
ASSM : Spatialisation et localisation de sources sonores

Objectifs : implémentation d'un système complet de spatialisation et de localisation de sources sonores. Le travail demandé pourra être réalisé au choix, avec Matlab¹ ou Octave² (alternative gratuite à Matlab) disponible sur les machines du CREMI.

1 Traitement du son avec Matlab

1.1 manipulation de sources sonores

Exercice 1 : Récupérez l'archive http://www.labri.fr/~fourer/Ens/1112/assm/td_spatial.tar.gz. Chargez simultanément tous les vecteurs s_i (enregistrés au format wav) dans une matrice de taille $I \times L$ avec I le nombre de sources et L la longueur maximale des vecteurs s_i .

Vous pourrez utiliser la fonction `[s,fs,nbit] = wavread(filename)` qui charge un fichier wav dans un vecteur s , récupère la fréquence d'échantillonnage f_s et le nombre de bits par échantillon. Vous pourrez utiliser les fonctions `plot()` et `soundsc()` respectivement pour visualiser et écouter les échantillons contenus dans un vecteur.

Exercice 2 : Un mélange linéaire instantané entre plusieurs sources peut s'exprimer ainsi :

$$s_{\text{mix}} = \sum_i a_i \cdot s_i(t) \quad (1)$$

En utilisant la matrice des sources sonores, créez le mélange instantané pour $a = (0.7, 0.5, 0.7, 0.6)$. Ecoutez le résultat obtenu. Essayez avec d'autres valeurs de a .

1.2 analyse et synthèse spectrale

Exercice 3 : Créez le fichier `spectrogram.m` qui contient la déclaration :

`function [Sw] = spectrogram(s,w, rec)` qui calcule la matrice S_w , c'est à dire le spectrogramme du signal s en utilisant la fenêtre d'analyse w de longueur `N=length(w)`, avec un recouvrement `1/rec`. (on choisira par défaut `rec = 3` et la fenêtre de Hann). Vous utiliserez la fonction `fft()` pour calculer la transformée de Fourier. Pour extraire les trames, il vous faudra calculer les indices i_{deb} et i_{fin} correspondants à la trame courante.

Vous pourrez alors calculer le spectre d'amplitude et de phase à l'aide des fonction `A = abs(Sw)`, `P = angle(Sw)` et l'afficher avec `imagesc()`.

Exercice 4 : Implémentez la fonction `[s] = inv_spectrogram(Sw, w, rec)` qui synthétise le signal s à partir de son spectrogramme. Vous utiliserez la fonction `ifft(Sw)` qui calcule transformée de Fourier inverse. Vérifiez que vous obtenez bien le même signal en affichant la différence entre le signal d'origine et le signal reconstruit.

¹<http://www.mathworks.fr/>

²<http://www.gnu.org/software/octave/>

2 Spatialisation

Pour chacune des méthodes proposées, vous commencerez par spatialiser une seule source isolée en essayant respectivement les angles $\{0, -\frac{\pi}{4}, \frac{\pi}{2}\}$ puis vous créerez un mélange instantané à partir des sources spatialisées en prenant par exemple : $\theta = (0, -\frac{\pi}{7}, \frac{\pi}{6}, \frac{\pi}{3})$.

2.1 amplitude panning

C'est une méthode fréquemment utilisé en studio d'enregistrement qui ne repose sur aucun modèle physique. Elle permet de simuler la position d'une source à un angle $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ en modifiant l'amplitude de la source sonore sur canaux gauche/droite relativement à θ :

$$s_g(t) = \cos\left(\frac{\theta}{2} + \frac{\pi}{4}\right) \cdot s(t) \quad (2)$$

$$s_d(t) = \sin\left(\frac{\theta}{2} + \frac{\pi}{4}\right) \cdot s(t) \quad (3)$$

Exercice 5 : Créez une fonction `[sg,sd] = amp_panning(s, theta)` qui calcule les signaux s_g et s_d en fonction de θ . Affichez puis écoutez le résultat obtenu avec `soundsc()` (vous devrez d'abord créer une matrice $L \times 2$ en concaténant s_g et s_d et en transposant la matrice obtenue : `[sg;sd]'`). Créez un mélange en appliquant les angles θ_i respectivement pour chacune des source. Ecoutez le mélange $s_g - s_d$. Commentez le résultat obtenu.

2.2 spatialisation binaurale par filtrage HRTF

On se propose maintenant d'appliquer un filtre lié à un modèle afin de calculer les signaux s_g et s_d correspondant à un modèle de tête. Les signaux sont obtenus en appliquant un produit de convolution entre le signal monophonique et la réponse impulsionnelle du filtre correspondant défini ainsi :

$$s_g(t) = (h_{g,\theta} * s)(t) = \int_{-\infty}^{+\infty} s(t - \tau) h_{g,\theta}(\tau) d\tau \quad (4)$$

$$s_d(t) = (h_{d,\theta} * s)(t) = \int_{-\infty}^{+\infty} s(t - \tau) h_{d,\theta}(\tau) d\tau \quad (5)$$

avec $h_{.,\theta}$ le filtre défini pour un angle θ .

Exercice 6 : Choisissez un fichier modèle (.mat) parmi ceux proposés (cf. dossier `model`) et importez les matrices `left` et `right` dont chaque colonne correspond à la réponse impulsionnelle du filtre h_θ pour θ allant de 0 à 355 par pas de 5 degré.

Ecrivez la fonction `[sg,sd] = spatialize_model(s, theta)` qui calcule s_g et s_d . Vous pourrez utiliser la fonction `conv(a,b)` qui applique un produit de convolution entre les vecteurs a et b .

D'autres modèles HRIR sont téléchargeables à l'adresse suivante : <http://interface.cipic.ucdavis.edu/sound/hrtf.html>.

2.3 spatialisation binaurale par modélisation paramétrique

On souhaite définir un filtre HRTF (Head Related Transfert function) à partir des indices ITD (Interaural Time Differences) et ILD (Interaural Level Differences) qui correspondent respectivement à une mesure perceptive de différence d'intensité en amplitude et au retard mesuré entre s_g et s_d . La spatialisation est définie ainsi dans le domaine fréquentiel :

$$S_g(f) = S \cdot 10^{-\Delta_a(f)/2} \exp(-i\Delta_\phi(f)) \quad (6)$$

$$S_d(f) = S \cdot 10^{+\Delta_a(f)/2} \exp(+i\Delta_\phi(f)) \quad (7)$$

avec :

$$\Delta_a(f) = \text{ILD}_\theta(f)/20 \quad (8)$$

$$\Delta_\phi(f) = \text{ITD}_\theta(f) \cdot 2\pi f \quad (9)$$

et :

$$\text{ITD}(\theta, f) = \beta(f) \cdot \frac{r(\sin \theta)}{c} \quad (10)$$

$$\text{ILD}(\theta, f) = \alpha(f) \cdot \sin(\theta) \quad (11)$$

On choisira $r = 0.0725\text{m}$ le rayon de la tête, $c = 335\text{m.s}^{-1}$ la vitesse du son. $\alpha(f)$ et $\beta(f)$ sont des fonctions d'échelle relatives respectivement à ILD et ITD.

Exercice 7 : Ecrivez la fonction `[sg, sd] = spat_param(s, theta)` qui calcule s_g et s_d en utilisant les indices ILT et ITD définis ci-dessus. Vous pourrez utiliser les fonction `alpha_spat(N)` et `beta_spat(N)` qui retournent un vecteur de taille N correspondant aux fonction $\alpha(f)$ et $\beta(f)$.

Exercice 8 : Comparez les résultats obtenus entre les méthodes (tests d'écoute) pour différents angles θ . Laquelle, vous semble la plus réaliste ?

3 Localisation binaurale

La localisation d'une source sonore peut s'effectuer à partir de l'estimation des indices ILT et ITD obtenus à partir du spectre.

On applique les étapes suivantes :

1. on calcule $\text{ILD}(t, f) = 20 \log_{10} \left| \frac{S_d(t, f)}{S_g(t, f)} \right|$.
2. on estime $\theta_L(t, f) = \arcsin \left(\frac{\text{ILD}(t, f)}{\alpha(f)} \right)$ (on coupera les valeurs $\left| \frac{\text{ILD}(t, f)}{\alpha(f)} \right| > 1$)
3. on estime ITD_L à partir de θ_L en utilisant l'équation (10).
4. on calcule $\text{ITD}_p = \frac{1}{2\pi f} \left(\angle \frac{S_d(t, f)}{S_g(t, f)} + 2\pi p \right)$
pour $p_1 = \lfloor pr \rfloor$ et $p_2 = \lceil pr \rceil$ avec
 $pr = f \cdot \text{ITD}_L - \frac{1}{2\pi} \angle \frac{S_d(t, f)}{S_g(t, f)}$
5. $\hat{\theta}(t, f) = \theta_{t, m}$ avec $m = \arg \min_p (|\theta_L - \theta_{t, p}|)$ et $\theta_{t, p} = \arcsin \left(\frac{c \cdot \text{ITD}_p}{\beta(f) \cdot r} \right)$
6. on construit un histogramme $h(\theta)$ en accumulant l'énergie du spectre correspondant en appliquant une gaussienne centrée en θ et réduite ($\sigma = 1$).

Ainsi, les angles de chaque source sont donnés par les maxima locaux de l'histogramme après l'application d'un seuillage.

Exercice 9 : Implémentez la fonction `[theta] = localize(s, fs)`, Vous pourrez vous aider des fonctions fournies dans l'archive pour le calcul de l'histogramme. Testez votre fonction sur un mélange binaural que vous aurez préparé en utilisant la spatialisation HRTF.