



SP3 : Segmentation et Authentification vocale

Rappel des objectifs :

proposer une méthode de reconnaissance du locuteur dépendante du texte, efficace en terme de temps de calcul et de ressources mémoire, apte à être adaptée en ligne à l'environnement sonore et prévue pour la fusion de modalités

LIA



SP4 : Segmentation et Authentification conjointes des deux modalités

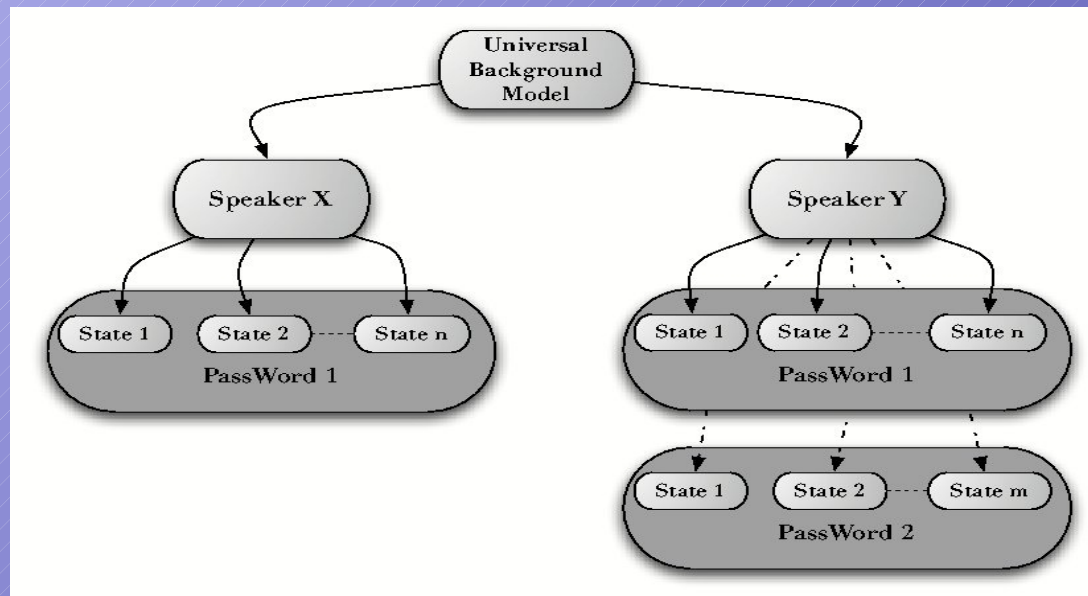
Rappel des objectifs :

Investiguer les possibilités d'un système bi-modal, audio+vidéo, selon les trois modes vu précédemment : Audio synchronisé par la vidéo, vidéo synchronisé par l'audio, impact du bimodal sur l'imposture

LIA

LIA_EBD

(modèle de reconnaissance de locuteur dépendant du texte et embarqué)



- Modèle à 3 niveaux (global, locuteur, structure du mot de passe)
- Occupation mémoire et temps de calcul réduits
- Alignement temporel prévu avec des marqueurs vidéo



Bases de données Audio-Vidéo

ValidDB :

- peu de locuteurs (76 hommes)
- seulement 2 phrases différentes
- contenu inadapté à la reconnaissance de mots de passe

MyIdea :

- deux sets de données
- BioMet : 4 sessions de données propres
- Banca : 4 sessions de 3 enregistrements, données propres et dégradées (indexées ?)

Myldea



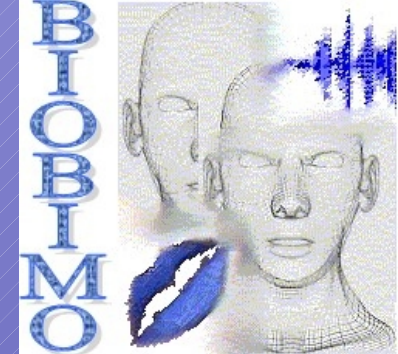
BioMet :

- 2 séries de digits
- 2 phrases phonétiquement équilibrées
- oui / non
- 10 phrases fixes
- 4 phrases mot de passe
- 4 phrases mot de passe imposteur
- 5 phrases aléatoires

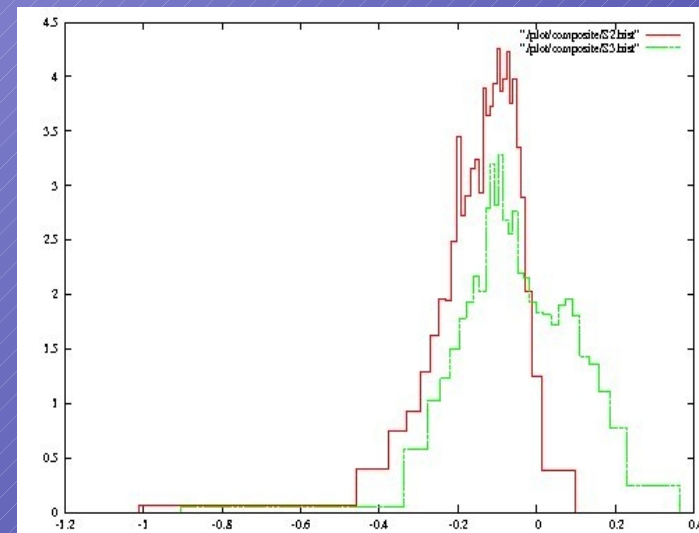
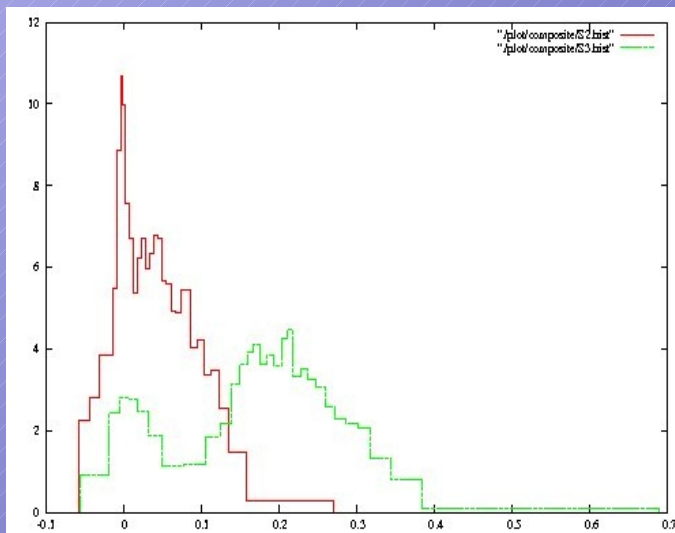
Banca :

- 12 digits
- nom et prénom / rue / ville / date de naissance
- 12 digits dans un autre ordre (imposteur)
- données imposteur (nom et prénom / rue / ville / date de naissance)

Travaux récents



- Développement du module de reconnaissance du locuteur dépendant du texte pour système embarqué
- Travail sur la séparation des phrases de ValidDB (reconnaissance du mot de passe, indépendant du locuteur)



Travaux récents



- Validb n'est pas adaptée à l'utilisation « mot de passe »
- Travail à mener : adaptation des modèles (locuteurs et mots de passe), algorithme d'adaptation, nombre d'états des HMMs et probabilités de transition
- Pour fixer la topologie de notre modèle, on valide la reconnaissance de mots de passe avec la base audio BJSON

Taux de reconnaissance des digits de la base BJSON

1 état	68,07%
5 états	71,46%
10 états	72,46%
15 états	74,64%