Project: Palestinian accent Recognition from speech

Students: Shahada Sammour, Nagham Addasi and Fatima Sabbah

Supervisor: Dr. Abualsoud Hanani

A brief summery about our project:

Human speech contains a wealth of information in addition to the message conveyed by the spoken words. This information includes the geographical, social and ethnic background of the speaker, his or her gender, age, emotional state and well-being. In this project we try to recognize the local Palestinian accents automatically from short speech (30-40 seconds) signal.

This technology has many applications. For example, regional accent recognition can be used to direct a speaker to a human or computer listener who is familiar with a particular regional accent. This might be a hotel or airport, or anywhere in the international telephone network. This technology could also be used to choose the computer speech recognition program that will work best for an individual, or the synthetic computer speech that he or she would most like to listen to. There are also potential forensic applications for computer programs that can detect an individual's geographical origins from a sample of his or her speech.

To our knowledge, this is the first work on automatic Palestinian accent recognition from speech. Therefore, the main challenge of this project is the unavailability of speech corpus for the Palestinian accents. Due to this and limited available resources, we have recorded for ten to fifteen volunteers from different seven cities in Palestine, namely Jerusalem, Al-Bireh, Hebron, Nablus, Lud, Tulkarem and Qalqilia. Each volunteer has been asked to talk about carefully selected topics and free speech about any topic he/she likes to talk about.

In the first stage of this project, we are trying to build a baseline system by applying the state-of-the-art techniques in the accent and dialect recognition area. These include acoustic and phonotactic based systems. Acoustic systems use well-known Mel Frequency Cepstral Coefficients (MFCC) to build accent models using a modelling technique such as Gaussian Mixture Models (GMM) and Support vector machines (SVM). In Phonotactic

systems, a sequence of phonemes is extracted form speech signal using one or more phone recognizer. A language model, such as n-gram, is built for each accent using the resulted phone sequences. The resulting models are then used to recognize the accent of speaker from short speech segment.