

Sujet 1

Exercice n°1 :

- 2 a) Expliquez brièvement les termes pertinent, discriminant et robuste
pertinent \rightarrow pas de redondance de données
discriminant \rightarrow dans les variables
robuste \rightarrow résistant au bruit
- 1 b) Expliquez brièvement les termes variabilité intra-locuteur, interlocuteur, coarticulation, bruit
intra \rightarrow même locuteur inter \rightarrow locuteurs différents
coarticulation \rightarrow influence du voisinage d'un phonème
bruit \rightarrow tout signal perturbateur
- 2 c) D'un point de vue articulatoire que différencie un phonème sourd d'un phonème sonore ?
Sourd \rightarrow vibration des cordes vocales sourd \rightarrow pas de vibration
- 2 d) Expliquez les étapes de numérisation du son ? à quel niveau se fait ce traitement ?
échantillonnage \rightarrow ~~codage~~ \rightarrow quantification \rightarrow codage
mais : code son
- 1 e) A quoi correspond la fréquence fondamentale du point de vue articulatoire et acoustique
la fréquence de vibration des cordes vocales
- 2 f) Expliquez brièvement les étapes du calcul d'un spectrogramme à partir du signal de parole
signal \rightarrow prétraitement (réduction de l'effet de bord, réduction de l'effet de bruit) \rightarrow fenêtrage \rightarrow FFT \rightarrow Spectrogramme
avant fréquence - parties haute et basse
stationnaire du signal passage au domaine fréquentiel
- 1 g) Le codage MFCC tien de 2 théories en traitement de parole expliquez brièvement
 \rightarrow Théorie additive \rightarrow Echel MEL
 \rightarrow Théorie de l'articulation \rightarrow séparation source filtre (usage central)

Exercice n°2:

h) Soit les figures 1 et 2, avec (justification) :

a. Précisez la fréquence d'échantillonnage du signal

11.25 Hz - 22.050 Hz

b. Y'a-t-il vibration des cordes vocales ?

Oui → signal périodique

c. Le signal est-il quantifié sur 8 ou 16 bits ?

16 bits → signal $6 \times [10506, 10506] \rightarrow [-128, 128]$

d. Quelle est la période du signal (en ms) et sa fréquence fondamentale F0 (en Hz)

$P = \frac{151}{21} \rightarrow 7 \text{ ms}$ $F_0 = \frac{1000 \text{ ms}}{7 \text{ ms}} = 143 \text{ Hz}$

e. D'un point de vue acoustique qu'elle est la nature du phonème prononcé

Voyelle → formants stables

f. Déduire la fréquence des formants F1 et F2

$F_1 \rightarrow 200 \text{ Hz}$ $F_2 \rightarrow 2200 \text{ Hz}$

g. A partir de la figure 2 déduire quelle est le phonème le plus proche

ly ou lh

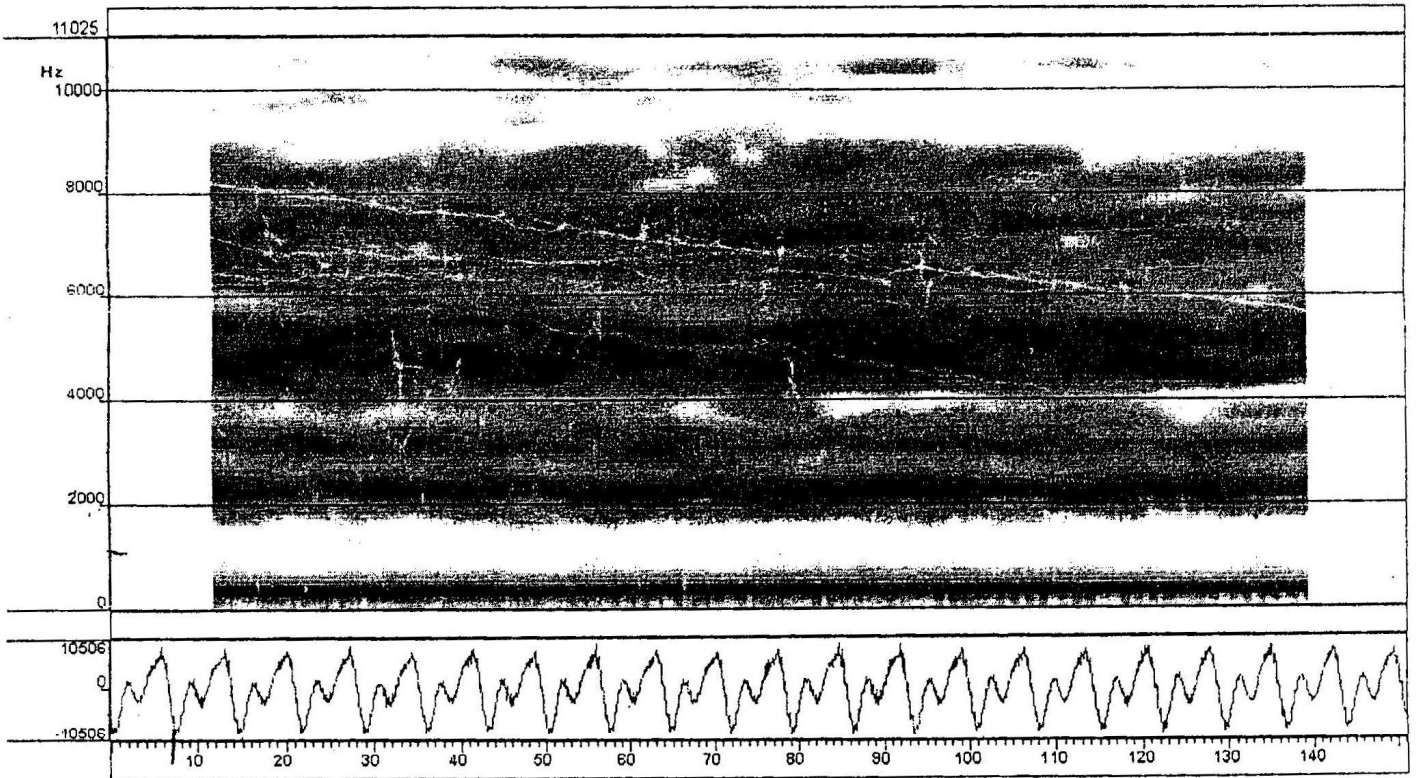


Figure 1 Visualisation Signal/Spectre lissé, en abscisses le temps en (ms) en ordonnées Fréquence en (Hz)

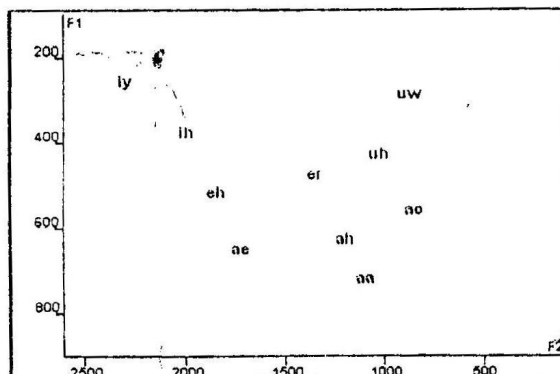


Figure 2

Sujet 2

Date : 16/06/10

EMD en Traitement de Parole

Durée : 1h

Master : RFIA

Exercice n°1 :

1 a) Parmi les 4 classes : voyelles, semi-voyelles, plosive, fricative quelle est la plus courte dans le temps ?, pourquoi

0,1 plosive, mode d'élocution plosive, Mode d'élocution

0,1 b) Parmi les 3 classes : voyelles, plosives, fricatives quelle est en général la plus énergétique? voyelles, voyelle

0,1 c) Cette énergie se concentre t elle en hautes ou en basses fréquences? basses fréquences, basses fréquences

0,1 d) Comment peut-on renforcer les fréquences à faible énergie? filtre passe haut, filtre passe haut

0,1 e) Quel est le rôle du codage dans une application de traitement de parole?

extraction d'informations

1,1 f) Expliquez les termes : Pertinent, Discriminant, Robuste

pertinent : pour le traitement de données

discriminant : séparabilité des classes

robuste : insensible au bruit

peut redondance de données
séparation des classes
robustesse de bruit

1,1 g) Citez les étapes du calcul des coefficients MFCC en précisant le rôle de chaque étape

prétraitement (réchantillonnage, normalisation)

filtrage (élimination effet de bord)

FFT (transformée en fréquence)

filtrage mel (passage à l'échelle mel - oreille humaine)

log (lissage spectral, réduction du nombre de coeff)

DTFT

Exercice n°2 (Chaque réponse doit être justifiée)

Soit un signal de parole contenant 8 phonèmes représenté par les spectrogrammes, figures 1 et 2

0,1 a) Donnez un titre aux figures 1, 2 et 3 (justifiez)

0,1 Figure 1 : Spectre bande étroite (harmoniques)

0,1 Figure 2 : Spectre ligne sériale (formants)

0,1 Figure 3 : Triangle vocalique (F1, F2, voyelles)

0,1 b) Précisez la fréquence d'échantillonnage du signal, ainsi que le nombre de bits utilisé en codage (justifiez)

22000 Hz, 11024 * 2 (effet miroir)

16 bits (144 * 2 = 288)

0,1 c) Sachant que les abscisses des figures 1 et 2 représente le temps en milliseconde, combien d'échantillons sont dans la

3,9 * 22000 = 85999

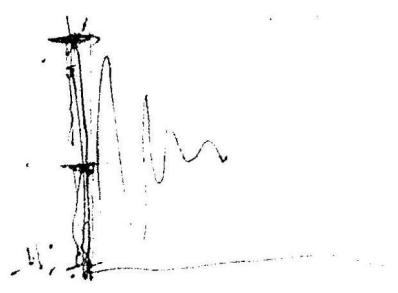
0,1 d) En utilisant une fenêtre de 32ms et une FFT de taille 1024 pour calculer le spectrogramme, combien d'échantillons sont dans la

fenêtre dans la matrice du spectrogramme (justifiez)

1024 / 32 = 32 (effet miroir)

4 e) Précisez la classe phonétique de chaque phonème avec justification.

- 1 : fricative (hautes fréq)
- 2 : voyelle (harmoniques, formants stables)
- 3 : plosive (court durée, hautes fréq)
- 4 : voyelle
- 5 : fricative
- 6 : voyelle
- 7 : semi-voyelle (largeur des formants, du a au e)
- 8 : voyelle



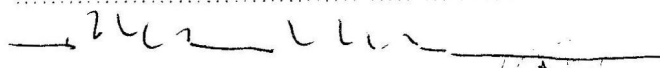
0,1 f) Parmi les phonèmes 1 et 5, lequel présente un caractère sonore, justifiez

5 - amériques

1 g) Positionnez les voyelles (leurs numéros) sur la figure 3.

h) Déduire la nature de chaque voyelle.

1 2 : o 4 : i 6 : a 8 : a



22/06/10
10:00