# Midterm 1: Speech Recognition

Suppose that we want to design a system to recognize the words for the ten digits: “zero,” “one,” “two,” , …, “nine.” One of the first things that we might do is analyze data values collected with a microphone for the ten corresponding sequences (or signals) to see if there are some statistical measurements that would allow us to distinguish these digits. The MATLAB data analysis functions allow us to compute these measurements easily. We could then print a table of the measurements and look for those that allow us to distinguish values. For example, one measurement might allow us to narrow the possible digits to three, and another might allow us to identify the specific digit from the three possible digits.

**Write a MATLAB program (midterm1a.m) to read and plot the Microsoft® wave files eight.wav, nine.wav, and zero.wav that contain the utterance of the word “eight”, “nine”, and “zero”, respectively.** **Use the subplot function to plot the utterance of “eight”, “nine”, and “zero” in one graph space.** All graphs should be titled clearly, for example, “Utterance of the word EIGHT”. The x-axis for both graphs should be labeled as “Index”.

**The program should also compute the following statistical measurements for the utterance data of “eight”, “nine”, and “zero”:**

1. The number of data in the file: **length**(x),
2. Mean: **mean**(x)
3. Standard deviation: **std**(x),
4. Variance (the squared value of standard deviation): **std**(x) .^ 2,
5. Average power (the average of the squared value of each data): **mean**(x .^ 2) , and
6. Average magnitude (the average of the absolute value of each data): **mean**(**abs**(x)).

The computation for “eight”, “nine”, and “zero” should be presented (in the command window) in the following format:

Utterance of the word EIGHT

number of data in file: n (with zero digit behind the decimal point)

mean: xxx.xxxxx (print 5 digits behind the decimal point for the rest of the values)

standard deviation: xxx.xxxxx

variance: xxx.xxxxx

average power: xxx.xxxxx

average magnitude: xxx.xxxxx

Utterance of the word NINE

……

Utterance of the word ZERO

……

**Write a copy of the result** presented in the command window **to a file** name **voiceSignature.txt**.

The program should also **save all the computed result in a .mat file**, so that other MATLAB programs may utilize the results.

**(Hint)**

* myArray = **wavread**(‘eight’) % to save data in eight.wav to myArray
* [myArray, Fs] = **audioread**(‘eight.wav’) % to save data from a .wav, .mp3, or .wma file to myArray at frequency Fs.
* You may record your own .wma files by going to Windows -> All Programs -> Accessories -> Sound Recorder, and click on the “Record” button (require sound card and mic).

Write another MATLAB program (midterm1b.m) to allow users to input a .wav file name. Compute the mean, standard deviation, variance, average power, and average magnitude of the .wav file input by the user.

Load all the variables saved by the first program to this program.

Generate three (3) Figure Windows. In the first figure, use the subplot function to plot the utterance of eight on top and the utterance from the user input file at the bottom. In the second figure, use the subplot function to plot the utterance of nine on top and the utterance from the user input file at the bottom. In the third figure, use the subplot function to plot the utterance of zero on top and the utterance from the user input file at the bottom. The graphs should be titled and labeled clearly.

Add the following tolerances to the statistical measurements saved by the first program. Use the lower and upper tolerance limits from the voice of “eight”, “nine”, and “zero” to determine if the .wav file (input by the user) contains the voice of an “eight”, a “nine”, a “zero” or none of the above.

For mean, accept **+/- 5%** of the calculated mean value.

For standard deviation, accept **+/- 10%** of the calculated standard deviation.

For variance, accept **+/- 8%** of the calculated variance.

For average power, accept **+/- 8%** of the calculated average power.

For average magnitude: accept **+/- 10%** of the calculated average magnitude.

The output of the MATLAB program should be presented to the command window in the following format:

Utterance comparison result for EIGHT

mean: xxx.xxxxx is within xxx.xxxxx and xxx.xxxxx

standard deviation: xxx.xxxxx is within xxx.xxxxx and xxx.xxxxx

variance: xxx.xxxxx is within xxx.xxxxx and xxx.xxxxx

average power: xxx.xxxxx is within xxx.xxxxx and xxx.xxxxx

average magnitude: xxx.xxxxx is within xxx.xxxxx and xxx.xxxxx

The voice in \*\*\*.wav is an EIGHT!

Utterance comparison result for NINE

mean: xxx.xxxxx is within xxx.xxxxx and xxx.xxxxx

standard deviation: xxx.xxxxx is NOT within xxx.xxxxx and xxx.xxxxx

…

The voice in \*\*\*.wav is NOT a NINE!

Utterance comparison result for ZERO

mean: xxx.xxxxx is within xxx.xxxxx and xxx.xxxxx

standard deviation: xxx.xxxxx is NOT within xxx.xxxxx and xxx.xxxxx

…

The voice in \*\*\*.wav is NOT a ZERO!

Your program should do the actual mathematical comparison of the acceptance criteria for “eight”, “nine”, and “zero”. If a voice file for the digit “one” is input by the user, your program should recognize that it is neither an “eight”, nor “nine”, or “zero”.

If the voice file received is the voice of “eight”, “nine”, or “zero”, **write the numeric digit to “voiceReceived.dat” file in the ASCII format**, so that other programs may read the input from this file and perform other actions.

**(Hint)**

* **load** myFileName.mat % to load the variables and values stored in myFileName.mat
* filename = **input** (‘Give me a file:’, ‘s’) % take a string as input and save the string as fileName
* Use variables to track comparison results.