
Damien Dessimoz

RECONNAISSANCE DE LOCUTEURS : COMPARAISON DE PERFORMANCES
ENTRE LA RECONNAISSANCE AUDITIVE PAR DES PROFANES ET DE
SYSTEMES AUTOMATIQUES.

Diploma Project : 2004

Résumé :

Cette recherche a comme premier but la comparaison des capacités de reconnaissance vocale de personnes sans connaissance en phonétique et de systèmes automatiques, et ce, dans des conditions forensiques, telles que des appels anonymes, des menaces téléphoniques et des paroles spontanées. Dans un deuxième temps, les caractéristiques utilisées dans l'approche auditive lors du processus de comparaison ont été identifiées, en vue de leur intégration dans l'approche automatique afin d'améliorer ses résultats.

Les enregistrements utilisés proviennent de la base de données « Polyphone IPSC02 ». Deux méthodes de transmission de la voix ont été prises en compte lors du choix des enregistrements pour évaluer l'approche auditive et l'approche automatique : le réseau fixe communauté et le réseau GSM. Quatre conditions d'évaluation ont ainsi été créées : deux conditions semblables (Fixe-Fixe et GSM-GSM) et deux conditions différentes (Fixe-GSM et Fixe-Bruit).

Nonante profanes francophones, dont les voix des locuteurs leur étaient non familières, ont été évalués dans l'approche auditive. En ce qui concerne l'approche automatique, deux méthodes de modélisation des caractéristiques décrivant l'enveloppe spectrale du locuteur ont été utilisées, la modélisation par mélange de gaussiennes (GMM) et la modélisation par quantification vectorielle (VQ). Les résultats de cette étude permettent de mettre en évidence de meilleures performances pour l'approche automatique dans des conditions semblables par rapport à l'approche auditive. Cette dernière approche est également moins sensible face à des modifications de conditions d'enregistrement. De ce fait, l'approche auditive obtient de meilleures performances dans des conditions différentes par rapport à l'approche automatique.

Dix caractéristiques auditives citées par les profanes ont pu être identifiées lors de recherche. De plus, une stabilité de ces éléments face à des modifications de conditions a pu être observée. Cette recherche a également permis d'améliorer les performances de l'approche automatique, en intégrant une caractéristique supplémentaire, la fréquence fondamentale f_0 , statistiquement indépendante des caractéristiques déjà utilisées dans l'approche automatique pour décrire l'enveloppe spectrale du locuteur. Cette démarche d'intégration de nouveaux éléments devrait permettre d'améliorer les performances de l'approche automatique dans des conditions différentes, conditions les plus souvent rencontrées dans le domaine de la reconnaissance de locuteurs en sciences forensiques.