

∞ INTRODUCTION GENERALE ∞

❖ Notre motivation

La parole est sans doute le moyen de communication le plus simple et le plus efficace chez les humains. Depuis le début de la recherche dans le domaine du traitement du signal, les chercheurs ont toujours eu une attention particulière pour le signal de parole. Grâce aux développements dans les domaines de l'informatique, de la linguistique, ... etc, le rêve de communiquer avec des machines est devenu de plus en plus réalisable. Ainsi, chercheurs et industriels se sont intéressés au développement des applications utilisant la parole comme un moyen de communication homme-machine. Dans les dix dernières années, deux applications du domaine du traitement du signal de parole ont connu une progression considérable : La Reconnaissance Automatique de la Parole (RAP) qui consiste à reconnaître le message prononcé et la Reconnaissance Automatique du Locuteur (RAL) qui consiste à reconnaître l'identité du locuteur derrière le signal de parole présenté.

Théoriquement, la reconnaissance du locuteur est un terme générique pour discriminer parmi plusieurs personnes en fonction de leurs voix. Les variations individuelles entre locuteurs ont deux origines essentielles. En premier lieu, les caractéristiques physiques de l'appareil de phonation influencent les formants, la valeur moyenne du pitch,... et cela indépendamment de la phrase prononcée. D'autre part, une même phrase n'est pas prononcée de la même façon par deux locuteurs.

L'automatisation des systèmes d'authentification, en vue de leur déploiement à grande échelle, constitue par conséquent un enjeu important. Dans ce cadre, Théoriquement, la Reconnaissance Automatique du Locuteur (RAL) est un terme générique pour discriminer parmi plusieurs personnes en fonction de leurs voix. Les variations individuelles entre locuteurs ont deux origines essentielles. En premier lieu, les caractéristiques physiques de l'appareil de phonation influencent les formants, la valeur moyenne du pitch,... et cela indépendamment de la phrase prononcée. D'autre part, une même phrase n'est pas prononcée de la même façon par deux locuteurs.

En RAL, on distingue deux spécialités : l'identification du locuteur et la vérification du locuteur.

- ❖ L'identification consiste à reconnaître un locuteur appartenant à une population de plusieurs locuteurs ; on compare (généralement) pour cela son expression vocale à des références connues.
- ❖ La vérification consiste à accepter ou à refuser une identité proclamée par un locuteur ; dans ce but on compare à un certain seuil la distance entre son expression vocale et la référence proclamée.

On distingue également, selon la phrase prononcée, deux types de reconnaissances :

- La reconnaissance indépendante du texte (free-text speaker recognition). Les techniques d'identification actuelles tendent à s'intéresser à ce type de reconnaissance.
- La reconnaissance effectuée sur la base d'un texte imposé. Cette dernière procédure, par contre, est la plus courante pour la vérification du locuteur (fixed-text speaker verification).

La variabilité intrinsèque à la parole est, elle aussi, une source de dégradation reconnue. Au delà de ces variabilités, une des limitations majeures des systèmes actuels tient à la modélisation du locuteur par des informations correspondant à l'enveloppe spectrale à court-terme. La recherche d'informations nouvelles constitue un des principaux thèmes de recherche actuels.

❖ **Problématique :**

Le problème de la vérification de l'identité d'un locuteur est un problème de classification binaire. Un utilisateur potentiel prétend avoir une certaine identité qu'il énonce ; il s'agit de décider si cette identité prétendue est vraie ou fausse, c'est-à-dire d'accepter ou de rejeter l'individu. Le système peut prendre cette décision à partir d'une ou plusieurs modalités biométriques.

❖ **Objectifs :**

Dans cette étude, on s'est fixé deux objectifs principaux :

- ✎ Conception d'un classifieur basé sur les machines à vaste marge (SVM).
- ✎ Application sur des bases de données audio réelles (TIMIT).

❖ **Structure du mémoire :**

Ce mémoire est organisé comme suit : dans le premier chapitre nous présentons les méthodes de classification de données. Après une brève introduction, la classification des données comme une tâche de décision. Les techniques de classification utilisés, tels que les réseaux de neurones et les machines à vecteurs de support seront introduits. Nous allons rappeler la notion du neurone formel ainsi que les propriétés générales des réseaux de neurones (perceptron multicouche) à apprentissage supervisé par rétro-propagation de l'erreur. Dans le deuxième chapitre, les principes des systèmes de reconnaissance automatique du locuteur et les techniques couramment employées sont présentés. Les différentes tâches applicatives de la RAL ainsi que les techniques d'évaluation des systèmes sont détaillées. Le chapitre suivant décrit les méthodes discriminantes de modélisation du locuteur, basées sur les machines à vecteurs de support(SVM). Enfin, le dernier chapitre présente les méthodes proposées pour l'identification du locuteur.