

## Sujet de stage, M2

### Attaques des systèmes de reconnaissance du locuteur par transformation de la voix

Les systèmes de reconnaissance du locuteur de l'état de l'art sont basés sur la représentation spectrale à court terme [1], qui peut être représentée de différentes manières (mfcc, lpcc, plpcc, etc). En plus de l'information spectrale à court terme les systèmes modernes utilisent l'évolution de cette information sur plusieurs trames. Dans ce stage nous souhaitons développer une stratégie d'attaque de ces systèmes basée sur la transformation de l'enveloppe spectrale. La première stratégie que nous souhaitons développer est celle qui utilise un mapping entre les enveloppes spectrales (court terme) d'un locuteur et celles du locuteur qui fait l'objet de l'attaque. Afin de réaliser ce mapping nous pouvons dans un premier temps utiliser ce que l'on appelle UBM-GMM pour structurer l'espace [2]. Des expériences récentes, faisant un mapping entre les trames du locuteur-imposteur et les moyennes locales des trames du locuteur faisant l'objet de l'attaque, ont montré certaines limites. Ce stage sera l'occasion d'étudier cet aspect en profondeur et proposer des solutions.

Pour aller plus loin dans cette idée, il est proposé d'utiliser les DNN [3] pour réaliser ce mapping. Les DNN s'avèrent très efficaces dans ce genre de problématiques. Le stagiaire proposera des idées sur l'architecture du réseau de neurone et sa taille. Une fois cette première tâche réalisée, le stagiaire mettra en œuvre des solutions d'attaque permettant de contrer l'utilisation de l'information dynamique par les systèmes état de l'art. Des tests objectifs en terme de taux d'erreur seront réalisés, ainsi que des tests d'écoute plutôt subjectifs.

[1] F. Bimbot, J.-F. Bonastre, C. Fredouille, G. Gravier, I. Magrin-Chagnolleau, S. Meignier, T. Merlin, J. Ortega-Garcia, D. Petrovska, D. A. Reynolds, "A tutorial on text-independent speaker verification", EURASIP Journal on Applied Signal Processing, 2004, Vol.4, pp.430-451

[2] Jean-François Bonastre, Driss Matrouf, Corinne Fredouille, « Artificial impostor voice transformation effects on false acceptance rates », Interspeech 2007

[3] Pavel Matejka ... « Analysis of DNN Approaches to speaker identification », ICASSP 2016.