

# Let's See Music

Étude Bibliographique - Caractérisation du Timbre  
22.07.2016

## 1 Introduction

L'expérience humaine et émotionnelle d'un phénomène passe par l'expérience des sens. Afin de maximiser le potentiel d'un phénomène, on peut chercher à maximiser son impact sensoriel, en termes d'intensité ou de diversité. Dans le cas de la musique, par exemple, l'expérience auditive est évidente, et nombre de personnes ont oeuvré et continuent aujourd'hui pour améliorer constamment la qualité de celle-ci. Cependant, on peut noter d'autres sens possibles, notamment le sens du toucher (avec les vibrations du sol par exemple) ou le sens visuel (avec la vision des musiciens par exemple).

*Let's See Music* s'inscrit dans cette démarche en proposant une visualisation spécifique et dynamique de la musique écoutée par l'utilisateur. [1] Afin d'améliorer l'expérience utilisateur, il pourrait être intéressant de différencier les différents instruments présents dans le morceau. La suite de ce document s'attache donc à décrire et caractériser un paramètre déterminant pour la reconnaissance d'un instrument : le timbre.

## 2 Contexte du projet

Initialement développé dans le cadre d'un projet scolaire au lycée, *Let's See Music* est un lecteur de musique permettant une expérience originale en proposant des animations spécifiquement et dynamiquement liées à la musique. Le logiciel, alors conçu en C++, permet de lire des fichiers au format .flac ou .mp3 et propose cinq animations différentes, le tout étant encapsulé dans une interface graphique conviviale et disposant même d'un module de reconnaissance vocale.

Ce projet entre aujourd'hui dans une nouvelle phase de conception et d'amélioration, et ce dans le but d'enrichir l'expérience utilisateur, avec toujours le même souhait profond : vous faire porter un nouveau regard sur vos musiques ! Dans ce but, une des fonctionnalités en phase d'étude est la distinction d'instruments. En étant capable de discerner les différents instruments présents dans le morceau, nous serions alors en mesure de proposer des animations encore plus en phase avec la musique.

## 3 Timbre d'un instrument

### 3.1 Définition

Par définition, le timbre est composé de l'ensemble des caractéristiques sonores permettant l'identification d'un instrument. [2] On associe classiquement la notion de spectre

à celle de timbre. Qu'est-ce que le spectre d'un instrument ? Ou plutôt devrions-nous dire : qu'est-ce que le spectre d'un signal ?

Le spectre d'un signal est la donnée de l'intensité de chaque composante fréquentielle présente dans le signal étudié. La théorie de Fourier précise que tout signal périodique est décomposable sous forme d'une série dite de Fourier, c'est-à-dire que tout signal périodique (dont les ondes stationnaires utilisées dans les instruments) peut être écrit comme une somme (potentiellement infinie) de signaux sinusoïdaux. Un tel spectre peut être estimé de manière satisfaisante en utilisant le module d'une opération de Transformée de Fourier Discrète, comme par exemple une FFT. On remarquera cependant qu'une telle opération a deux conséquences directes : la perte de l'information de phase d'une part, et la périodisation du spectre d'autre part. [3]

Dans le cas d'une note jouée, on appellera la première composante fréquentielle non nulle la fondamentale, et les suivantes seront appelées harmoniques ou partielles. On admettra dans la suite que la richesse d'un instrument, sa coloration et donc, dans une certaine mesure, son timbre, sont profondément liés à la présence d'harmoniques ou de partielles et à la répartition énergétique entre celles-ci. [9] Pour autant, la seule donnée du spectre suffit-elle à caractériser le timbre et donc à reconnaître un instrument ?

### 3.2 Marqueurs caractéristiques et représentation vectorielle

Diverses études se sont penchées sur le sujet de la proximité des timbres, et l'une d'elle révèle par exemple que la seule donnée du spectre du régime stationnaire du signal n'est pas suffisante pour identifier correctement un instrument. Ainsi, en jouant un son de piano à l'envers, c'est-à-dire en inversant l'attaque (régime transitoire initial) et l'atténuation (régime transitoire final), les chances d'identifier correctement l'instrument diminuent drastiquement [4], jusqu'à confondre piano et accordéon par exemple. De même, retirer artificiellement les régimes transitoires pour ne garder que le régime stationnaire augmente grandement la difficulté de reconnaître l'instrument utilisé. [9] L'attaque semble donc être un marqueur important, en plus du spectre. D'autres marqueurs pourraient encore compléter notre représentation en affinant le modèle, mais avec une influence plus ou moins grande. La notion de phase, par exemple, semble peu pertinente dans l'étude du timbre, et peut raisonnablement être ignorée ici. [9]

Puisque le timbre est une entité caractérisée par plusieurs grandeurs, elle peut être qualifiée de multidimensionnelle. On peut donc raisonnablement se pro-

poser d'utiliser une représentation vectorielle du timbre, permettant ainsi d'accéder à toute la puissance de l'algèbre linéaire. Cette représentation s'avère particulièrement efficace pour l'analyse de la perception humaine du timbre. [5] Diverses études ont été menées et il semblerait qu'une représentation dans un espace préhilbertien à trois dimensions soit la plus performante. Les trois dimensions proposées représentent alors les grandeurs suivantes : la densité spectrale d'énergie [3], la présence ou non d'énergie à basse amplitude et hautes fréquences (ie un signal non voisé, s'apparentant à du bruit), la présence ou non d'une synchronisation des régimes transitoires des hautes harmoniques ainsi que la fluctuation spectrale au cours du temps. [7]

Un instrument est caractérisé par son timbre, ici représenté par un vecteur caractéristique de dimension trois.

## 4 Application

### 4.1 Différenciation des instruments

Ainsi, nous pouvons d'ores et déjà réfléchir à l'intégration de ces notions au sein de *Let's See Music*. Nous pouvons imaginer que le morceau soit découpé selon des fenêtres temporelles, et que chaque fenêtre soit analysée afin de générer l'animation correspondante, et ce, en même temps que la musique est jouée.

Sur chaque fenêtre, il est facile de déterminer le spectre, via une simple FFT comme précisé en 3. Nous pouvons alors calculer les différentes coordonnées de notre vecteur caractéristique, et le placer dans notre espace. On notera que l'utilisation d'un spectrogramme peut être un bon moyen de visualiser les fluctuations temporelles de la densité spectrale d'énergie, y compris selon les états de la note (attaque, régime stationnaire, atténuation). À l'aide d'un algorithme de *clustering* [8], il est possible de faire des regroupements selon des instruments, au sens du maximum de vraisemblance (c'est-à-dire au sens du minimum de la distance).

On notera qu'il est alors nécessaire de disposer de centres (facilement initialisés via une opération de seuillage sur la vraisemblance par exemple, et facilement mis à jour via la minimisation du rayon moyen des groupes par exemple), et que les résultats peuvent dépendre de la norme choisie dans notre espace préhilbertien (des essais montrent cependant une influence assez limitée sur le résultat final).

### 4.2 Limites du modèle

Une complexité perdure néanmoins : afin d'avoir un modèle suffisamment précis, nous avons besoin de connaître les régimes transitoires, d'après 3.1. Autrement dit, nous devons pouvoir étudier chaque note dans sa totalité. Cela est malheureusement très peu trivial, d'autant plus que toutes les notes jouées n'ont pas forcément la même durée et peuvent se superposer.

Par ailleurs, la notion de vraisemblance entre timbres est très délicate. Deux instruments différents auront *a priori* deux timbres différents, par définition, y compris s'ils sont de la même famille. Deux violons auront par exemple

des dessins, des structures, des bois ou des cordes légèrement différents. [6] Un des intérêts de l'approche privilégiée par *Let's See Music* est qu'elle ne s'attache pas à reconnaître les instruments, mais seulement à les différencier. Cela permet de ne pas avoir à apprendre au système les différents timbres possibles pour chaque instrument, et permet au logiciel de gagner en autonomie.

Une autre difficulté apparaît alors : pour reconnaître des instruments, nous utilisons la notion de distance entre leurs vecteurs caractéristiques, mais est-ce réellement pertinent ? En effet, si deux instruments différents peuvent avoir un timbre similaire, un même instrument peut aussi avoir deux timbres légèrement différents. Le timbre varie selon l'état de la note (régime transitoire initial, régime stationnaire, régime transitoire final), la hauteur de la note, la technique de jeu (par exemple la technique du *palm mute* à la guitare, ou l'endroit du pincement des cordes) ou encore l'accord de l'instrument. [6] On peut ainsi se retrouver dans l'incertitude la plus totale, un instrument ressemblant trop à un autre, et un instrument ne se ressemblant pas assez à lui-même.

## 5 Conclusion

Distinguer des instruments de musique est un sujet extrêmement complexe, et l'identification d'instruments à partir de leur timbre a fait l'objet de nombreuses études. Nous avons néanmoins des pistes concernant la caractérisation et la représentation d'un timbre, pistes que nous pensons optimales au regard de ces mêmes études.

Cependant, le timbre est avant tout lié à notre perception humaine, et donc à notre ressenti, et l'étude du timbre se fait souvent sous une double approche, à la fois scientifique et socio-psychologique. Il n'est donc pas étonnant de conclure quant à l'incomplétude du modèle décrit dans ce document, même si celui-ci devrait néanmoins donner des résultats optimaux, au sens du rapport performances/-complexité.

Il reste encore à lever les difficultés rémanentes, soit en complétant notre modèle, soit en adaptant notre algorithme d'analyse musicale. Nous avons bon espoir d'atteindre des performances honorables permettant d'enrichir significativement l'expérience des utilisateurs de *Let's See Music*, pour peu que la difficulté d'analyse reste raisonnable.

## References

- [1] Killian Kems Arts, *Let's See Music*, consultable sur [killian.rkemps.fr/musique/lets-see-music](http://killian.rkemps.fr/musique/lets-see-music) (22.07.2016)
- [2] Wikipédia, *Timbre*, consultable sur [fr.wikipedia.org/wiki/Timbre\\_\(musique\)](http://fr.wikipedia.org/wiki/Timbre_(musique)) (22.07.2016)
- [3] PRANDONI Paolo & VETTERLI Martin. *Signal Processing For Communications*, CRC Press, 2008 - ISBN 978-1-4200-7046-0
- [4] LALITTE Philippe, *Aspects acoustiques et sensoriels du bruit*, 2008, consultable sur [lead.serv.u-bourgogne.fr/files/publications/000473-aspects-acoustique-et-sensoriel-du-bruit.pdf](http://lead.serv.u-bourgogne.fr/files/publications/000473-aspects-acoustique-et-sensoriel-du-bruit.pdf)

- [5] CHOWING John A., GREY John M., MOORER James A., RUSH Loren & SMITH Leland C., *Center for Computer Research in Music and Acoustics*, 1979, consultable sur [ccrma.stanford.edu/ aj/archives/docs/all/908.pdf](http://ccrma.stanford.edu/aj/archives/docs/all/908.pdf) (22.07.2016)
- [6] SÉRY Christelle, *De la notion de timbre, l'exemple de la guitare*, 2003, consultable sur [christellesery.fr/et\\_autres\\_jeux\\_files/De%20la%20notion%20de%20timbre.pdf](http://christellesery.fr/et_autres_jeux_files/De%20la%20notion%20de%20timbre.pdf) (22.07.2016)
- [7] GREY John M., *Multidimensionnal perceptual scaling of musical timbres*, 1976, consultable sur [aum.dartmouth.edu/ mcasey/m102/GreyMultidimensionalScaling.pdf](http://aum.dartmouth.edu/~mcasey/m102/GreyMultidimensionalScaling.pdf) (22.07.2016)
- [8] KANUNGO Tapas, MOUNT David M., NETANYAHU Nathan S., PIATKO Christine D., SILVERMAN Ruth & WU Angela Y., *An efficient k-means algorithm*, 2002, consultable sur [cs.umd.edu/ mount/Projects/KMeans/pami02.pdf](http://cs.umd.edu/mount/Projects/KMeans/pami02.pdf) (22.07.2016)
- [9] WEDIN Lage & GOUDE Gunnar, *Dimension analysis of the perception of instrumental timbre*, 1972, consultable sur [thirdworld.nl/dimension-analysis-of-the-perception-of-instrumental-timbre](http://thirdworld.nl/dimension-analysis-of-the-perception-of-instrumental-timbre) (22.07.2016)