

TP de Traitement du Signal Musical

Vocodeur de phase et étirement temporel

Philippe Depalle et Bertrand David

Vendredi 2 décembre 2016

1 TFCT

L'objectif de cette première partie est de mettre en oeuvre la Transformée de Fourier à Court Terme (TFCT) dans un système d'analyse/synthèse.

1.1 Généralités

On utilisera le programme `tfct.m` qui fournit un cadre pour le calcul de la transformée de Fourier à Court Terme, dont la définition dite "en convention passe-bas" est donnée en temps discret par :

$$\tilde{X}_0(b, f) = \sum_{n \in \mathbb{Z}} x(n)w(n-b)e^{-j2\pi fn}, \quad (1)$$

où $w(n)$ est une fenêtre d'analyse en temps discret, supposée sommable et réelle.

(a) Remarquer que l'expression 1, prise à f fixée, peut s'écrire comme une convolution et en déduire une interprétation de la TFCT en terme de filtrage. Expliciter le rôle de la fenêtre en terme de type et de longueur.

(b) Une autre définition de la TFCT, dite "en convention passe-bande" est donnée par :

$$\tilde{X}_{\text{loc}}(b, f) = \sum_{n \in \mathbb{Z}} x(n+b)w(n)e^{-j2\pi fn}. \quad (2)$$

Expliquer cette dernière appellation et calculer \tilde{X}_{loc} en fonction de \tilde{X}_0 . Laquelle des deux conventions correspond au calcul mené dans `tfct.m` ?

Dans toute la suite, le calcul de la TFCT sera effectué par TF rapide (FFT) en considérant uniquement des fenêtres causales de longueur finie M inférieure à l'ordre N de la TFD. Pour simplifier la TFCT sera maintenant indiquée par le numéro k du canal fréquentiel situé autour de $f_k = k/N$ et l'indice temporel p de la trame considérée, soit :

$$\tilde{X}_{\text{loc}}[p, k] = \sum_{n \in \mathbb{Z}} x(n+pR)w(n)e^{-j2\pi \frac{kn}{N}}. \quad (3)$$

où R représente le décalage temporel entre chaque trame d'analyse.

(c) A l'aide de `tfct.m`, pour $N = 64$ et $R = 1$, calculer le signal $x_k(p) = \tilde{X}_{\text{loc}}[p, k]$, pour $k = 12$. Est-il réel ou complexe ? Vérifier l'interprétation en terme de filtrage en reformulant l'implémentation de manière à utiliser la fonction `filter` de MATLAB. Écouter $\text{Re}(x_k)$ et commenter. À l'issue de cette question on revient à des valeurs de N et R plus typiques pour le traitement audio, telles que $N = 1024$ et $R = M/4$.

1.2 Reconstruction

La reconstruction s'effectue par une opération *d'addition-recouvrement*, qui s'écrit sous la forme :

$$y(n) = \sum_{p \in \mathbb{Z}} y_s(p, n - pR),$$

où $y_s(p, n) = \text{TFD}^{-1}[\tilde{X}_{\text{loc}}[p, k]](n) w_s(n) = \frac{1}{N} \sum_{k=0}^{N-1} \tilde{X}_{\text{loc}}[p, k] e^{j2\pi \frac{kn}{N}} \times w_s(n)$; avec $w_s(n)$ une fenêtre de synthèse, sommable et réelle.

(d) Montrer qu'une condition suffisante de reconstruction parfaite est $\chi(n) = 1 \forall n$ où $\chi(n) = \sum_{p \in \mathbb{Z}} w(n - pR) w_s(n - pR)$. Utiliser le programme `ola.m` pour vérifier cette condition lorsque la fenêtre produit $w_\pi(n) = w(n) w_s(n) = h^2(n)$ où $h(n)$ est une fenêtre de Hann et pour un recouvrement de 75%.

(e) Intégrer une partie de resynthèse effectuant l'addition-recouvrement décrite au (e) dans le programme `tfct.m` et qui vérifie la reconstruction parfaite.

(f) On cherche maintenant une condition approchée de reconstruction qui utilise une fenêtre de Blackman-Harris d'ordre 4. Ecrire la condition sur I qui assure le non-repliement en admettant que la transformée de Fourier de la fenêtre est négligeable en dehors du lobe principal. En déduire le recouvrement minimal, en %, nécessaire pour assurer cette condition. Mettre en oeuvre un schéma d'analyse-synthèse fondé sur cette condition et afficher l'erreur de reconstruction pour un signal donné. Que vaut le rapport signal/erreur en dB ?

1.3 Égaliseur à TFCT

On se sert des résultats précédents pour utiliser la TFCT comme un égaliseur à $N/2 + 1$ canaux réels (on suppose l'ordre N pair). On obtiendra alors les trames de synthèse comme :

$$y_s(p, n) = \text{TFD}^{-1}[\tilde{Y}[p, k]](n) w_s(n),$$

où \tilde{Y} est obtenu en affectant des poids à chacun des canaux soit $\tilde{Y}[p, k] = \gamma_k \tilde{X}_{\text{loc}}[p, k]$. La condition de reconstruction assure alors $y(n) = x(n)$ lorsque $\gamma_k = 1 \forall k$.

(g) Réaliser l'égaliseur et le tester.

2 Étirement temporel

Dans cette partie, on désire mettre en oeuvre l'étirement temporel à l'aide du phase-vocodeur. Les étapes supplémentaires à introduire dans le programme obtenu au paragraphe 1.2 sont :

- déroulement des phases d'analyse et estimation des fréquences instantanées dans chaque canal,
- calcul de l'incrément de phase en synthèse,
- construction du nouveau \tilde{Y} à partir des phases de synthèse et des amplitudes, et synthèse du signal de sortie.