

ESPACES INTERACTIFS VISUELS POUR LE CONTROLE DES SONS MUSICAUX

Jean-Michel Couturier

LMA-CNRS, 31, chemin Joseph Aiguier 13402 Marseille cedex 20

couturier@lma.cnrs-mrs.fr

Résumé

Cet article montre comment l'interaction gestuelle avec des interfaces graphiques peut rendre performant le contrôle des sons musicaux. Il expose auparavant la problématique du contrôle gestuel des processus sonores dans les instruments de musique informatiques. Un exemple de manipulation d'une corde virtuelle avec un écran tactile est donné en fin d'article.

INTRODUCTION

Les développements de l'informatique musicale ont apporté un renouveau dans la création d'instrument de musique, en séparant la production du son des gestes du musicien. La liberté qui en résulte possède cependant un inconvénient : la difficulté à contrôler efficacement les processus sonores. Les créateurs d'instruments de musique basés sur l'informatique cherchent à recréer le lien entre le geste et le son. Ce lien, selon la démarche utilisée, peut être direct ou peut passer par la manipulation d'objets d'interaction virtuels représentés visuellement.

Après un exposé de la problématique du contrôle instrumental en informatique musicale, nous étudierons ces deux démarches de la création d'instruments numériques, en développant plus particulièrement la seconde. Enfin, nous verrons comment le canal visuel peut être utilisé pour contrôler de façon originale la synthèse par balayage, grâce à une interaction directe entre l'utilisateur et une corde virtuelle.

1 CONTROLE INSTRUMENTAL EN INFORMATIQUE MUSICALE

Dans l'acoustique instrumentale, la production de sons musicaux est associée à des mouvements humains. Il y a une causalité physique entre les gestes effectués et les sons produits. L'arrivée des sons de synthèse et de l'informatique ont bouleversé cet état de fait : on peut dorénavant produire des sons, de la musique sans tenir compte des mouvements humains [1].

Depuis plusieurs années, les ordinateurs ont une puissance suffisante pour permettre de créer et de contrôler des sons de synthèse en temps réel. On peut donc effectuer un contrôle gestuel de ces sons numériques, donnant au musicien les moyens de s'exprimer comme il le fait avec les instruments traditionnels. Les instruments acoustiques traditionnels offrent de grandes possibilités de jeu, mais possèdent des limites dues aux contraintes physiques liées à l'instrument et à l'interprète, obligeant le plus souvent l'instrumentiste à s'adapter à l'instrument. Avec les systèmes électroniques et numériques, les contraintes physiques liées à l'instrument et à la production du son disparaissent, mais le lien entre le geste et le son est perdu ; ce lien qui était naturel est à réinventer.

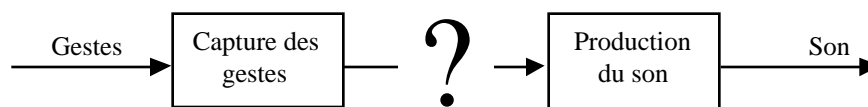


Figure 1 : Contrairement aux instruments acoustiques, le son issu des instruments électroniques et numériques ne dépend plus directement des gestes effectués. Il faut reconstruire la chaîne qui lie le geste et le son, ce qui tout en offrant une grande liberté pose de nouveaux problèmes.

Avec l'évolution des technologies, une quantité importante de contrôleurs gestuels a vu le jour et beaucoup de recherches ont été effectuées sur les contrôleurs et leurs liens avec les processus sonores. La correspondance entre ces périphériques et les paramètres de contrôle des sons de synthèse utilisés est couramment appelée *mapping*. Le

mapping est de toute première importance : il définit la manière dont les gestes exécutés par le musicien vont faire varier les paramètres de synthèse. Il n'est pas évident de le réaliser car si le geste est très fortement lié au son pour les instruments acoustiques, associer un geste adapté à un son de synthèse n'est pas simple. En effet, les paramètres de synthèse dépendent de l'algorithme de synthèse utilisé et ne sont pas toujours adaptés à être directement liés à un geste. Ces difficultés font que le *mapping* est une problématique forte de la recherche en informatique musicale [2] [3] [4].

La recherche sur le *mapping* doit prendre en compte le problème de l'interprétation en musique numérique, où apparaît le dilemme entre virtuosité et apprentissage [5]. Les instruments traditionnels demandent souvent un long apprentissage, mais ils permettent une grande virtuosité. L'opposé serait un dispositif qui, sans apprentissage, permettrait une interprétation parfaite. Cependant, une telle machine n'existe pas. Les recherches en informatique musicale et sur le *mapping* doivent trouver des compromis pour obtenir une virtuosité suffisante avec un temps d'apprentissage court. Ceci est d'autant plus vrai que la pérennité des nouveaux systèmes technologiques est si faible qu'il serait difficile de créer des instruments qui auraient un temps d'apprentissage aussi long que celui des instruments acoustiques traditionnels.

Alors, comment créer des interfaces homme-machine efficaces utilisant des mappings intelligents pouvant aller au-delà de ces problématiques ? Nous allons voir dans la partie suivante comment l'utilisation de systèmes multimodaux avec interfaces graphiques peut aider à la création d'instruments de musique performants qui prennent en compte les contraintes liées aux instruments numériques.

2 DES ESPACES VIRTUELS POUR CONTROLER LES PROCESSUS SONORES

Par construction, les instruments acoustiques sont des systèmes multimodaux : Les gestes du musicien mettent en œuvre des processus physiques qui produisent le son, et l'audition de ce son ainsi que le contact physique de l'instrument et sa visualisation vont aider l'instrumentiste dans son action (Fig. 1). Les instruments numériques cherchent à garder cette interaction, à ré-établir le lien entre le geste, le son créé et les divers retours (sonores, tactiles et visuels). Comme nous le verrons ensuite, différentes approches sont possibles pour concevoir ces instruments.

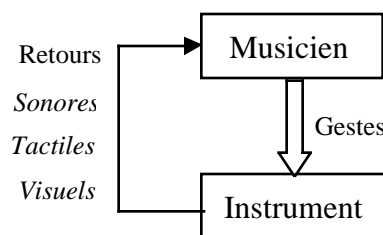


Figure 2 : Interaction entre l'instrumentiste et son instrument. Cette interaction est figée pour les instruments acoustiques et dépend de la physique de l'instrument et de son mode de jeu.

La première démarche dans l'élaboration d'instruments de musique numériques est de chercher à retrouver l'interaction instrumentale telle qu'elle est présente dans les instruments acoustiques. Cette démarche vise à associer des dispositifs permettant de capter les gestes du musicien et de les relier à des dispositifs produisant du son. Cela permet de créer un ensemble capteurs gestuels / système de production du signal sonore / haut-parleurs que l'on peut identifier comme un instrument de musique à part entière. La liberté offerte par la rupture entre le geste et le son est utilisée pendant la phase de conception de l'instrument. Une fois l'instrument créé, son fonctionnement est proche des instruments traditionnels, avec quelques fonctionnalités supplémentaires (réglages de l'instrument programmables, retours visuels, etc...). Ces instruments utilisent des capteurs gestuels plus ou moins proches des instruments acoustiques [6]. Les contrôleurs gestuels les plus proches des instruments réels sont ceux qui les imitent (comme les claviers et les contrôleurs de souffle), permettant ainsi de capter les gestes experts que les instrumentistes ont l'habitude d'effectuer. Ces capteurs peuvent servir à contrôler les mêmes changements sonores que leurs originaux : par exemple, on retrouve un clavier pour le contrôle de la hauteur et de l'amplitude dans la plupart des synthétiseurs commerciaux. Ils peuvent également quitter leur utilisation habituelle et manipuler d'autres paramètres du son, notamment le timbre. D'autres recherches ont été menées pour permettre la saisie de gestes musicaux différents et la création de nouvelles situations musicales. Ces recherches visent à développer de nouveaux périphériques ou à détourner des capteurs existants pour saisir les gestes des musiciens. Ces périphériques alternatifs, qui vont des simples curseurs et boutons rotatifs à des capteurs complexes comme

les caméras vidéos ou les systèmes à retour d'effort [7], permettent aux musiciens d'explorer de nouveaux modes d'expression artistique.

Une seconde démarche cherche à utiliser les avantages liés aux systèmes informatiques et aux espaces virtuels pour contrôler les sons musicaux. La prolifération des logiciels avec interfaces graphiques a donné à la vision, à travers l'écran d'ordinateur, une place prépondérante [8]. Si la vision est indispensable quand nous utilisons des programmes informatiques, il ne l'est pas pour le contrôle des sons musicaux (dans les formations classiques, la vue est souvent dédiée à la lecture des partitions). Inversement, le rôle du sonore en informatique est souvent limité. Cet état de fait ouvre de très grandes possibilités pour un contrôle des sons musicaux basé sur l'interaction avec des phénomènes perçus au travers du canal visuel.

Les logiciels actuels permettent grâce à un simple ordinateur personnel d'utiliser des modèles informatiques d'instruments et de matériel musical avec un encombrement réduit et sans avoir à acheter le matériel ou les instruments réels correspondants. De la même façon que les systèmes informatiques représentent des objets réels (comme la corbeille ou les dossiers) pour une plus grande convivialité des interfaces, une grande partie des logiciels musicaux utilisent des objets graphiques, en général bidimensionnels, simulant des contrôles réels : claviers, potentiomètres linéaires et rotatifs, boutons et molettes de commande, manettes 2D, etc.

Cependant, la manipulation de ces objets se fait généralement à travers la souris, qui, si elle est adaptée aux utilisations en bureau, est inadaptée au contrôle de sons en temps réel. La souris permet de manipuler seulement deux paramètres continus à la fois, dans un espace bidimensionnel, et un ou deux paramètres discrets, correspondant aux *clics* de souris. De plus, elle n'est pas le contrôleur idéal pour rendre compte de notre habileté manuelle. La tablette graphique, adoptée par un grand nombre de graphistes, est plus adaptée que la souris pour les manipulations musicales, mais possède elle aussi des limites en ce qui concerne le nombre de paramètres manipulables simultanément. Pour ces raisons, on ne peut pas manipuler les éléments proposés par ces logiciels de la même façon que leurs originaux réels, qui peuvent être tous utilisés en même temps : il faut souvent utiliser des contrôleurs externes. Mais dans ce cas, les contrôleurs font disparaître les avantages (flexibilité et immédiateté) du virtuel en figeant des paramètres sur des contrôleurs, les interfaces graphiques permettant le contrôle d'un nombre de paramètre illimité. De plus, brancher et configurer un périphérique externe est souvent difficile, prend du temps et ne permet pas de manipuler la totalité des paramètres du logiciel. Du coup, le plus souvent on est obligé d'avoir simultanément deux modes d'utilisation : préparation du logiciel à la souris et jeu instrumental avec des périphériques adaptés. Ceci est d'autant plus dommage que l'on trouve dans le commerce des versions logicielles d'algorithmes de synthèse et de traitement audio souvent aussi performantes du point de vue traitement du signal que les versions matérielles, avec des interfaces claires et lisibles. Pour palier à ces problèmes, les fabricants de logiciels développent parfois des « télécommandes » ou des périphériques spécifiques à leurs logiciels, comme des tables de commandes identiques aux tables de mixages. L'utilisateur retrouve ainsi le confort d'utilisation qui existe avec le matériel réel, faisant disparaître les limites de la souris, même si elle reste indispensable pour l'utilisation du logiciel.



Figure 3 : Contrôleur / carte audionumérique externe Digi 002 et Interface logicielle associée (Protools). Les curseurs du périphérique externe font varier les curseurs représentés à l'écran et effectuent les traitements sonores associés. Inversement, un déplacement des curseurs situés dans le logiciel va entraîner un déplacement des curseurs réels (avec la permission de Digidesign).

Les périphériques usuels posent un autre problème : la dissociation entre le manipulé et le manipulant. La manipulation se fait dans un espace défini, celui de la souris ou des contrôleurs externes, et les résultats de cette manipulation apparaissent dans un autre espace, l'espace virtuel représenté à l'écran. Cette dissociation est courante dans notre société (conduite d'une automobile, manipulation de télécommandes), mais dans le cas où on représente à l'écran des objets familiers que l'on toucherait directement s'il étaient réels, cette dissociation introduit une gêne : le lieu du geste n'est pas le lieu de l'interaction.

Des solutions existent, mais sont peu utilisées : ce sont les écrans tactiles et systèmes équivalents, qui associent le retour visuel au contrôle : l'utilisateur voit le geste qu'il fait et l'effet de ce geste. Ces systèmes procurent davantage de présence dans la manipulation ; on se retrouve dans une situation plus proche de la réalité : le doigt manipule directement un objet graphique représentant l'objet réel. De plus, ces systèmes conservent l'avantage de la souris sur des contrôleurs externes : il n'est pas nécessaire d'assigner les objets graphiques au contrôleur, l'interaction se fait directement avec l'objet graphique. Cependant, ils n'offrent pas la même sensation de toucher que les systèmes réels.

Enfin, nous avons deux mains : lorsque nous jouons de la musique avec des instruments acoustiques ou électroniques (comme les synthétiseurs analogiques, avec leurs nombreux potentiomètres), nous nous servons de ces deux mains [9]. La souris et les écrans tactiles usuels n'utilisent qu'une seule main [10], la deuxième main servant à presser le clavier (shift, control,...). Cela représente une importante limite pour l'expression musicale : les gestes musicaux usuels sont complexes et ne se limitent pas au déplacement d'un petit objet sur une surface, ou d'un seul doigt sur un écran. Ce problème est de taille, car si des interfaces spécifiques permettent de modifier plusieurs paramètres en même temps, les systèmes d'exploitation courants de l'informatique actuelle fonctionnent avec un seul curseur et permettent la manipulation d'un seul objet à la fois. Les périphériques conçus pour remplacer la souris ont la même limite, leur rôle étant de déplacer un curseur.

Parmi les personnes travaillant sur ce problème, on peut citer le groupe de recherche sur les « Interfaces Tangibles » du *MIT Media Lab*, qui a été créé afin de concevoir des interfaces avec interaction entre des objets réels et un environnement virtuel [11]. Ce groupe a notamment développé un système spécifique d'interaction musicale, appelé *AudioPad*, présenté aux *NIME 2002* [12] (Fig. 4).

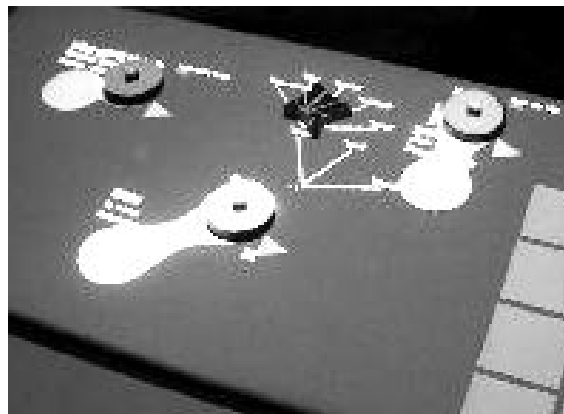


Figure 4 : Le système Audiopad est composé de cylindres et autres objets se déplaçant sur une surface plane sur laquelle est projetée la représentation d'un espace d'interaction virtuel. Il est basé sur des interactions bimanuelles combinant le choix des paramètres et leur modification, à travers la navigation dans un espace virtuel.

Ce système donne une idée de la puissance que peut avoir l'utilisation d'interactions avec des interfaces graphiques dans le contrôle simultané de paramètres multiples. Contrairement aux instruments traditionnels qui sont figés et dont la forme est connue, ici l'instrument évolue, les paramètres à modifier sont immédiatement disponibles et les objets d'interaction peuvent changer. Au-delà de la manipulation de paramètres à travers la représentation d'objets réels, d'immenses possibilités s'ouvrent, notamment l'utilisation de métaphores pour la représentation des sons dans le domaine visuel (cross-modalité) ou l'interaction avec des images en mouvement. La partie suivante expose un exemple de contrôle d'un algorithme de synthèse par une interaction directe avec une forme virtuelle.

3 CONTROLE VISUEL DIRECT DE LA SYNTHÈSE PAR BALAYAGE

La synthèse par balayage (en anglais *Scanned Synthesis*) a été développée par Bill Verplank, Rob Shaw et Max Mathews [13] et permet de générer des sons à partir de mouvements lents de systèmes mécaniques. Cette technique de synthèse génère une table d'onde dynamique et la lit pour produire des sons. Elle se présente de la façon suivante (Fig. 5) :

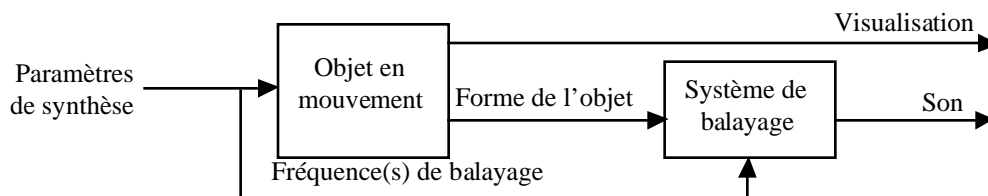


Figure 5 : Principe : La synthèse par balayage est composée d'un système dynamique lent, dont les fréquences de vibrations sont situées au-dessous de 15 Hz. Pour produire des fréquences audibles, la forme du système dynamique est balayée périodiquement à une fréquence correspondant à la fondamentale du son produit. Ce modèle permet de contrôler indépendamment le timbre et la fréquence fondamentale du son.

Les bases psychoacoustiques de la synthèse par balayage viennent de notre connaissance sur la perception auditive humaine et sur les capacités du contrôle corporel humain. Dans les années 60, Jean-Claude Risset [14] a montré que pour qu'un timbre soit intéressant il faut que son spectre change au cours du temps. Il se trouve que pour obtenir des timbres musicalement intéressants, ces variations doivent se situer sous les 15 Hz, fréquences auxquelles le corps humain peut se mouvoir. Les vitesses d'évolution de l'objet se situent dans cette zone. De plus ces vitesses sont suffisamment faibles pour que l'on puisse visualiser les mouvements de l'objet, ce qui offre la possibilité d'utiliser l'image de l'objet comme retour visuel et comme moyen de contrôle de l'algorithme.

J'ai implémenté cette technique en utilisant comme objet en mouvement un modèle physique de corde circulaire en différences finies (une collection de masses reliées par des ressorts et des amortisseurs). Un objet en C proposant un contrôle de haut niveau de la synthèse par balayage a été créé pour le logiciel Max-Msp sur Macintosh [15] et a été utilisé pour la réalisation d'un instrument complet qui fut présenté aux *NIME 2002* à Dublin [16].

Dans ces travaux, la corde était mise en mouvement par des forces ou lâchée depuis une position initiale fixe. La forme de la corde était utilisée uniquement comme retour visuel. Cependant, le lien entre le son et sa représentation visuelle est ici direct : on peut suivre l'évolution de la forme de la corde à l'œil nu, ce qui donne envie de modifier cette forme « à la main » et en temps réel. J'ai donc ensuite créé un dispositif permettant de manipuler directement la forme de la corde, comme on le ferait si elle était réelle. Le système utilise un écran tactile affichant les mouvements de la corde. L'utilisateur a la possibilité d'interagir directement avec cette corde en déplaçant son doigt sur l'écran. De cette façon, on peut interagir en temps réel avec la corde : la mettre en mouvement, lui donner une trajectoire, créer un obstacle sur son trajet, ...

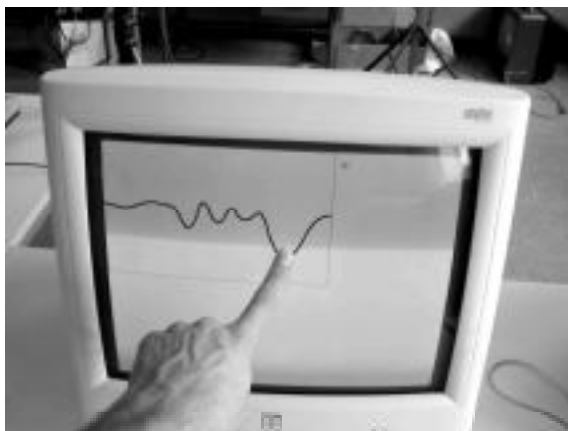


Figure 6 : Manipulation de la corde virtuelle de l'algorithme de synthèse par balayage à l'aide d'un écran tactile. Le principe de l'interaction est le suivant : Lorsque le doigt rencontre la corde, le programme agit de façon à ce que la corde ne traverse pas le doigt mais suive ses mouvements. Quand un doigt est posé sur l'écran, le programme détermine s'il est posé au-dessus ou au-dessous de la corde. En fonction de l'endroit où l'on a placé initialement son doigt, il faut appuyer sur la corde (en déplaçant le doigt vers le bas) ou la soulever (en déplaçant le doigt vers le haut) pour interagir avec elle.

Compte tenu de la complexité de la synthèse par balayage et du nombre de paramètres de contrôle possibles, ce dispositif permet d'utiliser cette technique de synthèse avec une aisance et une présence remarquable. Il offre une interaction proche de celle que l'on aurait si la corde était réelle, les seuls éléments ne pouvant pas être restitués étant l'inertie de la corde ainsi que la sensation physique du contact.

CONCLUSION

L'utilisation d'interactions directes avec des représentations visuelles offre un avantage non négligeable au contrôle des processus sonores. Cependant, on se rend compte que si les interfaces graphiques sont très répandues en informatique, les objets d'interaction qu'elles représentent visuellement sont difficiles à manipuler dans un contexte de jeu musical. Les travaux réalisés sur ce sujet sont prometteurs et donnent une idée sur l'étendue des possibilités qu'offre l'utilisation de ce type d'interaction.

Mes recherches futures iront vers la création d'espaces virtuels de contrôle des sons musicaux et vers la réalisation de périphériques plus efficaces pour utiliser ces espaces.

4 REFERENCES

- [1] I. Choi, "Gestural Primitives and the context for computational processing in an interactive performance system", CD-room *Trends in Gestural Control of Music*, éditeurs M. Wanderley and M. Battier, publication Ircam, 2000.
- [2] Wanderley M., Depalle P., "Contrôle Gestuel de la Synthèse Sonore", *Interfaces homme-machine et création musicale*, éditeurs H. Vinet et F. Delalande, pp. 145-163. Hermès, 1999.
- [3] A. Hunt, M. Wanderley, R. Kirk, "Towards a Model for Instrumental Mapping in Expert Musical Interaction", Actes de la *International Computer Music Conference ICMC'2000*, Berlin, Allemagne, ICMA, pp. 209-212.
- [4] D. Arfib, J.M. Couturier, L. Kessous, V. Verfaillie, " Mapping strategies between gesture control parameters and synthesis models parameters using perceptual spaces", à paraître dans *Organised Sound*.
- [5] D. Arfib, L. Kessous, "from Music V to creative gesture in computer music", Actes de la *VIIIth SBC Conference*, Curitiba, 2000.
- [6] J.C. Risset, "Nouveaux gestes musicaux : quelques repères historiques", *Les Nouveaux Gestes de la Musique*, ed. H. Genevois et R. de Vivo, pp.19-33, Parenthèses, 1999.
- [7] C. Cadoz, A. Luciani and J.-L. Florens, "Synthèse musicale par simulation des mécanismes instrumentaux, transducteurs gestuels rétroactifs pour l'étude du jeu instrumental", *Revue d'Acoustique*, vol. 59, pp. 279-292, 1981.
- [8] H. Vinet, "Concept d'interfaces graphiques pour la production musicale et sonore", *Interfaces homme-machine et création musicale*, éditeurs H. Vinet et F. Delalande, pp. 97-121. Hermès, 1999.
- [9] L. Kessous, " Instruments bi-manuels et espaces sonores", actes des *Journées d'Etudes Espaces Sonores*, Avril 2002.
- [10] M. Beaudouin-Lafon. "Moins d'interface pour plus d'interaction", *Interfaces homme-machine et création musicale*, éditeurs H. Vinet et F. Delalande, pp. 123-141. Hermès, 1999.
- [11] "Tangible Media Group", *MIT Media Lab*, <http://tangible.media.mit.edu/>.
- [12] J. Patten, B. Retch, H. Ishii, "Audiopad : A Tag-based Interface for Musical Performance", Actes de la conférence internationale *New Interfaces for Musical Expression (NIME-02)*, Dublin, Irlande, 24-26 Mai 2002.
- [13] B. Verplank, M. Mathews, R. Shaw, "Scanned Synthesis", Actes de la *2000 International Computer Music Conference*, Berlin, éditeur Zannos, ICMA, 2000, pp. 368-371.
- [14] J.C. Risset, "Analysis of instrumental tones", *Physics Today*, 22 (2), 1969, pp. 23-30.
- [15] J.M. Couturier, "La synthèse par balayage et son contrôle gestuel", Actes des *9^{èmes} Journées d'Informatique Musicale (JIM 02)*, Marseille, France, 29-31 Mai 2002.
http://perso.wanadoo.fr/gmem/evenements/jim2002/articles/P03_Couturier.pdf
- [16] J.M. Couturier, "A Scanned Synthesis Virtual Instrument", Actes de la conférence internationale *New Interfaces for Musical Expression (NIME-02)*, Dublin, Irlande, 24-26 Mai 2002.