

Musical Interaction with Artificial Life Forms: Sound Synthesis and Performance Mappings

James Mandelis and Phil Husbands

This paper describes the use of evolutionary and artificial life techniques in sound design and the development of performance mapping to facilitate the real-time manipulation of such sounds through some input device controlled by the performer. A concrete example of such a system is briefly described which allows musicians without detailed knowledge and experience of sound synthesis techniques to develop new sounds and performance manipulation mappings interactively according to their own aesthetic judgments.

KEYWORDS: artificial life, performance mapping, musical interaction, sound synthesis

I. Introduction

There has been growing interest over the last decade or so in using recently developed technologies from artificial life and the more avant-garde areas of artificial intelligence to provide new ways of generating and manipulating sounds (Miranda 1995a; Griffith and Todd 1998; Bilotta *et al.* 2001). This has opened up very interesting musical avenues where various processes for generating sounds, or whole pieces of music, or for controlling aspects of musical performance, can be thought of in terms of interaction with evolving artificial life forms.

Musical interaction with artificial life forms can be separated into two broad categories: interaction at note level and interaction at sound level. Interactions at note level usually produce complete musical pieces or music fragments made up of notes that comply with accepted musical and harmonic rules described in modern music theory. The interaction at sound level is concerned with the manipulation of parameters that define a sound using a particular sound synthesis technique (SST), or with parameters that define a particular deformation on an input stream (sound effects).

In the first case, the end result is usually constrained by expectations of adherence to a large number of rules that include considerations of structural coherence. In artificial life implementations, this is achieved either by limiting the music formed by the generation process to a set of legal or valid forms or by using a sub-system that checks, or judges, for such validity and rejects pieces that do not conform. In evolutionary terms this can be likened to natural selection. Additional user feedback can be used to steer the course of the evolution, and this can be likened to sexual selection where certain characteristics are transmitted to the next generation by being preferentially chosen by prospective mates.

In the second case of sound design, such rules tend to be either non-explicit or non-existent. This is partly because of the complexity and lack of transparency of SSTs. In this domain the subjective usually rules over the objective with personal aesthetics acting as the only guide. In artificial life implementations, this can be achieved by employing the user's aesthetic judgment to power the evolutionary processes that are used to develop the sound-generating forms (Dahlstedt 2001; Mandelis 2001, 2003; Yee-King 2000; Woolf 1999). This allows for more purist approaches in terms of artificial evolutionary paradigms – that is, it is not necessary to encode domain-specific knowledge (especially aesthetics-based knowledge) to constrain and guide the process. This is not to say that embedding formalised knowledge of this kind is a bad thing, but in an area such as sound design, where aesthetics are very difficult to formalise, the less constrained approach allows for a powerful exploration of sound space, turning up interesting and unexpected new forms that can be put to good artistic use.

As well as applying artificial life techniques to the generation of sounds, which are later used in a performance, it is possible to employ them in the closely related area of developing real-time sound parameter manipulation devices for use in performance. This paper concentrates on the unconstrained, exploratory, use of artificial life evolutionary techniques in these two areas.

II. Instrument Evolution and Performance Possibilities

A very useful framework for thinking about the core themes of this paper is that introduced by Mulder (1994) to describe the classification and development of musical instruments. The first step in instrument development, according to Mulder, involves traditional acoustic instruments that are manipulated in a certain way in order to produce their sounds. The next development is the use of electronics in order to apply sound effects on acoustic instruments. The manipulations remain essentially the same. His comments on the characteristics of these types of instruments are: "Limited timbral control, gesture set and user adaptivity. Sound source is located at gesture" (Mulder 1994: 243). These two steps are illustrated in figure 1.

The next step suggested by Mulder is that of Electronic Musical Instruments, where the essential manipulations of a piano (or other MIDI controllers such as

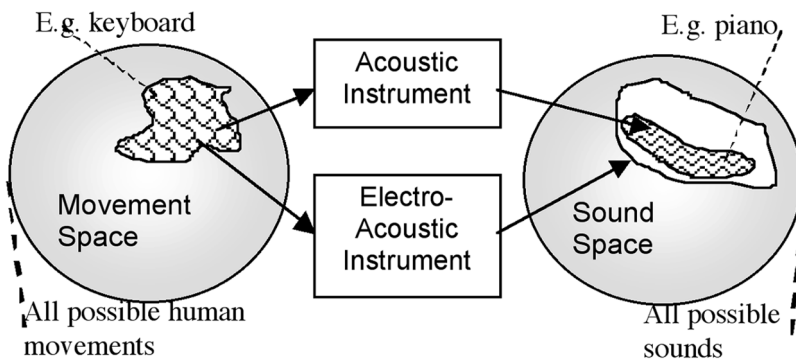


Figure 1
Steps 1 and 2 of instrument development (after Mulder 1994).

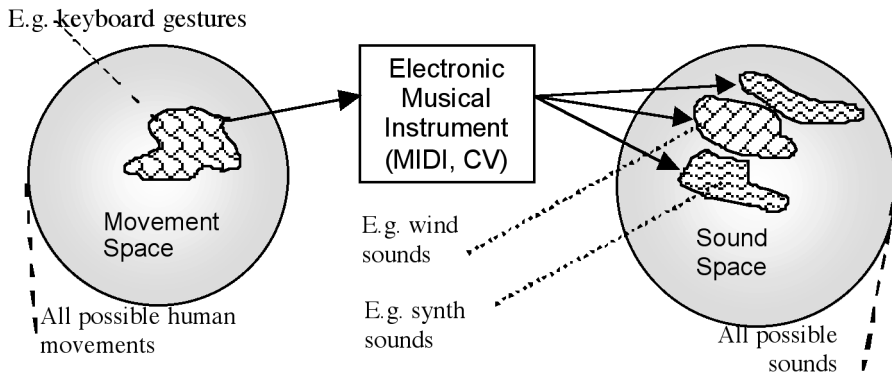


Figure 2
Step 3 of instrument development (after Mulder 1994).

wind, drums and MIDI-guitar) produce sounds that mimic other acoustic or electronic instruments (figure 2). His comments on the characteristics of these types of instruments are: “Expanded timbral control, though hardly accessible in real-time and discretised; gesture set adaptivity still limited. Sound emission can be displaced” (Mulder 1994: 244).

Mulder’s next step, illustrated in figure 3, involves virtual musical instruments (VMIs) where gestures from motion caption devices are used to drive sound engines. His comments on the characteristics of these types of instruments are: “Expanded real-time, continuous timbral control; gesture-set user selectable and adaptive. Any gestures or movements can be mapped to any class of sounds” (Mulder 1994: 245).

As a development of the last step, and as an extension to the overall classification, we suggest a new class. It involves VMIs produced by an artificial-life-based framework for adaptive generation of sounds and their gesture mappings.

Genophone, which is described in more detail in section IV, is an example of a system belonging to this new class of adaptive VMIs (Mandelis 2001, 2002, 2003).

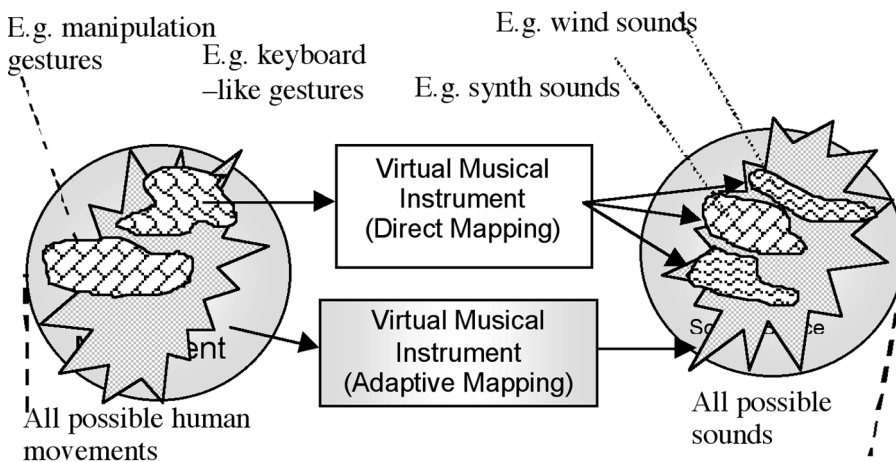


Figure 3
Step 4 (after Mulder 1994) and step 5 (after Mandelis 2001, 2002, 2003) of instrument development.

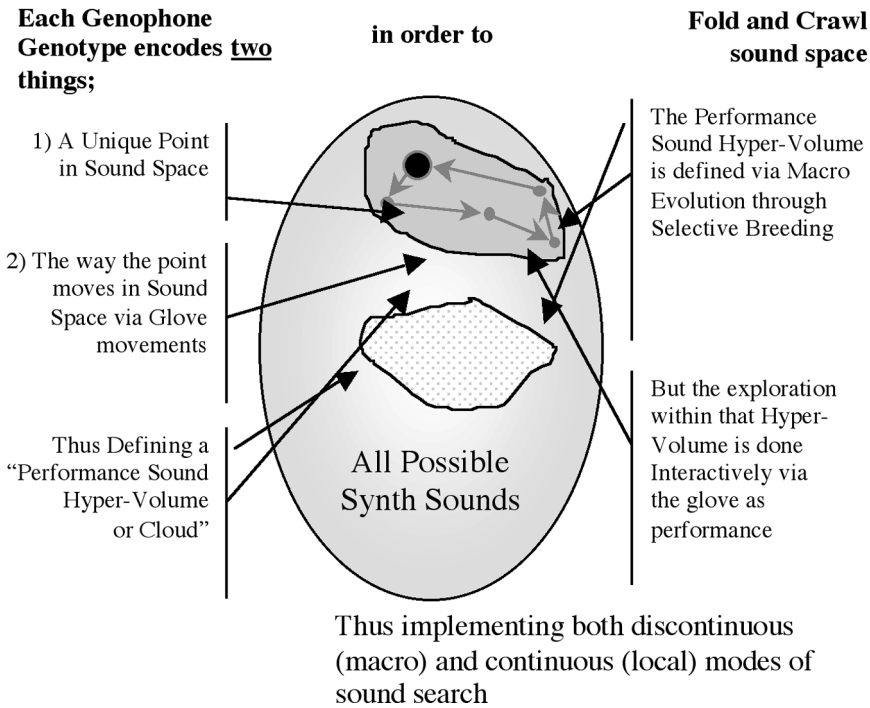


Figure 4
Genophone operation.

It exhibits the following characteristics: (a) expanded real time, (b) continuous timbral control and (c) gesture-set and sounds are user designed via an interactive artificial-evolution-based exploratory search process. Any gestures or movements can be mapped to any class of sounds where both the mappings and the sounds are subject to the same evolutionary forces applied by the user (figure 4).

III. Sound Synthesis and Performance

Music performed with traditional instruments is the production of sounds whose fundamental frequency corresponds to the note played in a given scale. As such, it is normally encoded in a musical score that describes mainly the notes to be played and when they should be played, together with some encoded information describing how these notes are played, for example *legato*, *fortissimo*, and so on. Identical scores can be interpreted in various ways giving rise to unique performances that are separated by the aesthetic values and the skills of the performer. Some of these differences are temporal, as in micro-fluctuations of the note timing (Longuet-Higgins 1982, 1984); others are qualitative as in modulations of intensity or timbre characteristics affected by skilful manipulations of the instrument. Today, with the widespread availability of music sequencers, the differences between the *execution* and *performance* of a piece are more evident than ever. We are all familiar with the mechanical sterile way a musical score can be executed by a computer with perfect (sic) timing and perfect (sic) pitch. Various commercially available systems have been developed that address this problem by intelligently

modulating the timing and the intensity of the notes in accordance with a particular musical style, therefore making a more live-sounding and pleasing musical performance.

This paper focuses on those aspects of musical performance differences that are not encodable in a traditional score, especially in the possibilities of novel expressivities provided by synthesisers and their exploration with artificial life paradigms. For a long time, synthesisers have been used to emulate traditional instruments and as such they sport pitch-bend and modulation wheels that aid in the expressivity of the instrument (Miranda 2002). Other parameters of the SST employed can be modulated by knobs and sliders, giving rise to the now widely accepted practice of “knob-twiddling” (especially in recent generations of musicians). Music makers have discovered, through trial and error, aesthetic values that can be expressed in a way that was not possible before: through the modulation of SST parameters. These new expressivities are circumscribed by the SST parameters available for real-time manipulation. Although individual SST parameters are often used for expressivity purposes, it is possible to manipulate multiple values simultaneously. Thus by varying an input parameter (i.e. knob, slider or other control device) a number of SST parameters can be simultaneously controlled, thus defining a meta-SST parameter. At this low-level stratum of performance possibilities there is no accepted way or model of how parameter changes can be implemented, as opposed to at the note level where well-established theories, models and rules are in place.

A particular timbre can be defined as a point in a P -dimensional parametric space, where P is the number of parameters used by the SST engine that produces that timbre. A musical performance can be thought of as an aesthetically pleasing trajectory (or set of trajectories) within that parametric space. For instance, if one of the parameters is the main oscillator frequency, then playing a monophonic melody can be thought of as moving the timbre’s point back and forth along that parameter dimension in intervals defined by the scale used. This particular parameter would normally be controlled by the keyboard key position (or equivalent); other parameters do not have such usage expectations associated with them but they can also be used to aid expressivity. Essentially, the problem is one of mapping a number of input parameters (I) (i.e. sliders, knobs, etc.) to a subset (S) of the total number of SST parameters (P), where $I = S = P$ (Krefeld 1990; Pressing 1990; Choi *et al.* 1995; Rován *et al.* 1997; Wessel and Wright 2000). If each controlled SST parameter has a unique relationship to an input (performance) parameter then a performance subspace is circumscribed within the parametric space, within which an I -dimensional trajectory can be defined as a performance if it satisfies some arbitrary aesthetic sensibilities. This mapping in effect defines an instrument with unique timbral characteristics and expressive manipulative behaviour: a virtual musical instrument (Machover and Chung 1989; Mulder 1994; Mulder *et al.* 1997; Wessel and Wright 2000).

IV. *Genophone*

To design the kinds of mappings and timbres described in the previous section is a complex and lengthy affair (Dahlstedt 2001); it involves an intimate knowledge of the SST involved that can be gained usually only after years of experience with the particular SST. *Genophone* (Mandelis 2001, 2002, 2003) is a system that has been

developed to facilitate the design and exploration of such virtual instruments without the need for such detailed knowledge and experience. It uses an artificial life paradigm in order to breed VMIs and their control mappings. A data-glove is used as an additional control device that provides five independent input control parameters that can be modulated simultaneously; also the normal performance control parameters are used, i.e. keyboard, velocity, after-touch, pitch-bend and modulation wheels. The system is shown in figure 5.

During a typical run of *Genophone*, two (or more) hand-designed VMIs are used as seeding parents, these then create a generation of offspring through the application of one (of several) genetic operators. Crossover operators mix parameter values from the two parents to create a new individual; mutation operators randomly change the value of one or more parameters encoded on an individual. After being previewed by the user, the offspring are assigned a relative fitness reflecting how much they are liked by the user. The previewing process involves a fragment of performance so that the user can experiment with the sounds, and the (glove) gesture mapping for manipulating them, that are encoded on the offspring in question. This fitness is used by some of the genetic operators to bias the resulting offspring towards the fitter members of the population. The new generation of offspring is then previewed by the user and a number of them are again selected as parents to create the next generation. Additionally it is possible to allow some other hand-designed parents to enter into the breeding process and contribute towards the next generation. This cycle continues until one or more individuals are deemed satisfactory as VMIs. This process is illustrated in figure 6.

Genophone has demonstrated that this technique is relatively quick and painless compared with any hand-design method, and that the breeding paradigm is a simple and intuitive one to grasp, while being very powerful. In practice, aesthetically interesting and useable VMIs are generated after a few cycles of the algorithm.

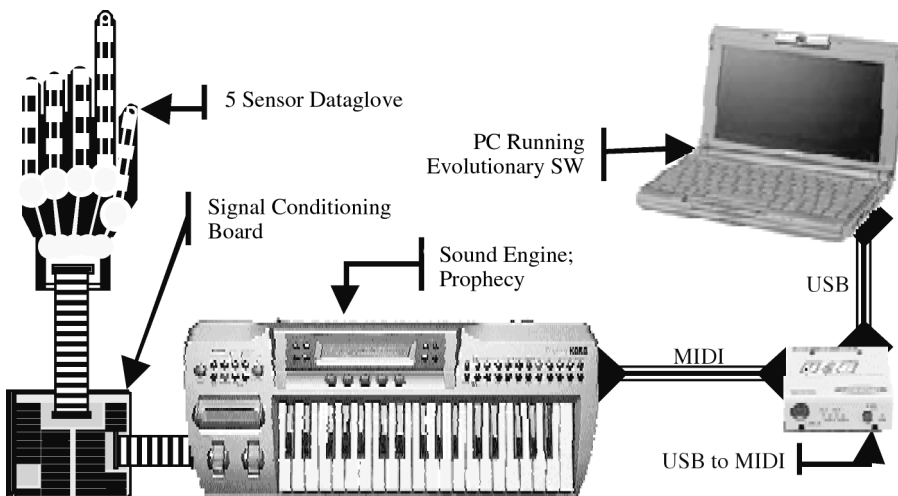


Figure 5
Genophone system.

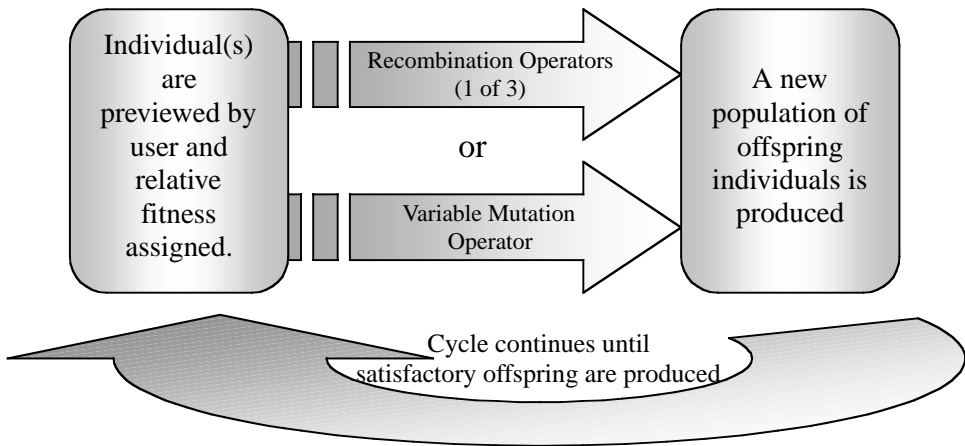


Figure 6
Genophone's evolutionary cycle.

V. Related Work

Perhaps the earliest application of artificial life techniques to sound design was Miranda's *Chaosynth* system, which dates from the early 1990s (Miranda 1995b). Here a cellular automaton was used to control some of the parameters used in a granular synthesis technique where a sound event is built from a series of micro sound bursts (granules). A cellular automaton is a grid of cells whose individual states (which might be switched on or off, or something more complicated) change every cycle according to some rule that takes into account the values of all neighbouring cells. In the case of *Chaosynth*, values emanating from particular regions of the grid are used to control the frequency and duration values for the individual granules used to make up the sound. Johnson (1999) later used an interactive genetic algorithm to explore the sound space afforded by granular synthesis techniques.

Recently, Dahlstedt has independently used an interactive evolutionary process, similar in outline to that employed in *Genophone*, to design sounds by manipulating the parameters of the underlying sound-generation engine (Dahlstedt 2001). This has been done in a generic way so that the system can be customised to operate with almost any hardware or software sound-generation engine. Dahlstedt points out that, as well as allowing musicians to design sounds without needing expert SST knowledge, evolutionary systems of this kind open up compositional possibilities based on "new kinds of structural relationships" which occur because "the sounds created during a breeding session are often audibly clearly interrelated" (Dahlstedt 2001: 241).

Evolutionary systems have also been used for less exploratory kinds of sound design. For instance, Garcia (2001) has developed methods to apply evolutionary search to the design of sound synthesis algorithms and demonstrated the efficacy of his approach by evolving various target sounds, including notes played on a piano, through the use of an automatic fitness function that measured how close the generated sound is to the target sound.

McCormack's *Eden* is another interesting example of a related application of

artificial life techniques in music (McCormack 2003). In this system, agents populate an artificial world in which they can move around and make and hear sounds. These sonic agents must compete for limited resources in their environment, which is directly influenced by the artwork's audience. The agents generate sounds to attract mates and, because of the influence of the audience on the virtual environment, particularly the growth rate of virtual food, to attract the attention of the audience. In this work McCormack has demonstrated the successful use of an open-ended automatic evolutionary process to generate a highly engaging interactive artwork.

VI. Conclusions

This paper has discussed the use of evolutionary artificial life techniques for the interactive exploration of sound-space and its extension to virtual musical instrument space. A concrete example of a system that has successfully demonstrated the efficacy of the approach has been briefly described. It has been argued that artificial life techniques can open up new creative and aesthetic possibilities.

References

- Bilotta, E., Miranda, E. R., Pantano, P. and Todd, P. (eds) (2001) *Proceedings ALMMA 2001: Artificial Life Models for Musical Applications Workshop, ECAL 2001*. Consenza, Italy: Editoriale Bios.
- Choi, I., Bargar R. and Goudeseune, C. (1995) "A manifold interface for a high dimensional control interface". In *Proceedings of the 1995 International Computer Music Conference*, pp. 385–392. Banff, Canada: ICMA.
- Dahlstedt, P. (2001) "Creating and exploring huge parameter spaces: interactive evolution as a tool for sound generation". In *Proceedings of the 2001 International Computer Music Conference*, pp. 235–242. Havana, Cuba: ICMA.
- Garcia, R. (2001) "Growing sound synthesizers using evolutionary methods". In *Proceedings ALMMA 2001: Artificial Life Models for Musical Applications Workshop, (ECAL 2001)*, ed. Eleonara Bilotta, Eduardo R. Miranda, Pietro Pantano and Peter Todd. Consenza, Italy: Editoriale Bios.
- Griffith, N. and Todd, P. M. (eds) (1998) *Musical Networks: Parallel Distributed Perception and Performance*. Cambridge, MA: MIT Press/Bradford Books.
- Johnson, C. (1999) "Exploring the sound-space of synthesis algorithms using in interactive genetic algorithms". In *Proceedings of AISB'99 Symposium on AI and Musical Creativity*, ed. Angelo Patrizio, Geraint Wiggins and Helen Pain, pp. 20–27. Brighton, UK: AISB.
- Krefeld, V. (1990) "The hand in the web: an interview with Michel Waisvisz". *Computer Music Journal* 14(2), 28–33.
- Longuet-Higgins, H. C. (1984) "The rhythmic interpretation of monophonic music". *Music Perception* 1(4), 424–431.
- Longuet-Higgins, H. C. and Lee, C. S. (1982) "The perception of musical rhythms". *Perception* 11, 115–128.
- McCormack, J. (2003) "Evolving sonic ecosystems". *Kybernetes: The International Journal of Systems & Cybernetics* 32(1/2), 184–202.
- Machover, T. and Chung, J. (1989) "Hyperinstruments: musically intelligent and interactive performance and creativity systems". In *Proceedings of the 1989 International Computer Music Conference*, pp. 186–190. Columbus, USA: ICMA.
- Mandelis, J. (2001) "Genophone: an evolutionary approach to sound synthesis and performance". In *Proceedings ALMMA 2001: Artificial Life Models for Musical Applications Workshop (ECAL 2001)*, ed. Eleonara Bilotta, Eduardo R. Miranda, Pietro Pantano and Peter Todd, pp. 37–50. Consenza, Italy: Editoriale Bios.
- Mandelis, J. (2002) "Adaptive hyperinstruments: applying evolutionary techniques to sound synthesis and performance". In *Proceedings NIME 2002: New Interfaces for Musical Expression*, pp. 192–193. Dublin, Ireland.
- Mandelis, J. (2003) "Genophone: Evolving Sounds and Integral Performance Parameter Mappings".

-
- In *Proceedings of EvoWorkshop 2003*, pp. 535–546. Lecture Notes in Computer Science, Vol. 2611. Springer.
- Miranda, E. R. (1995a) “An artificial intelligence approach to sound design”. *Computer Music Journal*, 19(2), 59–75.
- Miranda, E. R. (1995b) “Granular synthesis of sound by means of cellular automata”. *Leonardo* 28(4), 297–300.
- Miranda, E.R. (2002) *Computer Sound Design: Synthesis Techniques and Programming*, 2nd edn. Oxford: Focal Press.
- Mulder, A. (1994) “Virtual musical instruments: accessing the sound synthesis universe as a performer”. In *Proceedings of the First Brazilian Symposium on Computer Music*, pp. 243–250. Caxambu, Brazil.
- Mulder, A., Fels, S. S. and Mase, K. (1997) “Mapping virtual object manipulation to sound variation”. *USA/Japan inter-college computer music festival – IPSJ SIG Notes*, 97(122), 63–68.
- Pressing, J. (1990) “Cybernetic issues in interactive performance systems”. *Computer Music Journal* 14(1), 12–25.
- Rovan, J. B. Wanderley, M. M., Dubnov, S. and Depalle, P. (1997) “Instrumental gestural mapping strategies as expressivity determinants in computer music performance”. In *Proceedings of the AIMI International Workshop – Kansei, The Technology of Emotion*, ed. A. Camurri, pp. 68–73. Genoa, Italy: Associazione di Informatica Musicale Italiana.
- Wessel, D. and Wright, M. (2000) “Problems and prospects for intimate musical control of computers”. *Computer Music Journal* 26(3), 11–22.
- Woolf, S. (1999) “Sound Gallery: An Interactive Artificial Life Artwork”. MSc thesis, School of Cognitive and Computing Sciences, University of Sussex, UK.
- Yee-King, M. (2000) “AudioServe – An Online System to Evolve Modular Audio Synthesis Circuits”. MSc thesis, School of Cognitive and Computing Sciences, University of Sussex, UK.