Expressiveness and digital musical instrument design

Daniel Arfib, Jean-Michel Couturier, Loïc Kessous

LMA-CNRS (Laboratoire de Mécanique et d'Acoustique)
31, chemin Joseph Aiguier
13402 Marseille Cedex 20
France
[arfib,couturier,kessous]@lma.cnrs-mrs.fr

Abstract

In this paper, after giving some possible definitions for expressiveness, we examine the problem of expressiveness in digital musical instruments, which tends to involve using specific gestures to obtain an expressive sound rather than performing expressive gestures. Some of the particular features of digital musical instruments, such as pitch control, dynamic control and the possibility of exploring sound palettes, are described and some practical examples given. Lastly, several musical implications of the gestures used to obtain musical expressiveness are discussed, from the pedagogical and other related points of view.

1. Introduction

Expressiveness in music, as in all the arts, can have different meanings. Expressiveness is the capacity to convey an emotion, a sentiment, a message, and many other things. It can take place at various levels, from the macroscopic to the microscopic scale. In the case of musical performance, expressiveness can be associated with physical gestures, choreographic aspects or the sounds resulting from physical gestures. The design of a digital instrument must take its expressive possibilities into account. The notation used and the pedagogical aspects also need to be considered seriously if one wants other people to be able to use these instruments. The design of an instrument can also include didactic aspects, which can help beginners to get started.

Previous studies have been carried out on expressiveness in the artistic context. In particular, in [Camurri & al, 2001], the authors investigated expressiveness in gestures using computational modeling, and applied the findings obtained in artistic contexts where enhancing the expressiveness in interactive music/dance/video systems was one of the main goals. A multi-layer conceptual framework is presented by these authors, and examples are given showing how it can be used in interactive artistic performances. This framework is probably particularly suitable for applications where analysing the expressiveness of gestures is one of the main aims.

Expressiveness in the design of digital musical instruments is not restricted to producing expressive gestures: the gestures do not have to be "expressive" in themselves, but have to be able to generate expressive sounds. Here the same problems arise as with acoustical instruments: the gesture in itself may not be "beautiful", but the sound produced by the instrument should be esthetically pleasing. Our research does not focus on situations of the kind where the expressiveness of the gesture is extracted first. The links between gestures and sound processes are more direct and explicit here, and focus on sound production.

This paper describes expressiveness in digital musical instruments in terms of the characteristics of the sound and the adaptability of the instrument in question; some visual aspects are also discussed. The second section of the paper deals with three expressive features and the way we have implemented them in our digital instruments: pitch control navigation through sound palettes and dynamic control of sound parameters. The last section is about the implications of expressiveness during live performance: how to play music with a digital musical instrument.

2. Expressiveness and digital musical instruments

When inventing acoustical instruments, designers have to find the best compromise between the abilities of the human body and the physical constraints involved in sound production. The gestures used on acoustical instruments depend strongly on the physics of the instrument. In digital musical instruments, sounds can be generated without any physical constraints: the designers of instruments of this kind are free to choose whatever gestures they want and how they want these gestures to link up with the sounds produced. This linkage, which is called *mapping*, is one of the main aspects of computer music research [Hunt & al., 2003] [Wanderley, 2002]. Although commercial devices often include the MIDI system and controllers imitating conventional instruments (keyboards, breath controllers, etc.) to control the sound, the use of interfaces of novel or alternative kinds gives instrument designers greater more freedom in the mapping. It also gives performers better control over the expressiveness of their gestures. However, digital instruments are more than just musical controllers: the systems on which they are based also include synthesis algorithms and mapping strategies. The choice of synthesis algorithms, controllers and mapping systems will determine the nature of an instrument and its ability to play in various styles and configurations. Each step in the instrument's design will also determine how the audience will perceive the performer-instrument relationship on stage. Although the expressiveness of an instrument is mostly a question of sound, the visual aspects also play an important role, at the level of the gestures made by the performers, or the visual feedback possibly produced by the instrument.

2.1 Expressiveness and identity

One can speak about the expressiveness of a musical instrument to define something that one might also call its identity. This identity of acoustical instruments depends on the choice of synthesis algorithms, controllers and mapping systems. The identity of an instrument can be recognized from the sound produced at several levels: at the macroscopic level, which corresponds to the phrasing level, and also at the microscopic level, which corresponds to the sound object level.

The phrasing level

At the phrasing level, an instrument can be recognized even from a musical recording. Although the timbre is obviously the main feature used to identify an instrument, many people can tell the difference between the kind of musical phrasing produced on a keyboard combined with a sampler containing violin samples (even when it is a multi-layer sampler including many samples of each note, played at different velocities) and the musical phrasing played by a violinist. The levels at which this ability to discriminate operates even include that of the performers, since one can distinguish between two performers' interpretations. For example, the velocity curve and the micro-delays introduced relative to the strict timing indicated on the score can characterize the expressiveness of a particular pianist. These parameters can also be used to define styles, as often occurs when describing the options in the sequencer software programs which make it possible to adapt a sequence to a context.

Digital instrument design must allow performers enough flexibility as well as enough precision to be able to introduce nuances into their playing. A well-known weakness of the MIDI system [Moore, 1988], which does not satisfy this requirement, is due to the fact that it is based on a serial machine communication protocol; for example, chords become arpeggios and any irregularity in the latency will be detrimental to expressiveness.

The way we physically organize note control will influence the instrument's phrasing characteristics. The most appropriate mapping strategies and peripheral configurations depend on the level of musical expressiveness required, as well as providing tools giving an instrument its identity.

The note or sound object level

One can talk about the expressiveness of each note in a musical phrase. Each note can be modulated in terms of its tone, energy and spectrum. Violinists, guitar-players and other classical musicians use glissandi; they also use techniques such as hammering-on, pulling-off, and other pitch modulation techniques. The position of the bow along the string, its relative inclination and other aspects of a string-player's gestures have spectral implications. In computer music, especially when sound synthesis is used, spectral articulation, pitch modulation and energy control applied during the lifetime of a note also qualify as elementary gestures. From the beginning of computer music in the non-real time context, people have been modulating sounds by drawing curves and designing low frequency oscillators and jitter generators to make the sound seem more alive. Musical features such as vibrato and portamento are distinctive features of singing voices, but adding a similar mechanical vibrato to each note of a musical phrase will not be very expressive and will not even make it sound like a real voice. Spectral aspects such as vowel changes or the brightness of brass tones are also used by listeners to recognise an instrument and determine its naturalness.

Elementary musical gestures are linked to the phrasing level because phrasing involves making a lot of elementary gestures. Accurate data acquisition and transmission and the appropriate choice of sensor technologies and mapping strategies are required to be able to give sound expressiveness using elementary gestures.

2.2 Expressiveness as adaptability

The expressiveness of an instrument can also mean its ability to be used to play different styles of music. This can mean not only using different tone scales, tempered and otherwise, microtonal or natural scales, but also different ways of composing the whole phrase in terms of time, energy and spectrum. As far as pitch variations are concerned, the human voice and the violin provide good examples. Both instruments can play Western, oriental and Indian scales as well as classical music, jazz, contemporary and pop music. The expressiveness will be correlated with the ability of an instrument to allow the performer to adapt his playing to a context. The limits of this possibility are equal to those of the mapping system used. Musicians are not supposed to be restricted to a single musical style and are expected to be able to switch from one to another, crossing the frontiers between them. An 'expressive musical instrument' can therefore also be said to be an instrument which allows a performer to follow other musicians in various musical directions.

Adaptability includes several concepts involving musical properties. One of them is the concept of "emergence", which means that an instrument can be clearly heard and identified (when it is used as a soloist) against an orchestral background. This can be achieved in various ways by using an additional formant in the case of a voice synthesis, by specifying the directional characteristics or by enhancing the brilliance or the attack.

Adaptability also includes the possibility of changing the musical field within which the instrument is playing at any time, by changing the range of parameters or the sound palette.

2.3 Expressiveness and visual considerations

Expressiveness in musical performance can also involve visual aspects. Visual feedback can enhance the interactive processes between performers and their instruments, as well as helping the audience to understand how the performers master their instruments.

The use of video tools in music and performance

The development and spread of technological tools has led to performances where video images are combined with music.

One of the most commonly used approaches consists in composing a piece of music with a video counterpart, and playing it back in real time with gestural control. The video and the music can be controlled either by different performers or by the same performer. In the second case, the operator can use either different controllers or the same one to conduct the music and the video simultaneously. The artistic touch will then be a question of deciding the relative importance of the video and the music and the interactions between then. Introducing video into musical performances can greatly affect the way in which a performance is perceived by the audience.

Another approach is to use some of the components of the sound to control the video or some of the parameters of the instrument to illustrate the musical gestures. For example, Levin's AVES (AudioVisual Environment Suite) [Levin, 2000] and Jorda's FMOL [Jordà, 1998] are systems in which dynamic virtual objects are used to control both sound and visual feedback. The video serves as visual feedback helping the performers to play their instruments, as well as helping the audience to understand how the instrument works.

Interactive real-time visually displayed musical gestures

In live computer music, it is not always easy to understand the role of each of the performer's gestures. For example, since graphic tablets, joysticks and interfaces of other kinds can be used in very different ways to control sound, how is the audience supposed to know what the performer is really doing with them?

The presence of several performers on stage is an interesting special case. Visual devices could be used to help the audience to determine which performer is producing a specific phrase. Then the audience could focus on a particular performer and progress from an overall hearing to an analytic hearing. In this case, expressiveness is correlated with comprehensibility and unknown instruments can be very difficult to comprehend.

Visually displayed concepts and metaphors can also be helpful to the audience. This method consists in showing on a screen some of the components of the instrument, or presenting metaphorical images that illustrate the principles on which the instruments are based. Here the video will not only add visual effects to the performance but will also help the audience to understand the instrument by providing the performers and the audience with visual feedback.

Three different levels of visual feedback can be defined The first level involves the direct illustration of the parameters controlled by the players. For example, measurements of pressure or blowing force can be displayed in the form of a slider or the degree of illumination of a graphical object. The second level involves the visual representation of interpreted gestures, metaphors, concepts and principles. The representation can be either static or dynamic, depending on the mapping model used. Some experiments on visual feedback at this level will be presented below, using both static and dynamic models. The third level involves the representation of gestures in terms of their effects on the sounds generated. Audio signals are often used nowadays in music player software programs. Although the links between sound and video are rarely very strong, a relevant real-time sound analysis using appropriate sound descriptors and a suitable method of illustration based on a specific mapping procedure could provide efficient visual feedback - or at least, a correlated artistic picture of musical events.

Another approach, which might be said to fit in with the second level described above, could be to visually display the perceptual sound parameters (such as the loudness and brightness) used in the mapping chain to control the process of synthesis [Arfib & al., 2002b]. Visual feedback can be a part of the artistic composition or not, but the efficiency of the visual feedback will depend on its legibility. Combinations between the various levels are also possible and would be worth exploring. If visual feedback is to be used to improve players' performances, it will probably be necessary to take the flux of information provided by the visual feedback into account, as well as the interactions with sensory feedback of other kinds (such as the haptic and auditory feedback). Too much convergent information might, however, be difficult to integrate at the cognitive processing level.

3 Expressive features of some new digital instruments

In this section, it is proposed to describe three expressive features we have developed and implemented in our digital instruments: pitch control with the Voicer, sound palette navigation with the Photosonic Emulator and dynamic sound parameter control with the Filtering String.

3.1 *Expressive pitch control: experimenting with the Voicer*

The Voicer is an instrument simulating a vowel-singing voice. We have used a Wacom graphic tablet equipped with a stylus transducer and a game joystick to create an expressive solo instrument. The synthesis model consists of a sawtooth signal filtered by three cascaded second-order all-pole filters. This instrument makes it possible to simultaneously carry out expressive melodic control and vowel articulation. The joystick controls the vowel produced by varying the tongue hump position and the constriction of the vocal tract. Pitch and amplitude are controlled via the graphic tablet.
This section deals with the part of the instrument that controls the pitch. Starting with the problem of continuous pitch control in the MIDI system, we go on to discuss the dimensionality of pitch and pitch perception and control. The pitch control system used in the Voicer is then presented, followed by the visual feedback aspects.

Pitch and MIDI controllers

The MIDI standard [MIDI] has been a revolution in electronic music, providing a method of linking together synthesizers, controllers and computers; although this protocol was developed back in 1983, most controllers, synthesizers and software are still being equipped with this system MIDI makes it possible to transmit control data, continuous data and information about events. One of the particularities of the MIDI process is the fact that pitch control can be carried out in two ways. The first way consists in triggering a note with a given velocity and pitch (a NOTEON message). The second one consists in modulating the pitch of previously activated notes (a PITCHBEND message). Pitch bend values of 7 or 14 bits are generally used. Most of the MIDI controllers, such as the keyboard, wind controller and guitar controller, comply with this rule. The keys of the MIDI keyboard, for which MIDI was initially developed, trigger notes and the pitch is modulated by means of a pitch bender (a wheel, stick or lever). In the MIDI wind controller, pressing the keys selects a pitch, blowing triggers the note, the buttons at the back (near the thumb of the upperhand) can be used to select the octave, and there is a pitch bend wheel near the thumb of the lower hand. Lip pressure is sensed and can also be assigned to pitch bending.

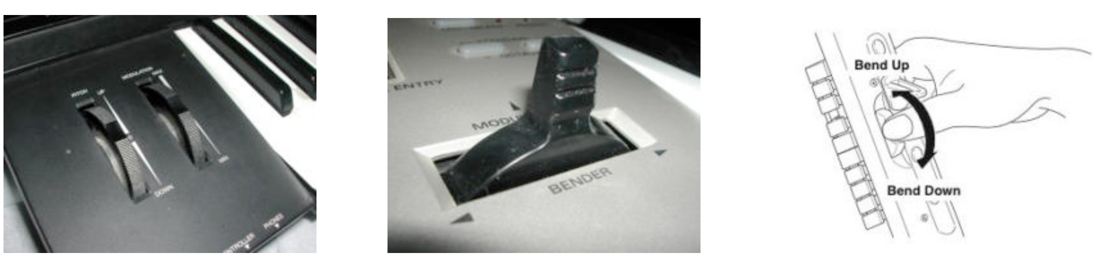

Fig 1. Three kinds of pitch benders

With the MIDI protocol and the features of MIDI controllers, one cannot obtain exact continuous pitch variations in a large range (of several octaves), without retriggering the notes. In addition, the two parts of the pitch control are generally controlled by different parts of the body. The MIDI system and the usual MIDI controllers are therefore not suitable for applications requiring the continuous control of pitch of which the voice and some acoustical instruments are capable.

Dimensionality of pitch control

With conventional instruments, we often have several ways of obtaining changes in pitch. A saxophonist can either use the octave key or change the pressure properties at the mouthpiece. The tuning difference between successive chords of a guitar is 4 or 5 semi-tones and the pitch can be changed from one octave to another by playing on another string. The piano keyboard consists of a series of cells, each containing 12 keys.

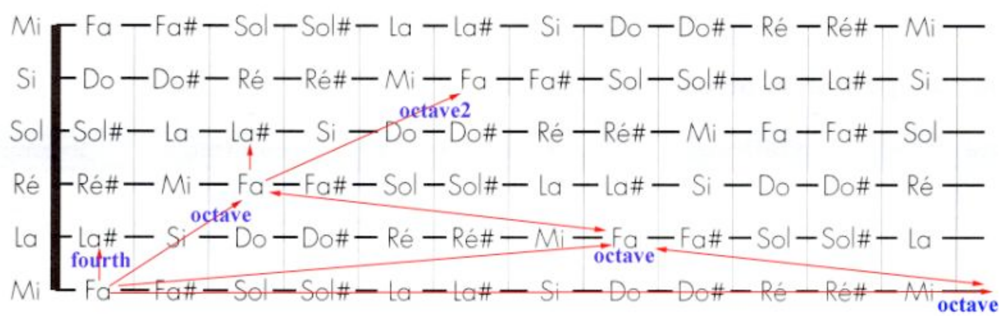

Fig 2. Organization of pitch control on guitar fingerboard

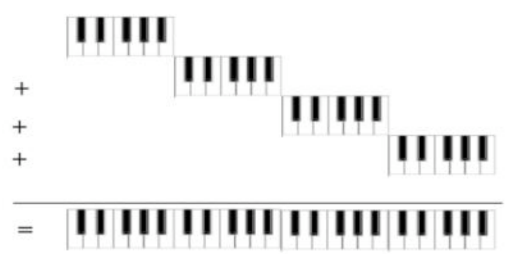

Fig 3. Representation of control keyboard as the repetition of a cell

The way in which pitch control is designed can affect expressiveness, especially at the phrasing level. Some of the gimmicks used in improvisation depend on the pitch control strategy implemented in an instrument.

Pitch perception and control

The human ear can perceive very small pitch variations. According to [Arom & al., 1997], musicians are able to discriminate adjacent intervals to within +/- 20 semitone cents, i.e. intervals less than one tenth of a tone apart. When musicians themselves tune their instruments, they are sometimes more accurate than +/- 10 cents. According to [Zwicker, 1990], using sinusoidal tones, a change in frequency of about 0.7 % is just noticeable at frequencies above 500 Hz. Musical tones are rarely sinusoidal tones, however; they have many harmonic components and the frequency changes in these harmonics can be detected at a lower frequency than the fundamental frequency. Finer intervals can also be detected when two notes coexist in the form of harmonic beats. An expressive instrument must match this accuracy in terms of data precision. Time precision is also required to preserve the form of a modulation; for example, a vibrato can be assumed to be sinusoidal although this is not exactly the case. Each performer produces his or her own specific pattern of vibrato The logarithmic pitch perception scales also have to be used to control pitch by performing linear gestures, even in the case of continuous pitch modulation devices.

$$frequency = 440 * 2^{(MIDI\_NOTE\_NUMBER-69)/12}$$

$$semi\_tone\_factor = \sqrt[12]{2} = 2^{1/12} = 1.0595$$

$$10\_semitone\_cents\_factor = (\sqrt[12]{2})^{10} = 1.0058$$

$$1/8\_semitone\_factor = (\sqrt[12]{2})^{8} = 1.0072$$

Fig 4. Formulae usually used to convert MIDI into frequency

Pitch control with the Voicer

The pitch control strategy used in the Voicer can be said to provide an answer to the following question: what happens if one wants to produce a glissando over a range of two octaves or more and to finish this gesture with a vibrato?

To control the pitch within each octave and from one octave to another, we divide the tablet's active space into 12 sectors (12 equal angular parts, each corresponding to a semitone on the chromatic scale). The pitch control is continuous and circular: turning the pen tip clockwise changes the pitch from low to high (with special features for vibrato and other pitch modulation gestures). We can go one octave lower or higher by pressing the lateral button on the stylus up or down. To facilitate gestures such as portamento and vibrato, the tuning control is more powerful at the limits of the angular sectors. The first mapping step consists in determining in which of the twelve angular parts the pen is located. Then we have to see how well centered it is. Lastly, we need to know how many turns have been made around the center of the tablet.

To reproduce pitch changes such as those made by a singing voice, the Voicer provides a pitch control that is more expressive than that of a keyboard or a wind controller equipped with a pitch bender. First, the Voicer was designed to provide an instrument with a similar level of expressiveness to that of a vowel-singing voice. A point worth noting about singing voice expressiveness is the importance of continuous pitch modulation. For example, a pitch bender can be configured in several ways: within a small range (+/- 1/2 semi-tone for example) or a large range (+/- an octave for example). The first type of range gives a closer control, but will only be a small range. The second type of range has its advantages but the precision of the control will be lower. The pitch control strategy used in the Voicer makes precise control possible in both small and large ranges and can be used to produce both fine vibrato and large portamento effects.

The intrinsic expressive abilities of the "controller part" of the Voicer make for accurate time and quantification data and good resolution, as well as for gestural precision. The high level of data accuracy and resolution result from the sensing technology and the system of

communication adopted. The graphic tablet used in the Voicer has greater precision and resolution than most pitch benders. Using a stylus in the preferred hand to control the pitch (and vibrato in particular) definitely seems to be a better means of achieving precision than using any kind of pitch bender with the non-preferred hand.
It would be interesting to make comparative assessments between players performing various tasks, using the Voicer control part, a wind controller, and a keyboard and the same vowel singing voice models to test the preliminary conclusions reached as the result of our experiments on the Voicer designer with the controllers described above.

Visual pitch control feedback

The visual feedback provided with the pitch control part of the Voicer takes the form of 12 angular sectors, which show up individually in different shades of blue. The sector pointed to by the stylus is also indicated by a red component; the red intensity depends on the pressure and is therefore associated with the loudness. The radius of the disc formed by the whole set of sectors depends on the number of turns performed and it therefore determines the octave played.

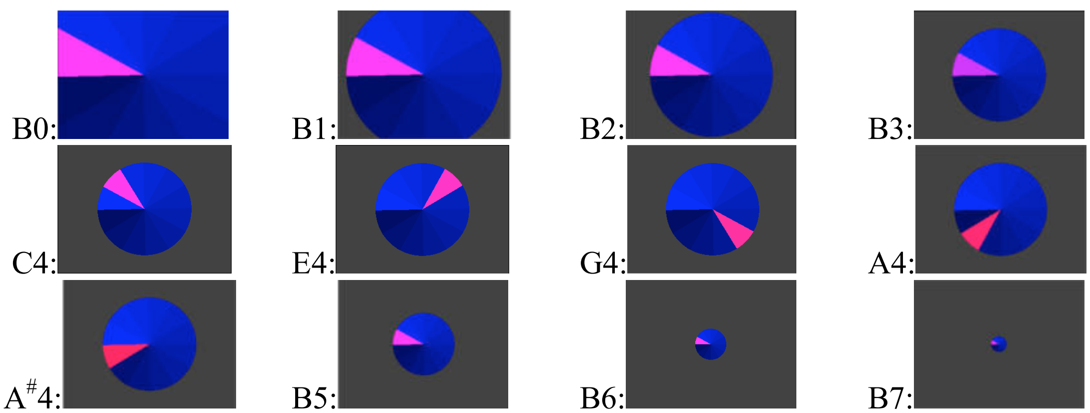

Fig. 5. Visual feedback for the Voicer

There are two possible ways of presenting the visual feedback to the performers: it can either be projected onto the screen front of them, or a pen-based touch screen can be used, providing either a direct or indirect relationship between manipulation and visual perception.

3.2 Navigation and spectral manipulation with the Photosonic Emulator

Some instruments can be used to navigate through pre-determined sound palettes, with the possibility of improvising within a palette and choosing one of several palettes. The possibility of selecting a sound palette and exploring it are essential features of these musical instruments. To explore sound palettes with gestures, several strategies are imaginable; one of them is to try to create a spatial representation of the sound palette and to transpose this representation into the physical space of the gesture. When using a two-handed instrument, one can keep the second hand for making the spectral changes in the sound resulting from the navigation. Combinations of this kind are used in photosonic instruments (optical instruments and their emulators) [Arfib & Dudon, 2002].

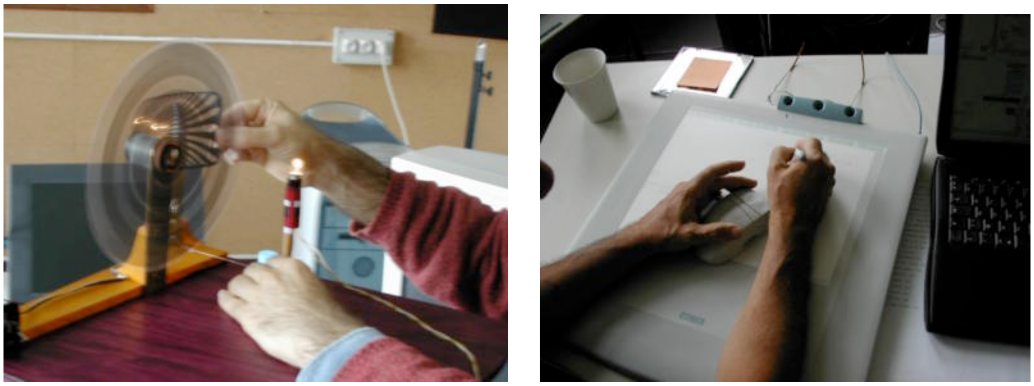

Fig. 6. The Photosonic instrument and its emulator

The photosonic instrument created by Jacques Dudon in the 1980s is an optical instrument played with two hands, one of which moves a light in front of a disk, while the other one interposes a graphic filter in front of a solar photocell. The first movement corresponds to an exploration of the sound palette inscribed on the disk, while the second movement corresponds to the “sculpture” of the sound by a filter which induces filtering (horizontal movements) and a Doppler shift (vertical movements). The photosonic emulator is a digital instrument that mimics these hand gestures. For this purpose, a special mapping procedure is carried out between the coordinates on a Wacom tablet and the parameters of the photosonic digital emulator.

Expressiveness depends here on these two basic gestures, as well as on various “microgestures” that can be recognizable immediately by ear. Some of these gestures are simply postures, which means that the positions of the two hands are fixed. They usually correspond to the filtering of the sound produced by a “ring”. However, the way in which this posture is reached and abandoned is most important. Even micro-variations can have effects. Some micro-movements always occur that affect some of the parameters involved in the hand position. Learning to move from one posture to another one is essential, and experience has shown that the photosonic emulator relies more on “modulation gestures” than on “decision gestures.

Some other gestures are more akin to movement in general. The non-preferred hand (left in the case of right-handed persons) governs the content of the sound, and can give the player the impression of exploring a palette, which can be linear or form a loop (in this case, making either zig zag movements or circular ones helps). Arches can also be described, which produce a silence before and after the movement. The preferred hand produces more subtle effects, however: horizontal movements serve to scan the range of filter possibilities, and vertical ones, as with the Doppler effect, make it possible to easily obtain pitch variations. These movements are linked to the sound in at least two ways: first of all, the mapping indicates the amplitude of the movements and, for example, the ambitus of a vibrato must be properly correlated with the trembling of the hand. Secondly, the effect obtained depends upon the filtering pattern used. New filters [Arfib & al., 2002] have been devised that can

make sounds resemble vocal sounds, so that the horizontal exploration no longer consists of scanning the central frequency of a band pass filter, but rather of making an interpolation between two vowels.
We have also developed a graphic interface in which two movable graphic objects are displayed on a screen showing a light and filter (see figure 7). We have used this interface for a specific implementation of the photosonic emulator including "the Pointing Fingers" controller [Couturier, 2003], which is a multi-finger touchscreen-like device.

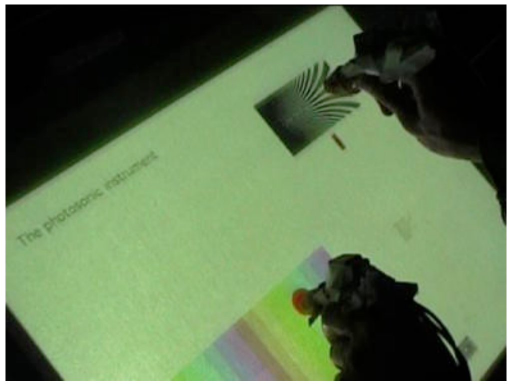

Fig. 7. A graphic interface for the Photosonic Emulator, controlled by the Pointing Fingers device. The left hand manipulates the light (red circle) on the rings and the right hand controls the position of the filter in front of the photocell.

This interactive mode provides a digital instrument that is really similar in terms of its appearance to the original optical instrument.

3.3 Expressive dynamic behavior: experiments with the Filtering string instrument

Another important feature of a digital musical instrument is whether it is endowed with static or dynamic behavior [Menzies, 2003]. Static behavior occurs when data triggered by the player's gestures instantaneously generate sound parameters: at any time, the sound parameters depend only on the gestural data collected at that time. With dynamic behavior, the sound variations are generated not only by the player's gestures but also by the effects of these gestures on the dynamic system. If a dynamic system is included in the mapping, the sound identity of the instrument will also depend on how the system is designed to respond to gestures. In this section, we will present a dynamic musical instrument, the "Filtering String", which illustrates the benefits of dynamic behavior in digital instruments.
The "Filtering String" instrument uses the shape of a slowly-moving string to control the gains in a filter bank [Arfib & al., 2002]. This instrument is based on the idea that using a dynamic system to control a simple synthesis will produce sounds with a richer spectral evolution in time than in cases where the synthesis is directly controlled by the user.
This instrument incorporates a Max/MSP object we created to provide a high level of control over a string model [Couturier, 2002], and has usually been used in scanned synthesis applications [Verplank, 2000] [Boulanger, 2000]. This object makes it possible to exert an overall control over the string parameters, since gestural data can be directly connected to it. The object's output is a list of parameters that correspond to the shape of the string; we have used this shape to control a filter bank.

Two categories of parameters play an important part here: the string parameters and the filter bank parameters. The filter gains are controlled by the string shape, whereas the other parameters of the filter bank can either be given constant values or be linked to the player's gestures. This instrument was designed to be able to control the sound via the interactions between the performer and a dynamic system, which is the reason why we decided to control only the string parameters with gestures. We have also provided different configurations of the filter bank parameters (Frequencies, Q, sound in); one can access these configurations and shift from one to another using selection gestures (on buttons).

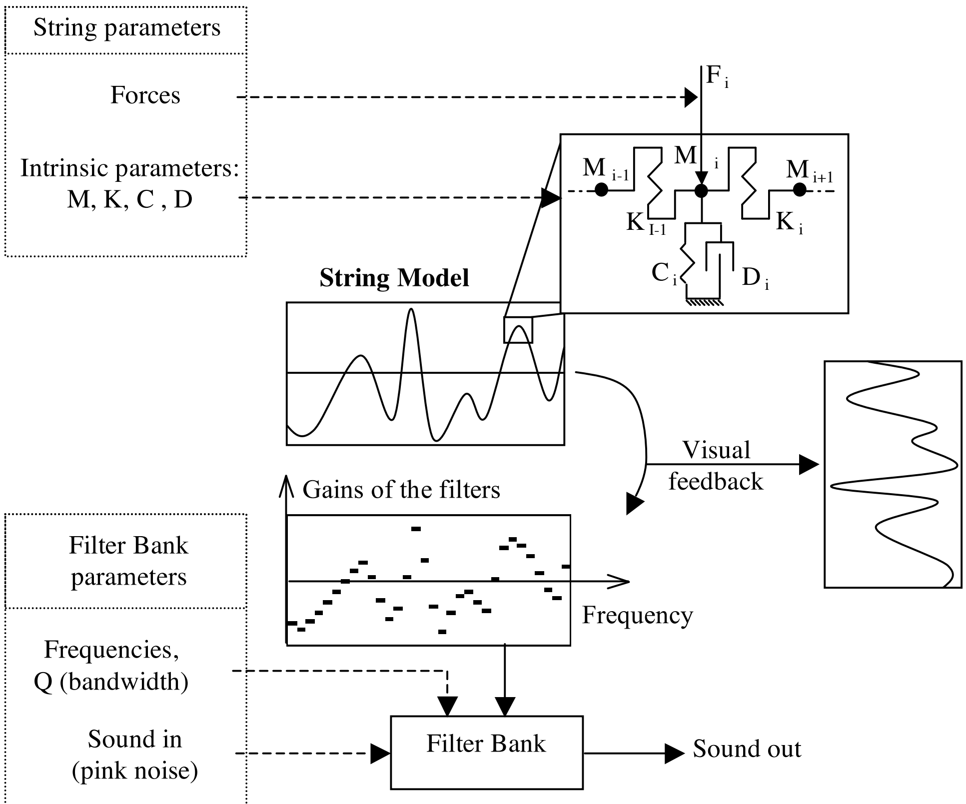

Fig. 8. In the "filtering string" instrument, a slowly-moving string modelled by a set of masses, springs and dampers controls the gains in the filters in a filter bank. Only the string parameters are controlled by gestures.

The sound identity of the instrument depends here on how the filter gains are dynamically driven, rather than on the spectral colour of a noise filtered by a filter bank. Depending on how the dynamic system used in the mapping moves and how it is controlled, it will contribute largely to the identity of the instrument in terms of the sound it produces.

The filtering string has also a visual identity, apart from the shape of the gestural controllers: in live performances, the string shape is displayed on a screen, and besides the artistic effects of this display, it helps the audience to understand how the instrument works. In addition, since the visual feedback is closely linked to the sound, he audience can see that the auditory and visual aspects of the instrument are part of the same process.

This visual feedback provided by the instrument is also important because it enables the performer to look at the dynamic device he is interacting with. It gives him a closer contact

with the instrument because, as in all interactions with physical objects, the performer can see what he touches.
The instrument is equipped with a graphic tablet [Wacom] controlling the string parameters (stiffness, tension, damping) and a touch surface [Tactex] controlling the forces applied to the string, along with a special mapping that enables the musicians to give the string whatever shape they want with their fingers. The users have to press the touch surface to apply forces to the string; the surface is divided vertically into two parts: pressures on the right side will apply forces towards the right side of the string and *vice-versa*. The horizontal pressure profile on the surface of the string corresponds to the forces that can be applied along the string.
The touch surface is used to energize the dynamic system, and the graphic tablet makes it possible to change its intrinsic parameters. The touch pad can be contacted using one or several fingers to press, slide, tap or lightly touch the surface. The effects of these gestures on the sound will depend on the values of the intrinsic string parameters controlled by the graphic tablet: for example, at low stiffness values, the string will move slowly and will not respond to fast movements on the touch pad.
Users have to learn the basic rules about the behaviour of the string before they can play the instrument successfully; these rules are easy to understand, and once they are known, the user can immediately play on the string.The instrument is usually played by alternating excitation gestures on the touch pad and modification gestures on the graphic tablet: the excitation gestures introduce energy into the dynamic system and the modification gestures drive the evolution of the sound [Cadoz & Wanderley, 2000]. The time taken by the dynamic system to lose its energy depends on the value of the damping parameter.

Fig. 9. When the user presses the surface, the forces applied cause the string to leave the equilibrium state; after being stimulated and driven by the graphical tablet, the string will evolve according to its own dynamics, before returning to its initial position.

When playing with other instruments, it is often necessary to tune the 32 frequencies of the filters. The musical work “le Reve du Funambule” is divided into several parts, and in the last part, the Filtering String is accompanied by the Photosonic emulator; both are tuned on the “Didymus” scale.
The Filtering string is an instrument in which expressiveness is strongly linked to the dynamic behaviour of the slowly-moving string.

4. Implications of expressiveness in musical performance

As a matter of fact, there is no pre-defined way of proceeding with expressiveness, and new instruments also set new challenges. One particularly strong challenge is how to play these instruments (the question of composition will not be addressed here). Pedagogy involves transmitting knowledge from an expert to a learner. Four aspects can be distinguished in the transmission of expressive sound production gestures: imitating gestures, performing gestures to copy specific sounds, interpreting a score or a gestural notation, and inventing new gestures. Each of these aspects will now be discussed and illustrated with reference to the instruments defined in section 3.

4.1 Imitating gestures

The first learning method consists of imitating a gesture performed by the teacher. This also presupposes that the latter person has actually mastered these gestures, at least in an archetypal form. The imitation can be decomposed into two parts: the skeleton of the gesture (its definition) and its body (its expressiveness).

The skeleton gesture can be described in biological terms, but its meaning belongs to the cognitive domain: we do not interpret a gesture only from the way it is made, but also from the underlying intentions. This means that defining a gesture as an exploration of a psychological space can make sense, and this also precludes simply defining a space and how it is explored. To give an example, producing a vibrato using a graphic tablet often involves performing a basic gesture: one must oscillate the pen tip in order to make a sinusoidal change in the frequency. This gesture can of course be learned without any sound production; however, the auditory feedback helps the learner to produce a good vibrato.

Expressiveness can be said to be adding emotion to a gesture. This of course depends greatly, in the case of the present instrument, on the mapping process used to convey the information resulting from the gestures into information liable to produce sound. Gestures do not necessarily have to be very demonstrative to be expressive, but they must make sense to the brain of the performer. To describe different modes of expression is to define nuances and ways of producing them. These new gestures are not really very different from the old ones, except that in traditional instruments, the expression depends on material constraints, whereas in gesture-controlled digital instruments, the expression depends on the mapping.

Navigation in the photosonic emulator can require, for example, “arches”, “circles”, and “scratching” gestures that are not particularly familiar to musicians. On the Voicer, the typical gesture used for pitch control is a circular movement instead of a keyboard one, and pressure and vibrato are associated with the same movement. The Filtering String is a complex case because the musician has to manage a dynamic system: the same gesture performed at two different moments can give rise to two different sounds; in addition, the multi-finger interface requires an entirely new vocabulary to be able to describe the movements.

4.2 Imitating sounds

The second method of learning (imitating sounds) seems to be more suitable for experts than for beginners. With this method, one attempts, for example, to imitate a Jimmy Hendrix excerpt not by learning the physical gestures but by imitating the sound until the specimen coincides with the model. Sounds and gestures are linked in a way that cannot easily be

described, and there is always some freedom in the gesture. This also means that new gestures can result from the intention to produce new sounds, so that the vocabulary of gestures continues to expand. With alternative musical controllers, visual feedback can guide players in the exploration of instrumental gesture vocabulary, by helping them to find the appropriate gestures for imitating a previously recorded musical performance.
As the Voicer emulates a singing voice, the process of memorizing and imitating its sounds is a somewhat ecological one in a way: one has to find suitable gestures so that the voice comes out the right way, with its two essential components: the pitch and the articulation control. Combining a gestural indication and a musical example is surely the best way one can teach someone to use the filtering string (and a video excerpt is definitely a must for this kind of learning). The photosonic filtering gesture can be learned by listening to the "harmonic content" of the resulting sound, and this clearly shows the existence of a loop between the auditory feedback and the gestures produced.

4.3 Interpreting a score

Learning to perform a gesture using a score or gestural notation is a step further in the dissociation between a model and a specimen. In this situation, the learner interacts with written indications, from which it has to be deduced what sounds or movements are required. In other words, the interpreter must try to produce sounds that match the composer's intentions. Traditional music is based on the use of musical scores, but there is also a general consensus about the style (for example, the blues must have a specific "groove" to sound like the blues). In order to write music in terms of gestures, one must find new codes.
There are many possibilities, starting at the physical level (for example, by drawing the trajectory of the hand), but the writing will be mostly at the metaphorical level: one must find terms or symbols that the performer is able to decode to grasp the musical intentions. In fact, this often leads to defining a vocabulary and a syntax linking together the items of vocabulary. A good test of these "languages" is to see whether it is possible to write automatic computer programs that will interpret a language and render some basic sound signals. The rest depends on the gift of human inventiveness, but even genius will not work if the basis of the language is not clearly translatable into sound.
As the Voicer is an instrument capable of melody, the pitch part can be written in the conventional way, while spectral modeling requires some additional indications. But other instruments are clearly breaking down the frontiers between interpretation and improvisation, depending on the amount of information given in the score. The structure of the works written for the Voicer so far has generally been quite specific as regards the atmosphere and the timing, and quite free as to the individual gestures required.

4.4 Inventiveness

The question of inventing new gestures and new forms of expressiveness is an important one: playing an instrument of the kind described above is not just reproducing something previously played or written by another person, but it is also discovering our own gestural capacities.
New gestures often develop while one is testing a new instrument. For example, if an instrument is equipped with an "octave" button, one can also impose a rhythmic pulsation by

simply using this button, not for the sake of the octave transposition itself, but rather to obtain the "click" resulting from this sudden change.

In addition, there is a feedback loop between the invention of new gestures and the invention of new instruments: when a new gesture is discovered while testing a prototype, one often has to adjust the mapping of the instrument in order to make the movement more natural. For example, if the ambitus of a vibrato gesture performed with a stylus on a graphic tablet is too small or too large, it will be unsuitable because the movement required will be unnatural, tiring, and/or too difficult to perform precisely.

This means that musicians' creativity has plenty of scope when playing new instruments, especially during the test period, when the instrument-maker can change the mapping of the instrument.

5 Perspectives and conclusion

Much research still remains to be carried out on the implications of the latest digital instruments and the gestures they require, as far as composition, interpretation and improvisation are concerned. How can composers write their music so that performers will express what they originally meant to say? This universal question has come to the fore, due to the fact that new gestures, notations, and musical practises are emerging from new instruments. As pointed out by [Ungvary & Vertegaal, 2000] and [Pressing, 1984], interpretation is an "In-time" process, while improvisation is based on an "Out-of-time" cognitive process. This distinction has a lot to do with the latest digital instruments where the range of freedom allows both interpretation and improvisation. Composition, which is a top-down process (the macro-structure governs the micro-structure), therefore clearly needs further definitions of domains and transitions, especially if the musical style is a spectral one. Our investigations on this highly complex topic are still in the early preliminary stages.

On the other hand, these instruments have been assessed only at the empirical level, and the only test to which they have been put has been the musical result. Efforts could be made in the future to find means of assessing the interactions between digital music instruments and their human performers.

In conclusion, we have attempted to show in this paper how designers of digital musical instruments have to take into account the expressiveness required by performers. This expressiveness has been described at different levels: at the theoretical level, where various aspects have been discussed; at the practical level, where one has to find good controllers, good methods of synthesis and good mapping procedures in order to be able to introduce expressive features into the design of the instruments themselves; and at the level of the applications, since instruments are designed to be played, and music therefore has to be written and performers have to practise in order to bring expressiveness into the scores and into their playing.

References

[Arfib & Dudon, 2002] Arfib, D., Dudon, J., "A digital emulator of the photosonic instrument", Proceedings of the 2002 Conference on New Instruments for Musical Expression (NIME-02) pp. 128-131, Dublin, Ireland, May 24-26, 2002.

[Arfib & al., 2002] Arfib D., Couturier J.-M., Kessous L. (2002) : “Gestural stategies for specific filtering processes”, proceedings of DAFx02 conference, Hamburg, 26-28 sept 2002, pp. 1-6.

[Arfib & al., 2002b] Arfib D., Couturier J.-M., Kessous L., Verfaille V. (2002) : " Strategies of Mapping between gesture data and synthesis parameters using perceptual spaces ", *Organised Sound* , Cambridge University Press, Vol 7(2), pp. 127-144, August 2002.

[Arom & al., 1997] Simha Arom, Gilles Léothaud, Fréderic Voisin, “Experimental ethnomusicology: An interactive approach to the study of musical scales” in Perception and Cognition of Music, Edited by Irène Deliège and John Sloboda, Psychology Press ISBN 0-86377-452-0.

[Boulanger, 2000] 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.

[Cadoz & Wanderley, 2000] C. Cadoz, M. Wanderley, “Gesture-music”, *Trends in Gestural Control of Music*, CD-ROM, edited by Marcelo Wanderley and Marc Battier, IRCAM, 2000.

[Camurri & al., 2001] Camurri C., De Poli G., Leman M., Volpe G., A Multi-layered Conceptual Framework for Expressive Gesture Applications Workshop on Current Research Directions in Computer Music, Barcelona, Spain, pp. 29-34, 2001.

[Couturier, 2002] 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, pp. 176-178, 2002.

[Couturier, 2003] Couturier J.M, Arfib D., “Pointing Fingers: Using Multiple Direct Interactions with Visual Objects to Perform Music”, Proceedings of the 2003 Conference on New Interfaces for Musical Expression (NIME-03), Montreal, Canada, pp. 184-187, 2003

[Hunt & Wanderley, 2002] A. Hunt and M. Wanderley. Mapping performance parameters to synthesis engine. Organised Sound, 7(2), pp.103–114, 2002.

[Hunt & al., 2003] A. Hunt, M. Wanderley, and M. Paradis. The importance of parameter mapping in electronic instrument design. Journal of New Music Research, 32(4) pp. 429–440, 2003.

[Jordà, 1998] S. Jordà, “A graphical and net oriented approach to interactive sonic composition and real-time synthesis for low cost computer systems”, Digital Audio Effects Workshop Proceedings, 1998.

[Levin, 2000] G. Levin, "Painterly Interfaces for Audiovisual Performance", Master Thesis, Massachusets Institute of Technology, 2000. http://acg.media.mit.edu/people/golan/thesis/

[Menzies, 2003] Menzies, D. "Composing Instrument Control Dynamics ", *Organised Sound*, Vol 7/3, pp. 255-266, April 2003

[MIDI] http://www.midi.org, http://www.midi.com

[Moore, 1988] Moore, F. R. 1988. "The Dysfunction of MIDI", Computer Music Journal, 12(1), pp. 19-28.

[Pressing, 1984] Pressing, J. "Cognitive process in Improvisation", in Advances in Psychology, vol. 19, W.R. Crozier, A.J. Chapman Editors, North-Holland ISBN 0444875018, Elsevier Sciences Publishers B. V., 1984.

[Tactex] Tactex, touch surfaces, http://www.tactex.com/.

[Ungvary & Vertegaal, 2000] Ungvary T., Vertegaal R. "Cognition and Physicality in Musical CyberInstruments", in "Trends in Gestural Control of Music", edited by Marcelo Wanderley and Marc Battier, IRCAM, 2000.

[Verplank, 2000] Verplank B., Mathews M., Shaw R., "Scanned Synthesis", Proceedings of the 2000 International Computer Music Conference, pp. 368-371, Berlin, Zannos editor, ICMA, 2000.

[Wacom] Wacom tablets, http://www.wacom.com/

[Wanderley 2002] M. Marcelo Wanderley, Editorial, Organised Sound, 7(2), pp. 83–84, 2002.

[Zwicker, 1990] "7.2: Just-Noticeable Changes in Frequency", p. 163, Zwicker, E. and Fastl, H, Psychoacoustics: Facts and Models, 1990, Springer-Verlag, New York.