ENHANCEMENT SPEED OF LARGE VOCABULARY SPEECH RECOGNITION SYSTEM

By

Eslam Eid Ali Mohamed El Maghraby

A thesis submitted in partial fulfillment

Of

The requirement for the degree of

Master of Science

In

Communications and Electronics Department

Department of Electrical Engineering

Faculty of Engineering

Fayoum University

Fayoum University, Fayoum, Egypt
2013

I

ENHANCEMENT SPEED OF LARGE VOCABULARY SPEECH RECOGNITION SYSTEM

By

Eslam Eid Ali Mohamed El Maghraby

A thesis submitted in partial fulfillment

Of

The requirement for the degree of

Master of Science

In

Communications and Electronics Department

Department of Electrical Engineering

Faculty of Engineering

Fayoum University

Fayoum University, Fayoum, Egypt
2013

ENHANCEMENT SPEED OF LARGE VOCABULARY SPEECH RECOGNITION SYSTEM

By

Eslam Eid Ali Mohamed El Maghraby

A thesis submitted in partial fulfillment

Of

The requirement for the degree of

Master of Science

In

Communications and Electronics Department

Department of Electrical Engineering

Faculty of Engineering

Fayoum University

Under the supervision of:

Dr. Amr Mohamed Refaat Gody
Communication and electronic Department-Faculty of Engineering, Fayoum University.

Dr. Rania Ahmed Abdel Azeem Abul Seoud
Communication and electronic Department-Faculty of Engineering, Fayoum University.

Fayoum University, Fayoum, Egypt
2013

III

ENHANCEMENT SPEED OF LARGE VOCABULARY SPEECH RECOGNITION SYSTEM

By

Eslam Eid Ali Mohamed El Maghraby

A thesis submitted in partial fulfillment

Of

The requirement for the degree of

Master of Science

In

Communications and Electronics Department

Department of Electrical Engineering

Faculty of Engineering

Fayoum University

Approved by Examining Committee

Assistant Prof.Dr. Amr Mohamed Refaat   **(Main Advisor)**

Prof. Dr. Salwa Hussein Abdel Fattah Elramly   **Examiner**

(Electronics & Communications Engineering Department, Faculty of Engineering, Ain Shams University)

Prof.Dr.Abd El - Fatah Ahmed Shata   **Examiner**

(Electrical department, Faculty of Engineering, Fayoum University**)**

Fayoum University, Fayoum, Egypt
2013

ACKNOWLEDGMENTS

First of all, I would like to thank my supervisor Associate Prof. Dr. Amr M. Gody for his guidance in academics. He is a great researcher with keen interests in science, constant efforts in doing things by hand and great personality. Not only he gave me a gate to enter this new field, he has also advised and encouraged me all the time. Without his deep understanding of the best supervising manner for a student, I could not achieve the goal of my thesis. He is the example I followed and will continue to follow. I would like to thank my committee members for reviewing this thesis and providing valuable suggestions. I extend my sincere gratitude and appreciation to all the people who made this thesis possible. My sincere thanks are to Dr. Rania Ahmed Abul Seoud and Prof.Dr. Salwa Hessen Elramly for being on my thesis supervision. Finally, I'm grateful to my family for their support and encouragement of my academic pursuit. I thank my sisters for supporting me in these years. I am very grateful for the love and the unconditioned support of my mother who makes my life more colorful and enjoyable.

ABSTRACT

The purpose of this thesis is to design an Arabic phone recognition system by using the newly feature extraction method Best Tree Encoding (BTE4). This feature is inherited from the wavelet packets decomposition of speech signal that is highly reliable even in the presence of noise. Phone recognition in continuous speech is a tough task with a low accuracy rate. By making some modification to BTE4 feature, the speed and the quality of the recognition process can be improved.

The stages of enhancement had been done in this thesis are: Firstly use the newly wavelet based feature extraction method Best Tree Encoding (BTE4) to design our Arabic phone recognition system and try to enhance it by changing the level of decomposition to reach the second generation BTE5. This generation increases the resolution of the system and the discrimination between different speech signal. As a result of that the system doesn't take much time to recognize a new speech signal even for a new speaker so this increases the speed of the system. Adding extra 2 levels for BTE5 to reach the third generation BTE7 gives more accuracy for the system. Due to that the adaptation time for new speech signal will be decreased and the speech of the system will also be increased. Secondly adding new information about the input signal by adding log energy components to BTE4 and BTE5, will increase the feature discrimination which in turn will affect both the recognition rate and the speed of the system. Thirdly changing the Entropy type form Shannon to log energy will decrease the time needed in calculation.

The experiment results explained that Best Tree Encoding indicates a good start for new promising features. BTE5 gives more success rate over BTE 4 for the same recognition process. Information is stored into 15 bits instead of 7 bits of the original BTE4. The results become more

optimistic to the Automatic Speech Recognition arena by approaching to 25% Success Rate compared to maximum of 20% of the original BTE4. This 5% difference in SR is efficiency enhancement to the original BTE4. In addition; BTE5 feature' vectors with size 15, gives 25% success rate (SR). MFCC with size 13 gives 39% SR for the same problem. This is very promising that BTE5 is approaching 71% of the SR of the most popular feature used in the applied area of automatic speech recognition.

The 7 levels wavelet packet analysis of speech signal is applied in the pre-processing phase of speech signal. Adding 2 more analysis levels over BTE5, gives more resolution and enhances the speed of the process to store information which in return supports to make more discriminative features for Automatic Speech Recognition problem. In addition to this, including energy distribution over the 4 equal portions of the bandwidth instead of the total energy, gives more discriminative information about the phones. BTE7 gives 22% enhancements over BTE4 and BTE5. BTE7 indicates more accuracy and fast recognition process than both BTE5 and BTE4.