exposed to the myriad combinations of personalities, ideas and products can be as confusing as it is uplifting.

Coalitions form fragment and reform. Like gravitates to like, positive sparks off negative. Convictions are confirmed, and demolished and re-confirmed. It is a primal sea teeming with single-minded individuals, ravenous for sensation.

And we are all in survival mode, for this is an emotionally threatening environment. Each one of us desperately seeks evidence that *we* are on the right track, and (secretly) that everyone else is in some way misguided. That somehow our personal 'feeling-of-knowing' will show us the way to the holy grail. Beneath every passionate articulation of what makes sense, and how things should or should not be done, there lurks the realisation of the near impossibility of the task we are asking of ourselves. Alongside every skillfully exposed hypothesis, stalking every confident sonic gesture, is the unsettling recognition that we are involved in a subjective lottery where the chances of winning are virtually nil.

This is of course, simply the reality of the art making process. We can acquire information relating to many aspects of the task. We can get to the bottom of the physics of our particular medium, and how the technology for manipulating it can be of use. We can engineer working contexts of great power and sophistication. But still we sense our ultimate impotence. It is this sense which compels us to act.
It is the maddening itch we must constantly scratch, an innate condition produced by our very humanity; the oxygen which feeds the furnace of our creative desire.
Which is why we gain such sustenance from being with other people who must also scratch.

So perhaps the most useful thing you come away with from a gathering such as this, is the realisation that the only way through to your particular goal is via an essentially self-defined and defining pathway formed by your own particular 'feeling-of-knowing'.

But if you succeed in 'following your bliss' as Joseph Campbell would have it, sooner or later you will find yourself in a region without landmarks of any kind, where none of your previously successful strategies are effective, where you stand or fall on your innate qualities. This is the time when you must investigate the central, elemental secret which is *you.*

As Annie Dillard (Author of 'Pilgrim at Tinker Creek') puts it:

'I think it would be well,
and proper, and obedient, and pure
to grasp your one necessity and not let it go,
to dangle from it limp wherever it takes you.

Then even death,
where you are going no matter how you live,
cannot you part'

**** EG 1.* audio** *(Telling of the secret from INPLACE)*
DUR: 6'.30"

So you find yourself in survivor mode, and you simultaneously rejoice and tremble. This is the paradox of the authentic creative act. It carries with it the possibility of naming the infinite, while simultaneously embracing the risk of dying to the world. The ability to survive in this extreme environment is some thing which cannot be taught at arms length. No number of learned papers or analyses will be of use. It is a place on the fringe of consciousness, beyond, or beneath any act of attention, on the outer periphery of our habitual, safe haven. It is an indescribable location which can be entered only through sentient experience. A 'survivor reality show' which abandons each competitor on their hermetically sealed island, with no cameras, no cell phone, no back up, no escape.

It is absolutely razorblade real.

Throwing oneself back upon oneself, inevitably generates an iconoclastic, self reliant attitude. Reinhold Messener (first person to climb Mt Everest solo, without oxygen) describes it quite categorically:

'Someone who is fulfilling his own ambitions does not need to listen to other people – he must be willing to accept the risk that is ever present when one converts dreams into reality, and he must be prepared to give each one of his dreams the chance to succeed. If he follows anyone else's standards, he won't discover new horizons.

What matters is to evolve one's own standards. Not to allow oneself to be coerced by any outside attitudes, and above all, to find one's own goals within oneself'. Identifying these goals is not always easy, even although, (or perhaps because) they are so intimately within. They lurk in a myopic haze, sensed rather than seen.

In the words of James N Powell (author of 'The Tao of Symbols').

'This obscure something, which has not yet presented itself in the form of coherent thoughts or words, we experience as an intense desire for expression, *a tumescent inevitability*.' (my italics).

Equally important to the 'feeling-of-knowing', or 'rightness' is its underside, or shadow, the 'feeling-of-*not*-knowing', or 'wrongness'. This is the feeling, which so often prevails, particularly in the early stages of the gestation of a work. It results in enormous, but necessary redundancy of both materials and effort. Everything has to be attempted, and usually it is rejected by this sense of wrongness.

The true measure of a potentially successful art maker is their ability to experience this state and despite their fear, panic, disorientation and exhaustion, not only stay intact, but determinedly initiate action. Acting under conditions such as these requires a lunge towards an unidentifiable target, which only materialises after and often because of the lunge itself. It is like jumping off a high cliff into thin air with no visual evidence that one will not fall. The jump transcends the ordinary, only if it is made in the face of seemingly fatal consequences.

The kernel of this act therefore lies in belief or faith in a positive outcome.

Successful artists believe they will succeed, in a totally unregulated environment, without any shred of a guarantee.

Unfortunately qualities such as belief and faith are notoriously elusive and can be regarded as inappropriate in data and information sharing contexts, which often characterise technology based activities. It is an ironic fact however, that data and information sharing, although of use in a restricted instrumental way, have little effect in the generating of belief and faith.

Of course one can achieve nothing without the tools. One cannot just 'think' a work into existence, and it would be churlish to diminish the importance of the interdependent continuum joining the creative impulse, the idea, and the technical means of expression.

But one only has to experience that ghastly moment when one 'runs out of steam', and a studio full of the latest technology becomes so much dumb junk, to realise that the spark which ignites the flammable material of the idea, quite clearly resides in the subjective interstices of the self. One's technique must therefore reach not only outwards, via the technology towards an extant product but also inwards, via stubborn patience towards the elusive identity of the creative impulse.

On the subject of technique: David Smith (American sculptor)

Define technique:

technique is what belongs to others

technique is what others call it when
you have become successful at it

technique as far as you are concerned
is the way others have done it

technique is nothing you can speak about
when you are doing it
it is the expectancy of imposters

they do not show a respect for themselves
or for what they are doing.

The problem lies in the fact that the outward pathway is situated directly centre stage, in fully focussed view, whereas the inward one is quickly shrouded in disturbing shadow. It is so much easier to make headway, when you can see where you are going! Before we know it only the clearly visible and tangibly measurable is acknowledged.

We should therefore be on our guard against the tendency (currently at large) which implies that the more one can measure and quantify the more one can qualify.

As Picasso said:

‘When we love a woman we don’t start measuring her limbs’.

Surely we should share our stories of failure and frustration, weep and gnash our teeth, beat our breasts and declare our fear; cling together under the shadow of our collective ghosts. We should give and receive mutual succour.
So often, alas, this seems to be an outlandish hope. Why must acknowledgement of the subjective aspect of the task, the complete bag of worms, remain unexamined; indulged if at all in off duty time in pubs, cafes or parties, or in the fleeting moments between conference sessions?

I have no answer, except to say that for this we can blame only ourselves, for it is a fascinating behavioural trait of most human beings that we participate in our own downfall. It is almost as though the replacement of analogue technology with digital has infected our intrinsic values so that we are now uncomfortable with anything which might be somewhere between on or off.

More and more we seem to reject that which resists the compulsion to sample, dissect and evaluate. Soon we will genetically engineer our vision so that everything in sight is in absolute focus, and the irritating phenomenon of peripheral uncertainty is abolished. What is more important, a faster computer, or better quality time to contemplate what one might enter into it? How many sacred streams do we cover over in the name of progress?

As Annie Dillard warns us:

‘ It is difficult to undo our own damage, and to recall to our presence that which we had asked to leave. It is hard to desecrate a grove and change your mind. The very holy mountains are keeping mum. We douse the burning bush and cannot rekindle it; we are lighting matches in vain under every green tree. Did the wind once cry, and the hills shout forth praise? Now speech has perished from among the lifeless things of earth, and living things say very little to very few……….

*** *EG 2*: audio** (*sheep talk section of INPLACE*) DUR: 4’00”

Annie Dillard continues:

Birds may crank out sweet gibberish and monkeys howl; horses neigh and pigs say, as you recall, oink oink. But so do cobbles rumble when a wave recedes, and thunders break the air in lightning storms. I call these noises silence and wherever there is stillness there is the still small voice, God’s speaking from the whirlwind, nature’s old song and dance, the show we drove from town.’

We should not be unduly worried by Annie Dillard’s use of the word God. Divine terminology is almost unavoidable in any discussion looking to make sense of the role of human beings in the world.

It is the Gods of the natural world however that she is invoking, harking back to a time when the connection between mankind and nature was acknowledged through particular cultural practices and attitudes. In those days the artist was Shaman, Tohunga, and Witch Doctor, with magical powers sustained by their access to the mystical over arching momentum of the natural world. Connective processes welded people with their environment, and blurred the boundaries we have now established between fact and fantasy. To many earlier cultures, the wind *did* cry and the hills *did* shout forth praise. Although we may not any more acknowledge it to the same extent, this magic is as alive and well in contemporary art making as it was in the broader social functioning of earlier times.

In his wonderful documentary film about Picasso entitled ‘Magic, Sex and Death’ John Richardson points out how the painter thought of his own practice as shamanistic. One only has to look at some of the sculptural works made from found objects to sense their voodoo quality. Picasso saw himself as a magician. The process of making art remains magical.

If any one *artist* was driven from town time and again, along with God’s voice, it was the Russian film maker Andrei Tarkovsky, perhaps because of his totally uncompromising views on the role of cinema as an art form.

In his book ‘Sculpting in Time’, he ponders the delicate, transcendental mechanism of the artistic image by describing the use in his film ‘The Mirror’, of a portrait attributed to Leonardo da Vinci. That elusive quality given off by an art work which transcends the simple combination of its contributing elements is nowhere better explained.

**** EG 3* video** *(Leonardo portrait: A Young Lady with a Juniper)*

'If you try to analyse Leonardo's portrait, separating it into its components, it will not work. At any rate it will explain nothing. For the emotional effect exercised on us by the woman in the picture is powerful precisely because it is impossible to find in her anything that we can definitely prefer, to single out any one detail from the whole, to prefer one, momentary impression to another.

And so there opens up before us the possibility of interaction with infinity, for the great function of the artistic image is to be a kind of detector of infinity towards which our reason and our feelings go soaring, with joyful, thrilling haste. Such feeling is awoken by the completeness of the image: it affects us by this very fact of being impossible to dismember.

In isolation, each component will be dead – or perhaps, on the contrary, down to its tiniest elements it will display the same characteristics as the complete, finished work. And these characteristics are produced by the interaction of proposed principles, the meaning of which, as if in communicating vessels, spills over from one into the other'.

Tarkovsky goes on to say:

'It is possible for us to see any number of things in Leonardo's portrait, and as we try to grasp its essence we shall wander through unending labyrinths and never find a way out. We shall derive deep pleasure from the realisation that we cannot exhaust it, or see to the end of it. A true artistic image gives the beholder a simultaneous experience of the most complex, contradictary, sometimes even mutually exclusive feelings. It is not possible to catch the moment at which the positive goes over into its opposite, or when the negative starts moving towards the positive. Infinity is germane, inherent in the very structure of the image.'

Did you know that the closer you get to the speed of light, the more mass you accumulate so that if you *did* succeed in reaching that ultimate value, your mass would become infinite and you would assume the condition of a 'black hole' with a gravitational pull so gross that nothing, including light itself, could escape. The speed of light is thus an insurmountable physical barrier. This is what Einstein would have us believe.

Tarkovksy's notion however, asserts that the seemingly inevitable consequences of rational relativity can be side stepped. That there is another 'cosmic' realm of travel. It involves the replacement of consecutive narrative structures by faithfully reproduced chunks of reality, juxtaposed so that their combination sparks the plasma drive of our deep intuition. This locates the phenomenon of the *epiphany* as the central mechanism of art. He believes the understandings which accompany the art experience happen synaptically, between things, as much as within them, resulting in an instantaneous numinous revelation.

The essence of this type of expression is undoubtedly the Haiku.

Coldly shining moon
Near the ancient monastery
A wolf is howling

The aim of a Haiku is to initiate an epiphanous experience in the reader by describing juxtaposed but syntactically linked prosaic images as succinctly as possible, within the constraints of a three line, five+seven+five syllable structure. It is the tension between the structural constraints, economy of expression, and the nature of the image itself, which produces the haiku phenomenon. The reader is transfixed directly by the result.

The dart carrying the emotion lodges in one's heart the moment one completes the reading. It is a dart made of ice, instantly melting on impact and leaving no trace. No matter how rationally the reader analyses the haiku in retrospect, the mechanism which releases the emotion remains elusively peripheral. It can never be intellectually described or justified. It is only available as poetic sensation.

And so we reach the nub of things. The fact that however much we dissect the mechanism of the artistic image and the processes which make it extant, both remain stubbornly inaccessible, and can never themselves be dispassionately available to us. Whether the source is in the disembodied soul, or the neurological pathways of the brain, the result is the same.

In the old days they prayed to the Gods, and they were bountiful. Now we must have faith in the magic of the creative process and if we are lucky, it too will be bountiful.

For me, as the job becomes progressively more difficult, the interrelationship between my innate requirement for artistic expression, (for the articulation of the unique *me*) and the sense of my place within a broader universal context, becomes less paradoxical and more symbiotic.

Camping at a coastal location, I note this in my diary: 'A thought today, occurred to me while just sitting watching the sea during the heat of midday. I saw the coast solely as it is, (a natural entity of water, land, sky and climate, all working as part of an environment which I take absolutely for granted) with no aesthetic, or philosophical overlay, just as a 'thing'. And I felt a slight tremor at the back of my mind, a sort of shiver of recognition saying that this is the real secret.

That reality, what we perceive with our physical senses, is all there is, and that anything I might feel as being below or beyond, is merely a construction stemming from my own psychological requirements. I think it was a glimpse of what it would be like if one could truly let go of one's personal point of view of the world and see it and yourself objectively, as something from which you are apart but to which you simultaneously belong.

It was like really getting outside of myself, or else getting fully inside of everything else. A sort of relaxing into a state of neutrality concerning it and me, somehow letting go of my feelings and just acknowledging it as being there. It was a mere hint, like a fragile odour immediately gone, or a mirage, or some kind of ghost sonority which leaves you wondering if you have heard it at all.'

The sensation of 'being outside of yourself', or 'inside of everything else' can happen for all sorts of reasons. In this case it was precipitated by an involuntary meditation. I was warm and comfortable, physically and emotionally relaxed, to the point where my head was empty of thoughts, and I was semi-somnambulistic. The sonic environment in such a location reinforces this state via the subtly undulating mantra of distant surf and air movement. I gradually became less and less aware of my own presence, so that the balance between my awareness of self and my surroundings, tended towards equilibrium.

This is an example of an inadvertent giving in, a relinquishing of effort, a slow lowering into an acceptance of the reciprocal connections between things rather than an emphasising of the boundaries.

More and more, when in the process of working, I wait for such a state to envelope me.

Lets give Annie Dillard the last word:

'The sea pronounces something, over and over, in a hoarse whisper; I can't quite make it out. But God knows I've tried '

*** *EG 4* audio** *(Sea song section 'INPLACE')* DUR: 10'.00"

Picking Up The Threads of a History More Extensive Than Previously Known: Percy Grainger's Work with Music Technology.

Warren Burt

Department of Creative Arts, University of Wollongong
Email: waburt@melbourne.dialix.com.au

Abstract

Although Percy Grainger's work with what he called Free Music, a music of gliding tones and beatless rhythms, is known to some, even to those who know about it, it has long been assumed that the work consisted of a few badly made recordings, a couple of drawings, and the remnants of a machine or two at the Grainger Museum in Melbourne. Research conducted last year for ABC Classic FM at the Grainger Museum, University of Melbourne has revealed that Grainger's work with Free Music was much more extensive than previously thought, with about an hour of sound recordings, and many other documents existing. The recordings reveal Grainger's open-ended and improvisatory approach to music technology. These recordings will be discussed, as well as showing the plans for his last machine, the Electric Eye Tone Tool, and demonstrating our contemporary reconstruction of it.

1 Introduction

Although Percy Grainger's work with what he called Free Music, a music of gliding tones and beatless rhythms, is known to some, even to those who know about it, it has long been assumed that the work consisted of a few badly made recordings, a couple of drawings, and the remnants of a machine or two at the Grainger Museum in Melbourne, as well as several computer and electronic realisations by others of his Free Music 1 and 2. Last year, for ABC Classic FM, I made a radio show about Grainger's Free Music, which included research in the archives of the Grainger Museum, as well as making a contemporary reconstruction of Burnett Cross and Percy Grainger's final machine, the Electric Eye Tone Tool. What I found was that, contrary to common knowledge, Grainger's work with music technology from the mid 1940s until his death in 1961 was deeply and continuously involving; work in which he anticipated, in his highly individual way, many of the techniques of music technology today; and work which was extensively, if a bit haphazardly documented. Far from the 2 or 3 minutes of recorded material that is the common perception of Grainger's recorded Free Music output, there is in fact, about an hour of recorded material, as well as a volume of sketches, scores, and plans. And some of the recordings, although probably just sketches, are compositionally astounding, and should be much better known than they currently are. This work involved a bit of detective work, and cross referencing, and was limited to the archives at the Grainger Museum at the University of Melbourne. A search of overseas collections might reveal even more material. Clearly, Grainger's work with Free Music was not a dead end, or even a magnificent failure, but part of a sustained research effort which left threads of inquiry to be continued, and which we are now just beginning to take up again.

2 The Beginnings

Grainger's first work with music technology was with recording technology, in the first decade of the 20^{th} century, when he was one of the first composers to make field recordings of folk musicians. In the early 1920s, he also worked with cutting player piano rolls by hand, in order to hear the rhythmically complex music of his *Sea Song Sketch* from 1907 played accurately. This early work with music technology planted the seeds to a free-wheeling approach to music technology that was to come to fruition in his work of the late 1940s and 1950s. However, this work was motivated by his idea of Free Music, which he had been developing since his teenage years in the 1890s. Central to Grainger's work with Free Music was the concept of the gliding tone. In the 1930s, Grainger worked with Leon Theremin and his theremins, and wrote at least three scores for theremin ensemble (*Free Music 1 and 2*, and *Beatless Music* of 1937, the latter an arrangement for six theremins of the 1907 *Sea Song Sketch*) and had Leon Theremin not returned to Russia in the late 1930s, Grainger probably would have done much more with him. (In terms of correcting the historical record, it might also be mentioned that Theremin's return to Russia was not the kidnapping that was reported for many years, but was staged by Theremin and his NKVD associates to look that way - full details are in Glinsky (2000) [1]. However, Grainger's frustration with the rhythmic control of the theremin, and his desire for more precise pitch control probably would have led him to construct devices for more accurate and repeatable control, which was indeed the path he eventually took with the physicist and engineer Burnett Cross.

Proceedings ACMC 2004
ISSN: 1448-7780
http://acma.asn.au

3 The Existing Recordings and Early Multitrack Experiments

When I approached the Grainger Museum for this research, they were very cooperative and helpful, but were in the middle of a building crisis, major structural flaws having just been found in the building. As a result, I had to work around various structural and time constraints, but was still able to uncover a lot of material. Grainger and Cross had made a number of 78 rpm recordings with a home record cutting unit between 1948 and 1954. All of the recordings in the Museum had been transferred to CD by Filmsound Australia. However, each CD was simply labelled with the names of whatever was on the jacket of the 78 rpm recording. Many of these labels were erroneous, or had been confused in the half century of the recordings lying in the archive, so I found I had to listen to every track on every CD to find that often mislabelled gem, which was in fact, a missing free music recording. There were also a number of drawings of various free music machines in the archives, and it was quite interesting trying to figure out which version of which machine was used for a particular recording. (Grainger and Cross were constantly rebuilding their machines as new ideas occurred.) A number of recordings also required close listening to find out exactly what was being done. This was especially the case with *The Lonely Desert Man Sees the Tents of the Happy Tribes* recordings. This is a series of six 78 rpm sides which contain various parts for reed organ, marimba, voices, and what sounds like either a cello or a Solovox - an early monophonic synthesizer made by the Hammond corporation. The recordings also contain a number of comments by Percy, Ella Grainger, Burnett Cross, and an engineer referred to as Howie, who was probably Howard Cross, Burnett's brother. From these comments, and by listening carefully to the kind of surface noise and equalization in each subsequent recording, I realized that what was happening was that Grainger was recording a part on one 78 rpm disc recorder, and then playing that back on another 78 rpm disc player, playing along with that, recording the result, and then playing that recording back while recording another part along with the mix. This was a primitive, but still effective way of multitracking.

Related to this is an amusing anecdote by Burnett Cross about a recording of *Bold William Taylor*, a folksong arrangement of Percy's. Cross had sung in amateur choirs and was a baritone. In the early 1950s, Grainger asked him if he would learn the voice part, "irregular rhythms, variations, Lincolnshire dialect and all." Cross did so, in his baritone range, at a slower tempo than required. In copying the record, they sped it up to the required tempo and pitch, and this sped up version was used by Grainger in a Knoxville, Tennessee performance with a group of live string and reed players. Cross reports that many years later, he met Peter Pears, the internationally famed tenor. Pears said, "Burnett Cross? Burnett Cross? The man who sang *Bold William Taylor*! Marvelous! Marvelous!" And so, through technology, Cross, the amateur baritone, became an internationally known tenor [2].

4 Early Attempts at Sequencing

Grainger not only anticipated multitrack recording, he also anticipated sequencing and interactive performance with sequencers. In the mid-1940s, he purchased three Hammond Solovoxes. These were monophonic synthesizers with some degree of timbral control. Grainger cut a piano roll by hand of his rhythmically complex *Sea Song Sketch*, and then placed the roll in his player piano. This piano had a mechanism which allowed the action of the piano to be moved enough so that the hammers would not strike the strings. This resulted in a silent piano whose keys would still move up and down. He then mounted the Solovoxes above and below the piano keyboard, and with an ingenious arrangement of strings, wires and pieces of wood and cardboard, had the moving keys activate the keys of the Solovoxes. (See Figure 1) The Solovox also had a volume control which a live player could manipulate. In the 1950 recording of Percy, Ella, and Burnett manipulating this contraption, Ella and Burnett are manipulating the volumes of the three instruments, Percy is turning the machine on and off, and playing another part live on a reed organ. The piece is about 30 seconds long, and the recording consists of 4 takes of the piece, with Grainger sometimes commenting between the takes. To me, this seems like an early example of having a sequencer control electronic instruments while live performers further interact with the electronic instruments in real time. A performance with, for example, the IRCAM ISP workstation and an acoustic instrument may be a bit more complex, but conceptually it inhabits much the same territory as this early experiment of Grainger's. It should also be mentioned that there are some recordings in the archive of *Early One Morning*, another folksong arrangement, played by Solovoxes and Reed Organ. Though not as mechanically innovative as the *Sea Song* recordings, these recordings do stand as some of the earliest recordings of live keyboard synthesizer playing.

5 The Butterfly Piano

With the work on the Butterfly Piano, Grainger's Free Music work takes a very big leap. For this is the first work in the series of Free Music recordings, work that seeks to go beyond the pitch - rhythm paradigm of traditional musical organization. Grainger wanted to obtain gliding tones. He was interested in how close discrete tones needed to be in order to obtain the illusion of a gliding tone. Using a "Knoxville" piano, a small 3 and ½ octave piano with full size keys and a real piano action, Grainger restrung it and retuned it so that the instrument played just over an octave of 36 tone equal tempered tuning, which he then dubbed The Butterfly Piano. (See Figure 2) A piano roll was cut by hand, with sequences Grainger was interested in hearing, and a vorsetzer unit was placed on the keys of the piano. Six recordings were made of this piano roll. In these, Grainger makes considerable comments about exploring the effects of closely tuned tones used in trills to obtain a timbrally richer effect than normal repeated notes could give. He also speeds up and slows down the speed of the roll, to explore the capabilities of the vorsetzer mechanism to produce differing results from the same roll. His work here is not only significant for its exploring of the psuedo-glides of 36 tone equal temperament, but also for his attitude towards his mechanism, the piano roll, which he regards as a source of variation in its own right, and not just as a reproducing mechanism.

6 The Reed Box Tone Tool

Following the work on the Butterfly Piano, Grainger and Cross turned their attention to one of Grainger's favorite timbres, the reed organ. Two sets of 126 reeds were retuned to 48 tone per octave equal temperament, giving just over 5 octaves of range in this tuning. The relative ease (until you try it! I speak from the experience of retuning an accordion into just intonation) of retuning free reeds prompted not only Grainger, but other microtonal experimenters, such as Harry Partch, to also work with making retuned reed organs in these same decades.

The reed box machine went through a number of incarnations, trying out different combinations of hoses, hills and dales paper rolls and so on. Figure 3 shows Version 6 from 1950. The first recordings of the reed box, from Jan 29 1951, sound as if this mechanism is being used, but the later recordings from the same day are far too complex in sound to have been produced in this way, so I'm not sure if the machine in this illustration was actually used in any recordings. Soon after this, however, Grainger and Cross abandoned the hills and dales paper roll approach for the reed box, and mounted the reeds in four large cardboard boxes, which they then played with hand-cut paper rolls. This machine was used for the recording that Percy introduces with the words "First Gliding Chords on Reed Box Tone Tool, September the 30th, 1951." This recording has been released on the CD of *Leonardo Music Journal No 6* [3]. However, some of the most amazing reed box recordings are from the earlier January 1951 session, in which a combination of up to 28 reeds, some in 12 tone equal temperament and some in 48 tone equal temperament are combined to make experiments in timbral synthesis. Some of the thickest and most amazing sounds ever produced by free reeds are in the January 1951 recording that Howard Cross introduces with a snicker and the words "This is done using both the upper and lower levels." This is some of the wildest tone cluster music ever, and it was made with harmonium reeds and hand cut paper rolls in 1951, some 10 years, for example, before Penderecki's *Threnody*, or 9 years before Bernard Herrmann's *Psycho.* For those who persist in thinking of Grainger only as the composer of *Country Gardens*, I can highly recommend playing them this amazing recording. In fact, sketch though it may be, I've come to like this recording as my favourite Grainger piece ever. Also of note in these recordings is the fact that the roll is once more treated as a resource for music performance possibilities. Pulled by hand, variations in speed and direction are both used. A passage might be heard slightly faster in one direction, then slightly slower in the other direction. Grainger and his co-workers were quick to seize on the musical possibilities of their machines, and again, were not just using them as simple reproduction devices.

7 The Oscillator Machines

Immediately following their work with the reed box, Grainger and Cross began to work with a Codemaster oscillator, assembling devices to produce smooth glides with it. The drawing called "Oscillator-Playing Tone-Tool, 3rd Experiment (early Nov 1951)" shows the hills and dales paper graph from the reed box experiments adapted to playing the oscillator. (See Figure 4) Some of the November 1951 "Oscillator Test" recordings may have been made with this machine. It was quickly superceded, though, by a two voice machine consisting of a solid paper roll with gliding patterns made with clothesline glued to the paper, and sliding devices moving controls for two oscillators up and down. (See Figure 5). And this machine was superceded by the famous Kangaroo Pouch Machine which still sits today in the Grainger Museum. (See Figure 6) In the current Museum installation, only the four paper rolls for pitch control are on the machine. The four narrower rolls which would control the volume of each voice are missing. Cross and Grainger had three voices working when they made their December 1951 recordings on the machine. On the recording, at mid-point, Cross's voice can be heard saying, enthusiastically, "And now in the other direction!" They then reverse the hand cranked roll and play their test pattern in a perfect retrograde.

8 The Electric Eye Tone Tool - Death and Rebirth

Shortly after these recordings were made, Grainger and Cross embarked on a much more ambitious project, The Electric Eye Tone Tool. This was to have been a seven voiced instrument, with seven sine wave oscillators controlled by variations in light on a series of 14 photocells. Patterns painted on a large plastic sheet pulled across the plate of the instrument caused the variations in light. In this way, Grainger's vision of a graphic notation for precise glides and intervallic leaps could be finally realised. (Figure 7) They had completed three working voices on the machine, according to Cross in his article about the making of this machine [4] and according to Richard Franko Goldman [5], Grainger in 1955 was writing graphs for selected passages from Wagner (the opening of *Tristan*), Scriabin, Grieg and other composers. Whether these were ever recorded, or whether any of Grainger's own experimental rolls exploring his own ideas were recorded is unknown to me. As for the machine itself, it suffered a very cruel fate, of which I have heard three versions. It was either destroyed in transit to the Grainger Museum, or it was disassembled for repairs at the time of Grainger's death, and the parts were thrown out by mistake, or it was disposed of by Grainger's executors, who were not aware of its significance. Whichever of these stories turns out to be true, the sad fact is that this machine, the crowning achievement of Grainger and Cross's work, no longer exists. However, Cross left rather detailed plans for the instrument, and a number of people over the past few decades, such as Caroline Wilkins and Rainer Linz, have expressed interest in rebuilding the machine from these plans. I, too, had toyed with this idea, but had never followed through on the idea until late 2003, when John Crawford of ABC Classic FM approached me with the idea of producing a show about Grainger's more radical thinking, as part of a Grainger Festival organised by the Adelaide Symphony Orchestra and ABC Classic FM. I proposed that he fund the reconstruction of the Electric Eye Tone Tool, and he eventually approved the idea, so that from August through October 2003, Paul Francis Perry, Malcolm Ellis, and myself built a new version. This version differs in some ways from the original. (See Figure 8) It is shorter than the original, contemporary halogen lamps providing a better light source than their bulbs, and it is narrower. Grainger's machine was 1.5 meters wide, ours is about 1 meter. This is because the 1.5 meter wide plastic sheeting Grainger used is no longer available. Similarly, we decided to use modern IC circuitry instead of the single transistor oscillators they used, and we ended up using easily available Dick Smith solar cells instead of their now obsolete selenium photocells. We built an upper set of control slits as in the Cross-Grainger machine, but for the first few pieces made with our machine, we found them not to be necessary. However, we feel that we've gotten as close to the performance of the original machine as one can get using currently available technology, and we feel that Percy and Burnett would be pleased with our machine as a continuation of their research. (Since Grainger and Cross were continually modifying their machines, we felt that a kind of "original instrument" aesthetic would be out of place in our rebuilding of the instrument. We want to move on from their work, not to repeat it.)

9 New Works for the Electric Eye Tone Tool

For the radio program, I decided that just playing Grainger's work would not be enough. I wanted contemporary composers to work with the machine. Accordingly, I commissioned Wang Zheng-Ting, Tristram Cary and Catherine Schieve to make works involving the machine. In our haste to finish the machine in time for the program, we made some errors in power supply design. As a result, for this program, our machine had a 50 hz frequency modulated wobble on all seven of its oscillators. Fortunately, all three composers felt they could live with that (Tristram was indeed less than happy, but still decided to go with the timbre as it existed) and made works for the machine. Interestingly enough, all three of them made works for the machine and accompaniment. Ting's work was the first to be made. He is a caligrapher as well as a composer and sheng player, so I thought he might be interested in doing some sort of caligraphic score. But when he moved his hands between the lights and the photocells, and heard the various twitters and sweeps that resulted from that, he decided to simply multitrack several tracks of hand made gestures, and then played sheng along with these tracks to make his piece, *Future Four Seasons.* Tristram's approach was different. We used AudioMulch to make a backing track which indeed did have pure stable sine wave timbres and Graingerian glides. We then cut envelope shapes out of duct tape and placed them on the plastic sheet. These made a short series of gestures, changing pitch and loudness on differing numbers of oscillators. In live performance, we played back the recording of the AudioMulch sine waves, while Tristram hand cranked the machine back and forth, assembling, quite elegantly, a series of modulated warbles in real time over the sine wave base. Catherine's approach was different yet again. Her principal compositional interest has, for many years, been the application of graphic notational techniques to music. An accomplished painter as well as a composer, she found that titanium white paint produced very controllable and repeatable gestures, while different colours and densities of paper and other materials produced gestures that varied considerably from performance to performance. Her piece, *Aviary*, also uses a backing track, made by recording the output of the machine while she was moving her paintbrush between the lights and the solar cells, and a live performance of a very colourful score, consisting of many different juxtaposed patterns of paper, ink and paint. (See Figure 9) In performance, following the Cross-Grainger practice, the roll is used first in one direction and then in the other, with speed variations produced by hand cranking as an essential part of the performance. Since completing this piece, she has made a second piece using the machine, *Repentistas*, for Violin, Electric Eye Tone Tool Two, Piano, Organ, and Toy Piano. In this piece, she worked exclusively with titanium white paint, making repeatable gestures based on the inflections of the Brazilian folk singers referred to in the work's title. (See Figure 10)

10 Future Developments

Future developments of the machine include not only re-designing of the power supply to provide more stable oscillators, and modifications to the ranges of the control knobs to provide a finer range of control over pitch and amplitude, but also to put the raw voltage output of the photocells themselves into a control voltage to midi converter, getting 14 midi continuous controller messages to apply to any desired set of synthesizer or sampler parameters. That these control patterns would then be able to be physically manipulated in real time by moving the roll back and forth is one of the ways that I feel the implications of the Cross-Grainger work will be developed. I would also like to commission other composers to work with the machine, and may even initiate a CD recording project of new pieces made for the revived Electric Eye Tone Tool. I would also be very interested in getting a commercial CD of the Grainger Free Music recordings released. In the appendix to this paper I have listed all the currently available recordings that I found in the collection of the Grainger Museum.

11 Acknowledgements

Thanks to John Crawford, Graeme Hinkley, Julian Day from ABC Classic FM; Barry Peter Ould, from Bardic Editions, Grainger's publisher; Brian Allison and staff at the Grainger Museum; Wang Zheng-Ting, Tristram Cary, and Catherine Schieve for thier compositions, and Paul Francis Perry and Malcolm Ellis for their help in rebuilding the EETT2.

References

1. Glinsky, Albert, 2000. *Theremin: Ether Music and Espionage,* Urbana and Chicago, University of Illinois Press.
2. Cross, Burnett, "Collaborating with Percy Grainger." *NMA* 7: 3-4.
3. CD Companion, *Leonardo Music Journal No. 6*, 1996. Cambridge, MA, MIT Press.
4. Cross, Burnett, 1991. "Grainger Free Music Machine" *A Source Guide to the Music of Percy Grainger,* Thomas P. Lewis, ed., New York, Pro/Am Music Resources, Inc. Pp/ 158-162.
5. Goldman, Richard Franko, "Percy Grainger's Free Music" *The Julliard Review II/3*:6-11.

Appendix

Listing of Recordings of Percy Grainger Free Music (and other) recordings from the collection of the Grainger Museum. Transcribed and tentatively identified by Warren Burt, Oct. 03 abbreviations: PG = Percy Grainger, EG = Ella Grainger, BC = Burnett Cross, WB = Warren Burt

1. PG Butterfly Piano Introduction No. 1 (0:56)
2. PG Butterfly Piano Introduction No. 2 (0:46)
3. PG Butterfly Piano Introduction No. 3 (1:33)
4. Butterfly Piano Roll Complete First Recording (1:29)
5. Butterfly Piano Faster with PG Intro (0:44)
6. Butterfly Piano Faster and Slower (0:54)
7. Sea Song Player Piano Roll Cut 1923 (0:37)
8. Sea Song with Solovoxes and Intro by BC Feb 1950 (1:53)
9. Solovoxes and Reed Organ Early One Morning 1950 (3:01)
10. Lonely Desert Man 01 PG Intro and Marimba (1:41)
11. Lonely Desert Man 02 Reed Organ part only (1:01)
12. Lonely Desert Man 03 Reed Organ Record & Marimba (1:10)
13. Lonely Desert Man 04 PG & EG Sing with Marimba Record (1:36)
14. Lonely Desert Man 05 PG & EG Sing with Mba Record Take 2 (1:35)
15. Lonely Desert Man 05A PG & Engineer Speak (0:17)
16. Lonely Desert Man 06 Mix with Recorded Cello? (1:15)
17. Reed Box 01 Jan 29 1951 First Recording (2:41)
18. Reed Box 02 Jan 29 1951 Second Recording (1:45)
19. Reed Box 03 Jan 29 1951 Top and Bottom Ranks - THICK (2:05)

20. Reed Box 04 Jan 29 1951 Timbral Synthesis (2:10) 21. Reed Box 05 Jan 29 1951 BC Explains Previous Take (1:21)
22. Reed Box 06 First Gliding Chords Sept 1951 (1:10)
23. Reed Box 07 Sept 51 Second Recording Roll forward and Backward (1:36)
24. Reed Box 08 Glides No Talk - From Cassette given to WB by BC (1:01)
25. Reed Box 09 Forwards - 1988 Recording (1:03)
26. Reed Box 10 Backwards - 1988 Recording (1:31)
27. Oscillator Test - 1 Voice with Hills and Dales Nov 2 1951 (1:29)
28. Oscillator Test - Two Parts (0:42)
29. Oscillator Test - Four Parts (1:19)
30. Oscillator Test - AM Oscillator - From Leonardo Music Journal CD (0:50)
31. Kangaroo Pouch Machine Dec 1951 BC Voice (1:22)
32. Free Music Number 1 - Les Craythorn Realization?? Only 3 voices (1:47)
33. Free Music Number 1 - Les Craythorn Realization 4 Voices - From BC Cass (1:47)
34. Free Music Number 2 - Conyngham realization - Move LP (1:12)
35. PG Sings and Plays Rufford Park Poachers (4:27)
36. PG Sings and Plays Lord Melbourne (3:12)
37. Various Introductions spoken 1988 by BC (7:29)

Figure 1: Sea Song Solovoxes and Player Piano Rig

Figure 2: Butterfly Piano

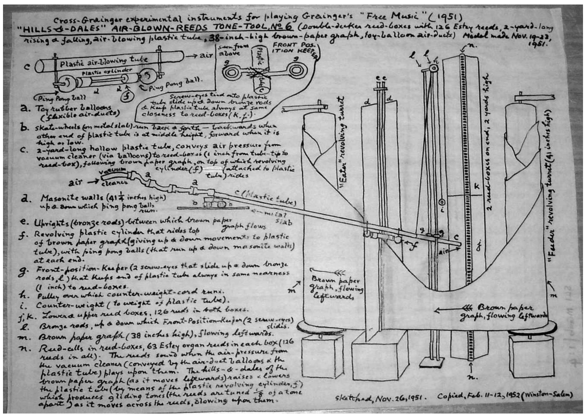


Figure 3: Reed Box Control Version 6

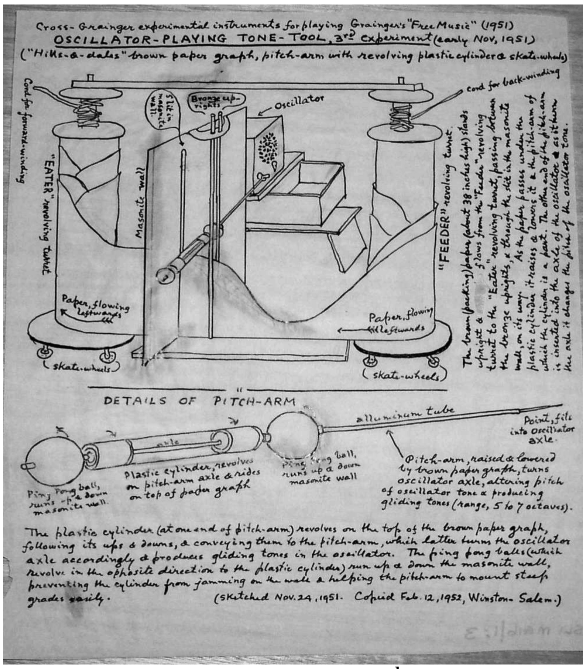


Figure 4: Oscillator Control, 3rd Experiment

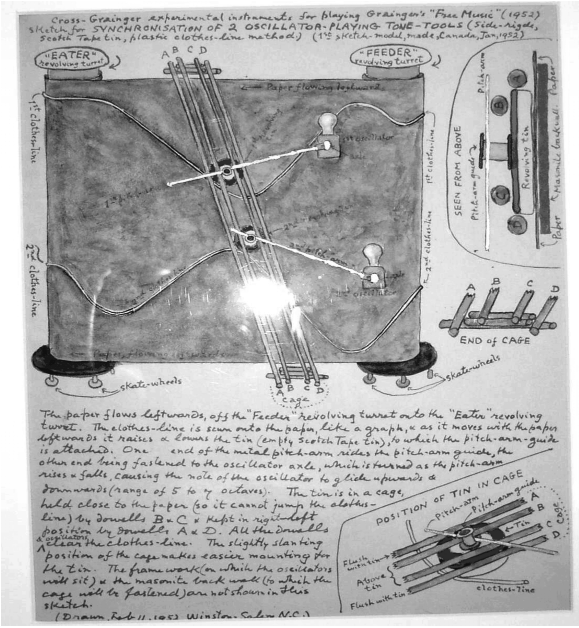


Figure 5: Clothesline Controller

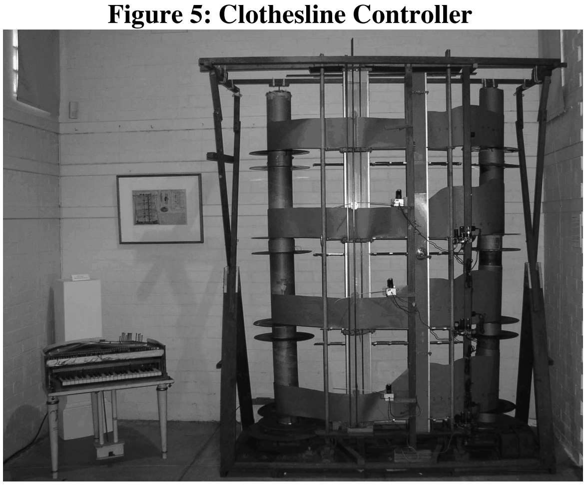


Figure 6: Kangaroo Pouch Machine, October 2003

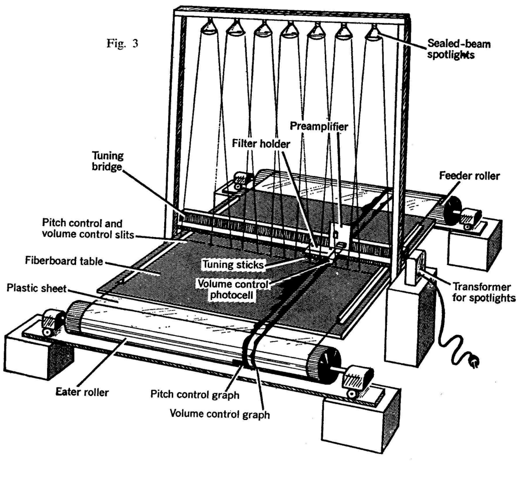


Figure 7: Burnett Cross Drawing – Electric Eye Tone Tool

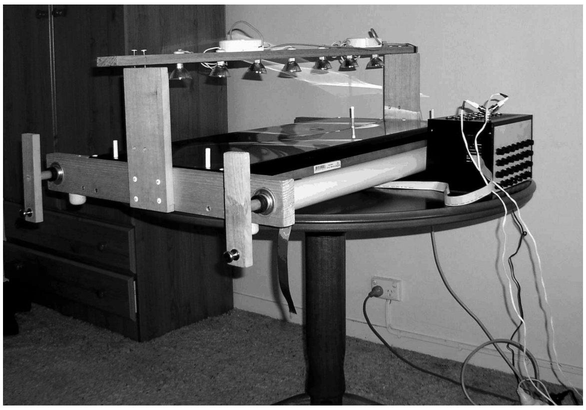


Figure 8: Burt-Perry-Ellis Reconstruction – Electric Eye Tone Tool Two

Figure 9: Catherine Schieve score for *Aviary* (detail)

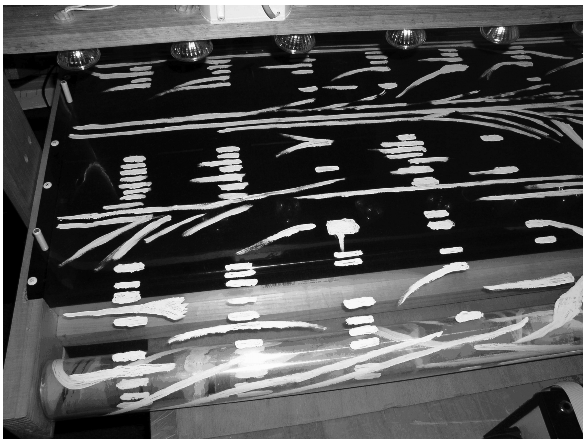


Figure 10: Catherine Schieve: score for EETT2 part for *Repentistas* (detail)

The Interactive Dynamic Stochastic Synthesizer

Andrew R. Brown †* & Greg Jenkins †

† Queensland University of Technology (QUT)
* The Australasian CRC for Interaction Design (ACID)
a.brown@qut.edu.au
g.jenkins@qut.edu.au

Abstract

Throughout musical history new sounds and instruments have opened new opportunities for music making. In this paper we outline a new interactive digital instrument that implements the dynamic stochastic synthesis algorithm devised by Iannis Xenakis. We discuss the history and operation of this synthesis process, previous implementations of it, and how our implementation is the first we know of designed specifically for live performance. Finally, the behaviour tendencies of the synthesis system and how these impact upon interactivity are discussed.

1 Introduction

The role of chance and random occurrence as a musical technique has played a significant part in Western compositional developments in the 20th century. Among the leaders of this development, that include John Cage, Gottfried Michael Koenig, Karlheinz Stockhausen, Lejaren Hiller & Leonard Issacson, and Luciano Berio, Charles Dodge and Brian Eno, is Iannis Xenakis.

Xenakis's interest in stochastic systems was informed by modern scientific theories including probability theory derived from quantum mechanics, and kinetic theories of gases. His conception of stochastic processes was as indeterminate functions with entropic tendencies or, more simply, processes that "evolve in different directions" and move between states of "order to disorder, or vice versa" [6:16].

Uniquely, Xenakis applied probability not only to the organisation of musical objects and structures but also directly to the generation of sound waves, with a process he called Dynamic Stochastic Synthesis (DSS). This synthesis process applies constrained random processes to the moment by moment fluctuations in air pressure that are fundamental to sound creation.

In this paper we outline an implementation of the dynamic stochastic synthesis algorithm that is intended for real-time performance, discuss some of the historical contexts that inform the design and discuss issues of interaction design that have arisen during its use in performance.

2 Dynamic Stochastic Synthesis

For most of his life Iannis Xenakis was concerned with the use of probabilities for creating music. He conceived of DSS as a way to apply stochastic functions to the generation of audio waveforms. His early compositional uses of stochastic processes were primarily in determining aspects of works at the level of note attribute selection and the organisation of musical form. It was not until later in his life, that he applied these processes directly to the dynamic stochastic synthesis process although he had previously worked with stochastic functions as elements of synthesis parameter control in processes including granular and FM synthesis. Xenakis termed the organisation of music at this sample-by-sample level, microstructure [7].

Dynamic Stochastic Synthesis involves multiple levels of probabilistic functions that determine the positions of break points in an envelope which in turn describes one cycle of an audio waveform, as shown in figure 1. The number of points in the waveform is predetermined by the user or programmer.

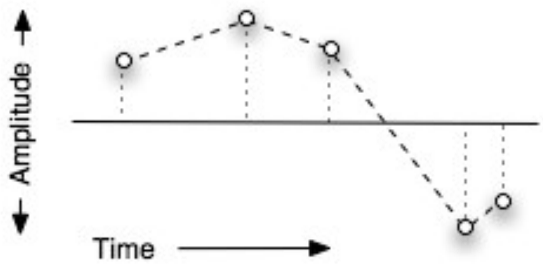


Figure 1. A breakpoint waveform description

The amplitude and time position of each break point are varied at each repetition by a random walk function. A random walk is a constrained stochastic function where each subsequent event is selected at random within a limited range above or below the previous value. The pitch (cycle duration) and dynamic (peak amplitude) are thus constantly varying from cycle to cycle. Each amplitude-time point varies independently and they all change concurrently. Slight variations produce a more stable 'phasing pitch glide' sound and larger variations result in a more 'noisy' timbre. The amplitude and pitch range can be compressed by setting maximum values for each, what Xenakis calls "elastic barriers."

Sample values for the waveform are calculated by interpolating between the stochastically generated amplitude-time points. The resolution of the interpolation is traditionally quite course, adding to the 'low tech' character of the DSS sounds (Xenakis, I. 1991; [2]).

Part of the appeal of this process to Xenakis was its computational efficiency and the fact that it did not rely on a harmonic or Fourier conception of the sound world. He comments that “the challenge is to create music, starting, in so far as it is possible, from a minimum number of premises but which would be ‘interesting’ from a contemporary aesthetic sensitivity, without borrowing or getting trapped in known paths” (Xenakis 1991:295).

Xenakis had been explicitly exploring stochastic systems in his works as early as 1957 with *Achorripsis* and, not long after, with the *ST* (Stochastic) series of compositions. He had been using random walks in his instrumental compositions since the 1970s, the most paradigmatic of which was *Erikthon* (1977) where both note level and larger structures were heavily determined by random walks [1] [3]. The extension to use stochastic functions to control audio signal generation was an insightful one, even if it appears as a natural extension of his work in hindsight.

The use of non-linear processes at the level of microstructure and the constant updating of the waveform, means that DSS has an ‘organic’ quality because it is in a constant state of evolution. The dynamic qualities of continual development and the emergent nature of the sound’s pitch, loudness and timbre result in a system that can produce a wide variety of outcomes, but is not always easy to control. In the early 1990s when Xenakis was able to realise a complete DSS implementation, almost 20 years after the initial conceptions, the system was a non real-time computer assisted compositional tool. This paper outlines some of our attempts to enhance the performability of DSS.

3 Previous DSS Implementations

The original implementation of dynamic stochastic synthesis was in the *Gendyn* program, written by Xenakis in BASIC with the assistance of Marie-Hélèn Serra [4]. The program generated completed works including *Gendy3* (1991) and *S.709* (1992) by employing stochastic processes at both the macro and micro structural levels.

The *Gendyn* program introduces complexities into DSS beyond those described in the previous section. At the level of microstructure these include the cascading of two random walks for each of amplitude and time rather than one, and the use of “mirrors” for maximum boundaries that reflect values above the maximum back into the desired range, rather than “brick wall” limiters that result in significant areas of peak values.

At the level of macrostructure, the *Gendyn* program controls the form of the piece enabling it to produce a complete work, not just a continuous sound stream. The program allows for multiple voices, 16 in the case of Xenakis’ implementation. Short periods of activity or silence in each voice are stochastically determined which produces a kind of chaotic rhythmic counterpoint. Xenakis refers these short periods as ‘fields.’ A probabilistic choice about the number and start location of these fields determines the texture and duration of the piece overall.

A faithful re-implementation of the *Gendyn* program was undertaken by Peter Hoffman in the mid 1990s. This version, which Hoffman called *The New GENDYN Program* [2], was written in C++ and Visual Basic and was efficient enough to run in real-time. Hoffman’s intention was to gain a full understanding of the *Gendyn* program and his implementation even reproduced some of the programming ‘errors’ in the original implementation that had important sonic results. Our implementation, described later in this paper, while building on Hoffman’s work has a different purpose. It intends to make the opportunities of DSS available to the computer music performer, and to add some extensions that broaden the sonic potential even further.

The *Stochos* program [1] employs many functions used by Xenakis including elements of dynamic stochastic synthesis. The application, implemented in Max/MSP, is a computer-assisted composition tool that enables the user to apply its functions to a composition’s macro and micro structure and to work in real-time or render audio files if the processing load exceeds real-time performance limits. As well as enabling DSS as one of the synthesis process in *Stochos*, Bokesoy and Pape apply the DSS concepts of stochastic line segment envelope curves, and mirrors as reflective limiters of stochastic data, as general tools throughout their program. While *Stochos* has a broader range of stochastic sonic functions than *Gendyn* or *New GENDYN*, they each explore the features of sonic space opened up by dynamic stochastic synthesis.

4 Sonic Behaviour

In a stable state where the random walk variations of both amplitude and time are reduced to zero, the DSS process acts as a stable oscillator at a fixed pitch. The precise sonic spectrum depends on the wave shape but the tendency is for a bright ‘buzzy’ tone, not unlike a sawtooth wave. Given enough points in the wave envelope and a lot of luck in freezing the points, a simple tone close to a sine wave is theoretically possible.

The addition of a small amount of amplitude variation creates a warmer chorusing or phasing effect. Allowing slight variations to the time positions of the points results in a wandering pitch glissando, at times not unlike the sound of a Blow Fly. Larger random walk steps in amplitude and time positions cause the sound to become frantic, a result reminiscent of asynchronous granular synthesis with random pitch shifting or of waveshaping. This is not surprising given that Hoffman describes DSS as “a non-linear stochastic distortion of the shape of the waveform over time" [2:32]. Further increasing of the random walk step size causes the timbre to approach Brownian noise or, with tightly constrained time mirror boundaries, white noise.

While the shape of the waveform has a significant effect on the timbre of the sound, the characteristic behaviour of DSS derives from this non-linear change over time. Hoffman suggests that "it is this change that makes up the specific quality of GENDYN sound by continually transforming its spectrum" [2:36]. He goes on to explain that DSS is a dynamic system where the dynamic behaviour of the system is fractal in nature (like all random walks) and, further, that the cascading of random walks means that the sound is governed by 'strange' attractors. When using the system for performance the challenge is to balance the interest generated by instability with a degree of control that enables musical expressiveness. The performer of DSS systems will have better control given an understanding of dynamic system tendencies and the importance of the relationship between the successive random walk states in creating different attractor diffusion patterns.

Another variable in DSS is the choice of stochastic distribution. For example, random functions usually generate a linear distribution, but Gaussian, Cauchy, Logistic and Poisson functions were favourites of Xenakis who explored the different tendencies of each in some depth. The general sound of the DSS system is similar regardless of the distribution function, but the behaviour of the system over time can be significantly affected by changes in stochastic distribution functions and their parameters.

The role of the mirrors in DSS is to constrain the boundary limits of the sonic space. In the case of amplitude change the mirrors act as dynamic compressors and also as peak level limiters. The reflective nature of the mirror boundaries also means that at more constrained amplitude ranges they have a timbral effect in producing jumps in the waveform's fundamental pitch within the overtone series because the reflections cause "transient inner symmetries" in the waveform [2:37]. The time mirrors act to constrain the pitch range, with more restricted boundaries forcing the pitch range into higher and narrower frequency regions. When generating broad spectrum signals (noisy sounds) the time mirrors can act like a high pass filter.

Because the waveform is made up from line segments between points, the resolution of the rendering of those into sample values can effect the sound. Lower sample bit depths result in a 'course' sound, because of associated changes in quantisation noise. This will also effect dynamic range, of course. Changes in sample rate can act as a crude low pass filter by changing the upper frequency limit.

Hoffman's analysis of the sonic behaviour of Xenakis's original algorithm and code were incisive and although his implementation could be altered in real time, the system was focussed more on reflective composition and analysis than real time performance. The aim of our implementation of DSS is to make these behavioural characteristics available to the performing computer musician in a way that he/she can best exploit these tendencies.

5 An interactive implementation

In order to explore dynamic stochastic synthesis in live performance a real-time system was required. In this section we outline the considerations and design of an Interactive Dynamic Stochastic Synthesizer (IDSS) program.

The need for a new implementation was several fold. Xenakis' original *Gendyn* software was not real-time, nor was it available. Hoffman's *New GENDYN* operates in real-time but is not focused on performance control and was designed for the Windows 95/98 platform. The *Stochos* system is possibly more complex than required or even suitable for performance and in any case was not published until after our implementation was written. Finally, the benefits of fully understanding DSS through implementing it in software are significant in their own right, and an informative aspect of this creative/research project.

The IDDS is a real-time application developed in Java and uses the audio system from the jMusic library [5] which includes the DSS algorithm which was added to jMusic as a consequence of this project.

The IDSS program features a graphical user interface (GUI), one channel strip of which is shown in figure 2. The top area of the interface has start and stop buttons, below which are sliders for each of the random walks (1 & 2) that specify the step size (Step1 and 2) and mirror positions (Max1 and 2) for amplitude (A) and time (T).

There is a visual display of the animated waveform and the secondary mirrors are also displayed here. Sliders to the left of and below the display box control the secondary mirrors for amplitude and time respectively. In the bottom left corner of the wave display area are triangular buttons that move secondary time and amplitude mirrors concurrently.

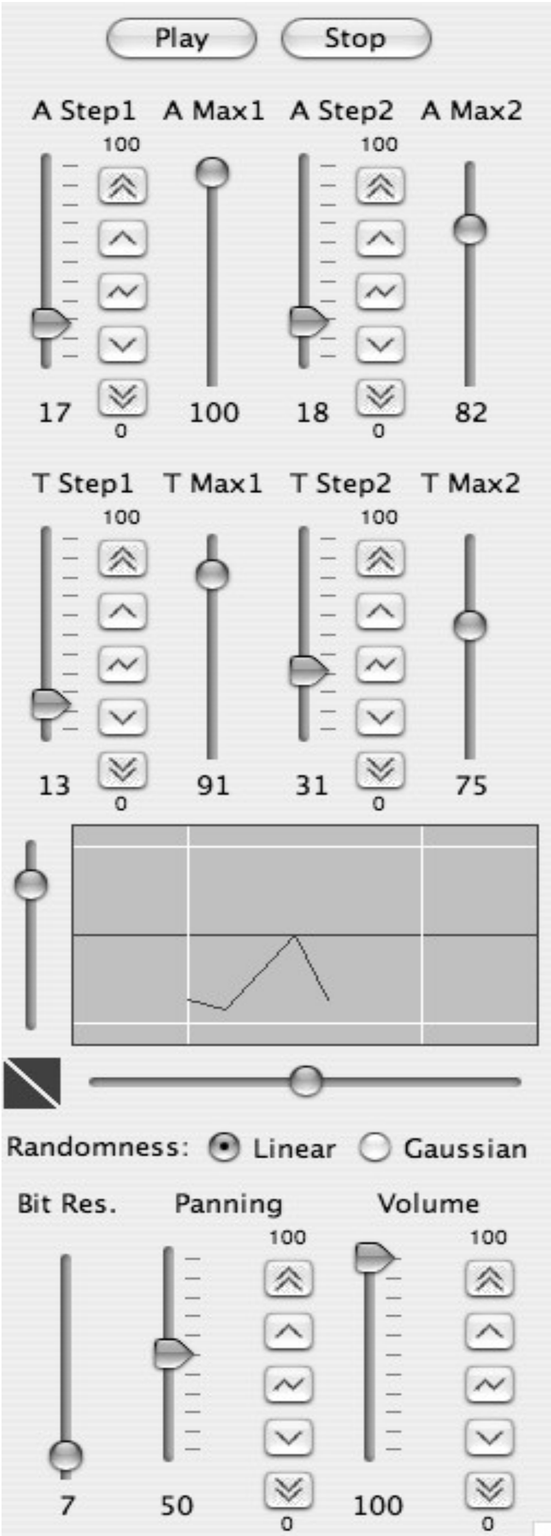


Figure 2. The graphical interface of one IDSS voice.

Below the wave display area are radio buttons for choosing the stochastic distribution type. Linear and Gaussian distributions are available. Below these are sliders to control sample bit depth, panning and volume gain for this voice.

To assist in the performability of the application, selected sliders have associated automated controller buttons. These buttons toggle fades (slow and fast, up and down) and the central button starts fader oscillation up and down repeatedly. The maximum and minimum values between which automation occurs are specified by numbers above and below the button array. These max and min values can be adjusted by clicking dragging on the number display. The automated faders are particularly useful in the control of multiple DSS voices as they offer the performer dynamic control over multiple parameters simultaneously when using the ubiquitous point and click controller (mouse/track pad). Computer keyboard shortcut commands for several functions also improve interaction in this regard. This level of multiple parameter control has been an important consideration of IDSS from its inception, in an effort to avoid the 'serial change' restriction often encountered in 'laptop' performances.

6. Interaction Considerations

During performance each IDSS voice produces alternate active and silent sound fields of variable duration, therefore the main task of playing the software is to vary parameters to stimulate and maintain an interesting polyphonic blend. Random walk step sizes and mirrors can be set so that a voice is quite stable or varies widely of its own accord. Each extreme has it's own challenges for performance. More stable settings involve less risk of drift into inappropriate regions of sonic space, but require more tweaking to maintain interest. Unstable settings can almost create a piece by themselves, not unlike the way Xenakis used his *Gendyn* program, however, control over the outcome is transferred to the algorithm to a large extent and performance consistency can vary widely. Composers using other DSS implementations have the luxury of throwing away unsuccessful outcomes. For example, Xenakis' first DSS piece, *Gendy3*, was named after it's full file name GENDY301, so we can assume that it took Xenakis more than three hundred attempts to arrive at a satisfactory outcome. The intent of our IDSS application is not a "sound -producing automaton creating complete composition 'out of nothing', purely through the structure-generating power of probabilities" [2:31]. but to enable a partnership between the processing algorithms and the live performer. Within this partnership, the processing algorithms determine the detail of the sonic microstructure while the performer is chiefly responsible for the macrostructure.

The live performance risk with generative music systems can be acute at times, and stochastic processes are often quite an imprecise tool. To help alleviate this risk the IDSS program can save and recall settings allowing the performer to act like DJ in scheduling and tweaking prepared sections of work. Unlike a pre-sequenced stream of MIDI information, or banks of sampled loops the generative nature of the real-time system maintains a degree of uncertainty even in this scenario. A skilled performer needs to not only understand every available parameter's operation and interrelation but also the degree of unpredictability likely to result from different relationships between settings.

Like other computer and electronic music performances the juggling of many sound sources is a critical skill. The software has been designed to operate with only the standard mouse and keyboard controllers. The automated controllers in the IDSS program assist the execution of several parallel tasks despite the often awkward one-mouse controller situation. The IDSS program supports assignable MIDI controllers, a feature that greatly enhances real time control and expressive capabilities by enabling the use of physical hardware controllers (knobs and/or slider boxes). With only two hands, such devices rarely permit simultaneous control of more than a few parameters so a combination of qwerty keyboard, mouse and MIDI control hardware is normally employed.

7. Conclusion

The application of a synthesis process in an interactive software environment requires a clear understanding of the operation, behaviour and potential of the process in order to adequately reveal the useful attributes to the musician and hide irrelevant details. We have outlined our approach to this task in relation to dynamic stochastic synthesis and its implementation in the Interactive Dynamic Stochastic Synthesizer. The design and development was informed by an understanding of the historical context, a familiarity with the technical aspects of the process, experimentation with DSS leading to analysis of tendencies in the synthesis process and methods of harnessing them for efficient use in live performance. However, as with all such efforts, our work is ongoing and further performances and exploration will suggest enhancements. Some already planned include stochastic macrostructure controls, meta controller assignment mapped to multiple parameters and an increased choice of probability distributions and interpolation functions.

References

[1] Bokesoy, S. and G. Pape (2003). "Stochos: Software for Real-Time Synthesis of Stochastic Music." *Computer Music Journal* **27**(3): 33-43.

[2] Hoffman, P. (2000). "The New GENDYN Program." *The Computer Music Journal* **24**(2): 31-38.

[3] Matossian, N. (1986). *Xenakis*. London, Kahn & Averill.

[4] Serra, M.-H. (1992). "Stochastic Composition and Stochastic Timbre: Gendy3 by Iannis Xenakis." *Perspectives of New Music* **31**(1): 236-257.

[5] Sorensen, A. and A. R. Brown (2000). *Introducing jMusic*. The Australasian Computer Music Conference, Brisbane, ACMA.

[6] Xenakis, I. (1971). *Formalized Music: Thought and Mathematics in Composition*. Bloomington, Indianna University press.

[7] Xenakis, I. (1991). *Formalized Music*. New York, Pendragon Press.

Granular Synthesis: Conception and Continuity

Timothy Opie

School of Music, Queensland University of Technology
Email: timopie@fastmail.fm
http://www.granularsynthesis.com

Abstract

There are a multitude of computer programs available to perform various tasks using granular synthesis. These programs were not spontaneous creations; they grew from an intellectual tree of knowledge that has been slowly developing for over 200 years. Inside each program resides the ideas of Jean-Baptiste Joseph Fourier, Albert Einstein, Erwin Schrödinger, Dennis Gabor, Iannis Xenakis, Curtis Roads, Barry Truax and others. These composers and scientists are the ghosts in the granular synthesis machine.

1 Introduction

In 1952 Karlheinz Stockhausen wrote to Herbert Eimert stating: "the possibility of realising a *sound-atom* is completely beyond me" [17]. Stockhausen was in a small studio basement in Paris working on his *Study on One Sound*, but his results were unsatisfactory. His dilemma came from the fact that his equipment was not adequate, although he did pinpoint an idea that still interests many musicians today.

The conceptualisation of the sound atom and microsound through granular synthesis has been a long time in development. This paper will identify the contributors to these ideas, and the role they played in identifying key aspects of granular synthesis. It will outline this intellectual tree of knowledge.

2 Knowledge: The Sound Atom

The origins of Granular Synthesis came from the combination of two very fundamental scientific ideas. One idea came from the beginning of the 19th century, the other from the beginning of the 20th century (See Figure 1). Whilst each idea has its own ghosts, I will start from these places as they are milestones from their own specific areas.

At the beginning of the 19th century, Jean-Baptiste Joseph Fourier formulated a theory which has become a key component in many parts of digital sound synthesis. His theory stated that any waveform could be deconstructed into a combination of simple sine waves. The practices of the Fourier Transform (FT) have since been refined and improved, yet the initial idea is still implicit [26].

One hundred years later, Albert Einstein established beyond reasonable doubt, the reality of atoms. Many previous people had thought about the existence of finite particles that combined to make up the world. Some of the ancient Greek philosophers unsuccessfully proposed the idea. Isaac Beeckman toyed with it. Isaac Newton proposed a more definite concept of a particle based view of the world which was, for a short time, accepted but then later rejected. However, it was Einstein's insight and development of the work from Max Planck that established the theory for what was to become known as quantum physics.

As part of this work, Einstein specifically proved that sound and light are made of discrete particles. Niels Bohr developed the Einstein atom into a physical model. Louis de Broglie and Wolfgang Pauli developed the idea on the behaviour of atoms. Werner Heisenberg, and then Paul Dirac worked out the probability matrices from which atomic behaviour can be predicted, and then Erwin Schrödinger took all of these ideas and related them to more easily understood physical ideas. His equation that explained quantum phenomena was so neat and self explanatory it set the way for a whole new breed of quantum physicists [5][11][8][32].

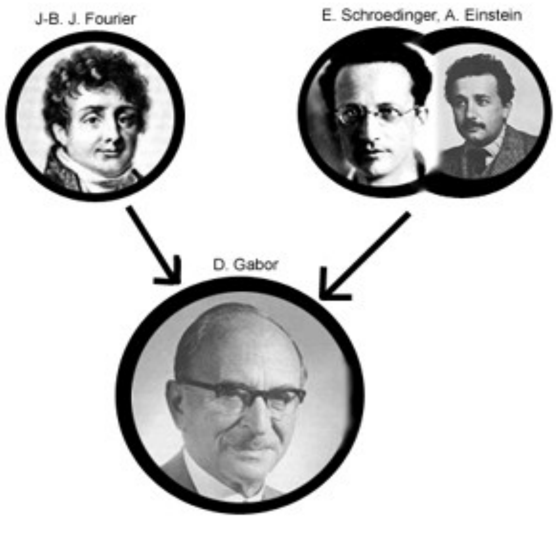


Figure 1: The primary knowledge base of granular synthesis.

Twenty years later at a conference amongst the Institute of Electrical Engineers, Dennis Gabor proposed an idea for reducing the bandwidth required for communication. Gabor was concerned that FT took no account of time. FT worked only in the frequency domain, with the expectation that the

Proceedings ACMC 2004
ISSN: 1448-7780
http://acma.asn.au

sound remained unchanged for an infinite amount of time. Using quantum physics as a guide he decided to quantise the sound. He segmented the sound into discrete portions along the time-domain, and then used FT to analyse the frequency domain, which he put into a grid called the Gabor Matrix. Viewing the matrix he could discard all unused output, send only the time segmented information across the bandwidth and then reconstitute it using an inverse FT within the correct time-domain segment. This idea has been used somewhat in signal analysis, data reduction and compression methods, but it is his definition and research into the division of sound, and how to divide it, that has mostly concerned musical composers [10].

The quantative approach used by Gabor to analyse sound signals differs from an elementary particle in quantum physics. He only wanted to determine an elementary particle in terms of the human perception of sound. He was interested in the threshold of the human ear, and what one perceives as a single elementary sound signal. After determining 21ms at 500Hz to be about the shortest duration of sound perceivable he set about looking at ways to capture and reproduce sounds of that duration. Applying a quantative methodology to sound analysis put him at the disposal of all quantum research that had been conducted over the past 40 years. Through a mathematical process involving FT, time-domain and quantum analysis, he quantised the elementary sound and created the analogy of the sound atom [10].

Gabor's next step was to use this information in a practical sense. He created some machines that could granulate and reconstitute a sound, although only in the time domain, which is where granular synthesis has mostly resided. The most popular of these machines was a kinematic frequency converter based on a tape device, but with multiple rotating heads that could be used to change the time domain of a sound by stretching or condensing the time base without affecting the pitch, or for changing the pitch without affecting the time base [10].

In 1950 Werner Meyer-Eppler gave a talk entitled *Das Klangfarbenproblem der elektronischen Musik* (*The Problem of Timbre in Electronic Music*) in which he made use of Gabor's ideas on musical quanta [26]. Stockhausen, Pierre Schaeffer, and Jacques Poullin were all in contact with Meyer-Eppler during this period, which obviously prompted them to start thinking about this concept. Stockhausen began working in his poorly equipped studio in Paris trying to superimpose sine waves on top of each other, and cut them down to size without even having access to a tape recorder [17]. Schaeffer and Poullin set about creating a similar device to Gabor. They named their device the *Phonogéne*. Later in the 1960s, Herbert Eimert used a similar device created by the company Springer called the *Tempophon*. Whilst their work in that time is interesting, it was Xenakis who came up with the next conceptual idea (See Figure 2).

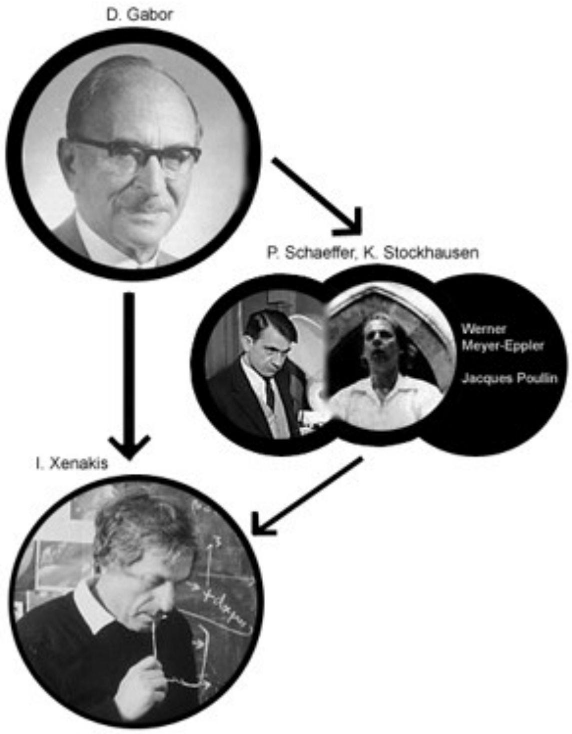


Figure 2: Scientific knowledge becoming a compositional tool.

3 Composition: Splicing

It is Iannis Xenakis who first conceived and executed a composition using granular synthesis. He went through Gabor's paper and then set about formulating an idea which takes the Gabor Matrix and turns it into a cube, a grid of Gabor matrices, which he commonly referred to as a screen. The cube was used for setting up 3D spatialised clouds of sound [33].

In 1959 he used this process for creating the piece entitled *Analogique B*. The clouds were determined stochastically using Markovian indeterminism, following a detailed method outlined in his book *Formalized Music*. Xenakis designated the grains using a number of different array matrices of parameters on a screen, the screen representing the audio and time spectrum. Each cell of the screen was given parameters associated to the grain structure in that audio-time region. Xenakis created books of screens to represent complex audio sounds. The sound source he used for this composition was recorded sine tones generated by tone generators onto analogue tape. The tape was spliced into 40ms fragments and then meticulously pieced together. *Analogique B* was created to complement *Analogique*

A, which was derived using the same stochastic methods, but written for two violins, two violoncellos, and two contrabasses [34][26]. Interestingly, Schaeffer was openly opposed to Analogique B because he didn't like the sound of it, and he did not like the fact that the composition was so scientifically based [19].

Xenakis' book *Formalized Music* and his workshops moved granular synthesis to the next stage; from a knowledge base, through a compositional idea, leading to a computer process.

4 Computer Process: Klang!

Xenakis' efforts to compose with granular synthesis by tape splicing are to be highly commended, but in order to make granular synthesis a more widespread tool for composition it needed to have a digital medium with which to manipulate the grains. The computer made it possible to incorporate many mathematical algorithms directly into the structuring process of the granular texture. In the early days of computing however this was still a very laborious task.

Motivated by Xenakis, the composer Curtis Roads became extremely interested in granular synthesis [22]. Roads first heard about granular synthesis when he attended a workshop conducted by Iannis Xenakis in 1972. Since first learning about granular synthesis, Roads has spent much time researching and writing on the topic [23][24][25][26] (See Figure 3).

Roads first began working with computer sound synthesis on an advanced mainframe, the dual-processor Burroughs B6700. This mainframe was exceptionally good for its time, but using it was still very time consuming. The Burroughs machine was running the program Music V, written in 1966 by Max Mathews, and had to be programmed using the language Algol [21]. All input was done through the use of punched paper cards. Due to storage limits on the mainframe, Roads was limited to one minute of monaural sound at a sampling rate of 20 kHz. It would take several days to process a minute of sound. The mainframe, as was typical, was also unable to produce audio. Roads would take the punch cards to the Burroughs machine. Once it had finished processing, the data would be written to a (large) digital tape. The digital tape would then have to be transferred to a disk cartridge which involved setting up an appointment at the Scripps Institute of Oceanography. Finally, the cartridge could be converted to Audio on a DEC PDP-11/20 at the Centre for Music Experiment [18][26].

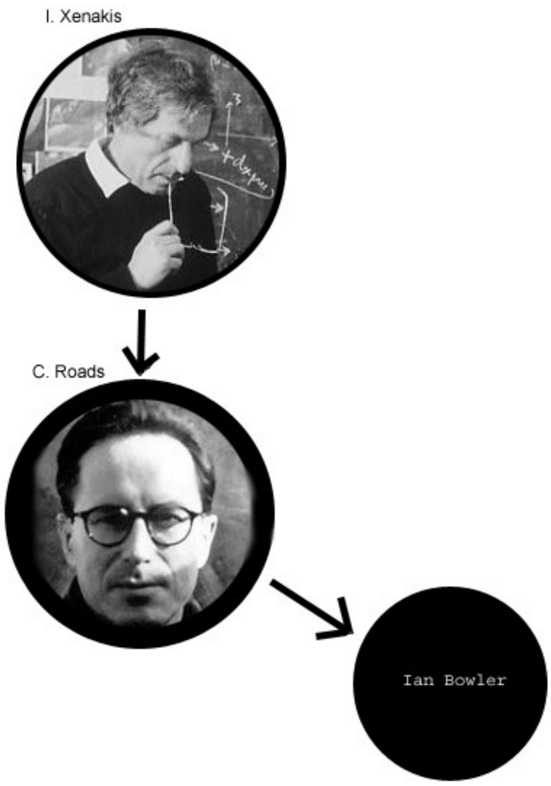


Figure 3: Composing on computers

Using this process Curtis Roads composed the first piece of computer generated music using granular synthesis. It was just a technical experiment entitled Klang-1, written in 1974. This piece of music had a duration of 30 seconds. It was an experiment testing three parameters: Grain Envelope, grain duration and density. Klang-1 featured 766 grains of sound. Each grain needed to be programmed onto a separate punch card. Roads treated each grain as if it were an individual note, just with a very short duration; specifying the frequency, start time and duration of each note. Roads fixed the duration of each grain at 40ms, which was identical to the grain duration of Xenakis' Analogique B. The densities ranged from 0 to 25 grains per second. The grain envelope was a simple Gaussian envelope. The grains were all generated as sine waves with frequencies from a 1000-tone scale ranging from 16.11Hz to 9937.84Hz. At the completion of this experiment of granular synthesis on a computer Roads made the following statement:

> I did not intend *Klang-1* to be anything more than a technical experiment. Yet I vividly recall the magical impression that the granular sound had on me as it poured forth from the computer for the first time [26].

The second piece of music, using granular synthesis, realised on a computer, was entitled *Prototype* and was written in 1975. This was an actual composition rather than an experiment. To compose this piece Roads wrote a program called *PLFKLANG* in *Algol* which could generate thousands of grain specifications based on seven high-level cloud parameters. Defining parameters as clouds is significant because it can define an entire texture of grains, whereas the grain is discrete as an individual particle [26].

Later versions of the computer program *Music V,* maintained and developed by Daniel Arfib, contained probably the first code written to perform analysis-synthesis in terms of the Gabor grain. Risset used the granular analysis-synthesis tools of Music V, on an IBM-PC, to perform time stretching of recorded samples in his piece entitled *Attracteurs Etranges* [22].

Following on from the work of Roads, Ian Bowler in 1985 proposed a process called Switched Waveform Synthesis, based on granular synthesis. Its purpose was to create different timbres within the constraints of a microcomputer [3]. This idea however, did not get developed very far.

5 Real-Time: Anytime

The next conceptual idea for granular synthesis came in 1986 through Barry Truax. Truax wrote a computer program that allowed him to use granular synthesis in real-time. It was a breakthrough that has made granular synthesis a much easier process to work with. It led the way to a barrage of programs that all focus on automating various parts of the granular synthesis process, allowing it to be played back in real-time (See figure 4).

Barry Truax started learning about granular synthesis after reading an article published in the *Computer Music Journal* by Roads in 1978 entitled *Automated Granular Synthesis of Sound* [23]. Truax's approach to granular synthesis was very different to that of Roads. Truax pioneered real-time granular synthesis right from the onset of his research. In 1986 he wrote a real-time application of granular synthesis implemented using his previously developed *POD & PODX* system. At the time Truax implemented the real-time granular synthesis algorithms in *PODX* he was using a DMX-1000 Digital Signal Processor controlled by a DEC LSI-11/23 microcomputer. Using both the DMX-1000 and the DEC LSI-11/23 meant that the sound generation could be handled by a processor optimized for audio, whilst the microcomputer would only have to be concerned with number-crunching and outputting a stream of numbers [27][28][30][31]. The computer was set up so that the granular synthesis parameters could be changed in real-time with a single key stroke. Truax was also able to store a range of presets and assign them to any unused key on the keyboard. The parameters could also be set using a *tendency mask*, which is a graphical representation of a range over time. The graphical tendencies masks would be processed as a combination of presets and ramps [29]. *Riverrun* was the first piece of granular synthesis based music Truax composed using the PODX real-time system.

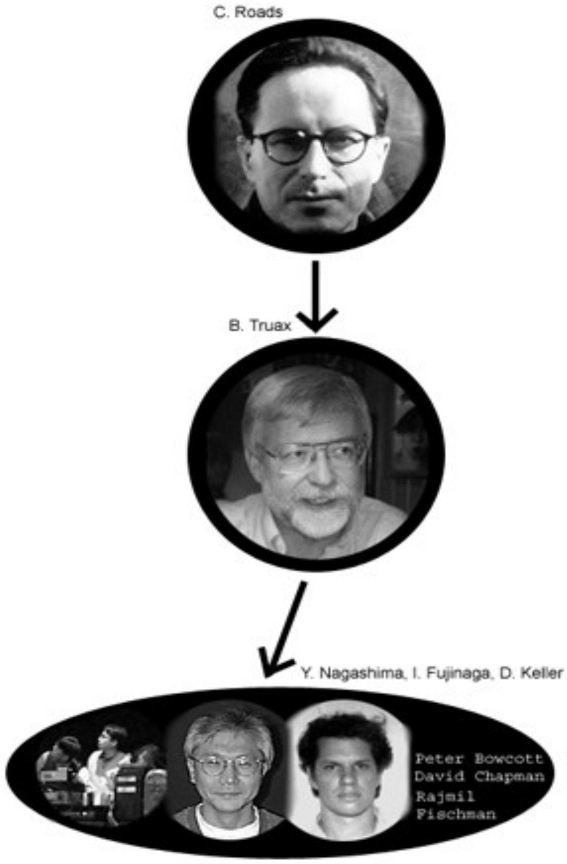


Figure 4: Computer Extensions

A most notable feature of the work by Barry Truax was his attention to control. The PODX system was the first in a whole wave of different control ideas for automated granular synthesis on a computer. There have been many other control methods designed over the last 15 years by a large range of composers.

Using genetic algorithms as a control method for grain distribution has been experimented with by Peter Bowcott, Ichiro Fujinaga, Horner, M. Hamman, and Goldberg [12]. Bowcott based his genetic model on Cellular Automation [2]. Bowcott has also used population and growth models upon which to control grain distribution [1]. David Chapman, Michael Clarke, Martin Smith, and Paul Archbold researched and created a computer language called *AMY* to help them control and study fractal based granular synthesis [4]. Ecologically-based granular synthesis, as a means of control, has been experimented with by Damián Keller and Barry Truax. What is interesting in their approach to sound organisation is their view of what they are trying to achieve. They pointed to a statement made by Dannenberg to the effect that there is a lack of research done in sound organisation for time scales ranging from 100ms to a several seconds. The micro-scale synthesis issues have been well researched, such as synthesis and analysis of single acoustical quanta. What Keller and Truax began researching were short recognisable natural sounds, such as bouncing, scraping, breaking and filling. They created algorithms to create these types of sounds in *Csound* and *Cmask* using granular synthesis [16]. Yoichi Nagashima has used neural networks to automate and refine the control of grain distribution in a project called PEGASUS [20]. Rajmil Fischman has been experimenting with the Schrödinger equation as a control method for grain distribution [8].

As we can see, there is an abundance of ideas for control and creation of granular synthesis available. The conception and continuity of granular synthesis has been uncovered. The question then becomes, how do we ourselves make use of these ideas? Where are these ghosts actually revealed?

If we are computer programmers, we can simply begin assimilating the maths and science and start programming. There are even some programming languages and libraries available, which will allow you to create your own program, but still work within a framework that simplifies a lot of common commands. A program such as *jMusic* gives very low level control of the sound and implementation structure. Some higher level programs and music composition languages available include *Supercollider, Csound, Max/MSP, AudioMulch*, and *Reaktor*.

For general usage and performance, setting up an entire programming language to perform a simple task might be like using a sledge hammer to crack a nut. In these cases there are a plethora of applications that perform specific granular synthesis tasks.

An extremely powerful program for time stretching using granular synthesis is *MacPOD*, based on the *GSAMX* functionality of *PODX* [30], and written by Damián Keller and Chris Rolfe for Macintosh [14][15]. The algorithmic compositional program called *AL & ERWIN* (AL for Algorithm and Erwin as in Erwin Schrödinger) is a great tool for experimenting with granular clouds based on the Schrödinger equation from quantum mechanics [7]. *Amber* by Matthew Mccabe is a program designed specifically for creating interesting and complex sounds [35]. *Cloud Generator* by Curtis Roads and John Alexander is a non real-time program for creating simple granular clouds and makes a good learning tool [36]. *Granulator* by Nicholas Fournel, and *Granulab* by Rasmus Ekman both allow the composer to experiment with granular synthesis in real-time. They give the composer control over the standard parameters such as the shape and size of the grains, and also the density of the overall texture [37][38].

With this array of programs available, and many more not even mentioned, it is no wonder that granular synthesis is seeping into the commercial music world. DJ Spooky and Fatboy Slim have both made good use of granular synthesis, especially with Fatboy Slim's, *The Rockafeller Skank,* making the pop charts across the world [6].

There are still more branches to add to this tree of intellectual and compositional knowledge in granular synthesis. They may be small branches, such as control algorithms, and implementation programs, or they may be larger conceptual ideas. Are we happy just using the branches currently available, or do we want to create our own branches?

6 Conclusion

Through the combination of the Fourier Transform and quantum physics, Dennis Gabor has allowed the idea of the sound atom to become a real building block in modern music. It was initially just a concept for communication, but through the work of Xenakis, Roads, Truax, and many others, it has become a compositional and musical tool with a large spectrum of uses.

As the concept of granular synthesis is developed further, there will be more ghosts, in the granular synthesis machine, that will help drive future programs and ideas.

References

[1] Bowcott, P. 1990. High Level Control of Granular Synthesis using the Concepts of Inheritance and Social Interaction. *ICMC Proceedings 1990*, pp 50-52. ICMA.

[2] Bowcott, P. 1998. Cellular automation as a means of high level compositional control of granular synthesis. *ICMC Proceedings 1998*, pp 55-57. ICMA.

[3] Bowler, I. 1985. Switched Waveform Synthesis. *ICMC Proceedings 1985*. ICMA.

[4] Chapman, D., M. Clarke, M. Smith, and M. Archbold. 1996. Self-similar grain distribution: A fractal approach to granular synthesis. *ICMC Proceedings 1996*, pp 212-213, Hong Kong. ICMA.

[5] Cohen, H. 1984. *Quantifying Music*. D. Reidel Publishing Company, Dordrecht.

[6] Fatboy Slim 1998. *The Rockafeller Skank*. Skint Records, Brighton.

[7] Fischman, R. 2003. Al & Erwin. `http://www.keele.ac.uk/depts/mu/staff/Al/Al_software.htm`

[8] Fischman, R. 2003. Clouds, Pyramids, and Diamonds: Applying Schrödinger's equation to Granular Synthesis and Compositional Structure. *Computer Music Journal*, 27(2), pp 47-69

[9] Franz, M. 1993. Physics part ii. *Mechanics*. Videocassette (VHS) (28 min.). Appleseed Productions (Vic.).

[10] Gabor, D. 1946. Theory Of Communication. *The Journal of the Institution Of Electrical Engineers*, 93(3) pp 429-457.

[11] Gribbin, J. 1984. *In Search Of Schrödinger's Cat*. Transworld Publishers, London.

[12] Hamman, M. 1991. Mapping complex systems using granular synthesis. *ICMC Proceedings 1991*, pp 475 478. ICMA.

[13] Harley, J. 1997. Iannis Xenakis:Electronic Music. *Computer Music Journal*. 22(2) PP 75-76.

[14] Keller, D. and C. Rolfe. 1998. The corner effect. *Proceedings of the XII Colloquium on Musical Informatics*, `http://www.thirdmonk.com/Articles/CornerEffect/CornerEffect.html`

[15] Keller, D. and C. Rolfe. 1998. *MacPod*. Real-time asynchronous granular synthesis software for the Macintosh PowerPC. Third Monk Inc. http://www.thirdmonk.com

[16] Keller, D. and B. Truax. 1998. Ecologically-based granular synthesis. *ICMC Proceedings 1998*. ICMA.

[17] Maconie, R. 1990. *The works of Karlheinz Stockhausen*. Clarendon Press, Oxford.

[18] Manning, P. 1995. *Electronic & Computer Music*. Oxford. Clarendon Press.

[19] Matossian, N. 1991. *Iannis Xenakis*. Halstan and Co. England.

[20] Nagashima, Y. 1992. Pegasus. *ICMC Proceedings 1992*. ICMA.

[21] Phillips, D. 2000. *Linux Music & Sound*. No Starch Press, San Francisco.

[22] Risset, J.-C. 1991. Timbre Analysis By Synthesis: Representations, Imitations, and Variants for Musical Composition. *Representations of Musical Signals*, pp 7-43. MIT Press, Cambridge.

[23] Roads, C. 1978. Automated Granular Synthesis of Sound. *Computer Music Journal* 2(2) pp 14-26

[24] Roads, C. 1988. Introduction to granular synthesis. *Computer Music Journal*, 12(2) pp 11-13.

[25] Roads, C. 1996. *The Computer Music Tutorial*. MIT Press, Cambridge.

[26] Roads, C. 2001. *Microsound*. The MIT Press, Cambridge.

[27] Truax, B. 1986. Real-time granular synthesis with the dmx-1000. *ICMC Proceedings 1986*, pp 231-235. ICMA.

[28] Truax, B. 1987. Real-time granulation of sampled sounds with the DMX-1000. *ICMC Proceedings 1987*, pp 138-145. ICMA.

[29] Truax, B. 1988. Real-time granular synthesis with a digital signal processor. *Computer Music Journal*, 12(2) pp 14-26.

[30] Truax, B. 1999. Granulation of Sampled Sound. `http://www.sfu.ca/~truax/gsample.html`

[31] Truax, B. 1999. POD & PODX System Chronology. `http://www.sfu.ca/~truax/pod.html`

[32] Wiener, N. 1964. Spatio-Temporal Continuity, Quantum Theory And Music, *The Concepts of Space and Time*. pp 539-546. D.Reidel Publishing Company, Dordrecht.

[33] Xenakis, I. 1971. *Formalized Music*. Indiana University Press, Bloomington.

[34] Xenakis, I. 1997. *Electronic Music*. Albany NY. EMF.

[35] McCabe, M. 2003. *Amber*. `http://www.gnu.org/directory/hobbies/music/amber.html`

[36] Roads, C. and Alexander, J. 1997. `http://www.create.ucsb.edu/htmls/code.html`

[37] Ekman, R. 2003. Granulab. `http://hem.passagen.se/rasmuse/Granny.htm`

[38] Fournel, N. 2002. Granulator. `http://www.nicolasfournel.com/granulator.htm`

Global Dynamics Approach to Generative Music Experiments with One Dimensional Cellular Automata

Dave Burraston and Ernest Edmonds

Creativity and Cognition Studios (CCS), Faculty of Information Technology, University of Technology, Sydney, Australia.
Email: dave@noyzelab.com, ernest@ernestedmonds.com

Abstract

One Dimensional Cellular Automata (CA) offer the use of emergent computation and behaviours as compositional aids to the generative music process. Global dynamics and rule clustering are important concepts in CA research, and CA music research to date has not addressed these topics. Global dynamics and rule clusters offer a new perspective of CA based on the topology of attractor basins. A methodological approach is described as a foundation for future experimentation. Reflective practice techniques will be used as experimental method. Evaluation of results will be judged against recognised criteria. A one dimensional CA is chosen for a foundation experiment in generative music production and visualisations of this will be presented. This work will give CA based generative music a significant shift of context and awareness for the generative artist.

1 Introduction

Algorithmic and computational processes are an important tool for the technology based creative artist producing generative music [3]. CA offer the use of emergent computation and behaviours as compositional aids to the generative art process. CA have been utilised in a number of novel applications in MIDI based computer music [2]. Complexity theory demonstrates that complex systems of simple units, such as the cells in a CA, produce a variety of behaviours. Complex systems such as CA produce global behaviour based on the interactions of these simple units.

CA were conceived by Stanislaw Ulam and John von Neumann in an effort to study the process of reproduction and growths of form [1]. CA are dynamic systems in which time and space are discrete. They may have a number of dimensions, single linear arrays or two dimensional arrays of cells being the most common forms. The CA algorithm is a parallel process operating on this array of cells. Each cell can have one of a number of possible states, sometimes expressed as **k**. The simultaneous change of state of each cell is specified by a local transition rule. The local transition rule is applied to a specified neighbourhood around each cell, sometimes expressed as **r**. The number of cells is given as a number **L**.

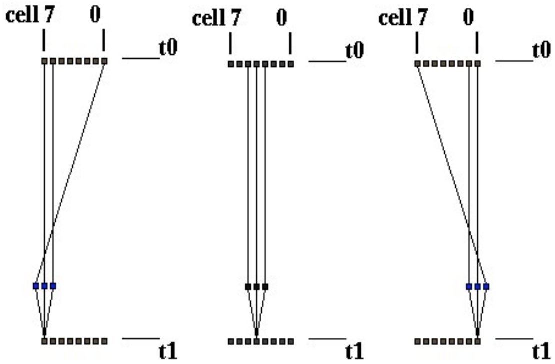


Figure 1. Wiring of periodic boundary cells (left and right) and cell 4 (centre) of an 8 cell (L=8) 1D CA.

CA are usually, but not always, infinite in length. Cells are commonly wrapped around at the edge of the array during the local transition rule computation, to achieve a conceptual infinite array. In this case the array is finite, but unbounded and is said to have periodic boundary conditions. One dimensional binary CA (**k=2, r=1**) have been classed by Wolfram with one of four behaviours [10] as shown in Table 1.

Class 1	**Class 2**	**0lass 3**	**Class 4**
Patterns disappear or become fixed	Patterns evolve to periodic structures	Patterns become chaotic	Patterns grow into complex forms.

Table 1. Wolfram's CA behaviour classes.

The wiring for the edge cells and cell 4 of an 8 cell 1D CA is shown in Fig. 1. Here we can see two time steps of the system from **t0** to **t1**. The transition rule is specified by an 8 bit binary number between 0 and 255 and an example for rule 110 is shown in Fig. 2. The 8 entries in a rule transition table are defined as **T7** to **T0** left to right.

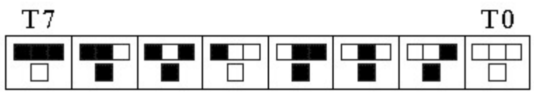


Figure 2. Transition rule table (rule 110).

The λ parameter, introduced by Chris Langton (1986), is an order (Class 1 and 2) – complexity (Class 4) – chaos (Class 3) parameter [7]. This is a kind of virtual tuning knob through the classes of CA rule space. Although the λ parameter appears quite useful, it should be used with care. Langton points

Proceedings ACMC 2004
ISSN: 1448-7780
http://acma.asn.au

out that it does have weaknesses and will not always be able to work correctly. The **λ** parameter is seen as a useful tool when dealing with CA of larger **k** and **r**, and future work may benefit from this.

Langton also supported and promoted work on the global dynamics of CA, formally defined by Wuensche and Lesser (1992), which offers a new perspective based on the topology of attractor basins, rule symmetry categories and rule clustering [11]. In this work an atlas of these basins is presented for a variety of small CA sizes up to about 15 cells depending on the particular rule. Here one can compare basin topologies and measures between rules to gain insight into different rule behaviours. Wuensche's Discrete Dynamics Lab (DDLab) software allows for the exploration of global dynamics [13]. A brief overview of Wuensche and Lesser's work on global dynamics, the Z parameter, and rule clustering is presented in the next section.

2 Global Dynamics and Rule Clustering

A CA state space consists of all possible global states. In a finite deterministic CA all state transitions must eventually repeat with period 1 or more. States are either part of an attractor cycle or lie on a transient leading to the attractor cycle. If a transient exists there will be states unreachable by any other states at the extremity. These extremities are called garden of Eden (**goE**) states. All transients leading to an attractor, and the attractor cycle, is termed the basin of attraction (**boa**) of that individual attractor. An example basin of attraction is shown in Fig 3. State space for a particular CA rule and size is populated by one or more basins of attraction, termed the basin of attraction field. The **boa** field may contain **equivalent** basins, where the states of other basins are rotationally symmetric.

DDLab constructs **boa** fields by the computation of pre-images of all states by a reverse algorithm, and can suppress equivalent basins during display. Measures can be taken of the number of basins in the field and garden of Eden states, the attractor periods, basin sizes, maximum transient and cycle period.

Order		Complexity	Chaos
Class 1 : Z = 0.25	Class 2 : Z = 0. 5	Class 4 : Z = 0.75	Class 3 : Z = 1

Table 2. Z parameter for Wolfram's classes.

The Z parameter is an order – complexity - chaos parameter varying from 0 (order) to 1 (chaos) and is calculated from the rule table. There are exceptions to the Z parameter, such as some chaotic behaviour at low Z values. The Z parameter is given for Wolfram's classification in Table 2.

Attractor basin topology reflects the dynamics of a CA rule and can be used as a method of identifying ordered, complex and chaotic behaviour. The **goE** density and in-degree frequency distribution of pre-images are two measures of CA dynamics [12]. The in-degree is the number of pre-images of a node in the **boa** field and is plotted as a histogram. Ordered rules are classified by very high **goE** density and large numbers of high in-degrees. Complex rules will have medium values for both. Chaotic rules have a low **goE** density and large numbers of low in-degrees.

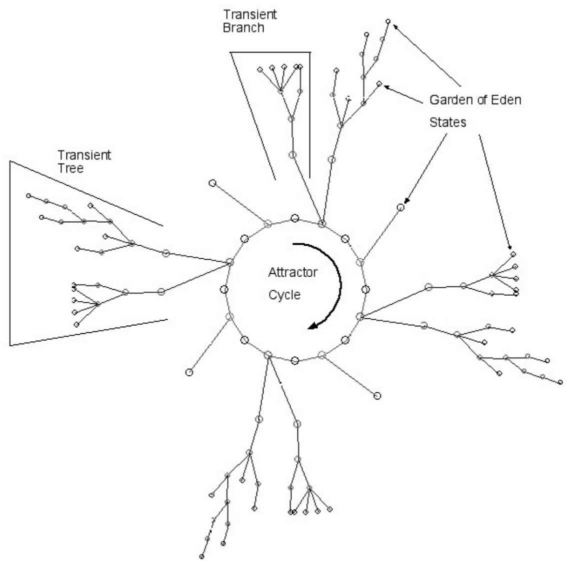


Figure 3. Basin of attraction.

The 1D **k2r1** rules can be grouped into 88 equivalence classes, by **negative**, **reflective** and composite **negative reflective** transforms, with a maximum of four rules per equivalence class [11]. In a rule cluster three rule table transformations are defined as shown in Table 3. The **negative** transform is inversion followed by four pairs of entries having their entries swapped. For example in the first pair **T0** becomes **T7** and **T0** becomes **T7**. The remaining pairs indicated are swapped in the same manner. The **reflection** transform involves pair swapping {**T6 <-> T3**} and {**T4 <-> T1**}. The **complementary** transform inverts the rule table contents. The rule cluster axis is shown in Fig 4. and the rule cluster layout is shown in Fig 5.

Negative	Rule table is inverted & T pairs swapped : {**T0** <-> **T7**}, {**T1** <-> **T6**}, {**T2** <-> **T5**}, {**T3** <-> **T4**}
Reflection	T Pairs swapped : {**T6** <-> **T3**} and {**T4** <-> **T1**}
Complementary	Rule table entries are inverted

Table 3. Rule Transformations.

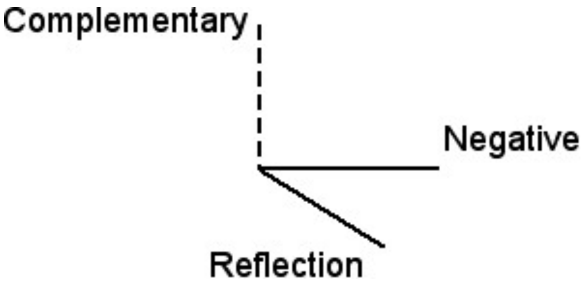


Figure 4. Rule Cluster Axis.

The lowest rule number (**R**) identifies a cluster and is always positioned in the top left corner. The negative (**Rn**) and reflection (**Rr**) transforms are identified, along with a further composite transform the negative reflection (**Rnr**). The complement transform (**Rc**) also has corresponding negative (**Rcn**), reflection (**Rcr**) and negative reflection (**Rcnr**) transforms.

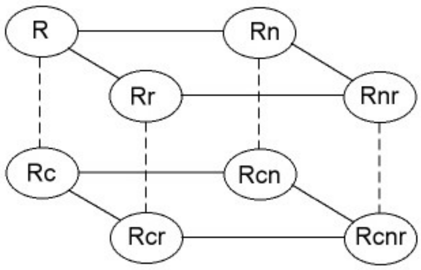


Figure 5. Rule cluster layout.

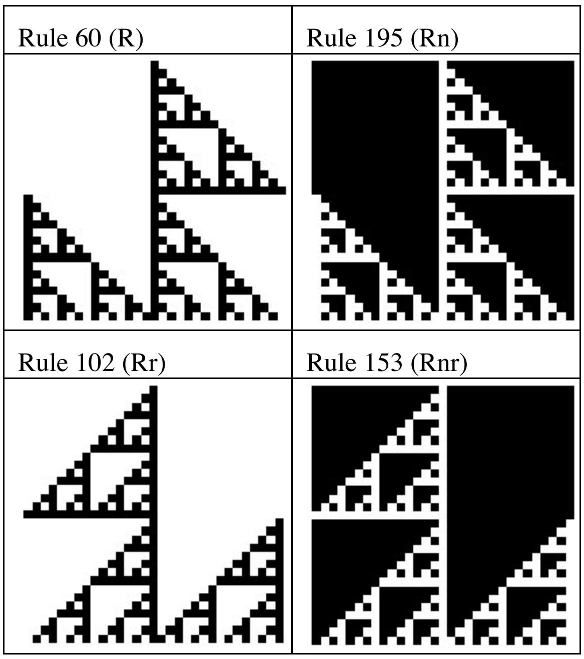


Figure 6. Rule 60 equivalent spacetime patterns.

Each rule cluster has two **boa** fields, one for the top layer (**R**) and one for the complementary bottom layer (**Rc**). Rule clusters contain 2, 4 or 8 different rules depending on whether the transformations result in the same rule number. For both layers the other transformed rules have identical **boa** measures and the states are simply related by being negative (**Rn**), mirror image (**Rr**) or both (**Rnr**), as shown in Fig 6. In some cases the clusters collapse further leaving no complementary rules, resulting in a single **boa** field. The symmetry categories are symmetrical, semi-asymmetrical and fully asymmetrical contained in a total of 48 rule clusters. Symmetric rules contain 2 or 4 rules per cluster, Semi-asymmetric rules have no collapsed clusters and always contain 8 rules. Fully asymmetric rules contain 4 or 8 rules per cluster. Examples of one of each possible cluster type for all categories is shown in Figs. 7a-7c.

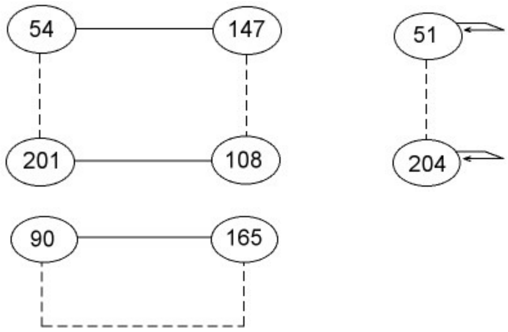


Figure 7a. Symmetric rule clusters, three types.

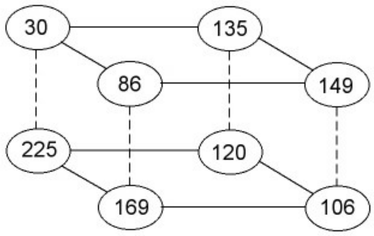
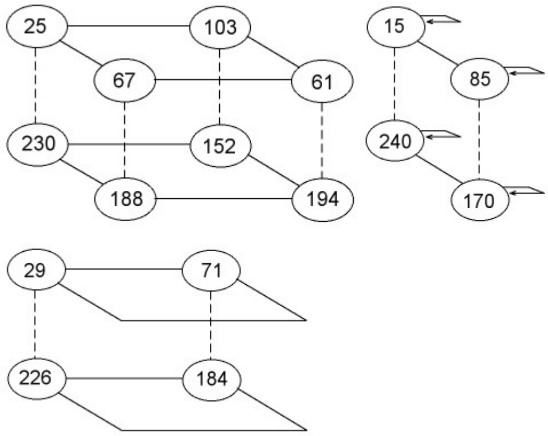


Figure 7b. Semi-asymmetric rule clusters all have this type.



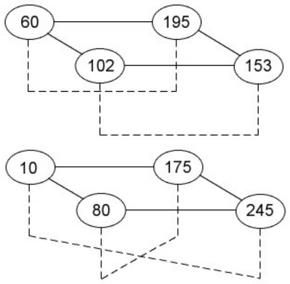


Figure 7c. Fully asymmetric rule clusters, five types.

Rule cluster 15 is important because it exhibits behaviour that is directly comparable to that of analogue sequencers. The cluster diagram for rule 15 is shown in Fig 7c. Rules 170 and 240 are known as the left and right shift rules, and form the complement part of the cluster. Any input pattern will be shifted left (rule 170) or right (rule 240) by one cell per CA step. Analogue sequencer modules commonly have 8, 12 or 16 steps and can be stepped sequentially one stage at a time from between 1 and the maximum number of steps available. This operation is from left to right and vice versa. This is analogous to rule 170 and 240 with a single active cell.

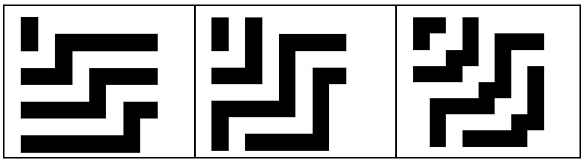


Fig 8a. Rule 15 (wobbly right) L = 8 with seeds of 128, 160 and 208

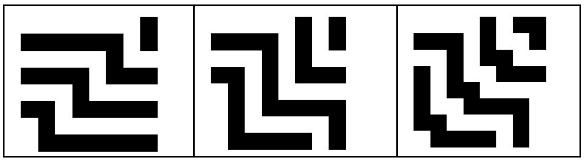


Fig 8b. Rule 85 (wobbly left) L = 8 with seeds of 1, 5 and 11

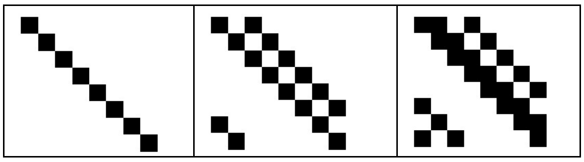


Fig 8c. Rule 240 (right shift) L = 8 with seeds of 128, 160 and 208

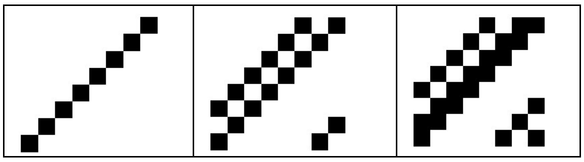


Fig. 8d. Rule 170 (left shift) L = 8 with seeds of 1, 5 and 11

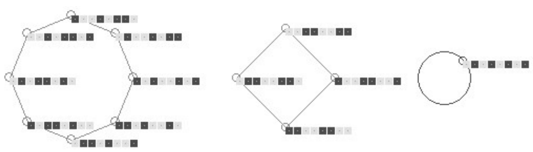


Fig. 9a. Three example basins for rule 15 (wobbly right) L = 8.

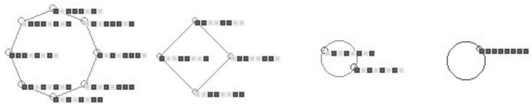


Fig. 9b. Four example basins for rule 240 (right shift) L = 8.

Example spacetime diagrams for each rule are shown in Figs 8a – 8d. In left shift and right shift rules the cycle length is never greater than L. Rules 15 and 85 have this shifting tendency but patterns appear “wobbly” and for some values of L, cycle length can be greater than L.

For rule cluster 15 all states exist on cycles and thus there are no goE nodes, and no transients or preimages on trees. Some example basins are shown in Figs 9a and 9b. Both rules are Z = 1 but appear to be ordered rules and are not chaotic as the Z parameter suggests.

3 Methodology

Reflection-in-action is a method of researching within a practical context [8]. The structure of reflection-in-action in practice based research, termed reflective practice, suggests move testing experiment as a useful methodology. In move testing an action is undertaken to produce intended change. Good results will affirm that move and bad results will negate it. Importantly unintended results that are good, should also affirm the move. Reflection on outcome and potential for future directions will be discussed.

Technical criteria for evaluating sound synthesis techniques were suggested by Jaffe (1995), however these are recognised as guides not rigid definitions [6]. These criteria were suggested in the context of sound synthesis, however some of the criteria applies to generative music production with MIDI. The most appropriate criteria in this context will now be discussed. How efficient is the algorithm? Defined as an extremely important, yet context dependant criterion. Three efficiency categories suggested are memory, processing and control stream. How sparse is the control stream? Related to efficiency and has particular impact on a real time system. How robust is the sounds identity? Asks if parametric adjustment alters the sound too extremely. What classes of sounds can be represented? Conversely this asks if a broad range of sound classes can be achieved by parametric adjustment.

Castagne and Cadoz (2003) revisited and extended these guideline criteria in the context of physical modelling sound synthesis [4]. The context of physical modelling is relevant because it addresses the entire musical creation process and does not just represent a sound synthesis perspective. Ten criteria for evaluating physical modelling were defined in four areas : computer efficiency, phenomenology, usability of scheme and environment for using the scheme. The efficiency area is similar to Jaffe's proposed scheme. An extended phenomenological criteria is diversity of context and relates to Jaffe's robustness and sound class criteria. In usability particular emphasis is placed on modularity. Modular principles are a very important criterion in obtaining generality, power and simplicity. The environment for using the scheme has two criterion. The first studies if generation algorithms already exist and how effective they are. The second asks if there is a friendly musician oriented environment for using this scheme. Results of this work will be judged on these criteria.

4 Generative Music Experiment

The move testing generative music experiment is defined as : **Action** : Construct visualisations and Max patch with a small, single rule CA. **Intended Change** : Simple, easy to use system with future scalability.

The Max patch was implemented with an 8 cell rule 30 using Bill Vorn's 1D CA Max external. This is part of Vorn's LifeTools, available on the internet, and forms part of the IRCAM library of Max objects [9]. Rule 30 has been chosen simply because this is the default rule in Vorn's external. Once the Max patch is built as a standalone application this will be the only rule available. Changing the rule is possible when using the patch within Max. A CA size of 8 cells was chosen to directly map to a simple 8 voice system.

The approach taken in building an experimental music system has some similarity to the University of York's Cellular Automata Workstation, for the Atari ST [5]. The work described here represents a different approach due to the global dynamics perspective, and the system is viewed as a modular foundation for future work.

A simple realtime system, named CA Simplistic Selector (CASS), was produced to map the CA cells to a set of adjustable MIDI events. This set of events will be termed a Cellular MIDI Event (CME) module. The system thus comprises of 8 CME modules, each module is connected to a unique corresponding cell. A block diagram of the experimental system is shown in Fig. 10.

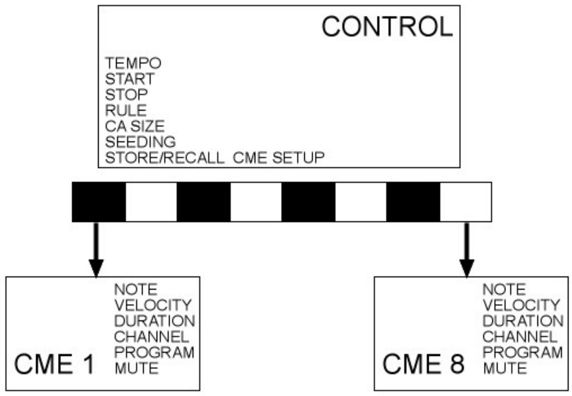


Figure 10. Block diagram of CASS experimental system.

Each CME is controllable in real time by the performer. The CME comprise of note, velocity, duration, MIDI channel and program controls. The system can thus be split up over MIDI channels as desired for multitimbral or one channel for single voice operation. A keyboard for each CME allows for note selection and playing within a seven octave range, a dialogue allows for note input for the full 0 – 127 range. MIDI velocity, channel and program can be input over their complete ranges. Duration is entered by a slider or dialogue from 10 to 10,000 milliseconds.

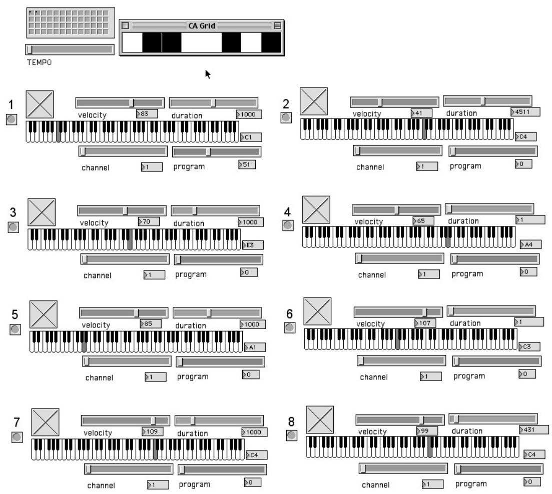


Figure 11. CASS screenshot.

The event onsets are controlled by the 1D CA, which may be seeded randomly, with a list or by a mouse. Individual CME's can be muted during performance. The performer has overall control of tempo and modifies the events to be triggered while the CA runs. This presents an interesting juxtaposition between performer and automated machine. A screen shot of the CASS system is shown in Fig. 11 and a close up of the CME module is shown in Fig 12. All of the CME settings can be saved and recalled for the whole system with the Max preset object.

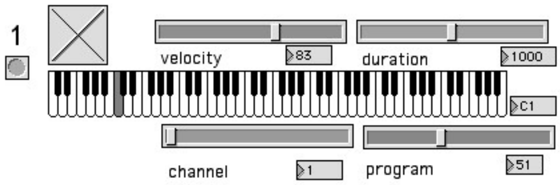


Figure 12. CME module.

The I/O functionality of Vorn's CA Max external is shown in Fig. 13. The inputs have been categorised as seed, clock and data. A number of seeding options are available and the output can be taken in the form of bangs from each cell or dumped as a list of active cells to outlet 2. Resetting of the CA by message is possible, either to an initial seed or to 0 for all cells.

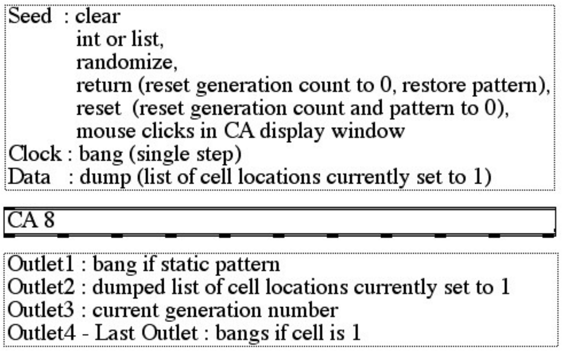


Figure 13. I/O functionality of Vorn's CA Max external.

The CME modules are wired directly to the cell outputs. The CA external will send out a bang from outlet 1 if the output remains static. This allows for the inclusion of automatic reseeding mechanisms to be built within a patch. The current generation count is output as an integer from outlet 3.

Rule 30 is from the semi-asymmetric category and the cluster is shown in Fig. 7b, in Section 2. A boa field for the eight cell Rule 30 was constructed with DDLab and contains 256 states. The main boa is of period 40 and contains 224 states as shown in Fig 14a. The longest transients in this boa contain 16 states and 2 states for the shortest. The remaining states are contained within four boa and are shown in Fig. 14b labelled in decimal. These are a period 8 attractor, two single state period 1 attractors and a two state period 1 attractor. DDLab gave the Z parameter as 1, indicating a chaotic rule. The boa field data was saved to a file and some of the measures obtained are shown in Table 4. The measures in Table 4. agree with the values in Wuensche and Lesser's atlas [11].

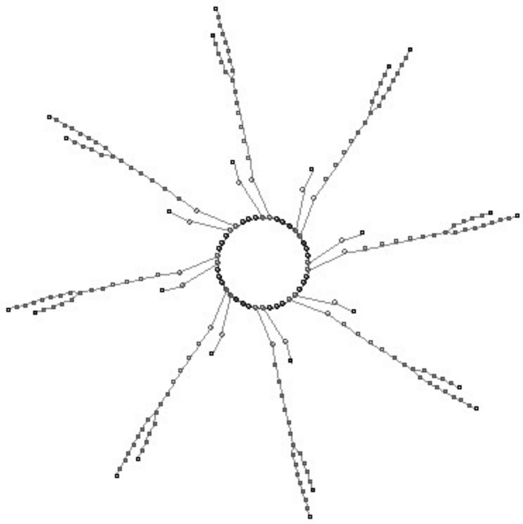


Figure 14a. Eight cell rule 30 basin of attraction period 40.

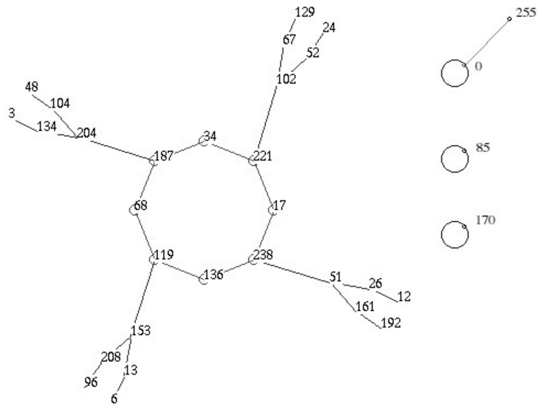


Figure 14b. Eight cell rule 30 basins of attraction with nodes labelled in decimal, period 8 (left) and 3 x period 1 (right).

typ e	no. basin s	period	no. states	% of states	go E	max transien t
1	1	1	2	0.781	1	1
2	1	40	224	87.5	24	16
3	1	8	28	10.9	8	3
4	2	1	1	0.781	0	0

Table 4. Summary of eight cell rule 30 boa measures.

DDLab can plot the state space matrix in decimal, showing the state space in terms of the left and right half of each states bit string. This results in an 8x8 grid showing every possible state in the field. The state space matrix of the entire field of rule 30 is shown in Fig. 15a. with each basin assigned a different shading. The state space matrix for attractor cycles only is shown in Fig 15b, with the same shading scheme.

0	16	32	48	64	80	96	112	128	144	160	176	192	208	224	240
1	17	33	49	65	81	97	113	129	145	161	177	193	209	225	241
2	18	34	50	66	82	98	114	130	146	162	178	194	210	226	242
3	19	35	51	67	83	99	115	131	147	163	179	195	211	227	243
4	20	36	52	68	84	100	116	132	148	164	180	196	212	228	244
5	21	37	53	69	85	101	117	133	149	165	181	197	213	229	245
6	22	38	54	70	86	102	118	134	150	166	182	198	214	230	246
7	23	39	55	71	87	103	119	135	151	167	183	199	215	231	247
8	24	40	56	72	88	104	120	136	152	168	184	200	216	232	248
9	25	41	57	73	89	105	121	137	153	169	185	201	217	233	249
10	26	42	58	74	90	106	122	138	154	170	186	202	218	234	250
11	27	43	59	75	91	107	123	139	155	171	187	203	219	235	251
12	28	44	60	76	92	108	124	140	156	172	188	204	220	236	252
13	29	45	61	77	93	109	125	141	157	173	189	205	221	237	253
14	30	46	62	78	94	110	126	142	158	174	190	206	222	238	254
15	31	47	63	79	95	111	127	143	159	175	191	207	223	239	255

Figure 15a. Eight cell rule 30 state space matrix, all states.

0							112		144					224	
	17	33							145			193			
	18	34	50	66											
		35						131							243
		36		68		100		132							
					85										
				70											246
7							119				183			231	
			56	72				136				200			
9	25														249
										170					
							123				187		219		
	28							140							252
											189		221	237	
14							126						222	238	
			63			111			159			207			

Figure 15b. Eight cell rule 30 state space matrix, attractor cycle states only.

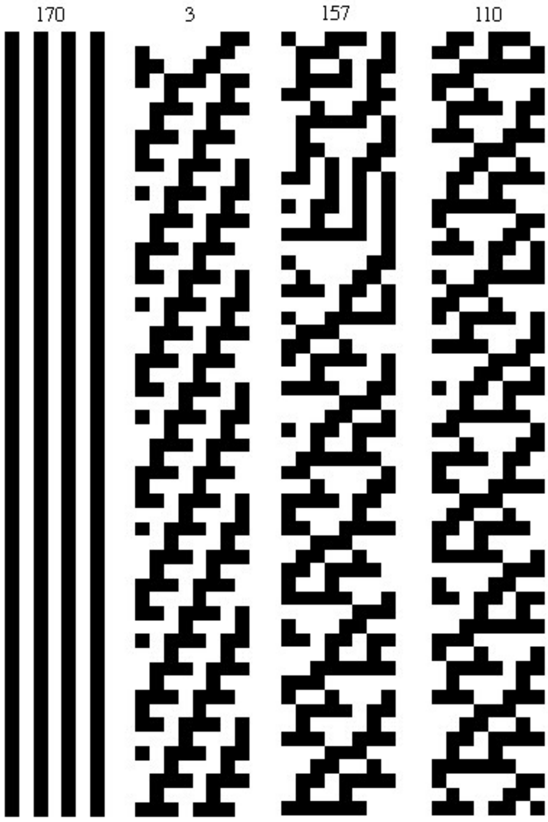


Figure 16. Spacetime plots of 8 cell Rule 30 with decimal seed states indicated above.

Spacetime plots of three attractor cycles from four example seed states are shown in Fig 16. Space is the horizontal axis and time progresses vertically downwards for 56 generations. Running the system for 56 generations allows for the maximum transient and cycle period to appear. From left to right these plots are; Period 1 (type 4) with seed = 170, period 8 (type 3) with goE seed = 3, longest transient to period 40 (type 2) with goE seed = 157 and shortest transient to period 40 (type 2) with goE seed = 110. The period 40 patterns show how the cycle will be entered at a different phase, and is also seen topologically on the boa diagrams. Generally speaking the period 8 and period 40 patterns appear to move right to left over time. The period 8 attractor looks reasonably ordered, whereas the period 40 visually has a slightly more random appearance in comparison.

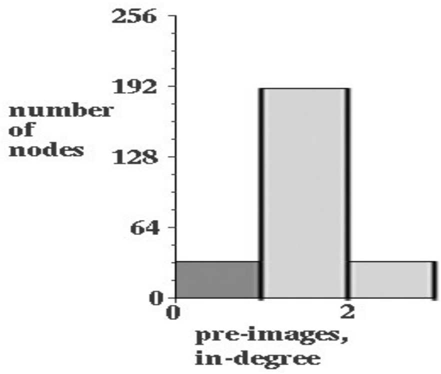


Figure 17. In-degree histogram for 8 cells rule 30, indicating a chaotic dynamic.

A pre-image in-degree histogram is shown in Fig. 17. The goE nodes with zero in-degree pre-images are in the left column, the remainder of the field has 1 or 2 in-degree pre-images. There is a low goE density, and high frequency of low in-degrees indicating by Wuensche's (1997) classification that this is a chaotic rule[12].

5 Evaluation and Reflection

Experimental results are now evaluated in terms of the four criteria areas previously introduced in section 3; computer efficiency, phenomenology, usability of scheme and environment for using the scheme. Reflections will also be given on the system in use as part of the generative composition process.

In terms of memory the 8 cell CA is very efficient. The CASS system is able to run and respond to user events in realtime over a broad tempo range. Vorn's CA external will not run in Max overdrive mode and this results in interruptions in MIDI output scheduling during some GUI events. This does not adversely affect studio use, but is sometimes noticeable in a performance context. Overall the control stream is light and messages are not sent when a cell is not active. This MIDI control stream is minimised by the simple and limited nature of each CME module. MIDI Channel and Program Change are only sent when changed by the user. MIDI Continuous Controllers such as Pitchbend, Mod Wheel and Pan have subsequently been informally tested without much additional efficiency overhead.

In terms of phenomenological criteria, successful diversity of application context was achieved, as was robustness. The main quality of the system is a temporal mapping of the CA to specified sounds. In the built application the seed state and tempo are the performance parameters affecting the CA. From random seed conditions the system is most likely to enter the period 40 basin. In general though the system will tend to be in the period 40 or period 8 basins from a random start, and the period one basins are not often reached randomly.

The aesthetic results produced during the composition process up to date have been both rhythmic and textural. Two compositions emerged from these experiments, exemplifying these approaches. "Ubendem, Wemendem" which uses rhythmic elements in the form of percussive sounds triggered by CASS. The majority of rhythm in this piece was created by selecting sounds while the system ran and storing them to build up a few simple patterns. "Cassorgize" is a slower tempo textural piece, using filtered organ sounds with a slow attack and long release. This approach involved selecting sounds first by assigning event parameters and then running the CASS system from a variety of different seed states.

Modular principles are a very important usability criterion in obtaining generality, power and simplicity. CASS has been successful by taking a basic modular approach. The modularity of this system could easily be extended by the addition of a simple switching matrix between the CA cells and the CME modules. This would allow for a single cell to trigger multiple CME's. The system should be able to scale easily both in CA size and the number of CME's, but this will have a bearing on efficiency criteria. It should be noted that when investigating the global dynamics with DDLab for larger numbers of cells there is a practical limit. The suggested limit in the manual is 16 cells for boa fields and 18 cells for a single boa [13].

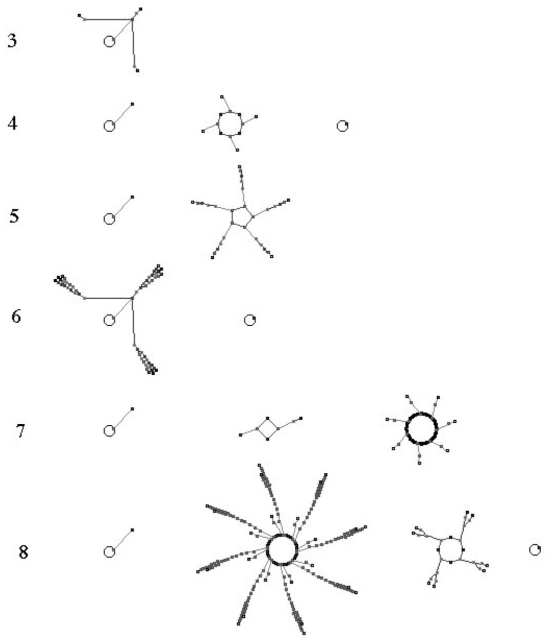


Figure 18. Rule 30 boa fields for 3 to 8 cells, equivalent basins are suppressed.

Rule 30 boa fields for sizes of 3 to 8 cells are shown in Fig. 18. A variety of basin types can be seen and DDLab allows for equivalent rotationally symmetric basins to be suppressed from the display. It will be interesting to examine both increasing and decreasing the number of cells compared to the current system. Examples of the boa fields for 9, 10 and 11 cell rule 30 are shown in Fig. 19.

Some I/O problems have been identified in the context of modular based generative music systems. The seed messages are passed as a list of active cell numbers and having an integer or binary seed is preferable. The cell output of the system would also benefit from being available in this format. Vorn's CA is currently adjustable from 4 to 256 cells, which is perfectly adequate. However, changing the number of cells involves changing the argument of the external within the Max patch. Similarly the rule can only be changed when the patch is being edited, using the Get Info command. The system would benefit by having these parameters available as inputs.

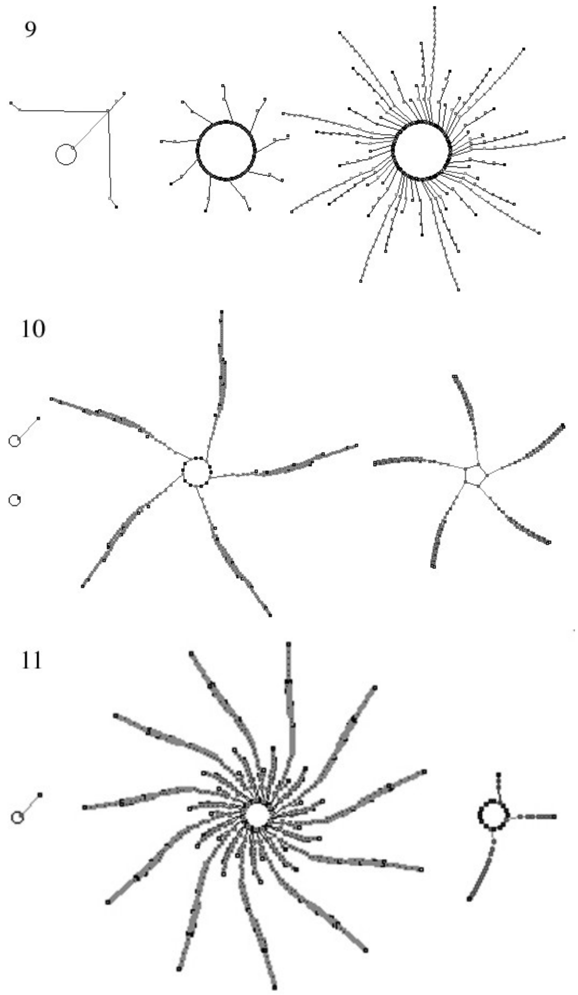


Figure 19. Rule 30 boa fields for 9, 10 and 11 cells, equivalent basins are suppressed.

The overall environment for using the scheme is simple and easy to use. Generation algorithms exist in the context of both visualising CA dynamics, and within the Max programming environment. Both have been effective for generative music research and production with the CASS 8 cell experimental system. Global dynamics is easily studied on a variety of computer platforms with DDLab for systems with small numbers of cells. DDLab has the added benefit of producing basin measures and diagrams in one software environment. Preparation of Fig. 8 and 16 spacetime diagrams to confidence check Vorn's CA external was aided by the ability to label basin nodes with state values, both in decimal and binary.

State space matrix plots are useful and offer a clear numerical view of the basin field. Spacetime diagrams can be easily produced in both DDLab and Mathematica 5. The Max programming environment allows for the investigation of experimental systems within a live and studio context. Vorn's CA is a Mac OS9 based external, currently untested and unlikely to work for OSX. This does not present a problem at present but future work will require an OS9 machine to be available or a new OSX based module. A 1D CA is available in Jitter, but this does not appear to implement periodic boundary conditions and an upgrade from Max is also required.

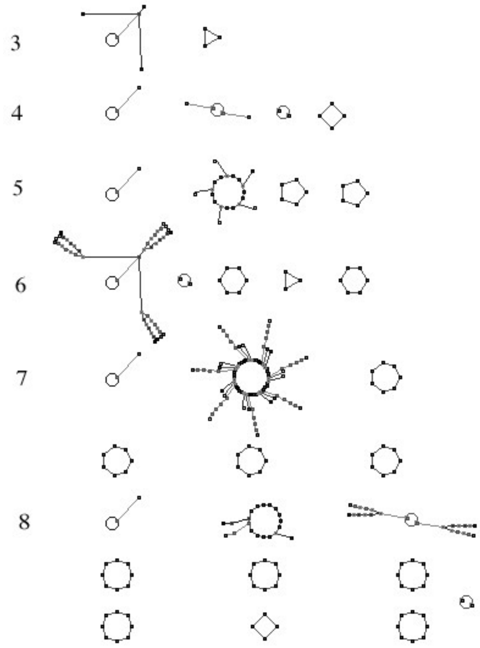


Figure 20. Rule 225 boa fields for 3 to 8 cells, equivalent basins are suppressed.

6 Conclusion and Future Work

In using CA for generative music we show that successful and sensitive application requires understanding of the science by the artist. Global dynamics and rule clustering are important concepts for the generative artist intent on using CA. The vast behaviour space of 1D CA will remain an untamed wilderness for music application unless these concepts are acknowledged, and actively researched in a musical context. The methodology presented in this paper will serve as the foundation for future music research in modular CA systems. Art and science are different yet fundamentally connected disciplines. Here the science has inspired and assisted the artist to achieve specific results. It is hoped that the artwork produced will inspire the science community and generate better understanding in that community of the musical applications of CA.

Vorn's CA external was created for his robotic art systems and not as a modular music component. The criticisms of its functionality are to be taken based on criteria for music composition, and not in terms of his individual application. The current work is made possible because of his placement of them within the public domain.

Rule cluster 15 is a fundamental connection to the traditional analogue sequencer and provides a familiar departure point for electronic musicians. Further work is planned at CCS on generative music research with other 1D CA and will be presented as it is completed. This will continue in a modular context with other parametric mappings and a selection of rules, including rule 30 and its complement rule 225 shown in Fig. 20. The CASS system has been used at Noyzelab and CCS in both performance, and studio production of generative compositions. It is a simple to use musical interface to CA dynamics, both within the Max environment and as a standalone application.

7 Acknowledgements

The authors would like to thank Andrew Wuensche for allowing us to present basin of attraction data and images made at CCS with DDLab. Figs. 1, 3, 9, 14, 15, 17, 18, 19 and Table 4 data were made with DDLab. Also thanks to Andrew Wuensche for allowing us to draw and present the rule clustering images. Figs. 4, 5, and 7, and rule 193 version of Fig. 6 originally appear in [11] and have been redrawn for this paper by Dave Burraston. Andrew Wuensche has provided clarification and valuable comments on this paper during its preparation. Figs. 6, 8 and 16 were made at CCS with Mathematica. Fig. 2 was made at CCS with NKS Explorer.

References

[1] Burks, A. Ed. 1970. *Essays on Cellular Automata*. Univ. of Illinois Press.

[2] Burraston, D., E. Edmonds, D. Livingstone, and E. Miranda. 2004. Cellular Automata in MIDI based Computer Music. *Proceedings of the 2004 International Computer Music Conference*. (accepted)

[3] Candy, L. and E. Edmonds. 2002. *Explorations in Art and Technology*. Springer.

[4] Castagne, N. and C. Cadoz. 2003. “10 Criteria for Evaluating Physical Modelling Schemes for Music Creation.” *Proc. of the 6th Int. Conference on Digital Audio Effects*.

[5] Hunt, A., R. Kirk, and R. Orton. 1991. “Musical Applications of a Cellular Automata Workstation.” *Proceedings of the 1991 International Computer Music Conference*. 165-168.

[6] Jaffe, D. 1995. “Ten Criteria for Evaluating Synthesis Techniques.” *Computer Music Journal* 19(1):76-87.

[7] Langton, C. 1986. “Studying artificial life with cellular automata.” *Physica D* 22:120-149.

[8] Schon, D. 2003. *The Reflective Practitioner*. Ashgate Publishing.

[9] Vorn, B. 1996. *LifeTools*. www.billvorn.com or from the IRCAM ftp server ftp://www.forumnet.ircam.fr/pub/max

[10] Wolfram, S. 1984. “Universality and Complexity in Cellular Automata.” *Physica D* 10D:1-35.

[11] Wuensche, A. and M. Lesser. 1992. *The Global Dynamics of Cellular Automata : An Atlas of Basin of Attraction Fields of One-Dimensional Cellular Automata*. Addison-Wesley.

[12] Wuensche, A. 1997. *Attractor Basins of Discrete Networks*. Cognitive Science Research Paper 461, Univ. of Sussex, D.Phil thesis.

[13] Wuensche, A. 2001. *The DDLab Manual*. Discrete Dynamics Inc. www.ddlab.com

Interactive control of higher order musical structures

Lindsay Vickery

Music Department
LASALLE-SIA college of the arts,
Email: lv@lindsayvickery.com

Abstract

This paper examines the challenge of using interactive procedures to control higher order musical structures, such as formal elements and texture. It proposes as one solution to this problem, a performance paradigm involving a cybernetic arrangement, placing a human performer and a machine in collaboration in which neither is exclusively privileged as the control source. Such a system would exploit the potentialities of the human improviser as an interactive subject in a manner practiced by some non-electronic interactive musical works and present the possibility of an open symbiotic interactive performance model in which control commands pass both *from* and *to* the human and machine components. The development of such works is explored in the context of Game Theory and non-linear compositions involving live players by Xenakis, Pousseur and Zorn as well as interactive works by Teitelbaum, The Hub, and the author's own compositions.

1 Introduction

The control and mapping of sound is an issue for a range of different interactive performance paradigms. These paradigms extend in a continuum of practices from the emulation of existing musical instruments (exemplified by the work of Kapur et al. [1], Cook et al. [2] etc) to the use of non-traditional sources of stimuli such as the movements of dancers or installation participants [3].

The control and mapping of lower order musical parameters in interactive works has received a substantial amount of attention in recent times[1]. The principal reason for this concentration may well be a very practical performative one. Designers of the systems that emulate existing musical instruments have sought to reproduce the 'one to one' relationship found in most acoustic instruments: an action results in a 'point' event. Although there has been consistent development towards a more subtle level of control '*of a note, not as a single static event, but as complex evolving sound with its own internal shape*' [10] through continuous control parameters, the underlying model is still based on the individual instrument and individual instruments are not primarily seen as being in control of higher order parameters.

In respect to interactive systems involving non-instrumental paradigms, Rovan has commented that '*an emergent integrity arises when the relationship between the dance and music systems is "believable"*' [11]. This comment is probably also true of the previous paradigm: in both cases audiences have an expectation that they will be able to "understand" the methods of interactive control. As Dobrian states:

> *Some have argued that it is less interesting to watch a performance on an interactive instrument, because the gesture-sound relationship can be so complex as to be incomprehensible, and in such a case it becomes an improvisation that is interesting only to the performer*. [10]

The most coherent mode of mapping, so called 'one-to-one' or 'direct' mapping, puts the performer in charge of what Stockhausen might term the musical 'foreground'. Such a concentration on musical foreground is reminiscent of the predicament of multiple serialism during the 1950s. The '*discrepancy between intention and result*' [12] in music of that period was noted by analyst-composers such as Ligeti [13] and Xenakis [14]. Both composers felt that the high level of foreground control in multiple serialism did not translate to higher order textural and formal parameters. Although there is arguably more to divert and distract the audience in interactive music, it faces a similar quandary: how to balance the need for recognizable 'gestural-auditory' coherence and at the same time provide 'auditory-structural' integrity and substance.

However, since higher order parameters are intrinsically more complex - being composed of groups of other parameters - the manner of their control is almost inevitably less evident than that of one-to-one interaction. It could be argued that in works in which higher order parameters are interactively controlled, questions of 'gestural coherence' are less important than the integral structural coherence of the whole performance. As in notated compositions, it is not as important for the audience to see 'under the hood' as it is to comprehend the overarching structural or textural shape of the work. As Rovan states:

> *for some it may be enough that the system of interaction "privately" affects the performer's expression within the piece. The audience is only vicariously part of the interactive experience*. [11]

[1] For example Lee, and Wessel, 1992 [4], Modler, and Zannos, 1997 [5], Bowler et al. 1990 [6], Rovan et al. 1997 [3], Mulder et al. 1997 [7], Wanderley et al. 1998 [8], Hunt and Kirk 1999 [9].

Concentrating exclusively on foreground parameters can, as is arguably the case in multiple serialism, lead to music that is too superficial to be successfully expressive in comparison with the composed score or musical improvisation. After all, listeners to piano music comprehend and enjoy it for many reasons other than a precise understanding of the workings of its mechanism.

This paper proposes an interactive performance paradigm located between the two discussed above, with control of higher order parameters as its principal concern. Such a 'cybernetic' arrangement would place a human performer and a machine in symbiotic relationship in the production of a musical discourse that encompasses orders of musical structure from low to high.

This arrangement would result in a system based upon interaction with the 'performative-instrumental' and musical-structural elements of the performance. It tackles the issue of interactivity from the composer's point of view and asks the questions: how are the elements of these higher order structural parameters understood; what methods exist to listen for them; and how might we harness these understandings in a real-time performance situation?

2 Understanding

Dannenburg has suggested that the absence of any general theory of semantics is one of the chief barriers to the codification of recognition and understanding in the field of music. He states that as a result it cannot be '*evaluated objectively in terms of how well it preserves semantic information across a change of representation*' [15].

His summary of the processes involved in musical understanding suggests that music is a predominantly self-referential language, often relying on '*repetition at different time scales*', of '*elements of the music* (that) *are repeated or transformed*'. He goes on to describe music as a form of complex domain knowledge in which '*listeners construct encodings when they listen to music, and that the encoding chosen by a listener will tend toward the shortest encoding possible*'.

The content of this encoding shorthand, according to Dannenburg is not in the form of musical elements such as melody, harmony, and rhythm, but rather '*the transfer of these elements within a composition, forming relationships and therefore structure*'. His discussion takes place in the context of cataloguing and retrieval of digital audio (in archiving for example) and centers on developing '*automated systems (that) can listen to music in audio form and determine structure by finding repeated patterns*.'

Our understanding of music is perhaps best understood from the same perspective. The field of cognitive science may provide an approach for connecting such automated systems with interactive mapping techniques by adopting a model resembling memory itself.

> *Human memory can be understood as a network of propositions and cognitive structures and can therefore be described in terms of elements or nodes and connecting relations*. [16]

The key issue in much interactive music is the ability of the performer(s) to randomly access material effectively. It is also a key issue in memory:

> *as short-term memory is limited, off-loading has to take place when we want to work on complex tasks. They see mapping techniques as a performance aid, serving as an external memory extension and as such realizing off-loading*. [16]

Mapping techniques using '*spatial, network-like visualisations for knowledge construction, organisation and presentation*'[16] might present an effective method of interfacing live performance (using analysis of digital audio such as described by Dannenberg) with the kind of complex relationships between elements found in pre-composed music.

Lazier and Cook (2003) describe a process they term '*audio mosaicing*' to both '*classify sound and to retrieve audio by content with the use of a variety of features as discriminators*'. The process '*can be used to extract global information from recordings or to compare smaller segments of sound on the local level (…) to retrieve and concatenate recorded sound*' [17]. They state that '*the ultimate goal (…) is to produce a cumulative perceptual effect desired by the artist from the combination of an ordered set of chosen segments*.' Although so far they have only used the process to '*create high quality syntheses of symbolic scores or re-syntheses of existing recordings*' [17], it represents a very potent prospect as the final phase in a cognitive based system of analysis, response and resynthesis between a live performer and computer.

Another approach might favour more rudimentary means of analysis such as the Max\MSP objects `fiddle~` [18] and `analyzer~` [19] to provide estimates of perceptual features such as pitch, loudness, brightness, noisiness, onsets, and Bark scale decomposition. All of these musical elements are useful measures of performance characteristics. Weyde has suggested that a complex system might evolve through the categorisation and mapping[2] of the performance at this stage, particularly for elements that are 'non-hierarchical', that is unpredictable.

> … *generally any musical information is well suited for mapping that can be expressed in terms of objects (e.g. notes, chords, motifs, sections) and their relations (e.g. similarity, succession, contrast, harmonic function)*. [16]

[2] This is a purely 'internal' mapping as opposed to the 'one-to-one' interactive mapping discussed earlier.

In such a system the data, which might include both statistical information about the material as well as actual samples of the material itself, would need to be managed. Dobrian outlines such a process in, stating that by:

> *Employing scheduling and storage techniques (extreme delay, capture and storage of data, reordering of events, etc.) one can shape a larger formal structure in real time.* [10]

The crucial issue here is how to design the control element, or 'rules' of such an interactive system to create an integrated and coherent structure combining the live performer and the computer. One possible direction to base such a system is the use of rule-based compositional processes such as Game Theory.

3 Models

Sward places the origin of game-based analysis in the 1920s when it '*began to be used in mathematics for predicting outcomes in economics and later human conflicts*' [20]. Its first employment for musical purposes was probably by Xenakis in works such as *Duel* (1959) and *Strategie* (1962) [12]. Xenakis '*developed a number of automated or semi-automatic compositional systems, thus foregrounding the systems themselves, derived from work in Cybernetics, including 'game' based compositional techniques*' [21]. Despite being envisioned as 'games' by the composer and including '*victory and defeat, which may be expressed by a moral or material prize, …and a penalty for the other*' [22], the 'games' in these works occur at a compositional and not at the performance level, with scores (a highly appropriate term) that are computer generated. Henri Pousseur's work *Repons* (1960) exhibits the first performative game qualities providing performers with '*a set of rules of play and … musical material which permits them to respond, with certain margins of improvisation, to all the situations into which the game puts them*' [23].

But perhaps the best known author of 'game' compositions is John Zorn who created some 27 such works between 1974 and 1992. These sets of rules, musical fragments and sometimes images, took a stance based firmly in Free Improvisation rather than notated composition and therefore radically different to Xenakis or Pousseur. The best known and most frequently performed of these works is *Cobra* (1984). Many of its rules, delivered by a 'prompt' to the performers on colour-coded filing cards, bear a resemblance to algorithmic commands. A small sample of the commands are: *Mouth 1* which directs performers to make a 'radical change' in what they are playing; *Mouth 3,* 'exchange' which directs those playing to stop and non-players to play; *Ear 3* signals a volume change; and *Head 1, 2* and *3* are 'Memory Cards' signalling that the current texture is to be remembered (it may be recalled at a later point).

For the most part the work is also devoid of explicit content[3], but is driven by a curiously 'democratic' or at least competitive form of real-time composition. To summarize, players in the group signal suggestions[4] for the next structural element (a card command) to a 'prompter' who chooses one of the suggestions and the holds up the appropriate card to enact the command. Given these conditions it is plain to see why the choice of players is so crucial. Zorn supports this perception:

> '*Each performance will be drastically different in sound and structure as the participants bring in their own private perceptions, past experiences, instrumental techniques, and interpersonal attitudes*' [24]

Importantly, the player choices are also clearly of a compositional nature, a situation that is demonstrated by the composer's actions in the case that he has not achieved a desirable line-up:

> *This even means that once he has chosen the players and the right chemistry turns out to be missing, he will not go ahead.* [24]

Franco Evangelisti founder of the improvisation ensemble *Gruppo di Improvvisazione Nuova Consonanza* signalled in 1959 that a new breed of performer would have to emerge:

> *one that was also a composer. This new figure would thus be able to link together certain musical elements which, in performances given by performers of a traditional type, are subject to the previous experience of the performer in question.* [25]

Cobussen suggests that:

> Cobra *is thus simultaneously reproducing the composer-conductor-performer hierarchy of traditional "classical" music and subverting that hierarchy from within the "composition" itself.* [24]

However, it could also be said that Zorn is simply abstracting the organizational model 'composer-conductor-performer(s)' to a point where it accepts a broader interpretation that allows for interaction between the three roles. It also broadens the notion that these functions are each prescribed to an individual, as there is a high degree of movement between three functions by *Cobra* participants.[5]

---

3 The exception is the *Cartoon Trades* cards, where the performer is expected to make short cartoon soundtrack-like gesture and pass it on to someone else, but even in this case there is no attempt to specify any exact sounds or motifs.

4 Hence the command titles such as *Ear 1* (signalled by the player touching the ear with the first finger).

5 Players, presumably when exasperated with the choices of the current 'prompt', are permitted (under certain conditions) to co-opt other performers and breakaway from

In many ways the structural organization of *Cobra* - from this high level, through the filing-card commands to the random access 'content' of individuals' improvising chops - resembles the kind of arrangement outlined in the first section of this paper for the interactive control of higher order parameters.

One of the key figures straddling the fields of improvisation and electronics is Richard Teitelbaum, whose work has centred upon the creation of 'automata': improvising electro-mechanical systems programmed by the composer to respond to his performance with a spontaneous interaction. The 'rules' here are of course based upon the composer's own experience as an improviser[6] and the impulse to create a mechanical reflection of his own playing has been likened using the popular culture icon of Dr. Frankenstein as well as the psycho-analytical example of Lacanian 'mirror-phase'.

> *In this can be seen the artists obsession with the process of creation and reproduction and humanities desire to reproduce itself, not only sexually but, if possible, through its artefacts and disciplines such as alchemy and A.I.*(Artificial Intelligence)*... The psycho-analytic aspects of this process - the artefact as a mirror of ourselves*. [21]

This comparison is of course highlighted by the performance arrangement of 'soloist and interactive electronics'. This arrangement also highlights another difficulty of interactive music: that sharing the interaction (as Zorn does between live musicians in *Cobra*) greatly reduces the degree of control for any individual performer. This is because it complicates the chain of causality in the interaction. In effect the performer is placed in the position ascribed to the audience by Dobrian in that the '*gesture-sound relationship can be so complex as to be incomprehensible*' [10].

Californian group *The Hub* sought to solve or at least embrace this problem:

> *The Hub is a computer network band. Six individual composer/performers connect separate computer-controlled music synthesizers into a network. Individual composers design pieces for the network, in most cases just specifying the nature of the data which is to be exchanged between players in the piece, but leaving implementation details to the individual players, and leaving the actual sequence of music to the emergent behavior of the network. Each player writes a computer program which make musical decisions in keeping with the character of the piece, in response to messages from the other computers in the network and control actions of the player himself*. [26]

At present interaction of this kind probably comes closest to emulating the kind of fluid command structure exhibited by Zorn's *Cobra*. *The Hub*'s composer/musicians are interacting directly with computers which not only restricts the audience's understanding of the interaction but tend to make the group resemble IBM office drones. However performances of this nature increasingly involve projection of the performer's desktops allowing a few more of the audience in on the game. But perhaps the most interesting direction is the opening up of performances for audience interaction via their own computers, thus eradicating a whole component of the 'composer-conductor-performer(s)-audience' organization.

Another potent emerging technology is that of biological art. Either through computer modeling of biological processes (generally called 'artificial life'' [27]) or interaction with actual biological material, it aims to 'breed' agents for specific purposes such as interacting with musicians or drawing portraits [28]. The 'rules' here, (strikingly similar in some respects to those of *Cobra*) can concern evolutionary imperatives such as mutation, mating, insemination and morphing [29]. In the case of Tissue Culture & Art's *Meart: The Semi Living Artist* (2001-) project, the intention is expressly stated to discover what happens 'when such a system starts to express qualities that are considered uniquely human aptitudes such as art?' [28].

Clearly akin to the 'Dr. Frankenstein' approach to creating an improvising partner, biological art presents an intriguing range of possibilities. Dahlstedt states that even in its current stage of development:

> *Interactive evolution as a compositional tool makes it possible to create surprisingly complex sounds and structures in a very quick and simple way, while keeping a feeling of control*. [29]

4 Towards a symbiotic human/machine interactive model: some examples

The following examples taken from the author's recent works to trace a trajectory towards models in which live performers might interact with technology to create what might be termed a symbiotic performance. The works show the development and exploration a range of issues related to interaction, particularly the balance of performer and machine freedom.

4.1 Performer Freedom

Like Zorn's game pieces these works spring from a desire to stimulate the spontaneous inventive qualities of Free Improvisation. The works employ a range of approaches including auditory stimulation through prerecorded or live sampling and visual

---

the group as a 'guerrilla' group in which they themselves become a proxy 'prompt'/conductor.

[6] He is a frequent duo partner of improviser and theorist Anthony Braxton.

stimulation through what I have termed 'Score-Films'.

Delicious Ironies (2001). *Delicious Ironies* was intended as a vehicle to provide an extremely unpredictable environment of sounds for the solo improviser. The intention was to use sound samples that were pertinent to the soloist, but also volatile and erratic enough to inspire an interesting response. For example different pieces in the series drew on morphologically related samples from film noir, boxing movies, record glitches or various extended techniques by the performers themselves.

The stream of samples that accompany the improviser in *Delicious Ironies* is controlled by nine layers of event generating objects. Each object emits the same formal structure iterated by 9 event generating objects each at different tempi. The fastest object sends cues at 9 times the speed of the slowest. Events sent from the objects are mapped to different aspects of sample playback: ie sample choice, playback speed, duration, volume, loop, pan and portamento amount.

Obviously, despite the fact that the events are 'played' by the computer in exactly the same way each time, the altered sample set generates an utterly different sounding piece each time. In addition to the aural 'surprises' in store as the computer performs the piece, the soloist may also be given a text cue to provide additional stimulation. Example 1 below shows a variety of sound suggestions (in the score they are spread out randomly over an A4 page). The performer is instructed that '*all sounds should be in transition towards something else; timbre and pitch should be in constant flux*.' [30]

Although *Delicious Ironies* was intended as an interactive work, the interaction is all in one direction. The computer's utterances act to prompt the live soloist and in the best circumstances in performance the two form an amalgam. However, despite it sometimes sounding to the contrary, the performer cannot influence the computer performance in any way.

bushmen discuss a travesty of justice in a jovial mood, washing machine malfunction, person shouting with no voice (extremely hoarse) in Italian, backwards talking little Japanese girl, distant aircraft, very very drunk karaoke singer sings birdie dance in fake German, ants communicate with antennae, weeping woman, lost homesick aliens make indiscriminate enquiries of passers-by, happy birds in birdbath, whispering sleeper, Aussie digger tells a joke, Godzilla attacking Hong Kong, Oxford don lectures on German existential philosophy, Texan gunshot victim bleeds to death, second grader recites times tables, soft shoe shuffle, late at night formula one driver relives his former glories in his head, Russian sailor tries to get his visa renewed, drycleaner explains unmovable spot, submarine pings sonar, quiz show contestant deliberates over curly question.

Figure 1. *Delicious Ironies Number 4*: vocal text cues

Splice (2002). *Splice* developed directly out of the control procedures employed in *Delicious Ironies*. It imposes the same formal structure but in contrast its contents comprise live sampling of the soloist's improvisation. It is an example of an encoded 'meta-music' - that is a compositional map that is without contents until a live performer adds them.

The soloist's samples, collected in real-time, are replayed according to the same scheduling as *Delicious Ironies* (ie playback speed, duration, volume, loop, pan etc.). In this case however the result is quite different: now the 'contents' of the computer's performance is in most respects determined by the soloist. Psychologically the process of interaction that occurs in *Delicious Ironies* is now partially reversed: the soloist is 'loading' the samples, however the computer still controls the timing of the actual sampling and playback. The computer's timing (although consistent) is opaque to the performer, creating a degree of uncertainty both about *what* has been sampled and *when* it will return. The transformation of the samples (through varied playback speed, volume and panning for example) adds a further layer of uncertainty.

The kinds of repetition and transformation exhibited by *Splice* are standard formal strategies for music. However, unlike a composer working with notated music, the computer is not discriminating in its choice of material: it sculpts any of the soloist's contributions into the same structure regardless of whether they are melodic, noise or even silence, so though it may sound different each time it is always the same shape.

In *Splice* the process of interaction - the detail of how the scheduling is determined - is not critical to the audience. It avoids the paradox outlined by Rovan - the simultaneous desire for interconnection and freedom of both agents in an interactive relationship - because the focus of the work is not on how the soloist and the computer components are linked but the aural outcome of the two components.

The Score-Film. The exploration of projected images as a medium was also motivated by the desire to stimulate improvising performers. Another important factor was perhaps the availability of extraordinary footage of biological and endoscopical material[7] created by the *Tissue Culture & Art Group* (Oron Catts, Ionat Zurr, and Guy Ben-Ary).

Figure 2. Still frame of saxophone interior from: *Keyhole* section of *Fantastic Voyage*, © Tissue Culture & Art and HEDKIKR

In our first collaboration, *Fantastic Voyage*[8] *(2002) TC&A*'s images were edited together by the author with the express purpose of being a score to be performed live by improvisation duo HEDKIKR (Darren Moore and Lindsay Vickery) in response to the film. In performance the image and sound form a symbiotic relationship in which the complex and sometimes extreme nature of both mediums – Free Improvisation and Biological Art - is rendered more comprehensible.

Subsequent explorations of interaction contexts with film have included: *Meart/Mesound* (2002), *Microphagia* (2002) *Cytoblasty* (2002), *Pig Wings* (2003) and *Sugar* (2003). *Meart/Mesound* drawing on the installation *Meart*, is probably the most radical and most potent of these collaborations.

'MEART is an installation distributed between two (or more) locations in the world. Its 'brain' consists of cultured nerve cells that grow and live in a neuro-engineering lab, in Atlanta. Its 'body' is a robotic drawing arm that is capable of producing two-dimensional drawings. The 'brain' and the 'body' will communicate in real time with each other for the duration of the exhibition.' [28]

For some the drawings it creates have a haunting primal quality, and for others they are like a foretaste of communications with alien life. In this initial musical interaction there was no feedback to the neuron culture and our interaction took place principally with the pneumatic blasts emanating from the installation.

More recent versions of the work involve a form of interaction in which the evolving drawing is compared to a portrait photograph and the neurons are stimulated to colour the dark regions more than the light ones, in a imitation of the process of human drawing. A similar form of interaction could occur musically, but has yet to be developed.

The works *Microphagia* (2002) and *Cytoblasty* (2002) follow a different path, by making the film itself interactive. In both of these works the interaction with the film is controlled by a dancer performing in a MIBURI MIDI jumpsuit [31]. Here the musician's interaction is again with the TC&A film images, however in this case they are unpredictable and non-learnable. One aspect of the success of *Fantastic Voyage* has been the high level of synchronization achievable after repeated performances while maintaining a high degree of freedom. A future direction for this work might be the memorization of the Score-Film (to attain a high level of synchronization) and then interactive non-linear projection, to provoke novel interpretations.

4.2 Computer Control

The examples cited so far explore the use of interactive paradigms in the field of free improvisation. The examples that follow develop some of the same ideas in the context of notated music and propose some potential directions for work that fuses aspects of all of these practices.

interXection (2002). *interXection* for Drum Kit and Ring Modulator was made for Tissue Culture & Art's film Pig Wings (2003), in which they 'differentiated bone marrow stem cells to grow pig bone tissue in the shape of the three solutions for flight in vertebrates.[9]' [32].

The music set out to highlight the analogy of the microphone as an audio microscope by magnifying barely audible sounds from the drum kit and processing through a Ring Modulator. The intention was that the Ring Modulator would bring the source sound's component harmonics into relief in a way that is analogous to a colour imaging microscope's

[7] The Score-film *Fantastic Voyage* features images made by *Tissue Culture & Art* of: a cell-sculpture, endoscope images of the drum-kit, the interior of the saxophone and saxophonist; a variety of assemblages of mouse cardiac cells, the ebbs and flows of a culture of fish neurons; a colour imaging a cell's structure and a close-up of veins at the back of a rat's eye.

[8] The work draws its title from the 1968 film about a secret US military organization that miniaturizes a medical team and injects it into the body of a comatose scientist from a shadowy evil foreign power.

[9] *Tissue Culture & Art* speculate in the work: '*the rhetoric surrounding the human genome project and xenotranplant-ation made us wonder if pigs would fly one day and if they will what shape their wings would take.*' [32]

rendering of biological samples. However its relation to higher order musical structures lies in is its translation of the formal devices from *Delicious*

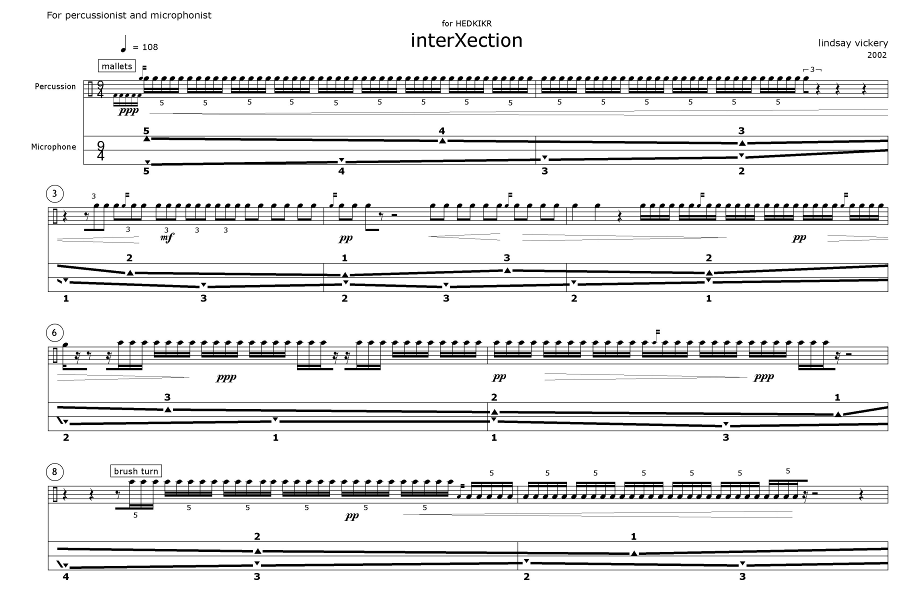


Figure 3. *interXection* score (excerpt): showing microphone position indications

Ironies and *Splice* into a notated form.

In *interXection* the same structure employed by *Delicious Ironies* and *Splice* is used to determine a number of parameters including the percussionist's tempo, roll speed, instrument, mallet type, accent, dynamics, rest position and length and the vertical and horizontal coordinates of the microphonist's mic in relation to the percussionist's current instrument. This microphone part, essentially notating two 'vectors', resembles the graphic interface for automation for Pro-tools© Effects inserts. (The comparison is an appropriate one, as it is possible to imagine recreating the microphone part 'post-production'.)

It is another example of the adaptation of electro-acoustic practices such as Stockhausen's analogies of spectral analysis in *Zyklus* (1959), and *Refrain* (1959), [33]. Lachenmann's 'instrumental-musique concrete' [34] in *Pression* for a cellist (1969) or Zorn's translation of studio techniques in the cut and paste live performances of *Naked City* (1989-1992).

Whorl (2004). Tempo is one of the least explored musical parameters in live performance. In non-solo performance each additional player decreases the ability to change tempo by many times. Accurate continuous changes in tempo (ie accelerando and rallentando) are generally regarded as non-specific commands (ie we are not taught to rall. over a particular, exact duration). These understandings are embedded in our musical perception to a high degree. Even in electronic music, where tempo variations *can* be precise, they often cause a perception in the listener of separate streams of sound rather than elements of a composite texture. This mirrors the way in which timbres are unpicked perceptually by the listener and attributed to different sources.

Whorl uses the same structural framework as *interXection* to independently control three live performers via headphones. Each player receives a separately varying click-track (with five tempi and connecting accelerandi and rallentandi), as well as instructions on what dynamic, musical material and pitch set to play. This arrangement creates an unusual set of conditions for the performers in which their listening skills are divided between synchronization with the computer generated click-track and 'ensemble' playing through listening to the other players. Initially in rehearsal this results in split focus – players tend to concentrate more on one task than the other. (Arguably this split focus also occurs to a degree when normal written music is first rehearsed.) Rehearsal of *Whorl* suggests that with familiarity traditional methods of group playing begin to take effect, for example aural coordination and visual observation of bodily gestures such as are generally used to co-ordinate nuances in chamber music.

Echo Transform (2004). *Echo Transform* combines the improvisation and sampling procedures of *Splice,* and the notated score of *interXection* with a click track to create yet another performance paradigm. Although there is a score, it is in many respects more of a guide for the performers than a traditional score. For example, the first and penultimate sections notate a single line with instructions about when players should fade-in and fade-out. The players are issued (individual) instructions to vary the line through commands to imitate, anticipate, interpolate and interrupt. Like *interXection* the score also asks the performers to 'morph' between states in a manner that is probably more akin to electro-acoustic music. For example the opening 'still' note transforms via widening vibrato to a tremelo over the first section.

As in *Splice* the sampling process is still visually 'opaque' to the performers (there is no visual sign that it is taking place and it is not notated in the score). However, due to the increased consistency of the process, (the computer will always reliably sample the same portions of the performance), the players can begin through repeated rehearsal to recognize the process aurally.

5 Conclusion

A key concern for all of the works discussed in this paper is the consideration of a wide range of musical parameters as important to the development of interactive music as an eloquent and powerful mode of musical expression. This process may involve a degree of discourse between performative aspects, such as the ‘degree of believability’ of the interaction, and broader concerns of the structural coherence of the music.

The examples cited indicate a direction for this research that may well involve initiatives such as: greater and more sophisticated degrees of interaction; the addition of biological agents in the system(other than humans); greater or more thorough integration of visual aspects such as interactive film; and overall, a greater level of interplay between these elements.

References

[1] Kapur A., A. Lazier, P. Davidson, R. S. Wilson, and P. R. Cook. "The Electronic Sitar Controller." In *Proceedings of the 2004 International Conference on New Interfaces for Musical Expression* (NIME), Hamamatsu, Japan, June 2004.

[2] Cook, P. R. 2003. "Perceiving Our Instruments: Psychoacoustics Meets Aesthetics in the Design of Perfomance Interfaces". In *Proceedings of the 40th Anniversary Celebrations for the Institute for Psychoacoustics and Electroacoustic Music* (IPEM40!).

[3] Rovan, J., Wanderley, M., Dubnov, S., and Depalle, P. 1997. "Instrumental Gestural Mapping Strategies as Expressivity Determinants in Computer Music Performance." In *Proceedings of the Kansei Workshop*, Genova, pp.68-73.

[4] Lee, M., and Wessel, D. 1992. "Connectionist Models for Real-Time Control of Synthesis and Compositional Algorithms." In *Proceedings of International Computer Music Conference 1992*, pp. 277-280.

[5] Modler, P., and Zannos, I.. 1997. "Emotional Aspects of Gesture Recognition by a Neural Network, using Dedicated Input Devices." In *Proceedings of the Kansei Workshop*, Genova, pp. 79- 86..Bowler, I., Purvis, A., Manning, P., and Bailey, N.. 1990. "On Mapping N Articulation onto M Synthesiser- Control Parameters." In *Proceedings of the International Computer Music Conference* 1990, pp. 181-184.

[6] Bowler, I., A. Purvis, P. Manning, and N. Bailey. 1990. "On Mapping N Articulation onto M Synthesiser-Control Parameters." In Proceedings of the 1990 International Computer Music Conference. San Francisco, International Computer Music Association, pp. 181-184.

[7] Mulder, A., Fels, S., and Mase, K. 1997. "Empty-Handed Gesture Analysis in Max/FTS". In *Proceedings of the Kansei Workshop*, Genova, Pp. 87-91.

[8] Wanderley, M., Schnell, N., and Rovan, J.B. 1998. "Escher - Modeling and Performing Composed Instruments in Real-Time." In *Proceedings of the International Conference on Systems, Man and Cybernetics 1998*, Pp. 1080-1084.

[9] Hunt, A., and Kirk, R. 1999. "Radical User Interfaces for Real-time Control." In *Proceedings of the Euromicro Conference*. Milan.

[10] Dobrian, C., 2001. "Aesthetic Considerations in the Use of “Virtual” Music Instruments", originally submitted for the CHI Workshop on New Interfaces for Musical Expression, presented in the national conference of the Society for Electro-Acoustic Music in the United States (SEAMUS), Iowa City, Iowa, USA, April 2002, and published in the Journa*l SEAMUS*, Spring 2003.

[11] Rovan, J.B., Wechsler, R., Weiss, F., 2001 "Seine hohle Form… - Artistic Collaboration in an Interactive Dance and Music Performance Environment". In *Proceedings of COSIGN 2001:* Computational Semiotics CWI, Amsterdam

[12] Griffiths, P., 1981. *Modern Music: The Avant Garde since 1945*, London: Dent.

[13] Ligeti, G., 1960. *Metamorphosis of Musical Form,* Die Reihe, No. 7, English Edition 1965, Bryn Mawr: Theodore Presser

[14] Xenakis, I., 1955. "*La crise de la musique sérielle*", *Gravesner Blätter*, Number 1, Pp. 2-4

[15] Dannenberg, R. B., and Hu, N., 2002. "Discovering Musical Structure in Audio Recordings". In Anagnostopoulou, Ferrand, and Smaill, eds., *Music and Artificial Intelligence: Second International Conference, ICMAI 2002*, Edinburgh, Scotland, UK. Berlin: Springer. Pp. 43-57.

[16] Weyde, T., 2004. "Visualization of musical structure with maps". In *Proceedings of the ESCOM Conference on Interdisciplinary Musicology*. Graz

[17] Lazier, A., and P. R. Cook. "MoSievius: Feature-based Interactive Audio Mosaicing," In *Proceedings of the International Conference on Digital Audio Effects (DAFX),* London, England, September 2003.

[18] Puckett, M., T. Apel, and D. Zicarelli, 1998. "Realitme audio analysis tools for Pd and MSP" *In Proceedings of the International Computer Music Conference*. International Computer Music Association, pp. 109-112

[19] Jehan, T. and B. Schoner (2001). "An audio-driven perceptually meaningful timbre synthesizer". In *Proceedings International Computer Music Conference*, La Habana, Cuba, pp. 381–388.

[20] Sward, R. 1981. *A Comparison of the Techniques of Stochastic and Serial Composition Based on a Study of the Theories and Selected Compositions of Iannis Xenakis and Milton Babbitt*. Ph.D Thesis, Theory, Northwestern University.

[21] Biggs, S., 1987. *Cybernetics in a Post-Structuralist Landscape*. Available [on-line] at <http://hosted.simonbiggs.easynet.co.uk/texts/cybernetics.htm>

[22] Xenakis, I., 1971. *Formalized music*, Bloomington: Indiana University Press, pp. 112-13.

[23] Butor, M., and Pousseur, H., 1971. "Repons". In *Musique en jeu*, Pp. 106-11, p. 107

[24] Cobussen, M., undated. *Deconstruction in Music*, Interactive Dissertation, Department of Art and Culture Studies, Erasmus University Rotterdam, The

Netherlands. Available [on-line] at <http://
www.cobussen.com/navbar/intro1.htm>
[25] Tortora, D., undated. *The Nuova Consonanza Improvisation Group*. Available [on-line] at <http://www.nuovaconsonanza.it/storia_pages_en/s_impro.html>
[26] Perkins, T., undated. *The Hub*. Available [on-line] at: <http://www.artifact.com/hub.html>
[27] Langton, C.G. (ed.). 1989. "ALIFE I", In proceedings of the *first international workshop of the synthesis and simulation of living systems*. Addison Wesley.
[28] Ben-Ary, G., Bakkum, D.,Bunt, S. Catts, O., DeMarse, T., Gamblen, P., Madhavan, R., Passaro, P., Potter, S., Shkolnik, A., Sweetman, I., Zurr, I., undated. *MEART - The Semi Living Artist*. Available [on-line] at http://www.fishandchips.uwa.edu.au/
[29] Dahlstedt, P., 2001. "Creating and Exploring Huge Parameter Spaces: Interactive Evolution as a Tool for Sound Generation". In *Proceedings of the International Computer Music Conference*, Habana, Cuba, 2001.
[30] Vickery, L., 2001. *Delicious Ironies: No. 4, performance notes*
[31] Vickery, L. (2002) The Yamaha MIBURI MIDI jump suit as a controller for STEIM's Interactive Video software Image/ine. In the *Proceedings of the Australian Computer Music Conference 2002*
[32] Catts, O., and Zurr, I., 2003. "The Pig Wings Project". In the program for *the Boston Cyberarts Festival* Exhibition at the DeCorva Museum and Sculpture park in Boston.
[33] Maconie, R., 1990. *The Works of Karlheinz Stockhausen* (second edition): Oxford: Clarendon Press
[34] Lachenmann, H., 1994. CD liner notes for *Helmut Lachenmann*, Col Legno Produktion GmbH: Salzburg

Fission or fusion: analysing the acousmatic reaction

David Hirst
Teaching, Learning and Research Support Dept, University of Melbourne
email: d.hirst@unimelb.edu.au

Abstract

This paper presents a procedure for the analysis of acousmatic music which was derived from the synthesis of top-down (knowledge driven) and bottom-up (data-driven) cognitive psychological views. The procedure is also a synthesis of research on primitive auditory scene analysis, combined with the research on acoustic, semantic, and syntactic factors in the perception of everyday environmental sounds. The procedure can be summarized as consisting of a number of steps: Segregation of sonic objects; Horizontal integration and/or segregation; Vertical integration and/or segregation; Assimilation and meaning.

1 Introduction

Not only do we recognize sounds, but we ascribe meanings to sounds and meanings to the relationships between sounds and other cognate phenomena. Music is a meaningful and an emotional experience. Acousmatic music uses recordings of everyday sounds, recordings of instruments and synthesized sounds. Acousmatic music may make use of traditional musical relationships, but there can also be unique abstract relationships between the sonic attributes of sounds and the perceiver of those sonic attributes that we don't find in traditional instrumental music. What is this "syntax" of acousmatic music and how does it interact with the semantic references afforded by some of the sonic material within acousmatic musical works?

We have previously reported on a bottom-up type of approach to the analysis of acousmatic music that adapted Bigand's model of Event Structure Processing of tonal music (Hirst, 2003) [6], and discussed the interpretation of some of Smalley's concepts (Hirst, 2002) [5]. In this paper we will examine some of the top-down processes that may operate in the perception of Western tonal music and also consider top-down processes in the recognition of environmental sounds. Through combining these two bodies of previous work, where there has been some research activity, we may shed some light on the perception of acousmatic music, where there has been only a meagre amount of research. The combination has resulted in the definition of a methodology for the analysis of acousmatic works – a procedure that requires a delicate balance between fission and fusion within the acousmatic reaction.

2 Defining the analytical methodology

The following procedure for the analysis of acousmatic music was derived from the synthesis of top-down (knowledge driven) and bottom-up (data-driven) views since it has been found impossible to divorce one approach from the other. The procedure is also a synthesis of the primitive auditory scene analysis approach of Bregman (1999)[2], with the research on acoustic, semantic, and syntactic factors in the perception of everyday environmental sounds carried out by Gygi (2001) [4], Howard and Ballas (1980) [7], and Ballas (1993) [1], together with theories of implication- realization (Narmour, 1989) [10] and symbolic interpretation (Pierce, 1931-35) [11].

The procedure can be summarized as consisting of a number of steps:

Segregation of sonic objects

1. Identify the sonic objects (or events).
2. Establish the factors responsible for identification (acoustic, semantic, syntactic, and ecological) and their relative weightings (if possible).

Horizontal integration and/or segregation

3. Identify streams (sequences or chains) which consist of sonic objects linked together and function as a unit that we could call a "pattern". The role of "trajectories" should also be considered. (The term "gesture" has sometimes been used in this context.)
4. Determine the causal linkages between the sonic objects within the chain-type pattern.
5. Determine the relationships between "pattern objects" – if this level of syntax exists. This amounts to the investigation of higher-order relationships within a "hierarchy".
6. Consider local organization in time – pulse, beat, accent, rhythm, meter.
7. Consider the horizontal integration of pitch, including emergent properties relating to timbre (vertical overlap).

Vertical integration and/or segregation

8. Consider vertical integration as a cause of timbre creation and variance:
    a. Timbre as a cause of integration and/or segregation.
    b. The dimensional approach to timbre, including emergent properties relating to pitch (horizontal overlap).

c. Texture resulting from contrasting timbres.

9. Also consider vertical integration or segregation in terms of the potential for psychoacoustic dissonance and musical dissonance (and consonance).

Assimilation and meaning

10. Consider the nature and type of discourse on the source-cause dominant to typological-relational dominant continuum, and the way it varies over time.
11. Consider implication-realization, and arousal and meaning on a moment-to-moment basis throughout the work.
12. Consider global organization in time – identify formal structures, like sectional or continuous organization, and the nature of the relationships between sections, i.e. hierarchical relationships.

3 Applying the methodology

Having outlined a methodology, and considered some preliminary issues, let us examine each of the steps in the above procedure in more detail.

3.1 Segregation of sonic objects

The first stages in our methodology are the identification the sonic objects (or events), and establishing the factors responsible for identification.

Building on the work of Bregman, McAdams has postulated several stages in the recognition of sound sources and events (McAdams, 1993) [9]. The first two stages are sensory transduction and auditory grouping of frequencies – both simultaneous and sequential. The third stage involves the analysis of auditory properties and features.

Auditory property analysis provides a group of abstract properties which act as input to the matching process with representations in memory. Whether through "comparison" or "activation", the result of the matching process will either be a match with a category, no match with a category, or too many matches.

There may be many factors responsible for categorical matching so the next step in our methodology seeks to explore the varied factors and the possible contribution made by each factor in sound event recognition.

3.2 Factors in the identification of environmental sounds

Acoustic factors

Gygi (2001) presents results of similarity studies conducted at the Hearing and Communication Laboratory at Indiana University. He points out that for "meaningful" sounds, there needs to be a consideration of their "psychological space". [4]

One experiment attempted to uncover the structure of such a psychological space by using multidimensional scaling (MDS) procedures.

Gygi conducted the experiment using 50 environmental sounds. Recordings of 100 instances of the environmental sounds were played to listeners who were asked to rate the similarity of each pair on a seven point scale from least similar to most similar.

A two-dimensional MDS solution was derived. Gygi searched for a meaningful interpretation of the two dimensions. He measured over twenty acoustic variables covering different aspects of spectral distribution. The highest correlations for Dimension 1 were spectral spread, and confidence in the pitch of the signal. Correlations with the second dimension were not so clear cut. The most significant correlations were to do with measures of rhythmicity and spectral mean.

Gygi's identification studies gives us a basis for isolating acoustic features of environmental sounds that we may want to single out for analysis.

Syntactic and semantic factors

In an article by Howard and Ballas (1980), they documented their experiments which tested syntactic and semantic factors in the classification of non-speech transient patterns [7]. What is important about this series of experiments is that they are concerned with *patterns* of sounds (i.e. sequences of sounds) – both pure tones and environmental sounds.

Their paper attempts to shed some light on syntactic (temporal structure) and semantic (knowledge of source events) factors in transient pattern recognition.

Howard and Ballas found that the grammatical group performed significantly better than did the non-grammatical group, and that listeners in the grammatical group had actually learned something about the syntactic rules (the grammar) used to generate the target patterns.

In one of their experiments, Howard and Ballas used real-world sounds to create patterns, and some of these sounds were required to be semantically sensible.

From the results of this experiment, Howard and Ballas observed that the grammatical groups performed at significantly higher level than the non-grammatical groups (approaching the significance of experiment one), and semantic instructions only enhanced the performance for those who also received syntactically structured target patterns (i.e. the grammatical/semantic group). In addition, the grammatical/semantic group performed significantly better than the grammatical/non-semantic group.

3.3 Horizontal integration and/or segregation

In this next phase of our analytical procedure we must identify sequences (streams or chains). Such a sequence would consist of sonic objects linked together and it would function as a single unit that we could call a "pattern". The role of "trajectories" should also be considered. (The term "gesture" has sometimes been used in this context.)

Next we need to determine the linkages between the sonic objects within each chain-type pattern. Which of the relevant factors are operating, and what are the relative weightings of those factors?

Then we shall determine the relationships between "pattern objects" – if this level of syntax exists. This amounts to the investigation of higher-order relationships within a "hierarchy".

Finally we must consider local organization in time – pulse, beat, accent, rhythm, and meter.

This phase is a lot like a melodic analysis in tonal music. Transformations in pitch or frequency define a melodic form (or gesture). However, we shall treat pitch as a quality that is emergent from the frequency spectrum of sonic objects within acousmatic music.

A sequence with large frequency transitions in a short space of time will not remain perceptually coherent. Alternations of high and low frequencies will cause segregation between the two different registers and the perception of separate streams will result.

Changes in timbre can effect the integration of a horizontal sequence, for example:

- Repeated and/or rapid changes in timbre can fragment a sequence.
- Less rapid shifts in timbre can be used to delineate larger horizontal units or "phrases".

A distinction can be made between form-defining sequential events and subordinate ornamental sound events. Ornamental events group with the form-bearing event they are subordinate to. Gestalt-like principles of primitive scene analysis seem to constrain these emergent events. The dependent event must be very close to the anchor event in frequency and time, otherwise the two will not group to form a larger event.

The choice of which event is the anchor and which is the dependent one depends on factors such as duration, intensity, and rhythm. Dependent events tend to "resolve" to stable anchor events.

Trajectories

Is a regular sequence of sounds (where the sequence is predictable) easier to recognize than an irregular one?

As we listen repeatedly to an auditory pattern, we learn the regularities in it. We can then anticipate segments of the pattern before they occur and integrate the sequence into a coherent mental representation. This preparation of our attention through repeated listening should assist the formation of sequential streams.

Using a rising or falling series of pitches as an example of a predictable sequence, Bregman poses the question: "Can primitive scene analysis follow and segregate such trajectories from their acoustic contexts?" (Bregman, 1999: 670) [2]

There is evidence both for and against such abilities in the formation of streams via trajectories. The concept of *trajectory* is important, but its application is far from straight forward.

3.4 Vertical integration and/or segregation

In this next stage we consider vertical integration as a cause of timbre creation and variance. We also consider vertical integration or segregation in terms of the potential for psychoacoustic dissonance and musical dissonance.

Timbre and texture

Bregman notes that timbre plays a role in the sequential organization of music. Timbre is a complex phenomenon and no one perspective can encapsulate its complexity. Bregman tackles timbre from several points of view.

Timbre as a cause of segregation and integration

As we saw above, large and rapid changes in timbre can cause the formation of separate, parallel streams as happens with a "compound melody".

Conversely, timbre can be used as sequential glue for musical phrases, and to delineate separate units or sections. Such delineation exploits scene analysis principles from nature where a sudden change in timbre usually implies a new event has begun. In contrast, a continuous change implies that a single event is changing in some way – through some form of incremental transformation.

The dimensional approach to timbre

Bregman points out that there is a use of timbre that is stronger than the accentuation and reinforcement of melodic forms.

Bregman highlights four approaches to the study of the dimensionality of timbre. Similarity studies have determined the number and nature of dimensions involved in the specification of timbre. Lerdahl (1987) has attempted some tests of the organization of music according to certain types of dimensionality [8]. Other studies have explored acoustic dimensions such as the formant frequencies of vowel sounds. Slawson (1985) has attempted to organize music according to these timbrel dimensions [12]. The physical modelling approach assumes our brains are built to form descriptions of environmental events rather than sounds in the abstract. In this view, changes in sound attributes correspond to changes in dimensions or components within a mental model of a physical system that produces the sound.

Timbre as the result of fusion, texture as a result of contrasting timbres

Knowledge of fusion and segregation principles can be used to assist the organization of musical texture, for example a number of musical sounds can be fused to create a global timbre, or a polyphonic texture can be created where two or more distinct sequential streams can be heard.

Tonal music uses scene analysis principles in a number of ways that are relevant to musical texture. Perhaps these principles could be adapted for the analysis of texture in acousmatic music.

An examination of the role of primitive scene analysis in counterpoint may be of some benefit too. In polyphonic music, the parts must not be totally segregated or totally integrated. Segregation between parts is improved by strong sequential organization within each part. Strong sequential organization is achieved through principles such as: using small pitch changes to favour sequential integration; avoiding "common fate" by prohibiting synchronous onsets and offset (different rhythms); avoiding parallel changes in pitch between two parts, i.e. encourage contrary motion or oblique motion; avoiding harmonic relations between simultaneous events in different parts, e.g. whole number ratios like 2:1, 3:2, etc.

Dissonance

Bregman distinguishes two types of dissonance: psychoacoustic dissonance and musical dissonance. Psychoacoustic dissonance is the sense of roughness caused when partials combine to produce a large number of beats at an unrelated rate [2]. Simple ratios like 3:2 (ca. seven semitones) will sound smooth, complex ratios like 45:32 (ca. six semitones) will seem rough.

Musical consonance is defined by Bregman as a cognitive experience. Stable and unstable combinations of sounds are defined by a musical style. Unstable combinations are points of tension, whereas stable ones are points of rest. Bregman notes that unstable combinations of tones often happen to be psychoacoustically dissonant in Western music.

Composes control such dissonant tones by using techniques such as: avoiding simultaneous start and stop times; capturing tones into separate streams by preceding each with tones close in pitch; capturing each into smooth and different trajectories; capturing each into its own repetitive sequence.

An increase in dissonance can be created by violating the above principles.

3.5 Assimilation and meaning

In this final stage we consider global organization and formal structures. This is also the stage in which we consider emotion, arousal and meaning, implication-realization, and the nature and type of discourses operating within the work.

Space does not allow us to consider these concepts within this paper, but the model displayed in Figure 2 draws upon the work of Dowling and Harwood (1986) [3].

4 Conclusion

The above approach is a reductionist one – tearing the musical materials apart. We must now put the work back together again in order to get a more holistic view.

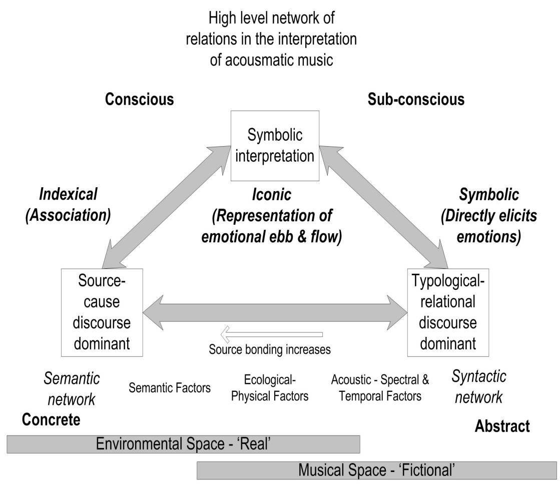


Fig. 2: Network of relations in the interpretation of acousmatic music.

Figure 2 is a pictorial representation of the concepts introduced in this paper and their relationships. Of fundamental importance is the differentiation between "Environmental Space" associated with the experience of "real-world" sounds and "Musical Space" associated with the "fictional sound world" created by the musical work. Note that there is scope for overlap between these worlds so that the composer can play with ambiguities and chimerical properties resulting from their combination.

In assessing the dominant forces at play within a work, a continuum has been constructed according to the type of discourse that may be dominant at any given time. Source-cause discourse (after Smalley, 1994) is dominant when sounds from the natural world are identified and related somehow within a semantic network [13]. These sounds will have strong source-bonding and the syntax of their relationships will be dominated, although not exclusively, by semantic and ecological factors.

As source-bonding decreases, the acoustic factors begin to dominate and the work adopts a typological-relational discourse based upon more abstract relationships within a syntactic network governed by implications and realizations. The continuum extends from the purely concrete reproduction of the environment (real space) to the totally abstract world of a fictional sound space with fictional sources (musical space).

References extend from the indexical to the symbolic, from simple association at one extreme to directly eliciting emotions at the other extreme, with a representation of emotional ebb and flow in between. The listener's experience along these various continua can vary from moment to moment within a given work, and this becomes one more dynamic aspect within the discourse.

What remains to be completed is to test this model of analysis on representative repertoire works.

5 References

[1] Ballas, J. A. (1993). Common factors in the identification of an assortment of brief everyday sounds. *Journal of Experimental Psychology: Human Perception and Performance* 19(2), 250-267.

[2] Bregman, A. S. (1999). *Auditory Scene Analysis: The Perceptual Organization of Sound*. (Second MIT Press Paperback edition) Cambridge, MA: MIT Press.

[3] Dowling, W.J. and Harwood, D.L. (1986). *Music Cognition*. Orlando: Academic Press.

[4] Gygi, B. (2001). *Factors in the Identification of Environmental Sounds*. PhD Dissertation: Indiana University

[5] Hirst, D. (2002). Developing Analysis Criteria Based on Denis Smalley's Timbre Theories. *Proceedings of the Australasian Computer Music Conference 2002*, pp 43-52. Australasian Computer Music Assoc. Melbourne.

[6] Hirst, D. (2003). Developing a cognitive framework for the interpretation of acousmatic music. *Converging Technologies: Proceedings of the Australasian Computer Music Conference 2003,* pp 43-57. Australasian Computer Music Assoc. Melbourne.

[7] Howard, J. H. & Ballas, J. A (1980). Syntactic and semantic factors in the classification of nonspeech transient patterns. *Perception & Psychophysics* 28(5), 431-439.

[8] Lerdahl, F. (1987). Timbral hierarchies. *Contemporary Music Review* 2: 135-60.

[9] McAdams, S. (1993). Recognition of sound sources and events. *Thinking in Sound: The Cognitive Psychology of Human Audition*. (ed. S. McAdams and E. Bigand). Oxford University Press, pp. 146-98.

[10] Narmour, E. (1989). The "genetic code" of melody: Cognitive structures generated by the implication-realization model. *Contemporary Music Review*. Vol. 4, Harwood Academic Publishers, UK, pp. 45-63.

[11] Pierce, C. S. (1931-1935). *Collected Papers (Vols 1-6)*. (C. Hartshorne & P. Weiss, Eds.). Cambridge, MA: Harvard University Press.

[12] Slawson, W. (1985). *Sound Color*. University of California Press, Berkeley, CA.

[13] Smalley, D. (1994). Defining Timbre - Refining Timbre. *Contemporary Music Review* Vol. 10 Part 2, Harwood Academic Publishers, Switzerland, pp. 35-48.

Configurable Hemisphere Environment for Spatialised Sound

Greg Schiemer, Stephen Ingham, John Scott, Aaron Hull, Damien Lock, Didier Balez, Gareth Jenkins, Ian Burnett, Guillaume Potard and Mark O'Dwyer

Faculty of Creative Arts, Faculty of Informatics, University of Wollongong

Email: schiemer@uow.edu.au; singham@uow.edu.au; john_scott@uow.edu.au; aaronh@uow.edu.au; damien@coscilia.com; bdidier@uow.edu.au; neathisdead@yahoo.com; i.burnett@elec.uow.edu.au; gp03@uow.edu.au; mo15@uow.edu.au

Abstract

This paper reports on a cross-disciplinary research initiative in spatialised audio at the University of Wollongong. It describes a multi-channel sound environment in the School of Music and Drama and its background. Some of the creative tools developed by software engineers working from the Whisper Research Laboratory in the Faculty of Informatics are presented together with strategies for usage by sound designers, composers and installation artists working with spatialised sound.

1 CHESS – a Configurable Hemisphere Environment for Sound Spatialisation

The CHESS system is an immersive sound projection system built to audition new forms of sound localisation. It consists of a loudspeaker array mounted in a spherical structure that can be configured to audition sound in a variety of digital formats. The user is able to configure many speaker positions including standard configurations and other customised speaker placements.

The CHESS system was built to serve digital audio research undertaken by engineering specialists from the Faculty of Informatics. It is used by Informatics staff and a PhD candidate from the Faculty of Informatics.

The system is located in the sound studios of the Faculty of Creative Arts where it is accessed by artists working in sound allowing them to monitor work produced in 3D audio. Within the Faculty of Creative Arts users currently include staff from the School of Music and Drama and the School of Journalism and Creative Writing. In time this will also include the School of Art and Design as new Post-Graduate programs are introduced.

The ability to reconfigure speakers in a variety of positions enables artists to create new simulated environments quickly and and experiment with them in production. Speakers are attached to hinged brackets mounted on the frame. These brackets can be positioned in a variety of ways e.g. raised, lowered or swivelled horizontally, allowing speakers to be repositioned around, above and below the listener.

2 History of CHESS

The CHESS system was set up with the help of an ARC RIBG grant in 2001. This was an initiative of the Digital Media Centre which sought to support new developments in interactive television. The initiative involves a collaborative research effort between the 3D-audio team from the Informatics Faculty led by Ian Burnett and staff in the sound, composition, music and production program in the Faculty of Creative Arts led by Stephen Ingham. In 2003 a further RIBG grant secured additional digital recording equipment and signal processing hardware developed by Lake Technology. Lake also made substantial financial contributions.

Didier Balez, a sculptor who works with steel and ceramics, and technician in the Visual Arts workshop in the Faculty of Creative Arts, constructed a customised steel structure to house the CHESS system working under the supervision of Ian Burnett and Stephen Ingham. The system was officially launched by the Pro-vice Chancellor (Research) in August 2003 at the opening of Sonic Connections, a 2-day showcase of research and creative work at the Faculty of Creative Arts [1].

The CHESS system uses a Pentium 4 PC communicating with a Mac G4 via a LAN. The PC provides a user interface for spatialising audio produced by the G4.

Spatialisation takes place using a java program that runs on the PC which transmits a stream of spatial coordinates to the G4. The G4 processes the audio which is fed to amplifiers and loud speakers. Figure 2 describes the current configuration of CHESS.

Figure 1. Configurable Hemisphere Environment for Sound Spatialisation (CHESS)

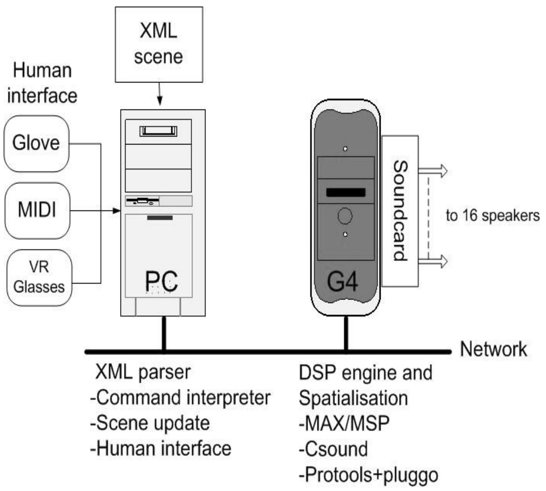


Figure 2. Current processor configuration used in CHESS

The java application running on the PC codifies these coordinates as a stream of spatial data called a 'scene'. Scenes are described in XML format and communicated to the G4 using the UDP protocol. The role of the G4 is to run signal processing algorithms to place sounds in a virtual auditory space.

These algorithms include:
Spatialisation: Ambisonics (1st to 4th order) or VBAP;

Delays and attenuations to account for the distance of sound sources and the properties of the medium;

Decorrelation, to create broad sound that can be used for beach or waterfall sound sources;

Calculation of reflections and obstructions to compute early reverberation;

Simulation of late reverberation by a Feedback Delay Networks algorithm controlled by perceptual parameters.

Using a PC to define the 3D scene and transmitting this to the G4, frees G4 resources for audio signal processing in user applications. Currently 24 sound sources can be spatialised by the 867MHz G4.

Figure 2 shows a sound card (Digi-001) connected to 16 speakers (1029a Genelec) which make up the hardware associated with the CHESS system. A Soundfield microphone is used to capture real 3D sound environments that can then be played back accurately in CHESS or be mixed with other sounds and environments.

3 Key Developments - Informatics

Research done by staff in Informatics over the past three years involves extensive publication and commercially sensitive research. Areas of published research include:

investigation of the apparent source width and shape of 3D sound [2][3][5][6];

Development of an XML scheme for describing 3D audio scenes [4][8];

development of a Java3D application by Mark O'Dwyer for parsing, rendering and visualising 3D audio XML scenes [2];

development of a java application for describing 3D audio scenes using XML [7][8];

development of Max objects for 4th order Ambisonics spatialisation and 5.1 panning (Potard, Informatics);

development of Max external objects to produce wide sound sources (Potard, Informatics);

development of software enabling B-format recordings to be played back in 5.1 and other spatialisation formats up to 16-Channel;

development of a set of pluggins to perform up to 4th order Ambisonics spatialisation on up to 16 speakers from within Protools LE;

developed tools for virtual positioning of sound objects using data glove and virtual reality glasses (O'Dwyer, Informatics);

comparisons made between the precision and source localisation characteristics of different spatialisation algorithms;

Figure 3 shows the Max panning object for the CHESS system.

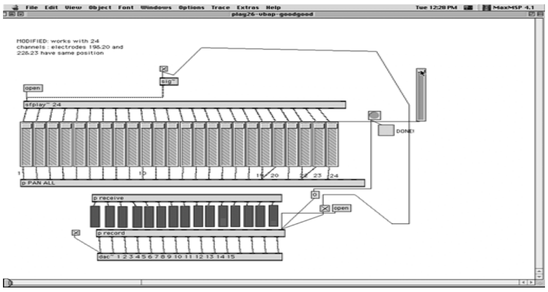


Figure 3. Max panning object for CHESS

3 User software

Currently Max/MSP is the principal tool for signal processing on the G4. Other musical applications have also been used including Csound and Protools using the Pluggo plug-in architecture. The chosen spatialisation algorithm is Ambisonics as it provides a format which is independent of a particular speaker configuration. Thus created scenes can be played on a different speaker configuration with out the need to regenerate scenes. The sweet spot and stability of sound sources are greatly improved using higher order ambisonics (HOA). Using 4th order ambisonics 25 signals are used to represent the sound field compared to 4 for 1st order (i.e. B-format).

Alternatively, the Max object, called VBAP (an acronym for Vector Based Amplitude Panning) can also be used to perform spatialisation of sound sources. Spatialisation using VBAP is less precise than ambisonics.

The Protools Plugin, called Pluggo RTAS (an acronym for Real Time Audio Suite) is relevant for Max users. It allows the CHESS system to play audio encoded in ambisonic B format by converting it to formats playable in any of the speaker configurations available in CHESS.

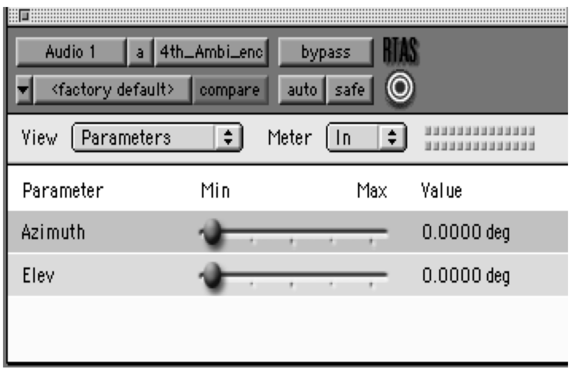


Figure 4. Protools Pluggo Plugin

This application has been used by Gareth Jenkins, a Master's candidate working under the supervision of John Scott in the School of Creative Writing, in collaboration with Guillaume Potard, to create a narrative work lasting 45 min entitled "*Owen's Waiting*". It was produced in Protools and encoded in 16 channels on the CHESS system using Pluggo plug-in. It was presented at Newcastle ElectroFringe in October 2003.

The first postgraduate users of the CHESS system from the Sound Composition Music and Production program are Masters candidates Damien Lock and Aaron Hull. Lock will investigate possible use of spatialisation in animation post production packages like 3D Studio max or Maya "Physics Engine" Plugins while Hull will further investigate possibilities for spatialisation in live performance using the Protools Pluggo Plugin.

Csound has also been used to produce spatialised audio with CHESS. Csound can also be used as a Csound object in Max/MSP and as a means for creating VST pluggins. This allows users familiar with Cubase VST to work with the CHESS system.

Csound was used by Greg Schiemer and Guillaume Potard for sonification of brain data in a work entitled "*Neural Dialogues*" which was submitted for the ICAD project entitled "Listening to the Mind Listening" [9]. Due to bandwidth limitations of the G4, it was necessary to work in two stages, one using Csound to create a 24-channel sound file which was panned using a Max program consisting of 24 VBAP object.

The Lake Huron Digital Audio Convolution workstation acquired in 2003 will overcome these bandwidth limitations. The Huron offers several spatialisation tools including binscape used to monitoring 3D sound environments in binaural headphones using Lake Theatrephone. The current Huron configuration will support 16 channels and with additional DSP cards will support up to 256 in realtime.

Currently CHESS is used as part of the "Listening to the Mind Listening" project as the main processing station for concert submissions [10]. The Lake Huron and custom Max/MSP patches are used to process submissions to a binaural format and for a custom 15-speaker array in the Studio, Opera House. Using the flexibility of CHESS, a scaled down replica of this speaker configuration has been set-up.

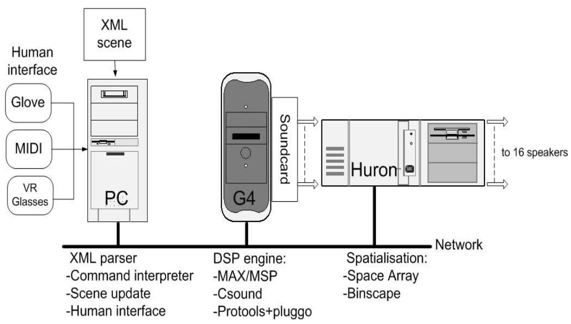


Figure 5. Projected future configuration using the PC/G4 with Lake Huron

5 Projections for the future

Although CHESS is designed as a reconfigurable speaker array it provides a relatively small audition space that is ideally suited for a single listener. The permanent nature of the frame supporting the speakers makes its relocation impractical. In its current state it does not address the need to project spatialised audio in public performance venues.

Further funding has been sought to extend the CHESS system for use in concerts and sound installations. The current system will continue to be used as a research and production space. Its extension will support presentation using spatialised audio in public exhibition spaces that provide a suitable location for larger audiences. This will also accommodate developments in new media performance where sound is an interactive partner with visual, haptic and kinaesthetic media.

The CHESS system represents a collaboration that crosses disciplinary boundaries. New audio tools have been developed and will continue to support digital audio research that leads the development of new standards. Development of new software extends the capabilities of sound production tools used by various composers and sound designers irrespective of their technical expertise or musical aesthetic. CHESS has encouraged new research along with unexpected and mutually beneficial outcomes.

6 Acknowledgements

We wish to acknowledge the following sources of cash support:

RIBG Round 2 University of Wollongong 2001;
RIBG Round 2 University of Wollongong 2003;
Lake Technology 2003.

References

[1] Sonic Connections, Faculty of Creative Arts, University of Wollongong, July 31-August 1 2003, curated by Greg Schiemer http://www.uow.edu.au/crearts/SonicConnections.html

[2] O'Dwyer, M., Potard, G. and Burnett, I. : "A 16-speaker 3D audio-visual display interface and control system", Poster in Proceedings of the 2004 International Conference on Auditory Displays, Sydney, Australia, 6-9 July 2004

[3] Potard, G. and Burnett, I. (2004) "Control and measurement of apparent sound source width and its applications to sonification and virtual auditory displays" in: Proceedings of the 2004 International Conference on Auditory Displays, Sydney, Australia, 6-9 July 2004

[4] Potard, G. and Burnett, I. (2004) "An XML-based 3D audio scene metadata scheme" in: Proceedings of the 25th AES conference, London, UK, 17-19 June 2004

[5] Potard, G. and Spille, J. (2003) "Study of Sound Source Shape and Wideness in Virtual and Real Auditory Displays" in: Proceedings of the Audio Engineering Society 114th convention, Amsterdam, March 2003, Preprint 5766

[6] Potard, G. and Burnett, I. (2003) "A study on sound source apparent shape and wideness" in: Proceedings of the 2003 International Conference on Auditory Displays, Boston, USA, 6-9 July 2003, pp 25-28

[7] Potard, G. and Ingham, S. (2003) "Encoding 3D sound scenes and music in XML", in: Proceedings of the International Computer Music Conference (ICMC 2003), Singapore, 2003

[8] Potard, G. and Burnett, I. (2002) "Using XML schemas to create and encode interactive 3-D audio scenes" in: Proceedings of DCW2002, Sydney Australia, April 2002, Lecture notes in Computer Science, Springer-Verlag, pp 193-202

[9] Potard, G. and Schiemer, G. (2004) "Sonification of the Coherence Matrix and Power Spectrum of EEG Signals" Listening to the Mind Listening, Concert of Sonifications, ICAD 2004 Sydney Opera House Studio Thursday 8 July, 2004.

[10] Listening to the mind Listening, ICAD 2004, http://www.icad.org/websiteV2.0/Conferences/ICAD2004/concert.htm

Exploration of the Height Dimension in Audio Reproduction

Jim Barbour

Media and Communications, Swinburne University of Technology
Email: jbarbour@swin.edu.au

Abstract

Multi-channel surround sound audio with added height loudspeakers offers a richer perceptual experience than traditional stereo or 5.1 reproduction, recreating truly three-dimensional aural soundfields. Based on human auditory perception, surround sound technology may assist the exploration of acoustic space as an expressive dimension.

1 Introduction

There is growing interest among composers, audio designers and music recording companies to use height loudspeakers to increase the sonic possibilities for audio reproduction, including enhanced realism of spatial location, spatial envelopment and improved sound stage detail. Music recording company MDG from Germany are releasing DVD-Audio discs with height information, and note:

> The reflections of even the subtlest sounds from the ceiling and floor will reliably tell us how large a room we are in. (MDG, 2001, 8)

Widespread acceptance of the ITU-R BS.775-1 specification[1] for five loudspeakers mounted on the horizontal plane plus a low frequency enhancement loudspeaker has seen the growth of home theatre systems incorporating DVD-Video and 5.1 surround sound. The introduction and growing adoption of the new commercial formats of DVD-Audio and SACD[2] now offer the ability to deliver audio recordings in uncompressed, high-resolution multi-channel formats. This paper proposes an extension of the 5.1 standard to include two full-range loudspeakers mounted above the listening position. Acoustic perception studies underpin this proposition and there are examples of elevated sources, recording techniques, delivery formats and commercial applications.

1.1 Binaural Discrimination

Blauert [6] described the head-related system of spherical co-ordinates using the horizontal plane, frontal plane and median plane, with angles of incidence for azimuth and elevation, Figure 1. For the horizontal plane, numerous studies have investigated the perceptual relationship between inter-channel differences, phantom image locations and panning law accuracy, including [12] [16] [27]. Sound source localization in the horizontal plane and the frontal plane primarily relies on binaural discrimination due to inter-aural amplitude, inter-aural time and inter-aural spectral differences, with the exception of the median plane where inter-aural differences are zero for a normal listener. Research into localization in the median plane, [5] [6] [21], concluded that three requisites were necessary for an auditory stimulus to be accurately located in a vertical space: (a) the sound must be complex, (b) the complex sound must include frequencies above 7000 Hz, and (c) the pinna must be present. Our ability to localize sound sources on the median plane relies primarily on pinna effects. According to Blauert [5][6] and others, the curves and ridges within the pinna reflect different frequencies depending on source elevation.

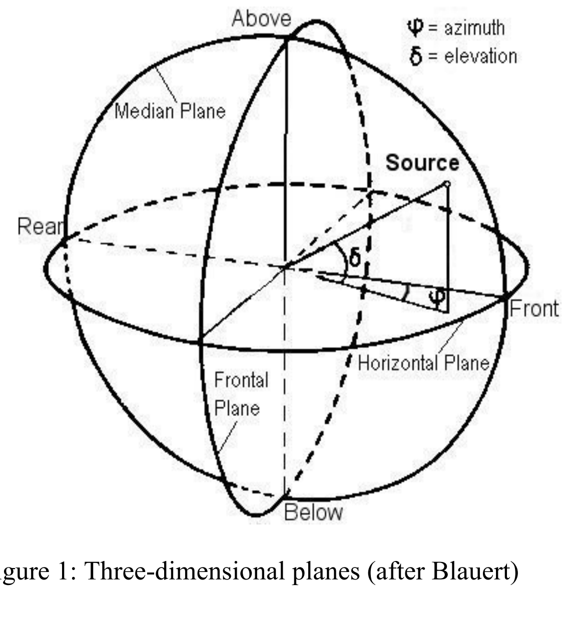


Figure 1: Three-dimensional planes (after Blauert)

2 Aural Perception of Elevated Sources

The author conducted a series of experiments, reported at an AES conference [3], investigating aural perception of phantom images in the vertical

[1] ITU-R BS 775-1 is the 5.1 arrangement of five equi-distant loudspeakers, front left and right, centre, sub, rear left and right

[2] Super Audio Compact Disc

Proceedings ACMC 2004
ISSN: 1448-7780
http://www.acma.asn.au

hemisphere generated by two loudspeakers using inter-channel amplitude differences. The key findings of these experiments are illustrative to further discussion of the exploration of height in reproduction. Using a custom built metal frame, two carefully matched loudspeakers were placed at different locations above and around the listening subjects. A short soundtrack consisting of male speech and reference pink noise was played and subjects indicated where they perceived the sound source to be. Results for each pair of loudspeakers show the median location for the perceived source and the deviation among the subjects tested. Where the deviation is large, it indicates a significant difficulty in accurately locating the source, referred to as localization blur. While more positions were tested than are presented here, following are the key results.

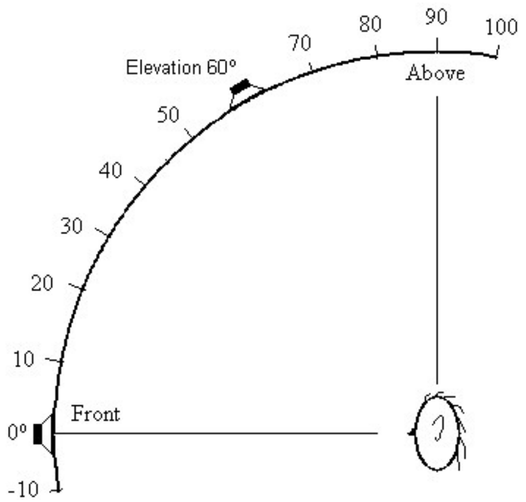


Figure 2: Loudspeaker locations for median plane test: 0°- 60°

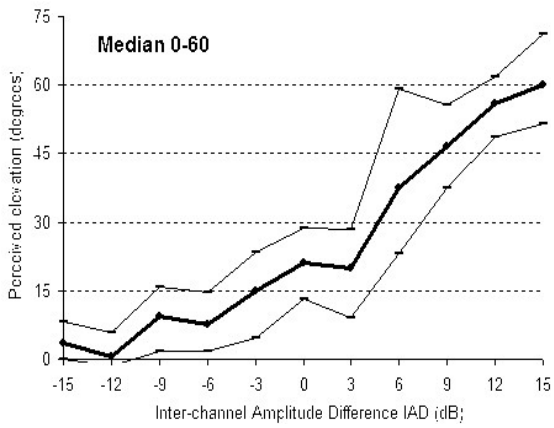


Figure 3: **Median Plane 0°- 60°**: median and standard deviation

For phantom images on the median plane with loudspeakers at 0° and 60°, Figure 3, the median of perceived locations was weighted towards the 0° elevation loudspeaker until the inter-channel differences were above the mid point, with wider deviations in the middle positions indicating a significant localization blur.

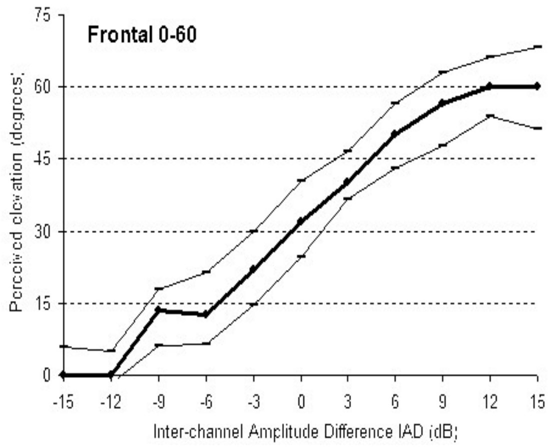


Figure 4: **Frontal Plane 0°- 60°**: median and standard deviation

However, for the frontal plane, Figure 4, the results were more evenly spread across the range of inter-channel differences, with a consistently smaller localization blur than for the median plane. In discussions with subjects, there was still a blur for the frontal plane but there was greater confidence in perception than for the median plane.

Figures 5 and 6 show results for tests with the loudspeakers mounted at 60° and 120° above the subjects, measured on the median and the frontal planes.

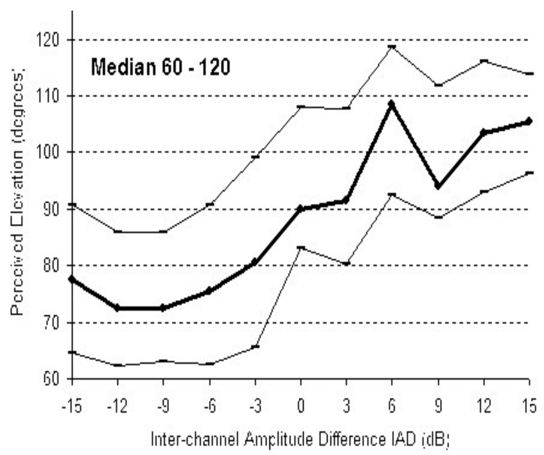


Figure 5: **Median Plane 60°- 120°**: median and standard deviation

For inter-channel amplitude differences on the median plane, Figure 5, the ability to perceive locations is very poor, with significant inaccuracy in median locations and significant localization blur. It could be concluded that the effects of the pinna are breaking down at these elevations.

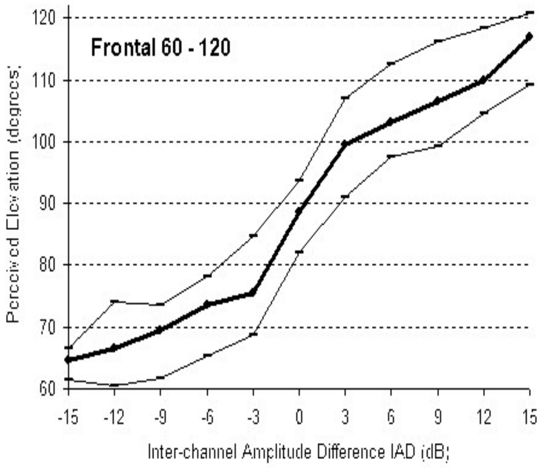


Figure 6: **Frontal Plane 60°- 120°**:
median and standard deviation

However, in the frontal plane, Figure 6, median locations correspond well with inter-channel differences and standard deviations are relatively small. This result would be consistent with inter-aural differences playing a large part in location, as the angle between the loudspeakers is equivalent to the 60° angle subtended by the horizontal stereo loudspeaker locations. However, the degree of localization blur for the vertically mounted pair is greater than for a horizontal stereo pair, according to the results published in [12], [16] and elsewhere.

Comparing front and overhead median plane perception, the front arc has a higher degree of accuracy in the perceived elevation and less localization blur. This difference between front and overhead localization on the median plane is consistent with pinna effects, where sounds from the front are reflected from the pinna into the ear canal whereas sounds from above and behind are not well reflected.

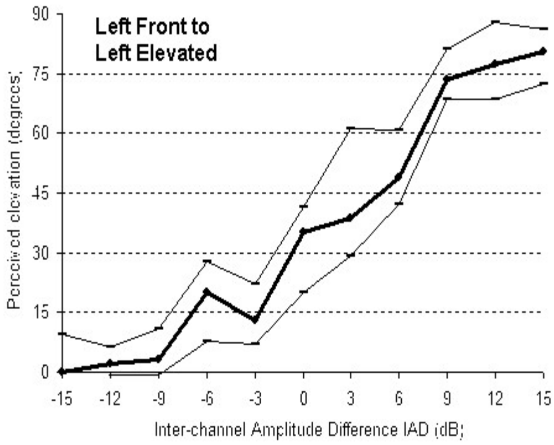


Figure 7: **Left Front to Left Elevated:**
(azimuth 90° elevation 60°)
median and standard deviation

Figure 7 presents the results for loudspeakers mounted at Left Front (azimuth 30° elevation 0°) and Left Elevated (azimuth 90° elevation 60°), showing there is a good degree of accuracy in perceived locations, though with some localization blur.

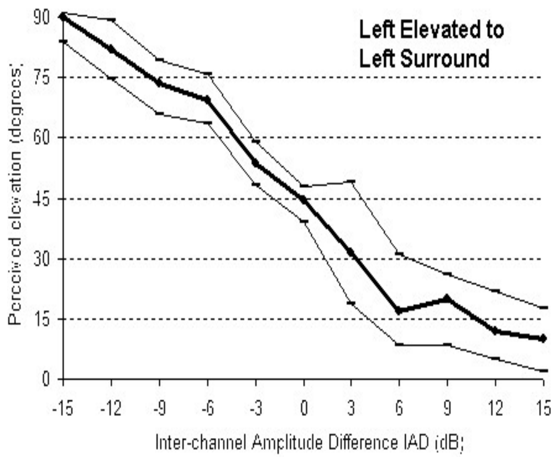


Figure 8: **Left Elevated to Left Surround:**
median and standard deviation

Figure 8 shows elevation perception behind the listener from the Left Elevated loudspeaker down to the Left Surround position. Similar to the front elevation perception of Figure 7, there is accurate median locations with localization blur increasing as the phantom image moves to the rear.

It was clear from the experiments conducted that localization on the median plane is poor for phantom images generated by two loudspeakers, with inaccurate median locations and significant localization blur. Further, for sources above the subject on the median plane the results were particularly poor. This suggests that a single loudspeaker placed on the median plane and used for height enhancement would be of little benefit for vertical localization.

Localization of phantom images was consistently better on the frontal plane than for the median plane. As the angle of elevation for the height loudspeaker increased, the accuracy of localization on the frontal plane improved and the localization blur decreased. This is consistent with the increasing influence of inter-aural differences as the angle of elevation increased. Also, accuracy of localization 'across the top', from 60°-120°, reflects similar results to the standard horizontal stereo positions. As a consequence, it is proposed that two loudspeakers be mounted on the frontal plane, azimuth 90°, evenly spaced either side of the median plane, elevation 60°, to generate effective localization and height envelopment.

3 An Acoustic Audit of Height

In our everyday environment, there are many sound sources with a significant angle of elevation. The author has been conducting detailed listening to establish an inventory of real sounds with elevation characteristics, based on Acoustic Ecology principles articulated by Truax [28], Schaffer [23] and others. When specifically considering elevated sound sources, there are natural environments like open fields, forests, caves and seascapes, and constructed environments including city and suburban streets, under bridges, inside buildings and underground tunnels and trains. Within each environment, there are natural elevated sources like wind effects, flora and fauna, and technology sources like flying machines, tall ground-based objects, ceiling mounted loudspeakers and many more. Also important in the context of this paper are the acoustic reflections from the upper reaches of halls, cathedrals and concert venues where aural events are experienced and recorded. An extensive listing of elevated sound sources is beyond the scope of this paper, but a few common examples in Table 1 will illustrate key characteristics.

Elevated sound sources in the Open:
Aeroplanes:
very high: low frequency spectrum drone, slow and small Doppler shift, sound well behind visual position, some localization blur
low, barely above roofline, eg helicopter: high frequencies clear, sharp transients, rapid and wide Doppler shift, distance effects of amplitude and frequency spectrum very clear, precise localization
Birds:
unusual to be high, eg crows: low frequency spectrum of calls, slow change in position and sound
flying low: aurally appear and disappear very quickly, rapid Doppler shift, wing flapping heard when very close, precise localization
stationary in trees: bright spectrum of calls, precise localization
Inside buildings:
footfall above and floor creaking: low frequency spectrum, poor localization
fluids in utility pipes: low frequency spectrum, possible localization
rain on the roof: dull, muffled sound, no specific location
Performance venues:
direct sound from high galleries at the side, front or rear: clear full spectrum, precise localization
acoustic reflections: amplitude and frequency spectrum highly variable, imprecise localization, distinct echoes occasionally heard

Table 1: characteristics of elevated sound sources

Key factors identified in locating elevated sources in everyday environments are fundamentally the same as for ground based or horizontal sources. These include HRTF effects of inter-aural time, inter-aural amplitude and inter-aural spectral differences along with torso influences, distance effects of amplitude, spectral changes and the direct to reverberation ratio, and motion effects of Doppler pitch shifts, amplitude and spectral changes. The principle additional factor utilized for elevation perception is pinna effects. As identified by the experimental tests described above, when an elevated source lies on the median plane and would present difficulties in accurate location, a listener would further interrogate the auditory scene for additional information to resolve the location, for example, head turning.

4 Architectural parameters

To consider spatial enhancement as a starting point for incorporating height loudspeakers, it is useful to look at architectural parameters that are indicators of high quality acoustic environments for music performances. In his book, *Concert and Opera Halls: How They Sound*, [4], Beranek analysed the objective acoustic characteristics and the subjective critical acclaim of many venues throughout the world. He conducted a detailed survey of musicians, conductors and critics worldwide to develop a scale for the assessment of the quality of performance spaces. Key questions included assessments of spaciousness, clarity, reverberation warmth and character, and acoustic support for the performer. Simultaneously, he conducted testing of the venues to determine physical measurements of key acoustic properties.

4.1 IACC

One important measurement in performance venues is the Inter-aural Cross Correlation coefficient, IACC, which measures the difference in reflection patterns from either side of a hall. In his book *Concert Hall Acoustics*, Ando [1] suggests that ‘all the available data indicates a negative correlation between the magnitude of IACC and subjective preference.’ (Ando, 1985, 77) This implies that a hall that has the greatest difference in architectural features between each side sounds most pleasing. Ando goes on to explain:

> Strong reflections from the ceiling and rear walls will increase IACC. But simply to absorb this is inadequate because we also need a sufficiently generous sound energy supply for the seats in the rear half of the hall. For this purpose, Schroeder proposed a highly diffusing ceiling to reflect most of the sound

> energy to the side walls, so that it would arrive at the listeners from a suitable direction for minimizing IACC. (Ando, 1985, 90)

Through careful consideration of the subjective assessments and objective measurements, Beranek established six orthogonal acoustic parameters and the range of variation of those parameters for the best venues in the world, acknowledging that the range of each variable was dependent on the style of music performed, for example, Baroque, Romantic, Modern or Opera. The percentage weighting indicates Beranek's assessment of the relative importance of each parameter.

Apparent Source Width (1-IACC)
prefer approx 0.7-0.9, 25% weighting
Early decay time: reverb to –10dB
0.3 – 0.4 seconds, 25% weighting
Surface Diffusion Index
0.9, aided by irregular surfaces, 15% weighting
Strength G for mid frequencies: 500-2000Hz
5dB louder than anechoic conditions, 15% weighting
Initial time delay to first reflection
Best at about 20ms, 10% weighting
Bass ratio: reverb time for 125-250Hz compared to 500-1000Hz
approx 1.2, 10% weighting

Table 2: Beranek's six parameters for acoustic excellence

In addition, Beranek suggested that there should be very little high frequency absorption in the hall, to aid brilliance in the sound, that the most important lateral reflections occur between 35°-75° from the front centre, and that any overhead reflections will add texture, especially when lateral reflections are low.

4.2 Early Reflections

By considering these parameters when composing or mixing for height, it is possible to make some important decisions that will influence the experience of a listener. If we wish to add the height dimension to a stereo or 5.1 recording, we need to consider the material being produced. If there is some reverberation already added to the recording, either from the performance venue or from electronic manipulation in a studio, it would not be desirable to add more reverberation from height channels. Since a listener in a venue is primarily hearing early reflections from above, we could process the recording through an 'early reflection generator' in post-production for reproduction through the height channels. By careful selection of the initial time delay and early reflection patterns, with no additional reverberation, a very realistic impression of height is achieved. Convolution using impulse responses recorded from vertical directions in different venues would also deliver realistic height impressions.

4.3Multiple Processors

When manipulating electronic sources for reproduction through height loudspeakers, the principles applied to horizontal reproduction would also apply. For example, realistic early reflection patterns and reverberation can be created electronically using multiple stereo processors, one for each of front, surround and height, or even a mono unit for each loudspeaker position. Multiple units would allow different parameters for initial delay, early reflections and reverb time to be set for each dimension, and also allow greater de-correlation of the reflections, which would decrease IACC. No matter where in space the source may be positioned, the reflection patterns and reverb characteristics could track position and motion adding realistic spatial enhancement.

5 Three Dimensional Soundfield: Recording and Reproduction

Recording natural sounds or performances for reproduction with height requires careful selection of microphones and their placement. There are numerous microphone techniques devised and tested for recording aural events for transmission by stereo or multi-channel formats, described in books, magazines, on websites and in conference proceedings, for example, Holman [13]. Depending on the taste of the producer and recording engineer, some incorporate coincident positioning of microphone capsules while others prefer spaced arrangements. For multi-channel recording incorporating surround information, there are additional microphones added to the stereo arrangements to capture spatial ambient information from the performance environment.

5.1 Critical Distance

A significant factor in venue recordings is the critical distance from the source, the distance at which the direct source amplitude matches the reverberant amplitude. Beyond this distance, any ability to specifically locate a sound source becomes virtually impossible, due to non-directional reverberation overwhelming the direct source information. Consequently, most multi-channel music recordings require some microphones to be relatively close to the source for reproduction through the front channels, allowing a listener to locate sources in the reproduced soundfield. Microphones positioned to record spatial envelopment are placed further from the source, carefully positioned to capture the characteristics of the environment for reproduction through the surround loudspeakers. However, most current techniques are predicated on capturing the horizontal soundfield, to be reproduced through a horizontal only loudspeaker system.

5.2 B-format

The Soundfield microphone [24] uses four microphone capsules arranged as a tetrahedron to capture the full three-dimensional soundfield at that point. After electronic manipulation in the unit's preamplifier, four audio signals are produced, equivalent to three bi-directional microphones and one omni-directional microphone. Known as the B-format signal, these four signals are referred to as W (omni-directional), X (front-rear horizontal), Y (left-right horizontal) and Z (up-down). After recording, the B-format signals can be sent to loudspeakers using the angles of azimuth and elevation to determine the corresponding amounts of X, Y, and Z. For example, if eight loudspeakers are positioned at the corners of a cube, each would receive equal amounts of W, X, Y, and Z but with different polarity depending on their location, eg front is +X, rear is –X, etc.. (For a more detailed consideration of the Soundfield microphone and its uses visit the Ambisonics website [2]).

5.3 Irregular Arrays

In this way, a soundfield can be reconstructed with great clarity, though the number of loudspeakers required is large, with the cube example considered to be the minimum necessary. Important to this paper's focus, the loudspeaker array should be regular, with each loudspeaker matched by another in the opposite position measured through the listening position. Irregular arrays like the ITU 5.1 standard are not well suited to B-format reconstruction, though one solution has been published as the Vienna coordinates, or G-format, described by audio researcher Angelo Farina [10] in a private email and described below in section 6.1. While the Soundfield microphone does measure spatial information accurately, it is not well liked by many recording engineers because of its susceptibility to critical distance compromises. If it is close enough to accurately record the direct signal, it is not in the ideal position to record the ambient field. Consequently, recording engineers are supplementing a Soundfield microphone with either close microphones to enhance the direct sound, or distant microphones to capture the ambient sound.

5.4 Height Information

The Soundfield microphone captures height information, though critical distance problems also apply in this dimension. While the Soundfield microphone simulates bi-directional responses, the off-axis null is only approximately -15dB. If the microphone is placed near the source, the loud direct sound will dominate the quiet reflected sounds from above. To overcome this imbalance, engineers are experimenting with spaced microphones placed in front of and above the performers to capture the height information, using cardioid or shotgun microphones pointed away from the source. For environmental recordings, the Soundfield microphone is excellent for capturing the three-dimensional soundfield, and is particularly successful for reconstructing motion perception of a moving sound source. A major difficulty highlighted by some users is a relatively loud noise floor, which interferes with quiet sounds.

6 Panning laws

Moving mono sources around the three-dimensional acoustic space created with height loudspeakers requires careful control of the amplitude and polarity sent to each loudspeaker. Investigations by Craven [8] into the psycho-acoustic relationship of sounds panned between adjacent loudspeakers pairs suggests that pair-wise amplitude panning should be replaced by panning laws which take account of psycho-acoustic principles. In essence, this requires that sounds panned to one loudspeaker should have a small amount of inverted polarity in the opposite channel. If applied to stereo loudspeakers, a sound panned fully to the right at 0dB would be approximately –14dB inverted in the left channel. Experienced listeners will immediately recognise this phenomenon as enhanced width, with the phantom image appearing further to the right than the right loudspeaker. When applied to multi-channel applications, particularly with B-format recordings, there is enhanced localization throughout the 3D soundfield.

6.1 Enhanced Height Perception

This principle may also be applied to B-format recordings replayed through an ITU 5.1 system. As suggested by Farina [10], rather than assigning the W, X, Y and Z signals to their 'true' loudspeaker locations, for example, using the 30° settings for the left front loudspeaker, Farina suggested that 45° settings be used to enhance the perceived width. Similarly, for the surround loudspeakers located at 110°, the W, X, Y and Z settings for 135° should be used. This also increases the 'opposing' loudspeaker characteristics inherent in true B-format reproduction. For the height loudspeakers located at azimuth 90°, elevation 60°, it was suggested by Farina to use the true B-format signal settings for the Z signal, but to add a small component of –Z to the five horizontal loudspeakers, for example, elevation down 30°, to enhance the perception of height. After experimental investigation by the author, it is agreed that using these settings, the soundfield is enhanced and localization improved for recordings made with the Soundfield microphone. While these psycho-acoustic panning laws are difficult to implement in hardware mixing consoles, researchers are developing panners as VST plug-ins for implementation in audio software programs, see the Ambisonics website [29].

6.2 Dual joystick spatializer

To further explore composing and mixing for height reproduction, the author has designed a dual joystick panner capable of sending a mono sound to any of six outputs. The configuration is matched to the author's preferred height loudspeaker positions of azimuth 90°, elevation 60°. It is currently an analogue device using six VCA's and control voltages produced by the two joysticks. The left joystick moves the sound around the four horizontal loudspeakers while the right joystick uses up-down movement to move the sound between the horizontal and the height loudspeakers, and the left-right movement to move between the two height loudspeakers. Using these two joysticks allows the producer or performer to play with the location of the sound throughout the spatial environment created by the six loudspeakers, and has proved to be fun to play with! Currently, the performance module has two sets of panners, four joysticks in total, and has proved useful for diffusion of stereo recordings. Future refinements could include using a Dumb Controller (after Fraiette [11]) to convert the control voltages to MIDI signals and interfacing it with software for more sophisticated movement, for example, multi-dimensional manipulation like matching the complex changes in the spectral patterns of a moving source, or adding more reverb as the height increases.

7 DVD-Audio and SACD delivery

For a reproduction system incorporating height loudspeakers to be commercially successful, there must be a readily available delivery format capable of carrying sufficient discrete channels of audio information to allow reconstruction of the three-dimensional soundfield. Two new delivery formats have been introduced in recent years, DVD-Audio and SACD, and both are capable of carrying six channels of high-resolution uncompressed audio. DVD-Audio is an extension of the DVD specification first made commercially successful with the DVD-Video application. While DVD-Video uses Dolby Digital AC-3 or DTS compression to deliver 5.1 audio, DVD-Audio uses virtually the entire bandwidth of the medium to carry multi-channel audio, based on PCM[3] recording technology. To deliver 24bit, 96kHz for six channels, DVD-Audio uses MLP[4] lossless compression [19], which is bit accurate on playback. Many music releases recorded specifically for this medium are using the six channels for 5.1 recordings, designed to be heard over home theatre systems. DVD-Audio also carries images, often song lyrics, band or recording information, as there is not sufficient bandwidth in the data transfer rates available for simultaneous video. Most DVD-Audio discs are dual layer, with the alternate layer carrying DVD-Video with AC-3 or DTS compressed 5.1 audio.

7.1 Universal Players

SACD is a competing format to DVD-Audio, using a different data recording technology, DSD or Direct Stream Digital, and is also capable of delivering six channels of high-resolution audio, equivalent to 24bit, 96kHz. SACD is also a dual layer disc, with the alternate layer carrying a stereo recording compatible with standard CD players. There is a limited capacity for images, but no video, and most SACD commercial releases are 5.1 recordings. While there is slow consumer adoption of either format and few releases available locally, the potential problem of two competing formats is being addressed by the latest generation of multi-format or universal players, capable of playing DVD-Audio and DVD-Video, SACD, standard CD and mp3 CD. Also, Minnetonka have released their Bronze authoring software for DVD-Audio at a realistic price of $US99, [18]. While this version does not allow access to the full potential of high-resolution audio using MLP, it is possible to burn discs with six channels of uncompressed 24bit, 48kHz resolution. As a consequence, it is possible for composers to use the DVD-Audio medium to deliver their compositions to listeners using up to six channels of standard-resolution audio.

8 Commercial releases

Several music recording companies are keen to explore height reproduction on their commercial releases, and have made important decisions concerning the format of their releases. After much discussion in trade forums, there is general agreement that music recordings do not require the centre channel or the LFE channel, both essential for cinema releases on DVD-Video. It is considered that a dedicated dialogue channel is not a critical requirement for music recordings, as the centre channel information is adequately reproduced as a phantom image created by equal left and right channels. Also, as most music listening will take place using full range loudspeakers, there is no need for the LFE channel to carry bass enhancement. Therefore, there are now two spare, high-resolution channels available for other information. These channels are being utilized for height information, with the method used for recording and reproduction varying between different companies.

MDG (www.mdg.de [17]) have released music with the description '2+2+2', referring to front, surround and height, and suggest that:

[3] Pulse Code Modulation

[4] Meridien Lossless Packing

> Three-dimensional portrayal of sound is thus an important step forward and in fact is an absolute prerequisite in the quest for natural music reproduction at home. (MDG, 2001, 8)

They have recorded music in a variety of venues using extra, dedicated microphones to capture height information. Their playback recommendations are for the front and surround loudspeakers to conform to the ITU 5.1 positions, and the height loudspeakers to be positioned directly above the front left and right loudspeakers, at azimuth 30°, elevation 30°. While this is not a large offset for the height loudspeakers, in auditioning MDG releases, the author's opinion is that there is an added level of spatial enhancement beyond the horizontal 5.1 system, particularly noticeable on recordings where the performance has been staged with risers for back sections of the orchestra and a choir behind and above the orchestra. And MDG believe:

> the listener.…enjoys an amazing sense of three-dimensional space and a logical, natural and stable sonic portrayal of instruments from almost all points within (and to some extent even outside) the area delineated by the speakers. (MDG, 2001, 9

Telarc [26] have released music with several different configurations for height reproduction. From information on their website, Telarc suggest that the height loudspeakers should be mounted on the side walls as close to the ceiling as possible, at the positions azimuth 90°, elevation approximately 45°. They also are uncommitted whether these height loudspeakers should be direct radiators, pointed at the listening position, or dipole radiators pointing front-rear. Auditioning several of their releases, it became clear they are not consistent with their production techniques. On one release, the height channels, formerly centre and LFE, were recorded as a stereo pair, with consistent imaging between the height loudspeakers. This recording reproduced beautifully on direct radiator height loudspeakers positioned as directed, adding a wonderful sense of spaciousness and envelopment to the recording. On another release, the centre channel carried substantially more reverberation than the LFE channel, and the combination of the pair did not appear to have a stable stereo image. This did not reproduce well on the height loudspeakers, with an imbalance due to clarity and direct/reverb differences.

9 Conclusion

The proposal for positioning full-range height loudspeakers on the frontal plane, either side of the median plane is supported by experimental investigation of acoustic perception. Architectural parameters that are indicators of high quality acoustic environments provide guidance for constructing three-dimensional acoustic spaces that will provide satisfying spatial envelopment and sound stage detail. Microphones and recording techniques provide excellent means to capture aural events for reconstruction through a sound system incorporating height loudspeakers. Investigations into psycho-acoustic based panning laws will assist in achieving realism in placing and moving sound sources throughout the acoustic space, and a joystick design provides a manual interface with the reproduction system. New delivery formats provide media for commercially releasing recordings and are already providing an opportunity to audition music recorded with height. Recordings and compositions by the author using height loudspeakers have created some enjoyable results for listeners when sounds are clearly located and/or are moving overhead. Also, different early reflection patterns and de-correlated reverberation from each height and surround loudspeaker provided the greatest perception of spatial envelopment. These results suggest that there would be significant interest among musicians and listeners for exploration of the height dimension in reproduction.

References

[1] Ando, Y., 1985, *Concert Hall Acoustics*, Springer-Verlag, Berlin.
[2] Ambisonics website, www.ambisonics.net
[3] Barbour, J. 2003, *Elevation Perception: Phantom Images in the Vertical Hemi-sphere*, AES 24th International Conference, Banff, Canada.
[4] Beranek, L., 1996, *Concert and Opera Halls: How They Sound*, Acoustical Society of America
[5] Blauert, J. 1969-70, *Sound Localization in the Median Plane*, Acustica, Volume 22, pp. 205-213
[6] Blauert, J. 1997, *Spatial Hearing*, MIT Press, ISBN 0-262-02413-6
[7] Chesky, http://www.chesky.com/, accessed 14 April 2004
[8] Craven, P., 2003, *Continuous Surround Panning for 5-speaker Reproduction*, AES 24th International Conference, Banff, Canada.
[9] Evans, M. 1998, *Obtaining Accurate Responses in Directional Listening Tests*, AES Preprint 4730, 104th Convention, Amsterdam
[10] Farina, A., private email 19/9/2003, farina@pcfarina.eng.unipr.it
[11] Fraietta, A., 2003, angelo_f@bigpond.net.au

[12] Griesinger, D. 2002, *Stereo and Surround Panning in Practice*, AES Convention Paper, 112th Convention, Munich.
[13] Holman, T, 2000, *5.1 Surround Sound, Up and Running*, Boston: Focal Press.

[14] ITU-R BS.1116-1, 1994-7, *Methods for the Subjective Assessment of Small Impairments in Audio Systems including Multichannel Sound Systems*
[15] King, R, 2000, "Orchestra Remains Up Front', *Mix Magazine*, August 2000, California: Primedia Inc.
[16] Martin, G., Woszczyk, W., Corey, J. and Quesnel, R. 1999, *Sound Source Localization in a Five-Channel Surround Sound Reproduction System*, AES Preprint 4994, 107th Convention, New York
[17] MDG, 2001, *Breakthrough in a New Dimension*, DVD-Audio liner notes, disc number 906 1069-5, www.mdg.de
[18] Minnetonka software, www.minnetonkaaudio.com
[19] MLP, www.meridian-audio/w_papers/mlp_jap_new.pdf
[20] Ratliffe, P. 1974, *Properties of Hearing Related to Quadraphonic Reproduction*, BBC Research Department Report 1974/38
[21] Roffler, S. and Butler, R. 1968, *Factors that Influence the Localization of Sound in the Vertical Plane*, Journal of the Acoustical Society of America, Volume 43, No. 6, pp. 1255-1259
[22] Roffler, S. and Butler, R. 1968, *Localization of Tonal Stimuli in the Vertical Plane*, Journal of the Acoustical Society of America, Volume 43, No. 6, pp. 1260-1266
[23] Schaffer, R.M., 1977, *The Tuning of the World*, Knopf, New York
[24] Soundfield microphone, www.soundfield.com
[25] Suokuisma, P. and Zacharov, N. 1998, *Multichannel Level Alignment, Part 1: Signals and Methods*, AES Preprint 4815, 105th Convention, San Francisco.
[26] Telarc, http://www.telarc.com/surround/
[27] Theile G. and Plange, G. 1977, *Localization of Lateral Phantom Sources*, Journal of the AES, Volume 25, No.4
[28] Truax, B., 1978, *Handbook for Acoustic Ecology*, ARC Publications, ASIN 0889850119
[29] VST plugin development detailed at www.ambisonics.net
[30] Wittek, H. and Theile, G. 2002, *The Recording Angle – based on Localization Curves*, AES Convention Paper, 112th Convention, Munich.

Wearable firmware: the Singing Jacket

Greg Schiemer

Mark Havryliv

Faculty of Creative Arts, University of Wollongong
Email: schiemer@uow.edu.au
mhavryliv@hotmail.com

Abstract

This paper discusses a design for a purpose-built interface that a wearer uses for controlling sound through movement. It presents work done so far on a mobile sound installation called the Singing Jacket and looks at ways in which the design can benefit from new developments in microcontroller technology. Several different microcontrollers are reviewed along with additional peripherals used for sensing movement and creating sound. These developments are presented against a backdrop of problems affecting musicians whose creative output is intimately bound up with changes that have been brought about by the ongoing development of electronic technology.

1 Electronic technology and disposable music

The last half of the twentieth century witnessed an unprecedented transformation in the way music is created, produced and distributed. Much of this change can be attributed to the revolution in electronics. The way musicians work has been transformed by changing technology and differs fundamentally from the way that musicians worked a century earlier. Electronic music composition involves new performance paradigms and frequently crosses disciplinary boundaries; its composers often perform their own music and a growing number are directly involved in the conception and implementation of new instrument designs and systems used for their creative work.

Electronic music composition poses a special problem. In the case of electronic composition, subsequent performance and reinterpretation is no longer possible when the technology associated with a particular work becomes superceded. This problem does not affect works composed using more conventional musical resources to quite the same extent.

Herein lies the paradox. Works by composers who have been proactive in the development of new forms of electronic music have become, by virtue of the accelerated rate of technological change, isolated from the very mainstream these works helped to bring about. Many of the early works of electronic music have become trapped within a period of technological development that has been superseded by more advanced technologies.

This phenomenon affects all composers of electronic music alike, both those engaging directly in building new electronic instruments as an integral part of the compositional process, and those working with new technology along more traditional lines.

Some examples of this phenomenon include music composed by Don Banks for ensembles that included keyboard instrument like the VCS-3 or Quasar M8 or theremin-based performances created by the Phillipa Cullen Dance ensemble in the early 1970s [7] or audiovisual environmental pieces works written by Martin Wesley-Smith for the Fairlight CMI [5].

There are two possible responses to this paradox:

1.1 The safe approach will be to focus all efforts on new musical activity using standard electronic musical instruments developed by instrument manufacturers. It would seem logical to argue that it is now possible to address this phenomenon by using instruments that have been engineered for the broadest cross section of musicians.

The problem is that by the time the product has been developed and matured to a point where it is useable by musicians in general, the technology on which it is based has been overtaken. The culprit is none other than the ongoing development of faster processors, new operating systems and, more importantly, new performance paradigms that grew out of earlier generation of technology.

1.2 The alternative response is more reckless: the musician develops generic technology and takes an active role in the ongoing development of non-standard instruments that appropriate consumer products manufactured for non-musical purposes. This approach does not sit easily with the musical constraints imposed by manufacturers of electronic musical instruments.

There are many precedents for musicians who take this approach. The new performance paradigms that continue to destabilise the technology on which standard electronic music is founded are a legacy of an earlier generation of musicians whose creative work embraced technological issues and laid the foundations of electronic music. The work of these musicians is documented elsewhere [18].

Proceedings ACMC 2004
ISSN: 1448-7780
http://acma.asn.au

2. Firmware design environments

Microcontroller technology has been responsible for much of the transformation that has taken place in music over the past decade and a half. Microcontrollers can be found in point-of-sales terminals, vending machines, domestic appliances, consumer goods, communications equipment, mobile phones and in engines and tyres of motor vehicles. Many are as powerful as the first computer music systems in the early 60s.

Microcontroller applications operate on firmware i.e. a dedicated software application that resides in non-volatile memory and runs automatically when power is applied. Firmware offers the advantage of a dedicated application that will run without the constraints of the disk operating system and the graphic user interface, or GUI.

While development of new composition systems has been focused on the GUI (the legacy of desktop computing) so far little has been done to create composition tools that focus on the interface between computer technology and the real world.

Notable exceptions to this include the MIDI Tool Box which uses a HC11 microcontroller [18][19], STEIM's SensorLab [21] and the Smart Controller [8] which uses field programmable gate array logic (GAL). While GAL is not the same as firmware in the strictest sense (i.e. conventional executable program instructions), it should be included in a summary of approaches attempting to address deficiencies in the user interface that standard composition environments in the desktop environment alone do not address.

2.1 The Singing Jacket

An example of a recent musical development using microcontroller technology is the *Singing Jacket* developed by Mark Havryliv.

The Singing Jacket is a mobile sound installation that is worn by a listener. It produces pre-composed sound that is influenced by the wearer's movement. The work creates a personal sound space within a larger acoustic environment.

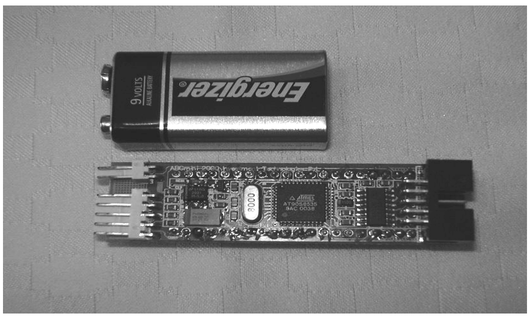


Figure 1.AVR AT90S8535 microcontroller

At the core of the Jacket is a small, power efficient microcontroller (AVR AT90S8535) which is programmed to produce sound that is influenced by a wearer's movement. The microcontroller is shown in Figure 1. The microcontroller has non-volatile memory on-chip which is used to store the program. The program determines the complexity of the interaction between the wearer's movement and the sound produced.

The first application for the Jacket generates sound by toggling (i.e. alternately switching between logic 1 and 0) adjacent output pins of the microcontroller. The pins are connected directly to a series of piezo-ceramic transducers to produce square waves.

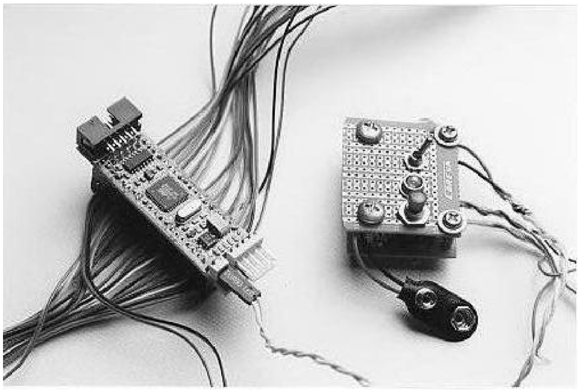


Figure 2. Cable harness connects microcontroller to transducers and switches (not shown)

Sound is output from 24 transducers. All frequencies produced are subdivisions of the processor clock, thereby producing ratios of the subharmonic series. Movement is detected using 8 mercury tilt switches that are sewn onto the fabric of the jacket arms.

The Singing Jacket is a departure from both the MIDI Tool Box and the Smart Controller in so far as it represents a recent paradigm shift that has taken place within the technology itself. Whereas the MTB uses a microcontroller designed for controlling motor vehicles, the Jacket uses AVR technology designed for battery-powered handheld appliances. Unlike the Smart Controller which is programmed using a graphic user interface the Jacket, like the MIDI Tool Box, is programmed entirely using low level text-based tools.

Further developments are planned for the Jacket. These will see the jacket become part of a development environment that is both wearable, self contained and able to connect to the outside world. Wearable implies a different level of mobility and power consumption than that associated with the first microcontrollers. It also implies a new kind of technology that is indistinguishable from everyday apparel [16].

3. Development issues - what will I wear tomorrow?

The kind of interactive works suggested by version 1 of the Jacket provided a point reference as we tried to identify the most appropriate choice of microcontroller interface.

In order to pursue these developments further, we asked two questions:
what technological trends are already evident ? and in which direction are they heading ?

We identified several trends:

- battery-powered laptops are now used for tasks previously performed using a mains-powered personal computer
- rechargeable handheld appliances are replacing mains-powered appliances
- hands-free devices are becoming a necessary extension of hand-held devices

The transition from mains powered to battery powered digital equipment is reflected in a number of microcontrollers we reviewed. We considered a range of processors designed for different purposes ranging from Motorola microcontrollers designed for automotive applications; to AVR microcontrollers used in a variety of household appliances; to the Ubicom microcontrollers used in mobile communications networks; to the ARM microcontrollers used in mobile handsets [1]. The increased memory capacity and reduced power supply requirements are shown in Figure 3; Figure 4 shows increasingly sophisticated on-chip peripherals and processor architecture; Harvard architecture – a design that uses an independent program bus and data bus – used in reduced instruction set computers (RISC) is introduced to conserve power. The increase in the processor speed is nevertheless accompanied by increased power consumption as shown in Figure 5.

Controller	Flash (bytes)	SRAM (bytes)	EEPROM (bytes)	ROM (bytes)	Pin-0uts	Supply Voltage (V)	Supply Current (mA)	Freq. Max (MHz)	No. of Power Saving Options
Motorola MC68HC908RF2	2k	128	-		12	1.8-3.3	4.3	4	2
Motorola MC68HC908JG16	16k	384	-		20	4-5.5	6.5	6	2
Motorola 68HC08KH12		384	-	12k	42	4-5.5	18	6	2
ATMega169 Butterfly	16k (+4Mbit data)	2k	512		32	2.7-5.5	10	16	5
ATMega16	16k	1k	512		32	2.7-5.5	12	16	6
Ubicom SX52BD/PQ	2k	262	2k		40	2.2-5.5	50	50	4
Motorola MC56F8323	32k (+8k data)	8k			46	2.25-2.75	60	60	4
ARM Processor AT91RM3400	128k	96k	128k		40	1.65-3.6	19	66	4

Figure 3.Memory capacities and electrical characteristics

Controller	Timers/PWM	Architecture	ADC	DAC	Coms
Motorola MC68HC908RF2	1 16-bit Timer	8-bit	-	-	
Motorola MC68HC908JG16	2 16-bit Timers	8-bit	1	-	SCI, USB
Motorola 68HC08KH12	2 16-bit Timer. 2 PWM Channels (16-bit)	8-bit	-	-	USB
ATMega169 (Butterfly)	2 8-bit Timers, 1 16-bit Timer. 4 PWM Channels (2 8-bit & 2 16-bit)	8-bit Harvard	1	-	USART, USI, SPI

ATMega16	**2 8-bit Timers, 1 16-bit Timer. 2 PWM Channels (8 & 16-bit)**	**8-bit Harvard**	**1**	**-**	**USART, SPI**
Ubicom SX52BD/PQ	**2 16-bit Timers, 1 8-bit Timer. 2 PWM Channels (16-bit)**	**Harvard (8-bit data, 12-bit program)**	**1**	**-**	**Programmable Serial Interface controlled by Firmware Libraries**
Motorola MC56F8323	**8 16-bit Timers, 6 PWM Channels**	**32-bit**	**2**	**-**	**FlexCAN, 2 SCI, 2 SPI**
ARM Processor AT91RM3400	**16-bit System Timer, 2 16-bit Timers 2 PWM Channels (16-bit)**	**32-bit Harvard**	**-**	**-**	**4 USART, USB 2.0, MMX, 3 Synchronous Serial Controllers (32 bit), ISO7816 Smart Card, RS485, Modem Control, Infrared, SPI**

Figure 4.Available on-chip peripherals.

3.1 New design

The preliminary prototype of the new jacket uses an ATMega16 microcontroller running at 16MHz [3]. The increased speed allows pulse width modulated (PWM) sine waves and other more complex wave forms to be produced. It also allows different waveforms to be independently produced and distributed through different speakers. Multiplexing has increased the number of switches available. The ATMega16 also offers larger storage for applications associated with wave form manipulation and generation.

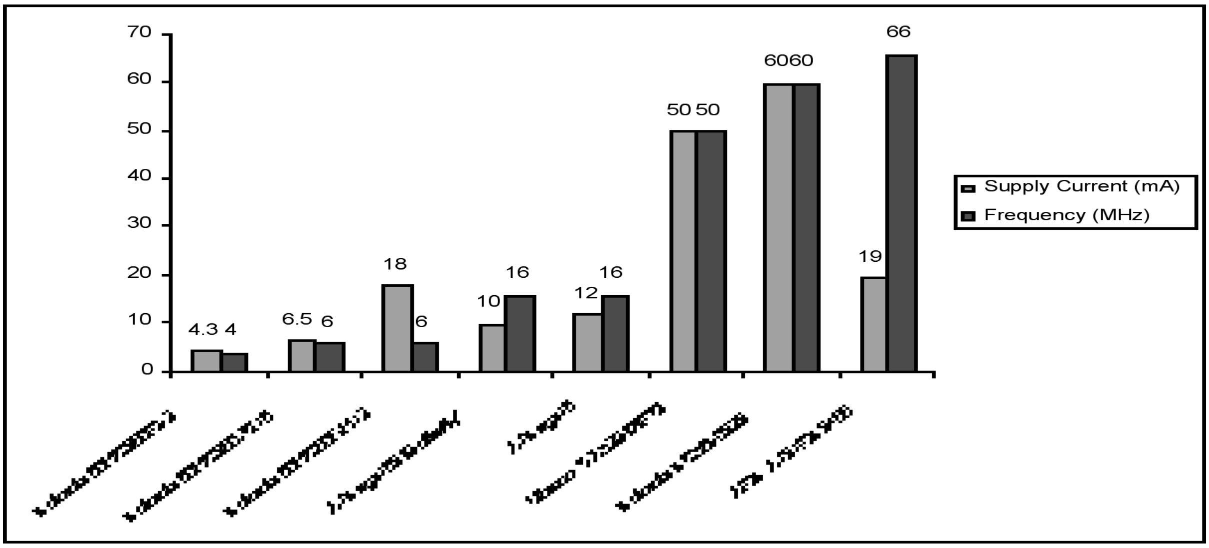


Figure 5.Current consumption versus frequency

3.2 Peripherals and communications

An important consideration for the new design is its user interface. Its communications capabilities must not compromise its mobility. Two microcontrollers are now being investigated.

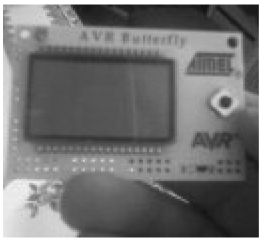


Figure 6.The AVR Butterfly chip

The AVR ATMega169, presented in a credit-card size badge called the AVR Butterfly, is powered by a 3-volt lithium battery [2]. It offers an assortment of peripherals ranging from LCD display, joystick, speaker, temperature sensor and light sensor. It also has USB capabilities, 4 Mbit of data flash and an RS232 level converter.

The Ubicom chip introduces the concept of virtual peripherals where software libraries provide the option of an SCI, SPI, Can bus, USB or Ethernet port. A new generation of this chip together with additional hardware allows implementation of 802.11 wireless LAN [20].

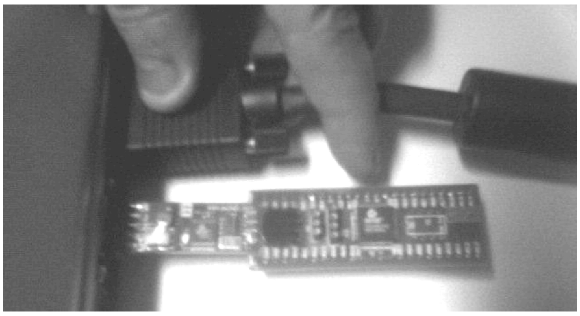


Figure 7. Ubicom SX56DB/PQ module with SX programming key plugged into RS-232 interface

3.3 E-field sensing

The Motorola e-field sensing chip (MC33794) offers new sensing capability. Though it was developed to monitor passenger safety in automotive air-bag controller design, in the context of the jacket it will provide a way to interpret the wearer's environment.

E-field sensing is closely related to one of the earliest electronic instruments, the Theremin. This chip originated in research from MIT Media Labs [15]. We used an MC33794 Evaluation Module based on a Motorola 68HC908QY4 microcontroller to test applications for controlling movement based on work begun by Phillippa Cullen in 1970.

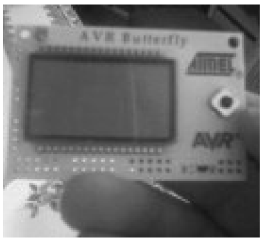


Figure 8 E-field evaluation module; MC33794 chip (top) reads up to 10 antennae.

A theremin signal traditionally monitors position of the body. One of our objectives was to derive velocity and acceleration as an alternative to using mercury tilt switches used in the jacket. This proved to be problematic due to limitations of the 68HC908QY4 used in this evaluation module [9]. An evaluation module based on Motorola's MC56F8300 offers a more sophisticated processor with an appropriate communications interface [10]. This approach to sensing will be pursued further once the MC56F8300 becomes available.

Figure 9.The Singing Jacket

3.4 Sound chips

The sound of the jacket can be enhanced using additional chips. Two chips have been reviewed: the Dream synthesiser and the Motorola Symphony chip. They communicate with a microcontroller via the SPI port.

The Dream Synthesiser (ATSAM9708) is a 128-voice wavetable synthesis chip [4]. Standard firmware includes equalisation, surround and an MPEG 2 audio decoder. The Dream chip includes the Roland GS Sound Set. Its MIDI ports communicate with a microcontroller via the SCI port. The chip's 3.3V supply voltage features a power-saving mode. It is well suited for mobile or wearable applications.

The Motorola Symphony chip (DSP56367) is a 24-bit Audio Digital Signal Processor [11]. The chip's DSP capabilities support current digital audio standards (Dolby, DVD Audio, 3D virtual surround). It operates with supply voltages between 1.8-3.3V, has two sleep modes and supports Enhanced Serial Audio Interface (ESAI) including I^2S, SPDIF, Sony and AC97.

The processing capabilities of the Symphony chip will be used to perform speaker compensation on the piezo-transducers. A precedent for this is the CSIRO's A4 chip frequency domain processor which convolve audio signal with the inverse of the frequency response [6].

4 Conclusion

The development of the jacket is seen as a possible prototype for an interactive musical instrument that will be a hands-free accessory for a mobile phone handset [17]. We think it is important for musicians to anticipate such developments and not lose an opportunity to influence directions for the future.

5 Acknowledgements

This project is indirectly related to a 3-year project entitled “Pocket Gamelan tuning musical applications for wireless internet” funded by the Australian Research Council.

6 References

[1] Atmel (2003a) "ARM7TDMI-Based Microcontroller AT91RM3400" San Jose, CA at: www.atmel.com/dyn/resources/ prod_documents/doc1790.pdf

[2] Atmel (2003b) "AVR Butterfly Evaluation Kit User Guide" San Jose, CA at: www.atmel.com/dyn/resources/ prod_documents/doc4271.pdf

[3] Atmel (2003c) "8-bit AVR Microcontroller with 16K bytes In-System Programmable Flash - ATMega16 ATMega16L Summary" San Jose, CA at: www.atmel.com/dyn/resources/ prod_documents/2466S.pdf

[4] Atmel (2004) "Sound Synthesis ATSAM9708 128-voice Integrated Sound Synthesiser" San Jose, CA at:www.atmel.com/dyn/resources/ prod_documents/doc1772.pdf

[5] Burt, W. (1991) Australian Experimental Music 1963-1990 Leonardo Music Journal, 1 [1] MIT Press pp. 5-10

[6] CSIRO Division of Radiophysics (1993) “The A4D2 Development System User Manual” Epping

[7] Ellyard, C. and MacLeod, I. (1975) “The Computer Arts – Reflections on “Australia 75” ” Proceedings of the Digital Equipment Computer Users Society, Melb. pp. 1231-1234 August.

[8] Fraietta, A. (2003) The Smart Controller/ shifting performance boundaries” in: Proceedings of ICMC 2003, Singapore pp. 53-56

[9] Motorola Semiconductors (2003a) "KIT33794DWBEVM Electric Field Imaging Device" Analogue Products Division Colorado at: www.jandspromotions.com/efield2003/ datasheet.pdf

[10] Motorola Semiconductors (2003b) "56F8300 Demonstration Board User Manual" Colorado at: e-www.motorola.com/files/dsp/doc/ data_sheet/MC56F8323.pdf

[11] Motorola (2004a) "Advance Information DSP56366/D Rev. 1.6, 01/2004 24-Bit Audio Digital Signal Processor" Colorado at: e-www.motorola.com/files/dsp/doc/ data_sheet/DSP56366.pdf

[12] Motorola (2004b) "MC68HC908RF2/D Rev. 3 MC68HC908RF2 Data Sheet" Colorado at: e-www.motorola.com/files/microcontrollers/doc/data_sheet/MC68HC908RF2.pdf

[13] Motorola (2002a) "MC68HC908JG16/D Rev. 1 MC68HC908JG16 Data Sheet" Colorado at: e-www.motorola.com/files/microcontrollers/doc/data_sheet/MC68HC908JG16.pdf

[14] Motorola (1999) "MC68HC08KH12 advance information" Colorado at: e-www.motorola.com/files/microcontrollers/doc/data_sheet/MC68HC08KH12.pdf

[15] Paradiso, J. (2003) "Dual-Use Technologies for Electronic Music Controllers: A Personal Perspective" in: Proceedings of the 2003 Conference on New Interfaces for Musical Expression (NIME-03), Montreal, Canada NIME Responsive Environments Group

[16] Post, R. and Orth, M. (1997) “Smart Fabric, or Washable Computing” in: Digest of Papers from the First IEEE International Symposium on Wearable Computers, Cambridge, MA pp. 167-8

[17] Schiemer, G., Alves, B., Taylor, S. and Havryliv, M. (2003) “Pocket Gamelan: building the instrumentarium for an extended harmonic universe” in: Proceedings of ICMC 2003, Singapore pp. 264-267

[18] Schiemer, G. (1999a) “MIDI Tool Box: An interactive system for music composition” (thesis with audio CD and CD-ROM) PhD (Electronics) School of MPCE, Macquarie University

[19] Schiemer, G. (1999b) “Improvising Machines: Spectral Dance and Token Objects” Leonardo Music Journal 9 MIT Press, pp. 107-114

[20] Ubicom (2002b) "SX48BD/SX52BD Configurable Communications Controllers with EE/Flash Program Memory, In-System Programming Capability, and On-Chip Debug" Mountain View, CA at: www.ubicom.com/pdfs/products/ sx/processor/SX-DDS-SX4852BD-14.pdf

[21] Waisvisz, M. (1993) Twenty-five years of STEIM, on overture in the program de zoetvooisde BLIKSEM (the sweet voice of lightning), den Haag

Developing a software meta-instrument for live electronic dance music

Rene Wooller

Creative Industries: Music, Queensland University of Technology
E-mail: r.wooller@qut.edu.au

Abstract

This paper discusses a two year long project to develop software that overcomes some of the limitations of electronic dance music when placed in a dynamic musical context. An analysis of pertinent aspects of this process, namely a musicological study of electronic dance music, development of theoretical principals and a subsequent rendering of the results into software code is presented. The success of the technology is reviewed and the future directions and possibilities of unimplemented theories are examined. The development and research methodology is briefly considered for potential usefulness in similar studies.

1 Introduction

In most electronic dance music venues, the music is delivered through relatively static mediums such as vinyl, CDs, or Computer sequencers. Although live musical manipulations can and do occur, most of the musicality is preconceived through studio production. LEMu is a means for more spontaneous and adaptable deliveries (Fig 1) of this music. The term "meta-instrument" is used to describe an instrument with some capacity for generalised control over a multi-voiced musical system.

2 Methodology

A project plan was designed, outlining four stages, each aiming to build a new version of the program with slightly different goals (see Fig 2 below). Stages one two and three were completed successfully.

An *iterative* or *rapid prototyping* type methodology (McConnell, 1998) was developed (Fig 3). Perhaps due to the small scale of the project and an initial lack of technical knowledge and experience, designs were generally only useful when closely linked with implementation. Often when new limitations, features and patterns were learnt, the designs would require modification. User requirements development and testing also seemed most useful when conducted in parallel with implementation.

Musicological research was an inspiration for designing many useful features, algorithms and theories (Wooller, 2003). Many of these are unimplemented, indicating there is much more still to be explored.

Future plans for similar projects could allow more time for technical research and implementation. More focus on technical research would enable the researcher to develop better software development skills. Designs would then be more informed and implementation much easier and effective. If implementation was an ongoing activity where the technical research could be applied, better results may be produced faster.

Figure 1: Live improvisation to video art

Research directions were governed by project requirements and plans (above). Information gathering was conducted in areas of programming technology, music software, musicology, and user-testing.

3 Programming Technology

Research into programming technology was conducted in parallel to all coding work and was also pursued through formal education and conference attendance. Various activities involved reading books, scanning and participating in developer lists and web forums, investigating source code and documentation of various APIs (Application Programming Interface or code library/toolkit), communication with experienced developers as well as attending lectures and tutorials.

Proceedings ACMC 2004
ISSN: 1448-7780
http://acma.asn.au

Stage One	Stage Two	Stage Three	Stage Four
Build Focuses ▪ Simplicity ▪ User Friendliness ▪ Simple, commercial EDM aesthetic	**Build Focuses** ▪ Retain Simplicity ▪ Extend functionality ▪ More complex EDM aesthetic	**Build Focuses** ▪ Extend functionality ▪ Retain Simplicity	**Build Focuses** ▪ Intelligence ▪ Retain Simplicity ▪ Extend functionality
Research Priorities ▪ Context (similar software etc.) ▪ Usability (REV festival)	**Research Priorities** ▪ Musicology ▪ Technology	**Research Priorities** ▪ Technology ▪ Usability ▪ Musicology	**Research Priorities** ▪ Usability ▪ Experiments
Build Capabilities ▪ Creates simple EDM ▪ Can be used by most people ▪ Effective on the short term	**Build Capabilities** ▪ Creates interesting EDM ▪ Can be used by most people ▪ Effective and engaging	**Build Capabilities** ▪ Creates interesting EDM with enough variation for a whole set ▪ Can be used by most people ▪ Can perform a range of customisable EDM manoeuvres	**Build Capabilities** ▪ Creates interesting EDM with enough variation to last an entire night (6-8 hrs) ▪ Can be used by most people ▪ Reacts to input in an intelligent way ▪ Is programmable to a small extent
Drawbacks ▪ Becomes boring quickly ▪ Not enough variation available to play more than ten minutes	**Drawbacks** ▪ Not enough variation to play a whole set without the music sounding too similar ▪ It may take a while for a user to	**Drawbacks** ▪ Takes a skilled user to perform for more than 5-10	**Drawbacks** ▪ Must be programmed well

Figure 2: Outline of Project Plan

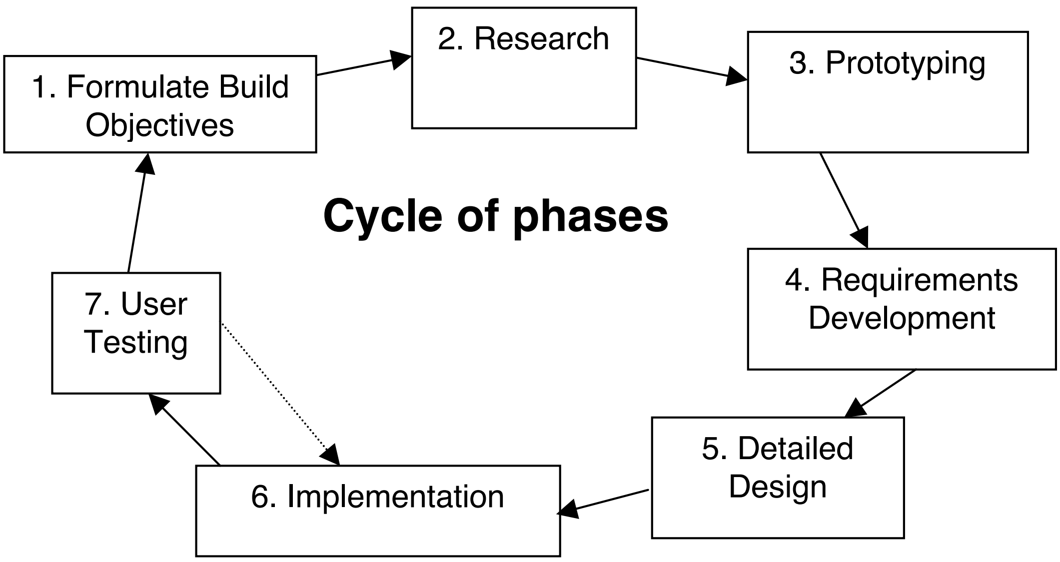


Figure 3: Outline of development methodology for a single stage

The reading and formal study of software development was mainly pursued during the first stages of the course and was useful for developing a foundational knowledge of programming practices. The books used (Lambert and Osbourne, 2002; Niemeyer and Knudsen, 2000) were useful only up to a point which was reached fairly early on. Perhaps more advanced books with general content such as design patterns (Grand, 1998) and pitfalls (Daconta et al., 2000) could have also been consulted at this stage. Familiarity with basic system structure and programming confidence was the result of sound knowledge of common algorithms, data types, interfaces and other programming foundations taught in the software development courses. Other aspects such as predicate calculus remain unused. Knowledge of important design principals were not imparted from these subjects, but were obtained instead through in-depth discussions with experienced music software programmers such as Ross Bencina (Bencina, 2003) at the Australasian Computer Music Conference and follow-up research online. This information became understood the most clearly during implementation.

As the first point of reference for understanding the tools being used, API documentation is instrumental. If it is comprehensive and easy to understand, documentation is usually the last point of reference as well. jMusic (Sorenson and Brown, 2004) and Java (Sun, 2004a) were particularly well documented. Despite this, the complexity of the Java API is such that other means are often necessary. Midishare documentation (Grame, 2004) was generally mediocre and communications with head developer (Letz, 2003) were very informative.

Investigating source code (if available) was found to be useful for obtaining a deep level of understanding of the underlying mechanics behind the library, although initially much more confusing than a book or tutorial. Also, by observing the language and structures implemented by experts (such as the java developers), certain programming techniques were learnt and the ability to analyze, infer logic and comprehend external software was developed.

An extremely useful resource was the java developer forums, found on the java website (Sun, 2004b). Many issues, especially concerning Graphical User Interface (GUI) development, have already been overcome by other developers, and the resulting solutions and discussions are available. Problems posted to these forums were discussed within hours.

4 Music Software

Other software that (partially) shared similar goals were noted (in order of similarity): Jam2Jam (Brown et al., 2004), KARMA (Kay, 2004), Reason matrix sequencers (Propellerheads, 2003), KOAN (SSEYO, 2004), Live (Ableton, 2004), Audiomulch (Bencina, 2004), AutoGam (Bachmann, 2001), Max/MSP (Cycling'74, 2004), Pure Data (Puckette, 1996) and Algorithmic Composer (Fraietta, 2001). Various GUI designs (since abandoned) were inspired from some programs. Since none of the software shared exactly the same concerns, either being focused more on education (Jam2Jam), multimedia (KOAN), performance modelling (KARMA), performance possibilities of the music studio (Abelton Live), audio/sound exploration (Audiomulch) or interaction programmability (Max/MSP, PD, etc), it was found to be more useful to develop fresh research, theories and software. The LEMu primarily differ from the main competitors Jam2Jam, KARMA and KOAN in that it is almost completely deterministic, rather than probabilistic.

5 Musicology

Musicological research focused on Electronic Dance Music (EDM) played in night-clubs. It was decided that research in this area should be limited to a narrower genre to enable greater depth. Drum and Bass was selected partly due to its rhythmic complexity. As well as this, other genres such as house, hardcore, progressive house and trance had already been examined by the author previously.

Prior experience with the limitations and advantages of various algorithms informed the way research was conducted. Analysis methods were prioritised if they might lead to algorithms that were more deterministic and transformational than probabilistically generative and/or recombinant. Deterministic algorithms are reliable and have the potential for a fine level of control suitable to musical instruments. Transformational algorithms seemed to reflect certain aspects of the improvisatory process more than the probabilistic and also seemed to be less widely used than recombinant algorithms. Although the discovery of deterministic compositional processes was prioritised, important stochastic ones were not overlooked and generalities involving probability were also developed. In this way, musicological analysis was mostly concerned with compositional techniques and systems rather than motific documentation. Research activities included rhythmic notation, reconstruction and spectral examination of a small selection of music (The Drum and Bass Arena; Grooverider Live; Drum and Bass Asassins), casual listening of a wider selection (see discogography) personal communication with producers (Joslin, 2003)and composers (Vella, 2003) and minimal text based research.

A diverse array of observed compositional relationships and processes led to definition of trends and development of relevant theories. The most important of these are outlined below.

The most prominent compositional technique in EDM is the layering of parts that cycle. Cycle lengths (in number of beats) are generally powers of two above two (4, 8, 16, 32, 64, 128 etc). This was observed throughout the whole selection of material. Two categories of cycles were determined: root cycles, and modification (mod) cycles (Wooller, 2003). A mod cycle is a cycle of change or modification. An example would be a filter sweep that occurs slowly over 16 bars, or a fill that happens on every eighth bar. A root cycle is simply a repeated pattern. It may or may not be transformed over time within a mod cycle. The foundational parts generally have shorter cycles. It is common for the root cycle of the drums to be one bar long and the mod cycle to be two bars long. This differs from the bass, where the root cycle is often two bars long with a mod cycle of eight or sixteen bars. This was observed in many tracks including: most of the album "Clockwork" (Stakka and Skynet, 2002), *Sea of Chaos* by Ant Miles, *Miami* by Bad Company, *Out of Control* by Spirit, *Yvon is On* by Total Science, *Sex Jive* by Alex Reece & Utah Jazz (from the Drum and Bass Arena, 2001) and *Watching Windows* (Roni Size/Reprazent, 1997). More extraneous parts such as vocal samples, melodic riffs, chord stabs, and sound objects tend towards longer cycles still. Examples of this can be found in *All Aboard* by Dilinja, *Survival* by Moving Fusion, *Reflections* by Ram Trilogy (from the Drum and Bass Arena, 2001) and *Watching Windows* (Roni Size/Reprazent, 1997).

Harmonic elements are often related to the western tonal system – this is evident in most of the music studied. There also exist many pitched parts that are non-harmonic and non-tonal, for example *All Aboard* by Dilinja, *Miami* by Bad Company, *Balderdash* by Marcus Intalex & ST Files, *Yvon is On* by Total Science (from the Drum and Bass Arena, 2001). The patterns of continuous elements such as pitch bend and filter frequency seem to be influenced by tools such as envelopes and LFOs (Wooller, 2003). For example, it is easier and quicker to draw a straight line with a graphical envelope than it is to draw tonal intervals. Continuos envelopes are observable in many tracks, including *Essence* by Ant Miles, *Gateman* by Digital, *Sex Jive* by Alex Reece & Utah Jazz (from the Drum and Bass Arena, 2001), although they are more obvious in type of commercial dance music found within "Dance Nation 3". An interesting characteristic of Drum and Bass is the pitch shifting of the bass which (if it is present) is almost always downwards. This trend is observable in *Silver Blade* by Dilinja (The Prototype Years, 1997), *The Flipside: Aphrodite remix* by Moloko (1998), *Gangsta Gangsta: Aphrodite remix* by NWA (1998) and "Grooverider live" (2002). Seemingly unplanned discord exists in many mixes and some tracks, suggesting that rhythmic elements are of greater compositional value in Club Drum and Bass. This was the case with *Need You* by Dylan & Ink when mixed by Grooverider on "Harder they Come"; *Miami* by Bad Company (Drum and Bass Arena, 2002); *Champion Sound - Bad Company Remix* by QProject (21st Century Drum and Bass, 2000); "Grooverider live"(2002) and many others.

Some techniques and trends relating to rhythmic creation and modification were determined from observation and inference. Two over three and four over three rhythmic patterns were extremely prominent throughout Drum and Bass. Elements that were based on or influenced by these patterns included the kick, snare, bass, sampled sounds, delay settings and the rate of various LFOs. These patterns are audible in many tracks within albums such as "Clockwork" (Stakka&Skynet, 2002), "Burningn'n Tree" (Squarepusher, 1997), "The Drum and Bass Arena" (2001) and to a lesser extent in "Drum and Bass Asassins" (2000). During most sections of music this rhythmic tendency seemed constrained by the necessities of a four/four rhythmic structure. For example, two over three or four over three patterns would generally be cycled within one (four beat) bar. However, more intense and less constrained sections displayed similar polyrhythms that extended beyond one bar and/or patterns that were transformed to follow the two/three over four feel. This technique is used in *Deep Inside* by Codename John, *Silver Blade* by Dilinja (both from The Prototype Years, 1997), *Watching Windows* by Roni Size (RoniSize/Reprazent 1997) and *The Sound* by Tronik100 (The Harder They Come, 2002). Theories of transformational processes were developed through a kind of lexical analysis of two patterns within the same part. One pattern was modified incrementally to make it the same as the other pattern. The number of modifications determined the closeness of their apparent relationship. The more important aspect of this analysis was the definition of possible transformational processes relating to rhythm (Wooller, 2003) including: selective repetition (likely the result of cutting and pasting technology), phase offset and rate change. To what extent these processes are used by composers can be partially ascertained by their effectiveness as software algorithms.

Part interactivity was also found to influence the composition of rhythmic patterns. Within Drum and Bass it is unlikely for the kick and snare to occupy the same beat, effecting a kind of binary opposition between them (perhaps reflective of certain hand drumming techniques). It was observed in the following tracks that the balance of the number of kicks and snares followed a general pattern that was roughly equal over long enough periods of time: *Hush Hush* by Shimon; *Mystic* by Calibre; *Just a Vision* by Solid State (all from Drum and Bass Arena, 2002); *Brown Paper Bag* by Roni Size (New Forms, 1997). This suggests that compositional connections between the kick and snare (or sounds that fulfil similar roles) and patterns of their general application could be modelled by algorithms. Relationships between the bass and the kick were also observed. Within "four on the floor" styles the bass sometimes seems to avoid the kick drum. For example, a very common pattern in commercial house is to place the kick on every beat and the bass on every offbeat. Many examples of this can be heard within "Dance Nation 3" (2003). Conversely, in Drum and Bass the kick will often be reinforced by the bass. This is especially apparent if the kick lacks bottom end - frequently the case, with drum samples often being pitched up and sped up. Sample heavy Drum and Bass displays this technique often: *All Aboard* by Dilinja, *Sea of Chaos* by Ant Miles, *Dubwise* by Accidental Heroes and *Two Faced* by Teebee (Drum and Bass Arena, 2002).

Tempo was also found to influence the cognition of rhythm, especially relating to the perception of streams (Vella, 2003). Complex rhythms were much easier to divide into streams at faster tempos like those found in Drum and Bass which are approximately 160-200 BPM.

A theory for the perceived meaning and roles of the various beats and rhythmic elements was formulated as a step towards logical definition of the subjective affect of Drum and Bass. This is briefly discussed below and in greater depth in the author's previous paper on Club Drum and Bass (Wooller, 2003).

6 Functionality and Algorithms

The research and theories discussed above inspired the conceptual development of music knowledge representations, a music software framework, music algorithms and various features of the application. Patterns of composition that occurred more widely and frequently within the research described above were given priority as developmental input. Practical implementation considerations also influenced these priorities.

The musicological framework has informed the software architecture design. LEMu is based on parts that cycle and are layered. All parts inherit features through a common ancestor but incorporate limitations and functionality that optimise the part towards its particular compositional role. They all have data storage and retrieval abilities and contain a "root pattern" which loops continuously. Although the length of the root pattern is fixed at thirty-two beats (eight bars), the perceived length can be controlled by the "scope". For example, if the scope is two beats long, only the first half of the first bar will be looped. This structure (rather than having a variable length root pattern) was chosen partly because it provided an easy way to mimic common compositional techniques in sample based music, for example: looping half a sample for seven bars and then playing the whole sample on the eighth bar. The "root pattern" relates the concept of "root cycles" discussed above. "Modification cycles" have also been implemented - in the form of automation. The length of automation is one hundred and twenty eight beats (thirty-two bars) and the perceived length is controlled by the "automation scope". Cycles of compositional change (mod cycles) can be created by automating various functions that apply pattern transformations.

Transformational functions that are common to each part include: repeat, phase offset, rate, velocity, quantise and note-length. They have all been identified through research as important compositional transformations. For example, repeat is operated with one notched slider. When the slider is in the default central position nothing happens. When it is moved to the left it repeats a selection from the start of the bar and when it is moved to the right it repeats a selection from the end of the bar. The position and length of the selection is influenced by how far the slider is moved. A detailed explanation of these functions is included in the feature guide (Wooller, 2004a) and the functional manual, which is accessible from within the program itself (Wooller, 2004b).

Tonal parts (currently the arpeggiator and bass) share limitations and features based on western tonality. All harmonic transformations such as pitch scaling and transposition, are applied to "scale degrees" or "pitch classes" and octaves, rather than specific pitches. Before the tonal note is outputted, it is rendered into a pitch by applying its scale degree and octave to the currently selected scale. Available scales are: all modes (Aeolian etc), harmonic minor, pentatonic, chromatic, whole tone and whole tone half tone. Melodic patterns can be inverted by reducing the pitch scaling factor below zero. The arpeggiator is non-traditional in that it has controls for a Low Frequency Oscillator that dictates a pitch contour of the arpeggio. The notes can be restrained into chordal or intervals notes of various inversions.

Each part has specialised generative algorithms that are used to create root patterns. The algorithm for the bass part follows the more commercial style of dance music – a note on every offbeat. This pattern is easily transformed into a four over three rhythm (Drum and Bass/Breaks style) bass line by increasing the rate by one notch. The algorithms for the kick, snare and high-hats operate stochastically within a specific range of possibilities gathered from rhythmic analysis. These algorithms are based on the concept of each beat having a distinct role within the pattern that it is part of, for example: primary, complementary, leading and secondary beats. If they are generated within certain boundaries of rhythmic placement and velocity, different roles are recognisable. For example, the primary kick beat has a high probability that it will occur on the one, but sometimes it will occur a sixteenth or an eighth later.

MIDI synchronisation is a recent addition which enables other music generators to play in time with LEMu. This greatly increases the ability to jam and adapt to external musical environments. Features that remain to be implemented are: beat role pattern recognition, part interactivity, control of musical states, pattern visualisations, support for sound banks, MIDI recording, and patchable components.

An extended implementation of the "beat roles" concept outlined above would be to add some attributes to the "Note" object that represent the importance and role (primary, complementary etc) of the note. The extended representation would make it possible to design pattern transformations that are far more complex and meaningful.

Part interactivity is an unimplemented feature that could utilise this note information effectively. For example, rather than simply following or avoiding the notes of another part, it could decide to follow only the important ones of a certain type. A number of complex and worthwhile interactive processes could be defined using this architecture.

7 Usability Developments

Within the field of cognitive ergonomics there are four criteria for determining the usability of a system (Clemen, 2003):

1. **Learnability:** the amount of learning necessary to achieve tasks
2. **Ease of Use:** the efficiency and effectiveness with which one can achieve these tasks
3. **Flexibility:** the extent to which a system can adapt to new tasks and environment requirements
4. **Attitude:** the positive or negative attitude of the user towards the system

Through various stages, LEMu development addressed all of these criteria. Features were strategically abandoned or re-worked to provide a balance between these elements. Over seventy features were managed in a spreadsheet where they were allocated relative levels of "importance" and "implementation time" (an estimation of the time it would take to implement). A "priority factor" was obtained by multiplying importance by time, enabling features to be organised into a hit list. This encouraged a consistent level of progress to take place. Perhaps if this system for feature tracking had been implemented at an earlier stage it would have reduced the chance of features being overlooked or unimplemented.

Major steps toward increasing the learnability of the system were made through comprehensive inclusion of many simple details such as tool tip texts, help files, sensible readouts and interface styles. The layout of the program was also considered, with components being organized from left to right in some kind of priority.

The current navigation system of the GUI (Fig 4) is fairly easy to learn; however it becomes increasingly cumbersome when more than five parts are used. A future version may organise part groups into tabs, rather than displaying a small panel for each part. The functions on the small panel are essentially useless, because the whole panel is activated on roll-over.

Other problems concerning ease of use were addressed throughout the project. Testing uncovered efficiency gains which were initially overlooked. Some examples of this include the ability to double click on items within explorer menus (such as the load pattern dialog) and various key short-cuts. Other aspects contributing towards effectiveness and efficiency were designed specifically: GUI navigation, MIDI input control and custom java components. jMusic source code was also made easier to use through solving various inaccuracies and incorporating unimplemented features within functions such as copy (definable phrase segments), cycle, elongate and quantise. Additions were also made to the jMusic library, including Midishare-jMusic interfaces, a Midishare sequencer, a graphical envelope editor, algorithms for removing rests, as well as additional note attributes such as degree, octave and scale.

The flexibility of functionality was developed to the extent of the task and environment requirements. LEMu is designed for: live performance and (as an aside) studio based composition. Various features such as MIDI output, MIDI controller input, remapping abilities and parameter controls allow a fair amount of diversity in the abilities of the program while remaining limited enough to be meaningful and accessible to the user.

Another aspect of flexibility for a software program is the ability to work under a range of software environments. A notable amount of development time was devoted to running LEMu at all the different sized screen resolutions above 600x480 and on various other machines running operating systems such as Mac OS X, Windows, and Linux. At a fundamental level, because the source code itself is available for modification, LEMu is extremely flexible to anyone who can program in java.

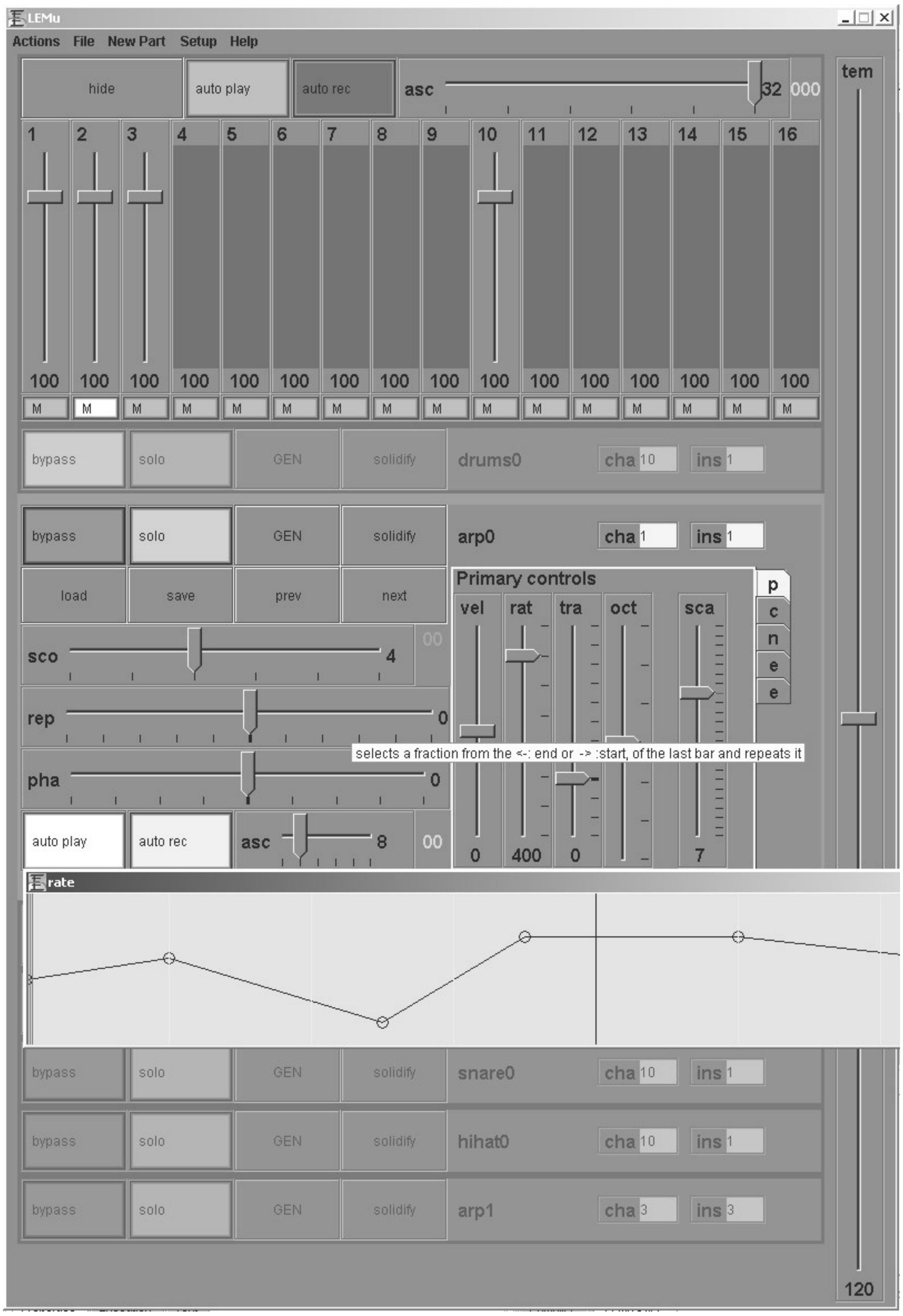


Figure 4: LEMu Graphical User Interface

User attitude seems to affect all of the other usability criteria to a notable extent. Because of this, some effort has gone towards making LEMu look at least semi-professional so that it does not suffer adversely from user prejudices. As well as this, the aesthetics of the interface have been designed to be relatively non-threatening and enjoyable to use.

8 Testing

Testing was conducted at a number of stages throughout the project. During stage one, the large public audience provided by the REV festival was an excellent opportunity to trial the simple version (Monro, 2002). This provided insights into ease of use and short-term learnability through observing unbiased manipulation of the program by a wide range of people. The relationship between GUI design and learnability was observed at first hand – people found the simple version much less threatening, seemingly due to the minimal amount of unknown functions that were immediately viewable. More attractive GUI designs seemed to increase the general learning capacity by encouraging enjoyment and confidence in the technology.

Throughout stage one and two, electronic music ensemble “The Cloud Kickers” was directed, using LEMu as the primary instrument. This provided a means to test performability, usability, long-term learnability and improvisation capabilities. Although a fair amount of successful material was developed and delivered (Cloudkickers, 2004), this ensemble would have been of much more use towards the end of the project, rather than the beginning. The version of LEMu that was primarily used through this period was quite old and badly designed and did not improve substantially over the year. Not a huge amount of skill development was attained by the LEMu users, mainly due to the poor usability inherent in that version. It was a useful project - however it seems that testing, implementation, and design should mostly occur in parallel, as they feed into each other.

Towards the end of the project, a solo experimental performance was executed with only moderate success (Barrett, 2003; Jaaniste, 2003). The problems were partly due to minimal preparation and technical timing problems that were later overcome. In retrospect, LEMu seems to be more effective in situations involving multiple musicians, whether they play another instrument or share control of LEMu. Single user operation is quite demanding, due to the number of parameters that necessitate control. The balance of flexibility with ease of use is similar to the problems of interactive installations as discussed by Clemen (2003), although LEMu is designed as a musical instrument rather than a public installation. As with conventional instruments, more rehearsal time is another effective way of overcoming the ease of use problem, and is evident when comparing the various musical endeavors.

During the last six months, a more formal testing process was initiated involving the development of a questionnaire and numerous installations on various machines. Answers from the question sheets were used to develop the feature tracking spreadsheet and judge priorities within it more accurately. A rapid development cycle would sometimes occur where testers gave instant feedback which was implemented directly, leading to more of the same. This was very effective for numerous smaller features and minor bugs, and a clear demonstration of testing and development working in parallel.

9 Conclusion

Through the research and development processes discussed, a software meta-instrument with documentation has been produced. Various effective ways of creating and manipulating different aspects of electronic dance music in real time have been explored in detail. From common musical transformation and generation techniques emerged useful software functionality. Through this process the project goals have been achieved despite minor problems. Many other theoretical functions and programming frameworks remain unimplemented. Their potential may be realised in future development. As an analysis and discussion of the undertaking it is hoped that this paper may be useful for similar projects.

Acknowledgments

Thanks goes to Andrew Brown for his much valued teachings over the last five years. All of the JMC posse. Tim Opie and Adam Kirby for their technical advice and philosophies. Andrew Sorenson for his hand-me-down computers and insights. Nick Coleman for expanding my techno brain. Greg Jenkins for cultivating an experimental music culture in Brisbane. All of the testers. Roger Wooller, Jan Buhmann, Cerae Mitchell, Ben Wooller, Richard Vella, Joel Joslin, Danny Muller and all of other people who I forgot to mention.

References

[1] Ableton 2004. *Live Tour - Performance* [web page]. Ableton, web address: http://www.ableton.com/index.php?main=live accessed on 15th of April 2004

[2] Bachmann, T. 2001. Autogam [software]. France: web address: http://autogam.free.fr/ accessed on 8th of March 2004

[3] Barrett, L. 2003. *Working Backwards* [web page]. Brisbane: Small Black Box, web address: http://www.smallblackbox.com.au/reviews.php accessed on 13th of April 2004

[4] Bencina, R. 2003. *Personal Communications*. Perth. email: rossb@audiomulch.com.

[5] Bencina, R. 2004. AudioMulch [software]. Melbourne: AudioMulch web address: http://www.audiomulch.com accessed on 16th of April 2004

[6] Brown, A.; Sorenson, A. and Dillon, S. 2004. Jam2Jam [software]. Brisbane: Exploding Art web address: http://www.explodingart.com/jam2jam.html accessed on 16th of April 2004

[7] Clemen, H. 2003. "Interfaces for Public Use Interactive Installations: Some Design Concepts, Problems and Possible Solutions." *Converging Technologies: Australasian Computer Music Association Conference.*

[8] Cloudkickers 2004. *Cloudkickers* [web page]. Brisbane: LEMu, web address: http://www.lemu.org/download.html accessed on 15 of April 2004

[9] Cycling'74 2004. MAX/MSP [software]. San Francisco: Cycling'74 web address: http://www.cycling74.com accessed on 15th of April 2004

[10] Daconta, M. C.; Monk, E.; Keller, J. P. and Bohnenberger, K. 2000. *Java Pitfalls,* New York: Wiley Computer Publishing

[11] Fraietta, A. 2001. Algorithmic Composer [software]. Sydney Australia: web address: http://www.users.bigpond.com/angelo_f/AlgorithmicComposer/ accessed on 16th of April 2004

[12] Grame 2004. *Midishare and Java* [web page]. Grame, web address: http://www.grame.fr/MidiShare/MainFrames.html accessed on 13th April 2004

[13] Grand, M. 1998. *Patterns in Java,* New York: Wiley Computer Publishing

[14] Jaaniste, L. 2003. *Broken Effort* [web page]. Brisbane: Small Black Box, web address: http://www.smallblackbox.com.au/reviews.php accessed on 13th of April 2004

[15] Joslin, J. 2003. *Personal Communications*. Brisbane.

[16] Kay, S. 2004. *What is KARMA?* [web page]. Karma-Lab, web address: http://www.karma-lab.com/KARMA/What_Is_KARMA.html accessed on 15th of April 2004

[17] Lambert, K. and Osbourne, M. 2002. *Java: A Framework for Programming and Problem Solving,* California: Brook/Cole Publishing Co

[18] Letz, S. 2003. *Personal Communications*. email: letz@grame.fr.

[19] McConnell, S. 1998. *Software Project Survival Guide,* Washington: Microsoft Press

[20] Monro, G. 2002. *REV Festival, Brisbane April 2002* [web page]. Wyoming: Gordon Monro, web address: http://www.gordonmonro.com/misc/rev2002.html#installations accessed on 15th of April 2004
[21] Niemeyer, P. and Knudsen, J. 2000. *Learning Java,* Seastopol, California: O'Reilly & Associates
[22] Propellerheads 2003. Reason [software]. Stockholm: Propellerheads web address: http://www.propellerheads.se accessed on 16th of April 2004
[23] Puckette, M. 1996. Pure Data [software]. San Diego: web address: http://www-crca.ucsd.edu/~msp/software.html accessed on 8th of March 2004
[24] Sorenson, A. and Brown, A. 2004. *jMusic* [web page]. Brisbane: QUT, web address: http://jmusic.ci.qut.edu.au/ accessed on 13th April 2004
[25] SSEYO 2004. *Koan Interactive Audio Platform - Vector Audio / Generative Music* [web page]. SSEYO, web address: http://www.sseyo.com/koan/koanVectorAudio_GenerativeMusic.html accessed on 15th of April 2004
[26] Sun 2004a. *Java Documentation* [web page]. Sun, web address: http://java.sun.com/j2se/1.5.0/docs/api/ accessed on 13th of April 2004
[27] Sun 2004b. *Java Forums* [web page]. Sun, web address: http://forum.java.sun.com accessed on 15th of April 2004
[28] Vella, R. 2003. *Personal Communications*. Brisbane.
[29] Wooller, R. 2003. "A Brief Analysis of Club Drum and Bass: Compositional Structures and Sonic Forms." *Converging Technologies: Australasian Computer Music Association Conference.*
[30] Wooller, R. 2004a. *LEMu Feature Guide* [web page]. Brisbane: LEMu, web address: http://www.lemu.org/files/featureGuide.html accessed on 15th of April 2004
[31] Wooller, R. 2004b. LEMu: Live Electronic Music [software]. Brisbane: LEMu web address: http://www.lemu.org accessed on 16th of April 2004

Discography

Cujo, *Adventures In Foam*, Ninja Tune, 2002

Goldie, *Timeless*, Full Frequency Range Recordings, 1995

Moloko, *The Flipside*, Echo, 1998

N.W.A, *Gangsta Gangsta & Dopeman Remixes*, Priority Records, 1999

Roni Size/Reprazent, *New Forms*, Talkin' Loud, 1997

Stakka & Skynet, *Clockwork*, Underfire, 2002

Squarepusher, *Burningn'n Tree*, Rough Trade Germany, 1997

Various, *Annual 2002*, Ministry of Sound, 2002

Various, *Dance Nation 3*, Ministry of Sound, 1997

Various, *Drum and Bass Arena*, React, 2001

Various, *Drum and Bass Assassins*, Beechwood Music, 2000

Various, *Grooverider Presents The Prototype Years*, Columbia Records, 1997

Various, *Grooverider – The Harder They Come*, Renegade Hardware, 2002

Various, *Grooverider Live*, Anonymous radio mix, 2002

Various, *21st Century Drum and Bass: mixed by Decoder*, React, 2000

The Object of Performance: Performativity in Contemporary Japanese 'Onkyo' Music.

Caleb Stuart

Electronic Arts, School of Contemporary Arts,
University of Western Sydney
Email: c.stuart@uws.edu.au

Abstract

Many contemporary approaches to the performance of music and sound have jettisoned traditional performative modes. Minimal or even non-existent theatricality has become commonplace in both laptop performance and live electronic compositions and improvisations. While there are a number of environmental and social reasons for this development, this paper discusses how these performances can be read as being highly performative by hearing performativity rather than seeing it. The paper will focus specifically on the electronic performances of the Tokyo based 'Onkyo' scene and the musicians Toshimaru Nakamura, Sachiko M and Otomo Yoshihide. These performers exemplify the use of minimal electronics, low levels of physical performativity in live performances and also the extremely quiet use of noise in their music.

1 Introduction: where did the action go?

The work After School Activity (2002) performed by the duo Brett Larner and Toshimaru Nakamura is an example of some of the complications and frictions created between live performance and studio recording by the musicians of the Onkyo scene. From the start it is clear the piece is recorded in the open air and not off a mixing desk. In the background we can here a train, birds, airplanes and finally a group of school children leaving a room.[i] The music itself consists of very quiet interjections from the musicians. Nakamura playing his no-input mixing board plays very quiet sine tones, not much louder than any background noises. Larner plays a koto, occasionally plucking the instrument with much dampening. When the musicians play they effect their instruments with as little physical gesture as possible. The document is then a recording of a performance with little sound, even less performativity, and no audience, placing the work between a live performance and a studio based recording.

The use of electronics in performance causes various negative responses from the audience, who feel a loss of spectacle and performativity in the action. This loss is felt as a result of the lack of gesture and visual cues from the performer and the loss of focus from the audience who have no visual object to ground their aural experience. A shift in focus from an understanding of the visual spectacle in performance to that of aural performativity is needed. Once this is understood the audience can approach the performance with a revisited focus and come to new understandings of the aural object and contemporary approaches to sound creation.

As I have written in "The Object of Performance: Aural Performativity in Contemporary Laptop Music,"[1] the aural object can be heard to exist in laptop performance through the presence of the sound itself. That is, through various aural devices sound becomes performative in itself, without requiring the visual to support it. With the use of the laptop in performance this is often heard through the utilisation of volume and also psychoacoustic sounds.

This paper will look beyond the computer in performance and look to the contemporary practice of live electronic improvisation and specifically the work of contemporary musicians within the Tokyo based scene often entitled 'Onkyo.' The music of Toshimaru Nakamura, Sachiko M and Otomo Yoshihide is often performed at the point of being inaudible. The audio, made up of various types of electronic noise, barely escapes the sound system. In addition there is often an extreme lack of gesture, the musicians hardly move and display no facial expression throughout the work.[ii] The work raises the issue of performativity in the live performance context, clearly questioning traditional values of the body in music, theatre and the spectacle.

Hoe can this music be heard as being live? Is it simply due to the improvisational nature of the performance or is there more to it? In this paper I will argue that the performance is best understood as an aurally performative work and that it cannot be heard as visually performative.

This shift in focus is not an actual change as music is after all an aural experience, but the audience has become both distracted and dependant on the visual spectacle and the physical gesture of musical performance. It is not that performance, gesture and movement are extra-musical, but that they are not the sole basis for music, nor are they necessary for the performance and reception of music and especially that of the sound and audio arts.

Proceedings ACMC 2004
ISSN: 1448-7780
http://acma.asn.au

It is also important to note from the outset that the electronic music scene addressed here is experimental in nature and is followed by a small sub-culture. In the last few years this sub-culture has become comfortable with the nature of live performance and with electronics being genuinely understood as instruments in themselves. The paper then looks to the scene for direction as to how a wider audience could begin to understand the nature of contemporary non-visual performance.

2 The Place of Performance

Performance's only life is in the present. Performance cannot be saved, recorded, documented, or otherwise participate in the circulation of representations of representations: once it does so it becomes something other than performance.
Peggy Phelan (1993)[2]

In her book *Unmarked: The Politics of Performance* Phelan argues that performance's own ontology is such that it only becomes itself through its own disappearance. The documentation of the performance is only then a 'spur to memory' nothing else.

The idea of live performance not being reproducible, of being completely created in the moment, leads to a spectator's gaze desperate to take in all the elements of the performance before its disappearance. Quickly the existence of any live performance plunges into the memories of the spectator, living in the mind and slipping into the unconscious. But without an object to hold, performance is often charged with emptiness and even valuelessness. While Phelan gives the visual gaze the full weight of her argument she does not go far enough for our purposes here. How is the spectator's gaze to deal with a situation where there is next to nothing for it to rest on, for it to consume? The spectator at a loss for the visual is confused as to what the object of the performance is? I believe the way out is to listen to the performance and through close listening the aurally performative nature of the work becomes clear.

3 Aural performativity

Performativity is directly linked to the body. We talk about the performance of the everyday in gender, sexuality, race and culture. In the performing arts and music the body is generally near the action, on display. The objection to music performed with electronics, or with a laptop for that matter, is that the audience is not being given the visual stimulus of a body. That is, the body of the musician is not directly and causally acting on an object physically to create a sound. The casual relationship between what we hear coming out of the speakers and the body of the performer is broken. The audience does not have anything to ground the listening experience in, to compare contrast or relate it to in regard to what they are hearing with their ears. We ask how is it that the near inanimate body of the performer is making the sound we hear, and we ask how is it that the array of boxes and wires, or the single box with a single cable can create the sounds being heard? The link between the slight twist of a knob or the sliding up of a fader does not always equate to an equal twisting or sliding of the sound we hear, the actual effect of such movements might not become aurally present for sometime, thus the direct link between gesture and sound outcome is broken.

In this situation a jump in thinking is needed. We need to grasp that the audience are the receptors of the performance and most importantly with this music they are the listeners - the performer is also a listener. This is especially clear in psychoacoustic experimental music as the bodies of all individuals in the space are receiving different aural information based on their position. The audience and performer are in the act of listening and the sound itself is physical in nature. It can be loud and painful, quiet and fragile and so on (we describe these sounds in a physical language). So the sound itself is interacting on the bodies present in a physical manner. We are literally moved by the sound waves and react to them as such. The air being pushed around the room causes us to have a physical experience as we hear the unfamiliar tones, which change with every slight movement of the head or we hear dense layers of audio that form complex relationships to each other. The performativity of the music is to be found in the act of listening and the performance of the audience in relationship to the sound they hear. There is no need then for us to see a performer physically interacting with an instrument to engage in this aural area.

It is important to note that the use of live electronics has a somewhat different relationship to history than that of the laptop.[iii] With this in mind it is important not to see the use of electronics in new music as a new or novel approach to sound production and performance. On the most basic level the technologies are not new, nor are they desirable domestic consumer technologies. These technologies are basic audio equipments such as a mixing desk and a sampler, they have been around for many decades and have been used in new musics for as long. There is a direct link between composer/musicians such as John Cage, David Tudor, Alvin Lucier and Gordon Momma, as similar problems of performativity were felt by the earlier musicians. This specific argument, however, is outside of the scope of this paper and is one for possible future examination.

4 Japanese Onkyo: improvising electronics in performance

Toshimaru Nakamura sits next to Sachiko M in The Studio at the Sydney Opera House. They sit as still as possible. Any movements towards the twisting of their instruments' knobs are done in an almost imperceptible manner. They sit with their heads down and without perceptible facial expression. High pitched sine waves and feedback is heard at extremely low volume. Brief interjections of pops and 'electronic' almost faulty sounds occasionally enter the space. The sound is so extremely quiet and fragile that one might not want to swallow in case they disturb the proceedings. In the space sound is made tactile, made palpable. Extra musical noises and sounds around the venue are greatly heightened, the humming of the city made perceptible. The performed sounds feels to the ear and the body to be full, they seem to surround and envelope and yet the volume level is very low. In fact the sound is so low that slight noises around the space are framed by this music, sounds not usually noticed are brought into clear focus amongst the deliberate sounds of the music.

This is a different way of listening to the sounds than the approach taken in John Cage's silence. We are not in this case hearing "all sounds" as music. Instead we are simply hearing sounds. These sounds are juxtaposed with the sounds created by the musicians but are not heard as part of the music, instead they are heard because of the music.

An important difference between the model of laptop music and contemporary electronic music is the absence of the hi-fi nature of the sounds produced. That is, in the laptop music of composers such as Pimmon it is of extreme importance that the music is played through the best possible sound system. The higher the quality of the system the better it will be able to cope with the extremes of the music. It can be argued that in laptop music it is literally the PA that actuates the music, that without the PA the presence of the performance, its aural performativity, will not be felt. This is not the case with the performers from the Onkyo scene. When performing at Off-Site musicians will commonly play through a guitar practice amp, each musician accessing one amp placed near their equipment set-up.[iv] There is no space or need for a PA at this venue. During performance these amplifiers are set to low volumes and some musicians will go as far as placing them face down on the floor.[v] It is also one of the reasons for the development of a very quiet style in the music, as Off-Site is both too small and too near other residential occupants to allow for high volume levels. Consequently the musicians needed to find an alternative way to produce sound.

The argument that aural performativity can be found in the quality and nature of the PA, the fact that a good PA and venue are needed to experience the performance at its truest, cannot be used here to argue for an aurally performative model for hearing this music. What then is the nature of the sound and what about it is aurally performative?

The effect of the sound can be fully experienced at low volume without the use of a PA, why then not merely use your home hi-fi equipment? Why can this music be heard as a live aural performance?

An example of the height of both quiet sounds, subtle shifts and minimal gesture comes from a performance by the duo Filament (Sachiko M and Otomo Yoshihide). In a performance at Kanagawa Kenmin Hall in Yokohama in September 2000 the duo play a single piece. The work is made up of various types of sine waves and high pitched metal tones. Sachiko M's sine waves are produced by an emptied sampler and Otomo's metallic tones from a turntable with a prepared cartridge-head, made up of a piece of metal soldered to the head which rests on the turntable plate. The turntable makes a high pitched metallic sound when running. The piece is performed at low volumes and yet because of the nature of the high tones and the sine waves the sound is extremely full and the work fills the space, seemingly coming out of the walls. The psychoacoustics of the sine wave has the effect of dramatically changing the sound levels with the slightest movement of the head. This effect makes it clear that what you, as an individual in the audience, is listening to is not the only version of the piece. On moving one's head you realise that there are many ways to hear this music and the audiences is needed to activate the sounds.

This is a similar experience to that of listening to the high volume work of Phil Niblock. In his performances he encourages the audience to move around the space. As they move the sound changes and shifts as different psychoacoustic effects are perceived and experienced. In the case of the Onkyo musicians it is clear that these types of experiences do not in fact need high volumes to be activated. Instead a similar experience can be achieved through the use of sine waves at low volumes.

While the audio at the Filament performance seems to engulf the listener, it is also note worthy that it was possible to hear, in the large concert hall, the clicks of the buttons of the sampler as Sachiko M gently and intermittently depressed them. These clicks served to illustrate the actual volume of the work. Atsushi Sasaki, a Japanese music critic, in the liner notes to the performance describes the sounds heard on the recording:

> To the constant sound of subtly strengthening and weakening sine waves are added the intermittent miniscule sound of circuitry, and a dull mechanical hum (too soft for us to recognize its origin). From beginning to end hardly anything happens here, but at the same time an amazing number of events occur. If one were to give a name to this "Onkyo" of just under 30 minutes, the words composition, improvisation, and performance would probably all be too extreme. But all of these are included in this quiet phenomenon.[7]

As Sasaki points out over the course of the performance almost nothing happened, slight shifts and the fragile hanging of small sounds seems to be almost at the brink of collapse, or perhaps less dramatically in simply fading from perception. On listening to the work it is important that the audience take note of the phenomena that is occurring around them. The work in many ways exists by its changes and by what is not there, or what is not presently there.[5]

While it is possible that this could be reproduced at home it is not so easy to listen closely and it is this very close listening that is the key to fully understanding the work. The live performance presents us with a situation where it is expected that we will listen, that we will be silent, that we will sit as still as possible. This is of course not possible in the domestic environment. If any music illustrates Adorno's loss of the original in the domestic consumption of sound it is the work of the Onkyo musicians.

Conclusion

Where does this music exist? How can we view it as spectacle, is there spectacle in the work? On listening and fully concentrating on this music we find that the sounds have a presence, a texture and an aural presence and that spectacle can be found in the very act of listening. The sounds seem to fill the space and on close listening interact with the audience, the room and its space. We feel the presence of the sound as if we might be able to grasp it, if only for a fleeting moment. While this effect can occur from listening to the recorded documentation of the performance or a CD release of a studio recording all the risk is removed. The presence of the performer/listener in the room is of key importance and the need for close listening. The act of listening then is clearly active and the performativity is carried out in an aural setting situated in the moment. When this is grasped by an audience, as it is in Japan at Off-Site, then the audience will fully hear the performance.

References

[1] Stuart, C. 2003. "The Object of Performance: Aural Performativity in Contemporary Laptop Music." *Contemporary Music Review* 22(4): 59-65.

[2] Phelan, P. 1993. *Unmarked: The Politics of Performance.* London: Routledge, p.146.

[3] Sasaki, A. 2001. "The Oscillating 'Will' and the Flickering 'Self'." Liner Notes to *29092000* (Amoebic, AMO-SAT-03).

Endnotes

[i] This is an effect of many recordings of live events in Japan. These recordings are almost always recorded on a microphone and often background noises from the venue can clearly be heard. Also more specifically with the music produced by the Onkyo musicians it is often the case that this background noise is louder than the music actually being performed.
See Caleb.K, "Interview with Taku Sugimoto, Tetuzi Akiyama & Utah Kawasaki," *Angbase* 2002, http://www.angbase.com/interviews/mongoose.html

[ii] This is not always the case, Otomo, for example, often has a heightened performance in his turntable and high volume works. Otomo when playing Onkyo style music, however, does play in stillness.

[iii] The laptop in performance can be seen as initially a kind of underground fashion as musicians such as those from Austrian label Mego joined a programming sensibility with new trends in experimental musics. Whilst the applications used may have obvious links to earlier computer sound programs and languages, basically the approach taken needs to be heard as a shift from academic computer music to contemporary digital music. This shift quickly moved to other producers who were briefly known as producing sound in a post-digital manner (for example Kim Cascone and Pimmon). It seems that from a contemporary perspective the use of the laptop in performance was a brief moment in the history of performance, one that has now moved on.

[iv] For a background on the scene in Tokyo see Clive Bell, "Site for Sore Ears," *The Wire*, July 2003.

[v] This has a number of effects including situating of the individual performers sounds. A common problem in live electronic performances between a number of musicians is that unless the audience knows the signature sounds of the musicians it is very hard to make a firm connection between the sounds being produced and which individual performer is producing them. The amp then literally make each individual set up into a discrete instrument.

Performance Theory And Practice In Stockhausen's Kontakte

Robin Maconie

email: maconie@xtra.co.nz

Stockhausen gained his early professional experience in radio, first as a student in the musique concrète studios of Paris Radio under Pierre Schaeffer, then as resident composer and later director of the electronic music studios of Cologne Radio.

Both Paris and Cologne studios operated under the wing of the Radio Drama department. Historically, aesthetically, and technically, electronic music is a form of radio drama. Radio drama was an entirely new form of sound composition. It explored the voices inside your head and how they interact with the real world. In the beginning, radio drama was broadcast live. It called for new performance skills. The microphone extended the range of conscious hearing. It brought together natural voices, artificial sounds, and music in one experience.

For about twenty years, until the arrival of tape after the war, radio led the way in production standards, because it went out live and did not rely on imperfect recording media. Not only computer music but music in every medium owes an enormous amount to the pioneers of radio. They were the first to recognize the challenges, and the first and most persistent in trying to solve them.

When we talk about performance theory and practice in electronic music, we mean not only how live music can be combined with music that comes out of a collection of loudspeakers; we also mean how the reality of a live performance can be reconciled with the realism of a music existing on tape, both as a documentary object, and as a subjective experience. But this is nothing new. If you go to a concert to hear a Mozart concerto, for example, the same process is going on there as well: reconciling the physical reality of the solo performer, representing the here and now, with the historical reality which is Mozart's score, and the psychological reality of what was going on in the composer's mind.

I am going to talk about performance theory and practice in relation to two major compositions of electronic music by Karlheinz Stockhausen, the *Gesang der Jünglinge* of 1955-57, and *Kontakte* for piano, percussion, and electronic sounds, from 1958-60. The earlier piece was composed for five-channel tape and incorporates the pre-recorded sound of a boy singing. It has no live performance component, although the boy's voice is intended to sound real and is recorded to a very high standard of realism. The impression is like a kaleidoscope, an aural vision of multi-faceted and coloured jewels moving about, that for a moment become recognizable, and then move into new patterns, always changing:

Music Example 1: *Gesang der Jünglinge*

It is a fundamental principle of Stockhausen's electronic music that what you hear is not just a loose collection of items picked up here and there from sound effects discs or by random experimenting with patches and so on. That sort of procedure can lead to attractive results, but that is that: after a while, you listen again, and bits start to flake off because there is nothing organic in the design to hold it together. Serialism gives you a structure for discovering a total range of sounds that share certain basic properties. If you are painting your house and want to work out what colours go together, a structured system of colour relationships will give you greater certainty and more flexibility of choice. With music it is the same. Understanding colour means you understand vision, how colour is perceived by the eye. With music you are involved with the hearing process, which is somewhat more complicated because it involves dynamic change in time, and thus memory as well.

The unifying element in *Gesang der Jünglinge* is the voice. Electronic music in the 1950s was closely identified with scientific research in speech recognition by computer, then a slow analogue device. Western science wanted to be able to program early computers to monitor suspicious telephone conversations, and in order to do that, they needed to work out the basic sound elements of speech, and rules for how they hang together. Composers were brought into this, often without fully realizing it, which is why so much early electronic music is speech-related. So the idea is: "Break down speech into its basic particles and then show how to put it together again."

Music Example 2: *Gesang der Jünglinge*

The easy way out of course would be to cut up a tape of real speech and then re-assemble it, which is what Berio does in *Omaggio à Joyce* and Eimert in *Musik und Sprache*. It works, but you don't actually learn anything. It's like running the film of an explosion backwards: everything comes together beautifully but you don't actually learn anything new

Proceedings ACMC 2004
ISSN: 1448-7780
http://www.acma.asn.au

about how structures fall to bits or how they can be put together again. What Stockhausen does is more systematic, and also more interesting. He wants to find out by analysis what the basic elements of speech are, and then fabricate artificial sounds that resemble them, so that when the process is reversed, and the voice is reconstructed, you know where you are coming from as well as where you are going. So you end up with a list of basic sounds from [fff] and [sss] to [ah] and [oh], and you discover that consonants are noises of specific bandwidth, and that [t] is the same as [sss] but shorter, and so on. It is a learning process, and you come out knowing a lot more about the nature of speech sounds than you did before.

Of course you can say that a good composer intuitively understands the human voice, so all this hard work is unnecessary. If only that were true. Even poets need to listen to poetry as sounds as well as read and write it as words on paper. Only that way do you reach a point where the sounds acquire their own level of meaning, quite separate from the sense content of the words themselves.

What Stockhausen and his colleagues learned from the experience of *Gesang der Jünglinge* and trying to simulate the human voice, was that it was not possible, with the technology then available, to reduce continuous speech to an alphabet of basic sounds. Not only that, it would never be possible. The voice doesn't work that way. Think about early computer imaging, where the best you can do is 20 x 20 resolution. But the image is fixed, so you know that it can be improved by increasing the resolution to 2400 x 2400. If the image is moving, all your camera needs to do is follow the bits that move and ignore the rest. With speech however everything is in continuous motion, and every element in the flow of speech affects everything else. So the intellectual model of speech as a jigsaw puzzle is never going to work, because there is always only one way to put a jigsaw puzzle together.

That the theory Stockhausen followed did not work in fact, is a discovery of the greatest importance. The theory was wrong. Now we know. That is what research is about. It doesn't change the fact that the composition is a work of genius, beautifully crafted, exciting to hear. It asks what is still a fundamental question: How does the voice work? What bits of information are involved? How do we recognize words as such? How do the sounds of words join together to make continuous sense? We still need to know.

Stockhausen's next major project in electronic music was *Kontakte* which means "Contacts." Here the focus is instrumental rather than vocal sound. As a student in Paris in 1952 he had spent long hours listening to and analyzing the evolution of recorded sounds of ethnic musical instruments: how the sound is initiated, reaches a steady state, and then dies away. This was years before the arrival of the most elementary analogue synthesizers. In comparison to speech, which is produced by an infinitely flexible structure, instruments are stiff resonators, for the sake of consistency of tone. That makes them simpler to define. There are fewer basic elements.

Once again, for practical reasons Stockhausen is not concerned with modelling the acoustic process, which is impossible when you are working with analogue tape. What he did instead was to create accelerated waveforms of different kinds. He was following in the path of the great Hermann Helmholtz, who experimented with sirens that could be programmed to a limited degree to vary the tone, and makers earlier in the century like Hammond, Martenot, and Trautwein, who employed tone wheels and simple electrical circuits to generate a range of simple wave forms: sine, square, ramp, etc. By creating a range of artificial wave forms out of accelerated loops of tape, in the form of rhythms, melodies, arpeggios, impulses, continuous tones, etc., Stockhausen ended up with a collection of continous tones of different timbres.

Music Example 3: *Kontakte* steady tones

This sort of work is hit or miss. You don't know in advance how the wave-forms will sound, so you work through a range of possibilities: some like drum patterns, some dynamic shapes, some accelerando, some made of more than one layer, and so on. And you play back the resulting continuous tones on a vari-speed tape machine, on a scale from low to high, and listen out for realistic sounding results.

Now if any of you have sampled a dog bark and played it back on a midi keyboard, you will have discovered that as the pitch goes up and down, the dog changes as well from a terrier to a chihuahua and then to a great dane. That is because every sound implies a structure. With musical sounds the structure is a musical instrument. In regular music, to achieve consistency of timbre you look for a family of instruments, with a miniature for the high notes and a larger version for the low notes. We have known about this for seven hundred years: a consort of recorders, a viol consort, a string quartet, a saxophone quintet, etc. It's all about scale transposition in physical terms as a way of achieving consistency in terms of tone colour.

In *Kontakte* however you end up, not with families of timbres, because you have only one tape of each wave form. Instead, you have groups of timbres that are identical in their internal structure, but sound very different according to pitch, duration, or treatment. In fact, they sound so incompatible it is a puzzle to understand how they can be related at all. For example, if you take a pure sine tone at 165 Hz, e below middle c, and hear it as an impulse, it can sound like a drum; but the same wave form at 512 Hz, c above middle c, if you let it sound, is like a flute. Take the same sound at 2000 Hz, chop it chort, and add reverberation, and it becomes a little Greek cymbal.

Example 4: *Kontakte* fugato

So what do you do? What theory can account for that? What Stockhausen decided was utterly brilliant, and in a way completely consistent with what he had been doing before, with the boy's voice, in *Gesang der Jünglinge*. He said, what this tells us is that the real musical instruments we know are related to one another in ways we had never realized before. They represent categories of thought that are specialized and consistent within themselves, but that are not in fact real. (This is the philosophical bit.) Not real, that is, in terms of the continuum of vibration as it really is. The music of *Kontakte* with live performers is already interesting because the audience is in the midst of an electronic sound world that is constantly whizzing and burping and rotating round your head, and you can hear that from time to time the players on the platform are trying to keep up with the electronic sounds, and identify them, play along, or even shoot them down. This is not your average electronic piece by Babbitt where the live voice and the electronic music go their own way. It is not like Berio's *Différences* where the live instrumental sounds are gradually replaced by electronically distorted versions, like a hall of mirrors. In *Kontakte* we discover that the electronic music is the higher reality, and the actual musical instruments are only cross-sections. You know that astronomers take the light from a star and create a spectrum, and look for black lines in the spectrum; these are the absorption lines, and they tell you what elements the star is made of. In *Kontakte* the electronic music is the spectrum of possibility, and the musical instruments the absorption lines, the elements.

Quadraloop: real-time audio looping software

Ivan Zavada

Sydney Conservatorium of Music, The University of Sydney
E-mail: izavada@conmusic.usyd.edu.au

Abstract

This paper will discuss the efficiency of a software designed specifically for the creation and generation of variable audio loops. Thus, while proposing a new computer tool for musical expression linked to a specific compositional process (looping), I wanted to explore and personalise a compositional language that would integrate a new form of expression in the professional music world. The development of the software Quadraloop gave me the opportunity to explore a new way to create music. Rather than using tools yielded by existing musical traditions, I created a link with these traditions through technical research and programming. By giving emphasis on a specific parameter, it was easier to develop a unique graphical environment to explore new musical possibilities.

1 Introduction

Composition procedures today are widely dominated by the presence of computers. The specialised field of computer music represents the outcome of the accumulated composition techniques of the twentieth century. This is true in the instrumental domain as well as in the electroacoustic music domain. The great musical inventions are now represented virtually and it is now possible to simulate various sound emissions and precise composition techniques in very elaborate ways. Henceforth, the act of composing combines technical innovation and musical expression. In the past years, new writing techniques were rarer, which gave the opportunity to merge various musical languages. A reflection or a fortaste of the globalising era? "A new way to make music rather than creating a new musical style"[1].

2 Quadraloop

The software Quadraloop was developed to adhere to the following criteria: to explore and exploit the variability of an audio loop in a real- time situation, in the context of electroacoustic composition. The application generates the necessary sonic material for the realisation of an entire musical work. The most important aesthetic criteria is repetition of course. This follows an already existent musical style called Minimalism. But the goal of this software is not to recreate the Minimalist musical style. There are several musical genres, which promote repetition as a vehicle of expression. This is why Quadraloop focuses on repetition as a phenomena and its different aspects. This software can therefore be used in various situations.

3 Musical applications

The three main functions of Quadraloop:

1- To compose a piece consisting of audio loops

2- To generate new sonic material in real-time

3- To create a loop for later usage

3.1 Composing

The basic construction of Quadraloop is done within Max/MSP's graphical programming environment (MacOS). The composer in the computer age is confronted with a new way to create. Not only does the composer need to be musically creative, but also needs to be technically inclined to successfully express new musical ideas. At least, this is the case in electroacoustic music. Did the act of composing shift toward programming and technical innovation? Maybe not to such an extent, however, it is at least necessary to be fluent in some technical or programming language as they are now indissociable with artistic creativity.

Quadraloop is designed to be used in several musical applications. The main application is the construction of an entire piece based on looping and repetition. It is possible to display a timeline with the representation of all the parameters affecting the imported or pre-recorded sound files. The visual score is then executed and it is possible to record the result. This then becomes a composition, in the sense of traditional electroacoustic (acousmatic) music.

3.2 Real-time Performance

An area in constant expansion is real-time performance and real-time manipulation of sound. Quadraloop emphasises this specific current in music and favours interaction between artist and technology. This important aspect in the conceptualisation of a virtual instrument enhances the role of the musician, in broadcasting artistic expression. Traditional electroacoustic music favours individual work in the studio. Nevertheless, "the possibility of *double immediate* widens the perspective of the artists, composers and musicians. Expand the instrument and project it in space,

Proceedings ACMC 2004
ISSN: 1448-7780
http://www.acma.asn.au

reinvent musical writings, widening the field of musical interpretation, such are the possibilities offered by these new technical and musical means"[2]. It is very interesting to use Quadraloop as an instrument generating sound instantaneously, without necessarily having a direct control over the final result. This leaves room for interpretation and adds an almost tangible human dimension to the performance.

3.3 Sound Design

The Quadraloop application can be used as an accessory to another music creation environment, as a tool for sound design. It is possible to create short musical instances, either melodic or rhythmic, for later insertion in other musical works. Because of the design of the interface, it is also easy to create new sounds otherwise impossible to simulate in more conventional graphic environments. For example, the variable indexing within a file permits a variable sweeping using both *loop start* and *loop end* indices. The following is a symbolic demonstration of what is called variable looping. The 26 letters of the alphabet represent the audio file, and the letters in boldface print designate the loop. Here, it is easy to observe the behaviour of the *loop start* and *loop end* indices:

State 1 : A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

State 2 : **A B C D E F G H I J K L M** N O P Q R S T U V W X Y Z

State 3 : A B C D E **F G H I J K L M** N O P Q R S T U V W X Y Z

State 4 : A B C D E F G H I J **K L M N O P Q R** S T U V W X Y Z

State 5 : A B C D E F G H I J K L M N O **P Q** R S T U V W X Y Z

State 6: A B C D E F G H I J K L **M N** O P Q R S T U V W X Y Z

4 Interface

The design of the graphical user interface is of great importance for efficient use in a musical application. This phase is determinant in moulding the aesthetic outcome of any virtual instrument. Visual organisation of the different parameters is a critical element to facilitate maximum usage of a process. The position and assigned priority of each parameter control gives a certain emphasis on specific manipulations. For example, in Quadraloop, there are three main parameters for changing sound. They are represented by three rotary knobs (X Y and P), index, length and pitch (see Figure 1).

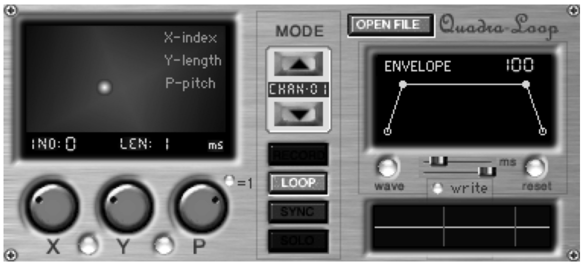


Figure 1. One of four looping modules in Quadraloop

The interface offers several graphical methods for processing sound, depending on the context and degree of desired control. A simple parameter of a loop can generate various musical results. In a two dimensional plane, for example, it is possible to control two parameters with a single click of a mouse. In the future, it would be nice to explore a three-dimensional space and have access to the three main parameters at once. This would certainly be possible with sensors, graphic tablets and external gestural controls.

5 Conclusion

As a composer, this project allowed me to develop new skills in an expanding field. I realised it was possible to invent your own creative musical instruments and share a personal artistic expression through formalisation of an abstraction. A structured approach may harvest a proper evolution of a distinctive musical idea. The framework of a technological process allows artists to enrich their vocabulary of expression and helps them find a new sense of musical direction. The potential of musical creation and artistic innovation is largely dependant on available tools and this has a growing influence on current technological developments. Quadraloop software works in this direction and enhances interaction with musicians. This novel graphical user interface for musical expression certainly blends well in the latest musical trends.

6 Demonstration

The possibilities with the Quadraloop software will be demonstrated with a real-time performance for violin (Ivan Zavada) and computer (Quadraloop software).

References

[1] Attali, J. 1989. *Bruits*. Fayard, Presse Universitaire de France, p. 33.
[2] Lévy, F. *Le temps réel en musique – CD-R*. IRCAM Publishing 2002.

Synthesis of Chinese *Dizi* Ornaments

Lydia Ayers and Andrew Horner

Computer Science Department, Hong Kong University of Science and Technology
E-mail: layers@cs.ust.hk and horner@cs.ust.hk

Abstract

The *dizi* is a Chinese transverse flute that produces a characteristic nasal buzzing tone. This project uses a line segment method for performing grace notes. To model a trill, we used one function table for frequency modulation and another for amplitude modulation. We combined the two designs to connect grace notes to a trill. This method produces realistic *dizi* ornaments that sound better than overlapping the tones and are easier to use than connecting them all with a line segment.

1 Introduction

The *dizi* is usually made of bamboo, with a cane membrane over one hole (see Figure 1) that produces a characteristic nasal buzzing tone rich in upper partials. Varying numbers of grace notes often precede or follow *dizi* trills. This project uses a combination of a line segment with frequency and amplitude modulation functions to make realistic *dizi* ornaments.

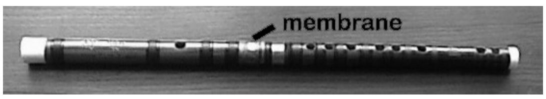


Figure 1. Dizi.

We studied the musical characteristics of *dizi* ornaments to answer the questions: What is the most accurate and efficient method to synthesize those characteristics using the group wavetable synthesis instrument design from our previous work? Is the method good enough to fool listeners into believing that humans played the synthesized examples?

After comparing the results, we concluded that the line segment method is an efficient method for modeling accurate slurred grace notes, but trills sound better using separate function tables for frequency and amplitude modulation, and the new method is easier to use. Combining the two methods is highly successful and the ornaments sound realistic enough to fool listeners into believing that humans played our synthesized examples.

2 Background

Previous research modeled the spectra of ordinary single tones of many wind instruments [1][2][3][4][5]. We previously synthesized individual *dizi* tones using a wavetable model that is easy to use and sounds like the original tones [6][7]. We used the same design to model the *dizi* tones in this project. Isolated synthesized tones may be indistinguishable from acoustic tones on listening tests, but musicians don't always play isolated tones! They connect tones to make "musical" performances. Since wavetable synthesis is straightforward and matching works well for isolated *dizi* tones, we will focus our discussion on modeling ornaments connected to the tones, such as grace notes and trills.

3 Synthesizing Grace Notes

Overlapping can simulate slurs when the decay of one tone can overlap the attack of the following tone. But this technique doesn't work on grace notes because the notes are too short, and they can even have clicks. Connecting together very short tones, therefore, often makes it obvious that the music is synthesized. Connecting the frequencies of two notes with a line segment and morphing the transition using parameter interpolation [8] obviously can give a smoother slur.

We created an instrument which cross-fades two unison signals sharing one frequency line segment and phase, but using their correct wavetables (see Figure 2). Amplitude envelopes cross-fade the two signals as with overlapped notes, but it is smoother because the frequency change in the transition more closely resembles one on a real *dizi*, and using the same phase (and the other frequency parameters, such as noise and vibrato) also makes the cross-fading itself less noticeable. We also increased the noise in the cross-fade as the amplitude decreases, so that the maximum amount of noise is in the middle of the cross-fade, at the point of minimum amplitude.

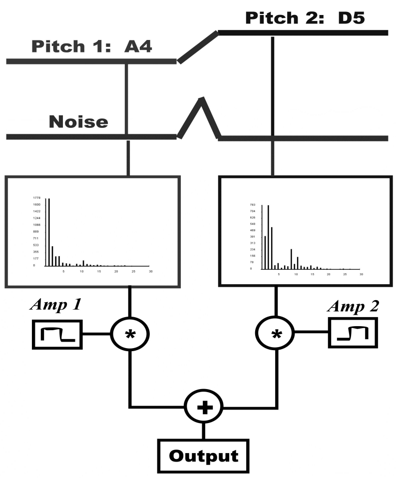


Figure 2. Cross-Fading Two Notes with Noise in the Transition.

Proceedings ACMC 2004
ISSN: 1448-7780
http://www.acma.asn.au

We can use the same method to play additional notes. Our design can play three slurred notes, allowing one grace note before and one after the base note.

4 Synthesizing Trills

But how can we slur varying numbers of notes in trilled phrases? In Csound, slurred groups of different numbers of notes need different numbers of input parameters. Requiring the maximum number of input parameters for every note statement and using placeholder numbers for those we don't need results in an extremely cumbersome design. A whole note trilled quickly in a slow tempo can easily contain more than 32 individual notes! Can a repeating function controlling the frequency of a single tone produce smoother trilled notes than overlapping tones does?

5 Trill Frequency Modulator

What is the average repeating pattern that best represents the basic frequency change for this trill as well as others? The function does not need to model pitch variation of the average tone, change of speed or jitter, so it can represent one average cycle of the trill, and adjusting the parameters randomly within their typical ranges can vary each cycle.

The frequency modulator must use a function that can approximate the frequency analysis graph of the trill shown in Figure 3.

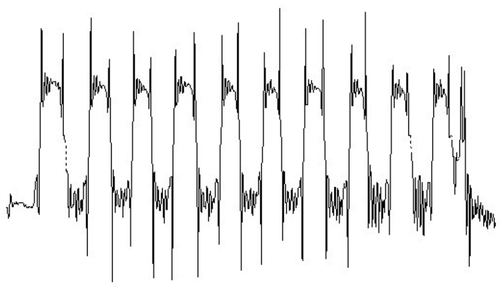


Figure 3. Frequency Analysis of an A5 to B5 Trill.

Figure 4 shows a close-up of one cycle of this trill. Averaging one cycle provides a good shape for a trill frequency function.

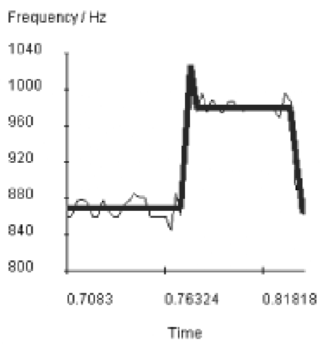


Figure 4. Averaging One Cycle of the Trill for a Frequency Function.

The function must oscillate between the frequency of the lower note and the frequency of the higher note. The cycles begin with a slight overshooting of the required frequency, perhaps 20% (see Figure 5), and we randomly vary the trill rate.

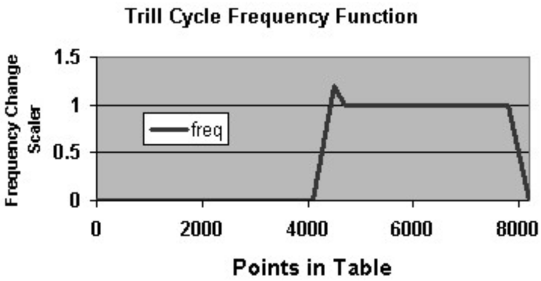


Figure 5. Trill Cycle Frequency Function.

6 Trill Amplitude Modulator

Figure 6 shows the waveform of two cycles of the trill. The waveform shows a drop-off in amplitude in the transitions between the trilled notes. It also shows that one of the notes is a little bit louder than the other.

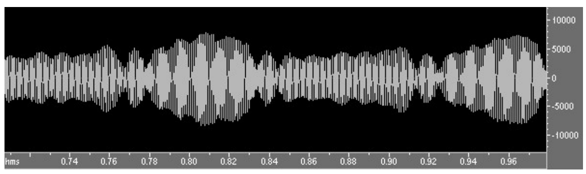


Figure 6. Two Cycles of Waveform of Dizi A4 to B4 Trill.

What is the optimal shape for the trill amplitude function? Analyzing the two trill cycles shown in Figure 6 provides more detail on the sharp amplitude drops during the transitions between the notes (see Figure 7). In addition, the amplitude shows an extra spike at the beginning and ending of the period.

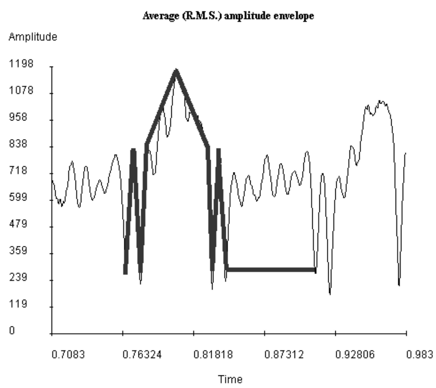


Figure 7. Averaging a Cycle of the Trill for an Amplitude Function.

To get a realistic trill, we must model the amplitude changes of the transition. Although the most important amplitude changes occur at the same time as the most important frequency changes, we cannot use the same function to control both the frequency and amplitude because the frequency peak is at the beginning of the cycle and the amplitude peak is in the middle. In addition, when notes in tremolos alternate by leap, each note must use its own wavetable for its spectrum. This requires cross-fading two signals while the first function alternates their frequencies.

The average amplitude of a trill cycle provides a separate function that must oscillate between the amplitude of the lower note and the amplitude of the higher note (see Figure 8). The cycles begin and end with a slight amplitude spike during the transition between the two notes.

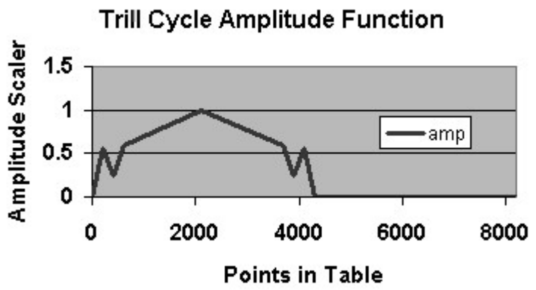


Figure 8. Trill Cycle Amplitude Function.

We stored the new trill shape in a function table and then used an amplitude modulator in that shape to alternately fade the amplitudes of the two signals in and out (see Figure 9).

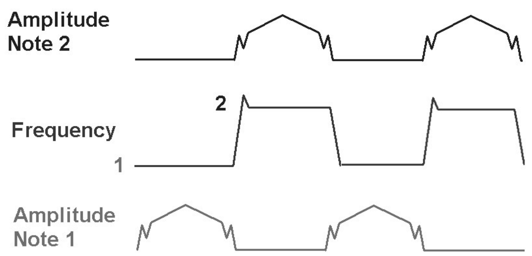


Figure 9. Combining the Functions to Produce a Trill.

This design permits several score parameters to control the trill rate using a line segment. For example, our design uses four score parameters to control an initial, middle and final tremolo rate, and the time required to change from the first tremolo rate to the second. A separate score parameter controls the changing amplitude of the trill as it would for a single sustained tone. This method is much easier than typing many notes and trying to manually adjust their start times, durations and amplitudes to get a naturally-varying rate and amplitude quality.

7 Combining Ornaments with the Trill

Trilled notes on the *dizi* may be preceded or followed by grace notes. A four-note design covers many of the possibilities (see Figure 10).

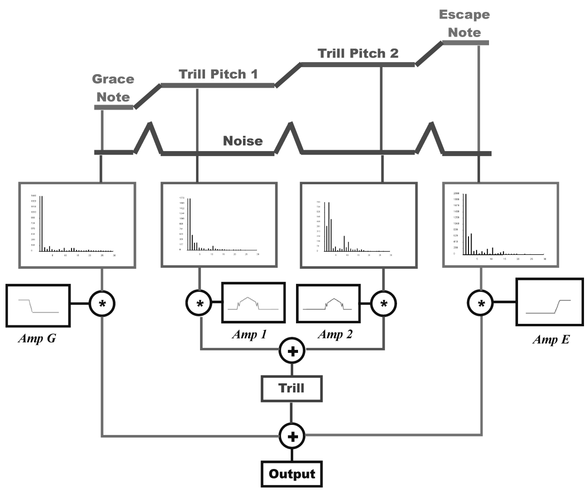


Figure 10. Summary of the Four-Note Design.

We assigned one note for the preceding grace note and one for the escaping grace note, with a trilled pair of notes between the grace note and the escape note. A single line segment, shared among all the notes, controls their overall frequency (see Figure 11). The tremolo rate is set to 0 during the grace notes, since they are not part of the tremolo, and it varies as previously described during the tremolo between the grace notes. The amplitude envelope of each signal is set to 0 during the time it is not active.

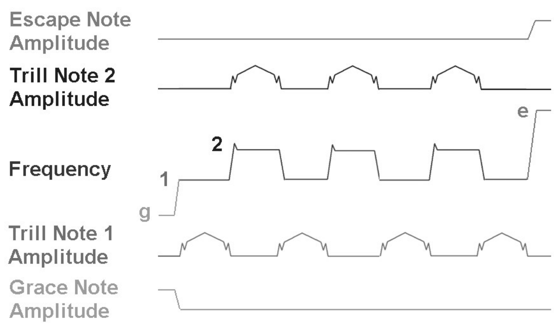


Figure 11. Adding Grace Notes Before and After the Trill.

8 Conclusion

After comparing the results, we concluded that the line segment method is an efficient method for modeling accurate slurred grace notes, but trills sound better using separate function tables for frequency and amplitude modulation, and the new method is easier to use. Combining the two methods is highly successful and the ornaments sound realistic enough to fool listeners into believing that humans played our synthesized examples.

9 Acknowledgements

Special thanks go to William **Chan** Chun Nin, who has been invaluable for recording *dizi* samples and consultation on playing the *dizi.* And, thanks to the RGC Grant # HKUST 6020/02H for funding this research.

References

[1] Risset, J. C., and Matthews, M. 1969. "Analysis of Instrument Tones." *Physics Today*, 22(2): 23-30.

[2] Morrill, D. 1977. "Trumpet Algorithms for Computer Composition." *Computer Music Journal*, 1(1): 46-52.

[3] Horner, A. and Beauchamp, B. 1996. "Piecewise Linear Approximation of Additive Synthesis Envelopes: A Comparison of Various Methods," *Computer Music Journal*, Vol. 20, No. 2, pp. 72-95.

[4] Horner, A. and Ayers, L. 1998. "Modeling Acoustic Wind Instruments with Contiguous Group Synthesis," *Journal of the Audio Engineering Society*, Vol. 46, No. 10, pp. 868-879.

[5] Horner, A. and Ayers, L. 2001. *Cooking with CSound, Part 1: Woodwind and Brass Recipes*, AR Editions.

[6] Horner, A. and Ayers, L. 1998. "Modeling Chinese Musical Instruments," in *Proceedings of the 135th Meeting of the Acoustical Society of America*, Seattle, WA, Vol. 4, pp. 2541-2542.

[7] Horner, A., Ayers, L. and Law, D. 1999. "Synthesis Modeling of the Chinese Dizi, Bawu, and Sheng," *Journal of the Audio Engineering Society,* Vol. 47, No. 12, pp. 1076-1087.

[8] Rodet, X. and Lefevre, A. 1997. "The Diphone program: New features, new synthesis methods and experience of musical use," *Proceedings of the International Computer Music Conference*, Thessaloniki, Greece, 418-419.

Artist Talk: Miriama Young Titlipur

Miriama Young

Music Department – Woolworth Building,
Princeton University, Princeton,
NJ 08540, USA
Ph: + 609 933 4043
Email: myoung@princeton.edu

Abstract

I would like to present an artist talk on a recent performance of my interactive piece for dancer, *Titlipur*. I will present the performance on DVD, accompanied with a discussion of the project itself. My presentation will focus on the demands and aesthetic of performance practice in working with new technologies, the technical construction and musical intentions behind the piece, the experience and observations of working in collaboration with video artist and dancer, and the aesthetic response of the dancer in her interaction with new technology of wearing sensors. I will also discuss the piece in the context of works that utilize such interfaces for musical expression, as a way to stimulate ideas for future research. I will elaborate on the way this piece connects with ideas of the body and its relationship to music and live performance.

Artist Talk: David Sanders Song of the Kokako

David Sanders

School of Music
Victoria University of Wellington
Wellington
E-mail: sanderdavi@scs.vuw.ac.nz

Abstract

Since the 1960s New Zealand composers have shown an interest in working native birdsong into their music. This paper examines the use that various composers – most notably Douglas Lilburn – have made of tape recordings of North Island kokako made by John Kendrick of the Wildlife Division of the Department of Internal Affairs (c.1969). In addition, the paper includes a presentation of a new electroacoustic work of the author, entitled *Kokako Triptych* – a work composed with a full knowledge of, and making deliberate reference to, these earlier compositions.

Artist Talk: Ryan Cockburn

Ryan Cockburn

74 Moana Street
Aramoana
Email: stereojock@yahoo.co.nz

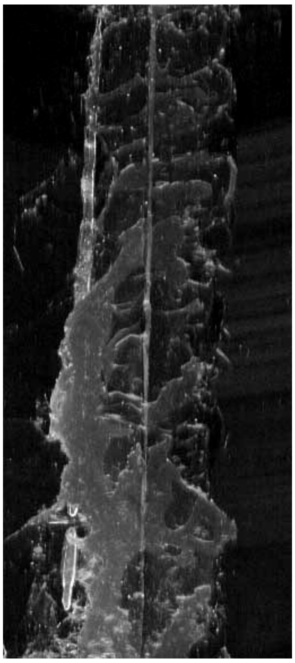
Introduction

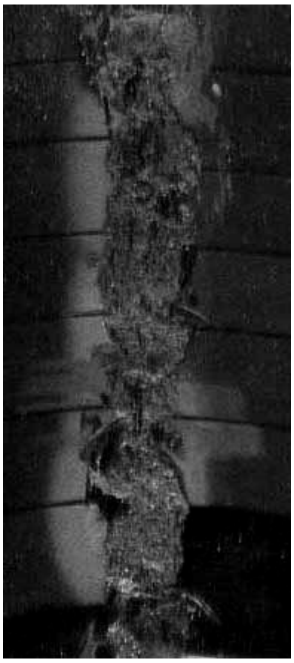
I am in my final year of study towards the completion of my master's degree at The Otago Polytechnic School of Fine Art. My field of study is sound art, with special focus on experimental turntablism and sculptural concepts of aurality.

Before I started my master's degree, I had spent a year painting and touring around the country in a rock band. Since arriving in Dunedin, I have performed many times as a turntablist and progressed towards constructing large sculptural sound objects. My special interest in turntablism stems from a childhood fascination with the medium and as an adult I am trying to find a niche for myself in this field.

The distinctness of the turntablism I practice has it's forebears in artists such as Milan Knizak, Christian Marclay and Janek Schaefer. Unlike conventional turntablism where a mixer is employed to mix between a collection of source material, I reconstruct the records by means of cutting them in different configurations and reattaching them back together. The method of reattachment determines the playability of that particular record.

My first forays into this field involved using different things to affect the journey of the stylus across the record. I tried wax, burning the records with cigarettes, stamping on and scratching them, etc. All of these methods worked in effecting the same recording, however I wanted to construct composite records from a variety of different recordings. At first I tried a variety of bonding methods to reattach the different sides of the records back together. I tried wax and silicone sealant, which didn't bond with the vinyl record, yet created a kind of bridge between them, as seen in the illustrations below. I then began by melting the records together by means of a hot knife, smearing the remnants of other records to fill the space between, as seen below. I would then sand the join to make the record more playable, as seen below. This would affect the sound, creating a rasping sound between the two different sides.

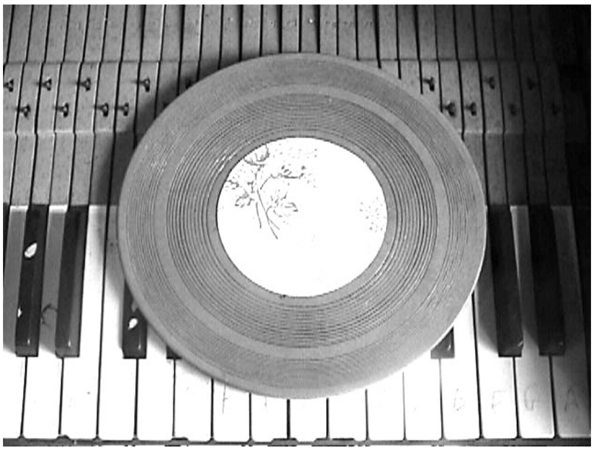




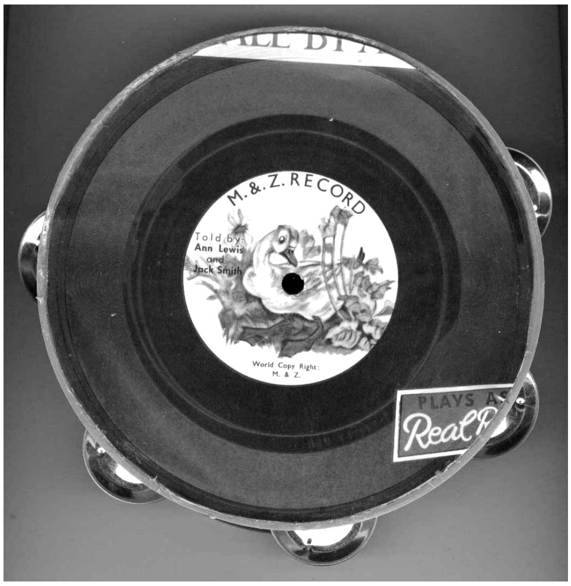
I used and performed with these records for a period and still employ them in my performances, however I became motivated to find a cleaner way of attaching the records together so that I wouldn't have the sound of the join so prominent in the sound produced. I then discovered that I could use a vinyl glue to reattach the pieces. This led to a revolution in my work. It meant that I could now cut and glue any configuration I wanted. I then experimented with circles and the effect on the sound if I placed them within the face of the record. This along with other configurations forms the basis of the source material that I use for my sound performances when I am performing in a strictly aural sense.

I have also made my own wooden and plaster records along with many other adaptations of records that I play with during my performances. Some of these are pictured below.

This is one of the plaster records I made from a mould of the slip-pad an old gramophone unit. Below is a wooden record that I lathed cut.



Below is a piece I made by attaching a record from an old children's story book to the top of a tiny toy tambourine.



I play regularly in Dunedin and have played in Wellington in 2002 as part of the Fringe Festival. I self release recordings of my work and hope to have lathe cut records of my performances created soon. The images I have shown above are meant to be small representation of the work that goes into the sound that you will see on the audition material. I have included these images as the documentation I have provided does not show what goes in inside and on top of the turntables I use.

Artist Talk: The use of Midi-Instruments as performance tools in Jazz and Fusion Improvisation

Colin Hemmingsen & Leigh Jackson

Conservatorium of Music
Massey University
Wellington
E-mail: chem@xtra.co.nz

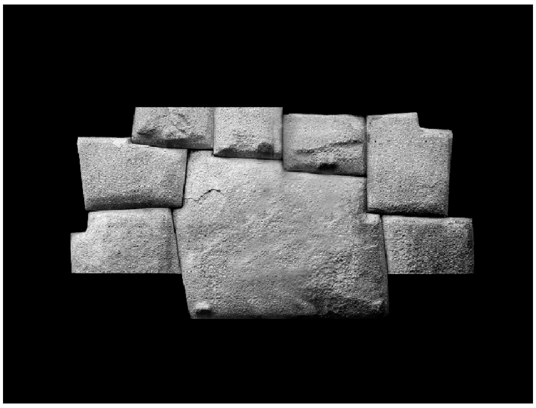
Abstract

This session focuses on midi-instruments and their use in jazz and fusion genres. Colin Hemmingsen and Leigh Jackson represent the Midi-Performance Research Centre recently established at Massey University in Wellington as part of the Conservatorium of Music. The talk will contain demonstrations of the unique functions and extended parameters that the WX-7 wind controller features with emphasis on the use of breath control to create expression. The use of midi-guitar in conjunction with the G4 powerbook to model tube amplifiers and to create solo sounds specifically suited to the guitar will be demonstrated.

Also the polyphonic capabilities of the midi-guitar and how sounds can be layered on separate strings to produce an expressive multi-instrument that can be played in real time. The session will explore the issues encountered specifically with the playing of these instruments in Jazz improvisation and how these were managed in performance during a recent recording project. Issues relating to latency and articulation and how these impact on the improviser's expression will also be discussed.
A short performance will include a practical demonstration of how these instruments function in Jazz and Fusion setting.

Proceedings ACMC 2004
ISSN: 1448-7780
http://www.acma.asn.au

Artist Talk: An experience of sinergy on Latinamerican indigenous instruments and electronic live media: the Orchestra of authoctonous instruments and new technologies of the National University of Tres de Febrero, Argentina



Alejandro Iglesias-Rossi

Universidad Nacional de Tres de Febrero
Buenos Aires, Argentina
E-mail: aiglesias60@hotmail.com

Abstract

Composers in the so called peripheral countries are at the crossroads between finding their own personal identity as creators and their cultural identity as members of a community that encompasses them. The challenge relates to getting to be oneself, discovering one's "uniqueness" in all its potency.

This process not only affects the creator but also influences and transforms the very geoculture he was born into.

The transculturization of elements (as in the case of avant-garde techniques and composing in the classical style of european origin) must be digested, internalized, in order to reappear with a special potency, a unique color that will broaden the fringes of knowledge, as one explores the unknown lands of creation.

This challenge is not only individual and "cultural" but also instrumental and operative, that is to say, it entails the election of the technique and fundamentally the means (the tools) the creator will choose, free from any "a priori", any prejudice that may restrain his visionary capacity.

A number of different subjects must be approached:
a) the antagonism between encyclopedic learning and one of wisdom;
b) not assuming the space one inhabits;
c) admiration for paradigms that are foreign to us, and that can only lead to insatisfaction.
d) a new regard to native instruments, giving them the same "onthological dignity" than european instruments have.

The act of committing oneself to the demands of finding a way to "be and do" that is well-rooted in the time and culture we belong to, erases the supposed dichotomy between contemporary creative techniques and cultural roots.

Finding that path, accepting the challenge, opens us to an unsuspected area of freedom. Composers must travel along this path, this process ripening not only in a personal way but also as members of a community, although in fact, both are the same.

Founded in 2004, the ***Orchestra*** is based on the experience of the Ensemble ***Fronteras del Silencio (The Borders of Silence).*** Created in 2001 by a group of composers-interpreters (Luciano Borrillo, Mariano Fernandez, Daniel Judkovski, Federico Martinez, Julieta Szewach, Daniel Vacs) under the direction of Alejandro Iglesias Rossi; the instrumentation of the Ensemble is an unusual one: 100 chrystal glasses, indigenous Southamerican instruments, prepared piano, percussion, violin, electric guitar and electronic media.

The Ensemble has performed widely in Latinamerica. Its recordings have been broadcasted by Radios around the world (Radio France, BBC Radio, Deutschland Radio-Berlin, RAI 3, Polish National Radio, Australian Broadcasting Corporation, NHK-Tokio, Societe Radio Canada-CBC, Danmarks National Radio, etc).

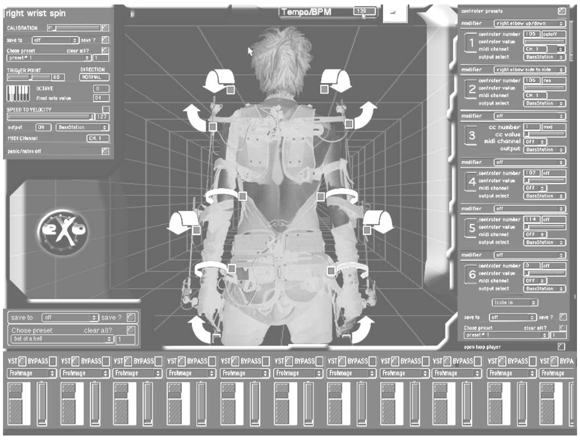
"Every performance of Fronteras del Silencio has an almost ritual character, with primigenial sonorities contrasted with electronic processes and a non-traditional instrumentation."
Radio Beethoven, Santiago de Chile on its XXth Century Program. (Comment on the Ensemble's performance at the XII Chilean Contemporary Music Festival, december 2002)

Studio Report: The ArMiDillo

Exo-sensory body instruments

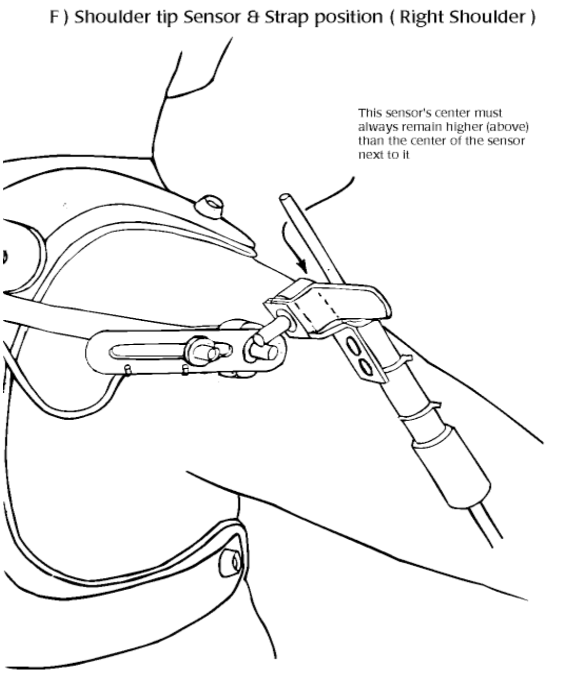
Lissa Meridan

School of Music
Victoria University of Wellington
PO Box 600, Wellington
E-mail: lissa.meridan@vuw.ac.nz



At the close of 2001, I was fortunate enough to spend a couple of months in San Francisco working on a research project. During my time there, I met and worked with Bruce Tupling (Vaporvent Studios) and Nir Bakshy from Animazoo, who were developing a MIDI-suit. The suit is now in production, and I am about to embark on another research trip to workshop the new suit and write the manual for the software.

Animazoo have sent me a demo version of the suit to work with prior to my visit, by which time the new model ArMiDillo will be released.



Exosense is based is San Francisco,CA and Brighton, UK.
info@exosense.net

http://www.exosense.net

"With Exosense movements are music. DJs and producers find a lot of interesting ways to make electronic music performances seem live and interactive, but it is rarely exciting to watch.

Now artists have the advantage of the Exosense, an exo-sensory body instrument, that allows music authoring in real time performance. The Exosense suit translates body movements into sounds, loops, lights and visuals, Completely merging performers and their art. .
Exosense is both visually pleasing and technologically advanced, bringing a futuristic and professional polish to any live performance

Remarkably light and comfortable Exosense suit is modelled on the human skeletal form, using rotational sensors placed on the body joints. The suit simply plugs into a midi interface and body movements send midi signals directly to sound modules and to popular audio sequencing software such as, Logic Audio, Cubase, Live and Reason.

Software included with Exosense enables users to control midi notes, continuous controllers(e.g. cutoff and resonance), pitch bend as well as play scales, trigger loops, samples and visuals. The software is designed with versatility in mind and allows musicians and developers to continue to expand the many possible applications of their Exosense.

Development originally started with the Environment of Emagic Logic™.
A complete working system was created with this robust sequencer and this in some ways provided the ideal 'breadboard' for the functional aspects of the suit and musical performance.

Current development of the latest version of the suit software far exceeds the previous prototypes and the latest build provides a complete solution for the exosense midi system. Using more powerful audio construction software the user definability and suit control has taken on new levels of precision."