

Titre : Identification du locuteur par un robot mobile
LABO : LIA, Avignon
Niveau : M2 informatique

Le sujet de stage proposé se situe dans le cadre de la reconnaissance automatique du locuteur par un robot mobile. Un des objectifs majeurs du travail de stage est de prendre en compte de manière aussi réaliste que possible les conditions réelles d'utilisation du robot, en l'occurrence le bruit ambiant et celui du robot lui même, les phénomènes de réverbération. Les durées des enregistrements de test est un aspect qu'il faut aussi prendre en compte.

En attendant la disponibilité des données de test et d'apprentissage (dont l'acquisition peut prendre du temps), la simulation des différentes perturbations acoustiques est une méthode tout-à-fait envisageable, cette approche s'est révélée efficace pour la reconnaissance de la parole ou encore la séparation de sources [Wang2018]. Le travail du stagiaire consiste à mettre en place un protocole expérimental prenant en compte les différentes difficultés intrinsèques à notre application (citées ci dessus).

Dans cette tâche le stagiaire évaluera avec A.I. Mergence (Entreprise fournissant le robot) les contraintes matérielles en terme de mémoire et de capacité de calcul de façon à proposer des modèles et des algorithmes sur mesure, par rapport aux capacités du robot.

Après avoir décidé et fixé le protocole expérimental, les données d'apprentissage et les données de test, le stagiaire procédera à la mise en place des techniques d'identification du locuteur de l'état de l'art, à savoir les i-vecteurs et les x-vecteurs [Bousquet2013] [Snyder2018]. Il s'agit de techniques existantes au sein du laboratoire, le travail du stagiaire consistera en l'adaptation au contexte matérielle de l'application visée (le robot).

Ce travail sera fait dans le contexte d'un projet ANR (ROBOVOX) qui vient d'être sélectionné pour en financement. L'ANR finance dans ce cadre une thèse sur 3ans, ce stage est un préambule à une thèse dans le domaine de la robustesse de l'identification du locuteur par les robots mobiles.

[Snyder2018] D. Snyder, D. Garcia-Romero, G.Sell, D. Povey, S. Khudanpur, « X-VECTORS: ROBUST DNN EMBEDDINGS FOR SPEAKER RECOGNITION », **In Proc. ICASSP 2018.**

[Bousquet2013] Pierre-Michel Bousquet, Jean-François Bonastre, **Driss Matrouf**, «Identify the Benefits of the Different Steps in an i-Vector Based Speaker Verification System», **Iberoamerican Congress on Pattern Recognition (CIARP) (2) 2013, 278-285**

[Wang2018] Z. Wang, E. Vincent, R. Serizel, and Y. Yan, "Rank-1 constrained Multichannel Wiener Filter for speech recognition in noisy environments." **Computer Speech & Language, 2018, 49, 37-51.**