

Study of haptic and visual interaction for sound and music control in the Phase project

Xavier Rodet,
Jean-Philippe Lambert,
Roland Cahen
Thomas Gaudy,
Fabrice Guedy
Institute of Research and
Acoustic/Music Coordination,
1 Place Stravinsky, 75004 Paris,
France
33 1 44 78 48 68
xavier.rodet@ircam.fr

Florian Gosselin
CEA-LIST
Route du Panorama BP6, 92265
Fontenay aux Roses Cedex, France
33 1 86 54 89 18
florian.gosselin@cea.fr

Pascal Mobuchon
ONDIM
14 rue Soleillet, 75020 Paris, France
33 1 40 33 88 04
mobuchon@ondim.fr

ABSTRACT

The PHASE project is a research project devoted to the study and the realization of systems of multi-modal interaction for generation, handling and control of sound and music. Supported by the network RIAM (Recherche et Innovation en Audiovisuel et Multimédia), it was carried out by the CEA-LIST for haptic research, Haption for the realization of the haptic device, Ondim for integration and visual realization and Ircam for research and realization about sound, music and the metaphors for interaction. The integration of the three modalities offers completely innovative capacities for interaction. The objectives are scientific, cultural and educational. Finally, an additional objective was to test such a prototype system, including its haptic arm, in real conditions for general public and over a long duration in order to measure its solidity, its reliability and its interest for users. Thus, during the last three months of the project, a demonstrator was presented and evaluated in a museum in Paris, in the form of an interactive installation offering the public a musical game. Different from a video game, the aim is not to animate the pixels on the screen but to play music and to incite musical awareness.

Keywords

Haptic, interaction, sound, music, control, installation.

1. INTRODUCTION

The objectives of the PHASE (Plateforme Haptique d'Applications Sonores pour l'Eveil musical) project are to carry out interactive systems where haptic, sound and visual modalities cooperate. The integration of these three interaction channels offer completely innovative interaction capacities, comprising gesture with force and tactile feedback, 3D visualization, and sound spatialization.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

NIME'05, May 26-28, 2005, Vancouver, BC, Canada.

Copyright remains with the author(s).

The objectives are scientific, cultural and educational:

- to study the cognitive processes implied in musical exploration through the perception of sound, gesture precision, instinctive gestures and the visual modality,
- to try out new ways of interactivity (haptic interaction, spatialized sound and 3D visualization, in real time),
- to propose new sorts of musical gestures, and by extension, to enrich the methods of training of technical gestures in any field (industry, handicraft, education, etc.),
- to improve characterization of sound and its relationship to music (exploration of the multidimensional space of timbre) and to offer a new kind of listening,
- to implement new methods of sound generation and to evaluate them with respect to a public of amateurs and professionals,
- to propose a new form of musical awareness for a large audience, specialized or not.

Finally, an additional objective was to test such a system, including its haptic arm, in real conditions for general public and over a long duration in order to study and to measure its robustness and reliability and to assess its interest for users. Thus, during the last three months of the project, a demonstrator of PHASE developments was presented and evaluated at the G. Pompidou museum in Paris, in the form of an interactive installation offering the public a musical game. Different from a video game, the aim is not to animate the pixels on the screen but to play music and to incite musical awareness. The teaching aspect was thus significant for children and for adults.

Section 2 presents studies on the use of gestures with haptic feedback for playing music, section 3 describes the haptic interface itself and section 4 explains the design and usage of the demonstrator.

2. Gesture interaction with sound and music

2.1 Gesture, music and metaphors

Gesture is the entrance point of the multi-modal systems considered in the Phase project. A series of studies were thus carried out to measure the relevance and the properties of some gestures in the specific context of musical control. In view of

gesture interaction, the play of a traditional instrument is the first idea which comes when thinking of playing music, but there are many others ways. More generally, we searched for which freedoms could be left to the player to express himself/herself in a given musical space. One can, for example, have access to rate, rhythm, timbre, tonality, etc. Various modes of gestural and musical playing have thus been studied. The purpose of the user is to play and control music. The user's gesture is carried out in the real space (containing his hand). Various feedbacks (haptic, sound and visual) influence his gesture. For example, to help the user to play a musical instrument which requires long practice before precise control is reached, haptic feedback can be used as a guide. Note that the training can also be facilitated by a scenarisation of the course of the demonstrator game.

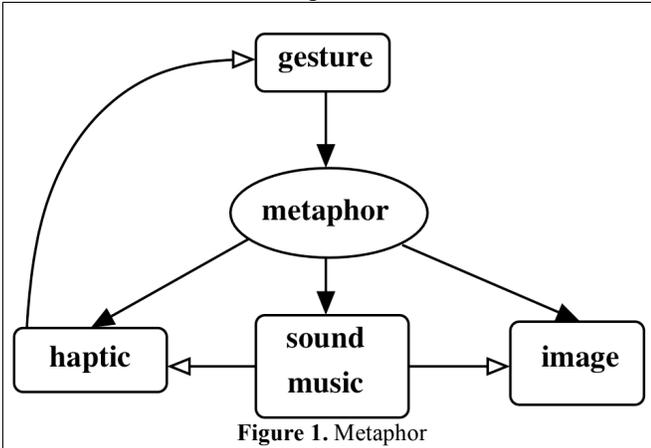


Figure 1. Metaphor

Opposite to vision and hearing, the haptic feedback loop is direct since gesture and force feedback are localized in the hand of the user. Among other important necessities, coherence of the various modalities must also be guaranteed between the real and virtual world:

- in the real world, the gesture is performed within a spatial universe with temporal dynamics or not,
- in the virtual world, a physics engine calculates the force feedback for the haptic device (the objects are defined by their form, mass and contact properties and they interact via forces).

Considering these constraints, all the modalities must be coherent for the user. This is achieved by using a *metaphor* which defines the link between the real world where the hand acts and the virtual music world (Fig. 1).

According to the context, one can define various metaphors, which lead to different implementations, more or less close to reality. Let us give only a few examples, which are combined in the demonstrator game in order to enrich the interaction:

- direct interaction resembles sonification. The physical model is taken as a starting point to produce a significant immersion for the player, since the modalities are then close to those usually perceived in reality. In this case, physical objects are also sound sources. The interactions between the user and these objects are similar to the play of a musical instrument.
- zones of interaction: the physical world is just a musical pretext, for instance of sound navigation. An example is the case of a microphone moving in space for audio mixing.
- re-play of a recorded music inscribed within the three-dimensional space.

A real world analog of this last metaphor is that of the needle

inside a vinyl groove: the musical course is then strongly related to the space course. According to whether the disc is spinning or not, various temporal controls are obtained : 1- One can control temporal pace directly if this one is traced in space; this is what occurs for example when one holds a *needle* in his hand and that one moves it along in the groove of a vinyl disc. 2- If, on the contrary, time moves along in space, one can control for example the relative reading speed as if the vinyl disc would spin. The speed of the needle is added to that of the disc, and creates accelerations or decelerations relative to the reference speed which is that of the disc. 3- One can control the flow of time in space by controlling the speed of the disc, i.e. the course of the music which is recorded on the disc. Note that music can be recorded not only just in the groove but as well in its neighborhood: for instance, the further from the groove the more modified the music is. In this way of playing a music recorded in a 3D domain, the musical course is strongly related to the course in space. This last metaphor has been particularly developed to be the heart of the game in the demonstrator.

Generally, in a given metaphor, each physical object is attributed some behavior, and visual and sound objects are then associated to them. This mapping strategy [13] uses several layers to connect physical objects, visual objects and sound objects.

2.2 Playing music

One of the very innovative aspects of the PHASE project is the aim of generating, handling and controlling sound and music. For this aspect, composer Roland Cahen has been deeply involved in the project [17]. Different ways of using gesture for playing music have been studied. On one side, the player can, for example, have access to rhythm, timbre, tonality, etc., and on the other side different gestures can be used for this control. Different modes of playing (from a gestural and musical point of view) have been identified: positioning oneself in the music, browsing the music, conducting the music, playing the music like an instrument, etc.

There are also different kinds of music that can be played (and re-played):

- well-known music can easily be identified and the control of a temporal position is easy, but the music is otherwise fixed and not adjustable;
- specific music can be composed to be modified by players;
- one can program music generators so that they interact with players.

Different musical interactive tools have thus been realized, that allows to:

- manipulate musical phrases, specifically created and easy to transform;
- generate a staccato flow enabling numerous programmed and interactive variations; Work on the sound material – recorded or generated – has also been realized using the granular synthesis for instance. This offers a privileged access to time allowing the system to re-play the music in different ways:
- one can control directly the temporal development with a gesture in space;
- one can as well control a relative tempo, or acceleration and deceleration.

Another approach consists in using a virtual physical world and listening to it in different ways. It is then tempting to exploit the different measures that come from the physical engine directly, particularly with the so-called physical model synthesis.

Finally, the musical link between the different physical and sound

objects needs to be defined. Writing and playing such sound and music, using dynamic controls remains a challenging task

2.2 Gesture and haptic feedback

On account of the complexity of the whole system, a module has been realized in order to abstract gesture control from various implementations. This allows to test musical or sound behaviors and interactive gestures without using the haptic system itself which involves technically complex realizations. For example, a graphic tablet or a simple mouse have been used for quick experiments. One obviously needs the complete device for final adjustments, but the abstraction module permits to validate concepts and implementations under simulation. Naturally, the absence of force-feedback must be taken into account in such a simulation.

In addition, the abstraction module allows to record gestures, to re-play them (in order to test, to regulate and to validate dynamic sound behaviors) or analyze them (out of their execution time), with or without the haptic arm. Thanks to the assistance of Frederic Bevilacqua, a study was carried out to compare a graphic tablet (Wacom Intuos A4 and aerographer) and a force-feedback arm (Virtuose 6D from Haption, www.haption.com). The goal was to identify significant parameters in the context of real-time sound and music generation. In addition, this study was used to evaluate the limitations of the graphic tablet to replace a haptic arm.

In this study, subjects are asked to listen to and represent short recorded musical phrases played on marimba. For each phrase, subjects listen to it three times and, immediately after, trace representations of the music, three times using the graphic tablet and three times using the 6D haptic arm. In the case of the arm, a simple horizontal plane is simulated, but subjects are not obliged to touch it while tracing a representation. Subjects can freely choose a representation but have to stay coherent for the six traces. Representations were recorded and then informally analyzed.

As one could expect, the absolute position and, to a lesser extent, the direction of gestures vary largely in absence of a haptic guide. After some low-pass filtering, absolute speeds and accelerations appear relatively similar for the two devices, for the same subject and the same sentence. The inertia of the haptic device used at the beginning of the project (Virtuose 6D) was however not negligible and the final model (Virtuose 3D 15-25) proved to be much better from this viewpoint.

Rather good coherence between various sentences for the same subject are observed, as well as between subjects in the case of short sentences. One can observe that representations are well correlated with the phrase structure. This is similar to the results of M. Wanderley [10], [11] in the case of clarinet gestures. Furthermore, the law of Fitts [16] and its derivative [12] have been checked on the representations. This led to the idea of using the various characteristics bound by the law (briefly, speed and radius of curvature) independently, thus in contradiction with the law, for expressivity purpose: to add musically expressiveness, the user can then choose to violate the law, for example by using a smaller or larger curvature radius than the speed of his hand would imply. The expressive possibilities of this principle proved to be very interesting in some musical sketches.

3. Haptic interface

3.1 Design Objectives

Haptic devices must allow natural and intuitive exploration of virtual worlds. Whether considering virtual reality, virtual art or CAD design, they share common design objectives [1], [2], [3]:

- **Performance criteria :** Ideally, an input device must be 'transparent'. The user must feel as if he is acting directly in the virtual environment. He must feel free in unencumbered space (which requires a large and singularity free workspace, low inertia and low friction). He must also feel crisp contacts against the obstacles (which requires a sufficient force feedback, a high bandwidth and a large stiffness).
- **Integration criteria :** The input device must be compact enough to be easily integrated in a workstation or exhibition area.
- **Maintainability criteria :** It must be designed with as much as possible commercially available components to minimize cost and simplify maintenance.

Although input devices are designed considering these qualitative criteria, some of these criteria are contradictory. As a matter of fact, existing devices exhibit very different performances [4], [5], [6] and are more or less adapted to different applications. In the design phase of a given device, it is however critical to know how efficiently it will fit a particular task. Precise requirements must thus be associated with each criterion.



Figure 2. Modified Virtuose 3D 15-25

Considering the fine and delicate nature of musical composition or instrument playing, it was decided to favour sensitivity and dexterity over maximum force capacities or large workspace. Therefore, a stylus handle for dexterity and an elbow support for precision have been preferred, however keeping a forearm-size working space. The range of motions and efforts associated with these kind of movements are around 200mm and 15N [7]. Based on the experience acquired on previous developments of haptic interfaces, position and force resolution requirements were set to 100 μ m and 0.5N. To render a realistic and palpable virtual world, the global stiffness requirements were set to 1500N/m. Moreover, the haptic interface must be rugged enough to be used in open public areas. It must also be sufficiently generic to be used by left or right handed people of different ages and sex. Considering previous requirements, a Virtuose 3D 15-25 haptic interface sold by French society Haption was chosen as it offers a rugged and validated basis that could easily be adapted to the PHASE project specific needs (Fig. 2). This haptic interface answers the design drivers given in table 1 as it is the simplified 3 DOF with force feedback industrial version of a 6 DOF with force feedback master arm previously

developed at CEA-LIST for tele-surgery applications which share same or close requirements [8].

3.2 Device adaptation to the PHASE requirements

The VIRTUOSE 3D 15-25 answers the PHASE project requirements considering workspace, force capacity and force quality. It allows to realistically render palpable objects or to guide the operator along a specified trajectory.

However, it doesn't offer a stylus handle nor tactile feedback. Several specific handles illustrated in figure 3 were thus designed at CEA-LIST. They share common requirements. They integrate on one side two buttons to be able to select some functions during the application and on the other side a tactile actuator, allowing high frequency force feedback and texture rendering. Considering the general public use, these buttons and actuators are either accessible from both sides or duplicated to allow left handed as well as right handed use.

Prototypes were then manufactured and tested by adults from ONDIM and IRCAM as well as pupils from the Atelier des Feuillantines in Paris. During these tests, the different handles were equipped with either a short pen (3.5cm long) or a longer one (15cm long) at their distal extremity and people were asked:

- on one side with the short pen to reproduce a tenth of basic mathematical figures (circles, squares, triangles, ...), to reproduce textures (hatching, ...) and finally to blot a drawing with the best possible accuracy,
- on the other side with the long pen to perform specific gestures of orchestra conductors.

The third design was found to be the most adapted. It was further refined then mounted on the Virtuose 3D 15-25 as shown on figure 2.

The tactile actuator is a linear electromagnetic actuator composed of a fixed coil and two moving magnets that are displaced in and out of the coil. The electromagnetic force is directly proportional to the product $N \times I$ (where N is the number of turns of the coil and I the applied current). The tactile actuator is symmetric to allow both right and left handed to operate the haptic interface and get a tactile stimulation. Both magnets are guided through a flexural bearing which is mainly composed of two sheets made of steel that have been laser cut to the appropriate shape (see Figure 3 an exploded view revealing the different components of the actuator: coil, magnets and flexural beams). The dynamic performance of such vibrotactile actuator are: Maximum stroke: $150\mu\text{m}$, Bandwidth: $>800\text{Hz}$, Dynamic force (peak): $\sim 50\text{mN}$.

A PWM (Pulse Width Modulation) power driver voltage to current amplifier has been used. The PWM power driver, available in 5×5 mm packages and are designed

to drive up to $\pm 1.5\text{A}$ output current. It is possible to limit by hardware the maximum output current to a safe value for the actuator (typically 300mA). The PWM switching frequency is adjustable up to 1MHz through an external resistor, which allows to use some very small energy storing inductances. The control voltage supply is $0-3\text{V}$. The bandwidth of the output current into the load signal is about 1150Hz , which is more than enough for the tactile application. A matrix of 8×8 micro electromagnetic actuator (VITAL interface) based on this principle has been previously developed at CEA LIST [9].

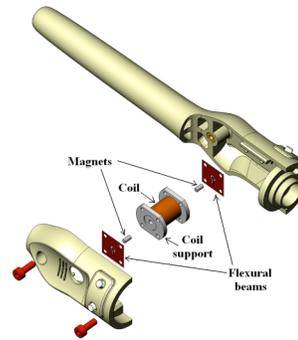


Figure 3. Tactile actuator

4. Multi-modal demonstrator

The developments made in the context of the PHASE project lead to the implementation of a multi-modal demonstrator intended for general public in real conditions and over a long duration.

4.1 Scenario

While in the large majority of cases, the music of video games is put at the service of visual or of the scenario and not the reverse, a part of our research was devoted to the use of the three modalities for the processes of musical awareness: how to design a device to sensitize general public with new listening and musical creativity modes?

To answer this question, a simple scenario has been conceived with a very intuitive beginning [18], [14]. It is based on the metaphor of a writing head and a playing head. This metaphor thus involves three main actors, the two heads and the human player, and allows for a significant amount of interaction possibilities:

- one hears and sees the writing head (WH) which generates and plays music according to the player's actions. Music is also inscribed in space in the form of a visual and haptic trace (in a sort of groove) that the WH leaves behind it as it moves.
- the player handles the playing head (PH). When on the trace in the groove, he feels and hears the music which he is re-playing. The PH pursues the WH.
- the trace links the PH and the WH together. The WH writes the trace in the groove which scrolls when the PH is inside it, like a needle in a vinyl disc groove. Thus, the movement of the PH is relative to the trace and the player controls indirectly the speed of the re-played music. To follow or even catch up the WH, the player must follow the music as exactly as possible along the trace.

The choice of this metaphor is justified for several reasons:

- an analogy with the real world: one can think of a vinyl disc the groove of which is engraved by the WH and re-played by a needle held in one's hand (the PH),
- a well-known game mode: a race between the WH and the PH,
- a musical principle of re-play, close to the fugue, the canon or the counterpoint,
- facility of gesture: the trace written inside a groove serves as a guide for the player.

The player must listen to the WH and play with it. This is a standard musical situation, except that the rhythm is not quantified: the relative speed of the two protagonists varies continuously. It can then become difficult to distinguish the two

musical actors. This is a musical issue in itself, for the system design, in order to facilitate the game, but also for the player who must listen both to the WH and to the PH in order to position himself/herself within the whole musical piece.

4.2 Implementation

The demonstrator is made up of: a computer (RTAI-Linux) which manages the position encoders and the engines of the haptic arm; a computer (Windows XP) which runs the software for simulation of physical interactions (Vortex) and centralizes Ethernet communications; a computer (MacOS X) which takes care of sound and music generation (MAX/MSP); a computer (MacOS X) which spatializes the sound on the restitution device (Spatialisateur); eight loudspeakers and two video projectors which ensure 3D

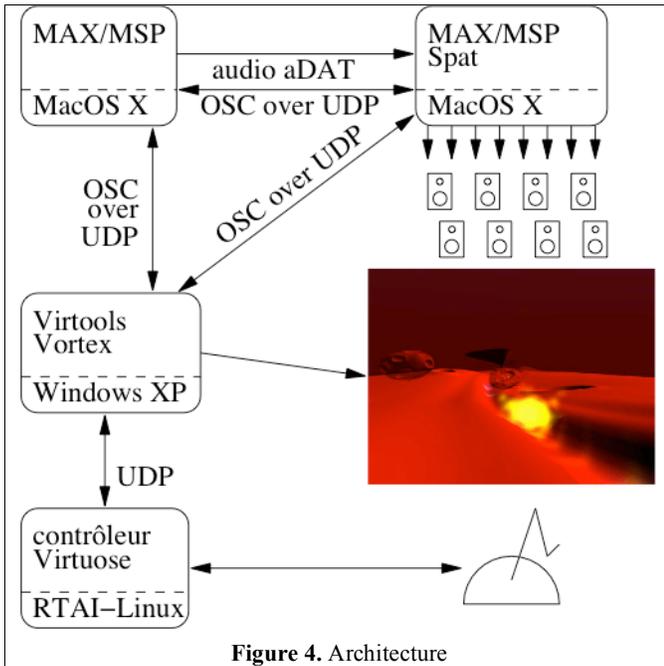
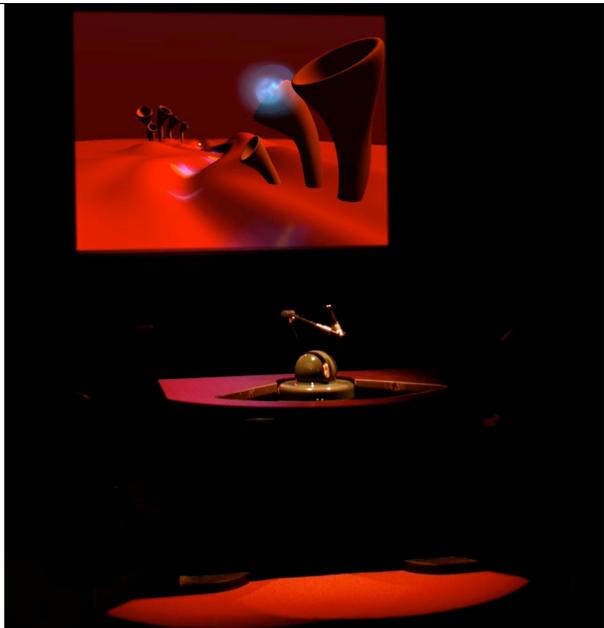


Figure 4. Architecture

sound as well as 3D visuals (Figure 4).

Since the demonstrator was conceived to be used by the general public for a long duration, particular care was brought to its



safety, its robustness and its reliability. Moreover, to allow for an elbow supported manipulation as previously specified, an ergonomic desk was designed and manufactured at CEA-LIST (Figure 5). It enables a precise and comfortable use without fatigue. It also allows an immediate understanding of the workspace of the arm, thus making the handling intuitive.

4.3 Results

The demonstrator was in free access at the G. Pompidou museum for three months. While visitors using the demonstrator were playing music, more precisely “interpreting” in some sense R. Cahen’s composition, their interpretation was recorded on a CD that the player could take with him at the end of his/her session (5 to 20 minutes). The demonstrator was an extraordinary success. The number of visitors is estimated at approximately 20.000, including children, adults, musicians, composers, teachers, etc. The reaction of the public was constantly enthusiastic, underlining the interest, the innovation and the pedagogical aspect. The realization of this interactive installation for general public in the form of a musical game integrating various metaphors, shows the validity of such a device and opens the way to many possibilities of gestural music control.

5. Conclusion

The PHASE project was first dedicated to research and experiments on the use of gesture control with haptic, visual and sound feedback, with the aim of playing music. Hardware, software and methodological tools were designed, in order to allow for the development and the realization of metaphors having a musical purpose. These developments are now available and usable. They were integrated in a demonstrator used in an open public area with a great success. It opens the way to many possibilities of gestural music control

6. REFERENCES

1. T.H. Massie, J.K. Salisbury, “The PHANToM haptic interface : a device for probing virtual objects”, Proceedings of the ASME Winter Annual Meeting, Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, Chicago, November 1994.
2. K. Young Woo, B.D. Jin, D.S. Kwon, “A 6-DOF force reflecting hand controller using the fivebar parallel mechanism”, Proceedings of the 1998 IEEE International Conference on Robotics and Automation, Louvain, Belgium, May 1998, pp. 1597-1602.
3. R. Baumann, R. Clavel, “Haptic interface for virtual reality based minimally invasive surgery simulation”, Proceedings of the 1998 IEEE International Conference on Robotics and Automation, Louvain, Belgium, May 1998, pp. 381-386.
4. D.A. McAfee, P. Fiorini, “Hand Controller Design Requirements and Performance Issues in Telerobotics”, ICAR

91. Robots in Unstructured Environments, Pisa, Italy, June 1991, pp. 186-192.
5. G.W. Köhler, "Typenbuch der Manipulatoren - Manipulator Type Book", Thiemig Taschenbücher, Verlag Karl Thiemig, München, 1981.
6. G. Burdea, P. Coiffet, "La réalité virtuelle" (in French), Hermès Publishing, Paris, 1993.
7. F. Gosselin, "Développement d'outils d'aide à la conception d'organes de commande pour la téléopération à retour d'effort", Ph.D. diss. (in French), University of Poitiers, June 2000.
8. F. Gosselin, A. Riwan, D. Ponsort, J.P. Friconneau, P. Gravez, "Design of a new input device for telesurgery", World Automation Congress 2004, Proceedings of ISORA 2004 10th Int Symp. on Robotics and Applic, June 28-July 1, 2004, Seville, Spain, Paper ISORA 120
9. M. Benali-Khoudja, M. Hafez, J.M. Alexandre, A. Kheddar, and V. Moreau "VITAL: A New Low-Cost Vibrotactile Display System" IEEE International Conference on Robotics and Automation ICRA 2004, 26 April – 1st of May 2004, New Orleans, USA.
10. M. Wanderley, "Interaction Musicien-Instrument : application au contrôle gestuel de la synthèse sonore" Thèse d'Université, Université Paris-6, 2001, Paris.
11. C. Cadoz and M. M. Wanderley. Gesture - Music. In M.M. Wanderley and M. Battier, editors, Trends in Gestural Control of Music, pages 71–94. Ircam - Centre Pompidou, 2000.
12. John Accot and Shumin Zhai, "Beyond fitts' law : Models for trajectory-based hci tasks", In ACM/SIGCHI : Conference on Huma Factors in Computing Systems (CHI), 1997.
13. A. Hunt, M. Wanderley, and R. Kirk, "Toward a Model for Instrumental Mapping in Expert Musical Interaction", in Proc. of the Int. Comp. Music Conf (ICMC), 1999
14. V. Gal, C. Le Prado, S. Natkin, and Liliana Vega, "Writing for video games", In Proceedings Laval Virtual (IVRC), 2002.
15. M. Wright and A. Freed, "Open Sound Control : A New Protocol for Communicating with sound Synthesizers", in Proc. of the Int. Comp. Music Conf (ICMC), 1997.
16. P. M. Fitts. The information capacity of the human motor system in controlling the amplitude of the movement. Journal of Experimental Psychology, 47 :381–391, 1954.
17. R. Cahen, "Générativité et interactivité en musique et en art électroacoustique", 2000
<http://perso.wanadoo.fr/roland.cahen/Textes/CoursSON.html/MusiqueInteractive.htm>,
18. R. Cahen, "Navigation sonore située", 2002
<http://perso.wanadoo.fr/roland.cahen/Textes/CoursSON.html/MusiqueInteractive.htm>