Sound to Sense, Sense to Sound
A State of the Art in
Sound and Music Computing

Pietro Polotti and Davide Rocchesso, editors
Appendix A

Controlling Sound Production

Roberto Bresin and Kjetil Falkenberg Hansen
Department of Speech, Music and Hearing, KTH, Stockholm

Matti Karjalainen, Teemu Mäki-Patola, Aki Kanerva, and Antti Huovilainen
Laboratory of Acoustics and Audio Signal Processing, Helsinki University of Technology

Sergi Jordà, Martin Kaltenbrunner, Günter Geiger, Ross Bencina
Music Technology Group, University Pompeu Fabra, Barcelona

Amalia de Götzen¹ and Davide Rocchesso²
¹Department of Information Engineering, University of Padua
²Department of Art and Industrial Design, IUAV University of Venice

Introduction

This appendix is a collection of five contributions by different authors. Each contribution is self-contained and can be read separately. The scope ranges from new musical instruments to sonification of interactive books. The ap-
Appendix starts with an introductory section, in which general concepts and problems related to sound control are presented. The authors outline the important issue of the choice of sound models that can provide suitable and responsive feedback in continuous control, as in the case of sound generated by body motion and sound as feedback in interaction. In the following sections, some recent and successful applications are illustrated. The first two, DJ Scratching and Virtual Air Guitar, focus on the control of virtual musical instruments, and in particular on control that generates sense through sound production. The last two sections, The reacTable and The Interactive Book, focus on the control of sounding objects, and are characterised by applications that control sounds that produce sense. The section sequence can also be seen as ordered according to an increasing level of sound model abstraction (from sampled sounds to cartoon sounds) and decreasing level of complexity of control gesture (from DJ scratching to simple sliding).

The authors of the different sections are:

A.1 Introduction
Roberto Bresin and Davide Rocchesso

A.2 DJ scratching with Skipproof
Kjetil Falkenberg Hansen and Roberto Bresin

A.3 Virtual air guitar
Matti Karjalainen, Teemu Mäki-Patola, Aki Kanerva, and Antti Huovilainen

A.4 The reacTable*
Sergi Jordà, Martin Kaltenbrunner, Günter Geiger, Ross Bencina

A.5 The interactive book
Amalia de Götzen and Davide Rocchesso
A.1 Background

There are a few emerging facts that are conditioning our present approaches to the study of sound. Sensors of many different kinds are available at low cost and they can be organised into networks. Computing power is generously available even in tiny and low-power processors that can be easily embedded into artefacts of different nature and size. New design strategies that take advantage of these technological opportunities are emerging: physical computing, natural interaction, calm technologies are some of the many buzzwords that are being proposed as labels for these new trends in design. For the purpose of sound-based communication, the concept of embodied interaction (Dourish, 2001) is particularly significant. Embodiment is considered a property of how actions are performed with or through artefacts, thus embracing the position that treats meanings as inextricably present in the actions between people, objects, and the environment. A key observation that emerges from embodied interaction examples is that human interaction in the world is essentially continuous and it relies on a complex network of continuous feedback signals. This is significantly important if one considers that most interfaces to technological artefacts that are currently being produced are developed around switches, menus, buttons, and other discrete devices. The design of graphical user interfaces has been largely inspired by ecological psychology and concepts such as direct perception and affordances (Gibson, 1979). When designing embodied interfaces, we call for a reconciliation of ecological psychology and phenomenology that looks, with equal emphasis, at the objects and at the experiences. By means of physical modelling we can represent and understand the objects. By direct observation we can tell what are the relevant phenomena, which physical components are crucial for perception, what degree of simplification can be perceptually tolerated when modelling the physical reality. Specifically, sound designers are shifting their attention from sound objects to sounding objects, in some way getting back to the sources of acoustic vibrations, in a sort of ideal continuity with the experimenters of the early twentieth century, especially futurists such as Luigi

\footnote{Parts of this section are extracted and modified from a recent work by Rocchesso and Bresin (2007).}
Russolo and his intonarumori. In the contemporary world, sounding objects should be defined as sounds in action, intimately attached to artefacts, and dynamically responsive to continuous manipulations. As opposed to this embodied notion of sound, consider an instrument that came shortly after the intonarumori, the theremin invented 1919 by Lev Termen. It is played by moving the hands in space, near two antennae controlling amplitude and frequency of an oscillator. Its sound is ethereal and seems to come from the outer space. This is probably why it has been chosen in the soundtracks of some science-fiction movies (see the documentary by Martin [2001]). Even though relying on continuous control and display, the lack of physical contact may still qualify the theremin as a schizophonic artefact, and it is not by coincidence that it is the only musical instrument invented in the twentieth century (the schizophonic age) that was used by several composers and virtuosi. Indeed, nephews of the theremin can be found in several recent works of art and technology making use of sophisticated sensors and displays, where physical causality is not mediated by physical objects, and the resulting interaction is pervaded by a sense of disembodiment.

A.1.1 Sound and motion

Sounds are intimately related to motion, as they are usually the result of actions, such as body gestures (e.g. the singing voice) or mechanical movements (e.g. the sound of train wheels on rails). In the same way as we are very accurate in recognizing the animate character of visual motion only from a few light points corresponding to the head and the major limb-joints of a moving person (Johansson [1973], we are very sensitive to the fluctuations of auditory events in the time-frequency plane, so that we can easily discriminate walking from running (Bresin and Dahl [2003]) or even successfully guess gender of a person walking (Li et al. [1991]). It is not a surprise that gestures are so tightly related with sound and music communication. A paradigmatic case is that of the singing voice, which is directly produced by body movements (see also Chapters 6.3 and 7 for overviews on gestures in music performance). In general, gestures allow expressive control in sound production. Another example is DJ scratching, where complex gestures on the vinyl and on the cross-fader are
used for achieving expressive transformation of prerecorded sounds (Hansen and Bresin, 2004). In the context of embodied interfaces, where manipulation is mostly continuous, it is therefore important to build a gesture interpretation layer, capable to extract the expressive content of human continuous actions, such as those occurring as preparatory movements for strokes (see Dahl, 2004). Body movements preceding the sound production give information about the intentions of the user, smoother and slower movements produce softer sounds, while faster and sudden movements are associated to louder sounds. Gestures and their corresponding sounds usually occur in time sequences, and it is their particular time organisation that helps in classifying their nature. Indeed, if properly organised in time, sound events can communicate a particular meaning. Let us consider the case of walking sounds (Giordano and Bresin, 2006). The sound of a step in isolation is difficult to identify, while it gives the idea of walking if repeated a number of times. If the time sequence is organised according to equations resembling biological motion, then walking sounds can be perceived as more natural (Bresin and Dahl, 2003). In addition, if sound level and timing are varied, it is possible to communicate different emotional intentions with walking sounds. In fact, the organisation in time and sound level of structurally organised events, such as notes in music performance or phonemes in speech, can be controlled for communicating different emotional expressions. For instance in hyper- and hypoarticulated speech (Lindblom, 1990) and in enhanced performance of musical structure (Bresin and Friberg, 2000) the listener recognises the meaning being conveyed as well as the expressive intention on top of it. Research results show that not only we are able to recognise different emotional intentions used by musicians or speakers (Juslin and Laukka, 2003) but also we feel these emotions. It has been demonstrated by psychophysical experiments that people listening to music evoking emotions experience a change in biophysical cues (such as blood pressure, etc.) that correspond to the feeling of that specific emotion and not only to the recognition. Krumhansl (1997) observed that sad music produced largest changes in heart rate, blood pressure, skin conductance and temperature, while happy music produced largest changes in measures of respiration. Music and sound in general have therefore the power to effect the variation of many physiological parameters in our body. These results could be taken into account in the
design of more engaging applications where sound plays an active role.

### A.1.2 Sound and interaction

An important role in any controlling action is played by the feedback received by the user, which in our case is the sound resulting from the user’s gestures on an object or a musical instrument. Therefore sound carries information about the user’s actions. If we extend this concept and consider sounds produced by any object in the environment we can say that sound is a multidimensional information carrier and as such can be used by humans for controlling their actions and reactions relatively to the environmental situation. In particular, humans are able to extract size, shape, material, distance, speed, and emotional expression from sonic information. These capabilities can be exploited to use sound as a powerful channel of communication for displaying complex data. Interactive sonification\(^2\) is a new emerging field where sound feedback is used in a variety of applications including sport, medicine, manufacturing, and computer games. There are many issues that have been raised in such applications, and answers are expected to come from interaction design, perception, aesthetics, and sound modelling. For instance, how do we achieve pleasant and effective navigation, browsing, or sorting of large amount of data with sounds? In the framework of the Sounding Object project, the concept of sound cartoonification has been embraced in its wider sense and applied to the construction of engaging everyday sound models. Simplified and exaggerated models have been proved to be efficient in communicating the properties of objects in actions, thus being excellent vehicles for informative feedback in human-artefact communication. For instance, it has been shown that temporal control of sound events helps in communicating the nature of the sound source (e.g. a footstep) and the action that is being performed (walking/running). The possibility of using continuous interaction with sounding objects allows for expressive control of the sound production and, as a result, to higher engagement, deeper sense of presence, and experiential satisfaction. Low-cost sensors and recent studies in artificial emotions enable new forms of interaction using

---

\(^2\)See [Hunt and Hermann \(2005\)](Hunt2005) for a recent overview of the field
previously under-exploited human abilities and sensibilities. For instance, a cheap webcam is sufficient to capture expressive gesture nuances that, if appropriately interpreted, can be converted into non-visual emotional cues. These new systems, albeit inexpensive and simple in their components, provide new challenges to the designer who is called to handle a palette of technologies spanning diverse interaction modalities. In the future, the field of interaction design is expected to provide some guidelines and evaluation methods that will be applicable to artefacts and experiences in all their facets. It is likely that the classic methods of human-computer interaction will be expanded with both fine-grained and coarse-grained analysis methods. On a small scale, it is often necessary to consider detailed trajectories of physical variables in order to make sense of different strategies used with different interaction modalities. On a large scale, it is necessary to measure and analyze the global aesthetic quality of experiences.
A.2 DJ scratching with Skipproof

A.2.1 Concept

*Scratching* has been established as a practice of treating the turntable and a sound mixer as a musical instrument, and is one of the usual DJ playing styles. Scratching and the related playing style *beat juggling* require much training for mastering the complex gesture control of the instrument.

Skipproof, a patch written for pd (Pure Data, Puckette [1996]), is both a virtual turntable and mixer, and an application for exploring the musical language of DJs. The name Skipproof is taken from a feature found on DJ-tools records called a *skip proof* section, where a sound (or set of sounds) are exactly one rotation long and repeated for a couple of minutes. If it happens that the needle jumps during a performance, the chances are quite good it will land on the same spot on the sound, but in a different groove. The main purpose is to “scratch” sound files using gesture controllers of different kinds.

Scratch performances are normally built up by the sequential executions of well-defined hand movements. Combinations of a gesture with the hand controlling the record and a gesture with the hand controlling the fader on the mixer are called scratch techniques. These have become common language for DJs, and they refer to them by name (*baby, crab, flare*) and their characteristics (*1-click flare, 2-click flare, reversed tear*). About one hundred techniques were recognised and more than 20 analysed in previous studies (e.g. Hansen [2002], Hansen and Bresin [2004]). The analysis focused on measuring the movement of the record and the fader. Based on the results, models of synthesised scratch techniques were constructed.

Originally, the software was intended to be a tool for reproducing and exploring the modelled scratch techniques with different characteristics. For instance, we could change speed and extent of the record movements while maintaining the fader gesture. The graphical user interface allows for easy experimenting with the models. We decided to develop the patch so it could

---

3 The crossfader normally has a very short span from silent to full sound, often less than a millimeter. It is more realistic to consider this fader as a switch than a fader.
be used also for controlling sound files in a turntable-like manner. The user is not restricted to play sound files, but can control sound synthesis techniques such as physics-based models. In the end, Skipproof is combining features from turntable, mixer, vinyl records and modelled DJ performances.

A.2.2 Strategy-implementation

Skipproof most important feature is to allow execution and control of the scratch technique models, and to do this with customisable interfaces and controllers. The graphical user interface, made with GrIPD (Sarlo, 2003), serves as a collection of controllers a DJ normally would expect to find in addition to new controllers, see Fig. A.1.

All sound samples that are used with the patch are 1.6 seconds long and looped, reflecting the idea of a skip proof track. The sounds are taken from a DJ tool record and are all favored for scratching.

There are currently 12 typical scratch techniques ready to use in the patch. These are models based on analysis of recordings made by a professional DJ (Hansen, 2002). Each technique consists of a forward–backward movement of the record and synchronised actions with the fader. Changing the way a technique is performed is dependent on the controller gesture, and this connection between gesture and sound output can be customised. When using several controllers, different mapping schemes will be realised (Hunt and Kirk, 2000).

In the first version of Skipproof, the techniques can be altered in extent and speed of the movement, the two most important parameters in scratching (Hansen and Bresin, 2004). Both extent and speed can be exaggerated, so for example scratch techniques are performed faster and with larger movements than in a real situation.

All the elements in the GUI can be accessed by any controller. Until now, Skipproof has been played by a number of devices, including computer peripherals and tablets, MIDI devices, Max Mathews’ RadioBaton (Boulanger and Mathews, 1997) and various sensors (e.g. LaKitchen’s Kroonde and Toaster.
sensor interfaces, [Coduys et al. 2004]. Among these controllers, the Radio-Baton in combination with sliders, buttons, magnetic field, flexion and light sensors has good potential to be a rewarding interface. Even a computer mouse in combination with other sensors is quite effective, in contrast to many new musical interfaces.

Recently, the concept of Skipproof has been successfully implemented and tested with the reacTable (see Section A.4). Scratch technique models were assigned to reacTable objects, where one object contained fader movement, one contained record movement and a third object contained the sound.

### A.2.3 Results-expectations

A professional DJ has been using Skipproof in two live concert situations, see Fig. A.2. The first time, the RadioBaton was controlling the virtual turntable and technique models. The RadioBaton sticks were replaced with treated gloves, so the musician could control the turntable speed and the techniques easily, moving his hand in a 3-D space over the antennae.

In a second concert, the DJ used again the RadioBaton, but now he could trigger the models based on the hand gesture, as compared to the earlier version with button triggers. The approach toward a defined area of the antennae was measured in speed and distance, and this gesture determined scratch speed and extent. For this performance, a light switch was used to replace the fader. A light sensor was placed pointing directly toward a lamp, and the DJ could break the light beam by placing his hand close to the sensor, or by waving the hand after the light with outstretched fingers. In that way, controlled and rapid sound on–off events were possible, just like with a fader.

The DJ commented later that although there was a lack of mechanical feedback from the interface, it opened up to new possibilities. Controlling recorded techniques was considered to be hard, especially to get the correct tempo. A scratch DJ uses physical markings on the vinyl (stickers, label) to see where in a sound the pick-up is, and this feature is moved from the controller to the GUI in Skipproof. This takes time getting comfortable with and is not at all optimal.
Users without DJ experience have found the set-up with RadioBaton as turntable and light switch as fader to be both intuitive and exciting, and quite fast they could perform with it in a simple fashion. Using the reacTable (see Section A.4) as the controller interface is a very interesting approach that addresses the need from non-professional DJs to perform simple scratch sequences.

DJs seem overall to be interested and intrigued by new technology and possibilities. It is possible to build a system that enhances DJ performances of scratching, and it is desirable to experiment with the equipment currently preferred by DJs. Performing with models of scratch techniques is still a novel approach that needs to be tested more.
Figure A.1: Skipproof graphical user interface (left) and synthesised techniques (right). The GUI contains a combination of standard controllers found on turntables and mixers, and novel controllers such as technique triggers, visual representation of sound, sliders for changing control parameters, and more. The synthesised techniques are saved as sets of two tables; the upper is the record movement and the lower is the crossfader movement.
Figure A.2: DJ 1210 Jazz scratching with Skipproof in a concert. The Radio-Baton is operated with his right hand. Beside the computer screen, there is a lamp for a light sensor, on the floor, there is a rack of foot switches. The DJ’s left hand is on the crossfader.
A.3 Virtual air guitar

A.3.1 Concept

A combination of hand-held controllers and a guitar synthesiser with audio effects is called here the Virtual Air Guitar (VAG). The name refers to playing an “air guitar”, that is just acting the playing with music playback, while the term virtual refers to making a playable synthetic instrument. Sensing of the distance of hands is used for pitch control, the right hand movements for plucking, and the finger positions may in some cases be used for other features of sound production. The synthetic guitar algorithm supports electric guitar sounds, augmented with sound effects and intelligent mapping from playing gestures to synthesis parameters.

Electronic and computer-based musical instruments are typically developed to be played from keyboard, possibly augmented by foot, breath or other controllers. In VAG, we have explored the possibility to make an intuitive yet simple user interface for playing a particular virtual (synthetic) instrument, the electric guitar. In addition to the synthetic instrument and related audio effects (amplifier distortion and loudspeaker cabinet simulation) we have explored three different controllers for player interface: data gloves in a virtual room environment, special hand-held controller sticks, and webcam-based camera tracking of player’s hands.

The first case (data glove control) is for flexible experimentation of possible control features, while the two others are intended for maximally simplified guitar playing, designed for wide audience visiting a science center exhibition. The two most important parameters needed in all cases are the pitch control (corresponding to fretting position) and the string plucking action. The pitch-related information is taken by measuring the distance of the two hands, which was found easier to use than the distance of left hand to a reference such as the players body. The string plucking action is most easily captured by the downward stroke of the right hand.

Here we present an overview of our strategy and implementation of the virtual air guitar, paying most attention to the user interface aspects from the
players point of view. We then shortly present our results and expectations. A more detailed description of the VAG can be found in Karjalainen et al. (2006).

A.3.2 Strategy-implementation

The VAG system is composed of (1) a guitar synthesiser with sound effects and audio reproduction, (2) a user interface consisting of handhold sensors, and (3) software to map user interface signals to expressive VAG playing.

The virtual instrument is a simulation of an electric guitar tuned to sound like the Fender Stratocaster, and it is realised using the Extended Karplus-Strong modelling technique described in Karjalainen et al. (1998). The preamplifier stages of a tube amplifier and loudspeaker cabinet were simulated digitally; the VAG also includes a delay unit and a reverberation unit integrated with the guitar synthesiser.

For the user interface, we have experimented with three approaches.

1. Data gloves and 3-D position-tracking in a cave-like virtual room (Fig. A.3). The finger flexure parameters can be mapped to sliding, string damping, selection of string or chords, or controlling different audio effects.

2. Small handhold control devices (Fig. A.4). The right-hand stick sends high-frequency pulses via a tiny loudspeaker, and the left-hand stick receives them via an electret microphone. The pitch is extracted from the propagation delay of the pulses, whereas the pluck is captured by an acceleration sensor microchip inside the right-hand stick.

3. Hand tracking by video image analysis (Fig. A.5). The user wears orange gloves, which correspond to blobs in video frames. A gesture extractor is informed by their location. A pluck is detected when the right hand passes through the imaginary guitar centerline. The guitar moves with the player, and the user’s grip calibrates its size. Vibrato, slide, and string muting are extractable gestures. The user can choose among two playing modes: rhythm guitar and solo guitar, i.e. she can strum four different

---

4Web documents on the project, including videos of playing the VAG, are available at: http://airguitar.tml.hut.fi and http://www.acoustics.hut.fi/demos/VAG
power chords, or freely play a guitar solo on a pentatonic minor scale, with additional techniques such as fret sliding and vibrato.

The VAG’s interface can use complex mapping rules and procedures from gesture to sound model. This mapping layer is called *musical intelligence* in the VAG system. The user’s gestures are first interpreted into a meta language that describes guitar playing techniques on an abstract, musical level. For example, moving the right hand over the imaginary guitar strings in a strumming motion is interpreted as a pluck event. The musical intelligence thus contains both the rules by which gestures are converted into these musical events, and the implementations of the events for a certain sound model.

### A.3.3 Results-expectations

The VAG is an entertainment device that is in line with the way of playing the air guitar - with showmanship, intensity and fun. The main effort has been to study how the hand positions and movements can be mapped to control typical playing of the electric guitar. This means a highly simplified user interface, which sets strict limits to what can be played by such a virtual guitar, certainly not satisfactory for a professional guitarist. It is, however, an interesting case of studying what can be done to demonstrate the basic features of playing a particular instrument with a given style, or to make inexpensive “toy” instruments for fun. This can then be augmented by extra functions in more complex controllers, for example using finger movements, foot pedals, etc. In such cases the challenge is to find expressive and intuitive forms of controls that may be useful also for professional musicians. Another way, to get rich sound from minimalistic controllers, is to use complex rule-based mappings from simple control signals to more advanced control signals for playing a virtual instrument.

Both the webcam and the control stick versions were placed on display at the Heureka Science Center in Finland in 2005 March. The webcam version became the most popular attraction of the music-related exhibition, being played over 60,000 times during the one year of the exhibition. It has also attracted international media attention, including numerous television shows,
radio programs, popular magazine articles, and online articles. Currently, a commercial Virtual Air Guitar game is in development.\(^5\)

The VAG is a good example of a virtual instrument that requires special controllers and playing strategies, different from keyboard-oriented control. The simple versions described above are intended for toy-like applications, such as games, or instructional devices to characterise the most essential features of plucked string instrument playing.

Among future challenges, there are studies on more expressive control interfaces, which could be useful also for professional musicians. With respect to real guitar playing, there is much more freedom to apply different gesture parameters for virtual instrument control.

\(^5\)http://www.virtualairguitar.com
Figure A.3: Soulful playing of a VAG in a virtual room using data gloves.
Figure A.4: Control sticks for playing the VAG. Right-hand stick (on the left) includes an acceleration sensor as well as a small loudspeaker to send the distance measurement pulse. Left-hand stick (on the right) receives the pulse by an electret microphone.
Figure A.5: Camera tracking of hand positions (orange gloves), as seen by the computer.
A.4 The reacTable

The reacTable is a multi-user electronic music instrument with a tabletop tangible user interface ([Jordà et al., 2007] [Kaltenbrunner et al., 2004]). Several simultaneous performers share complete control over the instrument by moving and rotating physical objects on a luminous table surface. By moving and relating these objects, representing components of a classic modular synthesiser, users can create complex and dynamic sonic topologies, with generators, filters and modulators, in a kind of tangible modular synthesiser or graspable flow-controlled programming language.

In recent years there has been a proliferation of tabletop tangible musical interfaces. This trend started with the millennium with projects such as the Audiopad ([Patten et al., 2002]), Jam-o-drum ([Blaine and Perkis, 2000]) or SmallFish 6, and nowadays so many “musical tables” are being produced that it becomes difficult to keep track of every new proposal 7. Is this just a coincidence or the result of a tabletop vogue? While arguably, not all the currently existing prototypes may present the same level of achievement or coherence, we believe that there are important reasons, perhaps often more intuited than stated, that turn musical instruments and tabletop tangible interfaces, into promising and exciting fields of crossed multidisciplinary research and experimentation.

A.4.1 Concept: Multithreaded musical instrument and shared control

New musical controllers and laptop performance Music controllers or new interfaces for musical expression (NIME) are experimenting an increasing attention from researchers and electronic luthiers. In parallel to this research bloom, the laptop is progressively reaching the point of feeling as much at home on stage as a saxophone or an electric guitar. However, the contemporary musical scene does not clearly reflect this potential convergence, and most

---

6SmallFish: [http://hosting.zkm.de/wmuench/small_fish](http://hosting.zkm.de/wmuench/small_fish)
7Kaltenbrunner, M. [http://www.iua.upf.es/mtg/reacTable/?related](http://www.iua.upf.es/mtg/reacTable/?related)
laptop performers seem hesitant to switch towards the use of new hardware controllers, as if laptop performance and the exploration of post-digital sound spaces was a dialog conducted with mice, sliders, buttons and the metaphors of business computing. The reasons for this apparent rejection can be probably found in the inherent nature of computer-based music performance. In traditional instrumental playing, every nuance, every small control variation or modulation (e.g. a vibrato or a tremolo) has to be addressed physically by the performer. In digital instruments nevertheless, the performer no longer needs to control directly all these aspects of the production of sound, being able instead to direct and supervise the computer processes which control these details. These new type of instruments often shift the center of the performer’s attention from the lower-level details to the higher-level processes that produce these details. The musician performs control strategies instead of performing data, and the instrument leans towards more intricate responses to performer stimuli, tending to surpass the note-to-note and the “one gesture-one acoustic event” playing paradigms present in all traditional instruments, thus allowing musicians to work at different musical levels and forcing them to take higher level and more compositional decisions on-the-fly (Jorda, 2005).

However, most of the music controllers currently being developed do not pursue this multithreaded and shared control approach, prolonging the traditional instrument paradigm instead. Many new musical interfaces still tend to conceive new musical instruments highly inspired by traditional ones, most often designed to be “worn” and played all the time, and offering continuous, synchronous and precise control over a few dimensions. An intimate, sensitive and not necessarily highly dimensional interface of this kind (i.e. more like a violin bow, a mouthpiece or a joystick, than like a piano) will be ideally suited for direct microcontrol (i.e. sound, timbre, articulation). However, for macrostructural, indirect or higher level control, a non-wearable interface distributed in space and allowing intermittent access (i.e. more like a piano or a drum), and in which control can be easily and quickly transferred to and recovered from the machine, should be undeniably preferred (Jorda, 2005).
TUIs: Making graphical user interfaces graspable  Even if Graphical User Interface (GUI) conception and design may be central in most HCI related areas, not many new music instruments profit from the display capabilities of digital computers, whereas in the musical performance model we are discussing, in which performers tend to frequently delegate and shift control to the instrument, all affordable ways for monitoring ongoing processes and activities are especially welcome. Visual feedback potentially constitutes a significant asset for allowing this type of instruments to dynamically communicate the states and the behaviors of their musical processes (Jorda, 2002; Jorda and Wüst, 2003); it is the screen and not the mouse what laptop performers do not want to miss, and it is in this context where tabletop tangible interfaces may have a lot to bring.

Tangible User Interfaces (TUIs) combine control and representation within a physical artefact (Ullmer and Ishii, 2001). In table based tangible interfaces, digital information becomes graspable with the direct manipulation of simple objects which are available on a table surface. Combining augmented reality techniques that allow the tracking of control objects on the table surface, with visualisation techniques that convert the table into a flat screening surface, a system with these characteristics favors multi-parametric and shared control, interaction and exploration and even multi-user collaboration. Moreover, the seamless integration of visual feedback and physical control, which eliminates the indirection component present in a conventional screen + pointer system, allows a more natural, intuitive and rich interaction. With these considerations, it may not be a coincidence that in recent years, an increasing variety of tabletop tangible musical controllers or instruments, such as Audiopad (Patten et al., 2002) or the reacTable, is being developed.

A.4.2 The reacTable: Strategy and implementation

The reacTable consists of a translucent luminous round table - a surface with no head position or leading voice and with no privileged points-of-view or points-of-control - in which physical artefacts or pucks can be moved and rotated. Each puck represents a synthesiser module with a dedicated function
for the generation, modification or control of sound. Six functional groups exist each one associated with a different puck shape (audio generators, audio filters, controllers, control filters, mixers and global functions); different pucks within a same group show distinct symbols on their surfaces) (see Figures A.6 and A.7).

Modular synthesis goes back to the first sound synthesizers, in the digital and especially in the analogue domains, with Robert Moog’s or Donald Buchla’s Voltage controlled synthesizers (Chadabe, 1975). It also constitutes the essence of many visual programming environments for sound and music such as Max or Pd (Puckette, 1996). The reacTable outdoes these models by implementing what we call dynamic patching (Kaltenbrunner et al., 2004). Connections and disconnections between modules are not explicitly indicated by the performer, but automatically managed by means of a simple set of rules according to the objects’ types, and their affinities and proximities with their neighbors. By moving pucks and bringing them into proximity with each other, performers on the reacTable construct and play the instrument at the same time. Since the move of any object around the table surface can alter existing connections, extremely variable synthesizer topologies can be attained resulting in a highly dynamic environment.

Figure A.6: Four hands at the reacTable
Additionally, all reacTable objects can be spun as rotary knobs, which allow controlling one of their internal parameters, and a second parameter is controlled by dragging the finger around the objects’ perimeter as shown in Figure A.8. Although the exact effects vary from one type of object to the other, rotation tends to be related with frequency or speed and finger dragging with amplitude. Finger interaction also allows to temporarily cutting (i.e. mute) audio connections, activating/deactivating discrete steps in objects such as step-sequencers or tonalisers (see D and G in Figure A.7) or even drawing envelopes or waveforms which will be “absorbed” by the nearest “wave or envelope-compatible” object.

In order to control a system of such complexity, with potentially dozens of continuous and discrete parameters, visual feedback becomes an essential component of the reacTable’s interface. It was decided to avoid any type of alphanumerical and symbolic information and to banish any decorative display. All shapes, forms, lines or animations drawn by the visual synthesiser bring relevant information, and all the relevant information of the system (i.e. both the instantaneous results of the performers actions as well as the current
state of all the system controlled processes) is permanently displayed. The lines that show the connections between the objects convey the real resulting waveforms that are being produced or modified at each node (in the case of audio connections) or the density and intensity of the values they transport (in the case of control event-based connections). Low frequency oscillators, metronomes and other objects that vibrate at visible rates (i.e. below a few Hz) are animated with their precise heartbeat periods. All object states and internal parameters are also permanently monitored, such as in the white 180 degree circular fuel gauges that surround any object indicating their rotational values, in the dots that show the position of the second parameter slider (see all the objects in Figure A.7 except D and G, and the two objects in Figure A.8), or on the finger-activated discrete steps that surround objects such as the step-sequencers or the tonaliser (see respectively D and G in Figure A.7).

A.4.3 Results and expectations: Interacting with the reacTable

The reacTable was designed with both casual users and professional performers in mind, and it seeks to combine an immediate and intuitive access in a relaxed and immersive way, with the flexibility and the power of digital sound design algorithms and endless improvement possibilities and mastery. Since
its first presentation in the summer of 2005, the reacTable has undergone a very active life outside of the laboratory. It has been exhibited in dozens of festivals, conferences or shows around the world, and has been played by several thousands users of all ages and different backgrounds (musicians, computer music and computer graphics experts; electronic music, digital art or computer game aficionados; teenagers, families with kids, etc.). The feedback has been very positive, often even passionate, showing that the reacTable can be very much enjoyable even without being fully understood.

In parallel to these public installations, many concerts have taken place during recent years, in festivals, clubs, art galleries or discotheques, which have turned the reacTable into an already mature musical instrument. Recently, it has even reached the rock stadiums as a regular instrument in Björk’s last world tour, thus becoming one of the very few new digital instruments to have successfully passed the research prototype state.

The ideal combination of these two very different test-beds fulfills our initial goal, which was to build a musical instrument conceived for casual users at home or in installations, as well as for professionals in concert, and these two complementary test fronts keep bringing very relevant information into the continual reacTable design process refinement.
A.5 The interactive book

A.5.1 Concept

A book is a very well known object which everyone has used at least once in life. It plays an important role in children education: most of us learned colors, names of animals and numbers just “reading” or, better, interacting with some nice, colored pull-the-tab and lift-the-flap books. In the last decades children’s books have been modified in order to use new interaction channels, inserting technology inside this old medium or using the book metaphor to develop new interfaces. It is quite clear that technology did not change too much for the book in thousand years: the history of books has seen new printing and composition techniques but the users are still basically dealing with the same artefact. Thus, the book as an object guarantees a high level of functionality.

Current commercial interactive books for children are very often similar to conventional colored stories with the addition of some pre-recorded sounds which can be triggered by the reader. The limitations of these books are evident: the sounds available are limited in number and diversity and they are played using a discrete control (typically a button). This means that sounds are irritating rather than being a stimulus to interact with the toy-book or allowing for learning by interaction.

Pull-the-tab and lift-the-flap books play a central role in the education and entertainment of most children all over the world. Most of these books are inherently cross-cultural and highly relevant in diverse social contexts. For instance, Lucy Cousins, the acclaimed creator of Maisy (Pina in Italy), has currently more than twelve million books in print in many different languages. Through these books, small children learn to name objects and characters, they understand the relations between objects, and develop a sense of causality by direct manipulation ([Hutchins et al., 1986; Schneiderman, 2002]) and feedback. The importance of sound as a powerful medium has been largely recognised and there are books on the market that reproduce prerecorded sounds upon pushing certain buttons or touching certain areas. However, such triggered sounds are extremely unnatural, repetitive, and annoying. The key for a
successful exploitation of sounds in books is to have models that respond continuously to continuous action, just in the same way as the children do when manipulating rattles or other physical sounding objects. In other words, books have to become an embodied interface (Dourish, 2001) in all respects, including sound.

A.5.2 Strategy-implementation

In recent years, the European project “The Sounding Object” was entirely devoted to the design, development, and evaluation of sound models based on a cartoon description of physical phenomena. In these models the salient features of sounding objects are represented by variables whose interpretation is straightforward because based on physical properties. As a result, the models can be easily embedded into artefacts and their variables coupled with sensors without the need of complex mapping strategies. Pop-up and lift-the-flap books for children were indicated as ideal applications for sounding objects (Rocchesso et al., 2003), as interaction with these books is direct, physical, and essentially continuous. Even though a few interactive plates were prototyped and demonstrated, in-depth exploitation of continuous interactive sounds in children books remains to be done.

Everyday sounds can be very useful because of the familiar control metaphor: no explanation nor learning is necessary (Brewster, 2002). Moreover, it is clear that the continuous audio feedback affects the quality of the interaction and that the user makes continuous use of the information provided by sounds to adopt a more precise behavior. For example, the continuously varying sound of a car engine tells us when we have to shift gears. In this perspective sound is the key for paradigmatic shifts in consumer products. In the same way as spatial audio has become the characterizing ingredient

8 http://www.soundobject.org
9 The conception and realisation of an early prototype of a sound-augmented book were carried on by the second author as part of the Sounding Object project. Later on, students Damiano Battaglia (Univ. of Verona) and Josep Villadomat Arro (Univ. Pompeu Fabra, Barcelona) realised the sketches that are described in this paper as part of graduation projects, under the guidance of the authors.
for home theaters (as opposed to traditional TV-sets), continuous interactive sounds will become the skeleton of electronically-augmented children books of the future. The book-prototype is designed as a set of scenarios where narration develops through sonic narratives, and where exploration is stimulated through continuous interaction and auditory feedback. Through the development of the book, the class of models of sounding objects has been deeply used and verified. The physical models of impacts and friction have been used to synthesise a variety of sounds: the steps of a walking character, the noise of a fly, the engine of a motor bike, and the sound of an inflatable ball.

The integration and combination of the sound models available from the Sounding Object project in an engaging tale has been studied. The first step was to create demonstration examples of interaction using different kinds of sensors and algorithms. During this phase the most effective interactions (i.e. easier to learn and most natural) have been chosen, and several different scenarios were prepared with the goal of integrating them in a common story. The scenarios use embedded sensors, which are connected to a central control unit. Data is sent to the main computer using UDP messages through a local network from sensors and the sound part is synthesised using custom designed pd patches. These pd patches implement a set of physical models of everyday sounds such as friction, impacts, bubbles, etc. and the data coming from sensors are used to control the sound object model in real time. In the following paragraph an investigation scenario will be described.

The steps scenario shows a rural landscape with a road; an embedded slider allows the user to move the main character along the road, and all movement data are sent to the computer, where the velocity of the character is calculated and a sound of footsteps is synthesised in real-time. The timing, distance, and force of the sound of each step is modified as a function of the control velocity. Fig. A.9 shows a preliminary sketch, while fig. A.10 shows the final prototype with the embodied sensor.

Our investigation shows that in the near future lift-the-flap books for children will be augmented by sounds that respond continuously and consistently to control gestures. The sample scenario shown in the previous paragraph demonstrates the effectiveness of sound as an engaging form of feedback and the possibility of embedding real-time physics-based models of everyday sounds in small inexpensive stand-alone systems. A relevant part of future work will concentrate on real-world tests with children that will enhance the playability/usability of prototype books. Another aspect which will be further developed is the embedding and the sophistication of the technologies used.
Figure A.10: Interaction through slider: the footsteps scenario prototype
Bibliography


