

L'hyperespace musical : définition, navigation, exploitation

M. Crucianu

Laboratoire d'Informatique, Université de Tours
École d' Ingénieurs en Informatique pour l' Industrie
64, avenue Jean Portalis, 37200 Tours, FRANCE
crucianu@univ-tours.fr
<http://www.rfai.li.univ-tours.fr/~crucianu/>

Résumé : L'exploration de larges collections de pièces musicales est contrariée par le contraste qui existe entre les méthodes de classement existantes, basées sur la notion plutôt simpliste de genre musical, et la richesse des caractéristiques subjectives perçues par les auditeurs. Nous tentons d'exprimer ces similarités subjectives dans la définition d'un hyperespace musical multidimensionnel, dans lequel de multiples techniques de navigation peuvent être développées. Nous nous intéressons aussi aux apports de l'apprentissage à partir de données dans l'évolution de la définition de l'espace musical et l'adaptation à l'utilisateur. Enfin, nous mettons en avant les multiples apports qu'un tel outil peut avoir pour toutes les parties concernées par la création et la diffusion musicale.

Mots-clés : hyperespace musical, apprentissage, hypermédia.

Abstract: It is usually difficult to explore large collections of music, mainly because there is such a large difference between common classifications, based on the rather simplistic notion of musical genre, and the richness of the human perception of music. We attempt here to embody the subjective similarities between pieces of music into a music hyperspace, which can be explored in various ways. We also study the ways in which machine learning techniques can progressively improve the definition of the music space and allow the navigation methods to adapt to the user. We eventually put forward the advantages such a tool can bring not only to music listeners, but also to musicians, music editors and distributors.

Keywords: music hyperspace, machine learning, adaptive hypermedia.

1. Introduction

Traditionnellement, la recherche de pièces musicales adaptées à ses propres goûts ou à des usages particuliers se fait à partir de noms d'artistes connus ou de genres musicaux (définis de façon plutôt simpliste). Depuis quelques années, différentes sociétés de commercialisation via Internet ont mis en place des outils de conseil du client basés sur les achats effectués par les autres clients ; le conseil prend le plus souvent la forme d'une courte liste des disques les plus achetés par les clients ayant commandé le disque affiché. Aussi, des systèmes à base de connaissances ont été mis au point par des chercheurs afin de conseiller les amateurs de musique dans leurs choix ; la portée de tels systèmes est limitée par la nature symbolique des représentations employées et par la faible accumulation de connaissances. Enfin, depuis peu, un outil comme Soundfisher (www.soundfisher.com) tente de proposer des recherches basées sur des similarités extraites par une analyse du signal, à travers une interface textuelle.

Notre objectif est de développer un outil permettant de trouver des pièces similaires, d'un point de vue perceptif, à d'autres pièces musicales, ou d'explorer de façon intuitive un « espace musical » à la recherche de pièces correspondant à des critères musicaux bien définis ou plus subjectifs. Notre proposition est originale notamment par les trois aspects suivants : le remplacement d'une recherche par genre musical par une recherche par similarité perceptive ou par positionnement dans un espace musical, la présentation des informations de similarité et de positionnement sous une forme visuelle (en deux dimensions ou plus), et l'introduction d'éléments d'adaptation à l'utilisateur dans le choix, la définition et l'exploitation des dimensions de l'espace musical.

La section suivante présente le principe de définition de l'espace musical et la navigation dans cet espace. Des détails concernant les différents types de similarités et de dimensions à prendre en compte sont exposés dans la troisième section. Dans la quatrième section nous discutons le rôle que peut avoir l'apprentissage à partir de données dans la définition et l'exploitation de l'espace musical. Les multiples apports potentiels de l'espace musical sont indiqués dans la cinquième section. Enfin, quelques directions envisagées pour l'extension de nos propositions sont mentionnées dans la conclusion.

2. Définition de l'hyperespace musical

L'espace musical est la représentation d'un ensemble important de pièces musicales sous la forme d'un nuage de points dans un espace multidimensionnel. Chaque pièce est un point du nuage et sa position est définie par les valeurs prises par des paramètres objectifs et subjectifs qui représentent les dimensions de l'espace musical (voir la section suivante). La distance entre deux points est une mesure de similarité entre les deux pièces musicales correspondantes.

Un tel espace musical peut être visualisé, par groupe de dimensions, à travers un outil de navigation classique, grâce à différents langages de représentation comme VRML, X3D ou Java 3D. Jusqu'à trois des dimensions de l'espace musical peuvent être représentées sous une forme spatiale, pour les autres nous pouvons nous servir d'autres caractéristiques visibles des points comme la couleur, la forme, la texture ou la taille. Le choix des dimensions à représenter et de la forme de représentation peut être laissé à l'utilisateur, mais le système peut lui fournir une assistance dans ce choix (voir la quatrième section).

L'association d'un ou plusieurs liens explicites ou implicites (passant par des traitements intermédiaires) à un point du nuage, un sous-ensemble de points ou une région de cet espace produit l'hyperespace musical. Les liens permettent des opérations de navigation diverses : écoute d'un fragment de la pièce musicale, obtention d'informations concernant la pièce (titre, interprète, éventuel album, accès direct à un distributeur, etc.), zoom avec changement manuel ou automatique des dimensions visualisées, affichage synthétique des caractéristiques communes des pièces du sous-nuage, etc.

Remarquons enfin que notre proposition n'a aucun point commun avec celle présentée dans [Bonardi et Rousseaux, 2000], malgré la similitude des noms. Ces auteurs développent une représentation spatiale à partir de la partition d'une pièce musicale et envisagent de lier un fragment d'une pièce à un fragment d'une autre pièce.

3. Quelles dimensions pour l'hyperespace musical

Principalement en raison de la nouveauté de cette approche, il est très difficile de sélectionner *a priori* un ensemble de dimensions approprié. De plus, il peut être très utile de laisser cet ensemble évoluer dans le temps pour profiter d'avancées dans la compréhension des relations entre les caractéristiques physiques des pièces de musique et leur perception subjective, ainsi que de régularités observées dans les choix des amateurs de musique. Nous nous contenterons ici de proposer une typologie des dimensions qui peuvent être prises en compte et de présenter quelques-unes des premières possibilités envisagées.

Nous considérons ici qu'une dimension n'est pas nécessairement définie par un ensemble continu, mais peut correspondre simplement à un ensemble ordonné de valeurs. Les dimensions continues seront représentées en priorité sous forme spatiale, alors que les dimensions discrètes peuvent très bien s'accommoder de supports de représentation discrets (couleur, forme, texture ou taille des points).

Nous pouvons distinguer les types de dimensions suivants :

- 1° Les paramètres objectifs extraits du signal, avec une correspondance subjective claire. Un assez bon exemple est le tempo d'une pièce musicale (il faut toutefois noter que pour une même pièce différents niveaux de tempo peuvent être définis). Ces paramètres ont une utilité générale indiscutable et des méthodes fiables permettant de les estimer doivent être développées.
- 2° Les paramètres objectifs extraits du signal, avec une interprétation subjective moins précise. Dans cette catégorie nous pouvons séparer les mesures de moyenne, comme l'amplitude moyenne ou le spectre moyen, et les mesures de variation (magnitude des variations, rapidité de variation), comme l'écart type des variations d'amplitude ou l'histogramme des rapidités de variation de l'amplitude. De tels paramètres peuvent être jugés utiles après l'analyse d'un nombre important de pièces musicales, ou peuvent contribuer à la définition de dimensions composites.
- 3° Les paramètres objectifs dérivés, obtenus par des méthodes d'analyse des données (analyses factorielles, *multi-dimensional scaling*, etc.) appliquées aux paramètres extraits du signal, sur un ensemble de pièces examinées. Le rôle de ces paramètres est de synthétiser des similarités constatées entre les pièces musicales.
- 4° Les paramètres subjectifs, obtenus par des méthodes d'analyse des données ou par des techniques d'apprentissage à partir des données collectées à partir des choix des utilisateurs (voir la section suivante). Ces paramètres reflètent en premier lieu la perception subjective d'un ensemble d'utilisateurs et peuvent éventuellement être liés *a posteriori* à des paramètres objectifs (voire expliqués par ces derniers).

Cette typologie correspond à une évolution qui part du contenu pour se rapprocher des usages : les groupes 1 et 2 constituent une caractérisation primaire du contenu, le groupe 3 reflète une caractérisation ensembliste du contenu, et le groupe 4 marque l'intervention des usages, l'adaptation à un groupe ou l'adaptation à un utilisateur. Nous pouvons constater ainsi que les dimensions des trois premiers groupes sont génériques (valables pour tous les utilisateurs), alors que dans le dernier groupe nous pouvons retrouver des dimensions personnalisables.

Il faut noter que chacune des dimensions des deux premiers groupes, prise individuellement, peut avoir une interprétation bien définie (donc une certaine autonomie), alors qu'une dimension appartenant à un des deux derniers groupes pourra seulement être interprétée conjointement avec d'autres dimensions. Cette distinction a une grande importance pour la sélection des dimensions effectivement affichées et la navigation dans l'espace musical.

Nous avons commencé par l'extraction du signal des paramètres suivants :

- 1° Le tempo, recherché actuellement dans les basses fréquences. Nous envisageons d'affiner la définition du tempo, en prenant en compte différentes bandes de fréquence.
- 2° L'amplitude spectrale moyenne dans les 32 bandes retenues pour la compression MP3. Nous obtenons ainsi 32 valeurs par pièce musicale, qui donnent (entre autres) une indication de sa richesse spectrale. Différentes méthodes de normalisation sont à l'étude, dans le but de faciliter les comparaisons entre plusieurs pièces. La construction d'une mesure de variation associée est envisagée.
- 3° L'histogramme des amplitudes, qui renseigne (entre autres) sur l'ampleur des changements de niveau sonore.
- 4° L'histogramme des *log-attack times*, qui fournit une indication concernant la violence des crescendo. Ces temps d'attaque sont déterminés à partir d'une enveloppe du signal sonore. Une analyse fréquentielle de cette même enveloppe devrait nous permettre d'évaluer la rapidité des changements de niveau sonore.

Ces paramètres doivent être exploités pour définir un nombre limité de dimensions représentatives. D'autres paramètres objectifs peuvent être déterminés, à partir des travaux concernant les relations entre les caractéristiques physiques des sons et leur perception subjective (comme par exemple [Misdariis et al., 1998]).

Les dimensions correspondant aux paramètres extraits du signal servent de point de départ dans la définition des autres dimensions de l'espace musical. Voici la procédure que nous avons retenue :

- 1° Définir au départ un maximum de dimensions basées sur des paramètres objectifs extraits du signal.
- 2° Traiter un nombre important de pièces musicales et sélectionner parmi les dimensions définies à l'étape précédente les plus intéressantes selon deux critères : un critère objectif de pertinence pour la discrimination entre pièces musicales, un critère subjectif de perception de la dimension.
- 3° Rechercher des correspondances entre les dimensions déjà déterminées et le classement par genre musical, dans le but d'améliorer la capacité de discrimination de certaines de ces dimensions.
- 4° À partir des préférences de groupes d'utilisateurs, définir des dimensions (par exemple, par des méthodes d'analyse factorielle) permettant de regrouper les pièces musicales écoutées par les mêmes publics et de discriminer entre des catégories de pièces musicales associées à des publics différents.

- 5° À partir de choix effectués par un même utilisateur, éventuellement en fonction de paramètres externes, déterminer des critères de regroupement personnalisés permettant de définir des dimensions personnalisées.

Ce processus est itératif et aura inévitablement une durée importante. Toutefois, l'exploitation de l'espace musical peut apporter des bénéfices aux utilisateurs rapidement, après la définition de seulement quelques dimensions de cet espace.

4. Rôle de l'apprentissage dans un hyperespace musical

Des méthodes d'apprentissage à partir de données ont été employées pour obtenir des classifications des pièces musicales, souvent liées aux regroupements classiques par genre musical (voir par exemple [Wold et al., 1996], [Dannenberg et al., 1997], [Soltau et al., 1998]). Notre objectif est d'étendre l'utilisation des techniques d'apprentissage à tous les aspects de la définition et de l'exploitation de l'espace musical qui peuvent en bénéficier. Dans ce domaine nous avons recensés les applications potentielles suivantes :

- 1° Suivant les difficultés rencontrées par l'utilisateur dans la discrimination entre pièces musicales, d'autres dimensions lui sont proposées pour améliorer les dissimilarités et/ou les similitudes qu'il perçoit. De nouvelles dimensions composites, personnalisées, peuvent être créées. Si les mêmes difficultés sont rencontrées par un ensemble d'utilisateurs, l'extraction de nouveaux paramètres peut s'avérer utile ; les données accumulées permettront de comprendre quelles dissimilarités ou similitudes sont mal représentées par les dimensions disponibles, et orienter ainsi la recherche. Aussi, à partir des dimensions préférées par une population d'utilisateurs, de nouvelles catégorisations peuvent être engendrées.
- 2° Le choix des dimensions explorées et des pièces à écouter donne la possibilité de segmenter la population totale des utilisateurs et d'identifier le comportement de chaque sous-population. Les corrélations constatées entre cette segmentation et d'autres caractéristiques des utilisateurs (age, catégorie socioprofessionnelle, etc.) offrent des moyens pour mieux comprendre ces utilisateurs afin d'adapter l'offre musicale.
- 3° Après une étape de segmentation grossière ou plus fine, le suivi du glissement des préférences de chaque sous-population permet d'orienter les nouveaux utilisateurs reconnus comme appartenant à une telle sous-population. Le mécanisme prédictif peut être poussé plus loin en identifiant des précurseurs de tendance dans chaque sous-population.
- 4° À partir de l'enregistrement des chemins (de courte durée) suivis par un même utilisateur dans l'espace musical, des similarités entre les chemins sont extraites et des corrélations sont recherchées avec des événements externes (jour de la semaine, heure dans la journée, etc.). L'objectif est de définir et de proposer à l'utilisateur des « parcours musicaux » personnalisés et adaptés aux conditions externes.

Les parcours musicaux sont des séquences de pièces musicales entières ou de segments de pièces. Les méthodes de traitement du signal employées pour la construction de l'espace musical peuvent servir aussi à cette opération de segmentation et à la caractérisation des segments obtenus.

- 5° Les chemins suivis peuvent aussi servir à segmenter une population, et cette information peut être exploitée suivant les points 2 et 3 précédents.
- 6° En collaboration avec des psychologues, des parcours musicaux spécifiques peuvent être proposés avec des objectifs précis (relaxation, changement d'état psychologique, etc.). L'adaptation à l'utilisateur (ou à un groupe d'utilisateurs), avec ou sans un retour directe de sa part, permet de personnaliser ces parcours.

5. Multiples apports d'un hyperespace musical

Bien que l'intérêt pour les utilisateurs de l'hypermusical transparaisse dans la présentation faite dans les sections précédentes, nous préférons insister sur quelques points précis. Nous examinerons par la suite les apports de l'espace musical pour les autres parties concernées par la création et la diffusion musicale.

Voici d'abord quelques-unes des possibilités offertes aux amateurs de musique :

- 1° Découverte des pièces musicales nouvelles à partir des pièces connues : soit par voisinage dans l'espace musical, soit par similitude selon certaines dimensions et dissimilitude souhaitée selon d'autres dimensions (par exemple, mêmes caractéristiques spectrales mais tempo plus rapide).
- 2° Possibilité de mieux situer les pièces écoutées, en relation avec d'autres ou par rapport aux dimensions proposées.
- 3° Découverte et exploration de nouvelles régions de l'espace musical, peu connues par l'utilisateur, à partir du positionnement de ces régions suivant différentes dimensions.
- 4° Extraction des caractéristiques communes des pièces musicales situées dans une région de l'espace musical (à partir d'une sélection essentiellement graphique) et identification de bons représentants de cette région.
- 4° Suivi des parcours musicaux répondant à leurs préférences ou développés pour eux dans un but précis.

La diffusion de musique sur Internet a permis, entre autres, aux artistes de se rapprocher de leur public en court-circuitant les intermédiaires traditionnels. En revanche, dans ce nouveau contexte il devient encore plus difficile pour un amateur de musique d'explorer de manière systématique, selon des critères musicaux, l'ensemble de la création disponible à un moment donné. Non seulement l'hypermusical propose une solution à ce problème, mais il offre aussi plusieurs autres facilités aux musiciens :

- 1° Les jeunes musiciens peuvent faire découvrir leur travail en offrant aux utilisateurs le moyen de le repérer par rapport à l'existant (re)connu.
- 2° Les jeunes musiciens sont capables de se positionner dans l'espace musical de façon optimale, par rapport à leurs spécificités mais en prenant en compte aussi les goûts du public, afin de mieux se lancer.
- 3° La création peut se développer en combinant différentes caractéristiques musicales recherchées par un certain public dans une période bien définie.

4° Une recherche musicale avant-gardiste peut explorer des régions de l'espace musical peu peuplées.

5° Des musiciens peuvent choisir de composer ou d'interpréter directement des parcours musicaux pour les utilisateurs. Les meilleurs albums actuels peuvent être vus comme des précurseurs réussis de tels parcours musicaux, malheureusement trop rares. Ces parcours peuvent prendre en compte les préférences de la population ciblée.

Enfin, voici des apports supplémentaires de l'hyperespace musical pour les maisons de disque et les sociétés de commercialisation de musique :

1° L'espace musical offre le moyen de mieux positionner les groupes à lancer, ou les nouveaux albums de groupes consacrés, par rapport à l'existant. Le positionnement dans l'espace peut fournir de meilleurs arguments pour les campagnes de marketing.

2° La recherche de nouveaux groupes peut insister sur les régions de l'espace préférées par les populations ciblées, ou sur les régions identifiées à partir des précurseurs de tendance.

3° De façon complémentaire, la recherche de nouveautés peut avoir lieu dans des régions de l'espace musical peu explorées, jugées éventuellement comme présentant un intérêt particulier de part leur positionnement.

4° La possibilité d'effectuer l'enregistrement et la commercialisation directe de parcours musicaux ciblés. Un ciblage des parcours musicaux peut être vu comme une évolution des actuelles compilations.

Les énumérations que nous avons mentionnées ici sont appelées à évoluer au fur et à mesure que l'exploitation de l'hyperespace musical devient de plus en plus large.

6. Conclusions et perspectives

Nous avons mis en avant l'hyperespace musical comme un outil permettant de trouver des pièces musicales similaires à d'autres pièces, ou d'explorer de façon intuitive des dimensions ayant des implications subjectives. Dans notre proposition, la recherche par genre musical est remplacée par la navigation dans l'hyperespace musical, suivant des similarités perceptives ou le positionnement par rapport aux dimensions disponibles. La présentation de ces informations de similarité ou de positionnement se fait sous une forme visuelle, plus facile à manipuler que la forme textuelle. Cette méthode de représentation des pièces musicales facilite l'introduction de techniques d'apprentissage qui améliorent et enrichissent les services rendus aux utilisateurs.

Le passage en revue de différents apports potentiels de l'hyperespace musical a permis de mettre en évidence l'utilité de cette approche non seulement pour les amateurs de musique, mais aussi pour les auteurs, les maisons d'édition et les distributeurs.

La mise en œuvre de ces propositions permettrait probablement d'identifier des extensions intéressantes de l'hyperespace musical. Nous évoquerons ici uniquement l'extension du principe d'organisation par dimensions perceptives aux collections d'images et de séquences vidéo. Dans ce cas spécifique, nous pouvons par exemple mélanger des dimensions

thématiques et des dimensions perceptives, ou définir des dimensions perceptives valables localement, pour un groupe thématique.

Références

Bonardi, A., F. Rousseaux (2000) Creating Hypermusic spaces, in *Proceedings of RIAO 2000*, Paris, France.

Dannenberg, R. B., B. Thom, D. Watson (1997) A Machine Learning Approach to Musical Style Recognition, in *1997 International Computer Music Conference*, International Computer Music Association (September 1997), pp. 344-347.

Misdariis, N., B. K. Smith, D. Pressnitzer, P. Susini, S. McAdams (1998) Validation of a multidimensional distance model for perceptual dissimilarities among music timbres, *16th International Congress on Acoustics*, Seattle, WA, June 1998.

Soltau, H., T. Schultz, M. Westphal (1998) Recognition of music types, in *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Piscataway, NJ, 1998.

Wold, E., T. Blum, D. Keislar (1996) Content-based Classification, Search, and Retrieval of Audio, *IEEE Multimedia*, pp. 7-36, 1996.