

# Captation et analyse du mouvement pour l'interaction entre danse et musique.

Frédéric Bevilacqua et Emmanuel Fléty

Ircam - Centre Pompidou  
1 pl. Igor Stravinsky  
75004 Paris – France  
frederic.bevilacqua@ircam.fr, emmanuel.flety@ircam.fr

***Résumé:** Nous commençons par un survol de divers systèmes interactifs entre geste et musique. Nous présentons ensuite les diverses approches que nous poursuivons pour l'analyse du mouvement et la transformation du geste en processus sonores.*

## 1 Introduction

En 2003, l'IRCAM a mis en place le "Pôle de recherche sur les technologies pour le spectacle vivant". Dans un premier temps, les sujets de recherche concernés sont principalement le geste, la voix et la spatialisation. Le "geste" regroupe à la fois les domaines de la performance (principalement danse), et du geste instrumental (musique).

Dans cet article, nous présentons les axes principaux de notre recherche sur le geste, en particulier dans le cadre de la danse. Afin de mettre en contexte notre approche sur la captation et l'analyse du mouvement, nous présentons premièrement un survol des systèmes interactifs entre geste et musique, dont certains ont été réalisés récemment à l'IRCAM. Nous décrivons ensuite les systèmes de captation et d'analyse du mouvement que nous mettons en œuvre.

## 2 Systèmes interactifs "geste-musique"

Les systèmes interactifs geste-musique<sup>1</sup> sont des systèmes où le geste est mis en relation avec un environnement sonore par un système de captation du mouvement. Nous ne pouvons pas, dans le cadre de cet article, couvrir entièrement ce sujet,

---

<sup>1</sup> Nous utilisons ici le mot interactif dans un sens large, en incluant des systèmes simplement "réactifs".

nous nous limiterons donc à décrire certains concepts et exemples.

De nombreux systèmes permettent à la fois un travail vidéo et sonore à partir du mouvement. Néanmoins, il est important de prendre note que les problématiques relatives aux interactions geste-son et geste-image sont fondamentalement différentes pour le spectateur, puisque la perception gestuelle et sonore se fait par des modalités perceptives différentes.

Nous faisons la distinction entre les dispositifs utilisés lors de performances, ceux qui sont utilisés comme outils d'écriture chorégraphique et/ou musicale et les installations interactives. Ces distinctions sont principalement liées à des choix artistiques, et non pas aux outils technologiques eux-mêmes, certains systèmes pouvant être appliqués dans les trois cas.

### 2.1 Performance

*Du geste instrumental au geste dansé*

Depuis la première démonstration du Theremin en 1920, de nombreux projets artistiques ont été réalisés utilisant un système de captation des gestes du "performer" pour contrôler un environnement sonore (pour des exemples récents, voir les conférences NIME<sup>2</sup> et le CD-Rom édité par l'Ircam [Wanderley 2000]). Il est utile de

---

<sup>2</sup> New Interfaces for Musical Expression, [www.nime.org](http://www.nime.org)

différencier les approches de type "geste instrumental" et "geste dansé", même si certains projets intègrent ces deux concepts.

Dans l'approche "geste instrumental", les mouvements du musicien-performer sont *a priori* définis ou contraints par l'instrument. Au contraire, dans l'approche "geste dansé", les mouvements sont en premier lieu dérivés d'une pratique de la danse. Dans ce cas, le système interactif met en principe à profit le langage ou en tout cas la maîtrise corporelle du danseur performer.

Parmi les musiciens/compositeurs qui ont développé une pratique instrumentale basée sur de nouvelles interfaces gestuelles [Wanderley 2000], on peut citer par exemple (parmi bien d'autres): Serge de Laubier et le "meta instrument", Michel Waisvisz et "the hands", Laetitia Sonami et "Ladys'glove" et Atau Tanaka avec la "Biomuse."

Plusieurs nouvelles interfaces instrumentales ont été développées à l'Ircam, généralement basées sur des paradigmes d'instruments déjà existants comme le violon (violin midi de Suguru Goto) ou en "augmentant" des instruments acoustiques par un système de captation, comme la "flûte acoustique avec captation midi, la clarinette & DataGlove de Butch Rovin. Nous travaillons actuellement sur un violon "augmenté", qui restitue la position, pression et accélération de l'archet. A noter également le cas des percussions virtuelles de Roland Auzet, développées à l'Ircam pour le spectacle Schlag!. Celles-ci représentent un exemple d'instrument (doublement) virtuel, puisque l'instrumentiste n'interagit avec aucune interface tangible, ses gestes étant captés par vidéo [Flety 2003].

Ces exemples relèvent principalement de ce que nous appelons un geste instrumental, dans le sens où la musique est sous "contrôle" du performer, les mouvements étant intimement liés au design de l'instrument ou de l'interface (même si celle-ci est virtuelle).

La catégorie des systèmes interactifs danse/musique se différencie clairement sur ces derniers points, comme nous allons le voir. Nous introduisons tout d'abord un exemple historique.

« Variations V »<sup>3</sup> (1965) de Merce Cunningham et John Cage est l'une des premières pièces où

<sup>3</sup> [www.artmuseum.net/w2vr/archives/Kluever/03\\_Variation.html](http://www.artmuseum.net/w2vr/archives/Kluever/03_Variation.html)

l'interaction entre la danse et la musique se base sur un dispositif de captation du mouvement en temps réel. Pour cette pièce, l'ingénieur Billy Klüver réalisa un dispositif permettant aux danseurs d'actionner des sons en traversant des faisceaux lumineux. Il est intéressant de noter que ce même concept se retrouve dans d'autres réalisations relativement récentes, comme la pièce "Electro Clips" de Christian Moeller<sup>4</sup> présentée à Ars Electronica en 1994, ou encore le système LaserWeb utilisé pour certaines pièces par la compagnie Troika Ranch<sup>5</sup>. Par ailleurs, les systèmes récents utilisant la captation vidéo peuvent être vus comme une extension de ce type de technologie.

Si John Cage et Merce Cunningham ont utilisé un "lien technologique" entre mouvement dansé et musique, leur but n'était sans doute pas de synchroniser son et mouvement de façon déterministe, mais plutôt de favoriser des "chance meeting" entre les aspects sonores et visuels. Il est également à noter que cette pièce incorporait des projections d'images de télévision manipulées par Nam June Paik. Cette pièce inaugure donc bien, sur plusieurs aspects, des travaux réalisés ces dix dernières années. Néanmoins, la technologie utilisée dans "Variation V" est essentiellement analogique. Les "nouvelles technologies", par l'utilisation de représentation numérique des images et des sons, offrent par contre des possibilités nouvelles d'analyse, de manipulation algorithmique et de transfert entre données gestuelles, visuelles et sonores [Manovich 2001].

Depuis le début des années quatre-vingt-dix, plusieurs compagnies comme Troika Ranch et Palindrome<sup>6</sup> ont développé à la fois des outils informatiques et une pratique artistique constante dans l'utilisation de capteurs attachés au danseur pour le contrôle interactif de processus sonores et visuels. Des centres de recherche (par exemple (liste non exhaustive) l'Ircam<sup>7</sup>, InfoMus Lab – Genova<sup>8</sup>, DIEM - Aarhus<sup>9</sup>, Ame-Arizona State University<sup>10</sup>) développent également des systèmes de captation, et des productions artistiques sont mises en place au sein de ces structures.

<sup>4</sup> [www.christian-moeller.com/](http://www.christian-moeller.com/)

<sup>5</sup> [www.troikaranch.org](http://www.troikaranch.org)

<sup>6</sup> [www.palindrome.de](http://www.palindrome.de)

<sup>7</sup> [www.ircam.fr](http://www.ircam.fr)

<sup>8</sup> [www.infomus.dist.unige.it](http://www.infomus.dist.unige.it)

<sup>9</sup> [www.musik-kons.dk/diem](http://www.musik-kons.dk/diem)

<sup>10</sup> [ame.asu.edu](http://ame.asu.edu)

Contrairement au "geste instrumental" dont la démarche s'inscrit dans une pratique historique d'invention de nouveaux instruments, les démarches basées sur le "geste dansé" posent de nouvelles questions: quel est le rôle du danseur par rapport à la musique: devient-il un interprète ? quel degré de perception doit avoir le public du processus technologique mis en place ?

Ces questions sont fortement liées aux considérations esthétiques de chaque compositeur et chorégraphe, et les réponses ne sont donc pas généralisables. Néanmoins, il est à noter que souvent, la fonction du danseur ou du performer n'est pas de contrôler l'environnement sonore de manière déterministe, mais d'être dans un véritable paradigme d'*inter-action*: le danseur ne devient pas musicien, mais interagit avec des processus musicaux pré-écrits. On peut se référer par exemple à la note de Butch Rován exprimant clairement ce type d'approche<sup>11</sup>.

Si ce genre d'oeuvre demande donc un travail de composition particulier, ce processus influence en général également le travail chorégraphique [Siegel 1998]. Néanmoins, il nous semble que l'état de l'art de la technologie est encore très limitant à ce niveau. En effet, les paramètres capturés sont souvent soit liés au positionnement de quelques articulations, soit relatifs à une quantité moyenne du mouvement, et sont donc difficiles à mettre en résonances avec un langage chorégraphique. Ainsi, les limites des systèmes de captation sont souvent source de frustration pour les danseurs/chorégraphes.

Ce point est très important, et constitue de fait une des motivations de notre travail de recherche à l'Ircam. Nous avons donc comme but de développer des systèmes adaptés à la performance scénique, permettant une interaction sur des paramètres de haut niveau, pouvant idéalement être mis en correspondance avec un langage chorégraphique. Plusieurs centres de recherche, dont l'Ircam, ont développé des programmes de

---

<sup>11</sup> Butch Rován, Program note for the piece *seine hohle Form* (2000). Commissioned by Palindrome Intermedia Performance Group.

"In the end, the goal in writing this piece was to create a work in which the music was not merely tied to the gestures of the dancers—which can be the challenge when working with interactive music/dance systems—but instead, create a work where the musical and dance components exist in a dynamic dialectic that allows for polyphony of sound and gesture."  
from <http://www.soundidea.org/rovan/notes.htm>.

recherche dans cette direction (voir par exemple EyesWeb-Infomus ou motion<sup>c</sup>-ASU)

## Outils d'écriture

Nous faisons référence ici à des systèmes utilisés lors de répétitions et/ou lors de l'écriture. Le système de captation du mouvement est un outil de travail qui n'est pas forcément utilisé lors de la performance.

Le spectacle de danse "Rew"<sup>12</sup> d'Hervé Robbe, dont la musique a été composée par Andrea Cera représente un exemple très intéressant de ce type d'approche. En effet, certains éléments musicaux ont été travaillés à partir de matériaux sonores générées par un système d'analyse du mouvement vidéo (Eyesweb) [Flety 2003], utilisé en temps différé. Les performances du spectacle n'incluaient pas de système de captation.

Il faut noter que dans le cadre de l'élaboration d'une pièce, un retour sonore en temps réel peut néanmoins être mis à profit pour permettre de favoriser un travail particulier du mouvement dans sa relation à la musique, même si le dispositif est absent lors de la performance. Ce type de travail semble encore relativement peu répandu. Néanmoins, cette approche semble intéressante si elle est couplée avec une analyse du mouvement automatique pertinente.

## Installation interactive

Nous mentionnons finalement les installations interactives, utilisant les mouvements du visiteur comme moteur de l'interaction. Sans vouloir couvrir ici ce sujet fort vaste, il est important de noter que de nombreux systèmes interactifs actuellement utilisés pour la danse dérivent de développements réalisés pour des installations. Les outils d'analyses vidéo VNS de David Rokeby<sup>13</sup>, en sont un exemple probant.

Néanmoins, nous voulons insister sur le fait que pour ce type d'installation, la problématique artistique du rapport geste/son est fondamentalement différente que pour une performance où le spectateur est passif. En effet, dans une installation, le visiteur perçoit l'interaction par ses propres mouvements, ou par

---

<sup>12</sup> Cera, A., Robbe, H. "Rew." – Music and Dance performance for 2 dancers and electronics. Premiered in Lisbon, June 2003.

<sup>13</sup> [homepage.mac.com/davidrokeby/home.html](http://homepage.mac.com/davidrokeby/home.html)

ses propres déplacements dans l'espace. Par contre, ce niveau perceptif est absent pour le spectateur d'une performance danse/musique interactive. Ainsi, la même relation entre geste et son peut s'avérer intéressante dans une installation mais de peu d'intérêt lors d'une performance de danse.

## 3 Captation et Analyse du mouvement

### 3.1 Systèmes de captation

Les technologies de captation ont très fortement évolué ces dernières années. En particulier, des systèmes très performants pour l'animation 3D (généralement référencé sous le nom anglais de *motion capture*) sont désormais disponibles. Ces systèmes sont très intéressants pour l'étude de mouvements [Bevilacqua 2002], mais restent très coûteux et difficilement compatibles avec des performances scéniques.

Nous développons donc des systèmes complémentaires, spécialement conçus pour la performance. Un système idéal serait compatible avec une plate-forme temps réel, facile d'utilisation et robuste. De plus nous essayons dans la mesure du possible de proposer des solutions modulaires, basées sur une architecture permettant de suivre dans une certaine mesure l'évolution technologique, et en conséquence d'assurer à la fois une cohérence entre les diverses applications (geste instrumental, geste dansé) et une certaine pérennité.

En tenant compte de ces critères, nous avons fait un certain nombre de choix. Pour plus de détails, nous référons le lecteur à [Fléty 2002, 2003]:

- architecture modulaire pour l'acquisition de données, compatible avec des modules compacts de communication sans fil [Fléty 2004].

- utilisation du réseau Ethernet et du protocole de communication OpenSoundControl (OSC) pour la transmission de données [Freed 1997]. Ceci implique donc un abandon (progressif) de la norme MIDI trop limitante pour des données gestuelles. Le protocole OSC permet en plus de communiquer entre un nombre grandissant de logiciels sous PC ou Mac, tels que Max/Msp<sup>14</sup>, EyesWeb<sup>15</sup> ou Director.

---

<sup>14</sup> www.cycling74.com

<sup>15</sup> www.eyesweb.org

En ce qui concerne la captation du mouvement dansé, nous poursuivons en parallèle deux approches complémentaires:

- Captation vidéo, basée sur la plateforme EyesWeb et Jitter.

- Utilisation de capteurs tels qu'accéléromètre et senseur de flexion placés sur le performer.

Ces deux systèmes permettent de mesurer des caractéristiques différentes du mouvement. La combinaison de ces systèmes permet donc d'obtenir un ensemble relativement complet et robuste de paramètres gestuels.

### 3.2 Analyse et reconnaissance du mouvement

Comme nous l'avons déjà dit, nous pensons qu'un effort de recherche important est nécessaire afin de dériver des paramètres interprétables du point de vue d'un langage chorégraphique. Il s'agit bien évidemment d'un but très ambitieux, qui nécessite une approche interdisciplinaire entre les arts et les sciences pour allier aux approches artistiques les progrès importants réalisés dans les domaines de la compréhension du mouvement humain (physiologie, sciences cognitives, médecine) et des techniques de traitement du signal et de reconnaissance automatique.

L'analyse du mouvement est réalisée par transformations successives des données transmises par le système de captation. Il est important de classifier les différents types de paramètres. Pour cela, nous proposons une hiérarchisation des paramètres du mouvement en plusieurs niveaux, de manière similaire à Camurri et al. [Camurri 2001]:

*Hiérarchisation des paramètres issus de la captation*

#### 1) Paramètres bruts de captation (bas niveau)

Les paramètres sont directement issus des capteurs ou du système vidéo et correspondent par exemple à la position ou à l'accélération d'un point particulier du performer. Même si ces paramètres peuvent sembler a priori intuitifs, ils se révèlent en pratique difficiles à contrôler ou obligent le performer à une représentation du geste très réducteur ou non adapté.

#### 2) Paramètres dérivés directs (niveau moyen)

Les paramètres bruts permettent de dériver, par des

opérations mathématiques, des paramètres qui ont une interprétation plus directe de la posture ou du mouvement. On peut par exemple calculer des valeurs de quantité de mouvement, centre de masse, ou autres valeurs statistiques de la posture.

Un grand nombre de paramètres peuvent être obtenus, accroissant la complexité de la représentation numérique du geste. De plus, ces paramètres ne sont en général pas indépendants, c'est-à-dire qu'ils contiennent une redondance d'information sur un geste donné. Il est donc judicieux de réduire le nombre de ces paramètres à un ensemble limité de paramètres indépendants, chacun décrivant des aspects distincts du mouvement (par exemple vitesse du centre de masse du danseur et symétrie de la posture).

Dans ce but, nous explorons actuellement plusieurs méthodes mathématiques qui permettent une telle réduction de paramètres (par exemple, analyse en composantes principales [Bevilacqua 2002] ou composantes indépendantes). Sans entrer dans les détails, il est important de noter que ces algorithmes doivent procéder à partir de données représentatives de l'ensemble des mouvements qui sont effectués. Cette approche est donc intrinsèquement liée à un contexte donné, d'une ou d'un ensemble de chorégraphies.

La physiologie et la biomécanique permettent également d'appréhender le problème des corrélations entre paramètres. Par exemple, Viviani a clairement montré que, dans certains contextes, la vitesse du bras peut être corrélée au rayon de courbure de la trajectoire [Viviani 1994]. Nous évaluons aussi l'utilisation de ce type de contrainte pour la réduction du nombre de paramètres.

### 3) Paramètres descriptifs de haut niveau

Par paramètres de haut niveau, nous faisons référence aux paramètres qui peuvent être mis en relation avec des paramètres perceptifs ou liés à des "langages" gestuels ou chorégraphiques.

Un des axes de recherche que nous poursuivons est de définir ces paramètres de haut niveau par rapport à un contexte donné, en utilisant des méthodes d'apprentissage automatique à partir de classes de mouvements choisis par un opérateur (par exemple le danseur/chorégraphe). Après cette étape, un algorithme de reconnaissance automatique peut être mis en place.

Nous ne cherchons donc pas, pour l'instant, à

trouver une formulation générale du mouvement. Néanmoins, notre but est de développer une méthodologie générale et une architecture souple permettant de formaliser et reconnaître des mouvements dans des contextes chorégraphiques donnés.

Pour cela, nous nous orientons vers des structures basées sur les Modèles de Markov Cachés [Rabiner 1989]. Ces structures modélisent le mouvement comme une succession d'état. La détermination de ces états ainsi que leur enchaînement s'opèrent par des modèles probabilistes. Cette approche convient généralement bien à des signaux temporels. Elle est très utilisée pour la reconnaissance de la voix et s'est déjà avérée intéressante pour les mouvements.

Les limites de ce modèle sont de devoir segmenter le mouvement en une suite d'état discrets. Néanmoins, la notion d'état est ici très générale, n'étant définie que par un ensemble de paramètres abstraits, choisis par l'utilisateur. En particulier, un état peut avoir une certaine durée (symbolisé par une lettre dans la figure 1), et ne correspond donc pas forcément à une posture fixe. De plus, certains travaux ont démontré la possibilité d'incorporer des notions de "styles" (définis comme des variations possibles d'un geste donné) sous forme d'un paramètre dans les Modèles de Markov Cachés, [Brand 2000].

## 3.3 Mapping entre geste et son

L'étude des diverses stratégies de "mapping" entre geste et son a récemment fait l'objet d'un intérêt accru [Wanderley 2001]. Le terme "mapping" couvre habituellement à la fois une partie de programmation et une partie artistique relative aux différents choix possibles pour une composition. La partie technique du mapping correspond en fait directement à l'analyse du mouvement que nous avons décrit<sup>16</sup>.

D'une manière générale, le mapping entre son et mouvement est souvent décrit comme une transformation d'un flux continu de données gestuelles en un flux de données sonores. Notre

---

<sup>16</sup> Wanderley avait proposé une stratégie générale de mapping, en introduisant un niveau intermédiaire entre données gestuelles et sonores [Wanderley 2001]. Cette formulation est en fait similaire à la structure que nous proposons. Le niveau intermédiaire proposé par Wanderley correspond à la dérivation des paramètres de moyens et haut niveau que nous décrivons.

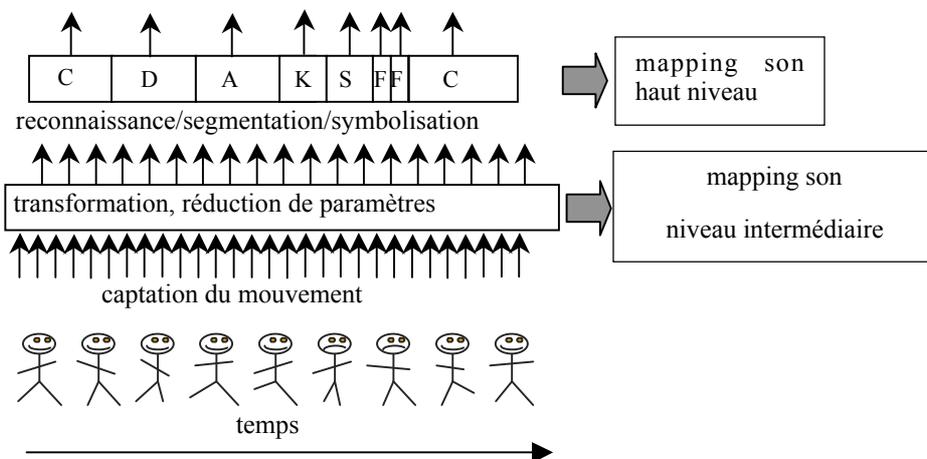


Figure 1. Analyse du mouvement et relation avec des processus sonores.

analyse du mouvement, générant des ensembles hiérarchisés de paramètres, implique implicitement une structure de mapping plus complexe. En effet comme cela est schématisé dans la figure 1, les différents niveaux de paramètres sont généralement relatifs à des échelles temporelles distinctes. Notre approche permet donc d'incorporer à la fois du contrôle continu de paramètres sonores et de générer des événements discrets dans le temps.

## 4 Conclusions et perspectives

Nous avons présenté divers éléments de notre recherche sur la captation et l'analyse du geste. Celle-ci s'inscrit dans une continuité de recherche et de développement d'outils pour l'interaction entre danse et musique. Cependant, les approches que nous avons décrites ne s'appliquent pas uniquement au mouvement dansé, mais s'inscrivent dans une stratégie globale, applicable également au geste instrumental.

## Remerciements

Nous remercions Norbert Schnell et Alexia Bevilacqua pour leurs commentaires et relecture du manuscrit.

## Références

[Camurri 2000], A. Camurri, P. Coletta, "Eyesweb: A real-time platform for interactive dance and music systems" *Proceedings of ICMC*, International Computer Music Conference 2000 – Berlin –Germany.

[Camurri 2001], A. Camurri, G. De Poli, M. Leman, G. Volpe, "A Multi-layered Conceptual Framework for Expressive Gesture Applications", *Proc. Intl MOSART Workshop*, Barcelona, November 2001.

[Bevilacqua 2002] F. Bevilacqua, J. Ridenour, and David J. Cuccia, "3D motion capture data: motion analysis and mapping to music", *Proceeding of the Workshop/Symposium on Sensing and Input for Media-centric Systems*, Santa Barbara CA, 2002.

[Brand 2000], M. Brand and A. Hertzmann. Style machines. *Proc. of ACM SIGGRAPH*, 183-192, 2000

[Fléty 2002], E. Fléty, "AtoMIC Pro: a Multiple Sensor Acquisition Device." *Proceedings of NIME*, New Interfaces for Musical Expression Conference 2002– Dublin, MIT Europe – Ireland.

[Fléty 2003], E. Fléty, M. Siguy, "EoBody : a follow-up to AtoMIC Pro's technology" *Proceedings of NIME*, New Interfaces for Musical Expression Conference 2003 – Montreal, McGill University – Canada.

[Fléty 2003], E. Fléty, "Collaboration Danse-Musique", *Rapport Interne Ircam*, 2003.

[Fléty 2003], E. Fléty, "Documentation Schlag!", *Rapport Interne Ircam*, 2003. Voir également [agora.ircam.fr](http://agora.ircam.fr)

[Fléty 2004], E. Fléty, N. Leroy, J.-C. Ravarini, F. Bevilacqua, "Versatile sensor acquisition system utilizing Network Technology", *submitted for*

publication, 2004.

[Freed 1997], A. Freed, M. Wright, M. “Open SoundControl: A New Protocol for Communicating with Sound Synthesizers” *Proc. of the 1997 International Computer Music Conference*, San Francisco, CA.

[Manovich 2001], L Manovich, *The Language of New Media*, MIT press, 2001.

[Rabiner 1989], R. Rabiner, "A tutorial on hidden Markov models" *Proceedings of the IEEE*, vol. 77, pp. 257-286, 1989.

[Siegel 1998], W. Siegel, "The Challenge of Interactive Dance – an overview and case study", *Computer Music Journal*, vol 22, 4, 1998.

[Wanderley 2000], M. M. Wanderley and M. Battier, editors. *Trends in Gestural Control of Music*, Ircam - Centre Pompidou, 2000.

[Wanderley 2001] Wanderley, M., “Performer-Instrument interaction: applications to gestural control of sound synthesis”. *Thèse de doctorat*, Université Paris 6 – France, 2001