

---

# Interaction musicale numérique

## Recréer les signaux du contrôle et de la perception

Florent Berthaut<sup>1</sup>, Myriam Desainte-Catherine<sup>2</sup>

1. Univ. Lille, UMR 9189 - CRISTAL - Centre de Recherche en Informatique Signal et Automatique de Lille, F-59000 Lille, France  
CNRS, UMR 9189, F-59000 Lille, France  
Centrale Lille, F-59000 Lille, France  
florent@hitmuri.net

2. Université de Bordeaux, LaBRI, UMR5800, F-33400 Talence, France  
CNRS, LaBRI, UMR5800, F-33400 Talence, France  
Inria, F-33400, Talence, France  
myriam@labri.fr

---

**RÉSUMÉ.** *Les instruments de musique numériques, de par leur diversité et complexité, ouvrent de nouvelles possibilités pour les artistes. Cependant, de nombreuses questions se posent quant à leur expressivité, leur facilité d'apprentissage, la limitation des retours vers le musicien ou encore leur perception par les spectateurs. Dans cet article, nous présentons le domaine de recherche de l'interaction musicale, et les principales pistes qui y sont explorées. Nous montrons comment ces travaux consistent avant tout à reconstruire les signaux du contrôle et de la perception.*

**ABSTRACT.** *Thanks to their diversity and complexity, Digital Musical Instruments open new possibilities for artists. However, they also raise many questions concerning their expressiveness, their learning curve, the limitation of feedback for musicians or their perception by the audience. In this article, we give an overview of the musical interaction research field, and describe the main research directions that it involves. We demonstrate that they amount to rebuilding the signals of control and perception.*

**MOTS-CLÉS :** *interaction musicale, instruments de musique numériques, interfaces, gestes, processus sonores, perception.*

**KEYWORDS:** *musical interaction, new interfaces for musical expression, digital musical instruments, gestures, sound processes, interfaces.*

---

DOI:10.3166/TS.32.345-363 © 2015 Lavoisier

### **Extended abstract**

Until the beginning of the 20th century, musical interaction was entirely acoustic. Musical sounds were produced by mechanical vibrations of parts of an instrument, in turn created by the energy of a musician's gesture. The development of electrical then digital musical instruments dramatically altered this principle. Thanks to their diversity and complexity, Digital Musical Instruments (DMI) open new possibilities for artists. However, they also raise many questions concerning their expressiveness, their learning curve, the limitation of feedback for musicians or their perception by the audience.

In this article, we give an overview of the research on New Interfaces for Musical Expression (NIME), and describe the main research questions that it involves. We demonstrate that these questions amount to rebuilding the signals of control and perception within DMIs.

In the introduction, we begin by providing a definition and model of DMIs. We then provide an overview of research which focus on the analysis and classification of DMIs, with the purpose of informing their design, evaluation and preservation. We also cover research on Digital Orchestras.

In section "Captation du contrôle", we present research on the various input modalities that are used as control signals in DMIs. In particular we describe classifications and sensing technologies for musical gestures, but also the use of sound, graphical interfaces and physiological sensors.

In section "Du contrôle aux processus musicaux" we focus on the essential concept of mappings, i.e. connections between control signals from the musician and sounds parameters that define the resulting music. We first describe the categories of mappings and their structures. Second, we discuss the control of musical structures and time. In particular, we present research on interactive scores, which allow for combining composition and real-time interaction.

In section "Signaux de la perception", we look at signals that are sent from the instrument back to the musician, to other musicians in an orchestra and to the audience. We demonstrate why these signals, which are implicit in acoustic instruments, need to be rebuilt in DMIs. In particular, haptic and visual feedback help the musician gain expertise with an instrument. They are essential for the handling of DMIs with multiple sound generating processes. We finally show how the audience and other musicians can be confused by the diversity and complexity of instruments with which they are not familiar, thus diminishing the perceived liveness of performances and preventing musical collaboration of the same level as in acoustic orchestras.

We conclude by providing a list of research questions that we believe are most important for the future of research in musical interaction.

## 1. Introduction

L'interaction musicale fut jusqu'au début du 20<sup>e</sup> siècle uniquement acoustique, le son étant produit par des vibrations mécaniques générées suite à un transfert d'énergie depuis les gestes des musiciens. L'apparition d'instruments électriques puis numériques a permis de diversifier les méthodes de production et de contrôle du son musical. La création de nouvelles interfaces, de nouveaux types de synthèse sonores, la gestion de processus sonore multiples, ainsi que des outils de prototypage rapide et de programmation visuelle, permettent maintenant aux musiciens de développer leurs propres instruments de musique, potentiellement uniques. Des orchestres d'instruments numériques se constituent également, ouvrant de nouvelles possibilités de collaboration musicale.

Ces instruments posent cependant un grand nombre de défis par rapport aux instruments acoustiques : comment développer des instruments expressifs et permettant la virtuosité, tout en facilitant leur apprentissage, comment améliorer leur compréhension par les spectateurs et les musiciens au sein d'un orchestre. Ces défis ont en commun qu'ils consistent à *recréer des signaux entre l'humain et l'instrument*, transmis physiquement dans le cas des instruments acoustiques. Ces problèmes sont au coeur de la recherche en interaction musicale.

Ce domaine de recherche est fortement pluridisciplinaire et situé à l'intersection de nombreux domaines, tels que les interfaces homme-machine, l'informatique musicale, les sciences cognitives ou encore le traitement du signal, auxquels il apporte de nouveaux contextes d'applications, de nouvelles questions mais également de nouvelles réponses. Cette diversité s'exprime notamment lors de la conférence NIME, à l'origine un atelier de la conférence ACM *Computer-Human Interaction* (CHI), et où sont depuis 2001 présentés chaque année un grand nombre de travaux de recherche.

Dans cet article, après avoir précisé la notion d'instrument de musique numérique, nous abordons les problématiques principales de ce domaine :

- La captation des contrôles du musicien : gestes, intentions, émotions.
- La reconstruction des signaux de contrôle entre les gestes et le son.
- La reconstruction des signaux de perception de l'instrument par les spectateurs et autres musiciens.

## 2. Instruments de musique numériques

### 2.1. Définition et structure

Parmi les différentes dénominations possibles pour les instruments comportant des composants logiciels, celle d'*Instruments de Musique Numériques* (ou Digital Musical Instruments / *DMI* en anglais) est maintenant celle retenue par la plupart des chercheurs. Leur structure est présentée sur la figure 1. Les DMI sont ainsi composés d'un premier segment permettant l'acquisition des gestes ou autres données provenant du

musicien et leur transformation en signaux de contrôle. Ce segment peut être appelé interface, dispositif d'interaction ou encore module de contrôle. Le deuxième segment associe les signaux de contrôle aux paramètres des processus sonores. Ces associations sont le plus souvent évoquées sous le terme *mappings*. Les *processus sonores* constituent un troisième segment, celui qui produit le signal sonore. Si plusieurs processus indépendants ou connectés participent à la production du son, on parlera d'instrument multi-processus. Finalement, depuis chaque segment de l'instrument, des signaux reviennent vers le musicien et/ou les spectateurs: ce sont les *retours sonores, visuels ou haptiques*. Dans les sections suivantes, nous détaillons les segments et signaux sur lesquels se concentrent les recherches en interaction musicale.

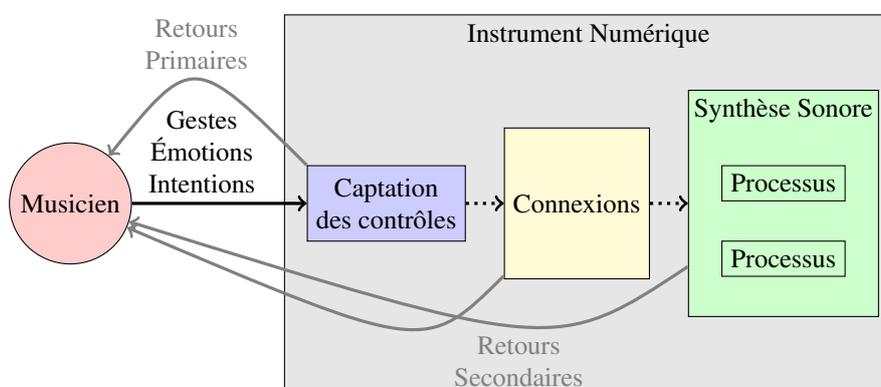


Figure 1. Structure d'un instrument de musique numérique

## 2.2. Analyse et évaluation

Au-delà de leurs structures, de nombreux travaux portent sur l'analyse et la classification des propriétés des instruments de musique numériques, avec pour objectifs leur conception, analyse et évaluation. Jordà Jordà (2003) définit ainsi une mesure de l'*efficacité* d'un instrument. Wanderley *et al.* (2000) proposent une classification selon deux dimensions. Birnbaum *et al.* (2005) définissent un espace de dimensions pour l'analyse des instruments.

On retrouve plusieurs axes communs à ces différentes classifications. Le premier de ces axes est la *complexité* en entrée telle que définie par (Jordà, 2003). Il représente la quantité et la complexité des gestes que l'utilisateur doit effectuer pour jouer de l'instrument.

Le deuxième axe est le *degré de liberté* (Jordà, 2003). Cet axe représente la palette expressive que l'instrument fournit au musicien. Il permet de distinguer des installations interactives dans lesquelles l'utilisateur n'a que peu de contrôle sur le son, des instruments à proprement parler. Il rejoint en cela la classification proposée par Pressing (1997) prenant en compte le rôle du son dans une application : artistique, environnementale ou informationnelle.

Le troisième axe est la *complexité en sortie*, c'est-à-dire les possibilités de résultat musical. Il est possible d'évaluer la diversité de sons que l'on peut obtenir avec l'instrument (hauteurs, amplitudes, timbres), ainsi que le nombre de processus (mélodies, rythmes) simultanés.

Le quatrième axe est la *courbe d'apprentissage*. Ainsi, une utilisation virtuose, c'est-à-dire exploitant toutes les possibilités de l'instrument (Dobrian, Koppelman, 2006), doit être possible après un temps de pratique raisonnable, mais l'instrument doit rester accessible aux utilisateurs novices (leur permettre d'avoir un impact, même très simple, sur le son) (Wessel, Wright, 2002) (Franco *et al.*, 2004).

Finalement, un axe important concerne les *retours disponibles*, qu'ils soient haptiques, visuels, auditifs, et est notamment utilisé dans (Birnbaum *et al.*, 2005). Ces retours sont essentiels car ils forment avec les gestes la boucle d'interaction qui va permettre au musicien de s'adapter à l'instrument.

Les instruments numériques étant de plus en plus souvent intégrés au sein d'orchestres numériques tels que les *laptop orchestras* (Wang *et al.*, 2009) ou *mobile phone orchestras* (Oh *et al.*, 2010), des classifications et outils d'analyse ont également été proposés qui prennent en compte les interactions, augmentées numériquement ou non, entre musiciens. Le *dimension space* de Birnbaum *et al.* (2005) comprend ainsi les dimensions de *distribution dans l'espace* et de *nombre d'utilisateurs simultanés*. Jordà (2005), Blaine, Fels (2003), Weinberg (2005) et Hattwick, Wanderley (2012) proposent des axes de description des orchestres numériques similaires, parmi lesquels l'interaction entre les musiciens, la flexibilité de leurs rôles, ou encore l'organisation temporelle et spatiale. Finalement, Berthaut, Dahl (2015) proposent une classification des modes de collaboration au sein des orchestres numériques, qu'ils divisent en modes de *coopération*, *communication* et *organisation*. Ils proposent également un ensemble de composants permettant de modéliser les orchestres numériques.

De nombreux travaux insistent sur la différence dans l'analyse des instruments de musique numériques par rapport aux interfaces homme-machine, notamment autour des notions d'*expressivité* (Dobrian, Koppelman, 2006) ou de *finalité de l'interaction*, cette dernière n'étant pas le résultat obtenu après une série d'opérations mais plutôt l'enchaînement des opérations elles-mêmes (Hunt, Kirk, 2000). Les outils de l'interaction homme-machine peuvent néanmoins être utilisés, notamment pour l'évaluation des instruments dans des tâches spécifiquement musicales (Wanderley, Orio, 2002).

Finalement, un autre objectif de l'analyse des instruments numériques est leur préservation et leur transmission. En effet, les matériels et logiciels évoluant continuellement, il peut arriver que certains instruments ou pièces ne puissent plus être joués après mise à jour de leurs composants. Bonardi et Barthélemy examinent la possibilité de retranscrire les structures et fonctionnement des instruments sous forme mathématique et fonctionnelle afin de pouvoir les réimplémenter sous différentes architectures (Bonardi, Barthélemy, 2008).

### 3. Captation du contrôle

Le premier segment d'un instrument de musique numérique comprend la captation des actions ou contrôles du musicien. De nombreuses modalités de contrôle sont utilisées, de la plus directe (potentiomètres ou boutons) à la plus indirecte (électroencéphalogramme).

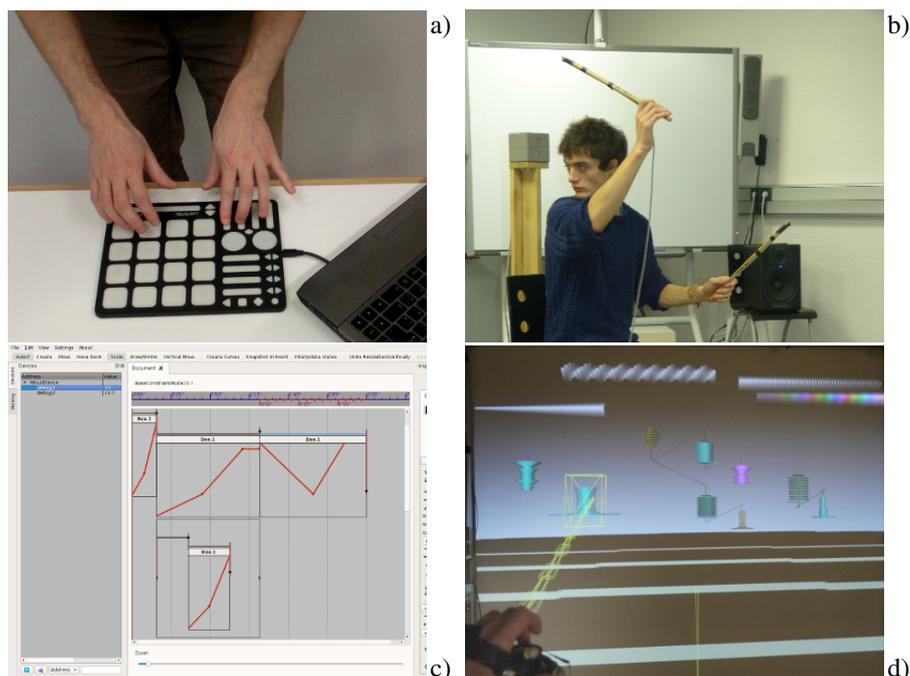


Figure 2. Différentes modalités de contrôle: a) Contrôleur MIDI Quneo avec capteurs de pression et position, b) Percussion aérienne avec suivi de position 3D, c) Interface graphique du logiciel de partitions interactives I-Score, d) Instrument 3D Immersif Drile

#### 3.1. Les gestes

La première modalité qui vient à l'esprit lorsqu'on se représente un instrument est celle du geste. Celui-ci peut être capté à l'aide de composants (de force, de déplacement linéaire ou angulaire, de vitesse, d'accélération) couplés ou actionnés par le musicien, ou extrait à distance par exemple à partir d'un signal vidéo. On peut définir le geste comme une action volontaire effectuée grâce à une partie du corps. On y inclura donc le souffle, le suivi de regard...

Si de nombreuses définitions et classifications du geste musical existent (Jensenius *et al.*, 2009), celle de Cadoz (1999) est intéressante en ce qu'elle offre trois catégories applicables aux instruments de musique numériques. Les gestes d'*excitation* sont

ainsi les gestes qui génèrent le son, par exemple gratter une corde sur une guitare. Les gestes de *modulation* vont permettre de modifier les caractéristiques du son produit, par exemple appuyer à différentes positions avec la main gauche sur la corde grattée afin d'en changer la hauteur. Finalement, les gestes de *sélection* permettent de choisir les éléments à manipuler sur l'instrument, par exemple les touches d'un piano. Il est à noter qu'un même geste du musicien peut impliquer plusieurs catégories, par exemple lors de modulation et d'excitation simultanées. Finalement les gestes ancillaires (Wanderley, Vines, 2006) accompagnent ces gestes de manipulation de l'instrument et sont essentiels à l'interprétation et à la transmission de l'émotion aux spectateurs (S. Dahl, Friberg, 2007).

Ces catégories peuvent être utilisées pour la conception d'instruments numériques, à la fois pour le choix des méthodes de captation, chaque type de geste ayant ses spécificités. Par exemple, les gestes d'excitation instantanée tels que les percussions requièrent un délai très court ( $< 20\text{ms}$ ) entre la sensation de contact avec l'instrument et le résultat sonore afin de permettre le jeu rythmique. Mais les catégories de gestes peuvent également informer les choix de *mappings* entre paramètres de contrôle et paramètres sonores, ainsi que l'explique la section suivante.

Une ressource importante pour la captation du geste est le *SensorWiki*<sup>1</sup> de l'équipe IDMIL à l'Université McGill, qui fournit une liste exhaustive de capteurs avec leurs caractéristiques ainsi que leurs utilisations dans le contexte des instruments de musique numériques.

### 3.2. *Le son comme contrôle*

De manière plus indirecte mais toujours volontaire, les instruments de musique numériques peuvent être contrôlés par un signal audio. Ce signal peut provenir de différentes sources, le plus souvent d'un instrument acoustique que l'on pourra alors dire *augmenté* (il peut l'être grâce à des capteurs gestuels également). Par exemple, des processus d'effets audionumériques appliqués au son de l'instrument acoustique peuvent être manipulés via les attributs même de ce son (spectre, hauteur, volume...) comme décrit dans (Verfaillie, 2003). Comme nous l'expliquerons plus en détails dans la section 4.2, le son d'un instrument peut également permettre de faire se dérouler des processus sonores, par exemple dans le cas du suivi de partition.

### 3.3. *Les interfaces graphiques*

Les interfaces graphiques constituent une autre modalité de contrôle. Bien qu'elles s'appuient souvent sur des capteurs gestuels afin d'obtenir la position de la main ou d'un stylet, elles constituent une strate supplémentaire entre le geste et le son, offrant un espace partagé au sein duquel les gestes et les sons sont représentés graphiquement.

---

1. <http://www.sensorwiki.org>

Ainsi, des outils graphiques permettent de manipuler la matière sonore visualisée, que ce soit un curseur et du texte dans le contexte du *live-coding* (Roberts *et al.*, 2013), ou une main virtuelle dans une interface 3D. Les interfaces vont du dessin musical (Levin, 2000), à la réalité virtuelle (Zappi *et al.*, 2011) (Berthaut *et al.*, 2010), en passant par des interfaces de composition 3D (Barri, 2009), ou encore des interfaces tangibles telles que la Reactable (Jordà *et al.*, 2005). Par rapport à des interfaces matérielles comprenant un ensemble de capteurs figés, les composants graphiques peuvent être recomposés, offrant ainsi une grande malléabilité des degrés de contrôle.

### **3.4. Les capteurs physiologiques**

Une modalité de plus en plus explorée est celle des mesures physiologiques. Contrairement aux autres modalités, elle peut ne pas être complètement volontaire, c'est-à-dire consciemment produite par les musiciens. Le but des mesures physiologiques est de permettre d'accéder en temps-réel aux émotions et intentions des musiciens. Ainsi, les électrocardiogrammes ou capteurs de conductance de la peau donnent des indications sur l'état d'excitation/stress ou relaxation du musicien. Les casques électroencéphalogramme (EEG) permettent à la fois de détecter les états émotionnels mais également d'autres signaux, par exemple les pulsations lumineuses perçues et l'intention de mouvements des mains/bras et jambes/pieds, fournissant ainsi des contrôles plus volontaires (Eaton, Eduardo, 2014).

## **4. Du contrôle aux processus musicaux**

Après captation, les contrôles venant du musicien doivent être connectés aux processus sonores, reconstituant ainsi le signal du contrôle. Les paramètres manipulés peuvent être de bas niveau (volume, fréquence de coupure d'un filtre...) ou de plus haut niveau (tempo, déroulement de séquences, routage du son entre effets audio-numériques...).

### **4.1. Du contrôle aux paramètres sonores**

La connexion entre contrôles et paramètres sonores est probablement la question la plus importante de l'interaction musicale. Cette connexion s'effectue le plus souvent en plusieurs étapes, comme décrit dans (Arfib *et al.*, 2002). La première de ces étapes est la traduction des données de contrôle, pour certaines d'entre elles relativement abstraites, en paramètres pertinents pour l'interaction musicale. Ainsi, au lieu de la position absolue de la main d'un musicien, sa vitesse ou encore les gestes de percussion effectués en l'air (détectés en calculant l'accélération) peuvent être préférés, par exemple de par la similarité entre l'énergie du geste et celle perçue dans le son. Les gestes peuvent également être analysés selon leurs qualités (Alaoui *et al.*, 2012), prenant inspiration des recherches en danse, ou selon les associations mentales avec des catégories ou paramètres musicaux (Godøy *et al.*, 2006).

La deuxième étape correspond à la connexion à proprement parler entre les paramètres du geste et ceux du son.

Selon Hunt, Kirk (2000), les *mappings* peuvent être de type *one-to-one*, *one-to-many* (divergents), *many-to-one* (convergenents), ou *many-to-many*, selon qu'ils connectent un ou plusieurs paramètres gestuels en entrée, à un ou plusieurs paramètres sonores en sortie. Hunt et Kirk montrent également que les *mappings* complexes (non *one-to-one*) sont plus intéressants et motivants, surtout sur le long terme. Ils favorisent l'exploration des possibilités sonores des instruments. De plus récents travaux ont exploré la possibilité de modifier ces *mappings* en temps réel durant une performance. L'instrument Drile (Berthaut *et al.*, 2010), via un modèle musical hiérarchique et une interface graphique 3D, donne ainsi accès à des *mappings* correspondant à différents niveaux de complexité en entrée et en sortie, s'adaptant ainsi au niveau d'expertise des musiciens. (Zappi, McPherson, 2014) présente un système dans lequel la modification des *mappings* fait partie du jeu musical, via une matrice matérielle donnant accès aux capteurs et aux paramètres sonores.

Les *mappings* peuvent ensuite être *explicites*, c'est-à-dire définis par des règles ou fonctions qui vont convertir et combiner les valeurs des paramètres gestuels et les assigner aux paramètres sonores. Ils peuvent également être *implicites*, via l'utilisation de modèles probabilistes ou d'apprentissage automatique (Caramiaux, Tanaka, 2013).

Sur le plan technique, plusieurs protocoles de communication permettent de connecter dispositifs d'interaction et processus sonores. Le protocole MIDI, standard publié par le consortium d'industriels MMA<sup>2</sup> en 1983, reste le protocole le plus utilisé, notamment par les interfaces commerciales. Cependant, ses limitations, notamment sa résolution faible (7 bits) et son format très lié au contrôle par clavier (127 notes avec niveaux de vélocité), ont motivé le développement de protocoles plus récents. Le plus répandu est l'OpenSoundControl, protocole s'appuyant sur un adressage de type url, des types de données très variés et à haute résolution (par exemple, nombres flottants 32bits ou blobs) (Wright, 2005). Il est maintenant géré par de nombreux logiciels musicaux, et certains constructeurs commencent également à l'intégrer à leurs équipements. L'absence d'un système standardisé pour la découverte des espaces de noms (par exemple quels paramètres sonores sont contrôlables pour un synthétiseur) freine néanmoins son adoption à plus grande échelle, malgré plusieurs tentatives telles que (Malloch *et al.*, 2013).

#### 4.2. Contrôle du temps et des structures

Pour définir l'interaction musicale, il faut introduire un troisième niveau s'ajoutant à ceux déjà mentionnés correspondant aux contrôles et à l'audio. Nous allons voir dans cette section comment ce niveau musical peut être décrit et quels sont les paramètres avec lesquels le musicien va pouvoir jouer. Enfin, les systèmes décrits ici se

---

2. MIDI Manufacturers Association.

différencient notablement des instruments introduits dans les autres sections, de par l'utilisation de partitions musicales, et font émerger un nouveau concept entre instrument et partition, que nous appelons *partition interactive*.

#### 4.2.1. *Partition musicale*

La structuration musicale la plus simple est celle de la partition. Prenons pour exemple le langage *Csound* (Boulanger, s. d.) qui a pour avantage de permettre de définir la synthèse sonore, les contrôles sonores ainsi que la partition. La partition est définie par un ensemble de "notes" qui sont datées et qui ont une durée par rapport à un tempo. Chaque "note" correspond à un appel à un "instrument", c'est-à-dire un processus audio paramétré par des contrôles. Ceci fournit un premier niveau de paramètres musicaux : les dates et les durées de chaque note, le tempo ainsi que le paramétrage du processus audio correspondant. En *Csound*, la partition est fixée lors de l'écriture et ne peut donc pas changer à l'exécution. Mais il existe des systèmes permettant des interactions avec cette partition.

#### 4.2.2. *Systèmes interagissant avec le temps*

Le *suiveur de partition* est un concept introduit et développé dans les années 80 (Vercoe, Puckette, 1985) (Dannenberg, 1984) (Cont, 2008). Il s'agit d'un système composé d'un musicien dont le son est capté par un micro et d'un ordinateur exécutant un logiciel de suivi de partition. Le logiciel connaît la partition jouée par le musicien et dispose de sa propre partition, dite numérique, qui est synchronisée avec celle du musicien. Le *suiveur de partition* effectue une correspondance entre ce que joue le musicien et la partition, et en déduit un tempo qui lui permet de jouer sa propre partition numérique comme s'il accompagnait le musicien. Ce genre d'interaction permet donc d'adapter le tempo du jeu de l'ordinateur à celui de l'humain et de les synchroniser sur des événements clefs.

Le *metapiano* (Haury, 2008) est un instrument composé d'un clavier restreint à une dizaine de touches connecté via le protocole MIDI à un ordinateur muni d'une partition incomplète. Celle-ci est composée de la suite des hauteurs des notes ainsi que des relations temporelles de synchronisation liant leurs débuts et leurs fins. Elle comporte aussi les rapports volumiques entre les différentes voix. La partition numérique ne comprend donc ni les volumes, ni les attaques, ni les durées qui devront par conséquent être fournis par le musicien au moment du jeu en live. Le musicien peut ainsi interpréter la pièce en jouant sur les variations agogiques, les articulations et la dynamique.

À l'instar du *metapiano*, le logiciel *i-score*<sup>3</sup> (Baltazar *et al.*, 2014) propose de jouer avec le temps d'une partition mais de façon plus générale. Alors que le *metapiano* est un instrument pour lequel on peut définir des partitions, *i-score* est un séquenceur permettant d'écrire sur une time-line des partitions dites *interactives* et de coordonner

---

3. [i-score.org](http://i-score.org)

divers types d'instruments multimédia, qu'ils soient musicaux, lumineux, graphiques, voire même pour des systèmes motorisés comme des robots par exemple. Dans *i-score*, une *partition interactive* est représentée par le matériau artistique qui est un ensemble de *notes* multimédia mises en forme dans le temps mais aussi par les interactions possibles qui s'effectueront en temps réel lors de la performance et qui donneront lieu à différentes exécutions. Une partition interactive est alors représentée par l'ensemble des solutions d'un problème donné sur des variables de dates et de durées au moyen de contraintes temporelles. Ces contraintes définissent les marges de liberté délimitant les domaines des données temps réel et donc l'ensemble des solutions que le musicien peut explorer tout en restant dans la cohérence musicale définie par le compositeur.

#### 4.2.3. *Partition incomplète et cohérence musicale*

Ces systèmes d'interaction avec le temps d'une pièce musicale ou plus précisément d'une partition introduisent un nouveau concept entre instrument et partition qui a été notamment pointé par des compositeurs (Philippe Manoury (1998)<sup>4</sup>, Brian Eno<sup>5</sup>) sous les termes de *partition virtuelle* ou de *unfinished pieces of music*. Ce dernier terme sied particulièrement bien à la modélisation informatique de ces systèmes. En effet, l'ordinateur dispose dans tous les cas d'une description incomplète de la pièce de musique.

Dans le cas du suiveur de partition, l'ordinateur connaît certes la partition théorique mais ne connaît pas l'interprétation que le musicien va en faire, notamment les variations de tempo. Il pourrait par contre connaître les consignes d'interprétation données au musicien ce qui pourrait l'aider à mieux anticiper son jeu.

Dans le cas du *metapiano*, le jeu du musicien est libre, ce qui lui donne la possibilité de déformer complètement la pièce en jouant un rythme éventuellement très différent de celui qui correspond aux hauteurs des notes introduites dans l'ordinateur. Cette possibilité du *metapiano* est une particularité dont jouent parfois volontairement les musiciens lors de leurs prestations, ce qui peut donner des résultats très éloignés de la pièce musicale originale.

Enfin, le séquenceur *i-score* dispose de la partition numérique, et de la cohérence musicale, donnée par les contraintes temporelles entre les "notes" de la partition. Celles-ci se présentent sous la forme de fenêtres temporelles pendant lesquelles l'événement (de début ou de fin de la note) est attendu. Ce mécanisme restreint les possibilités de jeu du musicien et donc de déformation de la pièce (Desainte-Catherine *et al.*, 2013).

Ces systèmes d'interaction avec le temps pourraient s'étendre à l'avenir pour donner l'accessibilité au jeu instrumental pour tous. La question de la cohérence musicale pourra se reposer alors avec un enjeu nouveau concernant la forme des retours à fournir au musicien novice. Pour l'aider à progresser dans le jeu de sa partition incom-

4. <http://etincelle.ircam.fr/733.html>

5. [http://archive.wired.com/wired/archive/3.05/eno\\_pr.html](http://archive.wired.com/wired/archive/3.05/eno_pr.html)

plète, comment lui signifier la marge de liberté qui est la sienne ou comment guider ses gestes pour améliorer la qualité de la pièce finale ?

### **4.3. Contrôle de la structure musicale**

Puisque les instruments de musique numériques, contrairement aux instruments acoustiques, permettent de contrôler plusieurs processus sonores simultanément, le jeu musical peut alors s'enrichir du contrôle de structures de production du son. On peut distinguer plusieurs niveaux de contrôle des structures musicales.

#### **4.3.1. Combinaison**

Le niveau le plus basique de contrôle est l'activation indépendante de processus dont les résultats musicaux sont combinés. C'est par exemple le cas du mode Session du logiciel Ableton Live<sup>6</sup>, utilisé par de nombreux musiciens électroniques afin de superposer des boucles musicales (rythmes, mélodies, échantillonnés ou synthétisés), à l'aide d'interfaces présentant une grille de boutons ou pads de percussion.

#### **4.3.2. Interconnexion**

A un niveau plus avancé, il est possible pour un musicien lors de la performance de contrôler comment les processus s'interconnectent, c'est-à-dire comment le résultat musical est construit par la succession de processus, par exemple de synthèse et traitement. Un exemple est la manipulation de synthétiseurs modulaires, dont une version modernisée est la Reactable (Jordà *et al.*, 2005).

#### **4.3.3. Construction**

Finalement, certains instruments s'appuient sur la construction de structures de processus sonores aux relations plus complexes. L'instrument Drile permet de construire des arbres de boucles musicales (Berthaut *et al.*, 2010), les manipulations effectuées sur un noeud se propageant à ses enfants. La libTuiles (Janin *et al.*, 2013) permet de créer des structures temporelles multi-échelles de manière dynamique, l'écoulement du temps dépendant de la composition de tuiles temporelles.

## **5. Signaux de la perception**

La section précédente analyse comment une partie importante des travaux en interaction musicale s'attache à reconstituer les signaux de contrôle, c'est-à-dire la transmission d'énergie entre le geste du musicien et l'instrument qui produit le son, qui sont perdus lors du passage du domaine physique au domaine numérique. De la même façon, la perception de l'instrument, que cela soit par le musicien ou par des observateurs d'une performance, est également altérée et les signaux correspondants doivent être recréés.

---

6. <https://www.ableton.com/>

### 5.1. Retours vers le musicien

Les signaux les plus importants sont ceux qui reviennent de l'instrument vers le musicien: les retours ou *feedback*. Ceux-ci peuvent être haptiques, sonores ou visuels. Vertegaal et Ungvary distinguent ainsi deux types de retours (Vertegaal *et al.*, 1996). Les *retours primaires* proviennent de la partie *captation* des instruments. Ce sera par exemple la sensation du contact avec un potentiomètre, ou le son de la frappe d'un pad de percussion. Les *retours secondaires* sont ceux émis lors du traitement numérique des contrôles ou de la production du son. Le son produit par l'instrument numérique entre dans cette seconde catégorie. Cependant, les instruments numériques doivent également reproduire les autres signaux de la perception vers les musiciens, à savoir les signaux haptiques et visuels.

#### 5.1.1. Retours haptiques

Les retours haptiques sont essentiels au jeu musical (O'Modhrain, 2001). Ils se répartissent en deux catégories. Les retours *passifs* ou tactiles, correspondant au sens du toucher, incluent par exemple les retours vibrotactiles qu'un instrument acoustique produirait naturellement. Les reconstruire améliore grandement l'expérience des utilisateurs selon (Marshall, Wanderley, 2006). Les retours *actifs* incluent par exemple les mécanismes de retour d'effort qui vont permettre de simuler la réponse de modèles sonores physiques tels que des percussions virtuelles (Berdahl *et al.*, 2005), facilitant ainsi la réutilisation de techniques de jeu existantes dans le cadre d'instruments numériques.

#### 5.1.2. Retours visuels

Comme nous l'avons précisé dans la section 3, les interfaces graphiques offrent une grande flexibilité dans la captation du contrôle musical. Elles permettent également de faciliter le contrôle de structures musicales et des instruments multi-processus. Par l'intermédiaire de dispositifs d'affichage plus ou moins complexes (des LEDs jusqu'à la vidéo-projection en passant par les écrans), elles informent le musicien de l'état des processus sonores ou encore de la configuration des connexions gestes-son.

Ces retours visuels nécessitent cependant de choisir des représentations appropriées pour les paramètres sonores et structures musicales. Les logiciels de mixage ou d'édition audio emploient des représentations statiques comme la forme d'onde ou le spectrogramme, ainsi que des représentations dynamiques telles que le vu-mètre, le spectre. Au-delà de ces visualisations permettant notamment d'avoir des mesures chiffrées des attributs sonores, plusieurs travaux ont exploré les préférences des utilisateurs pour la représentation de paramètres sonores perceptifs tels que la hauteur, le volume, la brillance à l'aide d'attributs graphiques (forme, couleur, taille, texture). Walker (1987) montre qu'elles sont ainsi globalement indépendantes du contexte culturel et souvent liés à des analogies physiques : (taille et volume) ou (taille et hauteur), (vitesse et volume), (couleur et brillance), (texture et timbre) ou (forme et timbre). L'éducation musicale peut également jouer un rôle, les musiciens habitués aux partitions associant naturellement position verticale et hauteur du son.

Les représentations de structures musicales les plus communes suivent une représentation horizontale du temps. Des partitions statiques sont munies d'un curseur dynamique donnant la position actuelle de lecture. Les piano-rolls proposent une subdivision verticale suivant les touches d'un clavier de piano et une subdivision horizontale suivant les mesures et temps. D'autres représentations existent, telles que les trackers (séquences défilant verticalement, une colonne par processus sonore), les représentations circulaires pour les séquences jouées en boucle, ou les représentations en graphes ou arbres pour les structures plus complexes, par exemple hiérarchiques. Dans tous ces cas, des composants visuels permettent de percevoir l'activité de chaque processus ou élément des structures musicales.

### 5.2. Perception par les spectateurs et l'orchestre

Les instruments de musique numériques, comme nous l'avons démontré dans les sections précédentes, offrent une grande liberté aux musiciens, leur permettant d'utiliser différentes modalités de contrôle ou encore différentes relations geste/son ou représentations du son. En contrepartie, cette diversité et cette complexité perturbent leur compréhension par les observateurs externes: les spectateurs et les autres musiciens au sein d'un orchestre. Ainsi avons-nous tous développé une certaine *familiarité* (Fyans *et al.*, 2010) avec les instruments acoustiques, qui provient à la fois de notre expérience personnelle des phénomènes physiques sur lesquels ces instruments s'appuient (l'amplitude du son dépend de la force appliquée, les objets plus larges vont donner un son plus grave...) et de notre contexte culturel. Ceci nous permet de prévoir et de comprendre le résultat des gestes d'un pianiste ou d'un batteur par exemple, sans pour autant percevoir le détail de ces gestes. Ceci nous permet également de percevoir l'engagement et l'émotion véhiculée lors des performances (S. Dahl, Friberg, 2007) (Vines *et al.*, 2011).

Dans le cas des instruments numériques, l'engagement est souvent perdu car les observateurs ne sont pas familiers avec les instruments et car ces derniers n'offrent pas suffisamment d'informations visuelles. Dans le contexte d'un orchestre, le problème correspond à celui de l'*awareness* (L. Dahl, 2012), c'est-à-dire la conscience des actions et des sons produits par les autres musiciens, sans laquelle le dialogue musical est impossible. Du point de vue des spectateurs, on parle de perte de perception de la *liveness*, que l'on pourrait traduire par *vitalité*, ou du besoin de *transparence* (Gurevich, Cavan Fyans, 2011) des interactions homme-instrument. Pour les spectateurs, les musiciens utilisant uniquement un ordinateur portable comme instrument pourraient très bien s'être contentés d'appuyer sur Lecture au début de leur performance puis être en train de consulter leurs courriels pendant le reste du temps ! *Il est donc primordial que les signaux de la perception des instruments soient reconstruits pour les observateurs externes de la même manière que pour les musiciens eux-mêmes.*

De plus en plus de travaux s'attachent à recréer cette compréhension des instruments numériques et de l'engagement des musiciens électroniques lors de performances. Il faut tout d'abord comprendre les paramètres perceptifs de la *liveness*.

Marshall *et al.* (2012) s'appuient sur les émotions perçues par les spectateurs comme une mesure possible et définissent plusieurs critères. Berthaut, Coyle *et al.* (2015) regardent eux plus particulièrement la perception de la causalité entre gestes et son dans les instruments numériques, en s'appuyant sur la théorie de l'*agency*. Ils montrent qu'il est possible de compenser la perte de causalité sur deux critères définis dans (Wegner, Wheatley, 1999) à l'aide de visualisations des mappings contrôle-son.

La deuxième étape consiste à appliquer ces résultats. Certains artistes et chercheurs tiennent ainsi compte de ces problématiques en modifiant leurs instruments ou en projetant une vue rapprochée des composants matériels ou des gestes. Dans le cas d'instruments graphiques se pose ainsi la question de la scénographie : comment placer les affichages afin d'assurer une perception correcte de la performance par les spectateurs (Berthaut *et al.*, 2014). Berthaut *et al.* (2013) proposent une approche différente. Le système Rouages permet de révéler les mécanismes des instruments numériques sans modifier l'interface pour les musiciens, en les augmentant à l'aide d'un dispositif de réalité augmentée. Ainsi, il amplifie les gestes grâce à des extensions virtuelles et affiche une structure simplifiée de l'instrument comprenant les principaux processus ainsi que les différents mappings entre capteurs et paramètres sonores. Le but est non pas de faire comprendre le fonctionnement complet de l'instrument, mais plutôt de révéler l'engagement du musicien électronique.

## 6. Conclusion

Tout au long de cet article, nous avons illustré ce en quoi les nombreuses possibilités musicales ouvertes par l'interaction musicale numérique entraînent des défis importants, qui ont en commun qu'ils consistent à recréer les signaux du contrôle et de la perception afin d'améliorer l'expérience des spectateurs et membres des orchestres. Les travaux de recherche en interaction musicale, notamment ceux présentés chaque année à la conférence NIME, dans laquelle les problématiques de captation, traitement et restitution des signaux d'interaction musicale occupent une place primordiale, constituent des réponses à ces défis, et proposent des outils à destination des musiciens, compositeurs et luthiers.

Parmi les perspectives de la recherche en interaction musicale, les pistes suivantes nous semblent les plus prometteuses : la question des interfaces purement gestuelles par rapport à l'objet physique ; la place de la composition et de l'écriture dans les nouveaux instruments (Fober *et al.*, 2013) ; les modèles physiques et les technologies associées permettant de lier retours haptiques et visuels au modèle sonore (Leonard *et al.*, 2014) ; les interfaces de réalité mixte intégrant le contenu virtuel aux performances et instruments physiques (Berthaut, Martinez *et al.*, 2015) ; les protocoles et outils pour orchestres numériques hétérogènes facilitant de nouveaux modes de collaboration ; la compréhension et l'amélioration de l'expérience des spectateurs.

### Bibliographie

- Alaoui S. F., Caramiaux B., Serrano M., Bevilacqua F. (2012). Movement qualities as interaction modality. In *Proceedings of the designing interactive systems conference*, p. 761–769.
- Arfib D., Couturier J. M., Kessous L., Verfaillie V. (2002). Strategies of mapping between gesture data and synthesis model parameters using perceptual spaces. *Organised Sound*, vol. 7, n° 2, p. 127–144.
- Baltazar P., Hogue T. de la, Desainte-Catherine M. (2014). i-score, an interactive sequencer for the intermedia arts. In *Proceedings of joined international computer music conf rence and sound and music computing (icmc-smc)*. Athena (Greece).
- Barri T. (2009, 18–21 May). Versum: audiovisual composing in 3d. Copenhagen, Denmark, Re:New – Digital Arts Forum. Consult  sur [Proceedings/2009/Barri2009.pdf](#)
- Berdahl E., Verplank B., Smith J. O., Niemeyer G. (2005). A physically-intuitive haptic drumstick. In *Proceedings of the 2007 international computer music conference*.
- Berthaut F., Coyle D., Moore J., Limerick H. (2015). Liveness through the lens of agency and causality. In *International conference on new interfaces for musical expression*.
- Berthaut F., Dahl L. (2015). Boeuf: A unified framework for designin and modeling digital orchestras. In *International symposium on computer music multi-disciplinary research (cmmr 15)*.
- Berthaut F., Desainte-Catherine M., Hachet M. (2010). Drile : an immersive environment for hierarchical live-looping. In *Proceedings of nime*, p. 192-197. Sydney, Australia.
- Berthaut F., Marshall M., Subramanian S., Hachet M. (2013). Rouages: Revealing the Mechanisms of Digital Musical Instruments to the Audience. In *Proceedings of nime*. Daejeon, South Korea.
- Berthaut F., Martinez D., Hachet M., Subramanian S. (2015). Reflets: Combining and revealing spaces for musical performances. In *International conference on new interfaces for musical expression*.
- Berthaut F., Zappi V., Mazzanti D. (2014, March). Scenography of immersive virtual musical instruments. In *Vr workshop: Sonic interaction in virtual environments (sive), 2014 ieee*, p. 19-24.
- Birnbaum D., Fiebrink R., Malloch J., Wanderley M. M. (2005). Towards a dimension space for musical devices. In *Nime '05: Proceedings of the 2005 conference on new interfaces for musical expression*, p. 192–195. Singapore, National University of Singapore.
- Blaine T., Fels S. (2003). Contexts of collaborative musical experiences. In *Proceedings of nime 03*, p. 129–134. Singapore, Singapore.
- Bonardi A., Barth lemy J. (2008, juin). The preservation, emulation, migration, and virtualization of live electronics for performing arts: An overview of musical and technical issues. *J. Comput. Cult. Herit.*, vol. 1, n° 1, p. 6:1–6:16. Consult  sur <http://doi.acm.org/10.1145/1367080.1367086>
- Boulanger R. (s. d.). *The csound book, perspectives in software synthesis, sound design, signal processing and programming*. MIT Press.

- Cadoz C. (1999). Les nouveaux gestes de la musique. In, p. 47-92. Éditions Parenthèses.
- Caramiaux B., Tanaka A. (2013). Machine learning of musical gestures. In *proceedings of the international conference on new interfaces for musical expression (nime 2013), seoul, south korea*.
- Cont A. (2008). *Modeling musical anticipation: From the time of music to the music of time*. Thèse de doctorat non publiée, UCSD.
- Dahl L. (2012). Wicked problems and design considerations in composing for laptop orchestra. In *Proceedings of nime 12*.
- Dahl S., Friberg A. (2007). Visual perception of expressiveness in musicians' body movements.
- Dannenberg R. (1984). An on-line algorithm for real-time accompaniment. In *Proceedings of the international computer music conference*.
- Desainte-Catherine M., Allombert A., Assayag G. (2013). Towards a hybrid temporal paradigm for musical composition and performance: The case of musical interpretation. *Computer Music Journal*, vol. 37, n° 2, p. 61–72.
- Dobrian C., Koppelman D. (2006). The 'E' in NIME: musical expression with new computer interfaces. In *Proceedings of the 2006 conference on new interfaces for musical expression*, p. 282.
- Eaton J., Eduardo M. (2014). Guide to brain-computer music interfacing. In, chap. On Mapping EEG Information into Music.
- Fober D., Letz S., Orlarey Y., Bevilacqua F. (2013, juillet). Programming Interactive Music Scores with INScore. In *Sound and Music Computing*, p. 185-190. Stockholm, Sweden. Consulté sur <https://hal.archives-ouvertes.fr/hal-00851956>
- Franco E., Griffith N. J. L., Fernström M. (2004). Issues for designing a flexible expressive audiovisual system for real-time performance & composition. In *Nime '04: Proceedings of the 2004 conference on new interfaces for musical expression*, p. 165–168. Singapore, Singapore, National University of Singapore.
- Fyans A. C., Gurevich M., Stapleton P. (2010). Examining the spectator experience. In *Proc. nime*, p. 451–454.
- Godøy R. I., Haga E., Jensenius A. R. (2006). Playing “air instruments”: mimicry of sound-producing gestures by novices and experts. In *Gesture in human-computer interaction and simulation*, p. 256–267. Springer.
- Gurevich M., Cavan Fyans A. (2011). Digital musical interactions: Performer–system relationships and their perception by spectators. *Organised Sound*, vol. 16, n° 02, p. 166–175.
- Hattwick I., Wanderley M. M. (2012). *A dimension space for evaluating collaborative musical performance systems*.
- Haury J. (2008). Notation musicale pour un clavier de deux touches. In *Documents musicaux, collection documents numériques*, vol. 11(3-4), p. 127-148. Lavoisier Hermes.
- Hunt A., Kirk R. (2000). Mapping strategies for musical performance. *Trends in Gestural Control of Music*, p. 231–258.

- Janin D., Berthaut F., Desainte-Catherine M. (2013). Multi-scale design of interactive music systems : the libTuiles experiment. In R. Bresin (Ed.), *SMC 2013*, p. 123-129. Stockholm, Sweden. Consulté sur <https://hal.archives-ouvertes.fr/hal-00813313>
- Jenseni A. R., Wanderley M. M., Godøy R. I., Leman M. (2009). Musical gestures. *Musical gestures: Sound, movement, and meaning*, vol. 12.
- Jordà S. (2003). Interactive music systems for everyone: exploring visual feedback as a way for creating more intuitive, efficient and learnable instruments. In *Proceedings of the stockholm music acoustics conference (smac03)*. Stockholm, Sweden.
- Jordà S. (2005). Multi-user instruments: models, examples and promises. In *Proceedings of the 2005 conference on new interfaces for musical expression*, p. 23–26. Singapore, Singapore, National University of Singapore. Consulté sur <http://dl.acm.org/citation.cfm?id=1085939.1085948>
- Jordà S., Kaltenbrunner M., Geiger G., Bencina R. (2005). The reactable\*. In *Proceedings of the international computer music conference*.
- Leonard J., Cadoz C., Castagne N., Florens J.-L., Luciani A. (2014). A virtual reality platform for musical creation: Genesis-rt. In *Sound, music, and motion*, p. 346–371. Springer.
- Levin G. (2000). *Painterly interfaces for audiovisual performance*. Thèse de doctorat non publiée, Massachusetts Institute of Technology.
- Malloch J., Sinclair S., Wanderley M. M. (2013). Libmapper: (a library for connecting things). In *Chi '13 extended abstracts on human factors in computing systems*, p. 3087–3090. New York, NY, USA, ACM. Consulté sur <http://doi.acm.org/10.1145/2468356.2479617>
- Manoury P. (1998). La note et le son : écrits et entretiens, 1981-1998. In *Musique et musicologie : Les dialogues*. L'Harmattan.
- Marshall M., Bennett P., Fraser M., Subramanian S. (2012, May). Emotional response as a measure of liveness in new musical instrument performance. In *Chi 2012 workshop on exploring hci relationship with liveness*. Consulté sur <http://www.cs.bris.ac.uk/Publications/Papers/2001542.pdf>
- Marshall M., Wanderley M. M. (2006). Vibrotactile feedback in digital musical instruments. In *Nime '06: Proceedings of the 2006 conference on new interfaces for musical expression*, p. 226–229. Paris, France, France, IRCAM — Centre Pompidou.
- Oh J., Herrera J., Bryan N. J., Dahl L., Wang G. (2010). Evolving the mobile phone orchestra. In *Proceedings of new interfaces for musical expression (nime10)*. Sydney, Australia.
- O'Modhrain M. S. (2001). *Playing by feel: incorporating haptic feedback into computer-based musical instruments*. Thèse de doctorat non publiée, Stanford, CA, USA. (Adviser-Chafe, Chris)
- Pressing J. (1997). Some perspectives on performed sound and music in virtual environments. *Presence*, vol. 6, n° 4, p. 482-503.
- Roberts C., Wakefield G., Wright M. (2013). The web browser as synthesizer and interface. In *Proceedings of the international conference on new interfaces for musical expression*, p. 313–318.

- Vercoe B., Puckette M. (1985). Synthetic rehearsal: Training the synthetic performer. In *Proceedings of the international computer music conference*.
- Verfaillie V. (2003). *Effets audio numériques adaptatifs*. Thèse de doctorat non publiée.
- Vertegaal R., Ungvary T., Kieslinger M. (1996). Towards a musician's cockpit: transducers, feedback and musical function. In *Proceedings of the international computer music conference*, p. 308–311.
- Vines B. W., Krumhansl C. L., Wanderley M. M., Dalca I. M., Levitin D. J. (2011). Music to my eyes: Cross-modal interactions in the perception of emotions in musical performance. *Cognition*, vol. 118, n° 2, p. 157–170.
- Walker R. (1987). The effects of culture, environment, age, and musical training on choices of visual metaphors for sound. *Perception and Psychophysics Vol 42(5)*, p. 491-502.
- Wanderley M. M., Orio N. (2002, September). Evaluation of input devices for musical expression: Borrowing tools from hci. *Computer Music Journal*, vol. 26, p. 62–76. Consulté sur <http://dl.acm.org/citation.cfm?id=1245194.1245202>
- Wanderley M. M., Orio N., Schnell N. (2000). Towards an analysis of interaction in sound generating systems. In *Isea2000 conference proceedings*.
- Wanderley M. M., Vines B. (2006). Music and gesture. In, p. p.167. Ashgate.
- Wang G., Bryan N., Oh J., Hamilton R. (2009). Stanford laptop orchestra(slork). In *Proceedings of the international computer music conference*, p. 505-508.
- Wegner D., Wheatley T. (1999). Apparent mental causation: Sources of the experience of will. *American Psychologist*, vol. 54.
- Weinberg G. (2005, juin). Interconnected musical networks: Toward a theoretical framework. *Comput. Music J.*, vol. 29, n° 2, p. 23–39. Consulté sur <http://dx.doi.org/10.1162/0148926054094350>
- Wessel D., Wright M. (2002). Problems and prospects for intimate musical control of computers. *Computer Music Journal*, vol. 26, n° 3, p. 11–22.
- Wright M. (2005). Open sound control: an enabling technology for musical networking. *Organised Sound*, vol. 10, n° 3, p. 193–200.
- Zappi V., Mazzanti D., Brogni A., Caldwell D. (2011). Design and evaluation of a hybrid reality performance. In *Nime '11: Proceedings of the 2011 conference on new interfaces for musical expression*, p. 355–360. Oslo, Norway.
- Zappi V., McPherson A. (2014). Design and use of a hackable digital instrument. In *Live interfaces*.

Article soumis le 1<sup>er</sup> juillet 2015

Accepté le 10 novembre 2015.

