# Détection d’onset

Cette partie traite de la détection d’onset. Elle correspond au tout premier bloc du schéma fonctionnel.

Un onset fait référence au début d’une note de musique ou d’un autre son, du moment que l’amplitude du signal passe de zéro à un pic initial. Ainsi, après avoir isolé ces notes unes à unes, nous pourrons faire une analyse des tons et des octaves pour chaque note.

|  |
| --- |
| /Users/apple/Desktop/Capture d’écran 2016-01-29 à 10.43.03.png |
| Figure 1 - Onset |

Il existe une dizaine de méthodes de détection d’onsets. Cependant, nous ne les avons pas toutes codées. Celles-ci sont directement appliquées au signal audio. Il en résulte un signal possédant plusieurs pics de plus ou moins fortes amplitudes. Enfin, une fonction de localisation des pics est nécessaire pour extraire ces fameux onsets. Au cours de ces étapes, il est possible qu’il se produise des fausses détections ou des onsets non-détectés.

Voici le schéma général d’un algorithme de détection d’onset :

|  |
| --- |
| /Users/apple/Desktop/Capture d’écran 2016-01-29 à 10.43.30.png |
| Figure 2 - Algorithme de détection d'onset |

Dans notre projet, nous avons testé plusieurs fonctions de détection d’onset puis nous en avons combiné pour améliorer leur détection. Voici les principales approches et des explications correspondant aux méthodes que nous avons implémentées.

## Méthode liée aux caractéristiques temporelles du signal audio

### Energie du signal

La méthode la plus évidente et simple prendrait en compte les variations d’énergie du signal audio (sans transformation) pour détecter les onsets. Cette méthode est efficace pour des signaux comme des percussions, qui présentent une forte amplitude dès lors qu’un son est produit.

On pourra par exemple prendre la norme au carré d’une trame de signal, ce qui nous donnerai son énergie, faire la dérivée première pour voir les variations d’énergie sur le signal. Enfin prendre un seuil à partir duquel l’énergie est suffisamment grande pour pouvoir déterminer un onset.

**Inconvénient :**  Ce type de méthode ne fonctionne pas de façon optimale car certains onset ne sont pas détectés par rapport au seuil que l’on se fixe (fausse détection/sur-détection). De plus, dans le cas de notre projet qui prend en compte les cordes de la guitare, les onsets produits par celle-ci n’ont pas une variation en amplitude comparable à des instruments percussifs.

## Méthodes liées aux caractéristiques spectrales du signal audio

### Spectral Flux ou Spectral difference (domaine fréquentielle)

Cette méthode se rapproche de celle présentée précédemment, mais elle prend cette fois-ci en compte un outil : la transformée de Fourier discrète.

On commence par faire la transformée de Fourier du signal audio. Dans notre projet, nous avons découpé le signal audio en plusieurs séquences. Ce sont ces séquences qui seront analysées et traitées les unes à la suite des autres.

La transformée de Fourier nous donne pour un signal audio en entrée une bande de fréquence en ordonnée et le temps, représenté par des échantillons k, en abscisse.

Fréquences

Matrice de nombre complexes

Echantillons k du signal

Voici un spectrogramme, représentant de façon intuitive la présence des fréquences dans le temps d’un morceau de 25 secondes :

|  |
| --- |
|  |
| Figure 3 - Spectrogramme d'un morceau de 25 secondes |

La matrice contient des nombres complexes, porteurs de l’information de la phase et de l’amplitude pour chaque échantillon k et pour chaque fréquence. Cette transformée de Fourier peut être définie comme un groupe d’oscillateurs sinusoïdaux de différentes fréquences avec des amplitudes et phases variant dans le temps.

Cette méthode détermine la variation de l’amplitude en puissance du signal pour chaque échantillon de fréquence (on ne prend plus l’énergie du signal audio mais l’amplitude de la transformée de Fourier).

La valeur absolue de la transformée de Fourier (amplitude) est dérivée pour obtenir la variation pour chaque fréquence entre l’échantillon k et k-1. On somme ensuite pour chaque échantillon de temps k ces variations pour pouvoir avoir une représentation temporelle de ces changements.

Une forte variation en amplitude à instant k nous donnera donc un pic.

**Inconvénient :** L’efficacité de cette méthode décroit pour des signaux non-percussifs (ce qui est également le cas pour la méthode présentée précédemment).

### Phase deviation (domaine de la phase)

Contrairement à la méthode du spectral flux qui prend en compte la puissance du signal, la méthode de déviation de phase utilise la variation dans la phase du signal.

Elle fonctionne de la même manière que spectral flux sauf que l’on prend la phase du signal (angle) suite à la transformée de Fourier (que l’on recadre entre –π et π). On étudie donc la phase du signal pour une fréquence donnée sur l’ensemble des échantillons k du signal.

La première différence de phase entre un échantillon k et k-1 nous donne la fréquence instantanée (première dérivée). Le changement de variation de la fréquence instantanée (dérivée seconde) est un indicateur d’un possible onset. En effet, au cours de la partie à l'état stable du signal, ces oscillateurs (voir FFT) ont tendance à avoir des fréquences constantes. Par conséquent, la différence entre deux valeurs de phase consécutifs doit rester constante.

La variation de cette différence de phase nous donnera l’indication d’un possible onset (voir (1) dans la bibliographie).

### Complex domain

L’idée de cette méthode est de considérer simultanément les effets de ces deux grandeurs, l’amplitude et la phase, en prédisant des valeurs dans le domaine complexe. On calcule l’amplitude et la phase d’un échantillon (valeur cible) à partir de deux échantillons précédents en supposant que l’état est stable.

La valeur cible est définie par la forme polaire du nombre complexe correspondant à un échantillon de la FFT comme :

Capture%20d’écran%202016-02-01%20à%2009.48.06.png

= (amplitude de l’échantillon k-1) \*exp(somme entre la phase de k-1 et la différence de phase entre k-1 et k-2).

Capture%20d’écran%202016-02-01%20à%2010.05.18.png

Cette valeur est donc estimée avec une amplitude et un taux de changement de phase constant.

La valeur courante possède l’amplitude et la phase correspondant à son échantillon.

**

En mesurant la distance euclidienne entre la valeur cible et la valeur courante dans le domaine complexe, on est capable de mesurer la stationnarité pour le kème échantillon.

La déviation de phase est donnée par la formule suivante (n représente k) :



L’égalité de droite correspond à la variation de la fréquence instantanée de l’échantillon k et celle de gauche correspond à la variation de la fréquence instantanée de l’échantillon k- 1.

La déviation de phase courante est alors donnée par :



La phase de la valeur ciblée est pour rappel :

Capture%20d’écran%202016-02-01%20à%2010.05.18.png

En simplifiant cette distance euclidienne (notamment en forçant la formule de la cible visée sur l’axe des réels, voir doc 4 sur l’onset detection), on arrive à la formule suivante :

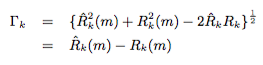


|  |
| --- |
| Capture%20d’écran%202016-02-01%20à%2010.40.26.png |
| Figure 4 - Simplification de la distance euclidiènne |

Lorsque la différence de phase entre deux échantillons successifs est nulle, c’est à dire quand l’égalité suivante est respectée :



alors :

****

Par conséquent, lorsque la valeur de la différence de phase est nulle (la prédiction est bonne), la différence complexe se réduit à la différence d’amplitude (différence d’énergie). Quand la différence de phase est non-nulle, un terme additionnel s’ajoute qui prend en compte la déviation de phase.

Finalement, on somme le terme Γk pour tous les échantillons k.

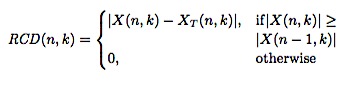
Capture%20d’écran%202016-02-01%20à%2011.27.58.png

Cette méthode donne une détection plus robuste que celles testées précédemment. (Voir articles (1), (2), (3), (4) et (5)).

### Rectified Complex Domain

Le problème avec la méthode Complex Domain est qu’elle ne fait pas de différence entre les augmentations et les diminutions de l’amplitude du signal, donc les onsets ne sont pas différenciables des offsets.

La méthode Rectified Complex Domain prend en compte les cas où la différence complexe représente une augmentation d’énergie, et donc un onset.

****

(Voir articles (2) et (4)).

## Méthodes liées à des analyses temps-fréquence/échelle de temps différentes (alternatives à la FFT).

Il existe d’autres alternatives aux méthodes d’analyse des coefficients de Fourier ou de l’énergie du signal comme l’utilisation d’échelles temps-fréquences différentes.

Une méthode utilise par exemple la décomposition en ondelettes (ondelette de Haar). Une autre est basée sur une représentation temps-fréquence de Cohen.

Cependant, nous n’avons pas développé ces méthodes car aux vues des résultats présentés dans les études réalisées sur l’onset detection, les méthodes mentionnées ci-dessus donnent déjà de très bons résultats.

(Voir articles (1), (7) et (8)).

## Probability Models

Une autre alternative possible que nous n’avons pas utilisé est de recourir aux méthodes statistiques. Elles sont basées sur l'hypothèse que le signal peut être décrit par un modèle de probabilité.

Un système peut alors être construit qui fait des analogies de probabilités sur les temps supposés des changements brusques dans le signal, étant donné les observations disponibles.

Le succès de cette approche dépend de la proximité de l'ajustement entre le modèle supposé, donc la distribution de probabilité décrite par le modèle, et la distribution des données observée, et peuvent être quantifiés utilisant des mesures de probabilité.

Ces modèles peuvent par la suite être perfectionnés par des techniques d'apprentissage automatique comme les réseaux de neurones.

(Voir articles (6)).

## Algorithme de détection des pics

Comme expliqué précédemment, on travaille échantillon par échantillon sur le signal. On commence par faire une transformée de Fourier discrète sur l’échantillon, on applique la fonction d’onset detection, puis on applique un filtre coupe-bande pour éliminer certains parasites sur le résultat de la fonction d’onset.

Après plusieurs tests avec les méthodes présentées ci-dessus, nous avons décidé de combiner deux méthodes. En effet on a pu constater que la méthode « Spectral flux » a tendance à oublier certaines onset alors que la méthode « Complex domain » au contraire nous donne des faux-positifs. Nous avons donc affecté à chacune de ces deux méthodes (que nous avons normalisé) un coefficient de pondération pour obtenir une fonction d’onset detection plus efficace.

Une fois la courbe représentant la fonction d’onset detection obtenue, nous avons établi une fonction donnant la moyenne locale sur le signal pour pouvoir récupérer les pics qui se trouvent au-dessus de cette moyenne.

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Partie sur l’OFFSET

Ci-dessous la comparaison entre deux fonctions d’onset detection sur une même chanson (avec un léger décalage dans le temps). On note que la méthode Complex domain donne des pics plus précis.

|  |
| --- |
| ../../test/CD_specFlux.jpg |
| Figure 5 - Comparaison de deux méthodes (Spectral flux & Complex domain) |

Ci-dessous la comparaison entre deux fonctions d’onset detection sur une même chanson (avec un léger décalage dans le temps). On constate que la méthode Complex domain donne des pics là encore plus précis, la méthode de phase deviation donnant des pics de faibles amplitudes et bruités.

|  |
| --- |
| ../../test/cd_phase_d.jpg |
| Figure 6 - Comparaison de deux méthodes (Phase deviation & Complex domain) |

Comme expliqué dans la partie « Peak detection algorithm », nous avons décidé de combiner les deux méthodes SpectralFlux et Complexe domain. Le résultat est le suivant :

`

|  |
| --- |
| ../../test/combinaison_methodes.jpg |
| Figure 7 - Combinaison de méthodes (Spectral flux & Complex domain) |

Notre fonction d’onset detection se présente dans sa forme finale comme ceci :

|  |
| --- |
| ../../test/final.jpg |
| Figure 8 - Algorithme final d'onset detection |

La courbe jaune est une moyenne locale, les pics rouges sont les onsets, les violets des offsets.

## Résultats :

|  |
| --- |
|  |
| Figure 9 - Evolution des résultats de la détection d'onset |

|  |
| --- |
| ../../Capture%20d’écran%202016-02-22%20à%2010.29.10.png |
| Figure 10 - Comparaison des pourcentages de détection d'onset par méthodes |

Après avoir essayé de nombreuses méthodes, la combinaison du SpectralFlux et de Complex domain nous semble la plus appropriée pour notre application. Nous arrivons à un taux de détection d’onset de l’orde de 97%.

**Principaux problèmes rencontrés :**

* Sur-détections
* Non-détections
* Lissage des fonctions d’onset
* Moyenne locale
* Onset trop rapprochés dans le temps

**Bibliographie :**

Méthodes principales :

**(1) - On the Use of Phase and Energy for Musical Onset Detection in the Complex Domain -**  Juan P. Bello, Chris Duxbury, Mike Davies, and Mark Sandler, Senior Member, IEEE - *IEEE SIGNAL PROCESSING LETTERS, VOL. 11, NO. 6, JUNE 2004*

**(2) - Simple Spectrum-Based Onset Detection** - Simon Dixon - Austrian Research Institute for Artificial Intelligence Freyung 6/6, Vienna 1010, Austria

**(3) - A Tutorial on Onset Detection in Music Signals** - Juan Pablo Bello, Laurent Daudet, Samer Abdallah, Chris Duxbury, Mike Davies, and Mark B. Sandler, Senior Member, IEEE *- IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING, VOL. 13, NO. 5, SEPTEMBER 2005*

**(4) - Onset detection revisited** - Simon Dixon - Austrian Research Institute for Artificial Intelligence Freyung 6/6, Vienna 1010, Austria - *Proc. of the 9th Int. Conference on Digital Audio Effects (DAFx-06), Montreal, Canada, September 18-20, 2006*

**(5) - Complex domain onset detection for musical signals -** Chris Duxbury, Juan Pablo Bello, Mike Davies and Mark Sandler *-* Department of Electronic Engineering Queen Mary, University of London Mile End Road, London E1 4NS, UK - *Proc. of the 6th Int. Conference on Digital Audio Effects (DAFx-03), London, UK, September 8-11, 2003*

*Pas dans la doc d’OD*

**(6) - Audio onset detection using machine learning techniques: the effect and applicability of key and tempo information -** CHING-HUA CHUAN **-** Department of Computer Science,University of Southern California Viterbi School of Engineering,Los Angeles, California, USA

**(7) - Auditory spectrum-based pitched instrument onset detection -** Emmanouil Benetos, *Student Member, IEEE* and Yannis Stylianou, *Member, IEEE*

**(8) - Wavelet-analysis-for-onset-detection** - Crawford Tait, William Findlay, Department of computing science, University of Glasgow, G12 8QQ

**(9) – Onset (audio)** - Article sur l’onset detection - <https://en.wikipedia.org/wiki/Onset_%28audio%29>

Autres articles intéressants :

**Influence of peak selection methods on onset detection** -

Carlos Ros , Ricardo Ribeiro, David Martins de Matos - ISCTE-IUL L2F/INESC-ID Lisboa

**Sound onset detection by applying psychoacoustic knowledge** - Anssi Klapuri - Signal Processing Laboratory, Tampere University of Technology P.O.Box 553, FIN-33101 Tampere, FINLAND

**Methodology and tools for the evaluation of automatic onset detection algorithms in music -** Pierre LEVEAU, Laurent DAUDET, Gaël RICHARD, Laboratoire d’Acoustique Musicale - 11, rue de Lourmel

75015 Paris - FRANCE