

Examen M2 IAAA option SAM

Durée 2h - documents autorisés - connexion internet interdite

29 janvier 2019

Exercice 1 (Sous-échantillonnage, environ 8 points). Vous justifierez brièvement vos réponses.

1. La figure 1 représente une image originale et deux versions sous-échantillonnées d'un facteur 0.3 :
 - (a) quelle est l'image originale ?
 - (b) quelle image contient du repliement spectral ?
 - (c) quelle image est un sous-échantillonnage sans repliement ?
2. La figure 2 représente des spectres de sons dans un contexte de sous-échantillonnage.
 - (a) quel est le facteur de sous-échantillonnage ?
 - (b) quel est le spectre du son avant sous-échantillonnage ?
 - (c) quel spectre contient du repliement spectral ?
 - (d) quel spectre correspond à un sous-échantillonnage sans repliement ?
 - (e) l'un des spectres n'a rien à voir avec les autres, lequel ?

Exercice 2 (Sur-échantillonnage, environ 6 points). En séance, nous avons vu comment sous-échantillonner un son —c'est-à-dire changer sa fréquence d'échantillonnage en la diminuant en affectant le moins possible le contenu du son— et les précautions à prendre pour éviter le repliement spectral. On souhaite ici sur-échantillonner un son, en passant à une fréquence d'échantillonnage plus élevée. On prendra comme exemple le cas d'un sur-échantillonnage d'un facteur 3 représenté sur la figure 3. Sur-échantillonner un son d'un facteur 3 revient à conserver un échantillon sur 3 et à déterminer la valeur des autres échantillons insérés entre les échantillons connus (ou au-delà du dernier). Une technique pour sur-échantillonner consiste à

- calculer le spectre du son original par transformée de Fourier discrète ;
- construire le spectre du nouveau son à partir du spectre du son original ;
- calculer la transformée de Fourier discrète inverse du nouveau spectre pour obtenir le nouveau son.

1. Si N est la longueur du son initial, quelle est la longueur du son sur-échantillonné ?
2. Pourquoi n'y a-t-il pas de problème de repliement spectral dans le cas du sur-échantillonnage, contrairement au sous-échantillonnage ?
3. Le spectre du son original est de taille N : quelle est la taille du nouveau spectre ?
4. Quel est l'intervalle de fréquences sur lequel est défini du nouveau spectre ?
5. Quelles valeurs affecter aux coefficients du nouveau spectre ?
6. Le spectre du son original est donné sur la figure 3 : dessinez le spectre du nouveau son en veillant à bien indiquer toutes les informations, notamment sur les axes.

7. Bonus (2 points) : le sur-échantillonnage d'un facteur entier (ici, 3) est simple à illustrer mais la technique ci-dessus peut s'appliquer plus généralement à des facteurs non-entiers. Donnez un exemple, expliquez les différences par rapport au cas précédent. Peut-on sur-échantillonner ainsi de n'importe quel facteur réel ? Dans le cas contraire, quels sont les facteurs de suréchantillonnage possibles ?

Exercice 3 (Détection d'événements sonores, environ 6 points). Pour détecter de manière robuste le début d'événements sonores tels que des notes de musique, la technique du flux spectral consiste à alterner des filtrages et des non-linéarités de la façon suivante :

- on considère un signal x de taille L ;
- on construit la matrice X de taille $K \times N$ obtenue en calculant le module de la STFT de x (aide : considérer la STFT comme un banc de filtres) ; un exemple est donné sur la figure 4a ;
- en général et comme illustré sur la figure 4a, un événement commence avec la naissance rapide de plusieurs composantes fréquentielles qui évoluent ensuite pour finalement disparaître : on peut faire ressortir le début de ces composantes en filtrant chaque ligne de X par un motif h (filtre FIR) tel que représenté sur la figure 4b ;
- on obtient une matrice Y de même taille que X dont une ligne est représentée sur la figure 4c : on somme alors toutes les lignes que la présence du motif dans les différents canaux se cumulent dans un seul vecteur z , représenté sur la figure 4d ;
- pour détecter le début des événements comme l'apparition soudaine du motif, on calcule la dérivée de ce vecteur, pour obtenir un vecteur dont les pics correspondent en général au début des événements que l'on souhaite détecter (voir figure 4e) ;
- une dernière étape de détection de pic permet de décider si on a un début d'événement à chaque instant.

On souhaite effectuer tout ce processus par un réseau de neurones qui prend en entrée un son de taille L et dont la sortie est un signal (vecteur) binaire dont les coefficients valent 1 lorsqu'un début d'événement est détecté et 0 sinon.

1. sur l'exemple du spectrogramme de la figure 4a, combien d'événements (notes de musique) voyez-vous et à quels instants commencent-ils ?
2. expliquez la forme du motif donné sur la figure 4b ;
3. comment calculer la dérivée d'un vecteur comme évoqué dans l'une des étapes ?
4. décrire aussi précisément que possible l'architecture d'un réseau de neurones permettant de mettre en œuvre ce système : types de couche, taille, nombre de filtre, poids, non-linéarité, etc. ; votre réponse peut comprendre du texte et des schémas, et vous pouvez vous concentrer sur les parties du système que vous maîtrisez le mieux ;
5. la dernière étape de détection de pic n'est pas décrite précisément : proposez une ou plusieurs solutions (vous pouvez traiter la question avec et/ou sans réseau de neurones) ;
6. en gardant cette architecture, on aimerait optimiser les poids des différentes couches en entraînant le réseau pour améliorer les performances en détection ; des annotateurs humains ont étiqueté des sons mais la localisation temporelle des événements n'est pas très précise dans ces données d'apprentissage (précision de quelques millisecondes) : par quel(s) mécanisme(s) peut-on aider le système à apprendre à détecter des événements avec une certaine tolérance sur l'instant du début des événements ?
7. question ouverte : comment pourrait-on faire une variante de ce système pour estimer le tempo, c'est-à-dire le nombre approximatif d'événements par seconde ? (il n'est pas demandé de construire un réseau de neurones)



(a)



(b)



(c)

FIGURE 1 – Sous-échantillonnage d'image.

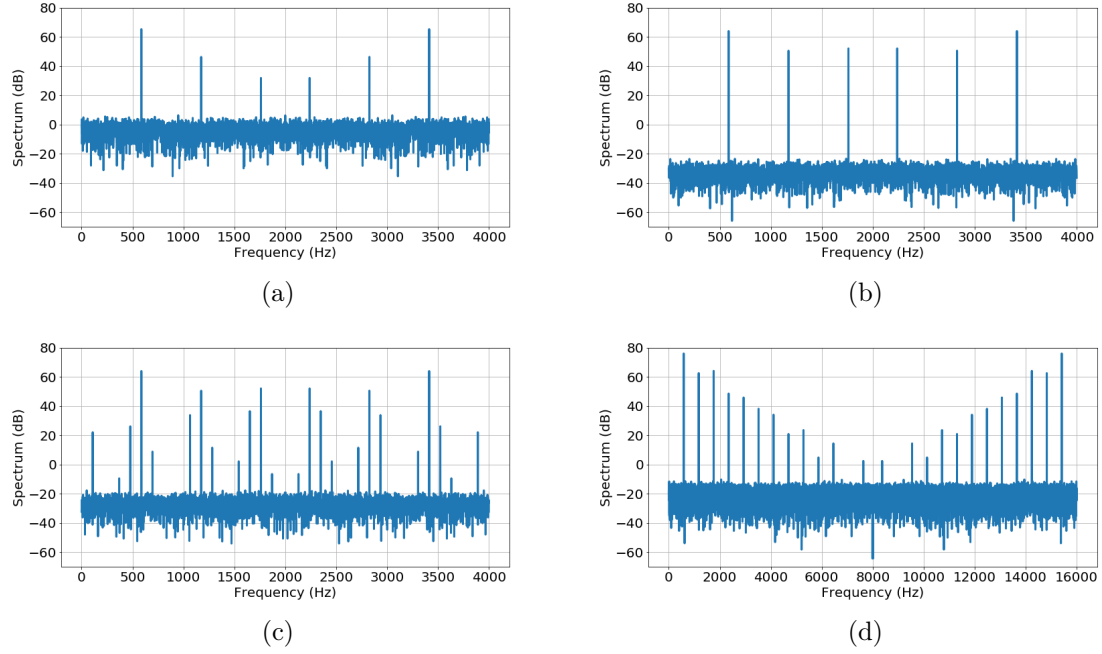


FIGURE 2 – Sous-échantillonnage de son : représentations fréquentielles.

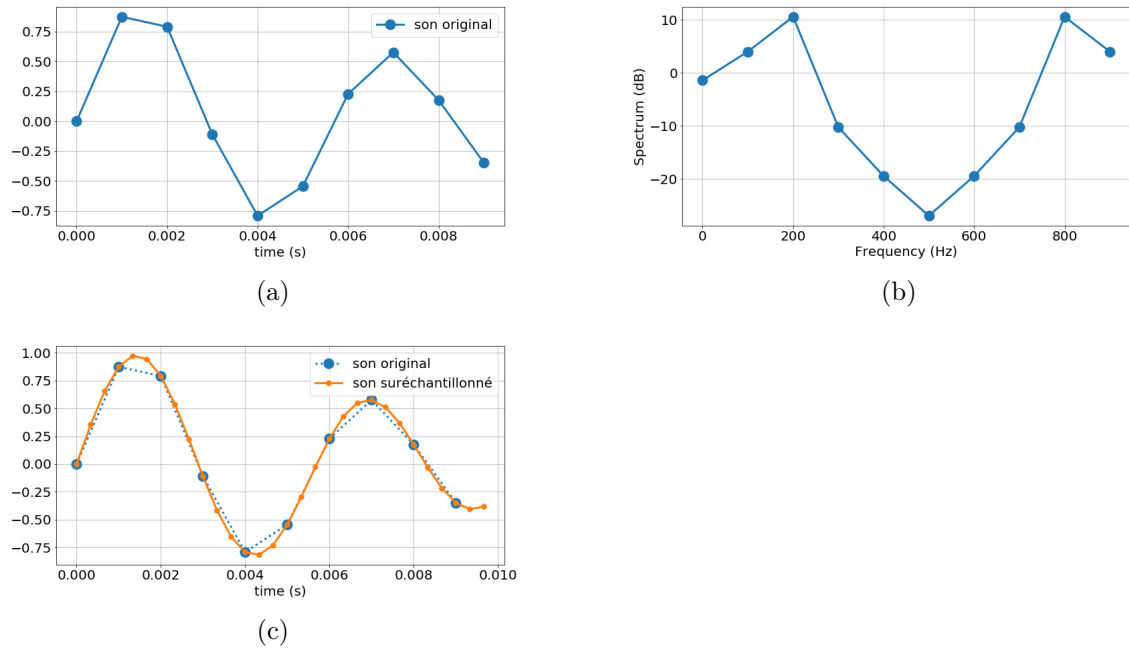
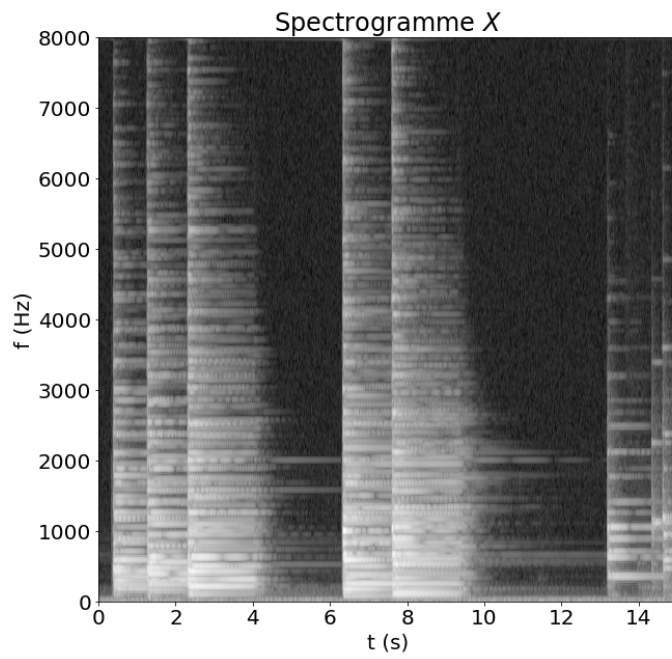
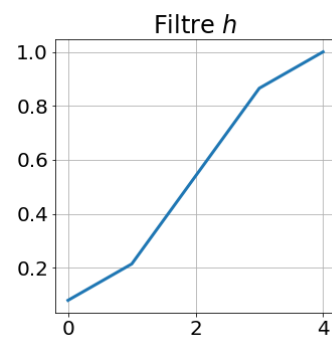


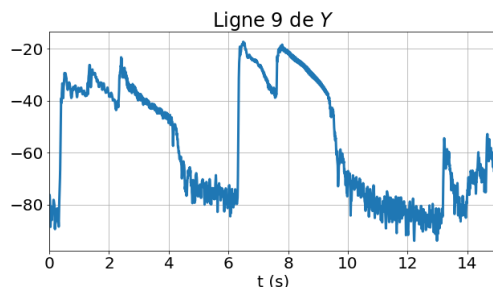
FIGURE 3 – En haut, son original de longueur $N = 10$ échantillonné à $f_s = 1000\text{Hz}$ à gauche et module de son spectre à droite. En bas, son suréchantillonné à la fréquence $3f_s$.



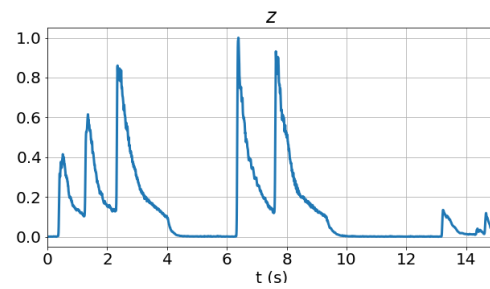
(a)



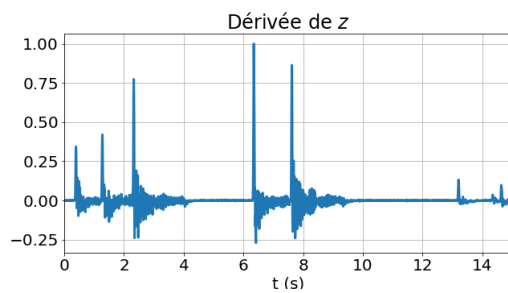
(b)



(c)



(d)



(e)

FIGURE 4 – Exemple de détection du début des notes dans un morceau de piano.