



M1 – Informatique

IAN – Introduction à l'Audio Numérique

Travaux Dirigés n°2

I. Spectrogramme

- Rappeler, grâce à des illustrations, le principe du calcul d'un spectrogramme (se fondant sur les calculs de spectres à court terme).
- Ecrire un algorithme qui permette de calculer le spectrogramme d'un signal, en fonction d'une taille des fenêtres d'analyse et avec un recouvrement de moitié.

II. Taux de passage par zéro

- Écrire une fonction en algorithmique pour calculer le taux de passage par zéro (ZCR : Zero Crossing Rate) d'une fenêtre (un vecteur d'échantillons normalisés).
- Écrire un algorithme faisant appel à la fonction précédente capable de traiter un signal par fenêtre glissante et produisant un vecteur correspondant au ZCR des fenêtres glissantes.

III. Energie à court terme

Sur le même principe que le ZCR (exercice précédent), proposer un algorithme qui calcule l'énergie à court terme d'un signal avec des fenêtres glissantes.

IV. Canaux MEL

L'échelle Mel a été conçue de telle manière que 1000 Hertz correspondent à 1000 Mels et qu'une variation constante dans le domaine des Mels soit perçue comme une variation constante de fréquence (en Hz) par les auditeurs.

Par exemple, en considérant une fréquence de base de 1000 Hertz, alors la fréquence à l'octave supérieure est 2000 Hertz. Cependant, un son à 2000 Hz ne sera pas perçu par un auditeur comme étant 2 fois supérieur à la fréquence de base. Par contre, le changement d'octave dans le domaine des Mels, 1 000 Hertz correspondent à 1 000 Mels. L'octave supérieure est donc à 2 000 Mels, ce qui équivaut à environ 3428 Hertz. Un son ayant une fréquence de 3428 Hertz sera donc perçu comme étant deux fois plus aigu qu'un son à 1000 Hz¹.

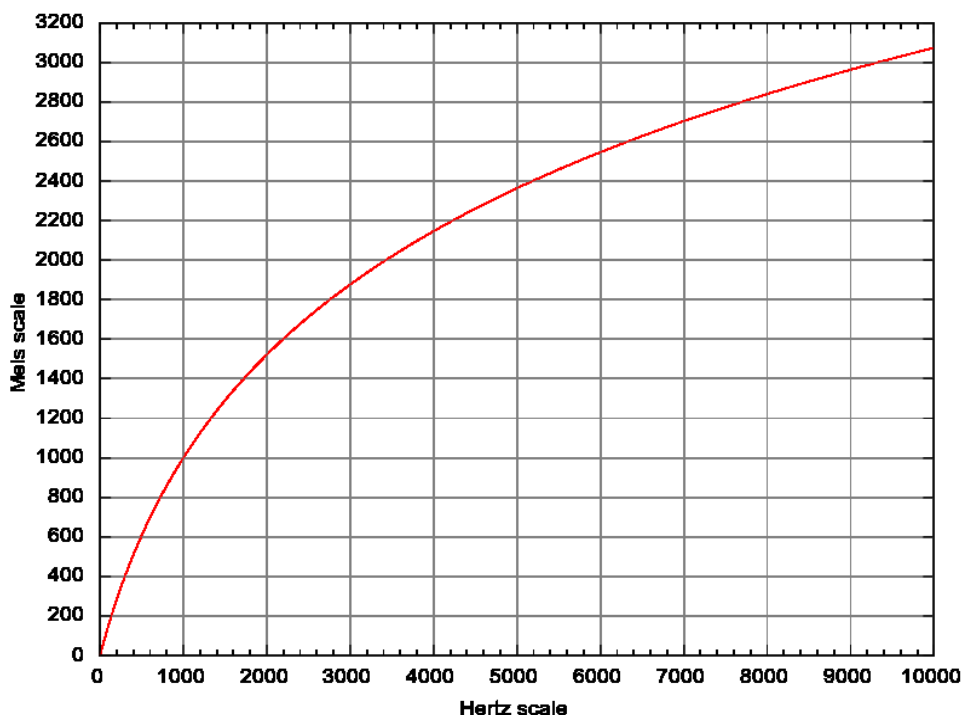
- Avec la formule Mel suivante² (avec f en Hz) :

$$Mel(f) = 2595 * \log_{10} \left(1 + \frac{f}{700} \right),$$

renverser la formule pour exprimer les fréquences en Hertz en fonction des Mels.

¹ Source : Wikipedia http://fr.wikipedia.org/wiki/%C3%89chelle_de_Mel

² Il existe plusieurs versions de la formule, celle-ci correspond à celle énoncée par Fant en 1949



b) En considérant une largeur de bande de fréquences de 0 à 8000 Hertz, si l'on souhaite découper en 14 zones homogènes selon l'échelle MEL, quelles vont être les fréquences en Hertz correspondantes ?

c) **Pour aller plus loin...** Considérons que les fréquences de coupure en Hertz soient réparties de la façon suivante (valeurs adaptées de la psycho-acoustique pour les systèmes de traitement de la parole) :

$f_{\text{coup}} = [0, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1150, 1300, 1500, 1700, 2000, 2350, 2700, 3100, 3550, 4000, 4500, 5050, 5600, 6200, 6850, 7500]$

Écrire un algorithme qui permette d'appliquer une pondération triangulaire en utilisant ces fréquences de coupures sur le spectre, pour produire les valeurs d'énergie de canaux MEL.