

# Contrôle gestuel de la synthèse sonore<sup>1</sup>

## Introduction

Nous proposons ici une réflexion sur le contrôle gestuel de la synthèse sonore à partir de l'étude des gestes des musiciens. Cette réflexion sera divisée en quatre parties principales : *définition et typologie des gestes, acquisition des gestes, interfaces matérielles analogiques-numériques, correspondance entre variables gestuelles et variables de synthèse*. Elle se termine sur un exemple d'environnement de synthèse en temps réel développé à l'IRCAM et qui met en œuvre les idées présentées tout au long de l'article.

## 4.1. Instrument composé ou virtuel

En désolidarisant cause et effet, l'incursion de l'électricité dans la lutherie a profondément modifié la nature du jeu des instruments de musique. Et bien qu'une prodigieuse diversification des moyens de production sonore en ait résulté, ce changement de nature n'a pas encore engendré de stratégies de contrôle permettant de développer la même subtilité et la même ampleur de jeu que celle que l'on trouve pour les instruments traditionnels.

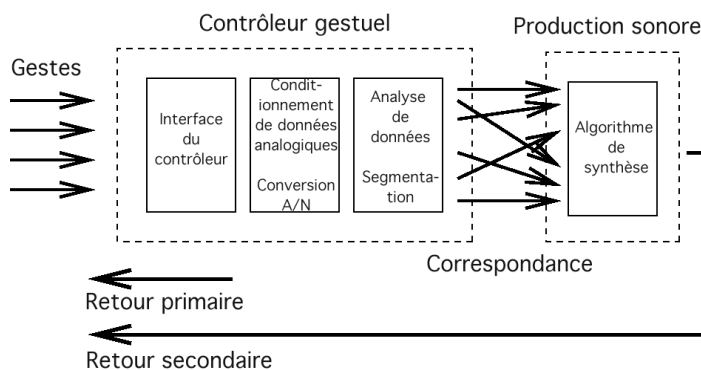
Afin d'améliorer la qualité de jeu des dispositifs de synthèse issus de cette nouvelle facture instrumentale, il importe tout d'abord de décrire puis d'analyser la structure type

---

<sup>1</sup> Marcelo M. Wanderley, Philippe Depalle et Xavier Rodet, Ircam - Centre Georges Pompidou, Groupe Analyse-Synthèse, 1, Pl. Igor-Stravinsky, 75004, Paris, France, (Philippe Depalle est également à la Faculté de Musique de l'Université de Montréal, 200, rue Vincent d'Indy, Outremont, Québec, Canada). Mél : [mwanderley@acm.org](mailto:mwanderley@acm.org), [depallep@ERE.UMontreal.CA](mailto:depallep@ERE.UMontreal.CA)

d'un *instrument virtuel* ou *composé*. Un tel instrument se décompose (cf. figure 4.1) en quatre composantes principales [DTW 97] qui sont :

- les *gestes du musicien* et les types de *retour* vers l'instrumentiste,
- le *contrôleur gestuel*,
- le *lien instrumental* qui traduit la *correspondance* entre variables *gestuelles* et variables de *synthèse*,
- et le *dispositif de synthèse sonore* proprement dit.



**Figure 4.1.** Structure type d'un instrument composé ou virtuel.

Parmi les *instruments virtuels* existants et incorporant un contrôleur gestuel, très peu possèdent un *lien instrumental* explicite. Notre propos est de montrer que l'intégration d'un ou plusieurs niveaux de correspondance entre paramètres permet d'améliorer le contrôle des synthétiseurs en utilisant les sorties des capteurs du contrôleur de manière plus élaborée.

Par ailleurs, l'étude des gestes et plus généralement du type d'interaction que développe un musicien avec son instrument peut être utile à la conception de dispositifs de contrôle gestuel qui simulent les interfaces d'instruments acoustiques traditionnels. Ceci rend alors possible l'utilisation d'un "vocabulaire gestuel expert" [RIS 99] déjà

existant, voire d'améliorer l'ergonomie de l'instrument ou d'en accélérer le temps d'apprentissage.

## 4.2. Gestes du musicien et types de retour

Afin d'élaborer des stratégies concernant le contrôle gestuel de dispositifs de synthèse sonore, il nous semble important de commencer par analyser la structure des gestes effectués par un musicien en situation de jeu instrumental traditionnel. Cette analyse peut devenir délicate si nous ne précisons pas tout d'abord ce que nous entendons par *geste*. En effet, la littérature regorge de définitions contradictoires [CW 99] de cette notion. Afin de se restreindre et pour éviter de discuter ici toutes les nuances possibles que peut revêtir ce mot dans le cadre de la communication homme-machine, nous nous limitons au seul domaine musical. Nous appelons ici *gestes de l'instrumentiste* les actions physiques effectuées par le musicien en situation de jeu instrumental. Par ailleurs, nous considérons également les actions physiques renvoyées par l'instrument sur son utilisateur. Nous appelons cette rétroaction : *retour instrumental*.

### 4.2.1. Catégorisation des Gestes

Très souvent, l'instrumentiste effectue simultanément plusieurs types de *gestes* lors d'une interprétation. Afin de définir ces différentes catégories, nous nous appuyons sur les résultats des nombreux travaux et réflexions déjà menés sur la notion de geste en musique, notamment ceux de François Delalande [DEL 88] et de Claude Cadoz [CAD 99b].

François Delalande propose une catégorisation du geste en trois classes : geste *effecteur*, geste *accompagnateur* et geste *figuratif*. La classe du *geste effecteur* regroupe l'ensemble des mouvements effectués pour produire mécaniquement les sons. La deuxième classe dite du *geste accompagnateur* rend compte des mouvements qui engagent le corps dans son entier : geste des épaules, mimiques, etc. Selon Delalande, ces gestes sont probablement aussi utiles à la production sonore qu'à l'épanouissement de l'imagination de l'instrumentiste. Enfin, les *gestes figuratifs* d'expression purement symbolique sont perçus par l'auditeur comme des articulations imprimant des balancements dans une mélodie. Delalande considère que même lorsque les *gestes figuratifs* sont corrélés aux *gestes effecteurs* (ceux de Glenn Gould dans le cas précis de son étude), cette relation demeure loin d'être claire.

Dans notre présente étude sur l'application du geste à la synthèse sonore nous nous intéressons plus particulièrement à la catégorie des *gestes effecteurs*. C'est ce type de geste que Claude Cadoz appelle *geste instrumental* [CAD 88].

Selon Claude Cadoz, les *gestes instrumentaux* se subdivisent eux-mêmes en trois classes, celle des *gestes d'excitation* (percussionnelle, continue ou entretenue), celle des *gestes de modification* (structurelle ou paramétrique) et enfin celle des *gestes de sélection* (séquentielle ou simultanée) [CAD 99a].

À ces trois classes, nous pouvons en rajouter une quatrième regroupant les gestes consistant à assurer des conditions normales de fonctionnement à l'instrument [CW 99]. Dans le cas d'une cornemuse par exemple, le geste du bras qui assure un niveau de pression suffisant pour le jeu entre dans cette quatrième catégorie. Nous proposons d'appeler ce type de geste, geste de *polarisation* ou de *maintien*. Il se distingue des trois autres types présentés ci-dessus dans le sens où il constitue un préalable essentiel à leur existence et à leur signification.

Notons enfin que nous avons effectué une analyse des *gestes accompagnateurs* (ou *ancillaires*) du jeu de clarinette et de leurs effets sur le son enregistré. Ces analyses sont publiées dans [WAN 99] [WDW 99]. Quant aux *gestes figuratifs*, nous pouvons citer en référence les travaux de S. Cannazza et al. sur l'expressivité du jeu de la clarinette [CPRV 97].

#### **4.2.2. Retour vers le musicien**

Afin de compléter la description du modèle d'instrument proposé, citons rapidement les deux types de *retour instrumental* : la *rétroaction primaire* essentiellement liée au contrôleur gestuel et la *rétroaction secondaire* directement liée à la production sonore [UV 99]. La *rétroaction primaire* est constituée du retour *visuel*, *tactilo-kinesthésique* et *auditif primaire*, ce dernier étant le bruit résultant du fonctionnement de l'instrument. Dans le cas d'un instrument à vent, il s'agit du bruit des clés. Quant à la *rétroaction secondaire*, elle est constituée par le *retour du signal sonore de l'instrument* aux oreilles du musicien.

Notons que le retour *tactilo-kinesthésique* (parfois appelé *tactilo-proprio-kinesthésique* [CAD 94]) revêt une grande importance dans le cas du jeu expert. Ainsi que le note J.-P. Roll [ROL 94], il existe dans presque tous les tissus du corps humain des récepteurs susceptibles d'être activés par des stimulations mécaniques. Les

sensations ainsi produites sont celles relatives à la *sensibilité superficielle* ou *cutanée* et celles relatives à la *sensibilité profonde* ou *proprioception*. La sensibilité cutanée est composée du toucher, de la sensibilité à la pression, aux vibrations et à la variation de température tandis que la *proprioception* désigne la sensibilité à la position dans l'espace, aux mouvements et aux forces exercées sur les muscles. Elle se divise elle-même en *kinesthésie*, sens du mouvement et en *statéthésie*, sens de la posture.

Les recherches menées dans le cadre de la synthèse sonore sur le *retour d'effort* entrent dans le cadre de la prise en compte de la *proprioception*. Puisque l'interface gestuelle est découplée du dispositif de production sonore dans un *instrument virtuel*, la *rétroaction* doit être explicitement simulée à l'aide d'*actionneurs* qui envoient une force de réaction en retour au geste appliqué<sup>2</sup> par l'instrumentiste.

### 4.3. Acquisition des gestes

Une fois répertoriés et identifiés, les gestes du musicien à prendre en compte pour le contrôle de la synthèse sonore doivent être captés, c'est-à-dire traduits en signaux électriques, puis numériques. Dans le cas des gestes instrumentaux, l'acquisition peut être *directe* ou *indirecte*<sup>3</sup>.

- Acquisition *directe* : différents types de capteurs mesurent l'ensemble des mouvements impliqués dans un même geste. C'est le cas, par exemple, de l'acquisition des mouvements complexes de manipulation d'un archet utilisé pour le jeu d'un instrument à cordes. En acquisition directe, la tendance est de capter des grandeurs physiques de base comme des forces, des déplacements, des accélérations, etc. Le plus souvent, chaque type de grandeur nécessite l'utilisation d'un type de capteur particulier.
- Acquisition *indirecte* : le geste est déduit de l'évolution des propriétés structurelles du son produit par l'instrument [EGO 95]. Le seul type de capteur utilisé dans ce cas est un microphone, c'est-à-dire essentiellement un capteur de pression ou de gradient de pression. Ce type d'acquisition fait ensuite appel à des techniques de traitement de signal temps-réel pour extraire l'information utile ce qui nécessite l'utilisation de dispositifs de calcul relativement puissants.

---

<sup>2</sup> Des exemples de recherche dans ce domaine sont disponibles sur les sites <http://www-acroe.imag.fr/> ou <http://haptic.mech.nyu.edu/>

<sup>3</sup> Il est même possible d'envisager une troisième forme d'acquisition consistant à prélever directement les informations musculaires (EMG) pendant l'activité physique dans le but d'identifier les gestes à partir de leur origine musculaire [TAN 93] [MAR 99].

Nous constatons que l'acquisition directe fournit des informations plus ou moins précises sur chacun des mouvements impliqués par le geste, tandis que l'acquisition indirecte mesure l'effet global du geste sur le signal sonore produit. En conséquence, il est parfois difficile de distinguer l'effet direct d'un mouvement sur le son instrumental des effets induits, tels qu'un filtrage dû au déplacement de l'instrument dans la salle où a lieu l'enregistrement.

Dans les sections 4.3.1. et 4.3.2. sont commentés les aspects relatifs à la capture directe des gestes puis, dans la section 4.3.3., sont évoquées les idées relatives à l'extraction du mouvement de l'instrumentiste par l'analyse des signaux sonores.

#### ***4.3.1. Capteurs et interfaces matérielles***

Le choix des capteurs représente une étape importante dans la conception d'un nouvel instrument. De nombreuses technologies de transducteurs sont disponibles sur le marché [FUC 99]. La plupart sont issues de recherches menées pour de tout autres objectifs que la production musicale. Néanmoins, plusieurs de ces technologies s'adaptent très bien au contexte de la nouvelle facture instrumentale. C'est le cas des capteurs utilisant des résistances sensibles à la pression, à une force (capteurs FSR de l'anglais, *force sensing resistors*). Par ailleurs, les capteurs à effet Hall ou utilisant l'effet piézo-électrique sont respectivement adaptés à la captation de faibles mouvements d'amplitude ou de force. Quant au captage de mouvements de plus grande amplitude, les principales techniques font appel à des systèmes vidéo, à ultrasons, à rayons infrarouges ou bien électromagnétiques<sup>4</sup>.

Il est néanmoins important de rappeler que le choix des capteurs n'est pas la seule étape importante dans l'élaboration d'un instrument. Bien souvent d'ailleurs, lorsque la motivation première de développement d'un instrument réside dans l'utilisation d'une nouvelle technologie de captage, le résultat sonore et musical s'avère décevant. En effet, la mise au point de la nouvelle technologie requiert alors l'essentiel de l'attention disponible au détriment d'autres éléments importants de l'instrument, comme le lien instrumental [WIN 95]. L'exploitation d'un nouveau type de contrôle aboutit souvent à une production sonore par trop démonstrative et qui, en étant trop connotée, manque finalement d'expressivité, surtout lorsqu'on la compare à la diversité de contrôle des instruments acoustiques traditionnels.

---

<sup>4</sup> Pour un survol des techniques de capteurs voir la page : <http://farben.latrobe.edu.au/motion/>

Une fois les capteurs choisis, il faut convertir les signaux électriques qu'ils produisent en signaux numériques. Après échantillonnage, les signaux sont le plus souvent formatés en informations de type contrôle MIDI [WW 96]. De nombreuses interfaces, appelées capteurs-MIDI et réalisant ces deux types d'opération, sont disponibles sur le marché. Afin d'étayer notre propos d'éléments concrets, nous dressons dans le tableau 4.4 une petite liste comparative des principales caractéristiques de quatre modèles<sup>5</sup>.

Nom de l'interface	ADB I/O	iCube	Sensorlab	AtoMIC Pro
Plate-forme	Macintosh	Indifférent, entrée MIDI	Indifférent, entrée MIDI	Indifférent, entrée MIDI
Fréquence d'échantillonnage maximale	< 90 Hz (15 Hz recommandé)	250 Hz (approximativement)	250 Hz (nominalement)	Latence maximale : 1 ms (hors convention MIDI)
Nombre d'entrées analogiques	4 chaque appareil [16 max. pour 4 ADBs]	24 en 12 bits 8 en 7 bits	32 en 8 bits 2x3 à ultrasons	32
Résolution maximale en entrée	8 bits	12 bits (24 entrées) 7 bits (8 entrées)	8 bits (32 entrées) 14 bits (2x3 ultrason)	8 bits
Type et nombre de sorties	Chaque canal configurable en entrée ou en sortie.	MIDI OUT & 8 sorties binaires	MIDI OUT	4 MIDI OUT indépendantes & 8 sorties numériques
Affichage	-	-	-	Écran LCD, 4 lignes de 20 caractères
Logiciels Fournis	Objet MAX, etc.	Objet MAX et Application Windows	Spider (Macintosh), interprète résident	Setup Patch MAX (Macintosh)
Dimensions (LxHxP) [mm]	122x29x65	121x34x94	200x35x100	165x38x225
Fabricant	BeeHive Technologies	Infusion Systems	STEIM	Ircam

Figure 4.2. Tableau comparatif de quatre modèles d'interfaces capteurs-MIDI.

<sup>5</sup> Une comparaison plus exhaustive est disponible en ligne, sur le site « Gesture Research in Music », à <http://www.ircam.fr/gesture>.

### 4.3.2. Contrôleurs gestuels

La plupart des *contrôleurs gestuels* réunissent plusieurs capteurs au sein d'une même unité. Ils constituent le premier bloc fonctionnel du modèle *d'instrument virtuel* (cf. figure 4.1). Nous pouvons alors distinguer trois catégories de contrôleurs gestuels, selon qu'ils imitent, s'inspirent ou se détournent des interfaces de contrôle gestuel des instruments traditionnels :

- Dispositifs de contrôle gestuel *imitatifs*. Les claviers "à toucher piano" et les guitares MIDI en sont des exemples types. Leur utilisation bénéficie pleinement de l'expertise gestuelle développée au fil du temps par nombre d'instrumentistes. Leur fabrication se fait dans un souci d'exacte ressemblance avec les dispositifs de contrôle des instruments originaux.



**Figure 4.3.** Interface capteurs-MIDI AtoMIC Pro.



**Figure 4.4.** SuperPalm, contrôleur MIDI de type violon développé par Suguru Goto (photo), Alain Terrier et Patrice Pierrot [WBD+ 98] [GOT 99].

- Dispositifs de contrôle gestuel *analogues*. Ces dispositifs rappellent la forme des interfaces rencontrées dans les instruments traditionnels. Mais dans ce cas, le but est surtout de conserver une analogie avec les interfaces traditionnelles pour les utiliser dans des situations musicales autres tout en exploitant éventuellement des propriétés différentes : claviers légers, ou le SuperPalm, par exemple. Ce dernier est un contrôleur qui rappelle la forme du violon mais qui peut être utilisé pour déclencher ou contrôler des événements de synthèse sonore n'ayant rien à voir avec la simulation des sonorités d'instruments à cordes. Il est par exemple utilisé pour contrôler de la synthèse granulaire.
- Dispositifs de contrôle gestuel *alternatifs*. Ceux-ci sont conçus pour utiliser des gestes qui ne font pas partie de l'expérience instrumentale musicale traditionnelle. Ils peuvent par exemple capter les déplacements du corps de



l'instrumentiste. De nombreuses interfaces *ad hoc* développées ces dernières années entrent dans cette catégorie, comme les fameux "gants de données" ainsi que les interfaces couramment utilisées dans d'autres domaines que la musique. Les tablettes graphiques en sont un exemple simple et caractéristique [WWF 97] [SDWR 99]. Dans la figure 4.5, nous montrons une réalisation à base de tablette graphique "augmentée" ou "instrumentée", c'est-à-dire à laquelle des capteurs de force et de déplacement ont été rajoutés (visibles sur le côté gauche, figure 4.5.a).



**Figure 4.5.a.** *Tablette graphique WACOM instrumentée par addition de capteurs de force type FSR (force sensing resistor).*



**Figure 4.5.b.** *Simulation d'un vibrato par mouvement du doigt sur tablette instrumentée [VIWR 99].*

#### 4.3.2.1. Comparaison des contrôleurs gestuels

La comparaison entre les différents types de contrôleurs gestuels s'avère difficile en raison de la diversité du mode de fonctionnement des capteurs tout autant que des signaux captés. Nous pouvons néanmoins profiter des travaux menés dans le domaine de l'interaction homme-machine [MCR90] [CMR91] sur la comparaison entre dispositifs d'entrée fondée sur leur décomposition en unités de captation primaires. Les unités de captation primaire sont classées en fonction de leur gestion de l'espace, de la nature de la grandeur physique captée et de la résolution à laquelle cette grandeur est captée. À ceci s'ajoutent les possibilités combinatoires entre les différentes unités de captation. Soit, de manière explicite :

- a) Les six degrés de liberté possibles : translation et rotation selon les trois directions X, Y et Z.

b) Les grandeurs physiques captées : - position P et sa dérivée dP, - force F et sa dérivée dF pour les translations, - angle A et couple T ainsi que leurs dérivées dA et dT pour les rotations.

c) La résolution de chaque variable : 1, 10, 100 ou infinie.

d) Le type de combinaison entre les unités de captation : (de l'anglais) « merge », « layout » ou « connect ».

La figure 4.6 montre l'application de la méthode proposée à la tablette graphique augmentée par les capteurs de pression. Nous pouvons facilement y voir les degrés de liberté, les types et les résolutions de chaque variable mesurée pour le stylo, le palet et les capteurs de pression (numéro 2 sur le graphique) et pression-déplacement. Par exemple, nous voyons que le stylo possède tous les degrés de liberté excepté la rotation selon l'axe Z.

⊖ Stylo  
○ Palet  
⊗ Capteurs additionnels

		Translation			Rotation			
		X	Y	Z	rX	rY	rZ	
P o s i t i o n	P	○	○	○	○	○		A
	dP							dA
F o r c e	F			⊗				T
	dF							dT
		1 10 100 inf	1 10 100 inf	1 10 100 inf	1 10 100 inf	1 10 100 inf	1 10 100 inf	

A  
n  
g  
l  
e  
  
C  
o  
u  
p  
l  
e

**Figure 4.6.** Présentation synthétique des caractéristiques des dispositifs d'entrée de la tablette graphique instrumentée [SDWR 99].

### 4.3.3. Techniques d'acquisition indirecte de gestes

À l'inverse des techniques d'acquisition directes qui captent les causes produisant le son instrumental, les techniques d'acquisition indirectes captent l'effet, c'est-à-dire le signal sonore lui-même. Puisque tout geste *effecteur* influence directement les caractéristiques du son produit, l'idée est de remonter la chaîne de production sonore afin de retrouver le geste effectué. Il s'agit, par exemple, d'identifier le doigté utilisé par un musicien à partir de la fréquence fondamentale estimée du son généré. En pratique, le son est capté par un microphone avec toutes les imperfections que cela comporte (qualité du microphone, influence de la salle d'enregistrements, bruits résiduels, etc.). Puis il est numérisé et soumis à des algorithmes de traitement du signal pour en extraire les paramètres du geste. Lorsque la technique de synthèse est fondée sur un modèle physique, les paramètres extraits sont des paramètres mécaniques ou acoustiques. Cependant, lorsque la technique de synthèse est fondée sur un modèle de signal, il n'est pas forcément nécessaire de remonter explicitement jusqu'aux paramètres du geste : il peut suffire d'estimer les variations de paramètres du signal sonore dont on est certain qu'elles sont directement produites par le geste en question. Ce peut être par exemple l'effet du geste sur les amplitudes et les fréquences des partiels. Cela revient d'une certaine manière à échantillonner le *lien instrumental* (Cf. 4.4).

La première étape consiste donc à extraire les caractéristiques les plus importantes définissant le signal sonore, c'est-à-dire la *puissance à court terme*, la *fréquence fondamentale*, l'*enveloppe spectrale*, ainsi que les *amplitudes*, *fréquences* et *phases* des partiels.

La *puissance à court terme*, qui rend compte du profil dynamique du signal sonore, informe sur la force avec laquelle le son a été produit mais également sur les éventuels mouvements de rapprochement et d'éloignement de l'instrument qui ont eu lieu. Pour retrouver des informations sur le geste de l'instrumentiste à partir de ce paramètre, il faut pondérer par la directivité de l'instrument et par le fait que certains sons sont plus faciles à produire que d'autres : à même « effort » d'un chanteur, l'émission des voyelles /i/ et /y/ se fait moins puissante que celles des voyelles /a/ et /o/, par exemple.

La *fréquence fondamentale* rend compte du profil mélodique du son et renseigne sur le doigté, la sélection dans le cas d'instruments à oscillateurs multiples, etc. Notons que

la recherche du geste à partir de la fréquence fondamentale peut s'aider de l'information de puissance à court terme : en effet, l'augmentation de la force sonore peut parfois entraîner une octaviation c'est-à-dire faire sauter la fréquence fondamentale d'une octave vers le haut.

L'*enveloppe spectrale* représente sous forme de courbe à évolution douce la répartition des amplitudes des partiels du signal sonore. Cette courbe est avant tout représentative de la conformation du résonateur de l'instrument. Ainsi, lorsque le résonateur est déformable, l'enveloppe spectrale fournit des informations sur les mouvements ayant engendré les déformations. Ceci se manifeste indubitablement dans le cas de la voix chantée, mais également de manière plus subtile dans le cas d'une trompette à sourdine. La pente de décroissance globale de l'enveloppe spectrale apporte également des informations sur la force sonore imprimée par le musicien.

Quant aux *fréquences, amplitudes et phases* des partiels, elles peuvent à elles seules apporter l'ensemble des informations fournies par les paramètres cités précédemment. De plus, elles permettent d'évaluer l'inharmonicité qui nous instruit sur la force sonore ainsi que sur le lieu de la frappe d'une corde ou d'une membrane, par exemple.

Afin de bien échantillonner les signaux captés et les paramètres extraits par analyse du signal sonore, il importe de connaître leurs bandes passantes. Celles-ci dépendent des contraintes mécaniques et physiologiques ayant gouverné leur production. Lorsqu'il s'agit d'un mouvement continu ou bien de modulation, les constantes de temps sont déterminées par des contraintes physiologiques. Il en résulte une bande passante comprise entre 100 et 1000 Hz. Il faut donc échantillonner à une fréquence allant de 200 à 2000 Hz.

Toutefois, certains signaux peuvent être à variations brusques, notamment les signaux de sélection et les signaux engendrant un changement dans le mode de fonctionnement de l'instrument. C'est le cas de la fermeture d'un clapet ou bien du contact brusque d'un objet sur un autre. Dans ces dernières situations, si le signal est régulièrement échantillonné, il convient de hausser la valeur de la fréquence d'échantillonnage.

#### **4.4. Correspondance entre paramètres de contrôle et paramètres de synthèse**

Comme nous l'avons souligné au paragraphe 4.1, un *instrument virtuel ou composé* doit intégrer un élément traduisant explicitement la correspondance entre les variables

gestuelles, paramètres issus du contrôleur, et les variables de synthèse que sont les paramètres d'entrée de l'algorithme de calcul des échantillons sonores. Nous avons appelé cet élément *lien instrumental* en traduction du terme anglais "mapping".

C'est le *lien instrumental* qui, en réalité, donne un sens au paramètre de sortie du contrôleur. Selon la nature de ce lien, un même mouvement représentera la fréquence d'un partiel, la fréquence fondamentale d'un signal périodique, le taux de bruit, l'intensité, l'inharmonicité ou bien des paramètres perceptifs comme la hauteur, la sonie, la rugosité, etc. Ce lien doit permettre au musicien de contrôler *l'instrument* selon l'approche voulue (de bas niveau, perceptive, ou abstraite) tout en assurant précision et expressivité. À cette fin, plusieurs stratégies de correspondance ont été proposées notamment dans [BPMB 90] [LW 92] [RWDD 97], mais la question reste ouverte.

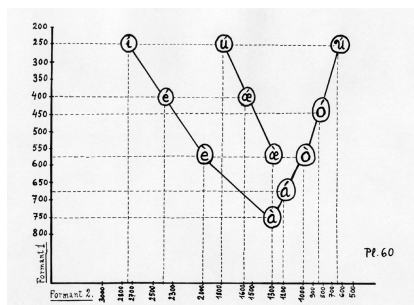
En ce qui concerne la simulation du jeu instrumental, il faut ensuite différencier deux cas selon que le synthétiseur contrôlé repose sur un modèle de signal ou bien un modèle physique. Dans ce dernier cas, une partie du *lien instrumental* est déjà intégrée à l'algorithme de synthèse sous forme de contraintes mécaniques ou acoustiques. Quant à l'accès de type jeu instrumental de synthétiseurs fondés sur des modèles de signaux, le plus simple est de représenter le *lien instrumental* par une "couche logicielle" implantant des règles de correspondance entre le vecteur des paramètres de sortie du contrôleur et le vecteur des paramètres d'entrée de l'algorithme de synthèse. Cette couche logicielle définit des paramètres de contrôle à différents niveaux (perceptifs, abstrait, musical, etc.) du type :

- Espaces instrumentaux (Cf. Fig. 4.7.b). Le musicien accède à des grandeurs perceptives définies comme des paramètres de représentation des sons d'instruments de l'orchestre [WES85] [VB 94].
- Triangle vocalique (Cf. Fig. 4.7.a) [STR 65]. Le musicien se déplace dans le triangle vocalique afin de contrôler un synthétiseur de voix chantée fonctionnant sur le principe des fonctions d'onde formantiques (FOF) [ROD 84].
- Forme géométrique. Une sphère virtuelle peut être déformée au moyen d'un gant de données connecté à l'ordinateur en tant que dispositif d'entrée [MFM 97].

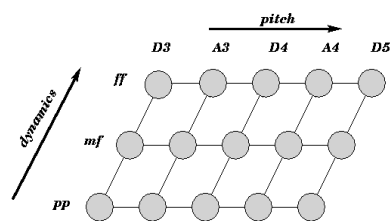
Mais indépendamment du paradigme choisi, un avantage du *lien instrumental* est de proposer à l'utilisateur la possibilité de manipuler des abstractions de plus haut niveau que des paramètres de fréquence, d'amplitude et de phase pour définir le spectre d'un

son dans le cas de la synthèse additive, ou bien des rapports de fréquence porteuse à fréquence modulante (c:m) dans le cadre de la synthèse FM.

Un deuxième avantage de la représentation du *lien instrumental* sous forme de couche logicielle réside dans la possibilité de l'organiser en deux (voire plusieurs) niveaux indépendants : le premier est alors fonction du choix du dispositif d'entrée (contrôleur) et le second du choix du type d'algorithme de synthèse. Il en résulte une grande facilité pour changer de contrôleur sans se soucier de la correspondance explicite entre paramètres intermédiaires et paramètres du synthétiseur. De plus, nous pouvons facilement définir plusieurs options d'implantation du premier niveau pour un même contrôleur. Par ailleurs, la possibilité de modifier la correspondance laisse entrevoir le choix entre plusieurs niveaux offrant un accès de complexité croissante. L'apprentissage d'un instrument virtuel serait alors grandement facilité par le passage progressif d'un niveau élémentaire à un niveau expert. Ceci est illustré par l'exemple décrit dans la section 4.5.



**Figure 4.7.a.** Distribution de onze voyelles françaises repérées selon les fréquences de leurs deux premiers formants (d'après [STR 65])



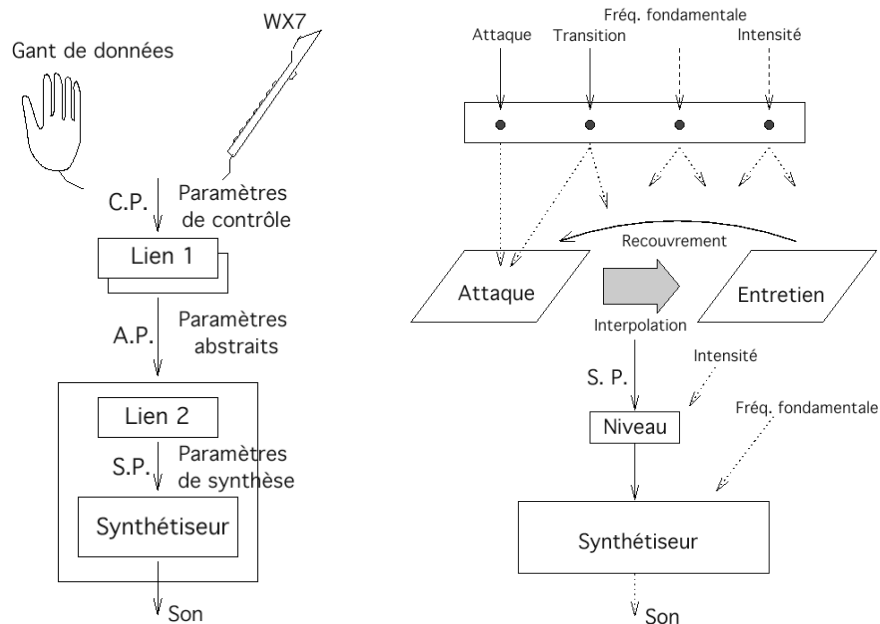
**Figure 4.7.b.** ESCHER - distribution des modèles additifs en fonction de la hauteur et de l'intensité des notes d'une clarinette.

Une étude dans cette direction a été proposée dans [WSR 98], où la correspondance entre les paramètres du contrôleur gestuel et ceux d'un espace bidimensionnel de timbres peut être choisie entre plusieurs modèles plus ou moins similaires au fonctionnement d'un instrument acoustique.

Ces remarques nous suggèrent que le lien instrumental peut être considéré comme le niveau où se définit le niveau d'interaction effective entre l'utilisateur et la machine. Ainsi, une forme de « zooming » continu entre deux niveaux de correspondance permet d'établir un contrôle passant d'un niveau macroscopique - phrasé, rythme - à un niveau microscopique - articulation ou contrôle précis du timbre.

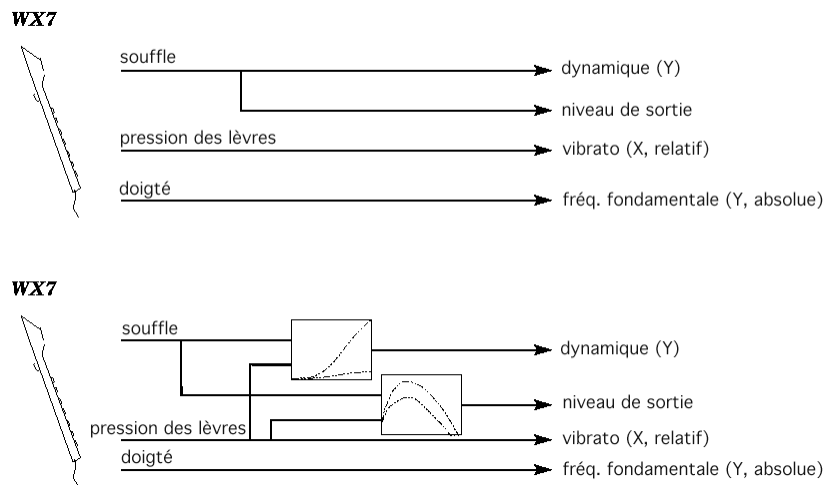
#### 4.5. Un exemple d'application

Nous présentons maintenant un projet développé à l'Ircam en collaboration avec Norbert Schnell et Joseph Butch Rován [WSR 98] qui implante dans le logiciel FTS [DDMS 96] un prototype de système de synthèse temps-réel prenant en compte les idées présentées dans les sections antérieures.



**Figure 4.8.** *ESCHER* À gauche: les deux niveaux de lien instrumental entre paramètres de contrôle et paramètres de synthèse. À droite : les paramètres abstraits (en haut) et l'explicitation du deuxième niveau de lien instrumental utilisé pour la synthèse (en bas). [WSR98]

L'idée principale est d'avoir un instrument virtuel flexible dans lequel plusieurs contrôleurs gestuels peuvent accéder à une même technique de synthèse. Utilisant des contrôleurs de type saxophone MIDI WX7 et un synthétiseur additif, nous avons fait des essais mettant en jeu plusieurs versions du *premier niveau* de *lien instrumental* allant d'un niveau de correspondance directe entre paramètres (un à un) jusqu'à un niveau expert. Ce dernier prend en compte les contraintes liées au mode de fonctionnement physique d'une anche de clarinette ou bien de saxophone [RWDD 97]. La correspondance directe offre un jeu instrumental simplifié, mais aux possibilités expressives inévitablement réduites tandis que le niveau expert rend le jeu plus proche de celui de l'instrument acoustique réel grâce à l'interdépendance entre certains paramètres comme le débit du souffle et l'ouverture de l'embouchure. Certes, l'instrument devient alors plus difficile à maîtriser pour un musicien non-expert.



**Figure 4.9.** Deux implantations du premier niveau de lien instrumental dans le cas d'une simulation de clarinette : jeu simplifié (en haut) et jeu expert (en bas) [RWDD 97].

#### 4.6. Conclusion

Dans cet article, nous avons présenté quelques idées concernant le contrôle gestuel de la synthèse sonore en articulant notre propos autour d'un modèle d'*instrument composé* ou *virtuel* constitué de quatre parties. Nous avons insisté sur l'intérêt



d'expliciter un *lien instrumental* approprié à une application spécifique. Après définition, classement et analyse des différents gestes effectués par un instrumentiste, nous avons présenté les différentes possibilités de captation de ces gestes, puis les technologies d'interfaces capteur-MIDI pouvant être mises en œuvre. Des éléments de comparaison entre contrôleurs gestuels, proposés dans autres domaines, ont été adaptés au cas des contrôleurs utilisés dans le cadre musical. Ils sont fondés sur la décomposition en variables de base comme la gestion des degrés de liberté spatiaux, des types de mouvement et de la résolution de chaque variable.

Nous nous sommes ensuite intéressés aux possibles *liens instrumentaux* entre paramètres du contrôleur gestuel et paramètres de l'algorithme de synthèse en situation de simulation de jeu instrumental. Nous avons montré combien le choix et la définition d'un *lien instrumental* jouent un rôle important dans les capacités expressives de l'instrument résultant.

Finalement, nous avons décrit un exemple pratique dans lequel ces différents concepts ont été utilisés pour donner lieu à un instrument virtuel interactif simulant un jeu de clarinette ou bien de saxophone.

#### **4.7. Remerciements**

Nous tenons ici à remercier MM. Patrice Pierrot et Alain Terrier pour leur collaboration sur le travail des capteurs ainsi que MM. Norbert Schnell, Joseph Butch Rován et Shlomo Dubnov pour leur collaboration au développement de l'environnement de synthèse ESCHER. Merci également à Aymeric Willier pour ses suggestions concernant l'acquisition des gestes par analyse de signaux biologiques.

Ce travail a été en partie développé grâce à un financement du *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq), Brésil, auquel nous adressons nos plus vifs remerciements.

## Bibliographie

- [BPMB 90] Bowler, I., Purvis, A., Manning, P. et Bailey, N., *On Mapping N Articulation Parameters onto M Synthesiser-control Parameters*, Proceedings of the International Computer Music Conference, ICMC, 1990, pp. 181 — 184.
- [CAD 88] Cadoz, C., *Instrumental Gesture and Music Composition*, Proceedings of the International Computer Music Conference, ICMC, 1988, pp. 1—12.
- [CAD 94] Cadoz, C., *Le geste canal de communication homme-machine. La communication « instrumentale »*. Sciences Informatiques - Numéro Spécial : Interface Homme-Machine, 1994, pp. 31 — 61.
- [CAD 99a] Cadoz, C., *Musique, geste, technologie*, Les Nouveaux Gestes de la Musique, H. Genevois et R. de Vivo (eds), Editions Parenthèses, 1999, pp. 47 — 92.
- [CAD 99b] Cadoz, C., *Continuum énergétique du geste au son : le geste instrumental et les interfaces ergonomiques*, article dans cet ouvrage.
- [CW 99] Cadoz, C. et Wanderley, M., *Gesture - Music*, Trends in Gestural Control of Music, M. Wanderley, M. Battier et J. Rován (eds), Ircam, 1999 (à paraître).
- [CPRV 97] Canazza, S., De Poli, G., Rinaldi, S. et Vidolin, A., *Sonological Analysis of Clarinet Expressivity*, Music, Gestalt and Computing, M. Leman (ed), Springer Verlag, 1997, pp. 431 — 440.
- [CMR 91] Card, S. K., MacKinlay, J. D. et Robertson, G. G., *A Morphological Analysis of the Design Space of Input Devices*, ACM Transactions on Information Systems, Vol. 9, No. 2, Avril 1991, pp. 99 — 122.
- [DDMS 96] Dechelle, F., DeCecco, M., Maggi, E., et Schnell, N., *New DSP Applications on FTS*, Proceedings of the International Computer Music Conference, ICMC, 1996, pp. 188 — 189.

- [DEL 88] Delalande, F., *La gestique de Glenn Gould*, Glenn Gould Pluriel, Louise Courteau Éditrice, 1988, pp. 84 — 111.
- [DTW 97] Depalle, P., Tassart, S., et Wanderley, M., *Instruments virtuels - Les vertus du possible*, Résonance, No. 12, 1997, pp. 5 — 8.
- [EGO 95] Egozy, E. B., *Deriving Musical Control Features from a real-time Timbre Analysis of the Clarinet*, Master's thesis, Massachusetts Institut of Technology, 1995.
- [FUC 99] Fuchs, P., *Les interfaces de la réalité virtuelle*, Collection « Interfaces – Les journées de Montpellier », édition révisée, 1999.
- [GOT 99] Goto, S., *Virtual Musical Instruments : Aesthetics and Technological aspects. The case of the Superpolm MIDI Violin*, Leonardo Journal, 1999 (à paraître).
- [LW 92] Lee, M. et Wessel, D., *Connectionist Models for Real-Time Control of Synthesis and Compositional Algorithms*, Proceedings of the International Computer Music Conference, ICMC, 1992, pp. 277 — 280.
- [MCR 90] MacKinlay, J. D., Card, S. K. et Robertson, G. G., *A Semantic Analysis of the Design of Input Devices*, Human-Computer Interaction, Vol. 5, 1990, pp. 145 — 190.
- [MAR 99] Marrin, T., *Searching for Meaning in Gestural Data : Interpretative Feature Extraction and Signal Processing for Affective and Expressive Content*, Trends in Gestural Control of Music, M. Wanderley, M. Battier et J. Rován (eds), Ircam, 1999 (à paraître).
- [MFM 97] Mulder, A., Fels, S., et Mase, K., *Empty-handed Gesture Analysis in Max/FTS*, Proceedings of KANSEI, the Technology of Emotion Workshop, 1997, pp. 87 — 91.
- [RIS 99] Risset, J.-C., *Nouveaux gestes musicaux : quelques points de repère historiques*, Les Nouveaux gestes de la musique, H. Genevois et R. de Vivo (eds), Editions Parenthèses, 1999, pp. 19 — 33.
- [ROD 84] Rodet, X., *Time-Domain Formant-Wave-Function Synthesis*, Computer Music Journal, Vol. 8, No. 3, 1984, pp. 9 — 14.
- [ROL 94] Roll, J.-P., *Sensibilités cutanées et musculaires*, Traité de Psychologie Expérimentale, M. Richelle, J. Requin et M. Robert (eds), chap. II/6, PUF, 1994, pp. 483 — 542.
- [RWDD 97] Rován, J., Wanderley, M., Dubnov, S. et Depalle, P., *Instrumental Mapping Strategies as Expressivity Determinants in Computer Music*, Proceedings of KANSEI, the Technology of Emotion Workshop, 1997, pp. 68 — 73.
- [SDWR 99] Serafin, S., Dudas, R., Wanderley, M. et Rodet, X., *Gestural Control of a real-time Physical Model of a Bowed String Instrument*, Proceedings of the International Computer Music Conference, ICMC, 1999 (à paraître).
- [STR 65] Straka, G., *Album Phonétique*, Les Presses de l'Université Laval, Québec, 1965.
- [TAN 93] Tanaka, A., *Musical Technical Issues in Using Interactive Instrument Technology with Application to the Biomuse*, Proceedings of the International Computer Music Conference, ICMC, 1993, pp. 124 — 126.

- [UV 99] Ungvary, T. et Vertegaal, R., *Cognition and Physicality in Musical CyberInstruments*, in Trends in Gestural Control of Music, M. Wanderley, M. Battier et J. Rován (eds), Ircam, 1999 (à paraître).
- [VB 94] Vertegaal, R. et Bonis, E., *ISEE : an Intuitive Sound Editing Environment*, Computer Music Journal, Vol. 18, No. 2, 1994, pp. 21 — 29.
- [VIWR 99] Viollet, J.-P., Isart, F., Wanderley, M. et Rodet, X., *On the Choice of Transducer Technologies for Specific Musical Functions: A study on gestural control applied to sound synthesis*, Rapport Interne, IRCAM, 1999.
- [WBD+ 98] Wanderley, M., Battier, M., Depalle, P. et al., *Gesture Research at IRCAM : A progress report*, Proceedings of the Journées d'Informatique Musicale (JIM), 1998.
- [WSR 98] Wanderley, M., Schnell, N. et Rován, J., *ESCHER : Modeling and Performing « Composed Instruments » in Real-time*, Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, SMC, 1998.
- [WAN 99] Wanderley, M., *Non-Obvious Performer Gestures in Instrumental Music*, Proceedings of the III Gesture Workshop, Springer-Verlag, 1999 (à paraître).
- [WDW 99] Wanderley, M., Depalle, P. et Warusfel, O., *Improving Sound Synthesis by Modeling the Effects of Performer Gestures*, Proceedings of the International Computer Music Conference, ICMC, 1999 (à paraître).
- [WES 85] Wessel, D., *Timbre Space as a Musical Control Structure*, Foundations of Computer Music, C. Roads et J. Strawn (eds), MIT Press, 1985, pp. 640 — 657.
- [WW 96] Williams, D. B. et Webster, P. T., *Experiencing Music Technology*, Chapitre 22, Schirmer Books, 1996.
- [WIN 95] Winkler, T., *Making Motion Musical: Gestural mapping strategies for interactive computer music*, Proceedings of the International Computer Music Conference, ICMC, 1995, pp. 261 — 264.
- [WWF 97] Wright, M., Wessel, D. et Freed, A., *New Musical Control Structures from Standard Gestural Controllers*, Proceedings of the International Computer Music Conference, ICMC, 1997, pp. 387 — 390.