

Différentes variantes GMM-SMOs pour l'identification du locuteur

Siwar Zribi Boujelbene****, Dorra Ben Ayed Mezghani*****, Nouredine Ellouze***

*** UR Signal Image et Reconnaissance de Formes, Ecole National des Ingénieurs de Tunis
- ENIT, BP-37 Campus Universitaire 1002, Tunis, Tunisie

*zribi_siwar@yahoo.fr

** dorra.mezghani@isi.rnu.tn

*** N.ellouze@enit.rnu.tn

Résumé. Dans cet article, nous présentons différentes variantes GMM-SMOs pour l'identification du locuteur en mode indépendant du texte. Pour mettre en œuvre les différents systèmes, nous avons opté une représentation multi-gaussienne de l'espace des caractéristiques basées sur l'algorithme Expectation Maximisation (EM). Ces nouvelles représentations constituent les vecteurs d'entrées pour entraîner les supports vecteurs machines (SVMs) par l'algorithme de type Optimisation par Minimisation Séquentielle (SMO).

1 Introduction

Les SVMs sont des nouvelles techniques de l'apprentissage statistique (Vapnik, 1995). Leur utilisation en reconnaissance du locuteur été très prometteuse. Dans ce papier, nous nous intéressons à mettre en place une combinaison des modèles GMMs et SVMs à différents noyaux. Les systèmes proposés ont la capacité de modélisation multi-gaussiennes (Reynolds et Rose, 1995) et l'efficacité de décision des SVMs de type SMO (Platt, 1999). Différentes variantes GMM-SMOs sont implémentés et comparés.

Notre article est organisé comme suit : les différents systèmes GMM-SMOs proposés sont présentés dans la section 2 et l'analyse des résultats est donnés par la section 3.

2 Systèmes GMM-SMOs pour l'identification du locuteur

Différentes variantes ont été conçues. Les points en communs des différents systèmes résident au niveau de la phase de paramétrisation, qui consiste à déterminer l'espace des caractéristiques extraite du corpus TIMIT, et de la phase de modélisation, qui consiste à discriminer les classes des locuteurs. Durant cette phase, nous avons utilisé la modélisation GMM basé sur l'algorithme EM ce qui permet de générer des nouvelles représentations de l'espace élaboré. Lors de la phase de décision, nous avons utilisés quatre variantes de noyau : noyau polynomial de degré 1, noyau polynomial de degré 10, noyau polynomial de degré 100 et noyau gaussien RBF. Pour étudier la similarité des caractéristiques d'une classe de locuteur avec toutes les autres classes des locuteurs, nous avons exploité la méthode un-contre-un.