



SP4 : Segmentation et Authentification
conjointes de deux modalités (voix/visage)

Protocole Constitution Base de Données audio

LIA

Généralités



- Acquisition possible avec différents systèmes.
- Même fréquence d'échantillonnage pour toutes les données (16kHz)
- Pas de contraintes d'environnement
- Pour un même locuteur : sessions espacées de plusieurs jours
- « Password Set » : 10 phrases courtes (type mot de passe de 1 à 3 secondes)

3 populations de locuteurs par genre sont nécessaires:

- population client
- population imposteur
- population monde

Population Client



20 locuteurs du même genre minimum

Chaque locuteur devra prononcer par session d'enregistrement:

- le « PasswordSet »
- 2 phrases longues de 15s (sans contrainte sur le contenu linguistique)
- énumération en continu de nombres de 1 à 30

Note: si la session d'enregistrement s'avérait être trop longue, l'énumération pourra se substituer à une des deux phrases longues de 15s.

10 sessions d'enregistrement seront nécessaires. La première sera considérée comme la session d'entraînement et devra être réalisée dans un environnement non bruité.

Population Imposteur



10 locuteurs du même genre minimum

Chaque locuteur devra prononcer par session d'enregistrement:

- le « PasswordSet »
- 2 phrases longues de 15s, sans contrainte sur le contenu linguistique
- énumération en continu de nombres de 1 à 30

Note: si la session d'enregistrement s'avérait être trop longue, l'énumération pourra se substituer à une des deux phrases longues de 15s.

Le nombre de sessions n'est pas imposé a priori, la variabilité (en termes de sessions et de locuteurs) doit être la priorité.

Population Monde



20 locuteurs du même genre minimum

Aucune contrainte hormis une durée minimale de 30 à 40 secondes.

Le nombre de sessions n'est pas imposé a priori, la variabilité doit être la priorité.



DISCUSSION

LIA