

Software Documentation

The Digit Speech Recognition Software in MATLAB consists of the following .m and .mat files with respective functions:

- **realtime_tester.m**

This is the main Speech Recognizer program, which utilizes the Training samples in the 'train' folder and recognizes live speech input in realtime. The speech should contain the voice of a single digit in a duration of about 5 seconds @ 8000Hz Sampling rate.

- **params.mat**

This MATLAB Data File stores the variables containing information about the Number of Sample Data available for each Digit in Testing and Training folder facilitating required looping code in other .m files. It also stores the Average MFC Coefficients of the Train Data, so that they are ready for use by the Recognizer program.

- **add_train_or_test_data.m**

This program is for adding more Training Samples or Testing Samples into the Speech Database of the Recognizer. Once a test/train speech is recorded, this program generates the MFCCs and stores them along with original time domain sound amplitudes. It also calculates Average Train MFCCs dynamically and updates parameters in *params.mat* so that the Recognizer can use them readily.

- **performance_evaluator.m**

This program reports the Accuracies on the testing data in 'test' folder based on the learning of the Recognizer on the training data in 'train' folder. Since it has to do Dynamic Time Warping (DTW) of each test sample with every train sample in ~500 dimensional space, it takes around 4 min to complete this evaluation process.

- **params_initialization.m**

This .m file is to be used to CLEAR all the parameters in the *params.mat* containing information on the Number of Samples for Test/Train, and Average MFCCs w.r.t Train data. So, this clears the total Train/Test Database. After this is run, please delete the Speech Samples inside Test/Train folders manually.

- **player_train_or_test_audio.m**

This .m file allows the playing of the recorded Train and Test Audio Signals of Digits 1, 2 and 3.

Digit Speech Recognition Project

Submitted for EE322M Course by:

VINEETH S. BHASKARA (120121007)