

TP4 -2013. L'analyse de Fourier 1D - Séparation de sources

Ce texte est basé sur l'analyse spectrale fenêtrée, c'est-à-dire l'analyse des signaux 1D à la fois en temps (lecture horizontale) et en fréquence (lecture verticale).

- (1) Télécharger les deux routines `analyse` et `synthese` dans le répertoire `TP2013-routines/TP4` et étudiez-en la syntaxe.
- (2) Chargez les deux signaux audio `Mix11.wav` et `Mix21.wav` depuis le répertoire `Audio` et écoutez les : vous y reconnaitrez (sur quelques notes d'un enregistrement musical) la superposition de deux instruments : synthétiseur et saxo. Avec la routine `analyse` et une fenêtre d'exploration (ici de Hamming pour gommer le phénomène de Gibbs lors de la troncature au travers de chaque position de cette fenêtre glissante), calculer les spectrogrammes des deux signaux `Mix11` et `Mix21`. Affichez les modules `f1` et `f2` de ces spectrogrammes en ne conservant que les 100 premiers canaux de fréquences (sur les 2048). Ces canaux de fréquences correspondent-ils aux basses fréquences ? aux hautes fréquences ? À quelle gamme de fréquences (en Hertz) correspondent ces 100 premiers canaux ?
- (3) Ces deux signaux `Mix11` et `Mix21` ont été obtenus à partir de deux signaux source `S1` et `S2` qui ont été combinés suivant respectivement

$$\text{Mix11} = \text{S1}/3 + 2*\text{S2}/3 \quad \text{Mix21} = 2*\text{S1}/3 + \text{S2}/3$$

Cela se voit-il à l'examen rapide des (modules des) spectrogrammes de `Mix11` et `Mix21` ? Écrivez un programme `separation1`

```
[f1bis, f2bis] = separation1(f1,f2,T);
```

qui renvoie, étant donnés deux (modules de) spectrogrammes `f1` et `f2`, deux nouveaux (modules de) spectrogrammes `f1bis`, `f2bis` tels que `f2bis = f1` aux pixels où `f1` domine `T*f2` (et `f2bis=0` sinon) et `f1bis = f2` aux autres points (`f1bis=0` ailleurs). Pour choisir `T` judicieusement, vous pouvez vous aider de l'information *a priori* que vous avez sur les contenus des signaux composés `Mix11` et `Mix21`. Plus précisément, vous pouvez calculer les deux rapports $(1/3)/(2/3) = 1/2$ (contributions comparées de `S1` dans les deux voix) et $(2/3)/(1/3) = 2$ (contributions comparées de `S2` dans les deux voix), ordonner ensuite ces deux rapports $1/2 < 2$ dans l'ordre croissant et prendre pour seuil `T` la moyenne géométrique de ces deux nombres. Pourquoi la moyenne géométrique plutôt que la moyenne arithmétique ? Comment, à partir des deux (modules de) spectrogrammes `f1bis` et `f2bis`, pourrait-on reconstituer les deux sources `S1` et `S2` ? Vous préciserez en particulier quelle phase (celle de `Mix11` ou celle de `Mix21`) il

convient de combiner avec chacun des deux (modules de) spectrogrammes `f1bis` et `f2bis`. Utilisez pour la reconstitution la routine `synthese`. Pour « écouter » les deux signaux `S1` et `S2` ainsi repomposées, utilisez la routine :

```
>> wavwrite(S1,Fs,B,'S1.wav');
>> wavwrite(S2,Fs,B,'S2.wav');
```

(la fréquence en Hertz `Fs` et le nombre de bits d'encodage `B` étant ceux des signaux `Mix11.wav` et `Mix21.wav`). Vous devriez constater que vous avez bien séparé les deux sources (ici synthétiseur et saxo). Cette démarche est à la base de la méthode de *séparation de sources* (ici en l'occurrence, les deux sources composant les signaux `Mix11` et `Mix21`).

- (4) Travaillez maintenant avec les deux signaux `Mix12` et `Mix21`. Ces deux signaux ont été, eux, obtenus à partir de trois signaux source `SS1`, `SS2`, `SS3`, qui ont été cette fois ci combinés suivant respectivement

$$\text{Mix21} = .4*S1 + .2*S2 + .8*S3 \quad \text{Mix22} = .6*S1 + .8*S2 + .2*S3$$

Cela se voit-il encore à l'examen rapide des (modules des) spectrogrammes de `Mix12` et `Mix22`? On souhaiterait rédiger un programme `separation2`

```
[f1bis, f2bis, f3bis, f4bis] = separation2 (f1,f2,T1,T2);
```

qui renvoie, étant donnés deux (modules de) spectrogrammes `f1` et `f2`, quatre nouveaux (modules de) spectrogrammes `f1bis`, `f2bis`, `f3bis`, `f4bis` en utilisant cette fois deux seuils de comparaison `T1` et `T2` (avec `T1 < T2`) au lieu d'un seul (`T`, comme à la question précédente).

- `f4bis = f1` aux points où `f1` domine `T2*f2` (et soit nul aux autres points);
- `f1bis = f2` aux points où `f1` est dominé par `T1*f2` (et zéros ailleurs);
- `f2bis = f2` aux points où `f1` est encadré par `T1*f2` et `T2*f2` (et zéro ailleurs);
- `f3bis = f1` aux points où `f1` est encadré par `T1*f2` et `T2*f2` (et zéro ailleurs).

Pour choisir `T1` et `T2` judicieusement (et donc concevoir ce programme en ayant en tête que l'objectif est la séparation des trois sources `SS1`, `SS2` et `SS3`), vous pouvez vous aider de l'information *a priori* que vous avez sur les contenus des signaux composés `Mix12` et `Mix22`. Plus précisément, vous pouvez calculer les trois rapports $.4/.6 = 2/3$ (contributions comparées de `S1` dans les deux voix), $.2/.8 = 1/4$ (contributions comparées de `S2` dans les deux voix), $.8/.2 = 4$ (contributions comparées de `S3` dans les deux voix). Rangez ensuite ces rapports dans l'ordre croissant $1/4 < 2/3 < 4$. Comment choisir les deux seuils `T1` et `T2` (inspirez vous de la question précédente)? Concevez alors le programme `separation2` demandé. Comment, à partir des quatre spectrogrammes `f1bis`, `f2bis`, `f3bis`, `f4bis`, pourrait-on reconstituer les trois sources `SS1`, `SS2` et `SS3`? Vous préciserez en particulier quelle phase (celle de `Mix12` ou celle de `Mix22`) il convient de combiner avec chacun des quatre (modules de) spectrogrammes `f1bis`, `f2bis`, `f3bis` et `f4bis`. Pourquoi vous apparaît-il plus raisonnable de conserver deux voix pour le signal `SS1`? Utilisez pour la reconstitution la

routine `synthese`. Pour « écouter » les trois signaux `SS1`, `SS2`, `SS3` ainsi recomposés, utilisez la routine :

```
>> wavwrite(SS1-voix1,Fs,B,'SS1-voix1.wav');
>> wavwrite(SS1-voix2,Fs,B,'S1.voix2.wav');
>> wavwrite (SS2,Fs,B,'SS2.wav');
>> wavwrite (SS3,Fs,B,'SS3.wav');
```

(la fréquence en Hertz `Fs` et le nombre de bits d'encodage `B` étant ceux des signaux `Mix11.wav` et `Mix21.wav`). Vous devriez constater que vous avez bien « séparé » les trois sources (ici synthétiseur, clavier, saxo). Ce que vous venez de faire ici constitue une première approche au principe de la séparation de trois sources (ici les trois sources `SS1`, `SS2`, `SS3` composant les deux signaux « mixtes » que sont `Mix12` et `Mix22`). Bien sûr, les choses ici se trouvent amplement facilitées de par le fait que les pondérations des trois sources dans les deux voix sont dans ce cas connues *a priori*.

(5) La routine

```
[Y,I] = sort(V,2,'descend');
```

renvoie, étant donné un vecteur ligne `V` à entrées, le vecteur ligne `Y` correspondant aux entrées de `V` rangées dans l'ordre des modules décroissants ainsi que les indices de ces entrées, soit :

```
for k=1:length(V)
Y(k) = V(I(k))
end
```

Réalisez une routine

```
>> T = Maxima (T1,T2,K);
```

qui, étant donnés deux tableaux `T1` et `T2` de même taille et d'entrées toutes positives et un entier strictement positif `K` donné,

- d'une part calcule un tableau `T` à `K` lignes et 2 colonnes dont les entrées `T(k,1)` et `T(k,2)`, $k = 1, \dots, K$ sont obtenues ainsi : on forme le tableau $\mathbf{Taux} := \sqrt{T1^2 + T2^2}$, puis on détermine dans ce tableau les `K` pixels (i, j) donnant les valeurs les plus énergétiques $\mathbf{Taux}(i, j)$; pour chacun de ces `K` pixels (numérotés de $k=1$ à `K` dans l'ordre décroissant des énergies de \mathbf{Taux} en ces pixels), on prend comme coordonnées du vecteur $\mathbf{T}(k, :)$ les valeurs en $(i(k), j(k))$ prises respectivement par `T1` et `T2` ;
- d'autre part affiche les `K` vecteurs ligne de `T` sous la forme d'un « nuage » de points dans le plan.

Testez cette routine sur les deux spectrogrammes `f1` et `f2` utilisés aux questions **3**, puis **4**. Voit-on apparaître l'organisation du nuage de points suivant des droites du plan ? Comment interprétez vous ces droites ? Comment sont liés les coefficients du mélange et les pentes de ces droites ?

(6) Écrire une routine qui à partir d'un tableau obtenu par

```
>> T = Maxima (f1,f2,K)
```

réalisé à partir des modules des spectrogrammes `f1` et `f2` de deux signaux correspondant à des enregistrements où se mêlent `m` sources instrumentales superposées, renvoie

```
>> L = coefficients(T,m);
```

où L préfigure la liste de ce que pourraient être les m rapports des contributions des diverses sources instrumentales S_1, \dots, S_m dans le « mélange » qui a conduit aux deux voix f_1 et f_2 . Pour cela, vous calculerez les K rapports d'intensité des coordonnées des divers vecteurs lignes de T , puis vous en formerez l'histogramme (routine `hist` de `MATLAB`), avant que de chercher les m pics les plus significatifs de cet histogramme.