

Les techniques de reconnaissance de Formes Application : à la reconnaissance de l'écriture Arabe

Rachid Zaghdoudi

Soutenue en: 2018

Abstract: Le domaine de la reconnaissance de l'écriture arabe manuscrite a reçu un intérêt de recherche intensif aux cours des deux dernières décennies et se concrétise aujourd'hui par un grand nombre d'applications opérationnelles. Malgré ces avancées, le problème reste ouvert à la recherche et à l'expérimentation à cause des caractéristiques morphologiques complexes du manuscrit arabe qui sont à l'origine de la défaillance des traitements. Les travaux de recherches présentés dans cette thèse s'inscrivent dans ce cadre et constituent une contribution au domaine de la reconnaissance de l'écriture arabe manuscrite. Dans un premier temps, nous proposons une approche globale basée sur une combinaison parallèle de classifieurs pour la reconnaissance hors ligne de mots arabe manuscrits. Le système développé s'articule autour de quatre modules distincts : prétraitement, extraction de primitives, classification basée sur deux types de classifieurs : SVM et k plus proches voisins flou et un module de combinaison. Des expériences ont été effectuées sur des images de la base de référence INF/ENIT ont montré l'efficacité du système proposé. En second temps, nous présentons un nouvel algorithme de segmentation en caractères des mots arabes manuscrits. Cet algorithme est fruit d'une coopération entre un ensemble de règles morphologiques spécifiques à l'écriture arabe (AMS) et un classifieur SVM mono classe, dont la tâche principale consiste à accepter ou rejeter les points de segmentation primaires générées par le module AMS. Il a montré de bons résultats de segmentation sur des mots extraits de la base de données de référence IFN/ENIT.

Keywords : reconnaissance de l'écriture arabe manuscrite ; combinaison de classifieurs ; séparateurs à vaste marge ; k plus proches voisins flou ; la transformée en cosinus discrète, histogrammes des gradients orientés ; segmentation.