

Design and use of some new digital musical instruments

Daniel Arfib, Jean-Michel Couturier, Loïc Kessous

LMA-CNRS
31, chemin Joseph Aiguier
13402 Marseille Cedex 20
[arfib,couturier,kessous]@lma.cnrs-mrs.fr

Abstract. This article presents some facts about the use of gesture in computer music, more specifically in home made instruments dedicated to performance on stage. We first give some theoretical and practical ideas to design an instrument, in the following areas: the sound, the gesture and the mapping. Then, we introduce three examples of digital instruments we have created, focusing on their design and their musical use.

1 Introduction

Creating a digital musical instrument is complex and some skills and methods can help the designer. During our research project about creative gesture in computer music, we have designed some particular instruments, now largely described elsewhere [2] [3] [4] [12] [13] [17]; this article introduces our global approach to design digital musical instruments.

In a first section, we describe our way “from sound to gesture” by giving some theoretical and practical ideas to design an instrument in the sound, the gesture and the mapping areas. In a second section, we describe the design and the use of three digital instruments we have created, the “Voicer”, the Scanned Synthesis Synthesizer, and the Photosonic Emulator.

2 Theoretical and practical ideas to design an instrument

2.1 Ingredients for sound space exploration

To create a digital musical instrument, two different options are possible; the first one is to conceive the instrument from gesture to sound, in other words to sonify a gesture, or a gestural fact. Concerning the second option, creative people such as composers usually start from the algorithms, explore their possibilities, imagine their expressive language and associated gestures, and only after choose the adequate peripherals. The definition of a sonic space and its exploration is equivalent to build

an “inverse chain”: the creative imagination of sound induces the determination of the instrumental play, and so induces adapted gesture and peripherals that support them. We try to be close to this idea.

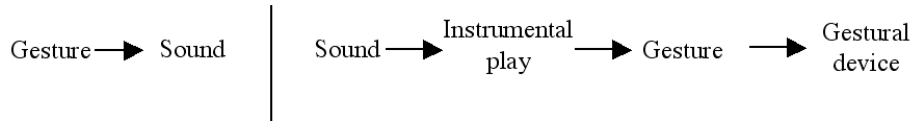


Fig. 1. regular (on the left) and inverse (on the right) chain between gesture and sound.

Gesture can intervene at different levels. At the global composition level, gesture drives the “macroform”, the global structure of a piece; at the sequence level, the sequences of notes are driven by gesture; at the ‘notes’ level, decision or selection gestures trigger events (abrupt changes in parameters); at the ‘curves’ level: modulation gesture drive curves (microforms) which are related to sound parameters; at the microscopic level, timbre is controlled by statistic or stochastic processes of micro events; finally, at the quantum level, clicks or dirac impulses are directly controlled. Our choice is to work at the level of notes and curves. This way is in continuity with a long work on synthesis models that use curves and triggering events. On the other hand, our work on gesture has led us to define selection, decision and modulation gestures, a natural way to control such events and curves.

Finally, design choices depend upon three elements: sound, gesture, instrumental play, which respectively correspond to the synthesis methods, the choice of peripherals and the type of mapping. The choice of a synthesis method or a digital audio effect is crucial in the design because it indicates the gesture metaphors. Gestural devices imply an ergonomic dimension that conditions the gesture. The choice of a mapping is the key point to obtain a musical instrument that one can learn and play: it transforms the information of the gestural device to the synthesis parameters but moreover it gives some coherence in the link between what we hear and what we do.

2.2 About sound

2.2.1 Definition of a sonic domain

In this path from sound to gesture, the first question is: what type of sound or what type of music does one want to play; what are the limits one imposes? These decisions will lead to the choice of a synthesis method and give an idea of the type of sensors that can do what is needed or not.

Sounds can be classified according to their means of producing, but also according to the impact they have on perception [24]. These sound typologies are often dichotic, and it is very interesting to put a name on what kind of sounds/music we want in terms of these typologies. As a matter of example, the separation between harmonic/inharmonic sounds is natural and a musical choice must rely on such

Design and use of some new digital musical instruments

categories. However, such a division is only a subclass of the noisy/timbral dichotomy, on the right side of this branch.

The determination of how the sound evolution will be treated is a matter of choice and complementarity. If the playing style is melodic, pitch must be in some way a part of the focus of the instrument with its different ways to control it (continuous or not). If sound timbre is the main focus, one must find the way to describe the expressiveness that one requires and how it can be brought out.

2.2.2 Synthesis methods

Synthetic sounds have fascinated the pioneers of computer music, who tended to find sounds that could sound atypical but quite musical. A good source to discover the computer music sounds are books such as the Csound book [6] or Miranda's inspiring one [22], records such as the Wergo series (including JC Risset's catalog [27]) and web sites where one can find patches that are compilable [15].

In wavetable synthesis, a wavetable is defined, and it is possible to read it with a variable index and to produce a periodic sound. The timbre of the sound obtained is fixed. The addition of jitter (small variations of random noise) or vibrato (a slow sinusoid) on the drive of frequency can make the sound alive [27].

Additive synthesis consists in adding partials having frequency and amplitude envelopes. Analysis techniques can also derive harmonic envelopes, such as the heterodyne filter [6] [23] or SMS.

In subtractive synthesis, a source-filter model is used, where the source is usually coming from a simple synthesis (waveform synthesis with regular shapes such as square, sawtooth, triangle, etc or noise) and the evolution of the sound comes from the filter. This method is successful because one recognizes the voice formants (peaks in the spectrum due to vocal tract resonances in speech). Analog synthesizers make a general use of filters and silent instruments (silent cello, violin, piano) use banks of resonant filter to simulate usual and unusual harmony boards [19] [26].

Initiating the FM synthesis, John Chowning [11] demonstrated that two oscillators associated in a modulation frequency scheme give rise to ray spectra in which partials vary with the modulation index. Non-linear distortion also named waveshaping is another modulative synthesis which has been worked out independently by D. Arfib and M. Lebrun [1] [18]. Distorting a sinusoid through a waveshaping function creates a harmonic spectrum. This spectrum evolves with the input amplitude of the sinusoid and gives an easy entry point for gestural control [2].

Granular synthesis consists in adding grains in order to obtain granular sounds [28]. The two most important parameters for the sound are the position of the grains and their shape. These grains can be replications of a mother-grain or a sample of recorded sound, dilated or compressed in time. In granular synthesis, many control functions are at a macrocontrol level, like synchronicity or random injection [29].

2.2.3 Digital audio effects

The limit between synthesis and transformation is very thin: for example, in a vocal synthesis with formants, one can replace the first part (source synthesis) by the reading of a natural sound and keep the filtering action and its control, making this way an electronic equivalent of the talk box [4]. Transformation techniques have been

classified this way in the DAFx book [38]: filters, delays, modulators and demodulators, nonlinear processing, spatial effects, time-segment processing, time-frequency processing, source-filter processing and spectral processing. One naturally finds the filtering and distortion (but here applied to sounds that are not sine waves), diverse granulation styles, time compression/expansion and purely spectral transformations using analysis-synthesis methods. The matter of controlling effects with gesture can be found in [32] and [30]; it is also possible to make instruments where the heart is a digital effect. Adaptive effects use a feature extraction to give or modify control parameters, or even the mapping of the effect [33].

2.3 About gesture

2.3.1 Think different about gesture

Working with digital sound and gesture is efficient if we can keep a symbolic link between the act and the resulting sound. This symbolism can be unfolded in two levels: the act itself as a function, and the movement as an expressive gesture.

As far as musical play is concerned, two strategies can be taken. In the first one, one can imitate the musician's play movements on traditional instruments: as an example, using the two hands on a guitar for pitch and trigger induces the notion of bimanuality, which is fruitful in the design of digital instruments. In the second attitude, one explores the universal symbolism of gesture, and discovers the intention behind a gesture and tries to extract it in a sensitive way.

Gestures must have a dynamic appearance compatible with human possibilities in terms of space and speed, and allow to touch the limits that one requires for the sonic domain; gesture sensors by themselves have idiosyncrasies, like the heaviness of some equipment or the need for a specific light environment for cameras.

Musical or expressive general gestures have been classified by some authors. As an example, three functions for the "gestural channel" have been defined by Cadoz [9], [10]: semiotic, ergotic, and epistemic. See also [14].

2.3.2 Sensor typology

The typology of sensors has been largely discussed in the field of Human Computer Interaction and even in the field of sound control. Paradiso [25] makes an historic and practical view of controllers; Wanderley [37] gives a survey of different methodologies for the conception of gestural interfaces and an evaluation of some alternative interfaces. Some classifications as equipped/non-equipped, continuous/discrete provide a great help in the conception stage of an instrument.

Sensors that are usable in gesture controlled musical instruments are of three types. Sensors that imitate acoustic ones, such as breath controller or keyboard, allow to sense expert gestures that the performer makes. These sensors can control the same acoustic changes as the original instruments. As an example, a keyboard is used to control pitch and amplitude in most commercial synthesizers. These sensors can also manipulate other exotic parameters, and this is more specifically powerful for spectral processing. Hybrid controllers allow to play acoustic instrumental sounds and they

also have other controls to modify other parameters. Alternate controllers are specifically made for new instruments or diverted from other non-musical applications. They range from as simple as linear or circular potentiometers or buttons, to video cameras and force feedback sensor. All of them enable musicians to step in new modes of artistic expression.

2.4 About mapping

Mapping binds the gesture to the sound, connecting the gesture data and the synthesis parameters. Constituting a bridge between the action and the perception, the mapping represents the essence of a digital instrument. Several approaches have been used [16] [36] [37]; we have developed a strategy at LMA that we explain in details in [5]. This strategy is based on a multilayer mapping using perceptual spaces. We have defined three layers above two perceptual spaces. These layers are: from sound perceptual space to synthesis parameters, from gesture data to gesture perceptual space and between the two perceptual spaces.

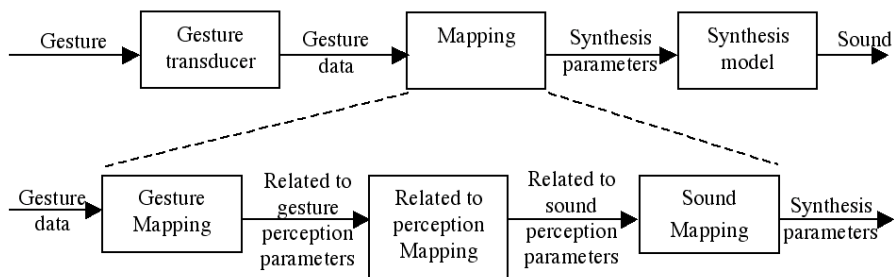


Fig. 2. The Mapping Chain can be divided in several steps with perceptual spaces.

The aim of this structure is having a higher flexibility and modularity in the realisation of a musical instrument: the definition of high-level parameters from gesture data and from sound synthesis parameters is realized upstream, and the mapping between the two perceptual spaces becomes very simple.

Independently of this multi-layer approach, mapping can have different qualifications: explicit/implicit, simple/complex or static/dynamic. The difference between static and dynamic is whether synthesis parameters values depend only on the actual gesture parameters, or whether there is a relation with past values [21].

Other details about practical questions on mapping can be found in [33], dealing with the use of adaptive effects.

2.5 About feedback

Feedback uses different modalities: sight, touch, hearing; to use feedback, the essential point is the coherence between the action and the perception. Obviously, the auditory feedback is the produced sound and is essential in music: generally,

musicians are hearing the sound they produce. The tactile feedback can be passive (physical behavior of the device) or active, like force feedback [8], whether the device reacts to the gestures or can be controlled. The visual feedback has the advantage of spatial discrimination that helps the user to dispose of much information simultaneously. The sight can be used in an informative level to help the performer, as for displaying the values of some parameters, but can also be part of the interaction with the instrument, like with graphical user interfaces.

3 Creating Instruments

3.1 Making Modular Instruments

One of the most important topics in the creation of our digital instrument is to use modularity: for a given synthesis method, changing a gesture controller with another has to be possible and easy. One can also use the same gesture device with different synthesis techniques. The mapping choices will decide if the instrument is modular or not. Our three layers model of mapping, with perceptual spaces, was created to reinforce modularity in digital instrument design. The key point is the middle space, which is the space to create instrument. The simpler this space, the easier is the creation of an instrument.

As shown previously, musical gestural devices can be classified into three types: the controllers that are imitation of acoustical instruments, the hybrid controllers and the alternative controllers. Our study was deliberately focused on alternative controllers. The reason is we are going from sound to gesture and we want to give to the controllers an expressiveness in accordance with the actions we want to apply on the sound. Imitative or hybrid controllers are often imposing the expressivities of their models. Starting with basic data (position, pressure, speed, ...) enables to obtain the expressiveness we want to control our sound processes. The acoustical instruments have strong physical constraints on the gesture sensing, due to the intrinsic link between gesture and sound production; the alternative controllers are free from sound production.

The musical instruments we create use a computer with software program, Max/MSP [20], which is a modular environment to realize virtual instruments and play them in real time. Many today software programs enable real time synthesis and many of them provide a certain kind of modularity at different levels. Some of them, like Max-MSP, Jmax or PD, are complete graphical programming environments in which complex structures can be realized by combination of several primary elements. In those oriented object environments, one can manage data from external gestural controllers as well as produce and modify sound. Moreover, the user of such programs can create its own bricks, to enlarge the number of object in the environment. The modularity and the evolutivity of those softwares make them very effective tools to design digital musical instruments according to our methodology.

3.2 Three examples of instruments

3.2.1 The voicer

The Voicer [17] is a digital musical instrument that imitates the vowels of a singing voice. the instrument is bimanual: the dominant hand drives the frequency and the other one navigates in a space of vowels. It uses a formantic vocal synthesis and it is driven by a graphic tablet [35] and a joystick. The choice of the vocal coloration is done by the position of the joystick. A source-filter synthesis model produces the sound. A 2D interpolator receives the x, y positions of the joystick and drives the filters in order to interpolate between 4 vowels. A specific mapping enables to drive the pitch on all the audio frequency range, to make vibratos and frequency slides as well as stable intonations.

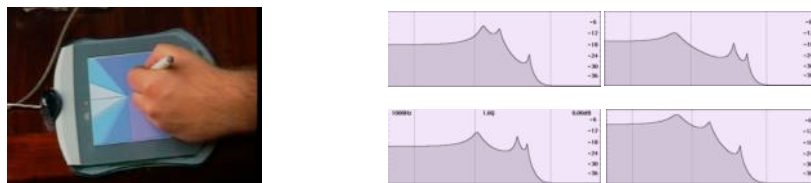


Fig. 3. A picture on the tablet helps the performer to control the frequencies; On the right, the frequency representation of the filter configurations corresponding to the 4 key vowels.

The “Voicer” is used in a musical work called “D’ici et d’ailleurs” as a soloist instrument in a jazz-like band; this work is articulated around an oriental mode in which the “Voicer” finds its expression field in accordance with its portamento, vibrato and micro-modulations possibilities as well as its vocal-like articulation.

3.2.2 Scanned Synthesis synthesizer

Scanned Synthesis [34] [7] is used in two different configurations in order to create two instruments. The principle of those instruments is the following: thanks to gesture by specific controllers, one interacts with a monodimensional shape (a string that can be circular) that moves at low frequencies, corresponding to the speed of our gestures. Gestures can put in motion the string or can modify its motion.



Fig. 4. The instrument uses a graphic tablet and a multi-touch tactile surface to control a string modeled by an assembly of masses, springs and dampers.

The shape of the string is used to produce sound in two configurations that are musically distinct: the Scanned Synthesis instrument, where the shape is scanned to

an audio frequency and the “filtering string”, where the shape is used to control the equivalent of a graphical equalizer.

In the Scanned Synthesis instrument [12], a tactile surface [31] is used to control dynamically the shape of the string (circular). Many presets have been created, providing different tones; those presets consist of the different values and the repartitions of these values on the string. The scanning frequency, that is the fundamental frequency of the sound, is controlled by the graphical tablet with the angular frequency control of the “Voicer”.

In the second implementation [4], the “filtering string”, we only use the shape of the string provided by our object, in order to modify the gains of a filter bank in a graphical equalizer. In this configuration, the graphical tablet controls the string parameters (stiffness, tension, damping) and the multi-touch surface controls the forces applied on the string, with a special mapping that enables the musician to give the string the shape he wants with his fingers.

In the work “Le rêve du funambule”, the input sound is a white noise that enables to produce complex sounds. The string shape is displayed on a screen during live performances and gives to the public a visual feedback that is totally representative of the sound process.

3.2.3 Photosonic Synthesis Emulator

The photosonic emulator [3] emulates the sound and gesture obtained on the photosonic instrument, an optical device created by Jacques Dudon at the Atelier d’Exploration Harmonique. The optical instrument uses a light, a photosonic disk upon which is inscribed a series of rings. A filter is interposed between the disk and a solar photocell. The sound is directly produced by the interruption of the light by the disk and the filter.



Fig. 5. On the left, the optical instrument; on the right, the tablet workspace is divided in two zones devoted to the control of the light and to the control of the filter.

The emulation of the instrument is realised in Max-Msp. The bimanual control is realised with two objects: a mouse and a stylus upon a graphic tablet; one makes the same gesture as moving a light and a celluloid filter in the photosonic (optical) instrument. This instrument is thus specific in the mimicking of the gestures of the “real instrument”. More than 800 disks have been created for the optical instrument, and some have been used in the emulator.

The photosonic emulator has been used in duet with the optical instrument: “Patchwork 605” emphasises a similarity between the palettes of the two instruments,

Design and use of some new digital musical instruments

“821, de l’autre côté de l’arc en ciel” display unusual sonic games due to new filters and rapid movements and “duo sur ondes fractales 905” makes up a dialogue on the exploration of a palette of fractal sounds. The emulator has also been used in “les dauphins” with an electronic cello giving an ecological flavour, and in “disque 729” a trio with electronic percussion and cello takes benefit of a long recursive delay to display a carpet of small photosonic sounds leaving place to some improvisation of the other instruments.

4 Conclusion

We have done here an overview of all the steps that we have ourselves followed in order to create our instruments and what we can do with them. We hope it will help others to understand the whole process of linking gesture to sound in order to perform on stage.

References

1. Arfib, D., “Digital synthesis of complex spectra by means of multiplication of non-linear distorted sine waves”. *Journal of the Audio Engineering Society* 27-10, 1979.
2. Arfib D., Kessous L., “Gestural control of sound synthesis and processing algorithms”, *Gesture workshop 2001*, ed. Ipke Wachsmuth & Timo Sowa, Springer-Verlag, Lecture Notes in Computer Science, LNAI 2298.
3. Arfib D., Dudon J. : “A digital emulator of the photosonic instrument”, *Proceedings of the 2002 Conference on New Instruments for Musical Expression (NIME-02)*, Dublin, Ireland, May 24-26, 2002, proceedings also on line at <http://www.nime.org>
4. Arfib D., Couturier J.-M., Kessous L. (2002) : “Gestural strategies for specific filtering processes”, *proceedings of DAFx02 conference*, Hamburg, 26-28 sept 2002, pp. 1-6.
5. Arfib D., Couturier J.M., Kessous L., Verfaillie V., “Mapping strategies between gesture control parameters and synthesis models parameters using perceptual spaces”, *Organised Sound* 7(2), Cambridge University Press, pp. 135-152, 2002.
6. Boulanger R., *The Csound Book*, MIT Press, 2000.
7. Boulanger R., Smaragdis P., Ffitch J., “Scanned Synthesis : An introduction and demonstration of a new synthesis and signal processing technique”, *Proceedings of the 2000 International Computer Music Conference*, pp. 372-375, Berlin Zannos editor, ICMA, 2000.
8. Cadoz C., Luciani A. and Florens J-L., "Synthèse musicale par simulation des mécanismes instrumentaux, transducteurs gestuels rétroactifs pour l'étude du jeu instrumental", *Revue d'Acoustique*, vol. 59, pp. 279-292, 1981.
9. Cadoz, C., *Les réalités virtuelles*, Dominos, Flammarion, 1994.
10. Cadoz C. «Musique, geste, technologie» dans *Les nouveaux gestes de la musique direction* Hugues Genevois et Raphaël de Vivo, Éditions Parenthèses, 1999, p 47-92.
11. Chowning J., "The synthesis of complex audio spectra by means of frequency modulation", *Journal of the Audio Engineering Society*", vol 1 n°2, p 526-534, 1977.
12. Couturier J.M., “A scanned synthesis virtual instrument”, *Proceedings of the 2002 Conference on New Instruments for Musical Expression (NIME-02)*, Dublin, Ireland, May 24-26, 2002.

13. Couturier J.M., "La synthèse par balayage et son contrôle gestuel", Proceedings of the 9th Journées d'Informatique Musicale (JIM 02), Marseille, France, May 29-31, 2002.
14. Crowley J. L., "Vision for man machine interaction", Proceedings of Engineering Human Computer Interaction, EHCI'95, Grand Targhee, USA, August 1995, Chapman&Hall Publ.
15. Gather J. P., Amsterdam Catalog of Csound Computer Instruments, <http://www.csounds.com/accci/index.html>
16. Hunt, A. Wanderley, M. Kirk, R., Towards a Model for Instrumental Mapping in Expert Musical Interaction. Proceedings of the International Computer Music Conference (ICMC'2000, Berlin), ICMA, 209-12, 2000.
17. Kessous L., "A two-handed controller with angular fundamental frequency control and sound color navigation", Proceedings of the 2002 Conference on New Instruments for Musical Expression (NIME-02), Dublin, Ireland, May 24-26, 2002.
18. Lebrun M., "waveshaping synthesis", *Journal of the Audio Engineering Society*, vol 27, pp. 250-266, April 1979.
19. Mathews M. and Kohut J., "Electronic simulation of violin resonances," *Journal of the Acoustical Society of America*, vol. 53, no. 6, pp. 1620-1626, 1973.
20. Max/MSP, <http://www.cycling74.com/products/maxmsp.html>.
21. Menzies D., "New Performance Instruments for Electroacoustical Music", Dphil Thesis, 1999, <http://www.zenprobe.com/dylan/pubs>
22. Miranda E., "computer Sound Design", 2nd edition, focal press ed., 2002
23. Moore, F. R., "Elements of Computer Music", Englewood Cliffs, NJ: Prentice Hall, 1990.
24. Palombini, C., "Pierre Schaeffer's Typo-Morphology of Sonic Objects", Ph.D. University of Durham, 1993, <http://www.sun.rhbnc.ac.uk/Music/Archive/Disserts/palombin.html>
25. Paradiso J., Electronic Music Interfaces: New Ways to Play J. Paradiso, IEEE Spectrum Magazine, Vol. 34, No. 12, pp. 18-30 (Dec., 1997).
26. Penttinen H., Härmä A., and Karjalainen M., "Digital Guitar Body Mode Modulation With One Driving Parameter", Proceedings of the DAFX-00 Conference, Verona, Italy, December 7-9, pp. 31-36, 2000.
27. Risset, J.C. (1969). An introductory catalog of computer-synthesized sounds. Reprinted with C.D. Wergo 2033-2, The historical CD of digital sound synthesis (1995), 88-254.
28. Roads C., "Introduction to Granular Synthesis", *Computer Music Journal*, 12, 2, 11-13, 1988.
29. Roads C., "Sound Composition with Pulsars", *Journal of the Audio Engineering Society* 49(3): 134-147, March 2001.
30. Sapir S., "Gestural Control of Digital Audio Environments", Special Issue: "Musical Implications of Digital Audio Effects", *Journal of New Music Research*, Guest Editor D. Arfib, Number 2, June 2002.
31. Tactex, touch surfaces, <http://www.tactex.com/>.
32. Todoroff T., "Control of digital audio effects", DAFX digital audio effects, U. Zolzer Ed., J. Wiley & Sons, 2002, pp. 465-497.
33. Verfaillie V., Arfib D. (2002) : "Implementation strategies for Adaptive Digital Audio Effects", proceedings of DAFX02 conference, Hamburg, 26-28 sept 2002, pp. 21-26.
34. Verplank B., Mathews M., Shaw R., "Scanned Synthesis", Proceedings of the 2000 International Computer Music Conference, pp. 368-371, Berlin, Zannos editor, ICMA, 2000.
35. Wacom tablets, <http://www.wacom.com/>
36. Wanderley M., Battier M., Ed, CD-ROM Trends in Gestural control of music, Ircam, 2000.
37. Wanderley, M. 2001. "Performer-Instrument Interaction: Applications to Gestural Control of Music", PhD Thesis, Paris, France, University Pierre et Marie Curie - Paris VI. http://www.ircam.fr/wanderle/Thesis/Thesis_comp.pdf.
38. Zolzer U., DAFX digital audio effects, J. Wiley & Sons, 2002,