

# Gestural Control of Sound Synthesis

MARCELO M. WANDERLEY, MEMBER, IEEE, AND PHILIPPE DEPALLE, MEMBER, IEEE

## Invited Paper

*This paper provides a review of gestural control of sound synthesis in the context of the design and evaluation of digital musical instruments. It discusses research in various areas related to this field and equally focuses on four main topics: analysis of music performers' gestures, gestural capture technologies, real-time sound synthesis methods, and strategies for mapping gesture variables to sound synthesis input parameters. Finally, this approach is illustrated by presenting an application of this research to the control of digital audio effects.*

**Keywords**—Audio systems, music, signal synthesis, user interfaces.

## I. INTRODUCTION

The evolution of computer music has brought to light a plethora of sound synthesis methods available in general and inexpensive computer platforms, allowing a large community direct access to real-time computer-generated sound. Both signal and physical models have reached a point where they can be used in concert situations, although much research continues to be carried on in the subject, constantly bringing innovative solutions and developments [1]–[3].

On the other hand, input device technology that captures different human movements can also be viewed as in an advanced stage [4], [5], considering both noncontact movements and manipulation.<sup>1</sup> Specifically regarding manipulation, tactile and force feedback devices for both

nonmusical<sup>2</sup> and musical contexts have already been proposed [6].<sup>3</sup>

We are then in a stage where such devices and sound synthesis methods can be combined to create new computer-based musical instruments, or *digital musical instruments* (DMI), producing gesturally controlled real time computer-generated sound. The ultimate goal is to design new DMIs capable of obtaining similar levels of control subtlety as those available in acoustic instruments, but at the same time extrapolating the capabilities of existing instruments.

In short, we need to devise ways to interact with computers in a musical context, i.e., to control multiple continuous parameters that allow the generation of sound in real time. This topic amounts to a branch of knowledge known as HCI. Various questions need to be addressed, such as the following.

- Which are the specific constraints that exist in the musical context with respect to general HCI?
- Given the various contexts related to interaction in sound generation systems, what are the similarities and differences within these contexts (interactive installations, DMI manipulation, dance–music interfaces)?
- How to design systems for these various musical contexts? Which system characteristics are common and which are context specific?

## A. HCI and Music

More specifically, gestural control of computer-generated sound can be seen as a highly specialized branch of HCI involving the *simultaneous control of multiple parameters, timing, rhythm, and user training* [7]. Hunt and Kirk consider various attributes that are characteristic of real-time multi-parametric control systems [8].

- There is no fixed ordering to the human–computer dialogue.

<sup>2</sup>For a survey on haptic devices, check the Haptics Community Web page at <http://haptic.mech.mwu.edu/>.

<sup>3</sup>Even so, many users still use the traditional pianolike keyboard as the main input device for musical interaction. This situation seems to be equivalent to the ubiquitous role played by the mouse and keyboard in traditional human–computer interaction (HCI).

Manuscript received April 26, 2003; revised October 29, 2003. The work of M. M. Wanderley was supported by grants from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brazil, McGill University (Individual New Researcher Grant), and the Fonds Québécois de Recherche sur la Nature et les Technologies (FQRNT—Professeur-Chercheur Stratégique).

The authors were with the Analysis/Synthesis Team, Institut de Recherche et Coopération Acoustique Musique, Paris 75004, France. They are now with the Sound Processing and Control Laboratory, Faculty of Music—McGill University, Montreal, QB H3A 1E3, Canada (e-mail: marcelo.wanderley@mcgill.ca; depalle@music.mcgill.ca).

Digital Object Identifier 10.1109/JPROC.2004.825882

<sup>1</sup>With the exception of extreme conditions, such as three-dimensional whole-body acquisition in large spaces.

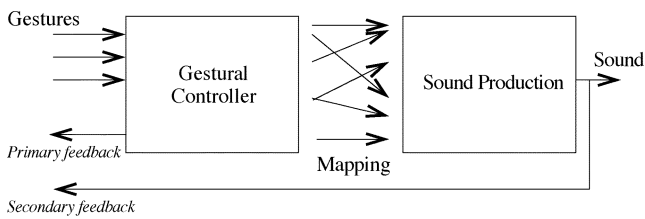


Fig. 1. A symbolic representation of a DMI.

- There is no single permitted set of options (e.g., choices from a menu) but rather a series of continuous controls.
- There is an instant response to the user's movements.
- The control mechanism is a physical and multiparametric device which must be learned by the user until the actions become automatic.
- Further practice develops increased control intimacy and, thus, competence of operation.
- The human operator, once familiar with the system, is free to perform other cognitive activities while operating the system (like talking while driving a car).

### B. Interaction in a Musical Context

In order to take into account the specifics of musical interaction, one needs to consider the various existing contexts—sometimes called metaphors for musical control [9]—where gestural control can be applied to computer music.

These different interaction contexts are the result of the evolution of electronic technology allowing, for instance, a same input device to be used in different situations: to generate sounds (notes) or to control the temporal evolution of a set of prerecorded notes. These two contexts traditionally correspond to two separate roles in music, those of the performer and the conductor, respectively. Technology has blurred the difference between traditional roles and allowed novel metaphors derived from other areas, such as HCI [9]–[12].

In this paper, we will focus on *instrument manipulation*, or performer–instrument interaction in the context of real-time sound synthesis control.

The approach suggested here consists on dividing the subject of gestural control of sound synthesis in four parts [13]:

- definition and typologies of gesture;
- gesture acquisition and input device design;
- synthesis algorithms;
- mapping of gestural variables to synthesis variables.

The goal is to analyze all four parts, which are equally important to the design of new DMIs.

## II. CONTROL OF DIGITAL MUSICAL INSTRUMENTS

In this paper, the term *digital musical instrument* [14] is used to represent an instrument that includes a separate gestural interface (or gestural controller unit) from a sound generation unit. Both units are independent and related by mapping strategies [15]–[18]. This is shown in Fig. 1.

The term *gestural controller*<sup>4</sup> can be defined here as the input part of the DMI, where physical interaction with the player takes place. Conversely, the sound generation unit can be seen as the synthesis algorithm and its input parameters. The mapping layer refers to the liaison strategies between the outputs of the gestural controller and the input controls of the synthesis algorithm.

This separation is most of the time impossible in the case of acoustic instruments, where the gestural interface is also part of the sound generation unit. If one considers, for instance, a clarinet, the reed, keys, holes, etc., are at the same time both the gestural interface (where the performer interacts with the instrument) and the elements responsible for sound generation. The idea of a DMI is analogous to “splitting” the clarinet in a way where one could separate these two functions (gestural interface and sound generator) and use them independently.

Clearly, this separation of the DMI into two independent units is potentially capable of extrapolating the functionalities of a conventional musical instrument, the latter tied to physical constraints. On the other hand, basic interaction characteristics of existing instruments may be lost and/or difficult to reproduce, such as tactile/force feedback.

### A. Gesture and Feedback

In order to devise strategies concerning the design of new DMIs for gestural control of sound synthesis, it is essential to analyze the characteristics of actions produced by expert instrumentalists during performance. These actions are commonly referred to as *gestures* in the musical domain. In order to avoid discussing all nuances of the meaning of gesture, let us initially consider *performer gestures* as *performer actions produced by the instrumentalist during a performance*, meaning both actions such as prehension and manipulation, and noncontact movements. A detailed discussion is presented in [19].

The importance of the study of gestures in DMI design can be justified by the need to better understand physical actions and reactions that take place during expert performance. Furthermore, gesture information can also be considered as a form of signal, i.e., they can be processed, transformed, and stored using *gesture editors* [20]. Gestures can also be synthesized using various models of movement [15], [21] or using rules in a similar way to speech synthesis [22].

In fact, instrumentalists simultaneously execute various types of gestures during performance. Some of them are necessary for the generation of sound [23], while others are not [24]–[26], although the latter are also present in most highly skilled instrumentalists' performances [27].

One can approach the study of gestures in a musical context by either analyzing the possible functions of a gesture during performance [20] or by analyzing the physical properties of the gestures taking place [28]. By identifying gestural characteristics—*functional*, in a specific context, or *physio-*

<sup>4</sup>The term *gestural controller* is used here meaning *input device* for musical control.

logical—one can ultimately gain insight into the design of gestural acquisition systems [29].

Regarding both approaches, one fundamental aspect is the existing feedback available to the performer, be it visual, auditory, or tactile-kinesthetic. Feedback can also be considered, depending on its characteristics, as follows.

- *Primary/secondary*, where primary feedback encompasses visual, auditory (clarinet key noise, for instance), and tactile-kinesthetic feedback,<sup>5</sup> and secondary feedback relates to the sound produced by the instrument [32].
- *Passive/active*, where passive feedback relates to feedback provided through physical characteristics of the system (a switch noise, for instance) and active feedback is the one produced by the system in response to a certain user action (sound produced by the instrument) [5].

### B. Gestural Acquisition

Once the characteristics of gestures are known, it is essential to devise an acquisition system that will *capture* these characteristics. In the case of performer–instrument interaction, this acquisition may be performed in three ways.

- *Direct acquisition*, where one or various sensors are used to monitor performer’s actions. The signals from these sensors present isolated basic physical features of a gesture: pressure, linear or angular displacement, speed, or acceleration. Each physical variable of the gesture to be captured will normally require a different sensor.
- *Indirect acquisition*, where gestures are extracted from the structural properties of the sound produced by the instrument [33]–[38]. Signal processing techniques can then be used in order to derive performer’s actions by the analysis of the fundamental frequency of the sound, its spectral envelope, its temporal envelope, etc.
- *Physiological signal acquisition*, the analysis of physiological signals, such as EMG [39], [40]. Commercial systems have been developed based on the analysis of muscle tension and used in musical contexts [41]–[45].

1) *Direct Acquisition*: Direct acquisition is performed by the use of different sensors to capture performer actions. Depending on the type of sensors and on the combination of different technologies in various systems, different movements may be tracked.

According to B. Bongers [5]:

Sensors are the sense organs of a machine. Sensors convert physical energy (from the outside world) into electricity (into the machine world). There are sensors available for all known physical quantities, including the ones humans use and often with a greater range. For instance, ultrasound frequencies (typically 40 kHz used for motion tracking) or light waves in the infrared frequency range.

<sup>5</sup>Tactile-kinesthetic, or *tactual* [30], feedback is composed of the tactile and proprioceptive senses [31].

Direct acquisition has the advantage of simplicity when compared to indirect acquisition, i.e., one can obtain independent streams of data representing individual control parameters. On the other hand, due to the independence of the variables captured, direct acquisition techniques may underestimate the interdependency of the various variables obtained.

2) *Sensor Characteristics and Musical Applications*: Some authors consider that most important sensor characteristics are *sensitivity*, *stability*, and *repeatability* [46]. Other important characteristic relates to the *linearity* and *selectivity* of the sensor’s output, its sensitivity to ambient conditions, etc. A more complete analysis proposes six descriptive parameters applicable to sensors [47]: *accuracy*, *error*, *precision*, *resolution*, *span*, and *range*.

In general instrumentation circuits, sensors typically need to be both precise and accurate, and present a reasonable resolution. In the musical domain, it is often stressed that the choice of a transducer technology matching a specific musical characteristic relates to human performance and perception: for instance, mapping of the output of a sensor that is precise but not accurate to a variable controlling loudness may be satisfactory, but if it is used to control pitch, its inaccuracy will probably be more noticeable.

In music, the use of commercially available sensors developed for other uses is the rule. Only a few researchers have proposed sensors specifically designed for musical use, for instance [48]. Various texts describe different sensors and transducer technologies for general and musical applications, such as [47], [5], and [49], respectively.

3) *Analog-to-Digital Conversion*: For the case of gesture acquisition with the use of various sensors, the signals obtained at the sensors outputs are usually available in an analog format, basically in the form of voltage or current signals. In order to be able to use these signals as computer inputs, they need to be sampled and converted in a suitable format, usually Musical Instrument Digital Interface (MIDI) [50] or more advanced protocols such as Open Sound Control (OSC) [51].

Various analog-to-MIDI converters have been proposed and are widely available commercially. The first examples had already been developed in the 1980s [52], [53].

Concerning the various discussions on the advantages and drawbacks of the MIDI protocol and its use [54], strictly speaking, nothing forces someone to use MIDI or prevents the use of faster or different protocols. As already pointed out, the limiting factor regarding speed and resolution is basically the specifications of the MIDI protocol, not the electronics involved in the design.

It is interesting to notice that many existing systems have used communication protocols other than MIDI in order to avoid speed and resolution limitations. One such system is the *transducteur gestuel rétroactif* (TGR) from ACROE [55]. Other papers have proposed different options to implement gesture acquisition interfaces, such as using hardware initially designed for audio processing [56], [57].

4) *Indirect Acquisition*: As opposed to direct acquisition, *indirect acquisition* provides information about

performer actions from the evolution of structural properties of the sound being produced by an instrument. In this case, the only sensor is a microphone, i.e., a sensor measuring pressure or gradient of pressure.

Due to the complexity of the information available in the instrument's sound captured by a microphone, various real-time signal processing techniques are used in order to distinguish the effect of a performer's action from other factors, e.g., the influence of the acoustical properties of the room or the intrinsic properties of the instrument.

Generically, one could identify basic sound parameters to be extracted in real-time [13] as follows.

- *Short-time energy*, related to the dynamic profile of the signal, indicates the dynamic level of the sound but also possible differences of the instrument position with respect to the microphone.
- *Fundamental frequency*, related to the sound's melodic profile, gives information about fingering, for instance.
- *Spectral envelope*, representing the distribution of sound partial amplitudes, may give information about the resonating body of the instrument.
- *Amplitudes, frequencies, and phases of sound partials* that can alone provide much of the information obtained by the previous parameters.

Several works on indirect acquisition systems have already been presented. They include both hybrid systems (using also sensors), such as the *hypercello* [58] and pure indirect systems, such as the analysis of clarinet performances [35], [36] and guitar [37], [38], [59].

5) *Sampling Gestural Signals*: Obviously, in order to perform the analysis of the above or other parameters during direct or indirect acquisition, it is important to consider the correct sampling of the signal. According to the Nyquist theorem, this frequency needs to be at least twice as high as the maximum frequency of the signal to be sampled.

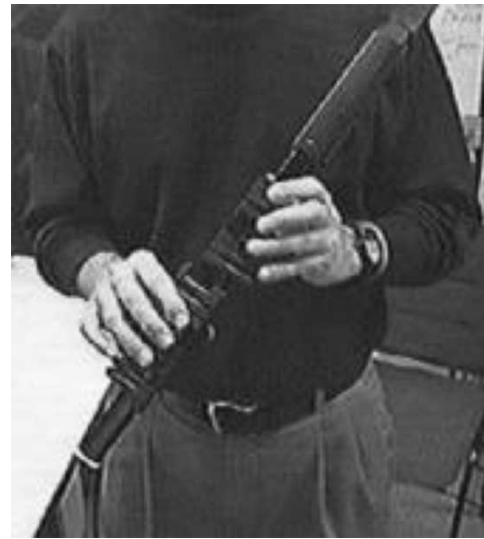
Although one could reasonably consider that frequencies of performer actions can be limited to a few hertz, fast actions can potentially present higher frequencies. A typical sampling frequency for gestural acquisition is 200 Hz [60]. Some systems may use higher values, up to 1 kHz [55], and other researchers considered the ideal sampling frequency to be around 4 kHz [56], [61].

### C. Gestural Controllers

Once one or several sensors are assembled as part of a unique device, this device is called a *gestural controller*.<sup>6</sup>

As cited above, the gestural controller is the part of the DMI where physical interaction takes place. *Physical interaction* here means the actions of the performer, be they body movements, empty-handed gestures, or object manipulation, and the perception by the performer of the instrument's status and response by means of tactile-kinesthetic, visual, and auditory senses.

<sup>6</sup>Called *input device* in HCI.



**Fig. 2.** J. B. Rován holding a wx7, an *instrument-like* (saxophone) controller by Yamaha.



**Fig. 3.** S. Goto and the SuperPolm, an *instrument-inspired* controller (violin).

Due to the large range of human actions to be captured by the controller<sup>7</sup> and depending on the interaction context where it will be used [12], its design may vary from case to case. Existing controller designs can be classified as follows [5], [12], [13].

- *Instrument-like controllers* (see Fig. 2), where the input device design tends to reproduce each feature of an existing (acoustic) instrument in detail. Many examples can be cited, such as electronic keyboards, guitars, saxophones, marimbas, and so on.
- *Instrument-inspired controllers* that although largely inspired by the existing instrument's design, are conceived for another use [62]. Fig. 3 presents one example of such controller, the *SuperPolm* violin developed by S. Goto, A. Terrier, and P. Pierrot [63], [64], where the input device is loosely based on a violin shape, but is used as a general device to control granular synthesis.
- *Extended instruments* are instruments augmented by the addition of extra sensors [58], [65]. Commercial augmented instruments included the Yamaha Disklavier, used, for instance, in pieces by J.-C. Risset

<sup>7</sup>According to A. Mulder, a *virtual musical instrument is ideally capable of capturing any gesture from the universe of all possible human movements and use them to produce any audible sound* [16].



Fig. 4. M. Battier manipulating the Pacom, an *alternate* controller designed by M. Starkier and P. Prevot.



Fig. 5. J.-P. Viollet using the WACOM graphic tablet.

[66], [67]. Other examples include the flute [68]–[70] and the trumpet [71]–[73], but any existing acoustic instrument may be extended to different degrees by the addition of sensors.

- *Alternate controllers* (see, e.g., Fig. 4), whose design does not follow that of an established instrument. Some examples include *the Hands* [52], graphic drawing tablets [74] (cf. Fig. 5), etc. For instance, an unorthodox gestural controller using the shape of the oral cavity has been proposed in [75].

For instrument-like controllers, although mostly representing a simplified (first-order) model of the acoustic instrument, many of the gestural skills developed by the performer on the acoustic counterparts can be readily applied to the controller. Conversely, for a nonexpert performer, these controllers present roughly the same constraints as those of an acoustic instrument,<sup>8</sup> i.e., technical difficulties inherent to the former will have to be overcome by the nonexpert performer.

Alternate controllers, on the other hand, allow the use of other gesture vocabularies than those of traditional instrument manipulation, thus being in principle less demanding for nonexpert performers. Even so, performers still have to develop specific skills for mastering these new gestural vocabularies [52].

These controllers can furthermore be classified into different categories.

<sup>8</sup>This fact can be modified by the use of different mapping strategies, as shown in [17].

- *Touch, expanded range, or immersive* controllers [76], depending on the amount of physical contact required from the performer. Mulder also [76] separates immersive controllers into *internal, external, and symbolic* controllers according to the possibilities of visualization of the control surface. In a different approach, Piringer [77] classifies immersive controllers into *partial or completely immersive* controllers.
- *Individual or collaborative* controllers [78], depending on whether the instrument is performed by one or multiple performers at one time.
- *Metaphorical or ad hoc* controllers, and so on.

### III. AN ANALYSIS OF EXISTING INPUT DEVICES

A reasonable number of input devices have been proposed to perform real-time control of music [4], [79], most of them resulting from composer's/player's idiosyncratic approaches to personal artistic needs. These interfaces, although often revolutionary in concept, have mostly remained specific to the needs of their inventors. As an illustration, four examples of gestural controllers are shown in Figs. 2–5.

The advantages and drawbacks of each controller type depends mostly on the user goals and background, but unfortunately systematic means of evaluating gestural controllers are not available [12].

A systematic approach needs to be held according to engineering principles: it is important to propose means to quantitatively evaluate existing designs<sup>9</sup> in order to identify their strong and weak points and eventually come up with guidelines for the design of new input devices [48].

Another unsolved discussion relates to the interest of ergonomically designed, easy-to-use, and intuitive interfaces for musical control. Several authors consider that new devices designed according to ergonomical and cognitive principles can become general tools for music performance [76], [81]–[83]. Other authors claim that effort demanding and hard-to-play instruments are the only ones that provide expressive possibilities to a performer, linking the concepts of *effort* and *expression* [84]–[86].

#### A. Design Rationale: Engineering Versus Idiosyncratic Approaches

The use of pure engineering/ergonomical approaches can be challenged by the comparison with the evolution of input device design in HCI. In fact, researcher W. Buxton [86] provocatively considered HCI and ergonomics as *failed sciences*. He argues that although a substantial volume of literature on input device evaluation/design in these two areas has already been proposed, current available devices have benefited little from all this knowledge.

The problem with both points of view—*engineering* versus *idiosyncratic*—seems to be related to their application context. While one can always question the engineering approach by stressing the role of creativity against scientific design [87], the proposition of scientific methodologies is a key factor for *the evaluation* of existing gestural controllers.

<sup>9</sup>A similar situation occurs in others areas, such as haptic devices [80].

Conversely, engineering methodologies, shall not prevent the use of creativity in design, although this can be a side effect of structured design rationales. But without a common basis for evaluation, the differentiation between input devices and simple gadgets turns out to be hazardous.

As stated before, the design of a new input device for musical performance is generally directed toward the fulfillment of specific and sometimes idiosyncratic musical goals, but is *always* based on an engineering corpus of knowledge. This technical background allows the choice of transducer technologies and circuit designs that implement the interface needed to perform the initial musical goals.<sup>10</sup>

Therefore, although the final development goals are musical and, consequently, any criticism of these goals turns into a question related to aesthetic preferences, their design is based on engineering principles that can, *and need*, to be evaluated and compared [48]. This evaluation is essential, for instance, in the selection of existing input devices for performing different tasks [74], but it can also be useful in the identification of promising new opportunities for the design of novel input devices [88].

### B. Gestural Controller Design and Evaluation

It may also be useful to propose guidelines for the design of new input devices based on knowledge from related fields, such as experimental psychology, physiology, and HCI [7].

Taking the example of the research in HCI, many studies have been carried out on the design and evaluation of input devices for general (nonexpert) interaction. The most important goal in these studies is the improvement of accuracy and/or time response for a certain task.

Also, standard methodologies for tests have been proposed and generally consist of pointing and/or dragging tasks, where the size and distance between target squares are used as tests parameters, following the relationship, known as Fitts' law [89].

In 1994, Vertegaal and collaborators presented a methodology, derived from standard HCI tests, that addressed the comparison of input devices in a timbre navigation task [32], [90]. Although innovative in this field, the methodology used consisted of a pure selection (pointing and acquisition) task, i.e., the context of the test was a navigation in a four-parameter timbral space [91], not a standard musical context in the sense of instrumental performance.

In a subsequent paper, Vertegaal *et al.* [81], [83] proposed an attempt to systematically match a hypothetical musical function (*dynamic*—absolute or relative—or *static*) to a specific sensor technology and to the feedback available with this technology. This means that certain sensor technologies would outperform others for a specific musical function. The interest of this work is that it allows a designer to select a sensor technology based on the proposed relationships, thus

<sup>10</sup>A description of several input device designs is proposed in [5], where Bongers review his work at STEIM, Amsterdam, The Netherlands; the Institute of Sonology, Den Haag, The Netherlands; and the Royal Academy of Arts, Amsterdam, The Netherlands. Other good reviews of various controllers has been presented by J. Paradiso in [4] and by Y. Nagashima in his home page at <http://nagasm.org>.

reducing the need for idiosyncratic solutions. An exploratory evaluation of this methodology was presented in [92].

Another attempt to address the evaluation of well-known HCI methodologies and their possible adaptation to the musical domain was presented in [12]. Although one cannot expect to use methodologies from other fields directly into the musical domain, at least the analysis of similar developments in better established fields may help find directions suitable for the case of computer music.

## IV. DEVELOPMENTS IN SOUND SYNTHESIS METHODS

On the other side of current trends regarding DMIs, the design of sound generation units benefits from various developments in sound synthesis. Sound synthesis is now a well-established topic which has been intensively studied for almost 50 years. It can be considered in its most general meaning as the study of sound representations that leads to appropriate implementation of sound generation devices. This is based on the conception and development of models for acoustic signals or instruments.

Several models for sound synthesis and processing have been proposed these last 40 years. They can be classified in two categories: physical models and signal models. The principle of physical models is to analytically describe the mechanical and acoustic behavior of musical instruments (or more generally of sound generators) in order to simulate them [2]. This results in an integro-differential equation system. Sound synthesis consists then in solving this system with specific initial and boundary conditions by using finite elements approaches or a modular decomposition followed by a simulation procedure. Physical models are especially useful for a realistic simulation of a given acoustic instrument and several models of musical instruments exist and are currently commercially available. The modular approach is much preferred in a musical context for its flexibility, despite some simplification and/or approximation. The most popular systems so far have been wave-guides [93], spring-mass systems [29], and modal synthesis [94]. As regards physical modeling, sound representation is idiosyncratic to the structure of the instrument that generates the sound. Disadvantages of physical models include the lack of analysis methods and the difficulties of developing interpolation or extrapolation mechanisms that preserve a high degree of perceptual coherence and remain controllable by performers gesture controls, with the exception of instruments that share a similar structure (violin, and viola, trumpet and trombone, etc.) or in very anecdotal situations [73].

On the other hand, the signal model is based on a phenomenological approach, which uses abstract mathematical structures for the coding of spectral and/or temporal properties of sounds [22]. There is no explicit reference to the instrument which produces the sound. Among the more popular signal models are the phase vocoder, the additive synthesis, the source-filter model, and FM synthesis. These models essentially code spectral characteristics as auditive perception is mainly related to the spectral content of sound

signals. In practice, signal modeling consists in an abstract structure, like a signal processing structure, designed to store information related to perceptual effects (such as formant parameters) within the structure's parameters (such as the coefficients of a second-order filter). In practice, most of the signal modeling synthesis techniques are based on the representation of a spectral information and code in more or less condensed structure frequencies, amplitudes, and phase of sound signals (e.g., additive synthesis, phase vocoder, source-filter synthesis) [1]. Signal models, with the notable exception of FM, present the advantage of having well-developed analysis tools that allow for the extraction of parameters corresponding to a given sound. Therefore, the morphing of parameters from different sounds can lead to continuous transformations between different instruments. Although not necessarily reproducing the full behavior of the original instrument, the flexibility allowed by signal models may be interesting for the prototyping of control strategies, since the mapping is left to the instrument designer.

#### A. Simulation and Extrapolation

There are usually two approaches in synthesizing sounds: we call them here *simulation* and *extrapolation*.

Simulation consists in an accurate reproduction of existing sound signals. This often implies the use of an analysis/resynthesis procedure. In the context of real-time application, simulation would be typically controlled by instrument-like controllers. Apart of the already mentioned economical aspect, the other interest of simulation is a scientific one: the idea is to validate sound synthesis models by showing that they are able to represent the largest class of sounds in order to increase our knowledge of musical instrument, playing conditions, and more specifically on the sound characteristics' evolution according to various instrumental gestures.

Extrapolation consists in using synthesis models, outside the normal range of parameters, or conditions other than those usually met in natural instrumental situations. The goal here is to create new sounds. In an extrapolation context, signal models have an edge on physical models, which results from the continuous behavior of their working mode: a continuous variation of parameters always provide the user with a sound signal (even though this sound signal might not be musically interesting). On the other hand, a physical model might abruptly stop generating sounds or might change its working mode. In the case of a clarinet, for example, a too high blowing pressure might push the reed against the mouthpiece and stop sound generation, instead of continuously increasing sound level.

## V. MAPPING OF GESTURAL VARIABLES TO SYNTHESIS INPUTS

Once gesture variables are available either from independent sensors or as a result of signal analysis techniques in the case of indirect acquisition, one then needs to relate these output variables to the available synthesis input variables.

Depending on the sound synthesis method to be used, the number and characteristics of these input variables may vary. For signal model methods, one may have: 1) amplitudes, frequencies, and phases of sinusoidal sound partials for additive synthesis; 2) an excitation frequency plus each formant's center frequency, bandwidth, amplitude, and skew for formant synthesis; 3) carrier and modulation coefficients ( $c : m$  ratio) for FM synthesis, etc. It is clear that the relationship between the gestural variables and the synthesis inputs available is far from obvious.

For the case of physical models, the available input variables are usually the physical parameters of an instrument, such as blow pressure, bow velocity, etc. In this context, the mapping of gestures to the synthesis inputs seems to be more evident, since the relation of these inputs to the synthesis algorithm are directly mapped by the multiple dependencies based on the physics of the particular instrument.

#### A. Systematic Study of Mapping

The systematic study of mapping in computer music is an area that has only recently received substantial attention.<sup>11</sup> Until now, few works have been proposed that analyze the influence of mapping on DMI performance or suggested ways to define mappings to relate controller variables to synthesis inputs. Examples include [8], [17], and [96]–[113].

#### B. Mapping for General Musical Performance

Although simple *one-to-one* or *direct* mappings are by far the most commonly used, other *mapping strategies* can be used. For instance, through the use of several mapping strategies, it has been shown that for the same gestural controller and synthesis algorithm, the choice of mapping strategy became the determinant factor concerning the expressivity of the instrument [17].

The definition of mapping strategies using instrument-like and perhaps instrument-inspired controllers can benefit from our knowledge of the physics of acoustic instrument. But in the case of an alternate controllers, the possible mapping strategies to be applied are far from obvious, since no model of the mappings strategies to be used is available. Even so, it can be demonstrated that the choice of mappings influences user performance for the manipulation of general input devices in a musical context.

An interesting work by A. Hunt and collaborators [8], [104], [114] presented a study on the influence over time of the choice of mapping strategy on subject performance in real-time musical control tasks. User performance was measured over a period of several weeks and showed that complex mapping strategies used with the multiparametric instrument allowed better performance than simpler mappings for complex tasks and also that performance with complex mapping strategies improved over time.

<sup>11</sup>A recent issue of the journal *Organized Sound*, on mapping strategies for real-time computer music, guest edited by the first author, presents an overview of the state-of-the-art concerning developments on mapping in computer music [95].

### C. A Model of Mapping as Multiple Layers

Mapping can be implemented as a single layer [96], [100] between controller outputs and synthesis inputs. In this case, a change of either the gestural controller or synthesis algorithm would imply the definition of a different mapping.

One way to overcome this situation is the definition of mapping as two (or possibly more) *independent layers*: a *mapping of control variables to intermediate parameters* and a *mapping of intermediate parameters to synthesis variables* [98], [102], [103].

This means that the use of different gestural controllers would necessitate the use of different mappings in the first layer, but the second layer, between intermediate parameters and synthesis parameters, would remain unchanged. Conversely, changing the synthesis method involves the adaptation of the second layer, considering that the same abstract parameters can be used, but does not interfere with the first layer, therefore being transparent to the performer.

The definition of those intermediate parameters or an intermediate abstract parameter layer can be based on perceptual variables such as timbre, loudness, and pitch, but can be based on other perceptual characteristics of sounds, [115], [116] or have no relationship to perception, being then arbitrarily chosen by the composer or performer [103].

Conversely, more than two layers of mapping can be defined. For instance, designs using three layers have been proposed in [111] and [117]. Multiple layer designs potentially facilitate mappings between layers, where one-to-one relationships may become more meaningful. This means that it is a tradeoff between simplifying the mappings and making the layer structure more complex.

## VI. AN EXAMPLE OF APPLICATION: GESTURALLY CONTROLLED DIGITAL AUDIO EFFECTS

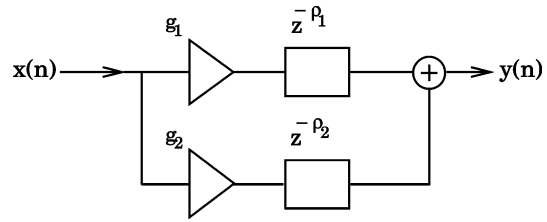
We have chosen to present an innovative example of application of the above discussion where we take into account performer expressive movements—a kind of gesture which is usually not considered in general input devices for musical expression—and use these as a variable to improve the naturalness of sounds generated by synthesis.

The implementation of the effects of these gestures results in a simple extension of the synthesizer, basically a sound processing module similar to existing digital audio effects units.

An extension of this research is the development of naturally controlled digital audio effects for musical applications.

### A. Ancillary Movements of Performers and Their Acoustical Effects

As discussed in Section II-A, musicians constantly perform movements not directly related to sound production [118]. For the case of a woodwind instrument performer, these movements can consist of postural adjustments, upward/downward movements of the instrument, and circular patterns, among others [119]. It has been shown that these movements are consistent and, therefore, can be considered as part of the performance [120].



**Fig. 6.** Symbolic representation of the two-path acoustical propagation system.

Furthermore, it is interesting to note that performer movements—for the case of woodwind instruments<sup>12</sup>—may influence the sound produced and recorded under close microphone conditions.

Considering the case of a clarinet, for standard recording conditions [121], movements of the instrument will cause significant amplitude modulations (and even cancellations) of sinusoidal sound partials due to the displacement of the sound source (the open holes) with respect to a microphone, depending on its position. The authors have presented a detailed report of the analysis of several clarinet samples recorded in various acoustically controlled conditions, including an anechoic chamber [26]. This was done in order to investigate and evaluate the effects of ancillary performer gestures on the timbre of the instrument. We have shown that the influence of ancillary gestures mostly results from the reflection off the floor, as compared to variations in the mouthpiece, directivity effects, or speed of performer movements.

The floor reflection, which is, in this case, the first reflection of the room reverberation, interferes with the direct sound of the clarinet. This effect can be represented by a simple model consisting of two delay lines each one including a variable delay  $\rho$  (expressed in samples), and a variable gain  $g$ . The first (characterized by  $\rho_1, g_1$ ) represents the propagation of the direct sound, while the second one ( $\rho_2, g_2$ , with  $\rho_1 < \rho_2$ ) represents the propagation of the sound that reflects off the floor. For a fixed position of the clarinet, the transfer function  $H(z)$  of this model (cf. Fig. 6) can be written as

$$H(z) = g_1 z^{-\rho_1} + g_2 z^{-\rho_2} \quad (1)$$

that factorizes into

$$H(z) = g_1 z^{-\rho_1} H_c(z) \quad (2)$$

where  $H_c$  is a comb filter

$$H_c(z) = 1 + \alpha z^{-D} \quad (3)$$

with  $\alpha = (g_2/g_1)$  and  $D = \rho_2 - \rho_1$ .

The magnitude of the frequency response of such a system exhibits an interleaved structure of evenly spaced soft peaks at frequencies  $f_p(k) = k(1/D)$  ( $k$  being an integer), and sharp dips at frequencies  $f_d(k) = (k - (1/2))(1/D)$ . As

<sup>12</sup>And any other instrument for which sound sources move with performer gestures, such as strings, brass, etc.



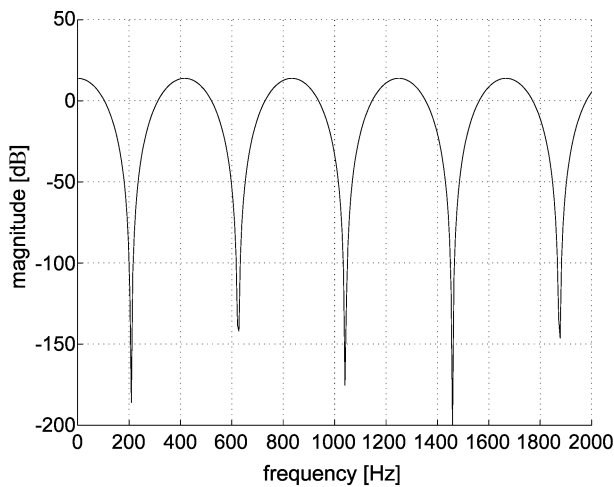


Fig. 7. Frequency response of the two-path system for a delay difference of 2.4 ms and  $\alpha = 1$ .

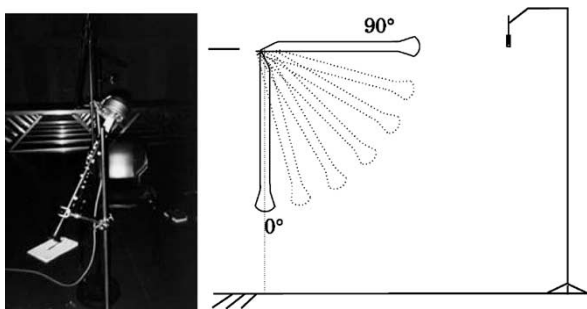


Fig. 8. Room response measurements with excitation provided by a loudspeaker connected to a clarinet tube.

an example,<sup>13</sup> Fig. 7 shows a plot of the frequency response where zeroes are distributed at odd harmonic locations of  $f_d(1) = 208$  Hz, while poles lie on harmonic locations of  $f_p(1) = 416$  Hz.

There are several factors that influence the specific values of the two gains  $g$  and the two delays  $\rho$ . These have been discussed in detail in [122].

### B. Gain and Delay Measurements

In order to validate the above model, gain and delay parameters were determined using experimental measurements through the estimation of the concert hall's impulse responses.

To achieve this goal, we have used standard techniques for impulse response estimations. For these measurements, the sound was generated by using a loudspeaker connected to a clarinet tube, all side holes closed. The temporal response of the global system (clarinet, microphone, and auditorium) was recorded for several clarinet orientation angles, as shown in Fig. 8.

Table 1 shows the values for the gain and the delay of the direct sound and of the first reflection, measured by a

<sup>13</sup>Assuming that the amplitudes of the direct sound and of the first reflection are equal (i.e.  $\alpha = 1$ ), and that the delay difference  $D = \rho_2 - \rho_1 = 106$  samples (i.e. 2.4 ms at a sampling rate of 44 100 Hz, which represents a distance difference of 0.792 m).

**Table 1**  
Measurements of Gain and Time Delay—Direct Sound and the First Reflection

Angle [degrees]	$g_1$ [dB]	$g_2$ [dB]	$\rho_1$ [ms]	$\rho_2$ [ms]
0	-49	-47	10.7	12.7
15	-48	-45	10.2	12.3
30	-47	-44	9.7	12.1
45	-46	-43.5	9.2	12.1
60	-43.5	-43.5	8.7	12.2
75	-42	-44.5	8.3	12.5
90	-39	-47	8.0	13.0

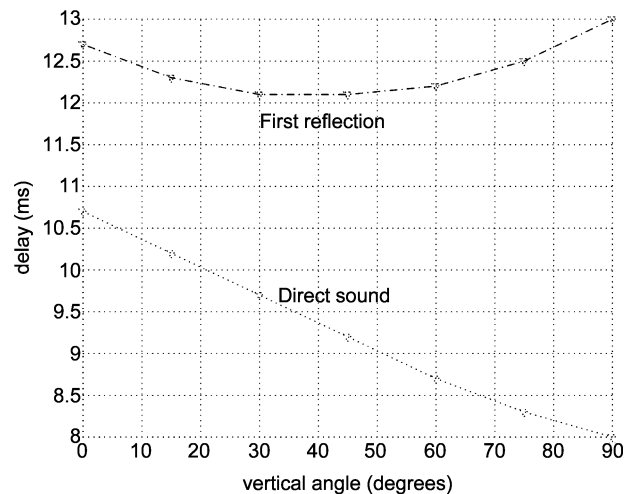


Fig. 9. Delay of direct sound ( $\rho_1$ ) and first reflection ( $\rho_2$ ) measured in the auditorium excited by the experimental device shown in Fig. 8.

microphone at 2 m away from the mouthpiece, at a height of 2 m.

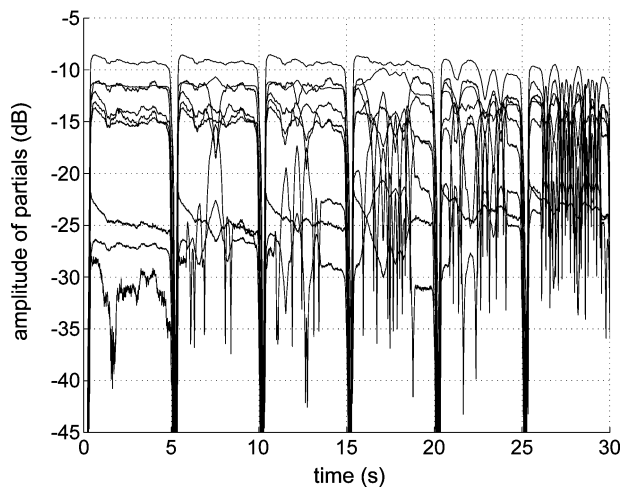
Fig. 9 shows the delays obtained for the direct sound ( $\rho_1$ ) and for the first reflection ( $\rho_2$ ), under the conditions described above.

When moving the clarinet tube from the horizontal to the vertical position, the delay difference evolves from 2 to 5 ms, which generates a harmonic structure of zeroes in the spectrum, the fundamental frequency of which decreases from 250 down to 100 Hz. For sound whose partial frequencies coincide with the positions of the zeroes of the system, a strong attenuation will be noticed. The same will also happen for the odd multiples of these frequencies.

Considering that the samples' recording conditions throughout this research comply with the standard clarinet recording procedures suggested in the literature (cf. [121]), and also that a clarinet player will most likely produce ancillary gestures during a performance (cf. [119], [120]), it is reasonable to expect that, in these circumstances, modulations are an integral part of the recorded sound.

### C. Real-Time Simulation

A real-time implementation of the model presented in Fig. 6 has been performed in jMax, the Institut de Recherche et Coordination Acoustique Musique's (IRCAM's) real-time synthesis and audio processing environment.



**Fig. 10.** Evolution of partials amplitudes for simulated motions applied to an original anechoic room sample ([0–5] s)—*D3 ff* standard performance. Arbitrary movements with increasing amplitudes were performed at ([5–10], [10–15], [15–20], [20–25], and [25–30] s).

The sound input  $x[n]$  to the model shown in Fig. 6 is a 5-s musical excerpt recorded in an anechoic chamber. We then simulate five different 5-s angular movements with a slider that controls the orientation angle (cf. Fig. 10). This angle is used for table lookup of gain and delay values for the direct sound and first reflection, as shown in Table 1.

This results in a timbre modulation that sounds similar to a flanging effect. This effect, which is often used in recording studios, consists of adding to a signal a slightly delayed copy of itself. This constitutes a comb filter structure that is very similar to the two-path acoustical propagation system presented in Section VI-A.

By changing the delay, one makes the dips sweep over the spectrum of the input signal, causing a very recognizable sound effect. Commercial flangers control the delay variations through the use of a low-frequency oscillator (LFO) waveform and present typical delay values evolving between 1 and 10 ms [123]. As these periodic variations may be perceived to be repetitive, some authors have proposed improvements by adding random variations to the LFO waveform [124].

Considering the structural analogy presented above, it seems that a further improvement in the control of flanger effects is to modify its delay and gain (or depth) parameters by performer gestures that naturally occur during instrumental performances. These gestures imply variations that are neither too repetitive nor random, and are tightly related to musical events being performed.

The amplitude modulation effect on sound partials can also appear in other circumstances, such as a beating effect in instruments having several slightly detuned strings associated to the same note, as in the case of a piano. Conversely, a similar modulation effect can be produced by a fixed comb filter applied to a time-varying spectrum, as in the case of sound coloration in auditoriums. The lack of such modulation in electronic sounds, in electric instruments, or in “sanitized” sounds recorded in absorbing rooms, is likely

to explain the success of flanger devices in modern studio technology.

## VII. CONCLUSION

This paper presented a critical review of various topics related to real-time, gesturally controlled computer-generated sound.

Starting from the specificities of the interaction between a human and a computer in various musical contexts, we have described in detail one interaction metaphor for music/sound control. We have focused on the specific case of real-time gestural control of sound synthesis, by presenting and discussing the various constituent parts of a DMI.

We have then claimed that a balanced analysis of these constituent parts is an essential step toward the design of new instruments, although current developments many times tend to focus either on the design of new gestural controllers or on the proposition of different synthesis algorithms.

We have finally illustrated these aspects on a case study that methodologically analyzed the effect of ancillary gestures on the sound produced. We furthermore proposed a model of this effect and presented simulation results that indicate the necessity of a more lively control of digital audio effects and computer-generated sound.

## ACKNOWLEDGMENT

The authors would like to thank X. Rodet, M. Battier, C. Cadoz, S. Dubnov, A. Hunt, F. Isart, R. Kirk, K. Ng, N. Orio, J. B. Rován, N. Schnell, M.-H. Serra, H. Vinet, and J.-P. Viollet for various discussions, ideas, and comments.

## REFERENCES

- [1] G. Borin, G. DePoli, and A. Sarti, “Musical signal synthesis,” in *Musical Signal Processing*, C. Roads, S. T. Pope, A. Piccialli, and G. DePoli, Eds. Lisse, The Netherlands: Swets & Zeitlinger, 1997, pp. 5–30.
- [2] V. Välimäki and T. Takala, “Virtual musical instruments—natural sound using physical models,” *Organized Sound*, vol. 1, no. 2, pp. 75–86, 1996.
- [3] G. Peeters and X. Rodet, “Non-stationary analysis/synthesis using spectrum peak shape distortion, phase and reassignment,” presented at the Int. Congr. Signal Processing Applications Technology (ICSPAT), Orlando, FL, 1999.
- [4] J. Paradiso, “New ways to play: electronic music interfaces,” *IEEE Spectrum*, vol. 34, pp. 18–30, Dec. 1997.
- [5] B. Bongers, “Physical interfaces in the electronic arts. Interaction theory and interfacing techniques for real-time performance,” in *Trends in Gestural Control of Music*, M. Wanderley and M. Battier, Eds. Paris, France: Institut de Recherche et Coordination Acoustique Musique—Centre Pompidou, 2000, pp. 41–70.
- [6] C. Cadoz, L. Lisowski, and J. L. Florens, “A modular feedback keyboard design,” *Comput. Music J.*, vol. 14, no. 2, pp. 47–56, 1990.
- [7] N. Orio, N. Schnell, and M. M. Wanderley. (2001) Input devices for musical expression: borrowing tools from HCL. *Proc. Workshop New Interfaces for Musical Expression—ACM CHI01* [Online]. Available: <http://www.nime.org/>
- [8] A. Hunt and R. Kirk, “Mapping strategies for musical performance—trends in gestural control of music,” in *Trends in Gestural Control of Music*, M. Wanderley and M. Battier, Eds. Paris, France: Institut de Recherche et Coordination Acoustique Musique—Centre Pompidou, 2000, pp. 231–258.
- [9] D. Wessel and M. Wright, “Problems and prospects for intimate musical control of computers,” *Computer Music J.*, vol. 26, no. 3, pp. 11–22, 2002.

- [10] A. Camurri, "Interactive dance/music systems," in *Proc. 1995 Int. Computer Music Conf.*, 1995, pp. 245–252.
- [11] M. M. Wanderley, N. Orio, and N. Schnell, "Toward an analysis of interaction in sound generating systems," presented at the Int. Symp. Electronic Arts (ISEA2000), Paris, France.
- [12] M. M. Wanderley and N. Orio, "Evaluation of input devices for musical expression: borrowing tools from HCI," *Computer Music J.*, vol. 26, no. 3, pp. 62–76, 2002.
- [13] M. M. Wanderley and P. Depalle, "Contrôle Gestuel de la Synthèse Sonore," in *Interfaces Homme-Machine et Création Musicale*, H. Vinet and F. Delalande, Eds. Paris, France: Hermès Science, 1999, pp. 145–163.
- [14] M. Battier, "Les Musiques électroacoustiques et l'environnement informatique," Ph.D. dissertation, University of Paris X, Nanterre, France, 1981.
- [15] S. Gibet, "Codage, représentation et traitement du geste instrumental," Ph.D. dissertation, Institut National Polytechnique de Grenoble, Grenoble, France, 1987.
- [16] A. Mulder, "Virtual musical instruments: accessing the sound synthesis universe as a performer," in *Proc. 1st Brazilian Symp. Computer Music*, 1994, pp. 243–250.
- [17] J. Rován, M. M. Wanderley, S. Dubnov, and P. Depalle, "Instrumental gestural mapping strategies as expressivity determinants in computer music performance," presented at the Kansei Technology of Emotion Workshop, Genova, Italy, 1997.
- [18] S. Sapir, "Interactive digital audio environments: gesture as a musical parameter," in *Proc. COST-G6 Conf. Digital Audio Effects (DAFx'00)*, 2000, pp. 25–30.
- [19] C. Cadoz and M. M. Wanderley, "Gesture-music," in *Trends in Gestural Control of Music*, M. Wanderley and M. Battier, Eds. Paris, France: Institut de Recherche et Coordination Acoustique Musique—Centre Pompidou, 2000, pp. 71–93.
- [20] C. Ramstein, "Analyse, représentation et traitement du geste instrumental," Ph.D. dissertation, Institut National Polytechnique de Grenoble, December 1991.
- [21] S. Gibet, P. F. Marteau, and F. Julliard, "Models with biological relevance to control antropomorphic limbs: a survey," in *Gesture and Sign Language in Human-Computer Interaction*, I. Wachsmuth and T. Sowa, Eds. Heidelberg: Springer Verlag, 2002, pp. 105–119.
- [22] P. Depalle, "Analyse, modelization et synthèse des sons fondées sur le modèle source/filtre," Ph.D. dissertation, Université du Maine, Le Mans, 1991.
- [23] C. Cadoz, "Instrumental gesture and musical composition," in *Proc. 1988 Int. Computer Music Conf.*, pp. 1–12.
- [24] F. Delalande, "La gestique de Gould," in *Glenn Gould Pluriel*: Louise Courteau, éditrice, inc., 1988, pp. 85–111.
- [25] J.-W. Davidson, "Visual perception of performance manner in the movements of solo musicians," *Psychology of Music*, vol. 21, pp. 103–113, 1993.
- [26] M. M. Wanderley, P. Depalle, and O. Warusfel, "Improving instrumental sound synthesis by modeling the effects of performer gesture," in *Proc. 1999 Int. Computer Music Conf.*, pp. 418–421.
- [27] C. Bahn, T. Hahn, and D. Trueman, "Physicality and feedback: a focus on the body in the performance of electronic music," in *Proc. 2001 Int. Computer Music Conf.*, pp. 44–51.
- [28] I. Choi, "Cognitive engineering of gestural primitives for multi-modal interaction in a virtual environment," in *Proc. IEEE Int. Conf. Systems, Man and Cybernetics (SMC'98)*, pp. 1101–1106.
- [29] C. Cadoz, A. Luciani, and J. L. Florens, "Responsive input devices and sound synthesis by simulation of instrumental mechanisms: the cordis system," *Computer Music J.*, vol. 8, no. 3, pp. 60–73, 1984.
- [30] B. Bongers, "Tactual display of sound properties in electronic musical instruments," *Displays*, vol. 18, pp. 129–133, 1998.
- [31] J. P. Roll, "Sensibilités cutanées et musculaires," in *Traité de Psychologie Expérimentale*, M. Richele, J. Requin, and M. Robert, Eds. Paris, France: Presses Universitaires de France, 1994, pp. 483–542.
- [32] R. Vertegaal and B. Eaglestone, "Comparison of input devices in an ISEE direct timbre manipulation task," *Interacting With Computers*, vol. 8, no. 1, pp. 13–30, 1996.
- [33] N. Bailey, A. Purvis, I. W. Bowler, and P. D. Manning, "Applications of the phase vocoder in the control of real-time electronic musical instruments," *Interface*, vol. 22, pp. 259–275, 1993.
- [34] C. Lippe, "A composition for clarinet and real-time signal processing: using max on the IRCAM signal processing workstation," in *Proc. 10th Italian Colloq. Computer Music*, 1993, pp. 428–432.
- [35] M. Puckette and C. Lippe, "Getting the acoustic parameters from a live performance," in *Proc. 3rd Int. Conf. Music Perception and Cognition*, 1994, pp. 328–333.
- [36] E. B. Egozy, "Deriving Musical Control Features From a Real-Time Timbre Analysis of the Clarinet," Master's thesis, Massachusetts Institute of Technology, 1995.
- [37] N. Orio, "The timbre space of the classical guitar and its relationship with the plucking techniques," in *Proc. 1999 Int. Computer Music Conf.*, pp. 391–394.
- [38] C. Traube, P. Depalle, and M. M. Wanderley, "Indirect acquisition of instrumental gesture based on signal, physical and perceptual information," in *Proc. 2003 Int. Conf. New Interfaces for Musical Expression (NIME03)*, pp. 42–47.
- [39] H. Lusted and B. Knapp, "Controlling computer with neural signals," *Sci. Amer.*, vol. 275, no. 4, pp. 82–87, 1996.
- [40] R. Picard, *Affective Computing*. Cambridge, MA: MIT Press, 1997.
- [41] A. Tanaka, "Musical technical issues in using interactive instrument technology with applications to the BioMuse," in *Proc. 1993 Int. Computer Music Conf.*, pp. 124–126.
- [42] B. Bongers, "An interview with sensorband," *Comput. Music J.*, vol. 22, no. 1, pp. 13–24, 1998.
- [43] A. Tanaka, "Musical performance practice on sensor-based instruments," in *Trends in Gestural Control of Music*, M. Wanderley and M. Battier, Eds. Paris, France: Institut de Recherche et Coordination Acoustique Musique—Centre Pompidou, 2000, pp. 389–405.
- [44] T. Marrin-Nakra, "Searching for meaning in gestural data: interpretive feature extraction and signal processing for affective and expressive content," in *Trends in Gestural Control of Music*, M. Wanderley and M. Battier, Eds. Paris, France: Institut de Recherche et Coordination Acoustique Musique—Centre Pompidou, 2000, pp. 269–299.
- [45] A. Willier and C. Marque, "Juggling gesture analysis for music control," in *Gesture and Sign Language in Human-Computer Interaction*, I. Wachsmuth and T. Sowa, Eds. Heidelberg, Germany: Springer-Verlag, 2002, pp. 296–326.
- [46] R. L. Smith, "Sensors," in *The Electrical Engineering Handbook*, R. C. Dorf, Ed. Boca Raton, FL: CRC, 1993, pp. 1152–1161.
- [47] P. H. Garrett, *Advanced Instrumentation and Computer I/O Design: Real-Time System Computer Interface Engineering*. Piscataway, NJ: IEEE, 1994.
- [48] N. Gershenfeld and J. Paradiso, "Musical applications of electric field sensing," *Comput. Music J.*, vol. 21, no. 2, pp. 69–89, 1997.
- [49] E. Flety, "3D gesture acquisition using ultrasonic sensors," in *Trends in Gestural Control of Music*, M. Wanderley and M. Battier, Eds. Paris, France: Institut de Recherche et Coordination Acoustique Musique—Centre Pompidou, 2000, pp. 193–207.
- [50] "MIDI musical instrument digital interface specification 1.0," Int. MIDI Assoc., North Hollywood, CA, 1983.
- [51] M. Wright, A. Freed, and A. Momeni, "Open sound control: state of the art 2003," in *Proc. 2003 Int. Conf. New Interfaces for Musical Expression (NIME03)*, pp. 153–160.
- [52] M. Waiswiz, "The hands, a set of remote MIDI-controllers," in *Proc. 1985 Int. Computer Music Conf.*, pp. 313–318.
- [53] M. Starkier and P. Prevot, "Real-time gestural control," in *Proc. 1986 Int. Computer Music Conf.*, pp. 423–426.
- [54] F. R. Moore, "The disfunctions of MIDI," in *Proc. 1987 Int. Computer Music Conf.*, pp. 256–262.
- [55] C. Cadoz and C. Ramstein, "Capture, representation and 'composition' of the instrumental gesture," in *Proc. 1990 Int. Computer Music Conf.*, 1990, pp. 53–56.
- [56] A. Freed and D. Wessel, "Communication of musical gesture using the AES/EBU digital audio standard," in *Proc. 1998 Int. Computer Music Conf.*, pp. 220–223.
- [57] P. Driessen and A. Schloss, "Toward a virtual membrane: new algorithms and technology for analyzing gestural data," presented at the 2001 Int. Computer Music Conf., San Francisco, CA.
- [58] T. Machover, "Hyperinstruments—A progress report 1987–1991," Massachusetts Inst. Technol., Cambridge, 1992.
- [59] C. Traube and P. Depalle, "Deriving the plucking point location along a guitar string from the least-square estimation of a comb filter delay," in *Proc. CCECE 2003: Can. Conf. Electrical and Computer Engineering*, 2003, pp. 2001–2004.
- [60] M. V. Mathews and G. Bennett, "Real-time synthesizer control," Institut de Recherche et Coordination Acoustique Musique, Paris, France, Tech. Rep. 5/78, 1978.

- [61] A. Freed, R. Avizienis, T. Suzuki, and D. Wessel, "Scalable connectivity processor for computer music performance systems," in *Proc. 2000 Int. Computer Music Conf.*, pp. 523–526.
- [62] D. Trueman and P. Cook, "BoSSA: the deconstructed violin reconstructed," in *Proc. 1999 Int. Computer Music Conf.*, pp. 232–239.
- [63] P. Pierrot and A. Terrier, "Le violon midi," Institut de Recherche et Coordination Acoustique Musique, Paris, France, 1997.
- [64] S. Goto, "The aesthetics and technological aspects of virtual musical instruments: The case of the SuperPalm MIDI violin," *Leonardo Music J.*, vol. 9, pp. 115–120, 1999.
- [65] B. Bongers, "The use of active tactile and force feedback in timbre controlling electronic instruments," in *Proc. 1994 Int. Computer Music Conf.*, pp. 171–174.
- [66] J. C. Risset and S. V. Duyne, "Real-time performance interaction with a computer-controlled acoustic piano," *Comput. Music J.*, vol. 20, no. 1, pp. 62–75, 1996.
- [67] J. C. Risset, "Évolution des outils de création sonore," in *Interfaces Homme-Machine et Création Musicale*, H. Vinet and F. Delalande, Eds. Paris, France: Hermès Science, 1999, pp. 17–36.
- [68] D. Pousset, "La Flute-MIDI, l'histoire et quelques applications," in *Mémoire de Maîtrise*. Paris, France: Université Paris—Sorbonne, 1992.
- [69] S. Ystad and T. Voinier, "A virtually real flute," *Comput. Music J.*, vol. 25, no. 2, pp. 13–24, 2001.
- [70] C. Palacio-Quintin, "The hyper-flute," in *Proc. 2003 Int. Conf. New Interfaces for Musical Expression (NIME03)*, 2003, pp. 206–207.
- [71] P. Cook, D. Morril, and J. O. Smith, "A MIDI control and performance system for brass instruments," in *Proc. 1993 Int. Computer Music Conf.*, pp. 130–133.
- [72] J. Impett, "A meta-trumpet(er)," in *Proc. 1994 Int. Computer Music Conf.*, pp. 147–150.
- [73] C. Vergez, "Trompette et trompettiste: un système dynamique non linéaire analysé, modélisé et simulé dans un contexte musical," Ph.D. dissertation, Institut de Recherche et Coordination Acoustique Musique—Université Paris VI, Paris, France, 2000.
- [74] S. Serafin, R. Dudas, M. M. Wanderley, and X. Rodet, "Gestural control of a real-time physical model of a bowed string instrument," in *Proc. 1999 Int. Computer Music Conf.*, pp. 375–378.
- [75] N. Orio, "A gesture interface controlled by the oral cavity," in *Proc. 1997 Int. Computer Music Conf.*, pp. 141–144.
- [76] A. Mulder, "Toward a choice of gestural constraints for instrumental performers," in *Trends in Gestural Control of Music*, M. Wanderley and M. Battier, Eds. Paris, France: Institut de Recherche et Coordination Acoustique Musique—Centre Pompidou, 2000, pp. 315–335.
- [77] J. Piringner, "Elektronische musik und interaktivität: Prinzipien, konzepte, anwendungen," M.S. thesis, Technical University of Vienna, Vienna, Austria, 2001.
- [78] T. Blaine and S. Fels, "Contexts in collaborative musical experiences," in *Proc. 2003 Int. Conf. New Interfaces for Musical Expression (NIME03)*, pp. 129–134.
- [79] C. Roads, *Computer Music Tutorial*. Cambridge, MA: MIT Press, 1996, pp. 617–658.
- [80] V. Hayward and O. R. Astley, "Performance measures for haptic interfaces," in *Proc. Robotics Research: 7th Int. Symp.*, 1996, pp. 195–207.
- [81] R. Vertegaal, T. Ungvary, and M. Kieslinger, "Toward a musician's cockpit: transducer, feedback and musical function," in *Proc. 1996 Int. Computer Music Conf.*, pp. 308–311.
- [82] A. Mulder, "Design of gestural constraints using virtual musical instruments," Ph.D. dissertation, School of Kinesiology, Simon Fraser Univ., Burnaby, BC, Canada, 1998.
- [83] T. Ungvary and R. Vertegaal, "Cognition and physicality in musical cyberinstruments," in *Trends in Gestural Control of Music*, M. Wanderley and M. Battier, Eds. Paris, France: Institut de Recherche et Coordination Acoustique Musique—Centre Pompidou, 2000, pp. 371–386.
- [84] J. Ryan, "Some remarks on musical instrument design at STEIM," *Contemporary Music Rev.*, vol. 6, no. 1, pp. 3–17, 1991.
- [85] —, "Effort and expression," in *Proc. 1992 Int. Computer Music Conf.*, pp. 414–416.
- [86] D. Buchla, W. A. S. Buxton, C. Chafe, T. Machover, R. Moog, M. Mathews, J. C. Risset, L. Sonami, and M. Waisvisz, "Round table," in *Trends in Gestural Control of Music*, M. Wanderley and M. Battier, Eds. Paris, France: Institut de Recherche et Coordination Acoustique Musique—Centre Pompidou, 2000, pp. 415–437.
- [87] P. Cook. (2001) Principles for designing computer music controllers. *Proc. Workshop New Interfaces for Musical Expression—ACM CHI01* [Online]. Available: <http://www.nime.org>
- [88] S. Card, J. Mackinlay, and G. Robertson, "A morphological analysis of the design space of input devices," *ACM Trans. Inform. Syst.*, vol. 9, no. 2, pp. 99–122, 1991.
- [89] I. S. MacKenzie and W. Buxton, "Extending Fitts' law to two dimensional tasks," in *Proc. Conf. Human Factors Computing Systems (CHI'92)*, pp. 219–226.
- [90] R. Vertegaal, "An Evaluation of input devices for timbre space navigation," M.S. thesis, Dept. Comput., Univ. Bradford, Bradford, U.K., 1994.
- [91] R. Vertegaal and E. Bonis, "ISEE: an intuitive sound editing environment," *Comput. Music J.*, vol. 18, no. 2, pp. 12–29, 1994.
- [92] M. M. Wanderley, J. P. Viollet, F. Isart, and X. Rodet, "On the choice of transducer technologies for specific musical functions," in *Proc. 2000 Int. Computer Music Conf.*, pp. 244–247.
- [93] J. O. Smith, "Acoustic modeling using digital waveguides," in *Musical Signal Processing*, C. Roads, S. T. Pope, A. Piccialli, and G. DePoli, Eds. Lisse, The Netherlands: Swets & Zeitlinger, 1997, pp. 221–263.
- [94] J. M. Adrien, "The missing link: Modal synthesis," in *Representations of Musical Signals*, A. P. G. DePoli and C. Roads, Eds. Cambridge, MA: MIT Press, 1991, pp. 269–297.
- [95] M. M. Wanderley, Ed., "Mapping strategies in real-time computer music," in *Organized Sound*. Cambridge, U.K.: Cambridge Univ. Press, 2002, vol. 7.
- [96] I. Bowler, A. Purvis, P. Manning, and N. Bailey, "On mapping N articulation onto M synthesiser-control parameters," in *Proc. 1990 Int. Computer Music Conf.*, pp. 181–184.
- [97] M. Lee and D. Wessel, "Connectionist models for real-time control of synthesis and compositional algorithms," in *Proc. 1992 Int. Computer Music Conf.*, pp. 277–280.
- [98] S. Fels, "Glove talk II: Mapping hand gestures to speech using neural networks," Ph.D. dissertation, Univ. Toronto, Toronto, ON, Canada, 1994.
- [99] T. Winkler, "Making motion musical: gestural mapping strategies for interactive computer music," in *Proc. 1995 Int. Computer Music Conf.*, pp. 261–264.
- [100] I. Choi, R. Bargar, and C. Goudeseune, "A manifold interface for a high dimensional control space," in *Proc. 1995 Int. Computer Music Conf.*, pp. 385–392.
- [101] P. Modler and I. Zannos, "Emotional aspects of gesture recognition by a neural network, using dedicated input devices," in *Proc. KANSEI Technology of Emotion Workshop*, 1997, pp. 79–86.
- [102] A. Mulder, S. Fels, and K. Mase, "Empty-handed gesture analysis in Max/FTS," in *Proc. KANSEI Technology of Emotion Workshop*, 1997, pp. 87–91.
- [103] M. M. Wanderley, N. Schnell, and J. Rován, "Escher—modeling and performing "composed instruments" in real-time," in *Proc. IEEE Int. Conf. Systems, Man and Cybernetics (SMC'98)*, pp. 1040–1044.
- [104] A. Hunt and R. Kirk, "Radical user interfaces for real-time control," presented at the Euromicro 99 Conf., Milan, Italy.
- [105] G. Garnett and C. Goudeseune, "Performance factors in control of high-dimensional spaces," in *Proc. 1999 Int. Computer Music Conf.*, pp. 268–271.
- [106] P. Modler, "Neural networks for mapping gestures to sound synthesis," in *Trends in Gestural Control of Music*, M. Wanderley and M. Battier, Eds. Paris, France: Institut de Recherche et Coordination Acoustique Musique—Centre Pompidou, 2000, pp. 301–313.
- [107] A. Hunt, M. M. Wanderley, and R. Kirk, "Toward a model of mapping strategies for instrumental performance," in *Proc. 2000 Int. Computer Music Conf.*, pp. 209–212.
- [108] I. Choi, "Gestural primitives and the context for computational processing in an interactive performance system," in *Trends in Gestural Control of Music*, M. Wanderley and M. Battier, Eds. Paris, France: Institut de Recherche et Coordination Acoustique Musique—Centre Pompidou, 2000, pp. 139–172.
- [109] M. M. Wanderley, "Performer-instrument interaction: Applications to gestural control of sound synthesis," Ph.D. dissertation, Université Pierre et Marie Curie—Paris VI, Paris, France, 2001.
- [110] C. Goudeseune, "Interpolated mappings for musical instruments," *Organized Sound*, vol. 7, no. 2, pp. 85–96, 2002.
- [111] A. D. Hunt and M. M. Wanderley, "Mapping performance parameters to synthesis engines," *Organized Sound*, vol. 7, no. 2, pp. 97–108, 2002.

- [112] S. Fels, A. Gadd, and A. Mulder, "Mapping transparency through metaphor: toward more expressive musical instruments," *Organized Sound*, vol. 7, no. 2, pp. 109–126, 2002.
- [113] D. Levitin, S. McAdams, and R. L. Adams, "Control parameters for musical instruments: a foundation for new mappings of gesture to sound," *Organized Sound*, vol. 7, no. 2, pp. 171–189, 2002.
- [114] A. Hunt, "Radical user interfaces for real-time musical control," Ph.D. dissertation, University of York, York, U.K., 1999.
- [115] D. Wessel, "Timbre space as a musical control structure," *Comput. Music J.*, vol. 3, no. 2, pp. 45–52, 1979.
- [116] E. Métois, "Musical sound information—Musical gestures and embedding systems," Ph.D. dissertation, Massachusetts Inst. Technol., Cambridge, 1996.
- [117] D. Arfib, J. Couturier, L. Kessous, and V. Verfaillie, "Strategies of mapping between gesture data and synthesis model parameters using perceptual spaces," *Organized Sound*, vol. 7, no. 2, pp. 127–144, 2002.
- [118] A. Gabrielsson, "Music performance," in *The Psychology of Music*, 2nd ed, D. Deutsch, Ed. San Diego, CA: Academic, 1999, pp. 501–602.
- [119] M. M. Wanderley *et al.*, "Non-obvious performer gestures in instrumental music," in *Gesture Based Communication in Human-Computer Interaction*, A. Braffort *et al.*, Eds. Heidelberg, Germany: Springer-Verlag, 1999, pp. 37–48.
- [120] —, "Quantitative analysis of nonobvious performer gestures," in *Gesture and Sign Language in Human-Computer Interaction*, I. Wachsmuth and T. Sowa, Eds. Heidelberg, Germany: Springer-Verlag, 2002, pp. 241–253.
- [121] A. H. Benade, "From instrument to ear in a room: direct or via recording," *J. Audio Eng. Soc.*, vol. 33, no. 4, pp. 218–233, Apr. 1985.
- [122] M. M. Wanderley and P. Depalle, "Gesturally controlled digital audio effects," in *Proc. COST-G6 Conf. Digital Audio Effects (DAFx'01)*, pp. 165–169.
- [123] S. Lehman. (1996) Flanging. *Harmony Central* [Online]. Available: <http://www.harmony-central.com/Effects/Articles/Flanging/>
- [124] P. Fernández-Cid and F. J. Casajús-Quirós, "Enhanced quality and variety for chorus/flange units," in *Proc. 1st COST G-6 Workshop on Digital Audio Effects (DAFx98)*, pp. 35–39.



**Marcelo M. Wanderley** was born in Curitiba, Brazil, in 1965. He holds the B.Eng. degree in electrical engineering from the Universidade Federal do Paraná (UFPR), Curitiba, Brazil, the M.Eng. degree in integrated analog circuit design from the Universidade Federal de Santa Catarina (UFSC), Florianópolis, Brazil, and the Ph.D. degree from the Université Pierre et Marie Curie—Paris VI, Paris, France, on acoustics, signal processing, and computer science applied to music.

From 1996 to 2001, Dr. Wanderley was with the Analysis/Synthesis Team at Institut de Recherche et Coordination Acoustique Musique (IRCAM), Paris, France, where he studied ways of designing new musical instruments based on computer-generated sound. Specifically, he focused on performer-instrument interaction and its applications to gestural control of sound synthesis. He is currently Assistant Professor and Music Technology Area Chair, Faculty of Music, McGill University, Montreal, QB, Canada. He is also the coordinator of the International Computer Music Association/Electronic Music Foundation *Working Group on Interactive Systems and Instrument Design in Music* and part of the advisory group of the EMF Institute. He has published several book chapters and papers and is the coeditor, with Prof. M. Battier, of the electronic publication *Trends in Gestural Control of Music*. He was also the Guest Editor of the Special Issue of *Organized Sound* on mapping strategies for real-time computer music. His main research interests include human-computer interaction, input device design and evaluation, gestural control of sound synthesis, and musical acoustics.

Dr. Wanderley was the Chair of the International Conference on New Interfaces for Musical Expression (NIME03) that was held at McGill University in May 2003.



**Philippe Depalle** was born in Vichy, France, in 1961. He received the B.Sc. degree in physical chemistry from Université de Paris Sud, Orsay, France, in 1982, and the M.S. degree and Ph.D. degree in applied acoustics from Université du Mans, Le Mans, France, in 1984, and 1991, respectively.

From 1985 to 1988, he was an Assistant Professor in Electrical Engineering, with École Supérieure d'Électricité, Metz, France. From 1988 to 1997, he was with the Analysis/Synthesis research team of the Institut de Recherche et Coordination Acoustique Musique (IRCAM), Paris, France. From 1997 to 1999, he was an invited professor at the University of Montreal, Montreal, QB, Canada. He is presently Associate Professor of Music Technology at the Faculty of Music, McGill University, Montreal, QB, Canada. His research interests include sound synthesis, sound processing, sound analysis, simulation of musical instruments, control of sound synthesis, and spectral analysis. His contributions to these topics are various: publications, presentations in international conferences, several pieces of software (including AudioSculpt, Additive, and participation in Diphone) and one patent on additive synthesis techniques. The fundamental component of his work is the systematizing of the "analysis/synthesis" point of view in the conception of computer music tools. He also served as a reviewer for the *Computer Music Journal* and the International Computer Music Conference. He also managed the team that "re-created" the voice of a castrato for the film *Farinelli* in 1994.

Dr. Depalle is a Member of the scientific committee of the International Conference on Digital Audio Effects (DAFX). He was also part of the organizing committee of the 2003 International Conference on New Interfaces for Musical Expression. He has served as a Reviewer for the IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING and for the IEEE Workshop on Applications of Signal Processing on Audio and Acoustics (WASPAA).