BIRZEIT UNIVERSITY

Faculty of Engineering and Technology

Electrical and Computer Engineering

DSP Fall 2020

Main Course Project

About the course project:

Teams of three students (maximum) must do this project. The best arrangement is to choose a division of the project so that each of you can work on separate but interlocking parts. Working individually or in team of four will not be accepted.

Learning teamwork is also one of the more general goals of this course, so team projects will pick up points for demonstrating a successful ability to work with others.

The projects will be graded based on a project report (of around 3-4 pages) as well as in-class short presentations (or discussion in my office) at the end of the semester.

Project submission must be via Moodle only (itc.birzeit.edu), but please use PDF format and **not** Word .DOC files if at all possible, since I often have formatting problems with Word files.

Your report must have the following structure, using these section headings and using IEEE paper format [template can be found on Moodle]:

Introduction: A general description of the area of your project and why you're doing it.

Problem Specification: A clear technical description of the problem you're addressing. Formulating a general problem (e.g., transcribing music) into a well-defined technical goal (e.g., reporting a list of estimated fundamental periods at each time frame) is often the most important part of a project.

Data: What are the real-world and/or synthetic signals you are going to use to develop and evaluate your work?

Evaluation Criteria: How are you going to measure how well your project performs? The best criteria are objective, quantitative, and discriminatory. You want to be able to demonstrate and measure improvements in your system.

Approach: A description of how you went about trying to solve the problem. Sometimes you can make a nice project by contrasting two or more different approaches.

Results and Analysis: What happened when you evaluated your system using the data and criteria introduced above? What were the principal shortfalls? (This may require you to choose or synthesize data that will reveal these shortcomings.) Your analysis of what happened is one of the most important opportunities to display your command of signal processing concepts.

Development: If possible, you will come up with ideas about how to improve the shortcomings identified in the previous section, and then implement and evaluate them. Did they, in fact, help? Were there unexpected side-effects?

Conclusions: What did you learn from doing the project? What did you demonstrate about how to solve your problem?

References: Complete list of sources you used in completing your project, with explanations of what you got from each.

The reason for this somewhat arbitrary structure is simply to help you avoid some of the more problematic weaknesses I've seen in past years. If you're having trouble fitting your work into these sections, you should probably think more carefully about your project.

<u>**Project description**</u>:

In this project, you need to design, implement and evaluate English vowel phonemes identification system using signal processing. English has 13 basic vowel phonemes, while Arabic has only six. The following 13 words contains 13 vowels in the form /h V d/, where V is the vowel.

heed, hid, head, had, hard, hudd, hod, heard, hoard, hood, who'd, hade, hide

The English vowels are different in their magnitude spectrum (Fourier transform). The general shape of the spectrum is distinctive for each vowel. Moreover, the first two format frequencies (F1 and F2) alone can be used recognize each vowel. The format frequencies are the frequencies corresponding to the peaks of the magnitude spectrum of the vowel. Therefore, F1 is the frequency corresponding to the first peak of the spectrum; F2 is the frequency corresponding to the second peak, and so on. In most cases and depending on the context, the vowels are different in their duration. A simple system can identify English vowels by estimating F1 and F2 for a given input vowel and then compare them against stored F1 and F2 of the 13 vowels.

<u>**Method1: Correlation**</u>: (Due in 16/11/ 2020)
use cross correlation to compare the signal segment of unknown vowel with the signal template (reference) of each of the 13 vowels and the recognized vowels the one that gives the highest correlation.
To get better performance, you need to get the signal segment aligned as accurate as possible.

<u>**Method2: Fourier transform**</u>: (Due in 25/11/2020)
Now, you need to develop the same system (English vowel identification system) but this time using Fourier transform. I.e. instead of taking cross correlation, you can take some details of the spectrum by quantizing spectrum using FFT . By this, each vowel is represented by a feature vector. These vectors can be used to identify vowel more accurately than just looking at two format frequencies; F1 and F2. In addition, you can append the vowel duration to the feature vectors. Therefore, you need to compute the mean vector of each vowel and store them in a lookup table. To recognize the input vowel, you compute its vector and match it with the stored vectors. The one that gives you the highest match is the recognized vowel. You can use cosine similarity, Euclidean distance, etc as a matching measure.

<u>**Method3: Filter bank**</u>: (Due in 5/12/2020)
Filter bank is an alternative method to the Fourier transform for estimating the spectrum features. Instead of using Fourier transform, you need to design a set of bandpass filters covering the signal bandwidth, and pass the vowel segment into each filter and compute the output. By, this you will get one output from each filter which represents the average power of the vowel in the filter band. These outputs are then used as feature vectors and a matching system, similar to the one described above in the Fourier transform method, is used to make the vowel recognizer.

Dataset:
You will be provided with sufficient samples of the 13 English vowels spoken by English native speakers (10 samples for each word). You need to divide this data into two subsets; training and testing. For example, you can use 7 for training and 3 for testing for each vowel. So, you build your system (three methods) on the training recordings and use testing recordings for testing your system and find its accuracy (percentage of the correctly identified vowels to the total testing vowels).

Tutorial on previous projects:
Eng. Ahmad Dar-Khalil made video tutorials for a past DSP projects and publish them on the YouTube. I highly recommend watching these videos before you start with your project.

https://www.youtube.com/playlist?list=PLnyw1IVZpaTsFgcU2QlK9x2jU8vIFaRBl

Project Deliverables: (submission deadline: 10/12/2020 23:55 on ITC)

1- Mini-report as described above.

2- System demonstration of each method as described above.

3- On a specific time before the final project submission deadline (initially 8/12/2020), I will provide you with a set of testing vowels with unknown labels. You will have two days to use your system to recognize them, and to send me the recognized vowels. The first five groups who get the highest scores will get a bonus.

You can use any programming language you prefer for implementing your project. However, I highly recommend MATLAB because it has many useful functions.

Abualsoud Hanani
November 1st, 2020