

Résumé

Le GSM (Global System Mobile) s'est imposé comme le moyen de communication téléphonique le plus répandu actuellement. La reconnaissance du locuteur à travers des enregistrements de parole de type GSM devient alors un sujet de recherche très important. Dans ce travail, nous avons mis en œuvre un système de reconnaissance du locuteur via une ligne de communication GSM. La ligne de communication GSM simulée et mise en œuvre est basée sur le codeur GSMEFR (Enhanced Full Rate). Le codeur GSMEFR correspond au protocole GSM 6.60 actuellement utilisé par les opérateurs de téléphonie mobile, il est basé sur l'algorithme ACELP (Algebraic Code Excited Linear Prediction). La simulation a été effectuée sous la boîte outils de **MATLAB version 7.4**. Le système de Reconnaissance Automatique du locuteur (RAL) développé est lui basé sur l'approche de modélisation par mélange de fonctions de densité de probabilité gaussiennes GMM (Gaussian Mixture Models). Une étude sur l'influence de la parole transcodée GSM sur les performances du système de RAL mis en œuvre a été réalisée avec différentes bases de données. Nous avons utilisé dans un premier temps la base de données ARADIGIT16K version original échantillonnée à 16KHz pour avoir la structure optimale de notre système de RAL. Ensuite, cette base de données a été sous-échantillonnée à 8KHz avec le logiciel **Praat** pour aboutir à la base de données bande téléphonique ARADIGIT8K. Cette base de données ARADIGIT8K a été passée à travers le codec GSM pour obtenir la base de données transcodée GSM (source uniquement) à savoir la base ARADIGIT_GSM. Le canal de transmission a été simulé par deux types de bruits, le bruit blanc additif AWGN (Additive White Gaussian Noise) et le bruit de Rayleigh à différents niveaux RSB (Rapport Signal Bruit). La base de données d'évaluation ARADIGIT16K est composée d'un vocabulaire de 10 chiffres arabes (de zéro à neuf) prononcés par soixante (60) locuteurs (31 locuteurs masculins et 29 locutrices), chaque chiffre est répété trois fois par chaque locuteur. Les résultats obtenus par les différentes bases de données ont montré une dégradation du taux d'identification du locuteur de l'ordre 6% pour la base transcodée GSM (source uniquement) et l'ordre de 16% pour la base de données transcodée GSM (source uniquement + canal simulé par le bruit AWGN) et de l'ordre de 24 % pour la base de données transcodée GSM (source uniquement + canal simulé par le bruit Rayleigh).