

Compte Rendu TP Séparation de Sources

Yichen ZHU

February 6, 2018

1 Généralité

L'objectif du TP est d'implémenter la méthode DUET pour la séparation des sources. L'hypothèse forte qu'on pose ici dans le problème c'est que le signal audio soit parcimonieux dans le domaine temps-fréquence.

2 Questions

1. On arrive identifier trois instruments à l'oreille : La basse, la guitare et le piano. On les distingue grâce à leurs différents timbres de son.
2. On trace le diagramme de dispersion dans le domaine temporel. On n'arrive pas à distinguer les directions pour les sources parce que plusieurs sources peuvent être activées en même temps.

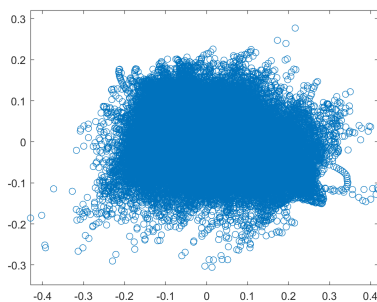


Figure 1: Représentation temporelle et représentation temps-fréquence

3. On trace la représentation temps-fréquence pour chaque channel de x . On peut remarquer que l'hypothèse forte qu'on a posée sur le signal audio est pertinente. Il y a beaucoup plus de coefficients zéros que les coefficients non nuls. Le signal audio est donc parcimonieux dans le domaine MDCT. On met les deux figures en même temps pour comparer, en gauche l'observation par les deux micros dans le domaine temporel et en droite dans le domaine temps-fréquence.

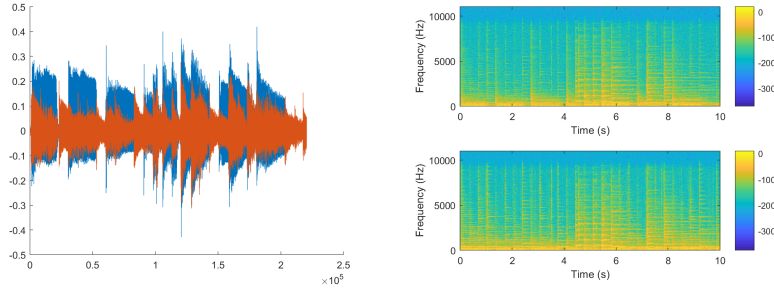


Figure 2: $x(t, m)$ et $X(f, n, 1), X(f, n, 2)$

4. On trace le diagramme de dispersion dans le domaine temps-fréquence et visuellement on a trois différentes directions. Cela s'explique avec notre hypothèse parcimonieuse, dans le domaine MDCT, on a beaucoup de coefficients nuls et on a trop peu de points qui se recouvrent lorsque les sources sont activées. Les angles entre les directions dépendent comment sont placés les micros.

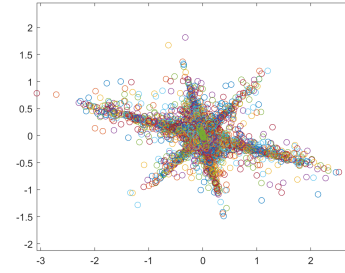


Figure 3: $Z(n, t)$ le diagramme de dispersion dans le domaine MDCT

5. On trace l'histogramme de l'argument de $Z(f, n)$ pour identifier la valeurs des trois angles pour les trois sources. On trouve facilement $\theta(1) = -1.325$, $\theta(2) = -0.275$ et $\theta(3) = 0.775$.

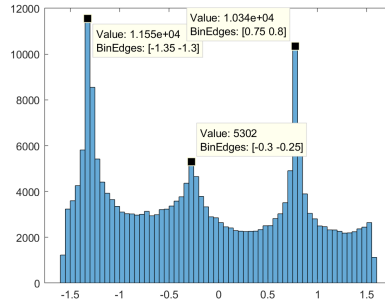


Figure 4: Histogramme des arguments de $Z(n, t)$

8. On écoute le signal reconstruit en stéréo. On a l'impression que certaines sources sont plus proches de l'oreille gauche et certaines plus proches de l'oreille droite.
9. On écoute ensuite les trois channels de s . On remarque que les sources peuvent être séparées mais sur chaque channel, on a l'impression qu'il y a deux sources activées (guitare + piano / basse + piano / guitare + basse).
10. En permutant les arguments, on a inversé le left channel et le right channel.