

Le projet *Ecoute active* à SONY CSL

François Pachet

SONY CSL-Paris, 6, rue Amyot, 75005 Paris,

pachet@csl.sony.fr

http://www.csl.sony.fr

Résumé

L'écoute active" est une activité de recherche consistant à explorer divers paramètres de contrôle agissant sur de la musique préenregistrée, et permettant d'enrichir la pratique de l'écoute musicale. Ces paramètres de contrôle définissent des espaces d'exploration dans lesquels des pièces de musique peuvent être écoutées par des auditeurs non experts en subissant certaines variations. L'objectif est à la fois d'augmenter le confort d'écoute des auditeurs, mais aussi de proposer des chemins d'accès à des musiques mal connues ou difficiles. Proche de l'idée de forme ouverte, l'écoute active s'en différencie en ce que nous nous intéressons à l'exploitation de répertoires musicaux existants, plus qu'à l'exploration libre ou la création de nouveaux matériaux musicaux.

1. Introduction

1.1 Les laboratoires SONY CSL

Le laboratoire de recherche SONY CSL-Paris (Computer Science Laboratory), créé en 1997 par Luc Steels¹, est une extension du laboratoire SONY CSL de Tokyo², fondé en 1988. L'ambition de ces laboratoires est d'identifier et développer les technologies informatiques qui pourront être intégrées dans les systèmes d'information et de loisir de demain.

Dans ce contexte, le thème central de ce laboratoire de recherche fondamentale est l'étude et la modélisation de la *complexité* dans les organisations humaines. Deux équipes le composent: une équipe travaille sur l'origine et l'évolution du langage, considéré comme un système dynamique complexe: comment

modéliser l'émergence de conventions langagières partagées (lexique, syntaxe, grammaire).

Une autre équipe s'intéresse à la musique sous l'aspect de la perception individuelle (comment perçoit-on la musique, comment naissent nos goûts et préférences musicales) et collectif (comment des communautés virtuelles échangent et propagent des préférences musicales). Plusieurs projets de cette équipe concernent en outre la recherche de nouveaux modèles d'interaction musicale qui permettent à l'utilisateur d'exploiter au mieux les nouvelles possibilités offertes par les nouvelles technologies du multimédia. Ces nouvelles possibilités sont essentiellement de deux types. D'une part l'accès à domicile d'énormes quantités de données multimédia (film, musique), à travers les réseaux. D'autre part la présence d'informations *sur* le contenu des données (meta-data), décrivant les données elles-mêmes (voir par exemple les nouveaux efforts de standardisation Mpeg-4 et Mpeg-7).

1.2 Systèmes d'écoute interactifs

Notre point de départ est la constatation que chaque nouvelle technologie apporte avec elle de nouveaux types de contrôle. Contrôle est ici à comprendre comme moyen d'action, pour l'utilisateur (auditeur, spectateur) lui permettant d'adapter, voire de modifier le contenu multimédia d'origine, à sa guise.

En ce qui concerne la musique, on peut ainsi dresser une "histoire des contrôles" parallèle à l'histoire des développements technologiques des appareils musicaux. A l'origine, le phonographe d'Edison ne comportait aucun contrôle³. L'électricité a contribué à l'écoute musicale en améliorant la qualité du son

¹ <http://www.csl.sony.fr>

² <http://www.csl.sony.co.jp>

³ En fait il en proposait un: on pouvait orienter le pavillon, et ainsi influencer, de manière primitive, sur la spatialisation du son...

produit, ainsi qu'en introduisant de nouveaux contrôles: volume et tonalité (graves/aigus). Plus tard, la stéréophonie a encore amélioré la qualité du son, et parallèlement ajouté de nouveaux contrôles (gauche, droite). Enfin, le son numérique permet des égalisations plus fines, et introduit un nouveau cortège de potentiomètres.

Aujourd'hui comme hier, les avancées technologiques s'accompagnent de nouveaux types de contrôles. Lesquels ? Peut-on concevoir des types de contrôles "meilleurs" que ceux développés jusqu'à présent ?

Le projet *Ecoute active* tente d'apporter une réponse à ces questions, et développe l'idée d'une technologie dont les contrôles associés sont conçus de manière à permettre une exploration musicale pertinente, et sémantiquement valide. *Ecoute active* se distingue des autres projets sur l'interactivité musicale par le point de vue centré sur l'auditeur, et la nature des interactions recherchées. D'une part, par opposition à la notion de "forme ouverte" (Eckel, 1997), nous nous intéressons à construire des environnements dans lesquels les variations musicales préservent le sens original, pour autant que l'on puisse le définir avec précision. Par ailleurs, nous ne cherchons pas à décomposer un signal numérique musical en différentes composantes perceptives, comme proposé par Lepain (1998) dans le cadre de la consultation interactive de documents musicaux (*écoute interactive*), mais tentons plutôt d'extraire des possibilités de variation d'un document donné qui "font sens" musicalement.

Ce projet consiste donc à chercher des moyens d'enrichir la pratique de l'écoute musicale d'oeuvres existantes, en donnant à l'auditeur des moyens de modifier son écoute selon différents critères. Ces critères ou paramètres définissent des espaces dans lesquels un auditeur peut se déplacer pendant l'écoute d'un morceau de musique, en faisant varier certaines de ses caractéristiques. Les points importants de notre contexte de travail sont 1) l'exploitation de pièces musicales existantes, préenregistrées, et 2) l'identification de variations sémantiquement cohérentes sur ces pièces musicales.

L'équipe "Ecoute active" développe plusieurs projets sur ce thème. Ce document en décrit les deux principaux: *MusicSpace*, projet portant sur la spatialisation interactive de sources sonores, et *RecitalComposer*, projet portant sur

la sélection interactive de musique. Ces deux projets sont illustrés ici à titre indicatif, des références renvoient à des documents plus détaillés.

2. Spatialisation

Le projet *MusicSpace* concerne l'étude de moyens de contrôle associés à la spatialisation sonore, en particulier dans une configuration de reproduction 3D. De nombreux systèmes de spatialisation ont été élaborés depuis plusieurs années, et certains d'entre eux atteignent aujourd'hui des performances excellentes (par exemple le spatialisateur de l'Ircam). Notre propos ici n'est pas de développer un nouveau système de spatialisation, mais plutôt de fournir des moyens de contrôle sur la spatialisation qui *fassent sens* pour un auditeur, en particulier un auditeur non professionnel.

Le point de départ de ce projet est le désir plus ou moins verbalisé de modifier la spatialisation de pièces de musique de manière directe et intuitive pendant l'écoute musicale. Ce désir correspond en fait à deux types d'action. D'une part il s'agit de pouvoir explorer différents points d'écoute pendant l'audition d'un morceau de musique: tourner autour d'un orchestre, se rapprocher d'un instrument ou s'en éloigner. D'autre part, il s'agit de modifier la configuration spatiale des différents instruments ou sources sonores: isoler un instrument ou un groupe d'instruments, reconfigurer à sa guise. Mais quel sens ces actions peuvent elles avoir dans le contexte d'une écoute non professionnelle, et comment les rendre possible sans altérer l'oeuvre écoutée.

MusicSpace est une proposition de réponse. *MusicSpace* est essentiellement une *interface* permettant d'écouter des pièces de musique (format audio et Midi), en proposant différentes formes d'interaction spatiale. L'interaction de base consiste à permettre le déplacement des différents instruments, ainsi que de l'auditeur lui-même, représenté par une icône (avatar). Ces déplacements induisent un mixage des différents instruments calculé en fonction des distances entre l'icône de l'auditeur et les icônes des différents instruments. L'auditeur peut ainsi dynamiquement se rapprocher d'un instrument (le piano pendant le chorus pour une pièce de Jazz par exemple), ou encore en supprimer temporairement un pour mieux écouter les autres.

Plusieurs versions de MusicSpace sont développées : une version 2D spécialisée pour le Midi (le système MidiSpace, voir Figure 1) et une autre 3D dans un environnement VRML (Figure 2).

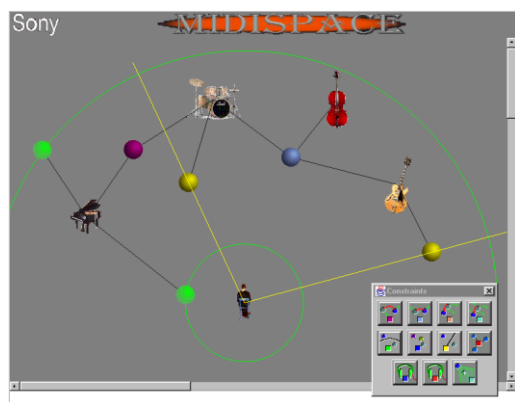


Figure 1. MidiSpace : un espace à deux dimensions pour explorer un morceau de musique.

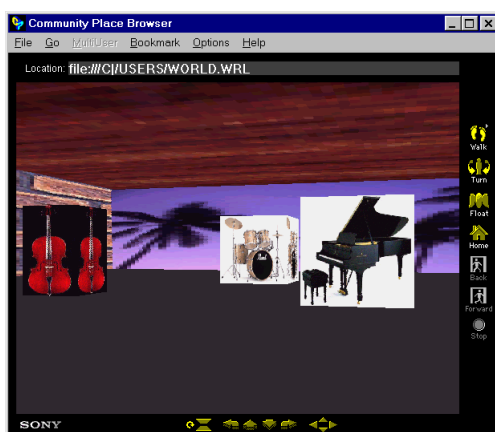


Figure 2. MidiSpace 3D sur un trio de Jazz.

La question principale qui se pose dans un tel environnement est alors celle de la cohérence: comment les déplacements de l'auditeur, et des sources sonores peuvent-ils faire sens ? Comment concilier l'idée même d'une ingénierie du son, avec celle d'un contrôle autorisé par l'auditeur ?

La proposition de MusicSpace consiste à mettre en oeuvre une représentation - sommaire mais efficace - des connaissances de l'ingénieur du son, nécessaires à garantir certaines bonnes propriétés du mixage. Ces propriétés sont de plusieurs natures. D'une part il s'agit d'obtenir un signal global satisfaisant (dans des limites de niveau raisonnables, ni saturé ni trop faible). D'autre part il s'agit de réaliser un certain nombre de propriétés qui dépendent du style, du morceau, voire de la

volonté du compositeur. Dans l'exemple simple d'un trio de Jazz, on peut par exemple considérer la basse et la batterie comme une entité à part entière (la section rythmique). Ceci se traduit alors par la pose d'une contrainte, appelée groupe, qui va assurer que le *rapport des distances* à l'auditeur est constant. Par ailleurs, on peut vouloir opposer le piano, considéré dans ce contexte comme instrument soliste, à la section rythmique. Ceci se traduit en MusicSpace par la pose d'une contrainte dite d'équilibre, entre le piano et la section rythmique. En termes techniques, cette contrainte assure un *produit constant* des différentes distances. MusicSpace propose ainsi une bibliothèque de contraintes, correspondant à des propriétés intéressantes du mixage. Ces propriétés s'expriment de manière géométrique, en fonction des paramètres de base des sources sonores: distance à l'auditeur, position angulaire, orientation, et de manière générale tout paramètre géré par le système de spatialisation ainsi contrôlé.

Ces contraintes sont alors interprétées par un solveur de contraintes. Le principe d'interaction est fondé sur un cycle de base simple: l'utilisateur déplace une source sonore (ou son avatar). Ce déplacement va en général provoquer des violations de certaines contraintes. Le solveur est alors déclenché, et calcule les déplacements des autres sources sonores nécessaires à rétablir les propriétés des contraintes. Une fois ces déplacements calculés, l'interface déplace automatiquement les icônes correspondantes.

Ce mécanisme simple mais efficace a plusieurs effets. D'une part il permet de produire un *langage d'expression* des actions de mixage. A ce titre, MusicSpace peut être considéré comme une interface de haut niveau pour des tables de mixage, et la bibliothèque de contraintes constitue une ontologie des actions de mixage. D'autre part MusicSpace s'inscrit dans un courant récent, matérialisé par exemple par le standard Mpeg-4, consistant à développer l'idée d'un format musical indépendant du système de reproduction sonore de l'auditeur, mais qui soit reconfigurable *in situ*, pour s'adapter aux conditions particulières de reproduction (par exemple nombre de haut-parleurs ou qualité des convertisseurs digital/analogique). Dans ce contexte, MusicSpace permet d'étendre les possibilités d'adaptation, tout en garantissant, de manière générique, des propriétés finales du mixage.

Enfin, MusicSpace permet d'envisager un cycle composition/mixage/écoute un peu différent du cycle classique: d'une part le système de contraintes permet de déléguer à la machine une part - la part la plus mécanique - des responsabilités de l'ingénieur du son, permettant à celui-ci de se concentrer sur les propriétés du mixage plutôt que sur les actions effectives sur la table, tout en donnant à l'auditeur une marge de manoeuvre nouvelle. D'autre part, et c'est sans doute la conséquence la plus intéressante, les compositeurs peuvent imaginer des musiques nouvelles qui intègrent d'emblée des possibilités de variation spatiale par l'auditeur.

3. Sélection de musique

Le projet *RecitalComposer* concerne le problème de la sélection musicale à grande échelle. Les progrès récents en matière de télécommunication, réseaux, et de cryptage des documents multimédia permettent en effet de transporter à domicile des catalogues entiers de films ou de musique. Dans le cas de la musique, les utilisateurs vont avoir, ont déjà dans certains cas, accès de manière interactive à des catalogues de l'ordre de 100.000 titres. Que faire avec ce choix immense ?

Les approches traditionnelles sont, ici encore, essentiellement technologiques. Le problème de la sélection musicale est soit considéré comme un simple problème d'accès à une base de données (système de requêtes plus ou moins précises), soit comme une instantiation de la technique dite de *filtrage collaboratif*. Dans les deux cas, l'utilisateur se voit proposé un moyen de navigation soit trop rigide et sans surprise (bases de données), soit biaisé par les traits caractéristiques une communauté particulière (filtrage collaboratif).

Le projet *RecitalComposer* propose une approche au problème de la sélection musicale, qui consiste à exploiter les propriétés des programmes musicaux, vus comme des séquences temporelles particulières. Au lieu de spécifier les propriétés des titres désirés, *RecitalComposer* permet de spécifier les propriétés des séquences. Le système choisit ensuite automatiquement les titres et leur emplacement dans la séquence.

Ces propriétés sont de nature diverse. D'une part on peut indiquer des propriétés de continuité entre titres successifs. Par exemple forcer une continuité de styles d'un titre à l'autre, ou de tempo (pas de saut trop

important). Par ailleurs on peut indiquer des propriétés dites de cardinalité sur l'ensemble de la séquence: au moins 40 % de Jazz, ou bien 60 de voix de femme, etc. Enfin on peut indiquer des propriétés de différence: forcer les auteurs à être différents, ou bien encore limiter le nombre d'instruments principaux à 2 ou 3.

La spécification d'un programme musical par de telles contraintes donne lieu à un problème combinatoire complexe (Pachet et al. 1998). La technologie des CSP (*Constraint Satisfaction Problem*) est utilisée pour résoudre ces problèmes dans des temps acceptables. Le système résultant, *RecitalComposer*, permet alors de proposer différents services pour la sélection musicale interactive. L'un d'entre eux, appelé *PathBuilder*, permet de spécifier un morceau de départ, et un morceau d'arrivée, à partir d'un catalogue de titres. L'utilisateur peut ensuite spécifier la longueur du programme (par exemple 25 morceaux), et le nombre de *discontinuités* entre 2 titres successifs. Une discontinuité dans ce contexte est une différence de valeur pour un attribut des titres.

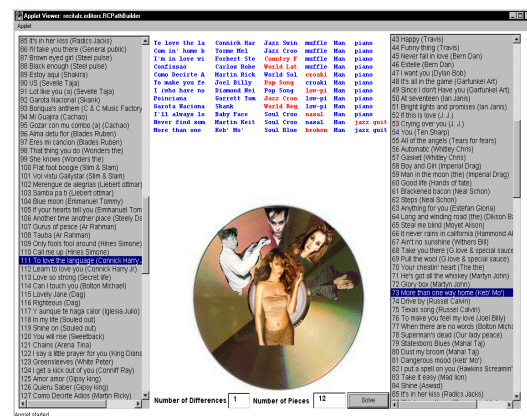


Figure 3. L'interface de PathBuilder. L'utilisateur choisit un morceau de départ et un morceau d'arrivée, ainsi que le degré de "discontinuités". PathBuilder trouve un chemin qui va de l'un à l'autre, à partir des titres du catalogue. Les discontinuités sont indiquées en rouge, les continuités en bleu.

4. Conclusion

Les nouvelles technologies apportent avec elles nécessairement des nouveaux moyens de contrôle. Ces contrôles ne seront effectifs que s'ils font sens pour l'utilisateur. Par opposition aux contrôles traditionnels, fondés sur la technologie sous-jacente (le bouton de volume de l'amplificateur), et donc dénués de sens musical, l'informatique permet aujourd'hui

d'envisager des contrôles plus sémantiques, et en ce sens d'imaginer un changement radical de la nature même de l'écoute musicale.

5. Bibliographie

Eckel G., "Exploring Musical Space by Means of Virtual Architecture", *Proceedings of the 8th International Symposium on Electronic Art*, School of the Art Institute of Chicago, 1997.

Delerue, O. Pachet, F. "MidiSpace, un spatialisateur midi expérimental", Journées d'Informatique Musicale, Agelonde, 1998.

Delerue, O. Pachet, F. "Constraint Propagation for real time spatialization", ECAI 98 Workshop on Constraints for Artistic Applications, Brighton, 1998b.

Lepain P., "Écoute interactive des documents musicaux numériques", *Recherches et applications en informatique musicale*, Hermes, 1998.

Pachet, F. Delerue, O. "A Constraint-based Temporal Music Spatializer", *ACM Multimedia Conference*, Brighton, 1998.

Pachet, F. Delerue, O. "Annotations for Real Time Music Spatialization", International Workshop on Knowledge Representation for Interactive Multimedia Systems, KRIMS-II workshop, Trento, Italy, 1998b.

Pachet, F. Delerue, O. "A Mixed 2D/3D Interface for Music Spatialization", First International Conference on Virtual Worlds (VW98), Springer Verlag, Lecture Notes in Computer Science, n. 1434, pp. 298-307, 1998c.

Pachet, F. Delerue, O. Constraint-Based Spatialization, First International Workshop on Special Effects, DAFX 98, Barcelona, 1998d.

Pachet, F. Roy, P. Cazaly, D. A Combinatorial Approach to Content-based Music Selection. Submitted to IEEE International Conference on Multimedia Computing and Systems, Florence, Italy, 1999.