ISSN 1744-1986

Technical Report N° 2004/22

Electronic Sound Synthesizer Interface Design

Allan Seago

11th June 2004

Department of Computing
Faculty of Mathematics and Computing
The Open University
Walton Hall,
Milton Keynes
MK7 6AA
United Kingdom

http://computing.open.ac.uk

ELECTRONIC SOUND SYNTHESISER INTERFACE DESIGN

Allan Seago
The Open University
Walton Hall
Milton Keynes
MK7 6AA
a.seago@londonmet.ac.uk

ABSTRACT

The focus of this project is to address a deep-rooted problem in the user interface of present-day hardware and software of electronic music synthesizers. While sound qualities such as pitch and loudness lend themselves relatively easily to user control, the timbre of sound is much less easily described or defined in any systematic way. Consequently, in all present day user interfaces so far examined for controlling or modifying timbre in electronic synthesisers, there is a wide semantic gap between the task and system languages used. In this project, we are carefully analysing current approaches to controlling timbre, and investigating principles, issues and problems that must be addressed to improve the design of user interfaces for sound synthesis. In subsequent stages of the project, a prototype of a new kind of interface will be designed, implemented and evaluated.

The research is still in its early stages, but a taxonomy of synthesizer interface design categories is identified and the conclusions drawn from a series of user tests summarised.

Keywords

Timbre, Synthesizer, Music, Usability, Semantic directness, innovative user interfaces.

1. CONTEXT

The problems presented in the design of effective interfaces for sound synthesis have been noted by a number of researchers working in electroacoustic music [11] [13], computer science [7], audio technology [1] [3] [12] and Human Computer Interaction [4] [6] [8]. While pitch and loudness can be readily controlled, the specification of more intangible qualities such as sound colour, texture and articulation (timbre) is more difficult. There is a considerable semantic gap between the vocabulary typically used by musicians to describe sound, (the user domain), which may draw on a lexis of emotional association or analogy with colour or texture, and the engineering terms required for sound synthesis (the system domain).

Similarly to the way in which facial recognition is rarely attributable to a single facial feature, but rather to a complex matrix of facial characteristics, verbal descriptions of sound rarely correspond to a single and easily defined set of acoustical attributes. Consequently, an effective user interface designed to enable sound creation and modification from user directives needs to provide a layer which mediates between these two levels.

However, the user interface of the typical commercial synthesiser tends to be very much expressed in system (i.e. engineering) terminology, obliging the user to become fluent with the synthesis method employed. The problem of designing a more user-oriented user interface has been approached in a number of ways: using techniques drawn from artificial intelligence [5]; knowledge based systems [1,2]; and by employing metaphors derived from physical models of acoustical mechanisms [12]. Artificial intelligence and knowledge-based approaches to mapping user and system domain languages have focussed on automatic timbre classification strategies or acquisition of synthesis values by parameter extraction. These techniques can be very helpful where a sample exists of the desired timbre, but they do not help very much where the user is trying to specify a mentally imagined timbre.

Physical models have been very successful in providing the means of emulating 'real' instruments. However, as has been pointed out by Wessel, Risset and others, cited in [2], this particular model restricts the musician precisely because of the physical world metaphor. An imagined sound for which the musician can find no physical analogue is clearly one that cannot be easily realised using this metaphor.

2. GOALS

The aims of the thesis are

- To present, using various evaluation techniques, critical analyses of user interfaces used in software and hardware synthesisers currently commercially available.
- To investigate, design and develop a software based user interface, which will enable the creation of a 'sound object', the sonic characteristics of which are determined by the user. The term 'sound object' is defined here as a pitched or non-pitched sound whose sonic characteristics may or may not evolve over time, but which nevertheless can be regarded as a perceptually unified object – for example, a single note played on a clarinet, the sound of a plucked string, a spoken word, or fragment of a word, the sound of breaking glass. Such an interface will enable a user to create a sound object of moderate

complexity, in terms of its spectral composition and its temporal evolution, using a user/system dialog which employs a search algorithm to guide a user through an n-dimensional timbral space. The approach taken here will be different from that taken by other researchers in the field. There exists a considerable body of research which seeks to devise systems that bridge the semantic gap already mentioned in the synthesiser user/system interface, using, in some cases, expert system or neural net technologies. The aim of the research proposed here, however, would be to avoid as far as possible the use of a verbal or symbolic layer in the interface; the user will instead engage with the sound itself as it evolves.

- The extent to which this goal has been successfully attained will be determined by the conducting various evaluations and user tests.

3. STATUS

The research as a whole is still in its early stages. The nature of the enquiry is essentially cross disciplinary, taking in HCI, psychoacoustics and computer science; the literature survey of current work in this area consequently needs to be cast quite widely. This work is continuing.

A short paper [9] presenting a critical analysis of synthesizer user interfaces for timbre has been submitted to HCI 2004.

A series of user tests has been conducted [10] in which users were asked to complete three ‘sound creation’ tasks on two different commercially available synthesizers, and the approach and steps they took were recorded and analysed.

In order to obtain empirical data which will guide the design of a new user interface, a number of listening tests are to be conducted to ascertain the degree of correlation between the parameters of a set of predefined sounds, and listeners’ perceptions of those sounds. The design and compilation of these listening tests is now well under way; an initial pilot will be conducted over the summer, and it is intended to perform them in September, 2004.

4. INTERIM CONCLUSIONS

A number of interim conclusions can be offered, which are listed here.

A taxonomy of synthesiser user interface types, based on the model of sound presented to the user, has been identified [9]. A distinction is made between user interfaces which allow visual representations of sound to be manipulated more or less directly (Direct Specification), and those that allow the manipulation of an architectural structure (User Specified Architecture), or the parameters of such an architecture (Fixed Architecture), which generates the sound. However, none of the core languages involved have been found to map appropriately to the task language of the musician.

The user tests [10] showed that successful negotiation of the interfaces of all three user interface types depends to a great extent on, firstly, the domain knowledge that the user brings to the task, and, secondly, the degree of practical experience that the user has of similarly configured instruments. Those users with experience of other synthesisers were able to transfer and apply this in their strategies for completing the test tasks. All users, however, showed a definite preference for a user interface which incorporated elements of Direct Specification, as opposed to a user interface which was of the Fixed Architecture type.

5. REFERENCES

[1] Ashley R (1986). A Knowledge-Based Approach to Assistance in Timbral Design. Proceedings of the 1986 International Computer Music Conference

[2] Chadabe, J (1997). Electric Sound. Prentice Hall.

[3] Ethington R. and Punch B. (1994) SeaWave: A System for Musical Timbre Description, Computer Music Journal 18:1 pp 30-39.

[4] Gaver, W. W. (1993). Synthesizing Auditory Icons. ACM INTERCHI '93, ACM.

[5] Miranda E. R. (1995). An Artificial Intelligence Approach to Sound Design. Computer Music Journal 19(2): 59-75. MIT Press.

[6] Mynatt, E. D. (1994). Designing with Auditory Icons. CHI '94, Boston, Massachusetts, ACM.

[7] Nicol, C. Brewster, S.A. Gray, P.D. (2004). A system for manipulating auditory interfaces using timbre spaces. CADUI 2004 (Jacob, R., Limbourg, Q. and Vanderdonckt, J., Eds.). ACM.

[8] Ruffner J. W. and Coker G. W. (1990) A Comparative Evaluation of the Electronic Keyboard Synthesizer User Interface, Proceedings of the Human Factors Society, 34th Annual Meeting Human Factors Society.

[9] Seago A, Holland S and Mulholland P. (2004). A Critical Analysis of Synthesizer User Interfaces for timbre. Paper submitted to HCI2004: Design for Life.

[10] Seago A. (2004). Analysis of the synthesizer user interface: cognitive walkthrough and user tests. Technical report TR2004/15, Dept of Computing, Open University.

[11] Slawson, W (1985). Sound Color. Berkeley: University of California Press.

[12] Smith J. O. (1992). Physical modeling using digital waveguides. Computer Music Journal, vol. 16 no. 4, 74-91.

[13] Wishart, T (1996) On Sonic Art Harwood Academic Publishers.