

# Analyse de Fourier appliquée à la séparation de sources

Ch. Dossal

Avril 2013

## 1 Introduction

Le but de ce TD est de séparer différentes sources à partir d'un enregistrement stéréophonique en exploitant les différences d'intensités des différentes sources entre les deux voix. Cette différence entre les voix peut être due par exemple aux différentes positions des instruments au moment de l'enregistrement si l'on suppose que chacune des voix est acquise par deux micros placés à des endroits différents.

Nous travaillerons sur des données *mélangées* artificiellement sans ajouts d'effets comme la réverbération. On parle de mélange instantané. Nous allons utiliser le spectrogramme vu lors d'un précédent TD.

On dispose de données stéréophoniques qui seront représentées sous matlab par une matrice  $S$  comportant deux lignes et 220500 colonnes (5 secondes de son échantillonné à 44100 Hz). On considère que chacune des voix est une combinaison linéaire de pistes monoinstrumentales avec des coefficients différents. On peut ainsi voir  $S$  comme le produit d'une matrice de mélange  $M$  de taille  $2 \times n$  par une  $P$  de taille  $n \times 22050$  contenant les  $n$  pistes instrumentales. Si  $n = 2$  et qu'on connaît la matrice  $M$  il est aisé d'effectuer la séparation juste par inversion de la matrice  $M$ . Notre but ici est de développer une méthode fonctionnant sans connaissance de  $M$  et avec plus de deux sources. Néanmoins pour commencer nous supposerons que  $M$  est connue et nous utiliserons deux sources avant de passer à trois.

Pour ce faire nous supposerons qu'à chaque instant les différents instruments jouent à des fréquences différentes. Pour chaque fréquence instantanée, nous allons comparer les deux voies et en fonction du rapport d'intensité entre les deux voies, attribuer la fréquence instantanée à l'un ou l'autre des instruments.

## 2 Séparation avec Matrice de mélange connue

### 2.1 Deux instruments

1. A l'aide de la commande *wavread* charger les deux sons *Mix11.wav* et *Mix21.wav* dans deux vecteurs *Smix11* et *Smix21*.  
Ces deux mélanges ont été obtenus à partir de deux sources *S1* et *S2* en faisant :

$$Smix11 = S1/3 + S2 * 2/3 \text{ et } Smix21 = S1 * 2/3 + S2/3.$$

2. Calculer et visualiser les spectrogrammes ainsi que les tableaux de phases de *Smix11* et *Smix21* à l'aide des programmes effectués lors d'un précédent TD. On prendra un recouvrement des fenêtres de 2000 points.
3. Ecrire un programme *Separation*

```
function [Spec3,Spec4,Spec5,Spec6]=Separation[Spec1,Spec2,T]
```

qui prend en entrée les deux spectrogrammes *Spec1* et *Spec2* et qui renvoie quatre spectrogrammes (modules de transformée de Fourier à fenêtres) *Spec3*, *Spec4*, *Spec5* et *Spec6* tels que

- $Spec3(i, j) = Spec1(i, j)$ ,  $Spec4(i, j) = Spec2(i, j)$ ,  
 $Spec5(i, j) = 0, Spec6(i, j) = 0$ , si  $Spec1(i, j) > T * Spec2(i, j)$ ,
- $Spec3(i, j) = 0$ ,  $Spec4(i, j) = 0, Spec5(i, j) = Spec1(i, j)$ ,  
 $Spec6(i, j) = Spec2(i, j)$  sinon

On pourra se contenter de traiter la moitié basse des *basses fréquences* si on veut limiter le temps de calcul.

4. Quelle valeur de *T* vous semble pertinente pour ce mélange particulier ?
5. Ecrire une fonction qui en fonction des paramètres du mélange vous donne une valeur raisonnable de *T*.
6. Ecrire un programme qui reconstruit les signaux séparés à partir de *Spec3*, *Spec4*, *Spec5*, *Spec6* et des tableaux de phases. Tester les deux tableaux de phases. Que pensez vous ?
7. Comme on connaît les coefficients utilisés dans le mélange, peut-on réaliser une séparation parfaite dans ce cas ? Comment ?

### 2.2 Trois instruments

8. Charger les mélanges *Mix12.wav* et *Mix22.wav*. On a ici  
 $Smix12 = 0.4 \times S1 + 0.2 \times S2 + 0.8 \times S3$  et  $Smix22 = 0.6 \times S1 + 0.8 \times S2 + 0.2 \times S3$
9. Calculer et visualiser les tableaux de phases de *Smix12* et *Smix22*.

10. Ecrire un programme *Separation2*

```
function [Spec3,Spec4,Spec5,Spec6,Spec7,Spec8]=Separation2[Spec1,Spec2,T]
```

qui renvoie 6 spectrogrammes et où  $T$  est un vecteur à 2 composantes qui va séparer en 6 les deux spectrogrammes en attribuant les fréquences instantanées en fonctions des rapports des fréquences instantanées de Spec1 et Spec2.

11. Ecrire un programme qui estime des valeurs raisonnables de  $T(1)$  et  $T(2)$  en fonction des paramètres du mélange.
12. Reconstruire les trois sons à partir de des spectrogrammes ainsi obtenus et des tableaux de phases.

### 3 Estimation de la matrice de mélange

Suposons maintenant que la matrice de mélange soit inconnue. Comment faire ?

13. Ecrire un programme

```
function S=vectoriser(Spec1,Spec2,N)
```

qui prend en entrée deux spectrogrammes et un entier  $N$  qui renvoie un vecteur  $S$  de taille  $2 \times N$  contenant les paires d'intensités les plus énergétiques des spectrogrammes (on pourra utiliser la commande *sort*). Ce programme doit aussi afficher sous forme de nuage de points ces différentes paires.

On peut espérer que ces points seront principalement alignés le long de  $n$  droites où  $n$  est le nombre de sources.

14. Tester le programme précédent sur les deux jeux de mélanges utilisés.
15. Comment interpréter ces *droites* ? A quoi correspond un point sur une des droites ? A quoi correspond un point en dehors de ces droites ? Comment sont liés les coefficients directeurs de ces droites et les coefficients du mélange ?
16. Ecrire un programme

```
V=Direction(S,nbinstru)
```

qui à partir de  $S$  et d'un nombre  $nbinstru$  d'instruments renvoie une estimation des coefficients de mélange. On pourra calculer les  $N$  rapports d'intensités et utiliser les pics de l'histogramme de ces rapports (commande *hist* de matlab).

Plus précisément on pourra calculer l'histogramme de  $U = \frac{R}{1+R}$  où  $R$  est le rapport d'intensité entre les deux voix. En effet les pics de l'histogramme de  $U$  sont plus visibles que ceux de  $R$ .

17. Ecrire un programme qui à partir de deux mélanges et d'un nombre d'instruments, reconstruit les fichiers wav des différentes sources.