

Le rôle du contrôle gestuel dans l'informatique musicale

Daniel Arfib, Jean-Michel Couturier, Loïc Kessous

[*arfib,couturier,kessous*]@*lma.cnrs-mrs.fr*
CNRS-LMA 31 Chemin Joseph Aiguier
13402 Marseille Cedex 20

RESUME

Exemples sonores à l'appui, il sera montré que quelle que soit la méthode de synthèse numérique utilisée en Informatique musicale, ce sont les variations des paramètres de contrôle, en particulier lorsqu'ils sont reliés aux gestes, qui valorisent le caractère musical des algorithmes utilisés.

I. INTRODUCTION

L'objectif de cet article est de présenter différentes méthodes numériques susceptibles de fabriquer du son, et de dispositifs pour le contrôler en temps réel, les donnant ainsi véritablement leur statut d'instruments de musique (nombreuses références dans [6] et [7]).

II. MODELES ET IMPLEMENTATIONS

II.1. modèles physiques et modèles de signaux

Les modèles de synthèse en Informatique musicale sont l'ossature qui permet de créer des sons sur ordinateur. Dès le départ de cette science, l'enjeu musical est fort : il s'agit de savoir ce qui permet à l'auditeur d'entendre des sons "musicaux" c'est à dire en quelque sorte ce qui berne le système avec des stimuli produits artificiellement.

Il est évident que l'histoire a été un point majeur, et que la succession de la synthèse additive, soustractive, par modulation, granulaire puis par modèles de guide d'ondes a marqué une typologie, quitte à faire des fractures là où n'étaient que des nuances. La synthèse additive est très présente dans les premiers temps de la recherche, et un auteur comme Risset en est un artiste très représentatif. La synthèse soustractive, pour la plupart conditionnée par sa faculté de reproduire une certaine vocalité, est issue en partie des recherches sur la voix artificielle ou les "tables d'harmonie" d'instruments. Ainsi Mathews a démontré combien la sonorité d'un violon électronique est conditionnée par l'ensemble des résonances incluse dans son modèle. Les synthèses par modulation de fréquence et par distorsion non linéaire ont donné un éclairage particulier par la réduction de données qu'elles permettent, permettant de "saisir" d'emblée les caractéristiques du timbre d'un son. Les modèles granulaires ont permis de se départir de la contrainte "somme de sinusoides" pour entrer dans la notion de flux, fut-il aléatoire. Les modèles par guides d'onde (fig.1) ont réintroduit dans la synthèse pure la notion de corde, de tuyau ou d'embouchure par une combinaison de lignes à retard, de conformation d'ondes, et de filtres. Parallèlement les recherches en

acoustique instrumentale n'ont pas cessé qui ont permis d'établir des équations directes ou inverses permettant de créer des sons à partir de données plus physiques.

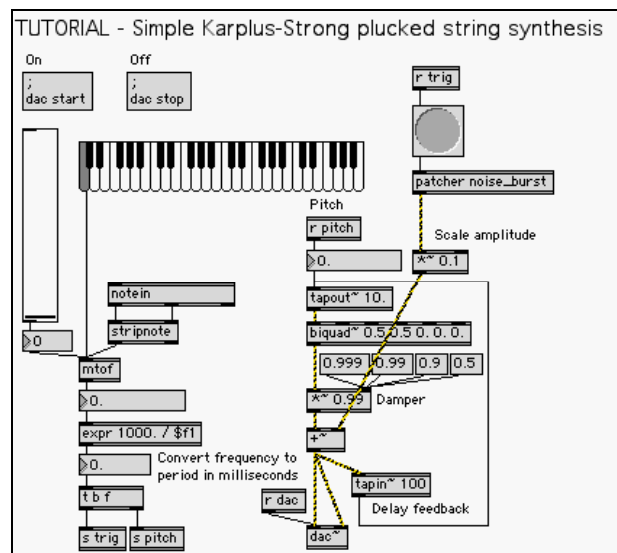


Figure 1. modèle guides d'onde en langage Max-Msp

L'usage musical est toutefois en grande partie indépendant de ces méthodes, en ce sens que le musicien sait se jouer des limitations inhérentes à celles-ci pour dégager un sens esthétique au son qu'il produit. L'imitation stricte du monde réel peut même parfois apparaître simplificateur et réducteur malgré une complexité des algorithmes grandissante. En fait l'important réside parfois plus dans les contrôles, ce qu'on l'appelle les "conduites" que dans les algorithmes eux-mêmes. Un exemple frappant est donné par le synthétiseur DX7, gloire des années 80 qui, à partir d'algorithmes rudimentaires : il s'agit d'utiliser 6 oscillateurs combinés en modulation de fréquence, a donné lieu à des banques de sons innombrables et variées ; variété cependant très typée, dans la mesure où le pilotage par clavier MIDI conduit à créer des profils de contrôle répétitifs d'une note à l'autre.

II.2. systèmes hors temps réel et temps réel

Une grande distinction historique a été faite entre les systèmes hors temps-réel et les systèmes temps-réel : les

premiers, représentés par des programmes tels que Music V ou Csound [1] assurent une polyphonie et une complexité extensible, tandis que les seconds représentés typiquement par le langage Max-Msp, répondent à une limite : chaque échantillon sonore doit être calculé en moins de 22 microsecondes (à 44.1 KHz de fréquence d'échantillonnage, signal mono), ce qui implique une complexité et une polyphonie limitées.

Cette comparaison ne doit cependant pas s'arrêter à une banale question de traitement de signal : la démarche musicale hors temps-réel est différente de celle du temps-réel : la première s'apparente à un acte de composition, la seconde à un acte d'interprétation ou d'improvisation. D'ailleurs il est intéressant de noter que le passage d'une partition MusicV à un programme Max-Msp a été réalisé pour une œuvre entière, "le Souffle du Doux" [2]. Deux étapes existent alors, la première étant la reproduction à l'identique de l'œuvre – il s'agit simplement de séquencer les appels aux instruments avec les bonnes valeurs – et la seconde étant la construction d'une lutherie électronique associant à un ordinateur un ensemble de capteurs-actuateurs permettant de jouer de ces instruments.

A titre d'exemple, nous décrivons maintenant deux implémentations issues d'archétypes de la musique numérique : le paradoxe de hauteur de Jean-Claude Risset utilisé dans " Mutation " et la distorsion non linéaire utilisée par Daniel Arfib dans " le Souffle du Doux ", décrits ici pour la simplicité de leur mise en œuvre et leur possibilité de contrôle gestuel ultérieur.

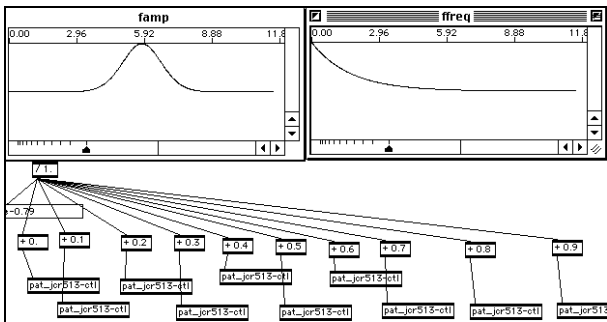


Figure 2. paradoxe de hauteur en Max-Msp

Les paradoxes de hauteur reposent sur une combinaison de sons sinusoïdaux à l'octave de l'autre, dont l'amplitude est fonction de la hauteur de ces composantes. Quand la fréquence fondamentale glisse, ces composantes décrivent donc une fonction croissante puis décroissante. La mise en œuvre en Max-Msp (fig.2) est quasi identique à celle utilisée en MusicV : on considère que les fréquences vont glisser sur 10 octaves, ce qui entraîne qu'il y a dix oscillateurs en parallèle, dont les amplitudes et fréquences sont gérées par deux fonctions : l'une est une exponentielle décroissante, assurant un glissando linéaire, et l'autre une courbe " en cloche " copiée de l'instrument Music V. L'instrument

fonctionne parfaitement quand il est activé par une évolution temporelle, ici donnée par un oscillateur en dents de scie dont la période est dix fois celle de la boucle de répartition. Un incrément positif assure un glissando descendant, et le contraire pour un incrément négatif.

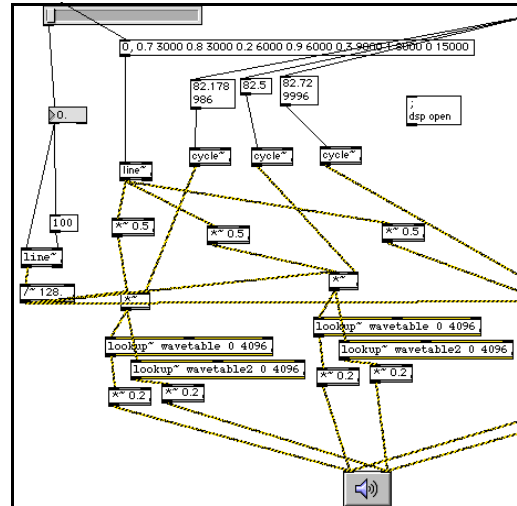


Figure 3. distorsion non-linéaire en Max-Msp

L'exemple de la distorsion non linéaire (fig.3) est démonstratif des stratégies possibles hors temps réel et en temps réel. Si l'instrument est ici classique - il s'agit d'utiliser la distorsion non linéaire de sinusoïdes par des fonctions de transfert non linéaire - les moyens de contrôler l'amplitude d'entrée de ces sinusoïdes est possible soit à travers une fonction préprogrammée, qui reproduit en toute fidélité un profil défini lors de la composition, ou bien en utilisant une entrée externe qui va permettre le contrôle gestuel.

III. LE GESTE CREATIF

III.1. les points de contrôle par le geste

Pour introduire un geste, il faut lier une variable de Max-Msp à un ensemble de paramètres à des valeurs issues des périphériques gestuels.



Figure 4. contrôle gestuel à l'aide d'un radio-Baton (g)

La première catégorie de gestes possibles est celle des gestes de décision. Un des plus importants est le séquençage temporel des événements, associé à un paramètre significatif. Dans le domaine MIDI, c'est l'activation d'une note (Note On) associée à la vélocité (vitesse de frappe d'une touche). Bien évidemment les programmes Max-Msp doivent être modifiés en conséquence pour activer la note au moment adéquat, et répartir l'influence du paramètre supplémentaire sur le déroulement du son, par exemple sur la courbe d'attaque d'une enveloppe, ou le niveau d'un index ou d'une amplitude. La seconde catégorie est celle des gestes de modulation. Ils proviennent souvent de périphériques MIDI qui donnent des valeurs variables en fonction de la position ou du mouvement : typiquement un potentiomètre est de cette nature, ou une tablette graphique. Des périphériques gestuels dédiés tels des gants, ou des capteurs de position, donnent aussi un flot de données de ce type.

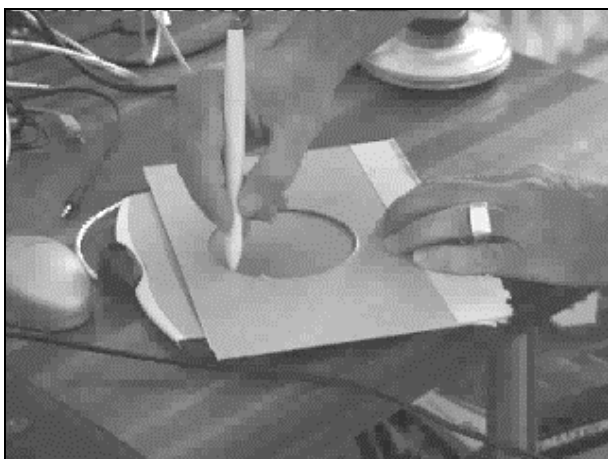


Figure 5. contrôle guidé d'un paradoxe de hauteur

La plupart du temps, on réalise une fonction de correspondance ("mapping" en anglais) entre cette ou ces valeurs d'entrée et les paramètres de synthèse du son. On peut s'aider, et cela semble la tendance forte du moment dans le contrôle gestuel, d'une couche de paramètres qu'on peut appeler psycho-acoustiques, puisqu'ils décrivent des phénomènes identifiables de l'écoute, qui peuvent être mis en relation à la fois avec les paramètres du geste et ceux du son. Ces paramètres intermédiaires, perceptifs, sont aussi un tableau de bord intéressant pour le feedback visuel du musicien.

Dans l'exemple de la distorsion non-linéaire, deux paramètres ont été "mappés" sur les coordonnées de deux baguettes d'un instrument appelé le Radio-Baton (fig. 4) ; ce sont l'index de distorsion et la fréquence des battements de l'effet choral provoqué par des fréquences légèrement différentes.

Dans l'exemple du paradoxe de hauteur, un guide de carton permet à l'utilisateur d'utiliser simplement la coordonnée angulaire d'un cercle comme témoin de la fréquence du glissement (fig. 5). A chaque tour le son a

glissé d'une octave, et le sens de rotation et sa vitesse permet de contrôler la vitesse et le sens du glissement. Une version utilisant des bruits filtrés en lieu et place de sinusoïdes a aussi été réalisé, et sa conduite est soit guidée soit libre.

Toutefois les gestes les plus intéressants associent à la fois un geste de décision ("trigger") et un geste de modulation : c'est comme si l'effet de surprise de l'initialisation du son devait se poursuivre par une possibilité de modelage temps-réel du corps du son qui lui fait suite. Cette piste concerne en plein l'interprétation en musique numérique.

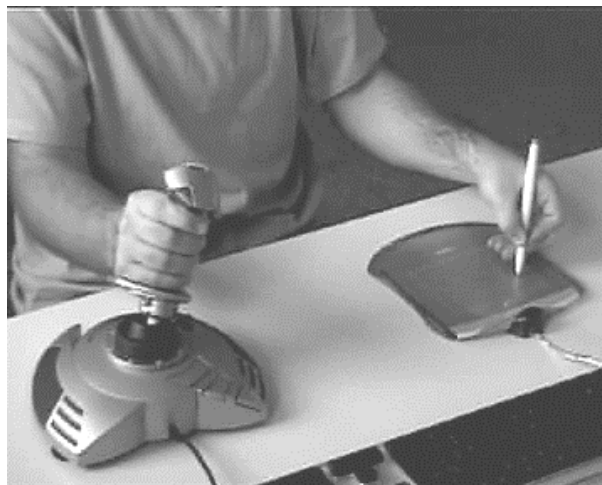


Figure 6. contrôle gestuel du « voicer »

Des instruments comme le "voicer" [3] ou la "synthèse par balayage" [4] exploitent cette combinaison. Le premier permet avec différents gestes (fig. 6) le déclenchement et l'arrêt du son d'une part et les changements de hauteur et de filtrage, (ce qui permet le passage d'une voyelle à une autre) d'autre part. Le second propose le contrôle gestuel d'un système dynamique lent dont les mouvements engendrent des variations dans le timbre du son (fig. 7). Après la mise en mouvement du système (geste de décision), le musicien peut contrôler l'évolution du timbre du son par des gestes reliés aux paramètres du système dynamique.



Figure 7. contrôle de la synthèse par balayage

III.2. les stratégies du geste

Afin de mieux faire comprendre que ces stratégies de correspondance entre son et geste ne sont pas uniques, nous prendrons l'exemple simple du vibrato, et allons voir deux stratégies radicalement différentes.

La première est d'associer la fréquence d'un son à une seule donnée, par exemple la distance de la main d'un capteur de distance. C'est exactement ce qui est le cas du Teremine, et dans ce cas c'est le tremblement de la main qui fait apparaître et disparaître le vibrato.

Mais il est aussi possible d'associer la fréquence et la profondeur du vibrato aux deux coordonnées d'un stylet ou d'une souris, auquel cas le geste est très différent : nous sommes désormais dans un espace de contrôle du timbre et non plus dans un espace gestuel de la fréquence.

Toute cette expérimentation sur le geste lié à la synthèse nous conduit maintenant à une réflexion sur la notion d'interprétation en musique numérique. Plusieurs pistes d'interprétation peuvent se dessiner :

- on peut vouloir uniquement moduler la sonorité des instruments de synthèse, un peu comme un contrôle de la couleur d'un son ou d'une séquence.

- on peut vouloir intervenir sur le déroulement temporel d'une séquence, c'est le rôle dévolu usuellement au tempo, aux retards dans l'exécution par un interprète. On peut noter ici qu'une préoccupation peut être non seulement sur l'activation des événements mais aussi de ce qui se passe au moment de leur désactivation.

- on peut aussi agir sur la modification du son à l'aide de modules de transformation sonore (plug-ins ou modules d'effets audio numériques écrites en Max-Msp)

- on peut aussi vouloir agir plus radicalement sur le macro-contrôle de séquences d'événements : ceci est le cas notamment en synthèse granulaire, où les densités de probabilités conduisent à des séquencements de micro-événements différents.

La démarche que nous avons est différente de celle utilisée habituellement en contrôle gestuel : usuellement on part des capteurs de geste, et on cherche des sons "qui vont avec". Ici on part du son, et on va vers le geste. Ce qui veut dire que partant d'une séquence sonore, on va trouver différents scénarios pour la mettre en geste, donc lui donner la possibilité d'être interprétée ou même improvisée. Ces scénarios peuvent être très variés. L'essentiel est de parvenir à un geste sensible et un son qui répond à ce geste.

Cela n'est pas un problème trivial, car il touche aussi au dilemme entre virtuosité et apprentissage. Les instruments traditionnels utilisent souvent les limites du geste humain : ils ne sont finalement pas si conviviaux, et même déforment parfois la physiologie humaine, mais ils permettent une virtuosité recherchée, fut-elle au prix d'un long apprentissage. L'inverse serait la machine sans apprentissage qui à partir d'un geste immature

produirait cependant une interprétation de rêve. La gestuelle informatique doit souvent se contenter d'être médiane dans ce débat entre apprentissage et virtuosité : on cherche à être jouable rapidement (des années d'apprentissage ne conviennent pas) mais aussi à disposer d'une virtuosité suffisante.

IV. CONCLUSION

Il est évident que la principale application de telles recherches est la constitution d'un instrumentarium numérique, permettant de jouer des instruments dans une configuration de musique moderne. En ce sens trois instruments complets sont déjà issus de cette recherche, qui tous trois utilisent la commande: le voicer permet une synthèse vocale [3], la synthèse par balayage permet d'utiliser une dynamique du geste [4], et l'instrument photosonique numérique qui permet d'émuler le comportement d'un instrument de musique optique [5]. Mais cette recherche a aussi trait à un domaine peu connu du dialogue homme-machine : celui de la Présence, et en particulier de l'ergonomie des systèmes conviviaux. Un instrument électronique se doit d'être le fidèle compagnon des instruments acoustiques, et représenter un niveau supérieur des possibilités de l'être humain de s'exprimer, au travers d'interfaces spécialement conçues pour lui donner l'impression d'être au cœur du son.

REMERCIEMENTS

Cette recherche a été aidée par le CNRS et le Conseil Général des Bouches du Rhône sous le titre "le Geste Créatif en Informatique Musicale"

RÉFÉRENCES

1. Boulanger Richard : *The Csound Book, Perspectives in Software Synthesis, Sound Design, Signal Processing and Programming*, MIT press, 2000
2. D. Arfib, L. Kessous, : *Gestural control of sound synthesis and processing algorithms*, "Gesture workshop 2001", ed. Ipke Wachsmuth & Timo Sowa, Springer-Verlag, Lecture Notes in Computer Science, LNAI 2298
3. L.Kessous, *A two-handed controller with angular fundamental frequency control and sound color navigation*, Proceedings of the 2002 Conference on New Instruments for Musical Expression (NIME-02), Dublin, Ireland, May 24-26, 2002
4. J.M. Couturier, *A scanned synthesis scanned instrument*, Proceedings of the 2002 Conference on New Instruments for Musical Expression (NIME-02), Dublin, Ireland, May 24-26, 2002
5. D.Arfib, J. Dudon : *A digital emulator of the photosonic instrument*, Proceedings of the 2002 Conference on New Instruments for Musical Expression (NIME-02), Dublin, Ireland, May 24-26, 2002
6. CD-ROM *"Trends in Gestural control of music"*, éditeurs M. Wanderley & M. Battier, publication Ircam, 2000
7. *Interactive Systems/Instrument Design in Music* www.notam02.no/icma/interactivesystems/wg.html