Cognition and Physicality in Musical CyberInstruments

Tamas Ungvary
Royal Institute of Technology
Stockholm, Sweden
ungvary@kacor.kth.se

Roel Vertegaal
Twente University
The Netherlands
roel@acm.org

Abstract

In this paper, we present the SensOrg, a musical CyberInstrument designed as a modular assembly of input/output devices and musical software, mapped and arranged according to functional characteristics of the Man-Instrument system. We discuss how the cognitive ergonomics of non-verbal and symbolic task modalities influenced the design of our hardware interface for asynchronous as well as synchronous task situations. Using malleable atoms and tangible bits, we externally represented the musical functionality in a physical interface which is totally flexible yet completely freezable.

Introduction

Musicians strive many years in order to connect their neural pathways to a vibrating segment of string, wood, metal or air. In many ways, learning how to play a musical instrument is dictated by the physical idiosyncrasies of the instrument design. A good instrumentalist typically needs to start almost from scratch when trying to play a new instrument. Even when musicians master their instruments, their sweet sorrow is not over. The chin marks of violinists, and the Repetitive Strain Injuries of drummers, bass players and pianists demonstrate the problems musicians face in the every day maintenance of their mastery. One might argue that the high learning curves and the physical contortion are symptoms of bad ergonomic design of traditional musical instruments. In this paper, however, we will take the *opposite* standpoint: there is a reason why acoustical instrument designs include physical hardship. Musicians need to achieve an extraordinarily sophisticated level of non-verbal communication. This functionality involves heavy cognitive requirements. From the point of view of *usability*, it is these cognitive requirements that dominate the physical design of the instrument. We should therefore approach the design of the physical Man-Instrument interface as a cognitive ergonomical problem.

In the four cognitive ergonomical criteria for assessing the *usability* of systems defined by Shackel [20], functionality is described by means of the concept *task*:

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

When what is important is expert achievement of the task result, *learnability* and *attitude* requirements are inhibited by the *ease of use* and *flexibility* requirements. The *ease of use* and *flexibility* requirements, in their turn, are conflicting. According to Polfreman [17], no single musical system is likely to fulfil individual task requirements. Systems should be customizable to other users and uses: the *flexibility* criterion. However, continuous flexibility of musical instruments would require constant adaptation and memorization from the musician. The cognitive load of dealing with a constantly changing system would never allow a musician to internalize his instruments and achieve the efficiency and effectiveness of the *ease of use* criterion [12].

It is these two conflicting issues, *flexibility* and *ease of use,* that we tried to address in the design of a computer music instrument. With the advent of computers powerful enough to perform processing of musical sound information, came the birth of computer music. What was special about the use of a computer for musical purposes was the ability to uncouple the input representation (physical manipulation) from the output representation (physical auditory and visual stimuli of a performance). In its most radical form, a programming language was used to specify the whole sound production process in a completely symbolic way. This was essentially an abstraction of the compositional process as it already existed in the classical tradition.

However, in the classical tradition, at the end of the compositional cycle there is the musician interpreting and communicating the symbolic representation to a human audience. We believe that although symbolic languages are extremely useful for describing the formal structure of a composition, the more formal they are, the more inappropriate they become for specifying the *whole* process of communicating non-verbal information. Before applying formal rules to non-verbal communication, we feel we need to learn more about what they should specify [8][23].

As a consequence, replacing a human interpreter with a formal language can be considered the foremost *usability* problem since the origins of computer music. With the advent of computers powerful enough for real-time processing of musical information, this problem was immediately addressed by the invention of computer music instruments. With such instruments, the human interpreter was basically back on stage, producing musical sounds using input devices to control real-time algorithms on a computer. In the design of computer music instruments, the ability to have a *loose coupling* between input device and the sound production process was again considered to be a key benefit. It allowed an indirection in the control of the sounding result by the performer, with generative computer processes adding to the richness of the music. It also allowed performers to use radically new input devices in radically new performance settings in ways not possible with traditional instruments. Indeed, we believe the artistic gains made with this approach were considerable. However, the usability problem shifted to the human interpreter: although the human communicator was now back in the cycle, uncoupling impaired the interaction of that communicator with his communication device: the instrument. The freedom of information structure in uncoupled instruments resulted in a mismatch of information flow across human input-output modalities. Traditional instruments seem far less affected by this problem. In the Hyperinstrument paradigm, Machover [14] tried to combine the qualities of a tight coupling in traditional instruments with the qualities of a loose coupling in computer devices.

Although we feel this was an important step towards recognizing the cognitive issues associated with the matching of input and output modalities, we felt that such augmentation of traditional instruments was, in many ways, a circumvention rather than a solution of the problem. Instead, we propose the new paradigm of *CyberInstruments*, which essentially consist of computer input and output modules, with algorithms in between. The modules are ordered such that modalities of human input and output are mapped with musical functionality performed by each module. In the design of SensOrg, our first CyberInstrument, we took a cybernetic approach in an attempt to solve the above cognitive ergonomical issues.

The Man - SensOrg system is seen as a *whole*, a whole of constituting elements with optimized mappings, rather than as a set of simple input-output relationships [4][18]. These elements include: rational and non-verbal intent, human actuator channel, input device, software functionality, output device and human perceptual channel, with information flowing across elements. In addition, feedback processes may occur at different levels between elements. We tried to use the *structure* of traditional instruments, rather than the instruments themselves, as an example of how such mappings might be achieved.

Cognitive Issues and Physical Design

We will first address some of the design considerations identified throughout the design cycle of the SensOrg Man-Instrument interface. We will then concentrate on the design of the physical Man-Instrument interface.

Achieving Nonverbal Communication: Symbolic and Non-Verbal Task Modalities

We consider the ability of music to directly communicate non-verbalizable information via non-verbal channels (in particular, as a form of paralinguistic audio) to be its most important functionality. Behavioural sciences have only recently started to address the role of non-verbalizable information in human functioning, perhaps relating it to specific hemispheric activity in the cerebral cortex [11]. Although it is unclear what the relation is between lower-level human emotion and higher-order associative intuition, these concepts for us define the essence of what is communicated in music. Although this has always been considered a speculative theory, Clynes [5][3] suggested early-on that passionate states of emotion correlate with patterns of muscular tension and relaxation in such a way that the direction of causal connection is no longer clear. We believe the same pattern occurs in many forms of non-verbal expression, from facial expressions, sighs, body position, gestures, paralinguistic speech, to touching one another [1]. Somehow, sensory-motor activity seems to be associated with the same cognitive functions that process non-verbal information. The efficiency of sensory-motor processing might be a requirement for managing the complexity of non-symbolic information in the process of expressing it, as well as in receiving it [7][25]. It is therefore that we consider sensory-motor tools essential in the process of musical expression. However, the above discussion does not imply that non-verbal communication has no rational structural elements. Although these elements are perhaps not of a highly conceptual level, *order* in the form of rhythmical structures, compositional sequences, etc., introduces a form of redundancy. According to Wiener, this redundancy may be essential to the *understanding* of information [35]. We believe that in the design of this order, analytical or verbalization processes can play an essential role. It is evident that in the communication of this design, symbolic representations are typically most effective.

We therefore regard the musical production cycle as a process in which non-verbal and symbolic task modalities complement each other, feeding back information from one to the other, and dominating at different stages of the process. It is in this light that we regard the traditional taxonomy of musical production modes: *composition*, *improvisation* and *performance* [22]. To us, this classification characterizes the time-complexity constraints of verbalization and non-verbalization in asynchronous and synchronous communication situations [6]. Composition maps onto asynchronous verbalization, while performance maps onto synchronous non-verbalization. Improvisation includes aspects of both. In the usability design of the SensOrg, the asynchronous verbalization constraint maps onto the *flexibility* criterion, and the synchronous non-verbalization constraint maps onto *ease of use* criterion. In order to communicate the verbalization process to the instrument, we needed to be able to specify symbolic relations. In an asynchronous situation, this is done by means of the computer equivalent of pencil and paper: a graphical user interface with visual feedback. These symbolic relationships are then mapped onto a sensory-motor representation in the form of a completely flexible set of physical interaction devices arranged in space. By *freezing* the physical representation of the internal state of the system, the human sensory-motor system can then be trained to achieve the *efficiency* and *effectiveness* required for expressing non-verbal information in synchronous situations. However, in order to continue support of the verbalization modality in synchronous situations, the physical interaction devices retain their capability to modify the symbolic relationships inside the system throughout, e.g., an improvisation.

Ease of Use: Reducing Problems of Cognitive Load and Recall by Freezing Functionality

As discussed above, our task modalities essentially reflect two ways of dealing with time-complexity constraints of information: complexity as-is (non-verbal mode) and complexity structured (symbolic mode). We believe *cognitive overload* (as a semantical form of information overload) might occur due to a mismatch between time-complexity constraints of functional information and time-complexity constraints of modalities that process that information. Miller [16] defines *information overload* as when channel capacity is insufficient to handle the information input. According to him, when the information input rate goes up, the output rate increases to a maximum and thereafter decreases, with the latter being a sign of overload. However, in our view, channel capacity depends on the interaction between the semantics of information *and* its rate (Schroder et al. 1967, see [10]). This yields a measure of *cognitive load* in the Wiener [35] sense, rather than *information load* in the Shannon-Weaver [21] sense (see Sveiby [24] for a discussion). Addressing problems of cognitive overload thus requires more than a simple reduction of information flow per channel by decreasing rate of information or by using multiple channels. It requires more than the selection of a channel on the basis of the load of other channels. It requires representing information in such a way that the processing of the meaning of that information is most efficient. Wiener suggests a negative relationship between entropy of meaning and entropy of information signal [35]. If this is correct, the usefulness of symbolic representations may be related to their ability to convey highly entropic semantics using little information. If we, however, assume a positive relationship between entropy of meaning and processing time required, we immediately see the benefit of non-symbolic representations. Thus, in designing a representation, the rate and entropy of the semantics that need to be communicated by the underlying function are important factors. This implies that a good mapping of the time-complexity constraints of a situation might ease cognitive load. Since in a cybernetic approach, we should regard human input/output as a feedback process, this mapping should not only occur in the design of system output, but also in the design of system input.

In an attempt to address some of the above issues in the hardware design, we collected a comprehensive set of input-output devices, carefully matching them onto the functionality of the system by identifying input/output channels associated with human processing of the information required by that functionality.

We used visual feedback for the more asynchronous symbolic functions; and auditory, tactile-kinesthetic feedback for the more synchronous non-verbal functions. We selected input devices in a similar fashion: buttons, faders, touchscreen and mouse for the more asynchronous symbolic functions; and buttons, faders, trackballs and touchsensors for the more synchronous non-verbal functions, in that order [32]. Later in this paper, we will discuss these mappings in more detail. Our mapping of I/O devices with software functionality also addressed the highly related issue of recall. We tried to introduce as much explicit knowledge into the real world as possible, attempting to reduce the requirements for knowledge in the head [37]. Essentially, we tried to externally represent the state of internal software functionality as much as possible. All I/O devices can be frozen into a unique spatial arrangement. Each device is coded by color, shape, orientation within groupings and textual information. For example, we put the touchscreen onto a picture of a Kandinsky painting.

The resulting device, the *Image-in-Kit*, is shown in Figure 1. By association of the position of virtual buttons with the arrangement of graphical objects on the picture, we tried to improve memorization of their function.

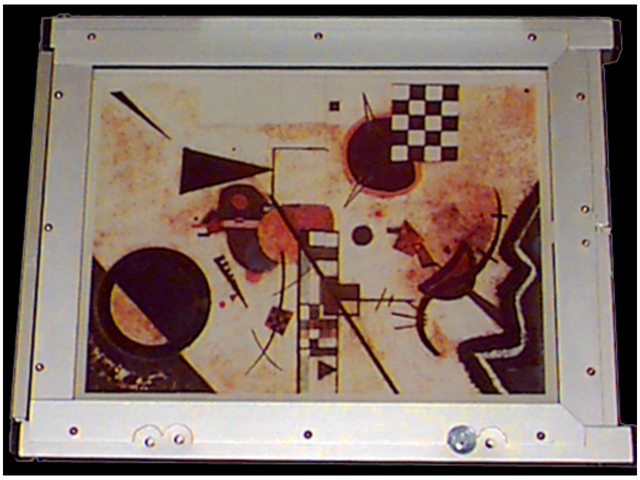


Fig.1. Video available in the original CD-Rom version. The Image-in-Kit: a touchscreen with *Gegenklänge* by Kandinsky.

Flexibility: Adaptation to Individuals and Task Situations by Malleable Functionality

In the design of the SensOrg, we wanted to combine the qualities of a tight coupling with the qualities of a loose coupling. As we have seen, in a loose coupling there is indirection, in a tight coupling there is not. The field of tension between tight and loose coupling is reflected in the conflicting requirements of the *ease of use* and *flexibility* criteria. We will now discuss how we made the system flexible, so that it could be adapted to different individuals and task situations such as compositional requirements. We could only choose to reflect the state of internal software functionality in the external devices if we also reflected the malleability of software functionality in the external devices. If the software functionality changes, the external devices should change and vice versa. If the software functionality stays the same, the external devices stay the same, as long as it is satisfactory. We did this by taking a modular approach to both software functionality and hardware devices. The software modules can be configured in an asynchronous, symbolic fashion by means of the graphical user interface. They can be driven in a synchronous, non-verbal fashion by manipulating the corresponding hardware modules. Similarly, hardware modules can be configured in a more asynchronous symbolic fashion by mapping them onto a software module, labeling them with a concept describing that functionality (with the device type being a label by itself), coloring them, positioning them freely within groups, and orienting groups freely within the instrument. They can be configured in a more synchronous, non-verbal fashion by selecting predefined configurations of software mappings using predefined buttons.

Apart from cognitive constraints, an important criterion for organizing hardware modules is the physical fit with human body parts. This is an extremely complex issue, where there are many individual differences. In addition, the task modality as related to musical functionality plays a role in this. Basically, the SensOrg hardware is so freely configurable, that it is almost totally adaptable to circumstances. It can accommodate individuals with special needs, including physical impairments. However, there are some basic functional and physical constraints which can be generalized across situations. The SensOrg is divided into two parts: one for the dominant hand, and one for the non-dominant hand. The dominant hand exercises mostly the more synchronous non-verbal functions, while the non-dominant hand exercises mostly the more asynchronous symbolic functions. This is because of the time-complexity constraints of information flow in these modalities. In the center of the dominant hand is the *FingerprintR*, a 3D sensor which conveys states of tension as exerted by subtle changes in force (see Figure 2). This is the most important device for the asynchronous non-verbal modality [31].

Fig. 2. Video. Video available in the original CD-Rom version. The FingerprintR knob as played by the fingers.

We will later discuss this issue in more detail. In order to meet the haptic feedback requirements of this process, the FingerprintR knob is concavely shaped, following the form of the finger with which it is played. This knob can be replaced to account for individual differences. In order to reflect the non-verbal intent in the muscle tension of the player, it is vital the upper-torso is in a relaxed position, while not relinquishing the ability to exert force. Since the SensOrg does not include devices operated by breathing force, the instrumentalist is typically seated like double bass players in an orchestra, so that his hand can be placed on the FingerprintR without necessarily exerting weight. Since the thumb opposes the other fingers, and can move more or less independently, the thumb of the dominant hand is used to control the more synchronous non-verbal button functions. In order to minimize the path and effort needed to press these buttons, they are placed below the FingerprintR knob. The area covered by the non-dominant hand is much larger. In the center of this area are groupings of faders and buttons. These are the most important devices for the more asynchronous symbolic functions.

Fig. 3. Video available in the original CD-Rom version. Flexipad with magnetic buttons and faders.

Button and fader modules stick to a position on a metal pad by means of small magnets. These pads (called *Flexipads*) can be positioned and oriented freely in space, and button and fader modules can be freely positioned on the pad (see Figure 3). Fader modules can be grouped so that they can be operated simultaneously with one hand gesture. Fader modules and button arrangements can be fitted to the hand by putting the hand onto a selection of devices, and then moulding the devices around the physical contour of the hand.

Overview of the SensOrg

Figure 4 shows how the discussed hardware modules fit together in the current implementation of the SensOrg CyberInstrument. All modules are mounted on gimbals attached to a rack with adjustable metal arms. This effectively allows them to be placed at any position or orientation. On the left, we see the Image-in-Kit touchscreen, with below it two Flexipads. On the Flexipads, modular structures of faders and buttons are shown. In the middle of the figure, we see the right hand subsystem with two FingerprintR knobs in the middle. Around these, two smaller Flexipads are arranged with real-time functionality. The above modules are the main physical ingredients of the SensOrg. Each hardware module is connected to software functions running on a PowerMac computer. The mapping of the input control data onto the musical parameter space is provided by means of the IGMA system, implemented in MAX [19]. This software front-end provides a framework for connecting hardware modules with musical functions which, for example, provide real-time high-level control of composition or sound synthesis algorithms. It also allows the output of such algorithms to be mapped to, e.g., a MIDI sound synthesizer producing an audible result. In the next sections, we will discuss our abstract design rationale for mapping input hardware modules to musical software functions for the more synchronous non-verbal tasks, omitting the implementational details. For a discussion of the IGMA software implementational details, and its functionality, we refer to [27]. We will first put forward a general design rationale for our mapping of input devices to the more synchronous non-verbal musical software functionality provided by IGMA. We will then study more closely the most prominent device for this type of functionality: the FingerprinterR.

Fig. 4. Audio clip available in the original CD-Rom version. The SensOrg. Click picture to hear the SensOrg composition *Fingerprint no. 1* by Tamas Ungvary.

Transducers, Feedback and Musical Function

When using a musical CyberInstrument in performance and improvisation task situations (using the more synchronous non-verbal musical functions), it is especially important that input devices be matched with the musical function they are to perform. Criteria for determining the suitability of an input device for a particular musical function are parameters such as: movement type sensed (position, movement, force); direction of movement (rotary or linear); degrees of freedom; resolution (continuous or discrete); agent of control (hand, fingers, lungs); and the type of feedback (tactile, kinesthetic, visual) [18][31].

When we refer to input devices as transducers, we explicitly incorporate their inherent output qualities in terms of the tactile, kinesthetic and visual feedback they provide. We refer to this feedback provided directly by the input device as *Primary Feedback*.

Figure 5 shows an elementary model that can be seen as providing a design rationale for our matching of musical functions with different transducers. We limited ourselves to rating the relationship between three relevant dimensions: synchronous non-verbal musical function, type of primary feedback, and a categorization of input devices based on movement type sensed (position, velocity, force); direction of movement (rotary or linear) and resolution (continuous or discrete)

For synchronous non-verbal musical functions, we restricted ourselves to a very simple categorization in which we define three types:

1) *Absolute Dynamical Functions***.** E.g., absolute selection of pitch, amplitude or timbre;

2) *Relative Dynamical Functions***.** E.g., modulation of a given pitch, amplitude or timbre;

3) *Static Functions***.** E.g., selecting pitch range, duration range, scale or transposition.

When we look at bowed instruments, control of pitch by finger position and control of timbre by bow position are examples of function 1. Relative control of pitch by vibrato and relative control of timbre by bow velocity and bow pressure are examples of function 2. Selecting different tunings, or putting a mute on or off are examples of function 3.

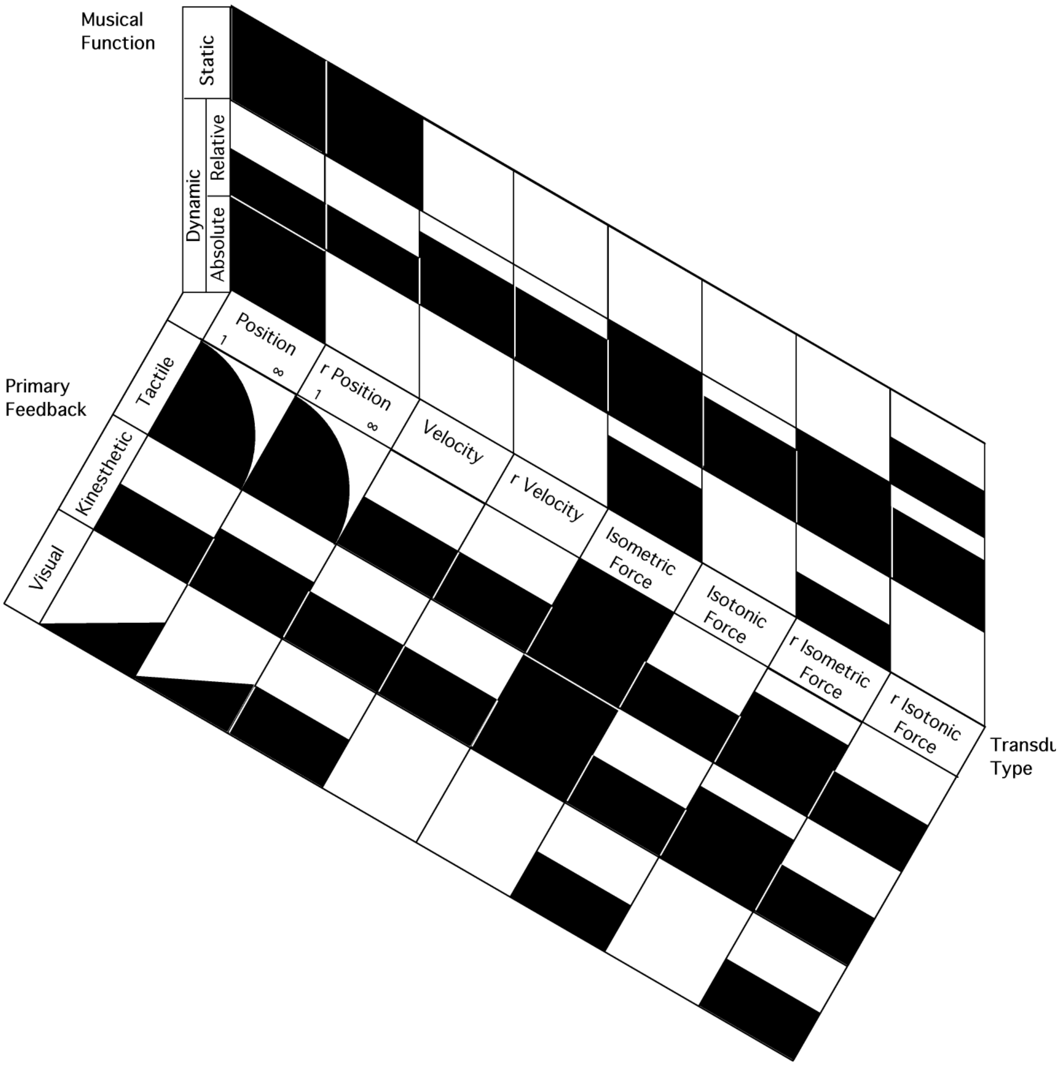


Fig. 5. Matching input device types with synchronous non-verbal functions.

	Physical Property	position	rotary position	velocity	rotary velocity	isometric force	isotonic force	isometric rotary force	isotonic rotary force
Resolution	discrete infinite	key button fader touch-screen tracker	rotary switch mod. wheel bend sensor rotary pot abs. joystick	mouse	dial track-ball	after-touch isometric joystick (Finger-printR)	accel-ero-meter	isometric joystick (Finger-printR)	pitch-bend wheel spring-mounted joystick

Table 1. A categorization of transducers.

Table 1 shows our categorization of transducers, which is largely based on a model by Mackinlay et al. [15]. In their model, input devices are decomposed into units which sense a particular physical property in a certain direction with a certain resolution. In this categorization we look at input devices from a human control perspective. We therefore make a distinction between isotonic and isometric force transducers. With isotonic force transducers, motion is needed to operate the sensor. With isometric force transducers, no motion is needed. This distinction is important when it comes to feedback properties. We categorize spring-mounted devices as isotonic force transducers since with these devices, the force exerted is directly proportional to the position sensed. In Figure 5, we only included the resolution parameter for position transducers.

As for Primary Feedback, we distinguish three types:

1) *Tactile Feedback* (sensed by the surface of the skin);
2) *Kinesthetic Feedback* (sensed internally by muscle and other receptors)
3) Visual Feedback.

We excluded primary auditory feedback since this is usually masked by the secondary auditory feedback produced by the musical result. We must stress that the relative importance of kinesthetic, tactile and visual feedback very much depends on the learning phase. For a musician, visual feedback of his movements is more important during the learning phase (governed by learnability), than during the expert phase (governed by ease of use). There is evidence which suggests that the inverse holds for kinesthetic feedback [12]. In the expert phase, tactile and kinesthetic feedback are important to allow a high level of precision for certain musical functions.

Selecting Input Devices for Synchronous Non-verbal Musical Functions

We started by rating our subjective notion of the effectiveness of each transducer type for operating each of the three musical functions. This is indicated in Figure 5 by the percentage of black filling in each square in the upper matrix. Our judgement was partially based on an identification of transducers and their typical musical function in traditional musical instruments. Independently of this, we then rated our subjective notion of the importance of each type of primary feedback in operating each transducer type. This is indicated in Figure 5 by the percentage of black filling in each square in the lower matrix. Note that for positioning devices, our rating strongly depended on resolution of the device (provided that this resolution is reflected in the physical feedback provided by the device, e.g., as fader clicks).

When we look at the pattern that emerges from our model in Figure 5, we can see that static functions seem best served by devices which can be left in a certain position. Depending on the resolution of the device, visual feedback of position can be regarded an important feature here (again, provided that the resolution is reflected in the physical feedback provided by the device). Consequently, position sensing devices (such as faders and buttons) seem the most appropriate here. For absolute dynamical functions, rotary devices seem less appropriate. Tactile feedback however, seems an important requirement. In positioning devices, this tactile feedback interacts with resolution. A good example of this is the use of frets on guitars for better control of absolute dynamical pitch selection. On a fretless bass guitar these frets are, amongst other things, sacrificed to better control relative dynamical functions (note that this requires a different fingering technique, which produces more tactile and kinesthetic feedback).

Relative dynamical functions seem best served by devices that sense a relative value (i.e., velocity or force), preferably with a large amount of tactile and kinesthetic feedback. Isometric joysticks (such as the FingerprintR) score well here, because they incorporate both qualities. We see an interesting pattern emerge: the more expressive the musical function, the more physical feedback of the device seems required. In an attempt to find explanations for this observation, we tried to analyse more closely our interactions with what appears to be the most physical device in the SensOrg, the FingerprintR.

Input, Physical Feedback and Musical Function: The FingerprintR

The FingerprintR was originally developed as a two-dimensional touch transducer for biocybernetic measurements by Manfred Clynes in the late sixties [3]. Its original construction, called Sentograph, comprised two sets of strain gauges mounted on a cantilevered arm of square cross-section, with the cantilever placed at right angles to the horizontal (*x)* and vertical (*z)* directions of measurement. A finger rest is mounted on top of the free end of the cantilever, which the user can press, thereby slightly bending the cantilever. The frequency range of measurement of the device is from 0-300 Hz, and the output 0 to 5 volts, corresponding to 8 to 1000 grams of weight on top of the finger rest. The resolution of the force measurement is better than .05 N, and the deflection of the cantilever is less than .04 mm / N. The original design was improved upon at the University of Uppsala, Sweden, to incorporate the measurement of back and forth finger pressure (*y*).

Also, the shape of the finger rest was made concave in order to obtain a better fit to the shape of the finger tip. In its current configuration, its three analog output signals are converted into MIDI controller messages by a modified Fadermaster with a resolution of 128 discrete steps in all directions [26]. After exerting pressure, the input parameters jump back to their rest value. The *x* and *y* dimensions of the device have a rest value 64, while the *z* component has a rest value 0. Clynes used his 2D Sentograph to measure essentic form, that is, the shape expressive actions may have in time during the expression of a particular emotional state, such as love, hate or anger [3]. During these experiments, muscle activity in the fore- and upper arm, shoulder and back was also recorded through measurement of the electrical activity produced in the muscular joints. Subjects were asked to repetitively express a particular emotion by means of finger pressure. Results indicated that in this way, given an emotional state, a corresponding and unique essentic shape might be recorded. Although the absolute position of the pressure components varied between trials and subjects, the shape of the pressure curves remained consistent between trials and even between subjects. Based on the results of these experiments, Clynes constructed a morphological language for the expression of emotional states.

In order to obtain a better insight into the nature of the FingerprintR as a musical instrument, we did some informal experimentation. We recorded a number of gestural phrases repeated over time. This in order to get a picture of the underlying form of the phrase and an impression of the accuracy with which these phrases could be repeated over time. The next sections will describe the experimentation and results, after which we will discuss the essence of interacting with this type of transducer.

Experimentation

During our experimentation, MAX was used to record the three directional parameters of pressure from a FingerprintR as controller messages with time stamps. Each parameter had a resolution of 128 discrete steps. The minimum of the *x* parameter corresponded to maximum pressure towards the left, while the maximum of this parameter corresponded to maximum pressure towards the right. The minimum of the *y* parameter corresponded to maximum pressure away from the body, while the maximum of this parameter corresponded to maximum pressure towards the body. The minimum of the *z* parameter corresponded to maximum downward pressure, while the maximum of this parameter corresponded to minimal downward pressure. In the first trial, a short circular gesture of about one second was repeated over one minute. The resulting recording was separated into 6 files of approximately 10 seconds length. In the second trial, a more complex movement pattern was recorded in the same way.

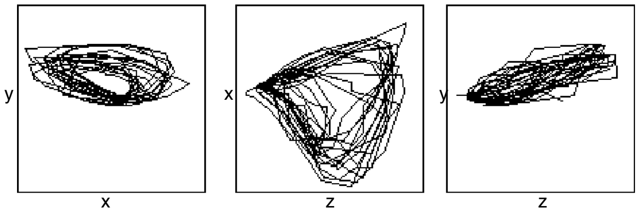
The resulting files were then processed in MAX to produce two different types of movement diagrams. The first type shows the three cross-sections of movement through the three-dimensional space constituted by the three pressure parameters, independent of time (Fig. 6a and 6b). The second type shows the relationship between the value of each pressure parameter and the velocity with which this value changes, again independent of time (Fig. 6c and 6d). This way, the underlying form of each pressure component can be visualized, providing a means of comparing the shape of each pressure components between trials in which this same shape is expressed. Results led to an analysis of the overall finger pressure pattern recorded by putting plastocine on top of the finger rest.

Results

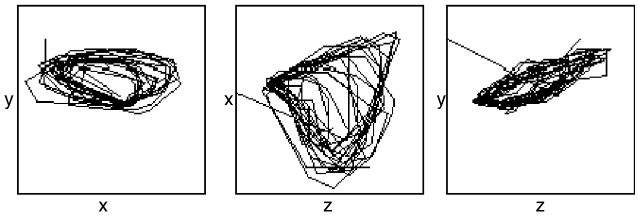
Figure 6a shows cross-sections of the movement of the three pressure components from an excerpt of 10 seconds length at the beginning of the first trial (the time series of this excerpt is shown in Figure 6g.).

Figure 6b shows an excerpt of movement of equal length at the end of the first trial. If we compare these images, we see that although a significant amount of jitter in absolute positioning can be observed, the shapes of the movements are quite similar.

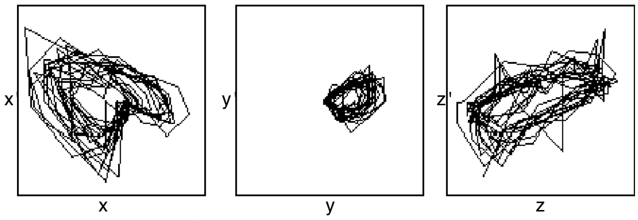
Figure 6c shows the dynamic behaviour of each pressure component. It shows that most of the differences between Fig. 6a and 6b can be explained by subtle differences in the dynamic behaviour of the individual components, particularly the z component. The x and y components are quite stable in both shape and scale. The z component shows a similar stability only in shape. The dynamical range of the x and z components is much greater than that of the y component. Closer examination of the overall finger pressure pattern in plastocine revealed that when moving on the x, the finger is rolled slightly horizontally. When moving on the y, however, the range in which this can be done without affecting the z parameter is much smaller. This difference is probably due to differences of the pivoting point of the finger tip between the x and y rotation of the finger tip. On the x, this pivoting point lies in the centre of the finger tip, while at the y, it lies towards the end of the finger tip. Figure 6e shows the overall movement pattern during the first trial. Figure 6f shows the overall movement pattern during the second trial, in which a more complex movement was made. If we compare the ranges of the components, we see that the range of y is consistently smaller than that of x and z. On the x and y dimensions, movement is attracted towards the centre, and on the z dimension, movement is attracted towards the minimum. This relates to the rest states of both the finger and device.



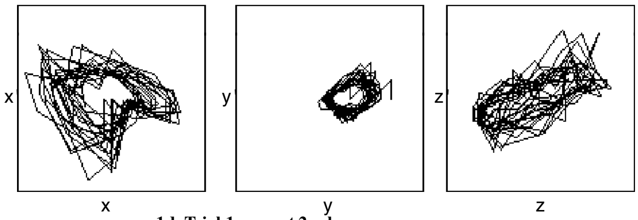
1a. Trial 1 excerpt 1: 3D space cross-sections.



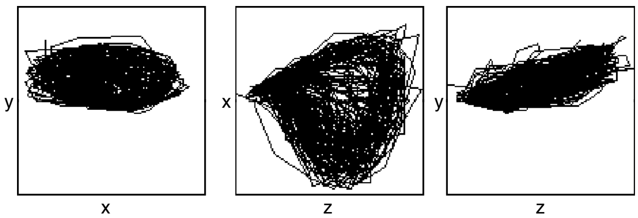
1b. Trial 1 excerpt 2: 3D space cross-sections.



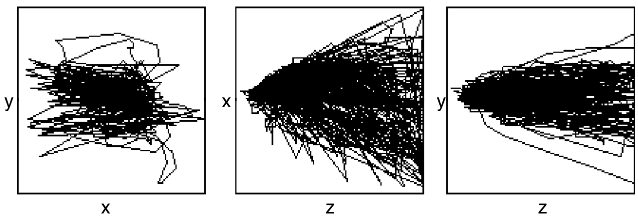
1c. Trial 1 excerpt 1: phase space x, y, z.



1d. Trial 1 excerpt 2: phase space x, y, z.



1e. Trial 1: 3D space cross-sections.



1f. Trial 2: 3D space cross-sections.

Fig. 6. Movement diagram of trial 1 and 2.

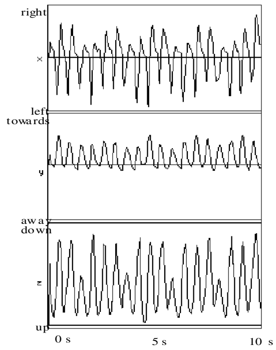


Fig. 2 Trial 1 excerpt 1 time series.

Discussion

What is the essence of the FingerprintR as a computer music controller? Our most important observation is that although the exact position of the individual pressure parameters at any given moment in time may not be controlled very accurately, the shape of the underlying phrase can be expressed quite accurately and consistently through time. We believe this capacity to accurately express the underlying shape of a musical phrase, without exact repetition of the physical parameters that constitute it, is a very important characteristic that gives the FingerprintR distinct qualities as a computer music controller.

From our experience with the FingerprintR, it seems rapidly accelerating and decelerating movements on the dimensions of pressure seem the most natural. An important explaining factor for this lies in the self-centering nature of the device, which corresponds to the rest state of the finger tip. For the x and y components, the interaction of the curved finger tip with the concavity of the finger rest might play a role. However, for the z component this is not a satisfactory explanation.

A more speculative, physiological explanation is given by Clynes [5]. He suggests that passionate states of emotion have underlying shapes of expression which feature rapidly accelerating and decelerating curves, which often correspond to patterns of muscular tension and relaxation. It is the force exerted by this muscular tension which is transduced by the FingerprintR. If we look at the nature of musical expression, we can indeed find some evidence for this notion. Highly curved acceleration and deceleration patterns are indeed predominant parameters for the communication of tension and relaxation in music.

We believe this capacity to produce rapidly accelerating and decelerating movements of musical parameters is another important characteristic of the FingerprintR as a computer music controller. The strong, almost synesthetic, sensation of reinforcement experienced while playing the FingerprintR might also be partly explained by this direct correspondence between the muscular tension that represents the emotion that needs expressing, and the expression itself. However, the nature of the transducer also plays a predominant role in this process. Particularly for the z component, there is a direct correspondence between the state of the input parameter and the state of the primary sensory feedback parameters: the experienced muscular tension and the experienced pressure on the tip of the finger. With the FingerprintR, the subtleness of expression can be influenced by refined control of the balance between the transfer of weight and the transfer of muscular force onto the device. This balancing of weight and muscular force to control subtleness of expression is not without precedent in musical instruments.

When we look at musical instruments in terms of the transducers that constitute a particular effect, we see that subtle modulation of pitch and loudness is often achieved by means of pressure control with tactile and muscular tension feedback. When examining the violin, for instance, we might identify at least 7 transducers: 4 absolute position transducers for pitch (finger positioning on strings), a position transducer for timbre (position of the bow relative to the bridge), a movement transducer, affecting timbre and whether the instrument sounds or not (bow movement), and a pressure transducer affecting dynamics (bow pressure) [18].

Most of these transducers can be freely combined in order to produce complex effects. However, another transducer might be added to the model. Although we are aware that the process of movement constituting vibrato on a violin is a complex one, which should be studied in greater detail before making any assumptions on the underlying principles, we believe that the subtle changing of the length of the vibrating part of the string in this process is not modelled satisfactorily by a change in absolute position of the finger on the string. During vibrato, the tip of the finger firmly presses the string against the fingerboard. Pressure is also exerted on the finger tip in the direction of the bridge. However, this does not result into an actual displacement of the finger. Our hypothesis is that as a descriptive component which adds to the traditional view of this kind of pitch modulation, the function of the pressure or force exerted on the finger tip in the direction of the bridge could be considered.

Clearer examples of subtle expression of emotion by means of pressure exerted with fine motor control might be the subtle raising of pitch of a guitar string by exerting upward finger pressure onto the string, or the use of embouchure in wind instruments.

Designing, Composing, and Educating for Physicality in Dynamical Expression

All in all, there seems to be an intimate relationship between non-verbal intent, synchronous non-verbal musical functionality, physical human actuator channels, input devices and the physical feedback these devices provide to the human perception channels as output devices in their own right. During musical performance, feedback makes it possible to regard these elements as one Gestalt. Not only can the muscle tension needed for subtle dynamical expression only be built up when movement is restricted, this tension in the muscles of a performer, as conveyed by muscular receptor and tactile feedback, has a function for the performer as a representation of the audible result for which this same tension provided the input. This of course only holds when the audible result has some causal relationship with the actual physical effort.

As such, physical effort is an important musical parameter for both the artist and the audience. Artists need to feel a piece as it is being created and performed. When one is writing a piece for instruments, the composer needs to consciously or subconsciously take into account the various physical aspects of the particular instrument. According to Waisvisz [13]: "One cannot compose the musical tension structure uniquely by formal rules; one can only compose for it. [...] One has to suffer a bit while playing." Consequently, when we design the physical interface of a computer music instrument, we need to carefully match transducers with the musical function they perform, taking the feedback requirements of dynamical expression into account. In addition, a good causal mapping can ensure that tension is properly translated to the audible result.

As for the audience, it perceives the physical effort as the cause and manifestation of the musical tension of the piece [13]. Even if pushing a button does not need a complex movement, like hitting a key on the piano it can be done in an expressive way. Educating proper instrumentalists for CyberInstruments is a way of achieving this, and perhaps one of the most important of our future goals.

Conclusions

In this paper, we presented the SensOrg, a musical CyberInstrument designed as a modular assembly of input/output devices and musical generator software, mapped and arranged according to functional characteristics of the Man-Instrument system. We have shown how structuring access to, and manipulation of information according to human information processing capabilities are essential in designing instruments for composition, improvisation and performance task situations. We regard this musical production cycle as a process in which non-verbal and symbolic task modalities complement each other, feeding back information from one to the other, and dominating at different stages of the process. We identified how these task modalities may be mapped onto the time-complexity constraints of a situated function: asynchronous verbalization vs. synchronous non-verbalization. By matching time-complexity constraints of musical functions, transducers, human I/O channels and body parts, we carved functional mappings between the more asynchronous symbolic elements on the one hand, and the more synchronous non-verbal elements on the other. To allow these mappings to be adaptable to individuals and situations, hardware as well as software configurations were designed to be totally flexible. Therefore, all physical interface devices are mounted on gimbals, attached to a rack with adjustable metal arms. To allow mappings to be effective, however, physical interface devices can be frozen in any position or orientation. The mapping of the input control data from these devices onto the musical parameter space is provided by means of IGMA software modules supporting real-time high-level control of synthesis and composition algorithms. To learn more about the development of the SensOrg, please visit our website: http://www.speech.kth.se/kacor/sensorg/main.htm

Acknowledgements

We gratefully acknowledge the support of our research by the Bank of Sweden Tercentenary Foundation, the Royal Institute of Technology and the Knut & Alice Wallenberg Foundation. Thanks also to Mario Zlatev, without whom this instrument would have remained a laboratory prototype, and Michael Kieslinger, who did a wonderful job programming the software functionality. Thanks also to Tibor Barany for the movie examples and web site.

References

1
Argyle, M. 1967. *The Psychology of Interpersonal Behaviour*. London: Penguin Books.

2
Cadoz, C., A. Luciani, and J.L. Florens. 1984. Responsive Input Devices and Sound Synthesis by Simulation of Instrumental Mechanisms: The Cordis System. *Computer Music Journal*. 8 (3).

3
Clynes, M. 1970. "Toward a View of Man." In *Biomedical Engineering Systems,* Clynes, M. and J. Milsum, eds. New York: McGraw-Hill, pp. 272-358.

4
———. 1982. *Music, Mind and Brain, The Neuropsychology of Music*. Plenum Press.

5
———. 1980. "The Communication of Emotion; Theory of Sentics." In: *Emotion, Theory, Research and Experience*. Volume 1: Theories of Emotion, R.Plutchik and H. Kellerman Eds. New York: Academic Press.

6
Dix, A., Finlay, J., Abowd, G. and Beale, R. 1993. *Human-Computer Interaction*. London: Prentice-Hall.

7
Edwards, B. 1979. *Drawing on the Right Side of the Brain*. Los Angeles: J.P. Tarcher.

8
Gabrielsson, A. "Music Performance." In Deutsch, D. (ed). *The Psychology of Music* (2nd ed. in press).

9
Gibet, S. and J.L. Florens. 1988. "Instrumental gesture modeling by identification with time-varying mechanical models." in *Proceedings of the 1988 International Computer Music Conference*. San Francisco, International Computer Music Association, pp. 28-40.

10
Hedberg, B. 1981. *How Organisations Learn and Unlearn*. I Nyström & Starbuck.

11
Iaccino, J. 1993. *Left Brain-Right Brain Differences*. Hillsdale, NJ: Lawrence Erlbaum.

12
Keele, S.W. 1968; "Movement Control in Skilled Motor Performance". *Psychological Bulletin* 70, pp. 387-402.

13
Krefeld, V. 1990. "The Hand in the Web: An Interview with Michel Waisvisz". *Computer Music Journal* 14(2).

14
Machover, T. 1986-1992. *Classic Hyperinstruments. A Composer's Approach to the Evolution of Intelligent Musical Instruments*. MIT Media Lab.

15

Mackinlay, J.D., S. K. Card, and G. G. Robertson. 1990. “A Semantic Analysis of the Design Space of Input Devices.” *Human-Computer Interaction*. 5: 145-190.

16

Miller, J. 1978. *Living Systems*. McGraw-Hill.

17

Polfreman, R., and J. A. Sapsford-Francis. 1995. "A Human Factors Approach to Computer Music Systems User-Interface Design". In *Proceedings of the 1995 International Computer Music Conference*. San Francisco, International Computer Music Association, pp. 381-384.

18

Pressing, J. 1990. "Cybernetic Issues in Interactive Performance Systems." *Computer Music Journal* 14(1): 12-25.

19

Puckette, M., and D. Zicarelli.1990. *MAX-An Interactive Graphic Programming Environment*. Menlo Park, Calif.: Opcode Systems.

20

Shackel, B. 1990. "Human Factors and Usability." In *Human-Computer Interaction: Selected Readings,* Preece, J. and Keller, L., eds. Upper Saddle River, NJ: Prentice Hall.

21

Shannon, C., and Weaver, W. 1949. *The Mathematical Theory of Communication*. Chicago: University of Illinois Press.

22

Sloboda, J. 1985. *The Musical Mind: The Cognitive Psychology of Music*. Oxford: Oxford University Press.

23

Sundberg, J., A. Askenfelt and L. Fryd. 1983."Musical Performance: A Synthesis-by-Rule Approach." *Computer Music Journal* 7(1): 37.

24

Sveiby, K. *What is Information?*. Sveiby Knowledge Management, Australia 1998. Available at: http://www.sveiby.com.au/Information.html.

25

Tenney, J. 1986 (reprint). *META/HODOS and META Meta/Hodos*. Hanover, N.H.: Frog Peak Music.

26

Ungvary, T., and P. Lundén. (1993) "Notes on a prototype of human-computer interface for artistic interaction: Sentograffito". *Proceedings of the Stockholm Music Acoustics Conference*, Royal Academy of Music, No. 79.

27

———, and M. Kieslinger. 1998"Creative and Interpretative Processmilieu for Live-Computermusic with the Sentograph." In: *Controlling Creative Processes in Music*. P. Lang (ed), Frankfurt am Main.

28

Vertegaal, R. 1995. "The Standard Instrument Space Libraries: Demonstrating the Power of ISEE". in

Proceedings of the 1995 International Computer Music Conference. San Francisco, International Computer Music Association.

29
———, and B. Eaglestone. 1996. "A Comparison of Input Devices in an ISEE Direct Timbre Manipulation Task." Amsterdam: Elsevier Science, *Interacting with Computers* 8(1). pp. 13-30.

30
———, and E. Bonis. 1994. "ISEE: An Intuitive Sound Editing Environment." *Computer Music Journal* 18(2): 21-29.

31
———, and T. Ungvary. "The Sentograph: Input Devices and the Communication of Bodily Expression." in *Proceedings of the 1995 International Computer Music Conference*. San Francisco, International Computer Music Association, pp. 253-256.

32
———. 1996. "Towards a Musician's Cockpit: Transducers, Feedback and Musical Function". In *Proceedings of the 1996 International Computer Music Conference*. San Francisco, International Computer Music Association.

33
———, and B. Eaglestone. 1998. "Looking for Sound? Selling Perceptual Space Using Hierarchically Nested Boxes." In *Summary of ACM CHI'98 Conference on Human Factors in Computing Systems*. Los Angeles: ACM SIGCHI.

34
———, B. Eaglestone, et al. 1994. "An Evaluation of Input Devices for Use in the ISEE Human-Synthesizer Interface." in *Proceedings of the 1994 International Computer Music Conference*. San Francisco, International Computer Music Association, pp. 159-162.

35
Wiener, N. 1948. *Cybernetics*. Cambridge, MA: The MIT Press.

36
Winkler, T. 1995. "Making Motion Musical: Gesture Mapping Strategies for Interactive Computer Music." in *Proceedings of the 1995 International Computer Music Conference*. San Francisco, International Computer Music Association, pp. 261-264.

37
Zhang, J., and N. Norman. 1994. "Representations in Distributed Cognitive Tasks." *Cognitive Science* 18, pp. 87-122.

