

Sujet de Stage Master 2 Recherche

Les bruits de bouche dans la reconnaissance automatique de la parole.

Encadrants : [Michel VACHER](mailto:Michel.Vacher@imag.fr) (Michel.Vacher@imag.fr) [Solange.Rossato](mailto:Solange.Rossato@imag.fr) (Solange.Rossato@imag.fr)
[Frédéric AMAN](mailto:Frederic.Aman@imag.fr) (frederic.Aman@imag.fr)

Laboratoire LIG/GETALP, Campus Universitaire
B.P. 53, 38041 GRENOBLE Cedex 9

Sujet du stage

Le sujet proposé entre dans le cadre du projet ANR CIRDO [1] dont l'objectif est la détection de situations de détresse par analyse des sons et de la parole couplée à l'analyse de la scène vidéo. Les études réalisées montrent que les systèmes de reconnaissance automatique de la parole doivent être adaptés au cas particulier de la voix âgée [2]. Les enregistrements obtenus lors de la réalisation de corpus de parole âgée montrent la présence fréquente de bruits de bouche en début ou en cours d'énoncé. La présence de ces bruits perturbe de manière significative les performances des systèmes automatiques.

Le but du stage proposé est de déterminer la méthode à utiliser pour prendre en compte ce phénomène et corriger ses effets, soit en intervenant au niveau des modèles acoustiques et du jeu de phonèmes, soit en intervenant au niveau du modèle de langage. Le moteur de reconnaissance est le système Sphinx [3][4].

Un corpus de voix âgées a été enregistré, il sera utilisé lors de l'étude mais il devra être complété par de nouveaux enregistrements.

Principales étapes de l'étude :

- prise en compte du domaine,
- étude bibliographique,
- choix et implémentation d'une méthode,
- validation de la méthode sur les données existantes.

Cette étude pourra faire l'objet d'une poursuite en thèse.

Mots clés : Habitat Intelligent pour la Santé et l'Autonomie, Assistance pour la Vie Autonome, Reconnaissance Automatique de la Parole, voix âgée, bruits de bouche.

Lieu du stage : LIG, équipe GETALP, 41 rue des Mathématiques, campus universitaire.

Indemnité de stage prévue.

Références

- [1] Le site du projet CIRDO, <http://iris.cnrs.fr/cirdo/>.
- [2] M. Vacher, F. Portet, S. Rossato, F. Aman, C. Golanski, R. Dugheanu, Speech-based interaction in an AAL context, *Gerontechnology*, 11(2):310, July, 2012.
- [3] CMU Sphinx, Open Source Toolkit For Speech Recognition, project by Carnegie Mellon University, <http://cmusphinx.sourceforge.net/>.
- [4] Carnegie Mellon University, Speech at CMU Web Page, <http://www.speech.cs.cmu.edu/>.