Fayoum University
Faculty of Engineering
Engineering Mathematics and physics Department

Novel Mathematical Model For Automatic Speech Recognition

By

Fatma Mohammed AbdEl_latif

A Thesis submitted in Partial Fulfillment
Of
The Requirements for the degree of

Master of Science in Engineering

Engineering Mathematics and Physics Department

(Engineering Mathematics)

Faculty of Engineering

Fayoum University

Egypt

2021

Fayoum University
Faculty of Engineering
Engineering Mathematics and physics Department

Novel Mathematical Model For Automatic Speech Recognition

By

Fatma Mohammed AbdEl_latif

Demonstrator of Engineering Mathematics

Master of Science in Engineering

Engineering Mathematics and Physics Department

Faculty of Engineering – Fayoum University

Supervised by

Prof. Dr. Hazem Ali Attia

professor of Engineering Mathematics
Engineering Mathematics and physics Department
Faculty of Engineering, Fayoum University

Signature: …….

Prof. Dr. Amr Mohamed RefaatGody

Professor of Digital Signals
Electrical Engineering Department
Faculty of Engineering, Fayoum University

Signature: …….

Dr. WaleedAbdelmagied Ahmed

Associate professor of Engineering Mathematics
Engineering Mathematics and physics Department
Faculty of Engineering, Fayoum University

Signature:…….

2021

Fayoum University
Faculty of Engineering
Engineering Mathematics and physics Department

Novel Mathematical Model For Automatic Speech Recognition

By
Fatma Mohammed AbdEl_latif

This thesis for M.Sc. Engineering Degree

Engineering Mathematics and PhysicsDepartment

(Engineering Mathematics)

Examination Committee

Name and Affiliation — Signature

Prof. Dr. Ramy Mohamed KamelTalaat
Emeritus Professor of Engineering Mathematics-
Faculty of Engineering - Fayoum University.

Prof. Dr. Hazem Ali Attia — Main Supervisor
Professor of Engineering Mathematics-
Faculty of Engineering –Fayoum University.

Prof. Dr. Amr Mohamed Refaat
Professor of Digital Signals- Faculty of Engineering-
FayoumUniversity

Dr. WaleedAbdelmagied Ahmed
Associate Professor of Engineering Mathematics,
Faculty of Engineering –Fayoum University.

Prof.Dr. Ahmed Hassan KamelMadian
Professor at Radiation Technology Research Center-
Atomic Energy Authority - Egypt.

Date of Examination: / /2021

Faculty of Engineering
Fayoum University
2021

ABSTRACT

Speech is the most natural form of human communication, and speech processing has become one of the most inspiring areas of signal processing. Speech recognition is the method of automatically identifying a person's spoken words based on information in a speech signal. An Automatic Speech Recognition (ASR) system takes a human speech utterance as input and returns a string of words as output. The definition of different types of speech classes, feature extraction methods, speech classifiers, and performance evaluation are issues that must be addressed when developing a speech recognition system. Feature extraction is the most essential aspect of speech recognition since it separates one speech from another. This research introduced an enhancement for a new feature extraction technique called Best Tree Encoding (BTE), which is based on wavelet packet decomposition and the best tree. The first method of enhancement of BTE is using various entropy functions. Shannon entropy (SE), Renyi entropy (RE) and Tsallis entropy (TE) are used in this research to extract features from the Wavelet Packet Tree (WPT) to obtain BTE. The second method of enhancement is using various mother wavelets to choose the best mother wavelet for Wavelet Packet Decomposition (WPD). Two Daubechies wavelets (db4, db5), two Coiflets wavelets (coif3, coif5) and one Symlets (sym4) with 4 decomposition levels are used in this research. The proposed model is tested and checked against the most widely used feature extraction technique Mel Frequency Cepstral Coefficient (MFCC) plus delta and delta-delta coefficients (39 parameters) to evaluate its performance. The TIMIT database is used in this research. All phones are divided into five classes Vowels, Fricatives, Silences, Nasals, and Plosives. The acoustical model has been implemented using Hidden Markov Model (HMM). No language model has been applied. The HMM Tool Kit (HTK) software is used for model implementation. The experiments show that in the first method of enhancement, The highest overall success rate of 75.85% was achieved using Tsallis entropy with entropy order 1.2. In the second method of enhancement, The highest overall success rate of 76.51% was achieved using db5 mother wavelet with Shannon entropy which is better than MFCC overall success rate of 71.76%. Comparing vector of 4 components of BTE to 39 component vector of MFCC makes it a very promising feature vector to be researching and developing.