

1 Variantes autour de l'exemple du cours

Il s'agit ici de programmer la NMF telle que décrite dans l'article de Lee & Seung fourni. Pour cela, on s'appuiera sur le script à compléter `test_nmf.m` qui gère la partie affichage.

1.1 Construction de l'exemple du cours

Cet exemple (DO-MI-DO/MI) est obtenu en construisant la matrice \mathbf{W} de la base spectrale et la matrice \mathbf{H} des activations de manière synthétique. Aucun son n'est donc calculé ici. Pour cela on pourra par exemple :

- utiliser la fréquence d'échantillonnage $F_e = 8000$ Hz, l'ordre de TFD : $N_{\text{fft}} = 512$.
- calculer les fréquences fondamentales F_0 du DO et du MI placés sous le LA 440.
- pour chaque note r , fabriquer un spectre harmonique (non négatif) de la forme $W_r(f) = \sum_{k=1}^K a_k w(f - kF_0)$ où w est une fenêtre usuelle (Hann de largeur 80 Hz par exemple) et a_k décrit l'amplitude des raies (par exemple décroissante, de la forme $a_k = \exp(-k/K)$). K est fixé à l'aide du critère de Shannon. En déduire la matrice spectrale \mathbf{W}_s .
- pour les activations, on utilisera une fonction du type $h_r(t) = \sum_p b_p h(t - p\Delta T)$ où $h(t) = e^{-t/\tau}$ pour $t \in [0 \Delta T]$ avec ΔT de l'ordre de 0.5 secondes et $h(t) = 0$ ailleurs. b_p vaut 0 ou 1. En déduire la matrice \mathbf{H}_s . Attention, pour se mettre dans des conditions réalistes, on supposera que la représentation est obtenue avec des trames de longueur $N = N_{\text{fft}}$ et un recouvrement de 75%. Soit un taux d'échantillonnage de $F_e/(N_{\text{fft}}/4)$ pour les activations temporelles. τ sera pris de l'ordre de $\Delta T/3$.
- construire alors la matrice temps-fréquence $\mathbf{X}_s = \mathbf{W}_s \mathbf{H}_s$

1.2 NMF classique

- Rappeler le principe des mises à jour multiplicatives pour faire décroître la fonction de coût $\mathcal{C}(\boldsymbol{\theta})$ où $\boldsymbol{\theta}$ décrit l'ensemble des paramètres du modèle (les coefficients de \mathbf{W} et \mathbf{H}). Comparer les règles de mise à jour données dans l'article avec une fonction de coût du type $\mathcal{C} = D_\beta(\mathbf{X}_s || \mathbf{W}\mathbf{H})$ où D_β est une β -divergence.
- Utiliser ces règles de mise à jour pour factoriser en produit de matrices non-négatives la matrice \mathbf{X}_s . Les matrices \mathbf{W} et \mathbf{H} seront initialisées avec des valeurs aléatoires (en respectant toutefois les ordres de grandeur et la normalisation).
- Tracer le graphe de \mathcal{C} en fonction de l'itération. Que constatez vous ?

1.3 Variantes

Testez votre algorithme en faisant varier les données et les paramètres :

- en répétant avec plusieurs tirages d'initialisations (quelles différences constatez-vous sur les résultats?)
- avec ajout de bruit (attention à respecter la non-négativité, utiliser `rand`)
- avec différentes valeurs de β (0, 1 et 2 typiquement)
- avec plus de notes et des indéterminations du type DO/MI, DO/SOL, DO/MI/SOL
- en fabriquant des notes synthétiques (formes d'onde) par somme de sinusoides et en calculant la TFCT du signal temporel obtenu.

2 Transcription automatique à l'aide de la NMF semi-supervisée.

Les notes isolées d'un piano vous sont fournies ainsi qu'un morceau réalisé à l'aide du même instrument. La NMF semi-supervisée consiste à définir la matrice spectrale \mathbf{W} à l'aide des notes isolées (fournies sous la forme de `xxx.wav` où `xxx` est le numéro midi, 21-108 pour la tessiture du piano), en calculant simplement une estimation de la densité spectrale de puissance associée. La seule mise à jour restante est donc celle de \mathbf{H} . On travaillera sur des signaux sous-échantillonnés à 22 kHz.

Le script fourni `build_signal.m` permet, à l'aide d'une syntaxe simple, de fabriquer un morceau à partir de notes définies par leur numéro midi, leur temps d'attaque, leur durée et la vélocité associée.

1. constituer la base \mathbf{W} correspondant aux 88 notes du piano
2. fabriquer un signal test correspondant à la gamme C3-C4.
3. programmer la NMF semi-supervisée à l'aide des résultats de la partie précédente et l'appliquer au signal test. Observer les activations temporelles obtenues.
4. En vous rappelant le TP sur le rythme, fabriquer une fonction de détection des attaques à appliquer sur les activations pour connaître les instants où les notes sont jouées.