

10ème Congrès Français d'Acoustique

Lyon, 12-16 Avril 2010

Contrôle gestuel du rayonnement acoustique de sons de synthèse

Hugues Genevois¹, Emilien Ghomi²

¹LAM – Institut Jean le Rond d'Alembert - UPMC, 11, rue de Lourmel, F-75015 Paris, genevois@lam.jussieu.fr

²LRI, Bât 490 Université Paris-Sud 11, F-91405 Orsay Cedex, emilien.ghomi@lri.fr

Outre leur grande richesse expressive, liée aux innombrables nuances de timbre et d'intensité dont le musicien sait tirer partie, les instruments traditionnels constituent aussi des sources acoustiques complexes. S'il est aujourd'hui possible de synthétiser et/ou de reproduire des timbres très variés, offrir la possibilité d'une expressivité instrumentale de qualité suppose non seulement de pouvoir contrôler précisément les processus de synthèse et de modification des sons, mais aussi leur mise en espace (localisation et rayonnement). La diffusion des musiques électroacoustiques s'est souvent appuyée sur des « orchestres » de haut-parleurs, tel l'Acousmonium du GRM et le Cybernophone de l'IMEB. Si de tels dispositifs permettent de créer des paysages acoustiques vivants et contrastés, ils sont peut-être moins adaptés à l'installation d'images sonores instrumentales. Aujourd'hui, en effet, les musiques électroacoustiques sont aussi jouées en direct, par des ensembles (orchestres de *laptops*, de joysticks, etc.). La cohérence et la lisibilité du jeu musical pose alors des problèmes nouveaux. Cela a notamment été mis en évidence par les fondateurs du *Princeton Laptop Orchestra* [1]. Dans le même temps, du côté des musiques amplifiées, des problèmes de même nature existent, et l'on peut penser que les systèmes développés par la firme Bose (*Cylindrical Radiator LI*) répondent à des besoins en partie comparables. Le dispositif présenté, proche de la *Timée* de l'IRCAM ou du *Spherical Loudspeaker Array* du CNMAT, vise à permettre au musicien de contrôler, dans et par le même geste expressif, la synthèse des sons et leur mise en espace.

1 Introduction

En inventant de nouveaux dispositifs de production musicale, le ^{XX}e siècle nous invite à revisiter le concept d'instrument de musique. De nombreux travaux, en musicologie [2,3], comme en informatique musicale [4] décrivent l'histoire de ces évolutions, les conditions techniques de leur apparition, et les formes originales "d'instrumentalité" liées à ces nouvelles lutheries. Dans le passage de l'instrument mécanique à l'instrument électrique, électronique et, plus encore, informatique, s'est opérée une rupture des chaînes causales traditionnelles qui garantissaient la cohérence temporelle, spatiale et énergétique des interrelations geste-instrument-son.

La synthèse sonore a fait des progrès considérables au cours des 50 dernières années ; ainsi, nous sommes aujourd'hui en mesure de synthétiser des signaux d'une grande variété, d'une grande richesse, soit pour simuler des instruments existants soit pour élargir la palette expressive à des timbres nouveaux. Pourtant, le travail du "luthier" en synthèse ne saurait s'arrêter à la génération des signaux, l'étape consistant à retraduire ces signaux calculés en sons audibles étant tout aussi capitale d'un point de vue perceptif et artistique. De précédents travaux, notamment à l'IRCAM et au CNMAT, ont montré l'intérêt de se préoccuper de la "re-matérialisation" des signaux, en particulier en offrant la possibilité de contrôler le rayonnement des sources acoustiques.

Notre présentation s'attachera, dans le droit fil de ces travaux ultérieurs, à proposer de nouvelles façons de faire évoluer les champs acoustiques synthétisés, liées au jeu instrumental. Si personne n'est en mesure de définir précisément ce qu'est ou n'est pas un instrument de musique [5], nous pouvons en revanche interroger la notion d'instrumentalité.

L'invention du téléphone et celle du phonographe, presque contemporaines l'une de l'autre, ont, en permettant la transmission des sons à distance et leur enregistrement perturbé notre relation "classique" aux sons, à la voix, à la musique. Désormais, les sons peuvent "voyager" dans l'espace et dans le temps. Ils ne sont plus nécessairement entendus à l'endroit et à l'instant de leur production. Les développements de l'électricité, puis du numérique, ont plus tard contribué à d'autres modes de "mise à distance" du corps dans la production musicale.

Les conséquences de ces découplages sur la relation musicien-instrument ont déjà fait l'objet d'études approfondies, et nous ne reviendrons pas sur ces points, sinon pour noter que ces technologies font passer d'un monde où la valeur est créée par le travail, l'énergie, à un monde où la valeur est liée à l'information, au symbolique, révolutions dont les conséquences artistiques, culturelles et sociales sont inévitables.

Mais revenons à la question de l'instrument et, pour cela, empruntons un instant à un autre domaine que celui de la musique des concepts susceptibles d'aider notre réflexion.

Le théâtre antique reposait sur la règle des trois unités : unités de temps, de lieu et d'action. Dans une certaine mesure, la relation "classique" du musicien ou de l'auditeur avec les phénomènes sonores reposait sur les mêmes bases. Or, nous l'avons vu précédemment, ces trois unités, qui garantissaient la cohérence de la relation instrumentale, se trouvent brisées par les technologies nouvelles. En conséquence, là où le musicien et l'auditeur étaient en relation directe avec les sources sonores "primaires", ils se voient aujourd'hui, avec les dispositifs de synthèse, confrontés à des sources "secondaires" produisant des simulacres dont rien ne garantit plus une quelconque cohérence temporelle, spatiale ou énergétique.

2 Synthèse et contrôle gestuel

Dès les débuts de la synthèse, certains inventeurs ont souhaité mettre l'accent sur les **qualités expressives** de leur instrument et leur **capacité à s'insérer acoustiquement** parmi les instruments de l'orchestre. C'est le cas notamment de l'Onde Martenot, dont il n'est pas interdit de penser que le succès, la longévité, et l'important répertoire sont en partie liés à ces caractéristiques.



Figure 1 : Ondes Martenot et diffuseurs

Nombreux sont aujourd'hui les travaux visant à redonner au musicien la possibilité de contrôler en temps-réel le processus de génération sonore. Parmi ceux-ci, nous avons principalement travaillé avec les trois suivants :

- le Méta-Instrument [6], conçu par Serge de Laubier et développé au sein de Puce Muse ;
- un joystick, dans le cadre d'un travail visant à développer des pratiques musicales collectives, donc nécessitant des interfaces gestuelles bon marché [7] ;
- une tablette graphique.

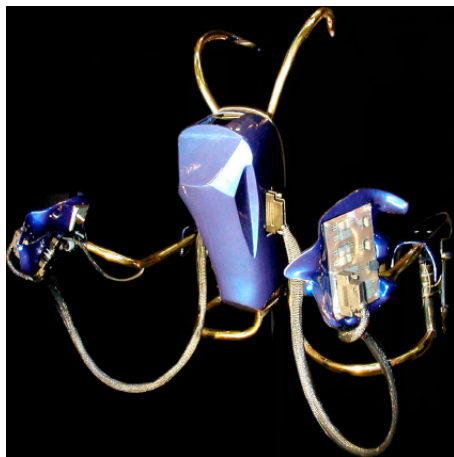


Figure 2 : le Méta-Instrument de 3° génération

Sur la base de ces trois types d'interfaces très différentes, sur le plan technique (nombre et qualité des capteurs, types de gestes) comme au niveau des prix, nous avons développé et expérimenté un concept de "modèle intermédiaire" destiné à piloter les processus de synthèse et de rendu 3D (sonores et visuels) à partir de bibliothèques de comportements. Ce travail a pu démarrer grâce au soutien de l'ANR (projet 2PIM) et se poursuivra dans le cadre du projet OrJo, soutenu par le FEDER et la Région Ile-de-France.

3 La notion de modèle intermédiaire

Nous l'avons rapidement évoqué en introduction : avec l'invention de la synthèse est née une possibilité de découplage total entre l'énergie produite par le musicien, et l'énergie acoustique en sortie du dispositif sonore.

Ce découplage, s'il peut poser des problèmes de cohérence lorsque l'on cherche à établir une interrelation "instrumentale" avec un dispositif, présente en même temps deux intérêts : il permet de définir des systèmes qui n'existeraient pas dans la nature, et d'imposer au système une évolution dans le temps qui peut changer radicalement sa réaction à une excitation.

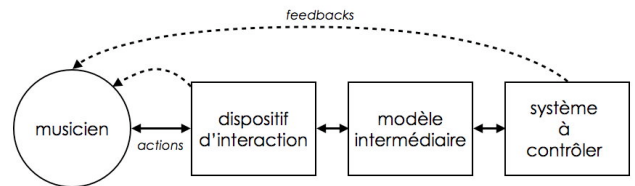


Figure 3 : du musicien à la synthèse

À partir de cette idée, nous décidons de regrouper ces systèmes complexes servant d'intermédiaires entre le musicien et le son produit sous le terme de "modèles intermédiaires". Un modèle intermédiaire est donc un système régi par un ensemble de lois définissant sa réaction à des excitations, et qui présente des éléments variables et des états accessibles [8].

L'utilisation d'un tel modèle implique le déploiement de ce que l'on nomme "mapping" et qui n'est autre qu'une mise en relation des gestes effectués par le musicien avec les paramètres du dispositif sonore.

Bien que la création musicale soit un domaine moteur pour ce type de recherches, le problème ne tarda pas à se poser de la catégorisation des opérations mathématiques effectuées entre le paramètre d'entrée et le paramètre du système. S'agit-il d'un modèle à part entière, d'une seule équation faisant intervenir plusieurs paramètres à la fois, d'une mise à l'échelle...? Effectivement, on remarque dans des articles comme celui de Momeni et Henry [9] que la distinction n'est finalement pas triviale entre mapping, analyse, modification d'échelle, modèle...

L'introduction du concept de modèle intermédiaire permet la création d'une architecture de développement s'appuyant sur une véritable séparation des différents étages de la chaîne de production audiovisuelle.

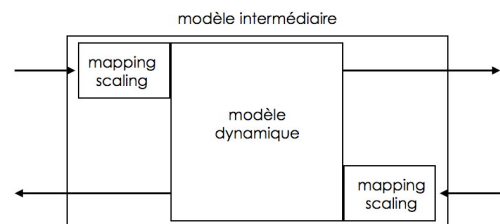


Figure 4 : schéma général d'un modèle intermédiaire

Dans l'architecture logicielle proposée, qui sert de support au développement de modèles intermédiaires indépendants a priori des systèmes à contrôler (synthèse audiovisuelle, autre modèle intermédiaire...), les couches

de mappings, permettant l'affectation des flux de données et leurs mises à l'échelle, sont décrites par des matrices.

Cette architecture permet une grande modularité, plus générale encore que celle offerte par les synthétiseurs modulaires. En outre, elle tend à gommer la distinction, souvent problématique, entre données de contrôle (par exemple gestuelles) et signaux (par exemple audio).

Par ailleurs, le même modèle intermédiaire peut servir à contrôler l'évolution de tous les paramètres de synthèse, y compris ceux qui conditionnent le rendu sonore (dynamiques, spatialités, etc.). Le même geste va ainsi permettre au musicien d'interagir globalement sur le résultat sonore, c'est-à-dire non les signaux calculés par l'ordinateur, mais les ondes acoustiques produites par l'ensemble du dispositif.

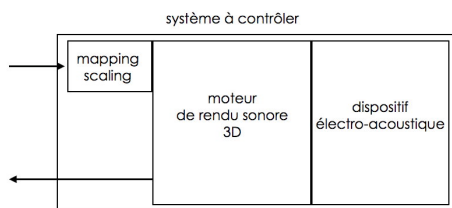


Figure 5 : synthèse et rendu sonore

De même que l'on parle de moteur de rendu en synthèse d'images, nous parlerons de moteur de rendu pour la synthèse sonore. Dans cette architecture, c'est le modèle intermédiaire qui est au cœur de l'interaction homme-machine [10]. C'est lui qui détermine en grande partie les caractéristiques instrumentales du dispositif et est chargé d'en assurer la cohérence.

Ce souci de cohérence (spatio-temporelle et dynamique) est d'autant plus nécessaire dans le cas du jeu collectif et, en particulier, des orchestres électroniques (Princeton Laptop Orchestra, Méta-Orchestre, etc.). [11]

En outre, cette approche permet de tendre vers une instrumentalité retrouvée, car elle apporte de nombreuses caractéristiques indispensables à un instrument de musique, comme par exemple une certaine perception des réactions du modèle, qui imposent des contraintes pouvant délimiter des modes de jeu et donner naissance à une liberté d'expression.

Les modèles de comportement sont variés, et leur étude n'a pas fait l'objet d'une réelle classification. En consultant diverses publications, on retiendra pour notre part :

- les modèles dits "physiques" par abus de langage, qui sont en général basés sur des liens masse-ressort issus du modèle masse-interaction qui a été rendu accessible en informatique par la formalisation de Cadoz [4] ;
- la "vraie" classe des modèles physiques qui peuvent être basés sur d'autres descriptions du monde physique (mécanique, thermodynamique, électromagnétisme, gravitation...) ;
- les modèles de signaux (faire passer un signal par une fonction de transfert, par exemple) ;
- les "modèles" basés sur un mapping dit dynamique (réseaux de neurones, apprentissage...) ;
- les modèles statistiques (stochastiques de Xenakis, Stochos et Cosmos de Sinan Bokesoy, les "vols d'étourneaux", les "bancs de poissons"...)
- les modèles géométriques (Ali Momeni, David Wessel) et ceux que nous appellerons "

topologiques" dont les principales variables sont spatiales : Syter (GRM), Hyperprism (Arboretum), Interpolator (Martin Spain), pMix (Oliver Larkin), Metasurface (AudioMulch).

Castagné et Cadoz définissent dix critères pour évaluer la pertinence d'un modèle physique [12] (basés sur les 10 critères de Jaffe pour évaluer un dispositif de synthèse) axés notamment sur l'aspect "plausible" du son résultant, le retour haptique possible, et surtout la capacité du modèle à s'accorder à la "réalité" intellectuelle construite par l'utilisateur, ce qui permettra de lui donner à percevoir les artefacts comme de "vrais" objets, et non plus des objets virtuels.

Les points communs de ces différents modèles sont une évolution complexe contrôlée initialement par des paramètres simples, la multiplicité des variables du modèle et la rapidité d'exécution. On peut par exemple concevoir un modèle masses-ressorts grâce auquel on récupère en temps-réel les coordonnées de milliers de masses en n'en contrôlant qu'une seule qui est reliée à toutes les autres.

Il est important de noter une distinction entre les modèles physiques dont la vocation est de générer directement le son (le mouvement du modèle est interprété comme une vibration qui génère une onde acoustique), de ceux qui visent à contrôler les variables d'un système de synthèse.

4 Contrôle du rayonnement acoustique

Examinons maintenant plus en détail les aspects relatifs au rayonnement acoustique du dispositif de synthèse, que nous souhaitons pouvoir contrôler musicalement, "instrumentalement".

Cela a déjà été remarqué et étudié, surtout depuis le développement des musiques électroacoustiques, les caractéristiques des dispositifs de diffusion sonore conditionnent largement notre perception des espaces acoustiques engendrés. [13, 14]

En ce qui nous concerne, nous nous attacherons à développer des modèles permettant une diffusion sonore vivante, localisée, de nature instrumentale, c'est-à-dire à l'image de celle des instruments de musique.

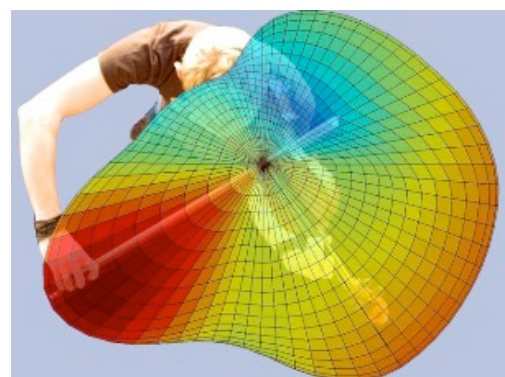


Figure 6 : source Gottfried K. Behler

Cet objectif a été très bien décrit dans différents articles, et réalisé en partie dans plusieurs centres de recherche et de création. Parmi eux, je n'en citer que deux : la Timée de l'IRCAM [15, 16] et le Compact Spherical Loudspeaker Array du CNMAT [17].

Ces dispositifs électroacoustiques permettent de reconstituer avec plus ou moins de précision des diagrammes de directivité complexes, ce qui est aussi notre but. L'originalité de notre démarche n'est donc pas dans le concept de source à directivité contrôlée.

Le cahier des charges que nous nous sommes imposé est le suivant :

- une directivité variable, vivante, et musicale ;
- un coût raisonnable ;
- un contrôle expressif.

Il s'agit en particulier d'étudier le contrôle simultané de la synthèse et du rendu sonore en mettant en œuvre une architecture logicielle basée sur le concept de modèle intermédiaire tel que nous l'avons défini.

Avant de nous lancer dans la construction de notre diffuseur, nous avons réalisé des essais avec une source sphérique dotée de douze haut-parleurs. Le dispositif dont nous disposons au LAM est utilisé en acoustique des salles.



Figure 1 : GSR (Globe Source Radiator) - Outline

Dans notre cas, nous avons fait en sorte de pouvoir adresser chacun des douze haut-parleurs indépendamment.

Cette première étape a permis de valider l'idée et de concevoir un modèle peu onéreux et plus simple à construire. S'il ne présente pas toutes les qualités requises, en termes de précision des figures de directivité possibles en particulier, il nous est apparu plus utile d'essayer d'obtenir une qualité de reproduction correcte avec des composants choisis peu onéreux.



Figure 1 : EAR (Enceinte Acoustique Radiante) - LAM

En effet, un tel dispositif doit pouvoir être dupliqué pour équiper un petit ensemble.

De plus, notre hypothèse, qui devra maintenant être vérifiée lors de tests plus approfondis, est qu'il importe plus de reconstituer des dynamiques spatiales vivantes et très corrélées à la synthèse, que de viser une précision spatiale coûteuse et pas nécessairement perceptible.

C'est à partir de ces hypothèses qu'a été conçu et réalisé la EAR (Enceinte Acoustique Radiante) sur laquelle ont été menés nos essais.

Après une première phase de mesure et de calibrage, les premiers algorithmes de restitution ont pu être couplés à des moteurs de synthèse simples mais pilotés, via des modèles intermédiaires du type réseaux masses-ressorts, à l'aide d'une tablette graphique.

Dans un second temps, des modèles statistiques seront implémentés (évolutions dynamiques de matrices de très grandes dimensions par des méthodes du type Runge-Kutta).

5 Conclusion

Dans les mois qui viennent, une exploration plus systématique des couplages entre interfaces gestuelles, modèles intermédiaires, moteurs de synthèse et de rendu devra être entreprise. Sur la base de quelques situations repérées comme permettant d'obtenir une interaction musicien-dispositif suffisamment riche et contrôlable, pourra alors commencer un travail d'évaluation des qualités instrumentales du dispositif.

Références

- [1] Smallwood S., Trueman D., Cook P., and Wang G., "Composing for Laptop Orchestra", *Computer Music Journal*, Spring 2008, Vol. 32, N° 1, 9-25 (2008).
- [2] Battier M., "L'approche gestuelle dans l'histoire de la lutherie électronique. Etude de cas : le *theremin*", in *Les nouveaux gestes de la musique*, H. Genevois et R. de Vivo (eds). Editions Parenthèses, 139-156, (1999).
- [3] Battier M., "La réinvention du son : de la lutherie électrique à la pensée du timbre des sons artificiels", *Les Cahiers de la Société Québécoise de Recherche en Musique*, vol 9, N° 1-2, Octobre 2007.
- [4] Cadoz C., "Musique, geste, technologie", *Les nouveaux gestes de la musique*, H. Genevois et R. de Vivo (eds), Éditions Parenthèses, 47-92, (1999).
- [5] Cance C., Genevois H. et Dubois D., "What is instrumentality in new digital musical devices? A contribution from cognitive linguistics and psychology", *Actes du congrès CIM09*, à paraître dans « La musique et ses instruments » (2009).
- [6] de Laubier S. et Goudard V., "Meta-Instrument 3: a look over 17 years of practice", *Proceedings of International Conference on New Interfaces for Musical Expression (NIME'06)*, (2006).
- [7] de Laubier S. et Goudard V. "Puce Muse - La Méta-Mallette", *Actes des JIM*, (2007).
- [8] Ghomi E., "Utilisation de modèles intermédiaires pour le mapping de paramètres de synthèse", *rapport de maîtrise ATIAM*, (2006).

- [9] Momeni A., Henry C., "Dynamic Independent Mapping Layers for Concurrent Control of Audio and Video Synthesis", *Computer Music Journal*, Spring 2006, Vol. 30, N°1, 49-66, (2006).
- [10] Beaudouin-Lafon M. "Moins d'interface pour plus d'interaction", *Interfaces Homme-Machine et Création Musicale*, H. Vinet et F. Delalande (eds), Hermès, 123-141, (1999).
- [11] Bahn C., Hahn T. and Trueman D., "Physicality and Feedback: A Focus on the Body in the Performance of Electronic Music", *Proceedings of the International Computer Music Conference*, Havana Cuba, (2001).
- [12] Castagné N., Cadoz C., "10 Criteria For Evaluating Physical Modelling For Music Creation", *DAFX Proceedings*, (2003).
- [13] Genevois H., "L'invention de l'espace : le son des musiques électroacoustiques et amplifiées", *Forum international « Musique et Architecture / Architecture et Musique »*, (2008).
- [14] Genevois H. et Orlarey Y. (sous la direction de), actes du congrès RMPD 95, *"le Son & l'Espace"*, Aléas ed. (1997).
- [15] Warusfel O., Derogis P., Caussé R., "Reproduction of directivity patterns using multi-loudspeaker sources", *Proc. 103rd Audio Eng. Soc. Convention*, New-York, (1997).
- [16] Misdariis N., Warusfel O., Caussé R., Nicolas F., "Le contrôle de la directivité par un système multi haut-parleurs", *Actes des JIM*, (2002).
- [17] Wessel D., "Instruments That Learn, Refined Controllers, and Source Model Loudspeakers", *Computer Music Journal*, Winter 1991, Vol. 15, N° 4, 82-86, (1991).