



جامعة محمد الخامس بالرباط
Université Mohammed V de Rabat

École Nationale Supérieure d'Informatique et d'Analyse des Systèmes
Centre d'Études Doctorales en Sciences des Technologies de l'Information et de l'Ingénieur

AVIS DE SOUTENANCE DE THÈSE DE DOCTORAT

Monsieur Ahmed BOUMEHDI

Soutiendra publiquement sa thèse de Doctorat en Informatique

Le Jeudi 11 Juillet 2024 à 9h00 au Grand Amphi à l'ENSIAS de Rabat

Intitulé de la thèse

AMÉLIORATION DE LA RECONNAISSANCE AUTOMATIQUE

DE LA PAROLE EN ARABE AVEC LES SEMI-SYLLABES

Président :

Pr. El Mostafa DAOUDI, PES, Faculté des Sciences, Université Mohamed Premier, Oujda

Directeur de thèse :

Pr. Abdellah YOUSFI, PES, FSJES-Souissi, Université Mohammed V de Rabat

Rapporteurs :

Pr. Mostafa BELLAFKIH, PES, Institut National des Postes et Télécommunications, Rabat

Pr. Mohammed MOURCHID, PES, Faculté des Sciences, Université Ibn Tofail, Kénitra

Pr. Hassan SATORI, PES, Faculté des Sciences, Université Sidi Mohammed Ben Abdallah, Fès

Examineurs :

Pr. Si Lhoussain AOURAGH, PES, ENSIAS, Université Mohammed V de Rabat

Pr. Hicham GUEDDAH, PH, École Normale Supérieure, Université Mohammed V de Rabat



Résumé: La reconnaissance automatique de la parole est un domaine passionnant et en évolution rapide qui vise à permettre aux machines de comprendre et de transcrire le langage parlé humain. Dans cette thèse, nous nous concentrons sur la reconnaissance automatique de la parole en langue arabe, une tâche complexe en raison des spécificités linguistiques de cette langue.

Dans le premier chapitre, nous présentons les principes généraux de la reconnaissance automatique de la parole, ainsi que les modules impliqués dans ce processus. Nous expliquons également l'extraction des caractéristiques acoustiques à l'aide des MFCC basées sur l'échelle Mel, ainsi que l'utilisation des MMC pour construire des modèles de reconnaissance automatique de la parole.

Dans le deuxième chapitre, nous introduisons une nouvelle approche pour améliorer la reconnaissance automatique de la langue arabe en mettant en place un modèle indépendant du vocabulaire basé sur une unité phonétique novatrice que nous avons appelée "semi-syllabe". Cette unité phonétique permettra de reconnaître n'importe quel mot en arabe, ouvrant ainsi de nouvelles perspectives dans le domaine de la reconnaissance automatique de la parole.

Le troisième chapitre décrit la mise en place d'un système de reconnaissance automatique de la parole basé sur les semi-syllabes. Nous explorons les outils logiciels utilisés pour développer ce système, en mettant l'accent sur l'importance d'HTK (Hidden Markov Model Toolkit) pour la modélisation acoustique et la prédiction des transcriptions. Nous détaillons également la structure des fichiers et les différentes étapes nécessaires pour construire le système.

Le quatrième chapitre est dédié à l'amélioration des semi-syllabes en effectuant plusieurs expériences.

Nous redéfinissons les semi-syllabes pour mieux représenter les gémissements et les semi-syllabes sans voyelle, tout en prenant en compte leur position dans un mot. Ces expériences démontrent l'efficacité de cette approche, conduisant à une meilleure représentation acoustique des semi-syllabes et une amélioration significative du taux de reconnaissance automatique de la parole en arabe.

En conclusion, cette thèse offre des avancées significatives dans le domaine de la reconnaissance automatique de la parole en langue arabe. En utilisant les semi-syllabes comme unité phonétique, nous avons développé un système de reconnaissance automatique de la parole capable de reconnaître n'importe quel mot en arabe, indépendamment du vocabulaire spécifique. Ces travaux ouvrent de nouvelles perspectives de recherche pour améliorer encore davantage les performances des systèmes de reconnaissance automatique de la parole en arabe, ce qui est essentiel pour répondre aux besoins des locuteurs arabophones et faciliter l'intégration de la technologie de la parole dans divers domaines d'application.



Mots-clés : Darija, HTK (Hidden Markov Model Toolkit), Langue arabe, MFCC (Mel Frequency Cepstral Coefficients), Reconnaissance automatique de la parole, Semi-syllabes, Wav2vec.

Abstract: Automatic speech recognition is an exciting and rapidly evolving field that aims to enable machines to understand and transcribe human spoken language. In this thesis, we focus on automatic speech recognition in the Arabic language, a challenging task due to the linguistic peculiarities of this language. In the first chapter, we introduce the general principles of automatic speech recognition and the modules involved in this process. We also explain the extraction of acoustic features using Mel Frequency Cepstral Coefficients (MFCCs) and the use of Hidden Markov Models (HMMs) to build speech recognition models. In the second chapter, we propose a novel approach to enhance Arabic speech recognition by creating a vocabulary-independent model based on a new phonetic unit called "semi-syllable". This phonetic unit enables the recognition of any Arabic word, opening new possibilities in automatic speech recognition. The third chapter details the implementation of a speech recognition system based on semi-syllables. We explore the software tools used to develop the system, emphasizing the significance of HTK (Hidden Markov Model Toolkit) for acoustic modeling and transcription prediction. We also provide a comprehensive explanation of the file structure and the essential steps involved in building the system. The fourth chapter is dedicated to improving the semi-syllables through several experiments. We redefine semi-syllables to better represent geminations and those without vowels, considering their position within a word. These iterative experiments demonstrate the effectiveness of this approach, resulting in enhanced acoustic representation of semi-syllables and significantly improved automatic speech recognition rates in Arabic. In conclusion, this thesis presents significant advancements in Arabic automatic speech recognition. By utilizing semi-syllables as a phonetic unit, we have developed a speech recognition system capable of recognizing any Arabic word, regardless of the specific vocabulary. These findings open up new avenues for further research to enhance the performance of Arabic automatic speech recognition systems, catering to the needs of Arabic speakers and facilitating the integration of speech technology in various applications.

Keywords: Automatic Speech Recognition, Arabic, Arabic language, Darija, HTK (Hidden Markov Model Toolkit), Language, Semi-Syllables, Acoustic Modeling, MFCC (Mel Frequency Cepstral Coefficients), Wav2vec.