

Thème : Reconnaissance Automatique de Parole et Vérification du Locuteur
=> 3 réponses au choix (a ou b ou c) dont une seule est correcte

1) Un système de reconnaissance de parole (RdP) dit "Speaker-independent" ou « multi-locuteurs » à grand vocabulaire (contenant plus de 10.000 mots) :

- a) permet de vérifier l'identité d'une personne à partir de quelques mots par elle;
 - b) permet à n'importe quel utilisateur a priori de prononcer un mot quelconque inclus dans ce vocabulaire étendu, ceci sans phase d'apprentissage préalable;
 - c) permet de reconnaître des mots de commande nécessitant un apprentissage préalable;
-

2) Quelle est la principale propriété de la Transformée du Cepstre (les MFCC en reconnaissance de parole et en vérification du locuteur) appliquée à chaque fenêtre d'analyse de signaux de parole échantillonnés ?

- a) De séparer l'excitation du conduit vocal de fonction de transfert égale à $H(z)=1/A(z)$;
- b) De renforcer le signal de parole pour sa transmission en appliquant directement le filtre $H(z)=A(z)$;
- c) De reconnaître directement le signal de parole ;

Où les MFCC sont les « Mel Frequency Cepstrum Coefficients ».

3) Pour reconnaître, à partir du vecteur d'observations noté O (signaux de parole analysés par la transformée cepstra de type MFCC), un mot donné correspondant au modèle de Markov caché (N états gauche-droit) noté M, quel critère l'algorithme de reconnaissance doit-il optimiser ?

- a) Max { $P(M/O)$ } ou Max { $P(O/M)$ } selon certaines hypothèses d'équiprobabilité sur les différents modèles M possibles du vocabulaire;
- b) Max { $P(M)$ } revenant à maximiser la probabilité d'occurrence de chaque mot possible;
- c) Max { $P(O)/P(M)$ } c'est-à-dire normaliser le maximum de vraisemblance;

où $P(O)$ est la probabilité de l'observation O, $P(M)$ est la probabilité du modèle de mot M, $P(M/O)$ est la probabilité a posteriori et $P(O/M)$ la probabilité a priori ou vraisemblance.

4) Pour réaliser un système de Vérification du Locuteur utilise-t-on ?

- a) Un modèle de type HMM à N états tel que $N \geq 3$ et on peut reconnaître le locuteur parmi une

population de M locuteurs différents;

- b) Une topologie de modèle avec un seul état observable et une densité de probabilité d'observation fondée sur un mélange de densités Gaussiennes, étant supposé l'existence d'un modèle universel du monde;
 - c) Une comparaison directe du spectre des observations avec celui des références;
-

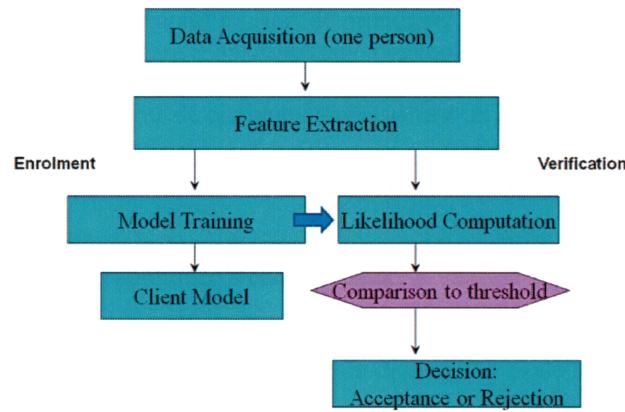
5) Pour contrôler les performances pures d'un système de reconnaissance en général, utilise-t-on :

- a) les valeurs directement des vraisemblances de sortie des modèles ;
 - b) une matrice de confusion donnant les taux de reconnaissance par mot ou item à reconnaître ;
 - c) le temps de réaction du système à reconnaître chaque mot ou item ;
-

Questions générales (documents autorisés)

1. Quel est le but d'un algorithme de Clustering?
2. L'algorithme des K-moyennes est un des algorithmes de Clustering les plus utilisés. Pourquoi l'appelle-t-on K-moyennes ?
3. Comment appliqueriez-vous l'algorithme des K-moyennes à un problème de partitionnement d'une base de données de chiffres manuscrits (de 0 à 9)? (SVP suivre les étapes ci-dessous pour compléter votre réponse).
 - a. Comment choisiriez-vous K? Justifiez votre réponse.
 - b. L'algorithme est-il sensible à l'initialisation? Qu'est-ce que cela signifie?
 - c. Comment pourriez-vous gérer le problème de l'initialisation?
 - d. Après l'initialisation, l'étape de clustering consiste à affecter chaque échantillon au cluster dont le centre est le plus proche. Comment est alors défini le nouveau centre ou prototype de chaque cluster?
 - e. Quel est le sens du critère optimisé?
 - f. Comment la convergence est-elle définie?
4. Comment utiliseriez-vous un Modèle à Mélange de Gaussiennes (Gaussian Mixture Model ou GMM) pour partitionner les données de la question 3? Quelle est la différence avec l'algorithme des K-moyennes?

5. Un Modèle à Mélange de Gaussiennes (Gaussian Mixture Model ou GMM) λ est entraîné sur les données d'un client. Quel est l'algorithme d'apprentissage utilisé?
6. Quels sont les paramètres définis dans le modèle client λ ?
7. Etant donné une séquence de vecteurs de paramètres $X=(x_1, x_2, \dots, x_T)$, quelle est la probabilité maximisée lors de l'apprentissage?
8. Expliquer les étapes principales d'un système de vérification d'identité biométrique, montrées ci-dessous:



9. Quelle est la règle de décision pour accepter ou rejeter un échantillon de test?
10. Quels sont les 2 types d'erreur du système biométrique?
11. Expliquer les moyens suivants d'évaluer les performances d'un système biométrique:
 - a. Equal Error Rate (EER)
 - b. Courbes DET (Detection Error Tradeoff Curve) et expliquer comment on obtient la courbe
 - c. Qu'est-ce qui est mieux pour évaluer un système biométrique: une courbe DET ou l'EER?