Approches audionumériques pour l'analyse musicale

Pierre Couprie

▶ **To cite this version:**

Pierre Couprie. Approches audionumériques pour l'analyse musicale. Musicologies nouvelles, 2018, Musicologies nouvelles, 5, pp.120-132. hal-02084867

HAL Id: hal-02084867
https://hal.science/hal-02084867

Submitted on 26 Apr 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Approches audionumériques pour l'analyse musicale

Pierre Couprie
IReMus, Sorbonne Université

Depuis quelques années, la musicologie s'est emparée de techniques provenant de l'acoustique et de l'informatique pour étudier certains aspects des œuvres qui résistent souvent à l'analyse musicale comme les sons complexes, l'interprétation, le processus créatif ou les structures dans la création contemporaine.

INTRODUCTION

En 2011, Philippe Lalitte proposait une classification des techniques d'analyse musicale assistées par ordinateur en trois catégories[1] : les techniques de manipulation et de modélisation des informations sonores et musicales, les techniques basées sur les possibilités multimodales et l'approche fondée sur le paradigme sub-symbolique. L'objectif de cet article est de prolonger ce travail en approfondissant les approches basées sur l'usage des fichiers audionumériques. Si ces techniques sont utilisées depuis de nombreuses années en acoustique et en informatique, elles n'ont été introduites que récemment dans le domaine musicologique.
Avant d'entrer dans le vif du sujet, il convient de poser les limites d'un tel travail. La première est bien évidemment son manque d'exhaustivité. Le choix des techniques présentées ici correspond à une première étape dans le domaine de la musicologie numérique systématique. De même, plutôt qu'évoquer une longue liste d'opérations avec leur référence, j'ai préféré en sélectionner quelques-unes et les expliquer à l'aide d'exemples concrets. Elles sont pour la plupart accessibles à tout musicologue à l'aide de logiciels gratuits, il pourra ensuite explorer des techniques plus complexes à mettre en œuvre, tant sur le plan de la manipulation informatique que sur celui de son usage en musicologie. La deuxième limite est liée à l'objet d'étude lui-même. Les données recueillies ou calculées à partir du signal audio et les visualisations qui en sont faites ne correspondent pas à notre perception musicale. Ainsi, le sonagramme n'est qu'une estimation du spectre, il ne renseigne en rien de la manière dont nous percevons le sonore ou le musical – et par conséquent, dont nous l'analysons. De plus, l'étude de fichiers audionumériques reste très empirique dans le domaine de l'analyse musicale[2], il faut réaliser de nombreuses expériences pour dégager les techniques qui sont réellement pertinentes. La troisième limite correspond à un problème lié à la visualisation statistique des données. Dans ce domaine, il est très facile de tirer de fausses conclusions en interprétant les mauvaises données, en généralisant à partir d'un cas singulier ou en cherchant des corrélations qui n'existent pas dans la musique. Il est donc important que chaque observation de ces représentations soit contextualisée dans l'objet musical lui-même. Enfin, la dernière limite se situe dans la réplication de ces expériences. J'ai tenté d'être le plus précis possible dans les réglages utilisés et dans les procédures employées afin de permettre une reproduction de ces représentations.
Cet article se déroule en deux parties, la première présente les logiciels, les techniques et leurs limites, et la seconde se concentre sur quelques méthodes de représentation analytique.

I. TECHNOLOGIES

Les logiciels

De nombreux logiciels sont adaptés à un usage en analyse musicale. Si l'analyse de certains corpus, comme ceux liés à l'informatique musicale ou à la musique électroacoustique, nécessite souvent l'usage de logiciels pour lire et interpréter les fichiers provenant des artistes, il n'en reste pas moins que ces mêmes logiciels permettent d'aider à l'analyse musicale. La figure 1 présente les logiciels audionumériques couramment utilisés dans ce domaine. Une partie d'entre eux a été spécifiquement développée pour l'analyse, la transcription ou la représentation musicale (bordure épaisse), mais un nombre important provient aussi de la création musicale ou de l'analyse acoustique du signal. Les ellipses grisées représentent les différentes pratiques analytiques :

---

[1] Philippe Lalitte, « Du son au sens : vers une approche sub-symbolique de l'analyse musicale assistée par ordinateur », *Musurgia*, vol. 18, n°1-2, 2011, p. 99-116.
[2] Stephen McAdams, Philippe Depalle, Eric Clarke, « Analyzing Musical Sound », *in* N. Cook, E. Clarke (éd.) *Empirical Musicology Aims, Methods, Prospects*, Oxford, Oxford University Press, 2004, p. 157-196.

1. la **représentation du signal** occupe une place centrale et de nombreux logiciels intègrent des représentations simples telles que la forme d'onde et le sonagramme. Audiosculpt[3] et SPEAR[4] sont des logiciels d'analyse-synthèse[5] permettant de modifier le son par filtrage spectral. The Snail[6], dérivé d'Audiosculpt, offre une représentation très précise des hauteurs musicales ;

2. l'**extraction de descripteurs**, détaillée dans la partie suivante, a gagné en importance depuis quelques années dans le domaine musicologique. Deux logiciels sont devenus une référence : les plug-ins Vamp dans Sonic Visualiser et PRAAT[7]. Si Sonic Visualiser peut être caractérisé comme la « boîte à outils » du musicologue[8], le second est surtout adapté à l'analyse de la voix seule[9]. Les deux logiciels permettent de travailler à partir de plusieurs descripteurs. Telemeta[10] est un système de gestion de contenu collaboratif contenant des outils de segmentation basés sur certains descripteurs audio très utilisés en ethnomusicologie. Music22[11] est un outil facilitant l'analyse de la musique modale. Enfin, EASY Toolbox est adapté à l'analyse de la musique électroacoustique[12] et la MIRToolbox[13] un ensemble de fonctions pour l'extraction de descripteurs audio dans Matlab ;

3. l'**analyse de corpus et l'archivage** sont des préoccupations assez anciennes dans le domaine musicologique, toutefois, le développement de logiciels spécifiques pour l'analyse musicale ou pour l'archivage est assez récent. J'ai déjà présenté Telemeta, CataRT[14], quant à lui, offre un mode de représentation des descripteurs sous la forme de nuage de points ;

4. les logiciels reliés à l'**annotation de spectre et à la transcription** offrent tous la même fonction : l'ajout d'annotations graphiques et/ou textuelles sur une représentation du signal audio. EAnalysis[15], TIAALS[16], Acousmoscribe[17], Acousmographe[18] ont été développés pour l'étude du corpus électroacoustique. ASAnnotation n'est plus maintenu depuis de nombreuses années et ne fonctionne plus sur les systèmes récents. Notons que Sonic Visualiser offre aussi des fonctions simples d'annotation sous la forme de marqueurs temporels avec une étiquette textuelle ;

5. dans le domaine de l'**annotation de partition**, seul iAnalyse[19] est disponible publiquement. Le logiciel MetaScore a été développé pour réaliser les guides d'écoute de la Philharmonie de Paris ;

6. la **représentation analytique** complète l'annotation du spectre et la transcription en proposant des représentations plus complexes. Lignes de temps[20], initialement développé pour l'analyse de la bande-son des

---

[3] http://forumnet.ircam.fr/fr/produit/audiosculpt/ (02/03/2018).

[4] http://www.klingbeil.com/spear/ (02/03/2018).

[5] L'opération d'analyse-synthèse permet de modifier l'image obtenue par une analyse du spectre et ensuite de re-synthétiser cette image afin d'obtenir un signal audio modifié. En musicologie, cette opération est utilisée lorsque l'on souhaite supprimer ou réduire l'intensité d'une partie du spectre afin de faciliter l'écoute du reste. Pour une explication du principe, voir Richard Cann, « An Analysis/Synthesis Tutorial: Part 1 », *Computer Music Journal,* vol. 3, n°3, 1979, p. 6–11.

[6] http://www.ircamlab.com/products/p2242-The-Snail/ (02/03/2018).

[7] Jeanne Miramon-Bonhoure, Amine Beyhom, « Petit manuel pratique pour débutants – 1ere partie : Analyses de systèmes intervalliques et d'échelles », http://ethnomusika.org/2011/11/14/praat-petit-manuel-pratique-pour-debutants-1ere-partie-analyses-de-systemes-intervalliques-et-d'echelles/ (02/03/2018).

[8] Nicholas Cook, Daniel Leech-Wilkinson, « A musicologist's guide to Sonic Visualiser », http://www.charm.rhul.ac.uk/analysing/p9_1.html (02/03/2018).

[9] Wim van der Meer, « Praat manual (for musicologists) », https://wimvandermeer.wordpress.com/praat-manual-for-musicologists/ (02/03/2018).

[10] http://telemeta.org (02/03/2018).

[11] http://music22.ghrab.tn (02/03/2018).

[12] Tae Hong Park, Zhiye Li, Wen Wu, « Easy does it: The Electro-Acoustic Music Analysis Toolbox », *10th International Society for Music Information Retrieval Conference,* Kobe, 2009, p. 693-698.

[13] MIRToolbox (02/03/2018)

[14] http://forumnet.ircam.fr/fr/produit/catart-standalone/ (02/03/2018).

[15] Pierre Couprie, « Le développement d'un outil d'aide à l'analyse musicale : bilan et perspectives musicologiques », *in* N. Marty (éd.) *Musiques électroacoustiques. Analyses – Écoutes*, Paris, Delatour, 2016, p. 157-169.

[16] Michael Clarke, Frédéric Dufeu, Peter Manning, « Introducing TaCEM and the TIAALS Software », *International Computer Music Conference. International Developments in Electroacoustics*, Perth, 2013, p. 47-53.

[17] Myriam Desainte-Catherine, Jean-Louis Di Santo, « L'acousmoscribe, un éditeur de partitions acousmatiques », *Electroacoustic Music Studies Network*, Buenos Aires, 2009, http://www.ems-network.org/ems09/papers/disanto.pdf (02/03/2018).

[18] Emmanuel Favreau, Yann Geslin, Adrien Lefèvre, « L'Acousmographe 3 », *Journées d'Informatique Musicale*, Rennes, 2010, p. 105-111.

[19] Pierre Couprie, « iAnalyse : un logiciel d'aide à l'analyse musicale », *Journées d'Informatique Musicale*, Albi, 2008, p. 115-121.

[20] https://www.iri.centrepompidou.fr/outils/lignes-de-temps/ (02/03/2018).

films, propose un outil de segmentation automatique de l'image et l'ajout de structures temporelles en fonction des plans ;

7. l'**analyse de l'interprétation** est un domaine musicologique en plein développement. Toutefois, rares sont les projets s'appuyant sur des outils numériques, voire le développement d'outils spécifiques. Le projet CHARM Mazurka, basé sur l'analyse d'interprétation de mazurkas de Chopin[21], a permis le développement de plug-ins spécifiques[22] pour Sonic Visualiser. Philippe Lalitte, quant à lui, a utilisé un ensemble d'outils d'extraction de descripteurs audio pour analyser différentes interprétations des *Dix pièces pour quintette à vent* de György Ligeti[23] ;

8. enfin, l'**analyse du processus créatif** se partage entre différents champs de recherche comme l'analyse des sources produites par les compositeurs, la recherche-création et l'analyse sociologique. Parmi les développements, Rekall[24] est probablement un des seuls à avoir abouti en tant que logiciel public.

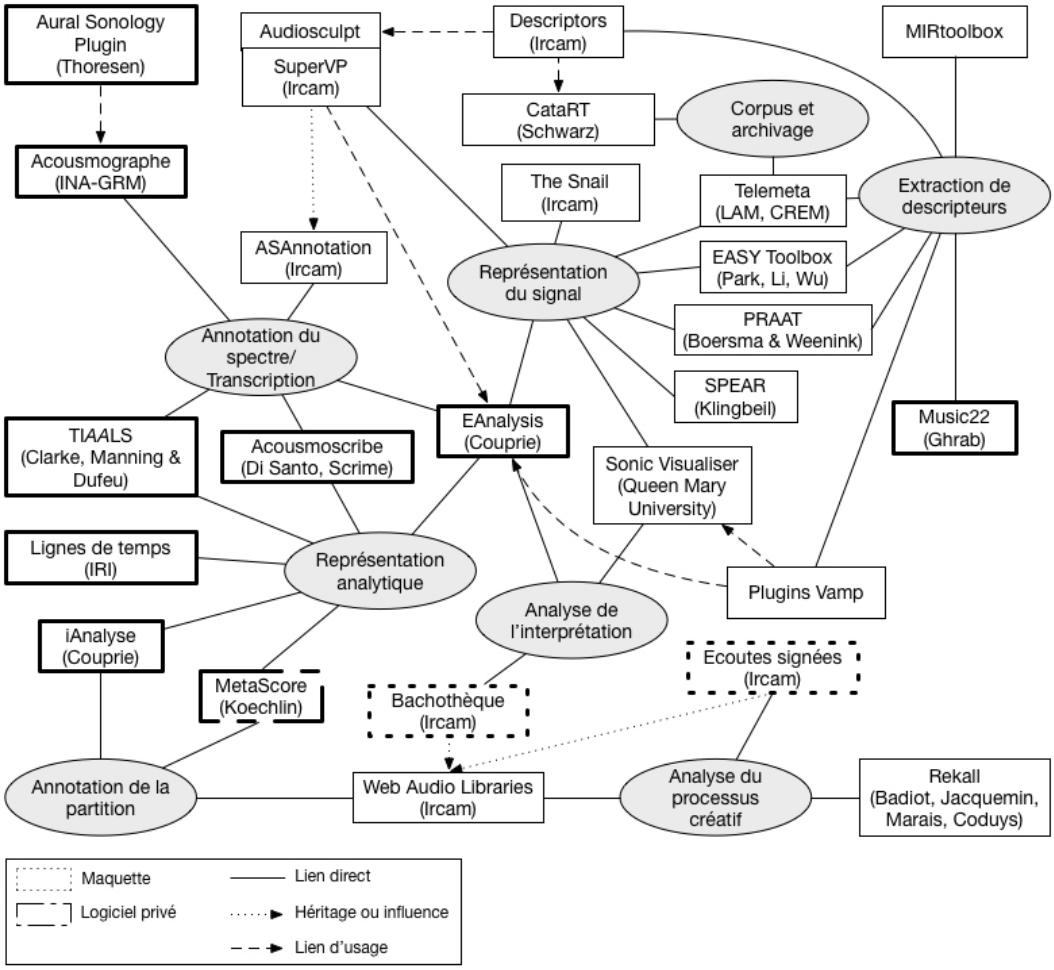

Figure 1 : Les logiciels audionumériques utilisés en analyse musicale.

Ce court panorama de quelques logiciels permet de mesurer le dynamisme des développements informatiques en musicologie numérique. Il convient maintenant de se pencher sur les méthodes de calcul employées pour extraire l'information des fichiers audionumériques.

[21] Nicholas Cook, *Beyond the Score*, Oxford, Oxford University Press, 2014, p. 176-223.
[22] Andrew Earis, Craig Sapp, « Tools developed by the CHARM Mazurka project », http://www.charm.rhul.ac.uk/analysing/p9_4.html (02/03/2018).
[23] Philippe Lalitte, *Analyser l'interprétation de la musique du XXe siècle*, Paris, Hermann, 2015.
[24] http://www.rekall.fr (02/03/2018).

L'extraction d'informations à partir du signal

La forme d'onde et le sonagramme

La forme d'onde et le sonagramme sont les deux représentations du signal les plus utilisées en analyse musicale. Dès les années 1970, Émile Leipp avait imaginé pouvoir reconnaître les sons en observant simplement leur sonagramme[25]. De même, dans les années 1980, Robert Cogan avait proposé une méthode d'analyse musicale fondée sur l'observation du spectre[26].

La forme d'onde représente la variation de l'amplitude du signal. Le sonagramme, quant à lui, est réalisé, par exemple, à partir d'une analyse de type FFT (*Fast Fourier Transform*) ou ondelette (*wavelet*), et il est calculé en quatre étapes (figure 2) :

1. le signal audio numérique est découpé en fragments en fonction de la taille de la fenêtre[27]. Les fragments sont légèrement superposés (pas de la fenêtre) ;

2. une fonction est ensuite appliquée à chaque fragment, c'est l'opération de fenêtrage qui dépend du type de fenêtre utilisée[28] ;

3. le calcul de la FFT est alors réalisé sur les échantillons du fragment ;

4. le résultat de la FFT permet d'obtenir une liste de magnitudes. Cette liste contient autant de valeurs que la taille de la fenêtre. Le sonagramme est réalisé sur la première moitié des valeurs (généralement de 0 Hz à 20 kHz ou à 22,05 kHz).

La taille de la fenêtre va déterminer la précision temporelle et spectrale du sonagramme. Les deux sont inversement proportionnelles[29]. Le tableau de la figure 3 montre une approximation de la précision spectrale et temporelle pour un fichier numérisé à 44.1 kHz. Si la précision spectrale est essentielle, notamment dans les basses fréquences, il faudra alors choisir une taille de fenêtre assez élevée (par exemple 4 096). Si, à l'inverse, la précision temporelle est plus importante, il faudra choisir une taille plus petite (par exemple 1024). Le choix de la taille de la fenêtre est donc essentiel et dépend de l'objectif.

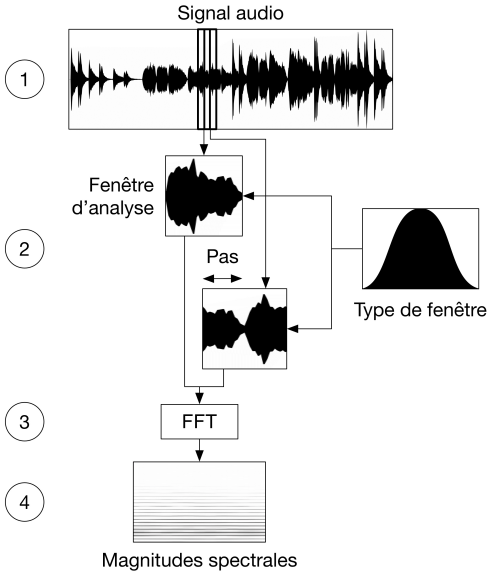

Figure 2 : Principe de calcul d'une FFT pour obtenir un sonagramme à partir d'un signal audio.

---

[25] Émile Leipp, *Acoustique et musique*, Paris, Mines ParisTech, 5/2010, p. 87-92.

[26] Robert Cogan, *New Images of Musical Sound*, Cambridge, Harvard University Press, 1984.

[27] La taille de la fenêtre est indiquée dans les paramètres d'analyse et est un multiple de 8. Cette taille correspond au nombre d'échantillons utilisés dans chaque fenêtre d'analyse sachant qu'un fichier à 44.1 kHz contient 44 100 échantillons par secondes.

[28] En analyse musicale, il est habituel d'utiliser une fenêtre de type « hanning » car elle permet d'obtenir un bon résultat sur l'échelle temporelle habituellement employée en musicologie.

[29] Le nombre d'échantillons dans la fenêtre détermine la taille temporelle de cette fenêtre – plus elle est large, moins bonne sera la précision temporelle. À l'arrivée, cette taille détermine aussi le nombre de tranches d'analyse fréquentielles (casiers fréquentiels) obtenues sur la valeur d'échantillonnage du fichier – par exemple, 2 048 casiers fréquentiels sur 44 100 Hz permettent d'obtenir des tranches d'analyse fréquentielle de 44 100 / 2 048 = env. 21,5 Hz. Seule la première moitié des casiers fréquentiels sera utilisée (de 0 Hz à 22 050 Hz).

Taille de fenêtre	Précision temporelle	Précision spectrale
1 024	env. 23 ms	env. 43 Hz
2 048	env. 47 ms	env. 21 Hz
4 096	env. 93 ms	env. 11 Hz
8 192	env. 186 ms	env. 5 Hz

Figure 3 : Précisions temporelle et spectrale d'une FFT pour un fichier numérisé à 44.1 kHz.

Le calcul du spectre par ondelette permet de pallier ce problème en apportant une précision dans les deux dimensions[30].
Si le sonagramme est un outil très utile en analyse musicale, particulièrement pour l'étude de corpus non notés, les descripteurs audio permettent de franchir une étape supplémentaire en reliant les caractéristiques sonores aux paramètres musicaux.

Les descripteurs audio

Plusieurs activités analytiques présentées dans la figure 1 sont basées sur l'extraction de descripteurs audio. L'utilisation des descripteurs s'est particulièrement développée à partir des années 1990 dans le cadre de l'analyse de contenu audio[31] (*Audio Content Analysis* ou ACA) et, plus largement, de la recherche d'information musicale (*Music Information Retrieval* ou MIR). Ces recherches englobent aussi toutes les données numériques musicales[32] : les informations extraites du signal audio, les partitions musicales, les informations contextuelles comme les notices ou les couvertures d'albums et, plus généralement, toutes les informations renseignant sur un usage musical spécifique. Elles entrent déjà dans la musicologie à travers différents domaines dont l'étude des musiques électroniques et de la *computer music*[33]. Il existe deux types de descripteurs audio :

1. les descripteurs de bas niveau, dont les données ne peuvent être directement interprétées, car elles consistent en une liste de valeurs non musicales. Ces descripteurs sont les plus simples à manipuler et varient peu en fonction des technologies employées. Ils nécessitent toutefois l'utilisation de représentations spécifiques afin de mettre en évidence les morphologies, les corrélations, les similarités, etc. ;

2. les descripteurs de haut niveau, qui, à partir d'ensembles de descripteurs de bas niveau, vont extraire des données plus complexes voire musicales comme la détection du tempo, des harmonies ou même de la structure musicale. Ces descripteurs sont directement interprétables (liste de hauteurs, de valeurs de tempo, de segments temporels, etc.). Toutefois, ils sont généralement développés pour des corpus bien spécifiques et la fiabilité des résultats diminue au fur et à mesure que l'on s'en éloigne.

Si la deuxième catégorie semble très intéressante pour la musicologie, et plus particulièrement pour l'analyse musicale, les descripteurs de haut-niveaux s'avèrent souvent plus complexes à utiliser et leurs résultats nécessitent bien souvent une correction manuelle. Dans cet article, je me concentrerai donc essentiellement sur l'usage de descripteurs de bas niveau.
Geoffroy Peeters propose de classer les descripteurs en quatre catégories[34] :

1. les descripteurs temporels comme l'amplitude RMS[35] ;

2. les descripteurs spectraux, calculés à partir d'une analyse du spectre sonore comme le barycentre spectral (*centroid spectral*), le *rolloff*, le *spectral skewness*, la variance spectrale, la *standard deviation*, etc. ;

3. les descripteurs d'énergie, permettant de caractériser les variations de plusieurs paramètres sonores ;

4. les descripteurs perceptifs, tentant de reproduire notre perception comme la sonie (*loudness*) ou les *mel-frequency cepstrum coefficients* (MFCCs).

[30] Il utilise des tailles de fenêtre dynamiques : larges pour les basses fréquences et élevées pour les hautes fréquences. Toutefois, peu de logiciels accessibles au musicologue offrent le calcul en ondelette.
[31] Alexander Lerch, *An Introduction to Audio Content Analysis. Applications in Signal Processing and Music Informatics*, Hoboken, Wiley, 2012, p. 1.
[32] Michela Magas, Xavier Serra (éd.), *Roadmap for Music Information Research*, MIReS, 2013, p. 14, http://www.mires.cc/files/MIRES_Roadmap_ver_1.0.0.pdf (02/03/2018).
[33] Kerstin Neubarth, Mathieu Bergeron, Darrell Conklin, « Associations Between Musicology and Music Information Retrieval », *in* M. Ogihara (éd.) *Proceedings of the International Society for Music Information Retrieval Conference* (ISMIR), Miami, 2011, p. 430.
[34] Peeters, *op. cit.*
[35] L'amplitude RMS (*Root Mean Square*) est une mesure de l'amplitude efficace du signal audio calculé en fonction du temps.

La figure 4 montre les étapes de réalisation d'une représentation à partir de descripteurs extraits d'un fichier audio[36]. Les descripteurs sont calculés à partir du signal ou de la FFT.

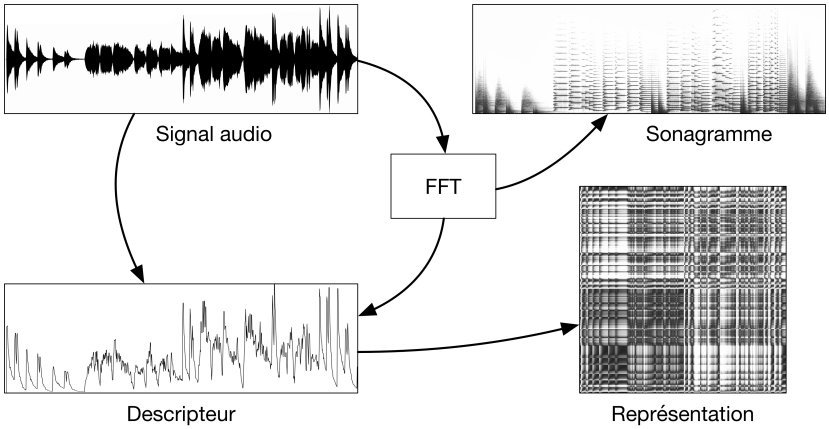

Figure 4 : Du signal audio aux représentations

Les deux premières catégories de Peeters offrent de très nombreuses applications en analyse musicale : l'étude de la variation de l'intensité, du timbre, de l'orchestration, des qualités d'enregistrements, de la forme musicale, etc. La quatrième catégorie contient des descripteurs qui permettront de compléter les représentations obtenues avec les premiers. Toutefois, les informations qu'ils apportent doublent bien souvent celles obtenues à partir des partitions comme dans le cas des MFCCs. L'usage des descripteurs nécessite donc une évaluation fine en amont pour éviter un long travail qui s'avérera ensuite inutile.
Dans la partie 3, je décrirai des exemples de représentations complexes réalisées à partir de descripteurs. Pour exemplifier la présentation précédente, voici un exemple simple d'utilisation de l'amplitude RMS et de la fréquence fondamentale (figure 5). Si le premier descripteur est de bas niveau et permet d'avoir une estimation de l'amplitude efficace du signal optimale pour la musicologie, le second est un descripteur de haut-niveau dont la fiabilité n'est pas toujours assurée. Il existe toutefois des cas où son calcul est relativement pertinent. L'exemple qui suit utilise un fragment de voix analysé à l'aide du logiciel PRAAT[37]. Cette figure représente trois visualisations : (de haut en bas) la forme d'onde, la fréquence fondamentale (en noir) et l'amplitude RMS (en gris). La réalisation d'une courbe à partir de la fréquence fondamentale permet de visualiser les inflexions fines de la voix. Cette analyse facilite la transcription, l'analyse des modes d'attaques et des variations de la hauteur ou de la qualité de la voix.

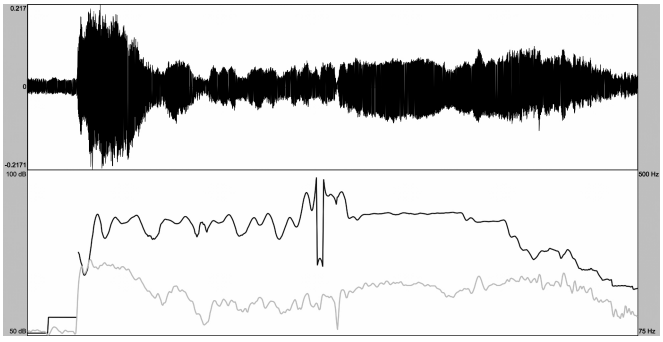

Figure 5 : Représentation[38] de l'amplitude RMS, de la fréquence fondamentale et de la forme d'onde réalisée avec PRAAT.

[36] Le signal audio est directement issu du fichier numérique. Afin d'obtenir des résultats pertinents, il est essentiel de travailler à partir de fichiers non compressés. Dans le cas d'enregistrements réaliser par le musicologue, il est aussi très important de limiter au maximum les bruits environnants qui risquent de perturber le calcul des descripteurs.
[37] Le logiciel PRAAT utilise un algorithme de détection de la fréquence fondamentale particulièrement efficace sur l'analyse de la voix. Dans l'enregistrement, la voix doit être largement prééminente par rapport aux instruments afin que la détection fonctionne correctement. PRAAT permet aussi de corriger manuellement le résultat, notamment de supprimer les sauts d'octaves ou les parties non chantées.
[38] L'extrait ici analysé (0'47-0'54) est emprunté à l'opéra *Langen mandra Wanara* du prince Danuredjo VII (1865 ?-1933), compositeur javanais (« Asie du sud-est », *in Le monde des musiques traditionnelles*, Paris, Ocora Radio France, 1994, CD HM41). J'ai utilisé les réglages par défaut du logiciel PRAAT.

La problématique du *workflow* en analyse musicale

Un des écueils les plus importants dans l'usage de technologies informatiques pour l'analyse musicale réside dans la gestion du *workflow*[39]. J'ai déjà analysé en détail[40] les problèmes posés par la conversion de fichiers informatiques entre différents logiciels afin de profiter pleinement des avantages de chacun. La figure 6 résume un *workflow* complexe utilisant trois logiciels pour extraire des descripteurs audio, en créer des représentations et présenter ensuite le travail dans un logiciel de PAO ou de mise en forme. La partie supérieure de la figure représente les étapes de l'analyse acoustique et musicale et la partie inférieure, les formats de fichiers utilisables dans chacune des parties. Les flèches et les encadrements épais soulignent les étapes principales pour le musicologue. Une des principales difficultés réside dans le transfert des données d'un logiciel à l'autre, souvent limité par les fonctions d'importation et d'exportation. Lors du développement du logiciel EAnalysis, j'ai veillé à intégrer l'importation de différents formats permettant ainsi d'en faire un espace de travail efficace pour le chercheur.

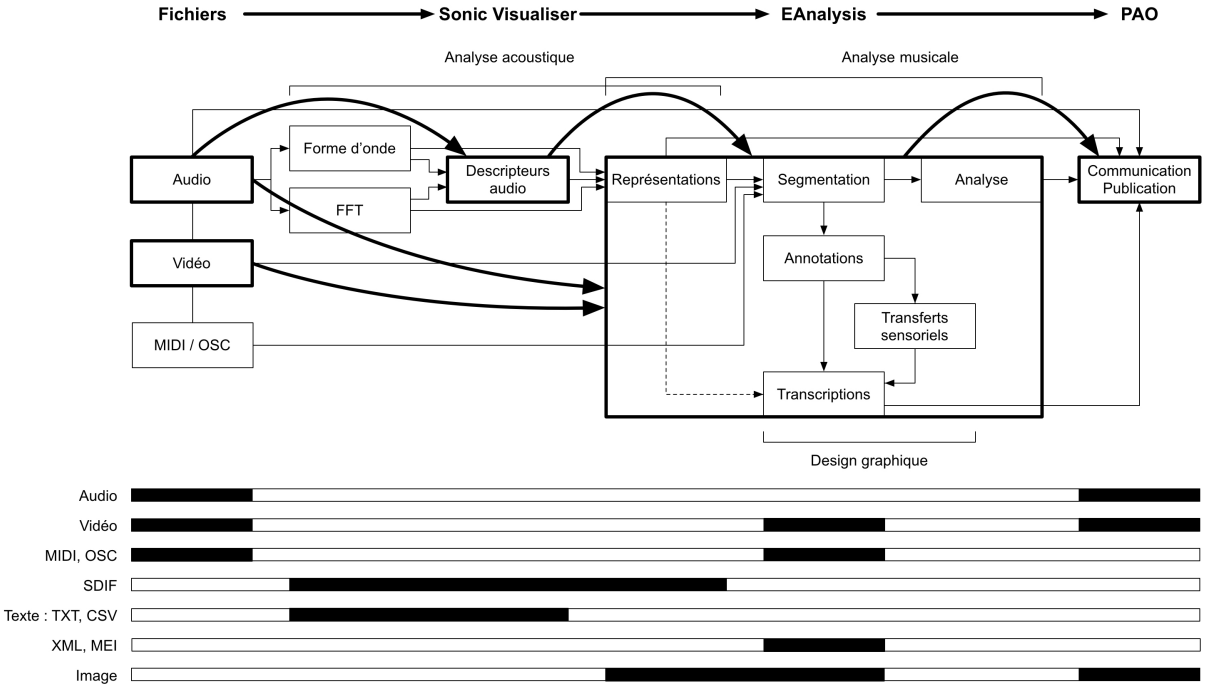

Figure 6 : Exemple de *workflow* en analyse musicale avec trois logiciels (Sonic Visualiser, EAnalysis et un logiciel de PAO).

En informatique, les données sont distinctes des interprétations, analyses et représentations que l'on peut en faire. Cette séparation facilite l'usage de méthodes d'analyse et de représentation différentes. Toutefois, en contrepartie, elle occasionne un risque de mauvaise interprétation des données. Par exemple, le descripteur de barycentre spectral (*spectral centroid*) produit généralement des valeurs très élevées sur les passages silencieux ou presque silencieux d'une œuvre. Il convient donc de les supprimer avant toute représentation afin que le résultat soit pertinent.

Après avoir présenté les technologies, les méthodes de calcul et les limites de l'usage de l'audionumérique en analyse, il convient de se pencher sur les représentations qui vont permettre une interprétation musicale du sonore.

---

[39] Le terme « *workflow* » (« flux de travaux ») désigne la succession des tâches à effectuer. Dans le cas présenté ici, le *workflow* décrit les différentes actions réalisées par le musicologue avec les logiciels de l'importation des fichiers à la création des représentations.

[40] Pierre Couprie, « L'analyse de la musique mixte : vers une redéfinition des "workflows" en musicologie », *in* M. Battier (éd.) *Regards sur les musiques mixtes*, *Portraits Polychromes*, n°23 hors-série, Paris, INA, 2016, p. 231-247.

II. TECHNIQUES DE REPRESENTATION EN ANALYSE MUSICALE

La réduction des données et les représentations composites

La réduction est utilisée lorsque l'on souhaite mettre en perspective plusieurs données pour analyser leur corrélation ou leur similarité. La visualisation d'un nombre trop important de jeux de données risque de rendre leur interprétation difficile. De plus, une mauvaise réduction de données risque aussi de causer des erreurs dans leur interprétation. La figure 7[41] présente un exemple de représentation composite – association de plusieurs visualisations – dont certaines sont réalisées à partir d'une réduction des données. Cette représentation est issue d'un projet de recherche sur l'analyse de l'interprétation de la musique acousmatique[42]. Les données proviennent de deux sources : les œuvres sur support audio et les données captées à partir de la console de mixage lors de l'interprétation de l'œuvre sur support. Cette figure comprend cinq visualisations de données :

1. le sonagramme ;
2. une représentation de l'évolution du timbre (BStD[43]) à partir de trois descripteurs audio sur une seule courbe ;
3. une visualisation de la position des 24 potentiomètres sur la table de mixage ;
4. la même visualisation, mais reportée sur le plan de l'implantation des haut-parleurs permettant d'étudier la répartition du son en fonction des différents plans sonores ;
5. une réduction de données de l'évolution des 24 potentiomètres sur trois courbes (trois groupes de huit potentiomètres[44]) : la moyenne quadratique[45] en épaisseur et la variance en trait fin. La moyenne quadratique permet d'obtenir une valeur plus pertinente que la moyenne arithmétique dans le cas de l'ouverture de potentiomètre. La variance offre une visualisation de la répartition des valeurs sur les huit potentiomètres de chaque groupe[46].

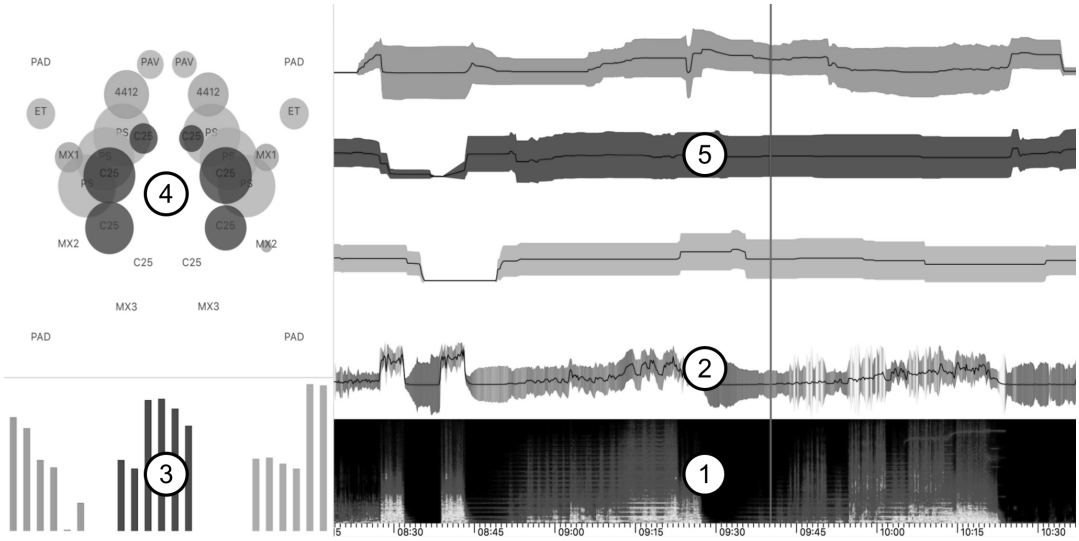

Figure 7 : Exemple de représentation composite utilisant la réduction de données en analyse de l'interprétation de la musique acousmatique.

[41] Il s'agit ici d'une image d'un instant d'une représentation animée.

[42] La musique acousmatique est composée en studio et interprétée sur un ensemble de plusieurs dizaines de haut-parleurs positionné autour et dans le public. L'interprète créé des plans sonores ou des déplacements en répartissant le son sur les différents haut-parleurs à partir d'une console de mixage. Le projet ici présenté est le résultat d'une collaboration entre la compagnie musicale Motus et l'Institut de recherche en musicologie (IReMus). Pour plus de détails sur le projet, voir Nathanaëlle Raboisson, Pierre Couprie, « Une expérience de captation et d'analyse de l''interprétation acousmatique », *in* P. Couprie et al. (éd.) *Journées d'informatique musicale*, Paris, Collegium Musicæ, 2017, https://jim2017.sciencesconf.org/data/Nathanaelle_Raboisson2017aa.pdf).

[43] Cette visualisation (*Brightness Standard Deviation*) sera présentée plus en détail dans la suite de l'article.

[44] Les potentiomètres sont groupés par huit, cela correspond aux trois groupes de haut-parleurs implantés dans la salle de concerts (deux couronnes de huit haut-parleurs et un groupe de haut-parleurs d'effets).

[45] La moyenne quadratique est le résultat de la racine carrée de la moyenne arithmétique des carrés des nombres.

[46] Dans ce cas, une variance élevée indique une disparité de réglages sur les huit potentiomètres de chaque groupe et une valeur basse indique une grande proximité de réglages.

Cette représentation complexe permet d'étudier les différents gestes mis en jeu dans l'interprétation de la musique acousmatique : les gestes intérieurs du compositeur (1 et 2), les gestes instrumentaux[47] produits sur la console de mixage (3) et les gestes de spatialisation perceptibles par le public (4 et 5). Cette représentation se compose de visualisations du sonore qui deviennent de plus en plus courantes dans le domaine de l'analyse musicale. Je vais maintenant présenter brièvement trois d'entre elles : le sonagramme, le diagramme de flux et la matrice similarité.

Les sonagrammes

Deux types de sonagrammes sont particulièrement adaptés à l'analyse musicale : le sonagramme linéaire et le sonagramme logarithmique ou par ondelette. La figure 8 présente deux exemples du même extrait audio :

1. la représentation linéaire des fréquences (à gauche) favorise la partie supérieure du spectre (au-delà de 5 kHz). Elle est parfaitement adaptée à l'étude des musiques électroacoustiques ou des œuvres essentiellement basées sur un travail spectral ;
2. la représentation logarithmique (à droite) met en valeur la partie basse du spectre qui correspond aux fréquences fondamentales des hauteurs musicales. Elle est adaptée à une analyse de la hauteur ou d'un suivi mélodique.

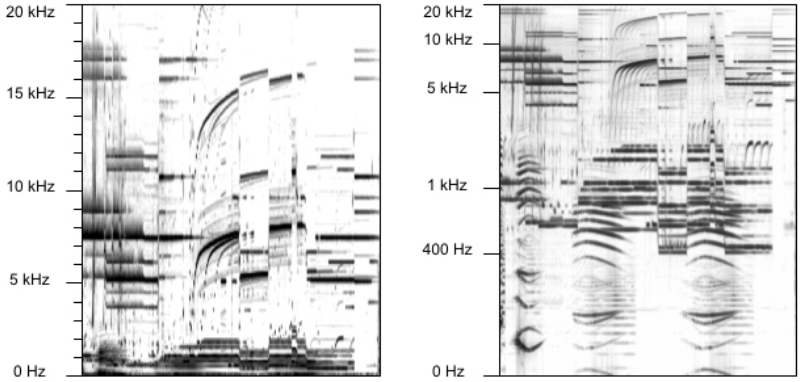

Figure 8 : Sonagramme linéaire[48] (à gauche) et sonagramme logarithmique (à droite) à partir d'un extrait de la partie 1 de *Son Vitesse-Lumière* de François Bayle.

Le logiciel The Snail, développé par l'Ircam pour l'accordage des instruments de musique, permet d'obtenir une représentation animée des hauteurs musicales (figure 9). Il utilise une représentation en spirale, chaque tour représentant une octave et les hauteurs étant distribuées sous la forme de tranches. Ce logiciel peut s'avérer très utile lors de l'étude de polyphonies de hauteurs ou dans le suivi de la variation de mélodies. Dans le cas d'*Adnos I* de la compositrice Éliane Radigue, je l'ai utilisé pour identifier les fréquences mobiles dans une polyphonie spectrale[49].

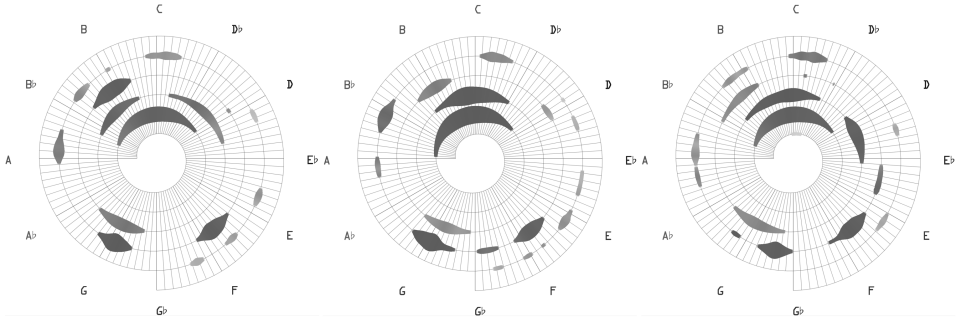

Figure 9 : Représentation des hauteurs de trois extraits d'*Adnos I* d'Éliane Radigue (38'01, 38'02, 38'03) réalisés avec le logiciel The Snail.

[47] Je reprends ici la terminologie du geste instrumental de Cadoz (Claude Cadoz, « Instrumental gesture and musical composition », *Proceedings of the International Computer Music Conference (ICMC)*, 1988, Cologne, p. 1-12).

[48] Réglages utilisés : taille de la fenêtre 2048, pas de la fenêtre 75%, type de fenêtre *hanning*.

[49] Pierre Couprie, « Immanence de la forme dans le cycle *Adnos* d'Éliane Radigue », *actes du colloque international Éliane Radigue*, Paris, en préparation.

Les diagrammes de flux

Les diagrammes de flux permettent de représenter des données temporelles sur plusieurs dimensions (par exemple, la position Y de la courbe, l'épaisseur et sa couleur). Les courbes de la zone 5 de la figure 7 utilisent trois diagrammes de flux résumant 24 séries de données en seulement trois courbes de deux paramètres : la moyenne quadratique (épaisseur) et la variance (position Y). Mikhail Malt et Emmanuel Jourdan[50] proposent de regrouper trois descripteurs audio sur une seule courbe afin de représenter l'évolution du timbre[51]. Ce graphique, nommé *brightness standard deviation* (BStD) combine l'amplitude RMS (couleur), la brillance (valeur Y) et l'écart-type ou la variance spectrale (épaisseur). La figure 10 représente l'évolution du timbre du *Prélude à l'après-midi d'un faune* dans deux interprétations (Ansermet en haut et Boulez en bas). Outre la différence de durée (env. 15 secondes), nous observons la variation du tempo (léger décalage des zones noires correspondant aux intensités fortes et des pics de la variance) et la plus grande uniformité de la variance dans la seconde interprétation qui traduit une variation moins importante de l'équilibre des pupitres[52].

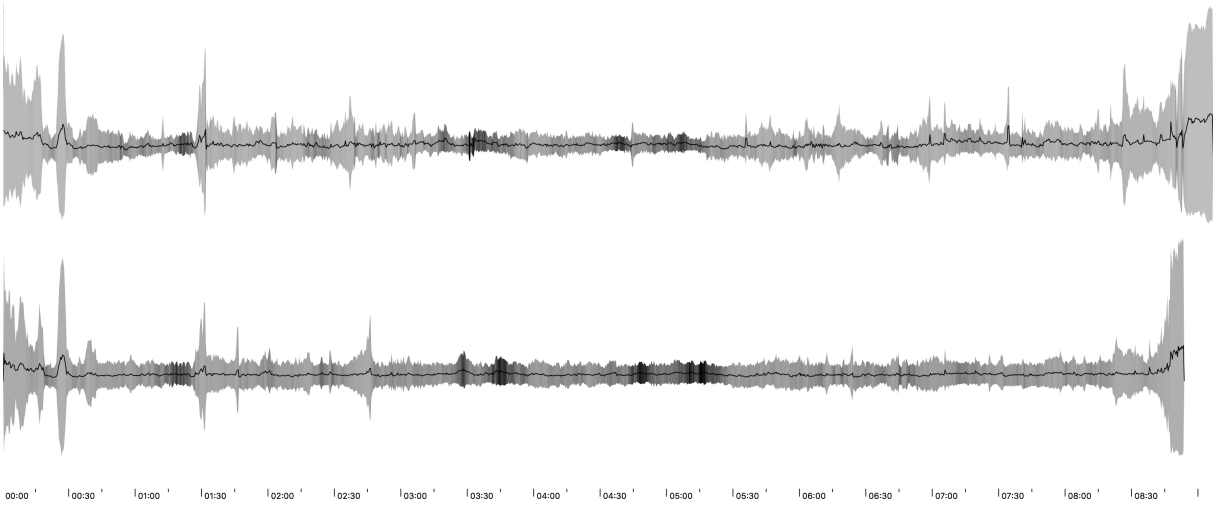

Figure 10 : BStD (amplitude RMS, brillance et variance spectrale)
du *Prélude à l'après-midi d'un faune* de Claude Debussy réalisée à partir de deux interprétations[53].

La matrice de similarité

En mathématiques, une matrice est un tableau contenant des nombres. Ces nombres peuvent être le résultat de différents calculs. Une matrice de similarité est réalisée à partir d'un calcul de distance entre une suite[54] de nombres. Prenons les valeurs obtenues avec l'extraction d'un descripteur audio – une suite de nombres –, il est alors possible de calculer les distances entre chaque valeur de cette suite. L'étape 1 de la figure 11 représente une matrice contenant les distances entre chaque valeur d'une suite de huit nombres présente en abscisse et en ordonnée. Comme le résultat obtenu est difficilement interprétable, on effectue ensuite une opération de *mapping* pour faire correspondre une couleur à chaque valeur de distance (étape 2). On obtient alors une matrice colorée (étape 3) sur laquelle il est facile de repérer des formes qui signalent les parties similaires – par exemple, les valeurs « 1 et 2 » du début de la liste sont similaires aux valeurs « 1 et 2 » présentes vers la fin, d'où les quatre blocs foncés formant les angles d'un carré sur la matrice de couleurs.

---

[50] Mikhail Malt, Emmanuel Jourdan, « Le "BStD" – une représentation graphique de la brillance et de l'écart type spectral, comme possible représentation de l'évolution du timbre sonore », *in* X. Hascher, M. Ayari, J.-M. Bardez (éd.) *L'analyse musicale aujourd'hui*, Paris, Delatour, 2015, p. 111- 128.

[51] Sur la multidimensionnalité de la perception du timbre, voir : Stephen McAdams, *Perception et cognition de la musique: les conférences Alphonse Dupront*, Paris, J. Vrin, 2015, p. 57-103

[52] Cet aspect peut aussi être en partie dû au travail de mixage de l'ingénieur du son.

[53] Interprétations (de haut en bas) : orchestre de la Suisse Romande (Ernest Ansermet) en 1957, Cleveland Orchestra (Pierre Boulez) en 1991. Paramètres de calcul : taille de la fenêtre 2048, pas de la fenêtre 75%, type de fenêtre hanning.

[54] J'emploie ici le terme « suite » pour désigner une liste de valeurs – par exemple « 1 2 3 4 » – et le terme « tableau » pour désigner plusieurs suites de valeurs organisées en tableau. En informatique, on les nomme respectivement tableau et matrice. Je continuerai toutefois d'utiliser le terme « suite » afin de faciliter l'explication.

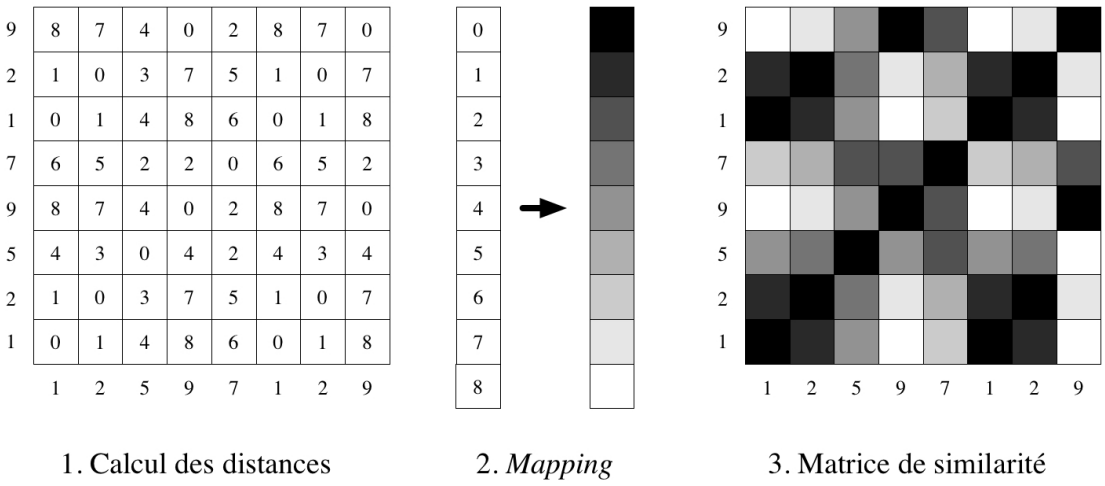

Figure 11 : Le principe de réalisation d'une matrice de similarité à partir d'une liste de valeurs.

En musicologie, les matrices de similarité sont particulièrement efficaces pour l'analyse formelle[55] ou pour faciliter la navigation dans des enregistrements audio de longue durée. On applique généralement le calcul de similarité au sonagramme ou à un ensemble de descripteurs audio comme ceux utilisés dans la BStD ou directement à la FFT. La figure 12 montre la matrice de similarité (au centre) réalisée à partir des descripteurs d'amplitude RMS, de barycentre et de variance spectrale calculés sur les 63 minutes du *Fil de la vie* de Pierre Henry. Comparée au sonagramme (en haut), il est aisé de constater qu'elle permet d'interpréter plus facilement les parties (en bas) sous la forme de colonnes verticales. De plus, en comparant les alternances de couleurs composant chaque colonne, il est aussi possible de repérer les parties qui sont probablement similaires sur le plan du timbre – par exemple les cases noires de la segmentation du bas. Charge ensuite au musicologue de confirmer ou d'infirmer la segmentation en cherchant les causes du résultat de la matrice.

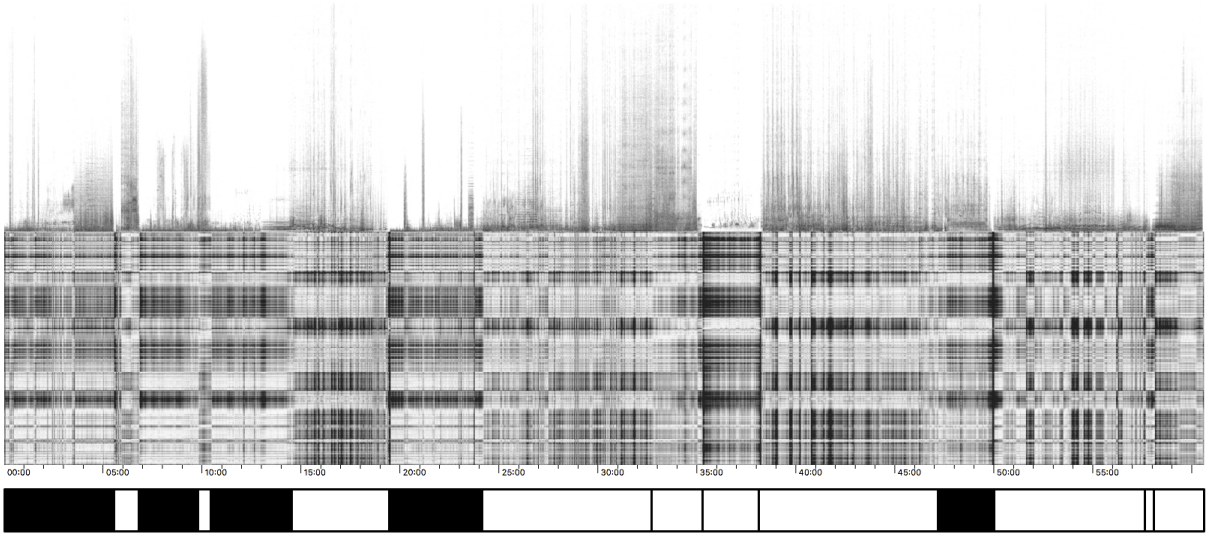

Figure 12 : Sonagramme (haut) et matrice de similarité[56] (centre) calculés sur *Le fil de la vie* de Pierre Henry et segmentation formelle (bas) interprétée à partir de la matrice.

[55] Par exemple : Momilani Ramstrum, « Philippe Manoury's Opera K... », *in* M. Simoni (éd.), *Analytical Methods of Electroacoustic Music*, New York, Routledge, 2006, p. 263-264.

[56] La matrice est calculée à partir de l'amplitude RMS, du barycentre et de la variance spectrale. Paramètres de calcul : taille de la fenêtre 2048, pas de la fenêtre 75%, type de fenêtre *hanning*. Le sonagramme est calculé avec les mêmes paramètres.

Les matrices de similarité permettent aussi de comparer des enregistrements[57] comme dans l'étude de Zattra et Orio sur les différences entre la version d'origine et la reconstruction de l'œuvre *Stria* de John Chowning[58].

CONCLUSION

Dans cet article, j'ai brièvement présenté quelques techniques permettant de créer des représentations à partir de calculs réalisés sur le signal audio. Malheureusement, il manque de nombreux travaux comme ceux réalisés par Chouvel, Bresson et Agon sur les sonagrammes différentiels[59] ou ceux de Cook et son équipe sur l'utilisation de la corrélation hiérarchique pour l'analyse de l'interprétation à partir d'enregistrement audio[60].
L'ensemble des travaux présentés permet toutefois d'envisager de nouvelles perspectives. Quatre domaines me semblent nécessiter une attention particulière. Le premier concerne la gestion des fichiers informatique. La musicologie systématique manque cruellement d'un format d'échange de fichiers qui serait reconnu par l'ensemble des logiciels proposant des fonctions d'analyse musicale. Le format MEI[61] contient quelques balises en analyse, mais elles restent très limitées, centrées sur le symbolique et fortement liées à la dimension temporelle de la partition ou de la transcription. Le deuxième domaine concerne les descripteurs audio qui sont souvent issus de projets commerciaux et adaptés aux musiques faites de notes et de rythmes. Or, même dans le corpus des musiques populaires, le travail sur le son engendre des spectres complexes ou bruités qui résistent bien souvent aux descripteurs. Par ailleurs, certains états de la matière sonore se situent entre deux catégories musicales – comme le parlé-chanté – et aucun descripteur n'est réellement adapté à leur étude. Un champ de développement de nouveaux descripteurs pourrait donc être envisagé. Enfin, deux derniers domaines restent encore très largement inexplorés : l'analyse de corpus et l'apprentissage machine.
L'utilisation de l'audionumérique et de l'informatique en analyse musicale, et plus largement en musicologie systématique, est probablement le vecteur qui permettra à la musicologie d'entrer dans une nouvelle étape de son évolution en devenant une science du XXI^{e} siècle.

Références

Claude Cadoz, « Instrumental gesture and musical composition », *Proceedings of the 1988 International Computer Music Conference (ICMC)*, 1988, Cologne, p. 1-12.
Richard Cann, « An Analysis/Synthesis Tutorial: Part 1 », *Computer Music Journal*, vol. 3, n° 3, 1979, p. 6–11.
Jean-Marc Chouvel, Jean Bresson, Carlos Agon, « L'analyse musicale différentielle: principes, représentation et application à l'analyse de l'interprétation », *in Electroacoustic Music Studies Network*, Leicester, De Montfort University, 2007, http://www.ems-network.org/IMG/JMChouvelEMS07/index.html.
Michael Clarke, Frédéric Dufeu, Peter Manning, « Introducing TaCEM and the TIAALS Software », *International Computer Music Conference. International Developments in Electroacoustics*, Perth, 2013, p. 47-53.
Robert Cogan, *New Images of Musical Sound*, Cambridge, Harvard University Press, 1984.
Nicholas Cook, *Beyond the Score*, Oxford, Oxford University Press, 2014.
Nicholas Cook, Daniel Leech-Wilkinson, « A musicologist's guide to Sonic Visualiser »,
http://www.charm.rhul.ac.uk/analysing/p9_1.html (02/03/2018).
Pierre Couprie, « iAnalyse : un logiciel d'aide à l'analyse musicale », *Journées d'Informatique Musicale*, Albi, 2008, p. 115-121.
Pierre Couprie, « Le développement d'un outil d'aide à l'analyse musicale : bilan et perspectives musicologiques », *in* N. Marty (éd.) *Musiques électroacoustiques. Analyses – Écoutes*, Paris, Delatour, 2016, p. 157-169.
Pierre Couprie, « L'analyse de la musique mixte : vers une redéfinition des "workflows" en musicologie », *in* M. Battier (éd.) *Regards sur les musiques mixtes*, Portraits Polychromes, n° 23 hors-série, Paris, INA, 2016, p. 231-247.
Pierre Couprie, « Immanence de la forme dans le cycle *Adnos* d'Éliane Radigue », *actes du colloque international Éliane Radigue*, Paris, en préparation.

[57] Dans ce cas, il faut d'abord procéder à un alignement temporel des valeurs (algorithme de déformation temporelle dynamique) afin que le calcul de similarité porte sur les mêmes index temporels.
[58] Laura Zattra, Nicola Orio, « ACAME – Analyse comparative automatique de la musique électroacoustique », *Musimédiane*, n°4, 2006, http://www.musimediane.com/numero4/LZattra/index.html.
[59] Jean-Marc Chouvel, Jean Bresson, Carlos Agon, « L'analyse musicale différentielle : principes, représentation et application à l'analyse de l'interprétation », *in Electroacoustic Music Studies Network*, Leicester, De Montfort University, 2007, http://www.ems-network.org/IMG/JMChouvelEMS07/index.html.
[60] Nicholas Cook, *op. cit.* , p. 176-223.
[61] MEI (*Music Encoding Initiative*) est un format de partition ouvert en XML plus complet que le format MusicXML et particulièrement adapté à la philologie musicale.

Myriam Desainte-Catherine, Jean-Louis Di Santo, « L'acousmoscribe, un éditeur de partitions acousmatiques », *Electroacoustic Music Studies Network*, Buenos Aires, 2009, http://www.ems-network.org/ems09/papers/disanto.pdf (02/03/2018).
Andrew Earis, Craig Sapp, « Tools developed by the CHARM Mazurka project », http://www.charm.rhul.ac.uk/analysing/p9_4.html (02/03/2018).
Emmanuel Favreau, Yann Geslin, Adrien Lefèvre, « L'Acousmographe 3 », *Journées d'Informatique Musicale*, Rennes, 2010, p. 105-111.
Tae Hong Park, Zhiye Li, Wen Wu, « Easy does it: The Electro-Acoustic Music Analysis Toolbox », *10th International Society for Music Information Retrieval Conference,* Kobe, 2009, p. 693-698.
Philippe Lalitte, « Du son au sens : vers une approche sub-symbolique de l'analyse musicale assistée par ordinateur », *Musurgia*, vol. 18, n° 1-2, 2011, p. 99-116.
Philippe Lalitte, *Analyser l'interprétation de la musique du XX^e^ siècle*, Paris, Hermann, 2015.
Émile Leipp, *Acoustique et musique*, Paris, Mines ParisTech, 5/2010.
Wim van der Meer, « Praat manual (for musicologists) », https://wimvandermeer.wordpress.com/praat-manual-for-musicologists/ (02/03/2018).
Alexander Lerch, *An Introduction to Audio Content Analysis. Applications in Signal Processing and Music Informatics*, Hoboken, Wiley, 2012.
Mikhail Malt, Emmanuel Jourdan, « Le "BStD" – une représentation graphique de la brillance et de l'écart type spectral, comme possible représentation de l'évolution du timbre sonore », *in* X. Hascher, M. Ayari, J.-M. Bardez (éd.) *L'analyse musicale aujourd'hui*, Paris, Delatour, 2015, p. 111- 128.
Stephen McAdams, *Perception et cognition de la musique: les conférences Alphonse Dupront*, Paris, J. Vrin, 2015.
Stephen McAdams, Philippe Depalle, Eric Clarke, « Analyzing Musical Sound », *in* N. Cook, E. Clarke (éd.) *Empirical Musicology Aims, Methods, Prospects*, Oxford, Oxford University Press, 2004, p. 157-196.
Jeanne Miramon-Bonhoure, Amine Beyhom, « Petit manuel pratique pour débutants – 1ere partie : Analyses de systèmes intervalliques et d'échelles », http://ethnomusika.org/2011/11/14/praat-petit-manuel-pratique-pour-debutants-1ere-partie-analyses-de-systemes-intervalliques-et-d'echelles/ (02/03/2018).
Kerstin Neubarth, Mathieu Bergeron, Darrell Conklin, « Associations Between Musicology and Music Information Retrieval », *in* M. Ogihara (éd.) *Proceedings of the International Society for Music Information Retrieval Conference* (ISMIR), Miami, 2011, p. 429-434.
Michela Magas, Xavier Serra (éd.), *Roadmap for Music Information Research*, MIReS, 2013, p. 14, http://www.mires.cc/files/MIRES_Roadmap_ver_1.0.0.pdf (02/03/2018).
Nathanaëlle Raboisson, Pierre Couprie, « Une expérience de captation et d'analyse de l'interprétation acousmatique », *in* P. Couprie et al. (éd.) *Journées d'informatique musicale*, Paris, Collegium Musicæ, 2017, https://jim2017.sciencesconf.org/data/Nathanaelle_Raboisson2017aa.pdf
Momilani Ramstrum, « Philippe Manoury's Opera K... », *in* M. Simoni (éd.) *Analytical Methods of Electroacoustic Music*, New York, Routledge, 2006, p. 239-274.
Laura Zattra, Nicola Orio, « ACAME – Analyse comparative automatique de la musique électroacoustique », *Musimédiane*, n° 4, 2006, http://www.musimediane.com/numero4/ LZattra/index.html.