Onset detection

Méthode de spectral flux et déviation des phase → méthodes sensibles à la distorsion de phase (déviation des phase) et aux variations liées à la phase et à l'amplitude des signaux de bruits.

La méthode consiste à observer les changements en fréquences et en amplitude du signal audio sachant que pour des niveaux stables à l'échelle locale, un signal audio possède une fréquence et une amplitude constante.

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, en abscisse.

Rappel sur la méthode de spectral flux :

Cette méthode détermine la variation de l'amplitude en puissance du signal pour chaque échantillon de fréquence. La valeur absolue de la transformée de Fourier et dérivée pour obtenir la variation pour chaque fréquence entre l'échantillon n et n-1. On somme ensuite pour chaque échantillon de temps ces variations pour pouvoir avoir une représentation temporelle de ces variations.

Rappel sur la méthode de déviation de 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 $-\pi$ et π.

La première différence de phase entre un échantillon n et n-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.

La valeur moyenne de la valeur absolue des changements (valeur absolue de cette dérivée seconde) nous donne une représentation temporelle de possibles onset.

Complex Domain Onset Detection :

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.

La valeur cible est définie comme :

$$\hat{S}_k(m) = \hat{R}_k(m)e^{j\hat{\phi}_k(m)}$$

= (magnitude de l'échantillon précédent) *(somme entre la phase précédente et la différence de phase entre cette phase précédente et celle encore avant).

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

$$S_k(m) = R_k(m)e^{\phi_k(m)}$$

En mesurant la distance euclidienne entre la cible visée et le vecteur courant dans le domaine complexe, on est capable de quantifier la stationnarité pour le $k^{ème}$ échantillon.

La déviation de phase est donnée par la formule suivante :

$$\tilde{\varphi}_k(n-1) - \tilde{\varphi}_k(n-2) = \tilde{\varphi}_k(n) - \tilde{\varphi}_k(n-1)$$

L'égalité de droite correspond à la variation de la fréquence instantanée de l'échantillon n et l'égalité de gauche correspond à la variation de la fréquence instantanée de l'échantillon n-1.

La déviation de phase courante est alors donnée par :

$$d_\varphi = \text{princarg}[\tilde{\varphi}_k(n) - 2\tilde{\varphi}_k(n-1) + \tilde{\varphi}_k(n-2)]$$

La phase de la valeur ciblée est donc :

$$\hat{\phi}_k(m) = \text{princarg}[2\tilde{\varphi}_k(m-1) - \tilde{\varphi}_k(m-2)]$$

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 :

$$\begin{aligned}\Gamma_k(m) \quad &= \quad \{\hat{R}_k(m)^2 + R_k(m)^2 - \cdots \\ &\qquad 2\hat{R}_k(m)R_k(m)cos(d_{\varphi k}(m))\}^{\frac{1}{2}}\end{aligned}$$