

# Non-obvious Performer Gestures in Instrumental Music

Marcelo M. Wanderley

Ircam - Centre Georges Pompidou  
1, Place Igor Stravinsky  
75004 - Paris - France  
mwanderley@acm.org

**Abstract.** This paper deals with the gestural language of instrumentalists playing wind instruments. It discusses the role of non-obvious performer gestures that may nevertheless influence the final sound produced by the acoustic instrument. These gestures have not commonly been considered in sound synthesis, although they are an integral part of the instrumentalist's full gestural language. The structure of this paper will be based on an analysis of these non-obvious gestures followed by some comments on how to best classify them according to existing research on gesture reviewed in the introduction; finally, the influence of these gestures on the sound produced by the instrument will be studied and measurement and simulation results presented.

## 1 Introduction

Sound synthesis tools provide many different and powerful options in inexpensive general-purpose platforms. A number of synthesis paradigms are available, either signal or physical models<sup>1</sup>, many of them in the form of public-domain software. With CD audio quality as a standard and fast synthesis algorithms, the time is right for considering in depth how best to control these environments.

A number of hardware interfaces have been proposed to perform this task, 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. One consequence is that many people still consider the piano-like MIDI<sup>2</sup> keyboard as the main interface for sound synthesis, something equivalent to the present role the mouse and computer keyboard play in traditional human-computer interaction (HCI).

Related to the design of alternative gestural controllers, the role of the gestures one uses to perform this control is a rich although fairly unexplored area

---

<sup>1</sup> Signal models intend to reproduce perceptual characteristics of a sound. Most common examples are: additive synthesis, subtractive synthesis, etc. Physical models, on the other hand, try to simulate, modify and extrapolate the functioning of an instrument by means of a computer algorithm. See [3] or [8] for an introduction.

<sup>2</sup> Musical Instrument Digital Interface.

of research, i.e., the range of possible gestures meaningful in music consists of a huge set of possibilities that are still not completely understood. This paper is proposed as a contribution to this discussion, by reviewing some of the previous related work on the theory of gesture and then focusing on particular gestures of wind instrument performers.

## 2 On Gestures and Music

The term gesture is a rather general idea that may have many (different) meanings in music. As an example, a composer may use the term musical gestures to designate a sequence of events within a space of musical parameters; sometimes it can also have some relation to a form of thinking (a movement of thought)<sup>3</sup>. Whatever the definition, it is easily seen that these ideas are independent of any direct physical meaning in the sense of manipulating an instrument. A performer, on the other hand, may consider performance gesture as the technique used to play an instrument, where it encompasses not only the gestures that actually produce an excitation to/ modification of the instrument, but also the accompanying body movements and postures. Computer musicians or computer music performers using electronic means to produce/control sounds may have a concept of gestures as specific isolated movements related to specific physical variables, such as pressure, velocity, acceleration, etc. that may be captured by sensors and transformed into digital signals input to computers.

Zooming out of the musical universe, one can also identify different ideas expressed by the same term gesture. In order to consider different proposals, let us first consider gestures as divided into two groups:

- Gestures where no physical contact with a device or instrument is involved. These have been called free, semiotic or naked gestures [5].
- Gestures where some kind of physical contact takes place. These have been called ergotic [5], haptic<sup>4</sup> or interactive [5].

Much of the research on empty-handed gestures relate to gestural communication, such as that on co-verbal gestures, or on sign-language recognition. In a musical context, one could cite other gestures (not always empty-handed), including conductors' gestures and on broader terms, dance movements, in the case of dance-music interfaces. These gestures will not be discussed here. Moreover, research on gesture where some form of contact takes place will be reduced to the study of instrumental gesture [5].

Let's now review some specific research that may influence our discussion on performer gestures.

---

<sup>3</sup> I would like to thank Mikhail Malt for his comments on the subject.

<sup>4</sup> One may here differentiate the uses of the term *haptic*. According to Baecker et al. [1] it has its origin in the Greek language and means *having to do with contact*. It is therefore used to represent any type of computer input where contact takes place. The term haptic is nevertheless widely accepted nowadays as designing the research encompassing touch- and/or force-feedback devices.

## 2.1 Natural gestures

J. Cassell uses the term *gesture* to address hand gestures that co-occur with spoken language. Natural gestures are, according to Cassell, types of gestures spontaneously generated by a person telling a story, speaking in public, or holding a conversation [17]. These gestures are idiosyncratic (speaker dependent) and influenced by external factors: situational, educational, and cultural. The author and collaborators have identified different types of natural gestures and also devised ways to recover their temporal structure.

## 2.2 Gestures in HCI

Since the gestural control of computer music is a special type of man-machine communication, it is worth considering definitions of gesture in the light of the research on Human-Computer Interaction. By doing that, one may quickly perceive that the concept of gesture may have a slightly different meaning: *a gesture is a motion of the body that contains information* [12]. Kurtenbach and Hulteen actually do not consider the act of pressing a button (or a key) as a gesture, since *motion of a finger on its way to hitting the key is neither observed nor significant. All that matters is which key was pressed*. C. Hummels and co-workers [11] have noticed the narrowness of this definition and have proposed that the word *convey* be used instead of *contain*, widening this definition to encompass human movements other than those related to empty-handed gestures.

## 2.3 Instrumental gestures

Back to the musical universe (and considering only gestures where some physical contact takes place), important work relative to gestural control of music has been developed by C. Cadoz and co-workers [10] [4] [14] [5] [6]. Cadoz established guidelines for the study of a specific type of gesture that is meaningful in instrumental music, which he has defined as *instrumental gesture* [4] [5]. According to him, instrumental gestures are specific to the gestural channel<sup>5</sup> and are defined as the ones applied to a material object, where physical interaction with the object takes place; the physical phenomena produced during this interaction convey some form of information and can be mastered by the subject. Cadoz also proposed a three-tier functional classification of gestures as [6]: *Excitation*, either instantaneous or continuous; *Modification*, parametric or structural; and *Selection*.

## 3 Ensemble of Performer Gestures

What do we mean by *performer gestures*? Those that have to do with the actual way of playing an instrument, i.e. the instrumentalist's own technique, both

---

<sup>5</sup> Gestural channel is defined as a means of action on the physical world as well as a communication means in a double sense: emission and reception of information [5].

instrumental gestures in the sense of Cadoz’s classification and those that may not be directly performed to produce a note.

One attempt in this direction has been proposed by François Delalande [7] in a study on the playing technique of Glenn Gould. He suggested the following gesture classification:

- *Effective gestures*, those that actually produce the sound;
- *Accompanist gestures*, body movements such as shoulder or head movements;
- *Figurative gestures*, gestures perceived by the audience through the produced sound, but without any direct correspondence to a movement. Examples would be changes in note articulation, melodic variations, etc.

In the light of Delalande’s classification, performer gestures in the sense stated above will be related to the first two items. The first one usually represents the traditional controller outputs, such as fingering. The second item, *accompanist gestures*, is the main interest of this article, but only when it directly relates to sound production.

### 3.1 Example of the Clarinet

Considering Delalande’s *effective gestures* or Cadoz’s *instrumental gestures*, let us analyze the three most common gestures of a clarinet player: blowing, lip pressure, and fingering.

Applying the typology proposed by Cadoz, one could classify breath pressure as an *excitation* gesture, lip pressure as a *parametric modification* gesture and fingering as a *selection* gesture.

These results would clearly be a simplification of the real instrument behavior, since they do not take into account the subtle interdependencies of the above functions, such as the case of the reed’s physical behavior [2] [15]. Nevertheless, they do represent the case of a MIDI wind controller, such as the Yamaha WX7, where an independent MIDI stream is output according to each of the above functions.

Although the three described gestures could account for a reasonable quality synthesis of a clarinet using a good synthesis model, one notes from the analysis of clarinet performances that instrumentalists do not only make *effective* movements but actually also express themselves by means of body movements<sup>6</sup>. We will use here the term *ancillary*<sup>7</sup> to designate only those gestures applied to the instrument, while Delalande considers *all* body movements as *accompanist*, changes in body posture and instrument movements during the performance.

---

<sup>6</sup> These movements will be considered as gestures, in the sense of Delalande’s *accompanist gestures*.

<sup>7</sup> Thanks to Mark Goldstein for the suggestion.

## 4 Non-Obvious Performer Gestures

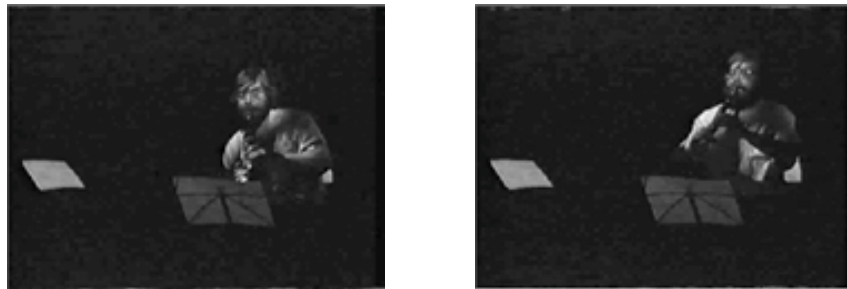
Let us consider here *non-obvious* or *ancillary* gestures as a class of wind instrument performer gestures that are produced by means of moving the instrument during the performance - *lifting it up/putting it down, to one side or another, fast tilt-like gestures, etc.*

For the time being, two main groups of non-obvious gestures will be devised according to their amplitudes ranges: *large-amplitude* and *small-amplitude* gestures.

This paper mainly deals with gestures that present a fairly large range of movement. Although these may be produced consciously – because of composer’s explicit requirements (some pieces by K. Stockhausen, for instance), as a visual effect that is immediately perceived by the audience, or as part of a communication language between players in an ensemble – the goal of this study is to analyze ancillary gestures produced by the musician while playing.

In order to develop a formal analysis of these gestures, three different clarinet players were observed in different circumstances:

- A video of French clarinetist Alain Damiens rehearsing a solo clarinet piece by Pierre Boulez, *Domaines*, produced at Ircam in 1985.<sup>8</sup>
- A video of Marc Battier’s<sup>9</sup> clarinet piece *Mixed Media*<sup>10</sup>, recorded during a concert in Kobe, Japan, 1993.
- Different acoustic and electronic performances of American clarinetist / composer Joseph Butch Rovin, during his residency at the Institute, 1996-1998.



**Fig. 1.** Alain Damiens playing *Domaines* by Pierre Boulez. Two shots taken less than a second apart - *Cahier D, Original*. Note the displacement of the clarinet and also the change in posture.

In order to complement the information obtained from the videos, we have further analyzed clarinet samples recorded by French clarinetist Pierre Dutrieu

<sup>8</sup> I would like to thank the copyright owners who kindly agreed on the usage of this video for this research.

<sup>9</sup> Thanks to Marc Battier for providing the film.

<sup>10</sup> Unfortunately, the name of the player could not be found.

in the framework of the Studio-on-Line project<sup>11</sup>; discussed with Dutrieu and also with French alto-saxophonist Claude Delangle, and finally with German musicologist and (amateur) clarinetist Peter Hoffmann.



**Fig. 2.** Three shots showing a fast upwards gesture. This sequence originally lasts one second - *Cahier C, Original*. Note (*left*) the *standard* posture, (*center*) the lowest point, and (*right*) the final point of the movement.

An analysis of the first video reveals certain gestural patterns. There are mainly three movements occurring at specific moments:

- Changes in posture at the beginning and during phrases.
- Slow continuous gestures, usually in an upward direction during long sustained notes, generally increasing in amplitude with an increase in the note's dynamics.
- Fast sweeping movements of the bell that mainly accompany short staccato notes.

Analyzing the data from the two other clarinet players, it can be seen that continuous movements may also be found in sustained notes. Nevertheless, not many fast gestures were found in the second video, but changes in posture seemed to be more frequent. Also, the amplitude range of these movements varies from player to player.

Finally, it has also been noticed that even when playing an electronic controller, the third player had the tendency to produce similar movements to the ones with an acoustic instrument.

#### 4.1 Analysis

Due to the small quantity of analysis samples available, it cannot be stated that any of these gestural patterns will be reproduced in every circumstance, although this first analysis does suggest that these basic gesture patterns may exist, at least in the case of the same player. As an example, a phrase is repeated twice in the first film, in the introductory credits and later during the piece (*Cahier D, original*). It is interesting to note that the player reproduced the same ancillary gestures when performing the sequence each time.

<sup>11</sup> A sound database with complete recordings of most orchestral instruments. For more information, see <http://www.ircam.fr/>

One can imagine that there may be different causes for these (different) kinds of gestures. Some of the observed gestures seem to be produced in order to express extra information than that conveyed by the sound, and we suppose that they may therefore be influenced by cultural and situational factors, such as musical style, room size and type, size of audience, etc. Others may be the result of particular technical difficulties encountered when playing the instrument, and thus have their origin in human physiological characteristics. Finally, some of the low-amplitude movements result from the simple fact that it is humanly impossible to play a wind instrument absolutely immobilized.

After discussions with clarinet and saxophone players, it seems not exact that these gestures are only produced in order to *consciously* express extra information, such as lifting the bell in order to sound brighter. Moreover, the performers reported that they are not aware of most of these movements and even were rather surprised when watching the sample movies or the sound analysis results.

Against the hypothesis of technical difficulties is the notion of *expert* performance [13]. We are here considering top performers who have long overcome most technical difficulties. Another point related to the relative role of motor control is that recent studies [9] on the roles of motor control and mental representation in children performance have shown that once a mental representation of the piece is developed, similar execution time profiles have been found between performances of the same piece in different instruments, independently from the different motor control skills required.

## 4.2 Further Considerations

I would also like to discuss in this paper the possible benefits gained from the research on co-verbal gestures. Although performer gestures are of a fundamentally different nature, this research may at least benefit from the techniques developed for the study of these gestures.

Considering it in more detail, one could state that some ancillary performer gestures accompany (augment or complement) the information that is conveyed by the primary channel (the sound) and give extra (visual) clues on the performer's musical intentions to the audience. Another point is that ancillary performer gestures may influence the primary information received by the audience, and this influence may also be perceptible.

It remains to be discussed whether these gestures will present universal, recognizable patterns among different players and what might be the exact influence of the environment on them.

## 5 Influence on the Instrument's Sound

Let's now show why it is important to study ancillary gestures by analyzing their acoustical influence on the sound. This will be done by presenting several experiments performed to understand their effects. These experiments<sup>12</sup> have

---

<sup>12</sup> Performed in close cooperation with Olivier Warusfel, Philippe Depalle, René Caussé, Federico Cruz-Barney, Gérard Bertrand, Joseph Rován and Peter Hoffmann.

been devised and performed in special circumstances in order to isolate specific phenomena each time.

## 5.1 Experiments

The main experiments performed were:

- Recording of clarinet sustained notes (7 seconds in average) in an anechoic chamber<sup>13</sup> under two conditions: empty anechoic chamber and anechoic chamber with the inclusion of a wooden floor. Both were performed three times:
  - Instrument kept immobilized by a mechanical apparatus;
  - Instrument played in a normal way (conventional, non-exaggerated);
  - Instrument played in an exaggerated way - i.e. large-amplitude movements.
- Room response measurements where the excitation was provided by a loudspeaker connected to a clarinet tube placed at different angles.
- Recordings with a clarinet immobilized by the same mechanical apparatus used during the anechoic chamber recordings, but in an auditorium with variable acoustics<sup>14</sup> in a reverberant configuration.
- Clarinet recordings from the Studio-on-Line database, where each note is available in three different dynamics (pianissimo, mezzo-forte, and fortissimo) and recorded by six different microphones - two reference (2 meters), one close (1 meter), one internal and two far microphones (around 15 meters away).

## 5.2 Discussion

A detailed quantitative analysis of the obtained results is beyond the scope of this article and is presented elsewhere[16]. We will here directly present some conclusions in order to show the influence of these gestures in the sound.

Movements of the instrument with respect to a close fixed microphone will mostly cause variations in the attributes of the direct sound and the first reflections captured by the microphone. Actually, both the amplitude and time of arrival of the direct sound and of the first reflection will be modified with a change in position, causing the modulations found in some of the analyzed samples.

This effect can be further explained by the analysis of the figures below. Figure 3 shows a D3 *ff* recorded in a reverberant auditorium: the left side shows a sample from the Studio-on-Line database, where one can notice strong partial amplitude variations. The right side shows the same note recorded with an internal microphone. Note that the amplitudes of the sinusoidal partials are fairly

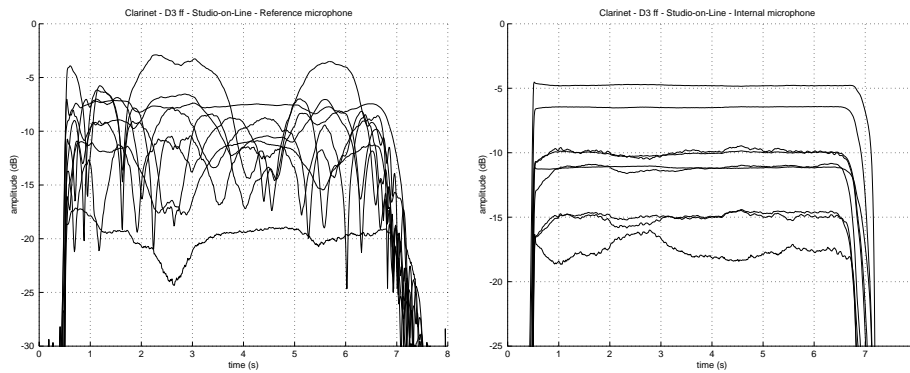
---

<sup>13</sup> A special room where acoustically absorbent material is placed in order to avoid sound reflections normally produced by walls, ceiling, and floor.

<sup>14</sup> The Espace de Projection at Ircam. Its acoustics can be modified from dry to strongly reverberant.

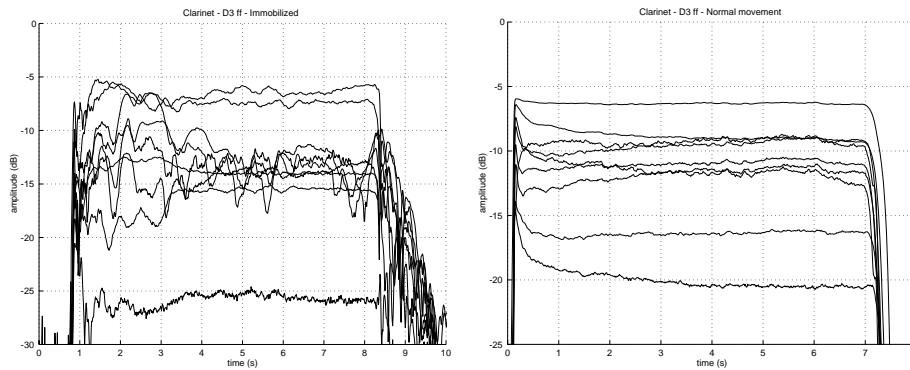


constant in this case. This amounts for a certain stability of the embouchure, what could be expected since we're dealing with expert performers.



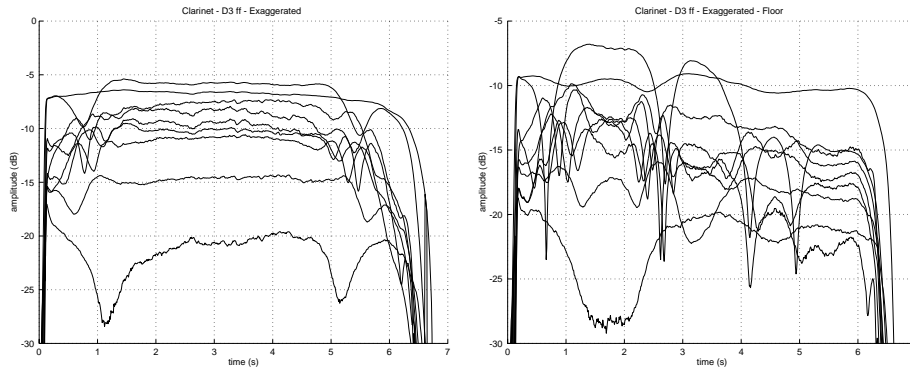
**Fig. 3.** Clarinet recordings in a reverberant auditorium (Espace de Projection - Ircam) - D3 recorded *ff*, standard playing - Studio-on-Line database - (P. Dutrieu): (*left*) reference microphone (2 meters in front of the instrument); (*right*) internal microphone.

The left side of figure 4 shows the note recorded in the same auditorium with a clarinet immobilized by a mechanical apparatus. Except for minor fluctuation in one of the partials, it is clear that the same variations did not occur to the same extent. Finally, the right side of this figure shows a recording of the same note in an anechoic room. Again, the partials present a rather constant amplitude in time.



**Fig. 4.** Clarinet recordings - D3 recorded *ff*: (*left*) reverberant auditorium - clarinet immobilized by mechanical apparatus (Peter Hoffmann); (*right*) empty anechoic chamber - clarinet played in a standard manner (J. Rován).

Figure 5 shows the analysis of D3 recorded in an anechoic chamber when large movements were performed. One may notice small differences in the amplitude values of the sinusoidal partials. This shows that the instrument’s directivity alone does not play the major role in these variations. Furthermore, the right side of the figure shows the analysis of the same note recorded with a wooden floor placed between the instrument and the microphone, again with large movements. The floor deepens in the amplitude variations to more than 12 dB even for low frequencies, producing effects that are similar to the ones found in the original sample (figure 3).



**Fig. 5.** Anechoic room recordings (J. Rován) - D3 recorded *ff* and exaggerated movement: (*left*) empty anechoic room; (*right*) wooden floor placed between the instrument and microphone.

### 5.3 Simulation

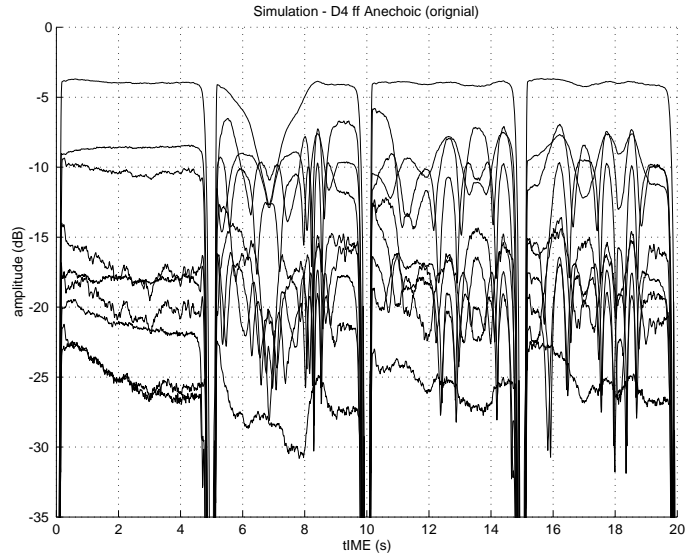
In order to verify the previous analysis, we have made tests with a model of these variations in the FTS sound synthesis environment, where results of the room response measurements were implemented as parameters of continuous delay operations<sup>15</sup>. The results are shown in figure 6.

The input of the system consists of a sound file recorded in an empty anechoic chamber with the clarinet immobilized. A slider is used to simulate a one dimensional movement of the clarinet from an angle perpendicular to the floor to an horizontal position. Since the measurements presented before did not show major influences of neither the embouchure nor the instrument’s directivity, the simulation model just takes into account the influence of the first reflection.

One can see that using this simplified model strong amplitude variations may be produced, depending on the clarinet movement simulated<sup>16</sup>.

<sup>15</sup> No spectral changes due to the reflected sound have been implemented at this stage.

<sup>16</sup> One must not forget that the effect of early reflections is completely tied up to the microphone type and position and may even be disregarded in the case of averaging multiple recordings or using distant microphones.



**Fig. 6.** Simulation of the effect of different movements by modification of direct sound and first reflection attributes according to the clarinet's angle. The first quarter of the picture displays the amplitudes of the original sound's sinusoidal partials (anechoic chamber). The remaining ones represent different movements applied to the instrument.

## 6 Conclusions

In this article different topics of the research on gesture applied to music have been discussed, mainly in the case of instrument performance.

I commented on the broad range of ideas the word gesture may have in a computer music context and then moved on to review some of the previous work relating gestures and music. After analyzing the existing classifications of gestures in music in the light of the behavior of acoustic wind instruments, it was pointed out that many gestures that are not usually referred to as important in sound synthesis should be considered, since they may actually affect the sound captured from the instrument.

It has been verified by measurements and simulation that the variations of the direct sound and floor reflection attributes for a specific movement may cause sinusoidal partial amplitude modulations that play an important role in the resulting sound for a particular microphone position.

## 7 Acknowledgements

Many thanks to Stephen McAdams, Xavier Rodet and Mark Goldstein for comments on previous versions of this manuscript, and also Philippe Depalle, Olivier Warusfel, François Delalande, Pierre Dutrieu and Federico Cruz-Barney for useful discussions and suggestions.

This work is supported in part by a scholarship from the CNPq, Brazilian National Research Council.

## References

- [1] Baecker, R. M., Grudin, J., Buxton, W. A. S., Greenberg, S.: *Readings in Human-Computer Interaction: Toward the Year 2000*, Morgan-Kaufmann, 2nd Edition, Part III, ch. 7 (1995) 469–482.
- [2] Benade, A. H.: *Fundamentals of Musical Acoustics*, Second edition, Dover (1990).
- [3] Borin, G., De Poli, G., and Sarti, A.: Musical Signal Synthesis, in *Musical Signal Processing*, C. Roads, S. T. Pope, A. Piccialli, and G. De Poli (eds), Swets & Zeitlinger B.V. ch. 1 (1997) 5–30.
- [4] Cadoz, C.: Instrumental Gesture and Musical Composition, Proc. Int. Computer Music Conf., ICMC, (1988) 1–12.
- [5] Cadoz, C.: Le geste canal de communication homme-machine. La communication “instrumentale” *Sciences Informatiques - Numéro Spécial: Interface Homme-Machine* (1994) 31–61.
- [6] Cadoz, C.: Musique, geste technologie, in *Cultures Musicales: Les Nouveaux Gestes de la Musique*, H. Génevoix and R. De Vivo (eds), Parentheses (1999).
- [7] Delalande, F.: La gestique de Glenn Gould, in *Glenn Gould Pluriel*, Louise Courteau Editrice (1988) 84–111.
- [8] Depalle, P., Tassart, S., Wanderley, M.: Instruments virtuels - Les vertues du possible. *Résonance*, 12 (1997) 5–8.
- [9] Drake, C.: Aux fondements du geste musical, *Science et Vie - Numéro Spécial: Le Cerveau et le mouvement*, A. Berthoz (ed), (1998) 114 - 121.
- [10] Gibet, S.: Codage, représentation et traitement du geste instrumental, *PhD Thesis*, Institut National Polytechnique de Grenoble (1987).
- [11] Hummels, C., Smets, G., and Overbeeke, K.: An Intuitive Two-Handed Gestural Interface for Computer Supported Product Design, in *Gesture and Sign Language in Human-Computer Interaction*, I. Wachsmuth and M. Frölich (eds), Springer Verlag (1998) 197–208.
- [12] Kurtenbach, G. and Hulteen, E. A.: Gestures in Human-Computer Communication, in *The Art of Human-Computer Interface Design*, B. Laurel (ed), Addison Wesley, (1990).
- [13] Lehmann, A. C.: Efficiency of Deliberate Practice as a Moderating Variable in Accounting for Sub-expert Performance, in *Perception and Cognition of Music*, I. Deliège and J. Sloboda (eds), Psychology Press, (1997).
- [14] Ramstein, C.: Analyse, représentation et traitement du geste instrumental, *PhD Thesis*, Institut National Polytechnique de Grenoble (1991).
- [15] Rován, J., Wanderley, M., Dubnov, S., Depalle, P.: Instrumental Gestural Mapping Strategies as Expressivity Determinants in Computer Music Performance. Proceedings of the Kansei - The Technology of Emotion Workshop, A. Camurri (ed), Genoa - Italy (1997) 68–73.
- [16] Wanderley, M., Depalle, P., and Warusfel, O.: Improving Instrumental Sound Synthesis by Modeling the Effects of Performer Gesture, Proc. Int. Computer Music Conf., ICMC, China (1999).
- [17] Wilson, A. D., Bobick, A. F., and Cassell, J.: Recovering the Temporal Structure of Natural Gesture, Proceedings 2nd Int. Conf. on Automatic Face and Gesture Recognition (1996).