

Bi-modalité & Résistance aux impostures



RAPPEL:

- Système AUDIO «assisté» par la vidéo
- Dépendance au texte

Impostures Considérées



AUDIO

- Imitation de la voix
- Utilisation du mot de passe d'un autre utilisateur

AUDIO-VIDÉO

- vidéo de l'utilisateur + voix d'un imposteur
 - avec le mot de passe de cet utilisateur
 - avec un énoncé différent

Structure Temporelle & Impostures AUDIO



Taux d'Égales Erreurs dans les différentes configurations pour des modèles à 512 Gaussiennes
Intervalle de confiance +/- 0,09

	GMM Baseline	LIA_EBD AUDIO seul
Tous les textes	3,06	2,68
Texte imposteur = mot de passe	3,68	3,55
Texte imposteur différent	2,11	1,06

Configuration LIA_EBD: 512G, 512 poids adaptés

Structure Temporelle & Impostures AUDIO



Taux d'Égales Erreurs dans les différentes configurations pour des modèles à 128 Gaussiennes.
Intervalle de confiance +/- 0,09

	GMM Baseline	LIA_EBD AUDIO seul
Tous les textes	3,67	1,72
Texte imposteur = mot de passe	4,35	2,23
Texte imposteur différent	2,72	0,50

Configuration LIA_EBD: 128G, 128 moyennes adaptées

Bi-modalité: Apport de la VIDEO



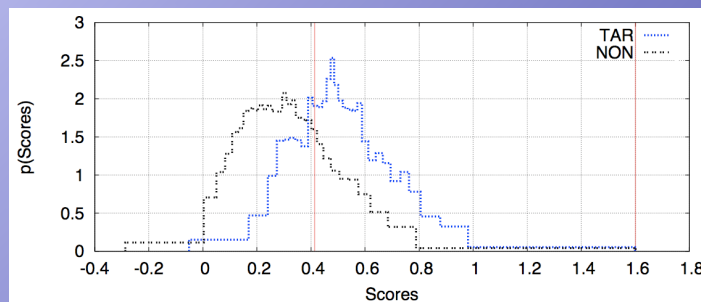
Taux d'Égales Erreurs dans les différentes configurations pour des modèles à 128 Gaussiennes. Intervalle de confiance +/- 0,09

	GMM Baseline	LIA_EBD AV - LIVE	LIA_EBD AV - Stored
Tous les textes	3,67	3,39	3,24
Texte imposteur = mot de passe	4,35	4,50	4,28
Texte imposteur différent	2,72	1,79	1,83

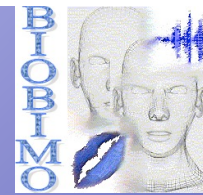
Configuration LIA_EBD: 128G, 128 poids adaptés

Voix + Mot de Passe du client MAIS Vidéo Imposteur

Configuration LIA_EBD: 128G, 128 poids adaptés



L'ajout de la vidéo a un effet visible sur les scores



Conclusions

- L'Architecture audio améliore les performances
- L'ajout de la vidéo apporte une information
- Si le système audio est «trop» performant, la vidéo n'apporte plus rien (dégrade éventuellement)

Perspectives

- Trouver une information vidéo plus «précise»
- Améliorer l'interaction entre audio et vidéo