

Master RFIA

Examen de rattrapage en
Traitement de la parole

Durée : 1h

Questions : (veuillez répondre directement sur la feuille)

1. Quelle est la différence entre les méthodes de reconnaissance globales et analytiques ?
 - globale : considère le mot comme unité de reconnaissance, indécomposable
 - analytique : considère différents unités minimale de reconnaissance (syllabe, phonème, ...)
2. Listez les différentes étapes d'un système de reconnaissance de la parole, précisez le rôle de chaque étape ?
 - 1) Prétraitement : réception du signal et préparation de ce dernier aux étapes suivantes
 - 2) Codage : extraction de l'information utile du signal prétraité.
 - 3) Classification : grâce à une base d'exemples, on passe à l'étiquetage et la prise de décision quand à la nature de la forme
3. Quelle est le rôle d'un module de décodage acoustique phonétique ? DAP
Le rôle d'un DAP est de transformer le signal acoustique, en une suite d'unités phonétiques sans connaissance des niveaux supérieurs
+ Signal \rightarrow segmentation \rightarrow paramétrisation \rightarrow ^{trait} acoustique \rightarrow classification
4. Quelle est la différence entre la phonétique articulatoire, acoustique et auditive ?
 - * articulatoire : étudie l'exploitation des organes articulatoire de l'appareil phonatoire
 - * acoustique : étudie l'onde sonore produite par les organes de la phonation
 - * auditive : étudie l'appareil auditif (réception de l'onde sonore).
5. D'un point de vue articulatoire que caractérise une voyelle, une fricative, une plosive, une semi voyelle ?
 - voyelle : vibration des cordes vocales, passage de l'air librement après la glotte. (fixation)
 - fricative : petite ouverture, passage de l'air est rétréci mais pas interrompu. (fixation)
 - plosive : passage de l'air est fermé puis ouverture subite.
 - semi-voyelles : coarticulation, passage d'une voyelle à une autre
6. D'un point de vue acoustique (spectrogramme) que caractérise une voyelle, une fricative, une plosive, une semi voyelle ?
 - voyelle : les harmoniques
 - fricative : grand nuage noir dans les hautes fréquences,
 - plosive : silence suivie d'une barre d'explosion
 - semi-voyelle :
7. D'un point de vue articulatoire, que caractérise une consonne sonore, orale et une consonne sourde ?
sonore : vibration des cordes vocales, sonette ; pas de vibration des cordes.
8. D'un point de vue acoustique, que caractérise une consonne sonore d'une consonne sourde ?
 - consonne sonore : ressemble aux voyelles, ayant quelques harmonique (production d'onde périodique)
 - consonne sourde : pas de structure harmonique
9. Quelle est la fréquence d'échantillonnage d'un signal de 4 secondes contenant 44100 échantillons ?
 $44100/4 = 11025$
10. Quelle est la différence entre un signal quantifié à 8bit et un signal quantifié à 16bit ?
 - sur 8 bit : 256 valeurs possibles : -128 \rightarrow +128.
 - sur 16 bit : 65536 valeurs possibles : -32768 \rightarrow +32768.

11. Pour une voyelle, qu'elle est la caractéristique parmi l'amplitude, la fréquence, le timbre qui est la plus susceptible de modifier sa nature? Expliquez.

Le timbre : car contrairement à la fréquence et l'amplitude qui garde l'aspect générale du signal signalé, le timbre change complètement son aspect.

12. Que représente la fréquence fondamentale (F0), et quel genre d'informations on peut en déduire.

- F0 correspond à la période de l'onde
- la fréquence de cette onde détermine la hauteur du son ; les harmoniques sont des multiples entiers de F0

13. Expliquez comment construire un triangle vocalique d'une langue, et son rôle.

- Spectre à large bande → détermine les 2 premiers formants de chaque voyelle prononcée
- On trace les 2 axes du F1 et F2 et on marque les points correspondant aux voyelles.

rôle : lors de la reconnaissance d'une voyelle, on calcule la distance entre (F1, F2) et les points déjà présent sur le TV.

14. Expliquez brièvement l'échelle Mel et son rôle dans le domaine de la reconnaissance de la parole.

c'est une échelle logarithmique, qui modélise la perception de l'oreille humaine suivant les fréquences

Echelle Mel et une courbe logarithmique qui est linéaire puis qui prend une direction logarithmique $\frac{Mel}{Hz}$

15. Quelle est la différence entre un spectrogramme à bandes étroites, un spectrogramme à bandes larges et un cepstre.

- Spectre à bandes étroites : large fenêtre (10-32ms) Bonne résolution fréquentielle, Analyse temporelle moins fine
- Spectre à large bandes : petite fenêtre (3-6ms) Bonne " temporelle, Analyse frq moins fine
- Cepstre : transformation du spectre pour dégager la fct de transfert du conduit vocal.

16. Expliquez brièvement le rôle des formants et la théorie qui s'y attache sur la production des voyelles.

- Les formants (essentiellement les 2 premiers) sont utilisés dans la construction du triangle vocalique
- les formants produits à la source par les cordes vocales subissent un processus d'atténuation et de renforcement pour la production de voyelles.

17. Expliquez brièvement le principe de la transformée de Fourier, et les avantages qu'elle offre la visualisation fréquentielle

Avantages :

- rapide à mettre en oeuvre
- offre une méthode d'analyse temps-fréquence.

18. Donnez une brève définition du bruit et expliquez les différents types dans le domaine de la RAP.

bruit : toute perturbation qui engendre la modification du signal de la parole :

↳ bruit additif : stationnaires, sont simplement des bruits additionnés au signal original

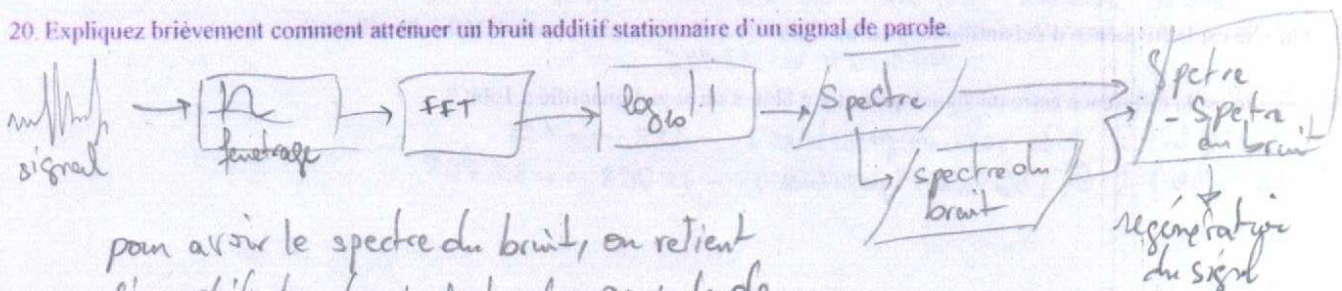
↳ bruit convolutif : le signal bruité est un produit de convolution entre signal original et signal du bruit

↳ bruit physiologique : spécifique à la morphologie de chacun.

19. Citez les facteurs qui entraînent une variabilité inter-locuteur et intra-locuteur

- intra-locuteur : l'état physique ou morale de la personne, humeur, émotion
- inter-locuteur : différence du timbre de voix entre une personne et une autre, cause physiologique

20. Expliquez brièvement comment atténuer un bruit additif stationnaire d'un signal de parole.



pour avoir le spectre du bruit, on retient l'amplitude du signal dans les périodes de silence, celle-ci détermine l'aplanissement celle du bruit.