Physics Department
Faculty of Science
Ain Shams University

Design and Implementation of a Stand-alone Voice Recognition System

A Thesis

"Submitted for Partial Fulfillment of the Requirements for M.Sc. Degree in Physics"

By

Mohammed Ramadan Saady Abd El Tawab

Demonstrator at Basic Science Department,
Faculty of Computers and Information,
Fayoum University,
Cairo, Egypt.

Supervised by

١- **Prof. Dr. Ashraf shamseldin Yahia**
Professor of Electronics and Electromagnetics,
Physics Department, Faculty of Science,
Ain Shams University.

٢- **Emeritus Prof. Dr. Hatem M. El Borai**
Professor of Electronics and Circuits,
Physics Department, Faculty of Science,
Ain Shams University.

٣- **Dr. El-Sayed A. El-Dahshan**
Assistant Professor of Computational physics,
Physics Department, Faculty of Science,
Ain Shams University.

٢٠١٦

Abstract

For the past several decades, the biometric security systems varied between those based on behavioral and others based on physiological features of people. One of the behavioral biometric security systems is that based on voice features of people and is called speaker recognition system. As a result of the advance of machine learning and computer technology, speaker recognition has rapidly evolved and has become very popular in the recent years. It is being intensively researched and found many applications where it saves lot of troubles which appear as a result of using other biometric security systems, so it is considered a high security system.

Many methods were presented for the aim of designing an automatic speaker recognition system and other methods were presented for the aim of developing reliable speaker recognition systems. In general, any speaker recognition system involves four basic steps: a) data base formation, b) pre-processing, c) features extraction, and d) classification or matching of extracted features.

In this work we have developed a method for speaker recognition using the English Language Speech Database for Speaker Recognition (ELSDSR) database which is compose of audio files for training and others for testing. The developed

method starts by pre-processing the training audio files of the database. Then the Wavelet Packet Transform (WPT) is employed on the pre-processed files for feature extraction purposes. For the excessive number of features provided with the WPT, the energy corresponding to each WPT node is calculated to reduce the dimensionality of the wavelet coefficients by removing redundant features and to form features vectors. The features vectors are sent to the Feed Forward Back-propagation Neural Network (FFBPNN) system.

Experimental results showed the effectiveness of the developed method by using the test audio files. Our results have also showed that the rate of correct recognition of the developed method is about ١٠٠٪ when using the training files and ٩٥,٧٪ when using a one testing file for each speaker from the ELSDSR database. The proposed method showed efficiency better than the well-known Mel Frequency Cepstral Coefficient (MFCC) and the Zak transform.

Table of Contents

List of Figures

List of Figures (cont….)

List of Tables

List of Abbreviations

ADALINE	ADAptive LInear Neuron
ANN	Artificial Neural Network
ANNs	Artificial Neural Networks
ART	Adaptive Resonance Theory
ASRS	Automated Speaker Verification System
CWT	Continuous Wavelet Transform
Db	Daubechies
db٢٠	Daubechies ٢٠-tap
DFT	Discrete Fourier Transform
DTFT	Discrete Time Fourier Transform
DWT	Discrete Wavelet Transforms
ELSDSR	English Language Speech Database for Speaker Recognition
FFBPNN	Feed-Forward Back-Propagation Neural Network
FFT	Fast Fourier Transform

IDWT	Inverse Discrete Wavelet Transform
IMM	Informatics and Mathematical Modeling
LMS	Least Mean Square
MADALINE	Multiple- ADAptive LInear NEuron
MFCCs	Mel Frequency Cepstral Coefficients
MRA	Multi-Resolution Analysis
RBF	Radial Basis Function
SIS	Speaker Identification System
SOM	Self-Organization Map
SRS	Speaker Recognition System
STFT	Short Time Fourier Transform
SVM	Support Vector Machines
SVS	Speaker Verification System
WPT	Wavelet Packet Transform