Examen de l'UE ATIAM/TSM
"Applications et ouvertures"

Geoffroy Peeters, Umut Simsekli, Roland Badeau

Contrôle sans document - 2h

1 Questions cours "Estimation multi-pitch"

Question 1
En quoi le cepstre réel peut-il être rapproché de l'auto-corrélation ? Démontrez-le mathématiquement.
Question 2
a) Qu'est-ce que la méthode de la somme spectrale pour estimer la fréquence fondamentale ?
b) Donnez-en l'expression mathématique.
c) Quelle est sa principale hypothèse/limitation ?
Question 3
Quelles sont les quatre opérations appliquées à un signal audio pour reproduire la décomposition opérée par l'oreille interne ?
Question 4
Expliquez le principe du lissage spectral dans la méthode d'estimation multi-pitch de A. Klapuri.

2 Questions cours "Estimation de structure"

Question 1
a) Démontrer mathématiquement que le cepstre réel permet de séparer la contribution du pitch et de l'enveloppe spectrale.
b) Comment s'appelle ce modèle de production du son ?
Question 2
Soit un signal audio contenant une note de musique à la hauteur C4 de série harmonique d'amplitudes $a_0 = 1$, $a_1 = 0.75$, $a_2 = 0.5$, $a_3 = 0.25$, $a_4 = 0.1$, donnez les valeurs de sa représentation en chroma.
Question 3
Quelles sont les deux hypothèses utilisées pour représenter la structure temporelle d'un morceau de musique ? Expliquez chacune d'elles.

3 Questions cours "NMF" : *regularized β-NMF*

In certain NMF applications, we are required to enforce sparsity constraints on the factor matrices W and H. This is often achieved by considering the following optimization problem:

$$(W^\star, H^\star) = \underset{W\geq 0, H\geq 0}{\arg\min} \Big[\sum_{f=1}^{F}\sum_{n=1}^{N} d_\beta(v_{fn}; \hat{v}_{fn}) + \lambda_W \sum_{f=1}^{F}\sum_{k=1}^{K} w_{fk} + \lambda_H \sum_{k=1}^{K}\sum_{n=1}^{N} h_{kn}\Big], \tag{1}$$

where $\lambda_W > 0$, $\lambda_H > 0$, $\hat{v}_{fn} = \sum_{k=1}^{K} w_{fk} h_{kn}$ and $d_\beta(\cdot;\cdot)$ is the β-divergece that is defined as follows:

$$d_\beta(v;\hat{v}) = \frac{1}{\beta(\beta-1)}\Big(v^\beta + (\beta-1)\hat{v}^\beta - \beta v \hat{v}^{\beta-1}\Big). \tag{2}$$

Question 1
What are the roles of $\lambda_W > 0$ and $\lambda_H > 0$ in this problem?
Question 2
Derive the multiplicative update rules for this particular problem.
Question 3
Let us assume that we obtain the optimal factors, $W^\star$ and $H^\star$. How can we use $W^\star$ and $H^\star$ for audio source separation?

4 Questions cours "Méthodes à haute résolution"

On rappelle que la méthode MUSIC consiste à diagonaliser la matrice de covariance $\boldsymbol{R}_{xx}$ du signal, et à déterminer les pôles $\{z_k\}_{k\in\{0\ldots K-1\}}$ en tant que solutions de l'équation

$$\|\boldsymbol{W}_\perp^H \boldsymbol{v}(z)\|^2 = 0 \tag{3}$$

où $\boldsymbol{v}(z) = \big[1,\, z, \ldots,\, z^{n-1}\big]^T$, et la matrice $\boldsymbol{W}_\perp$, de dimension $n \times (n-K)$, contient les vecteurs propres de $\boldsymbol{R}_{xx}$ associés aux $n-K$ plus petites valeurs propres (engendrant ainsi l'espace bruit).
Question 1
On suppose que tous les pôles du signal sont sur le cercle unité. Vérifier que l'équation (3) implique qu'ils sont également solutions de l'équation $P(z) = 0$, où

$$P(z) = z^{(n-1)}\, \boldsymbol{v}(1/z^*)^H \left(\boldsymbol{W}_\perp \boldsymbol{W}_\perp^H\right) \boldsymbol{v}(z).$$

Question 2
On définit la matrice $\boldsymbol{P} = \boldsymbol{W}_\perp \boldsymbol{W}_\perp^H$, et on remarquera que

$$P(z) = [z^{n-1},\, z^{n-2}, \ldots,\, z,\, 1]\, \boldsymbol{P} \begin{bmatrix} 1 \\ z \\ \vdots \\ z^{n-1} \end{bmatrix}.$$

Vérifier que la matrice $\boldsymbol{P}$ est à symétrie hermitienne et positive, et démontrer que $P(z)$ est un polynôme de degré $2(n-1)$, dont les racines de module $\neq 1$ peuvent être regroupées par paires (si z est racine, $1/z^*$ l'est aussi).
Question 3
L'algorithme *root*-MUSIC consiste à calculer les $2(n-1)$ racines de $P(z)$. D'après vous, comment pourrait-on en déduire les valeurs des pôles z_k ? (on prêtera attention au fait que $K < 2(n-1)$).
Question 4
La méthode *spectral-MUSIC* vue en cours et testée en TP consiste à chercher les K maxima du pseudo-spectre $P(e^{i2\pi f}) = \frac{1}{\|\boldsymbol{W}_\perp^H \boldsymbol{v}(e^{i2\pi f})\|^2}$. Vérifier que les valeurs du pseudo-spectre pour les fréquences $f_k = \frac{k}{N_{\text{fft}}}$, où $k \in \{0, N_{\text{fft}} - 1\}$ et $N_{\text{fft}} \geq 2n-1$, peuvent être obtenues à l'aide de la TFD de longueur N_{fft} du signal constitué des coefficients du polynôme P, complétés par des zéros. Quel est l'inconvénient de cette approche par rapport à *root-MUSIC*?