Charlotte McGinn

EE 102

# Speech Classification Report

Task 1:

For Task 1, I wrote a function called "audioConvert.m" which converts the specified index in the trainAudioRecords voice array to a .wav audio file. I called this function on the first two rows on the trainAudioRecords array and discovered that the first voice was female and the second voice was male. Comparing this against the labels in the trainAudioLabels array, I concluded that "1" corresponds to a female voice and "0" corresponds to a male voice.

Task 2:

For Task 2, I utilized the svmtrain to train an SVM model using the training data provided, and then classified the provided testing data with svmclassify to test the accuracy of the SVM model. I then calculated the percent error between the SVM classifications and the true test labels provided, and got a percent accuracy of 56.70%. While the SVM classifier is correct more than 50% of the time, it has a reasonably large rate of mislabeling data (43.30%).

Task 3:

a) Voice 1
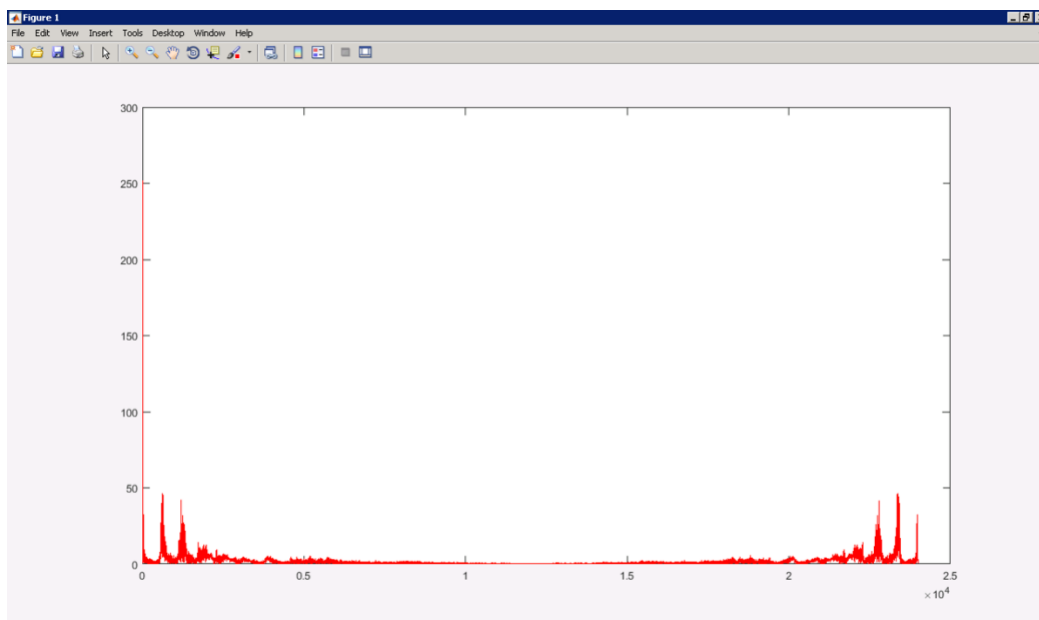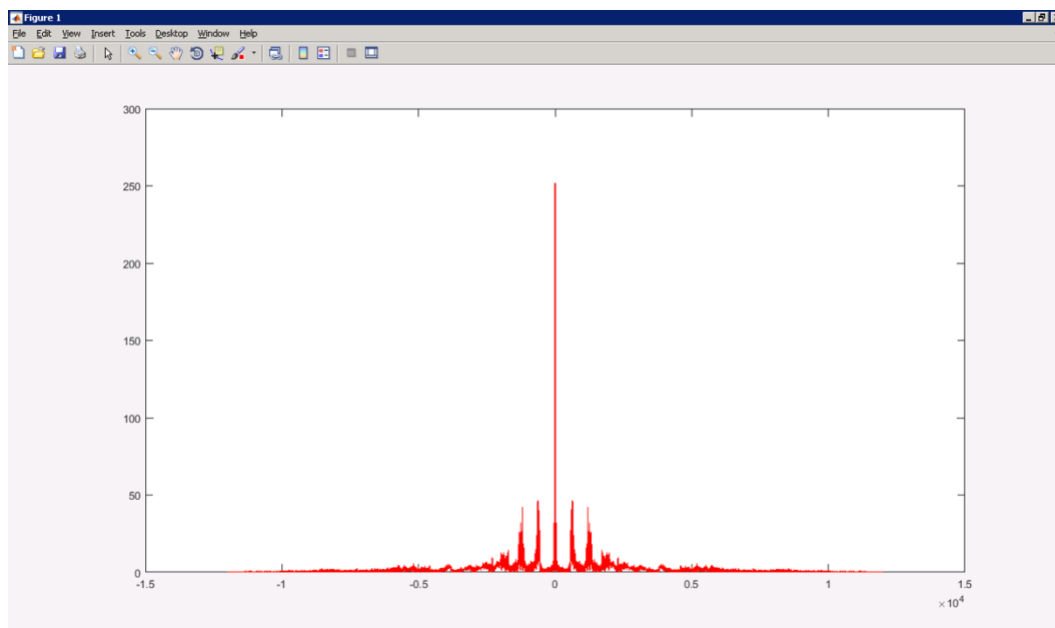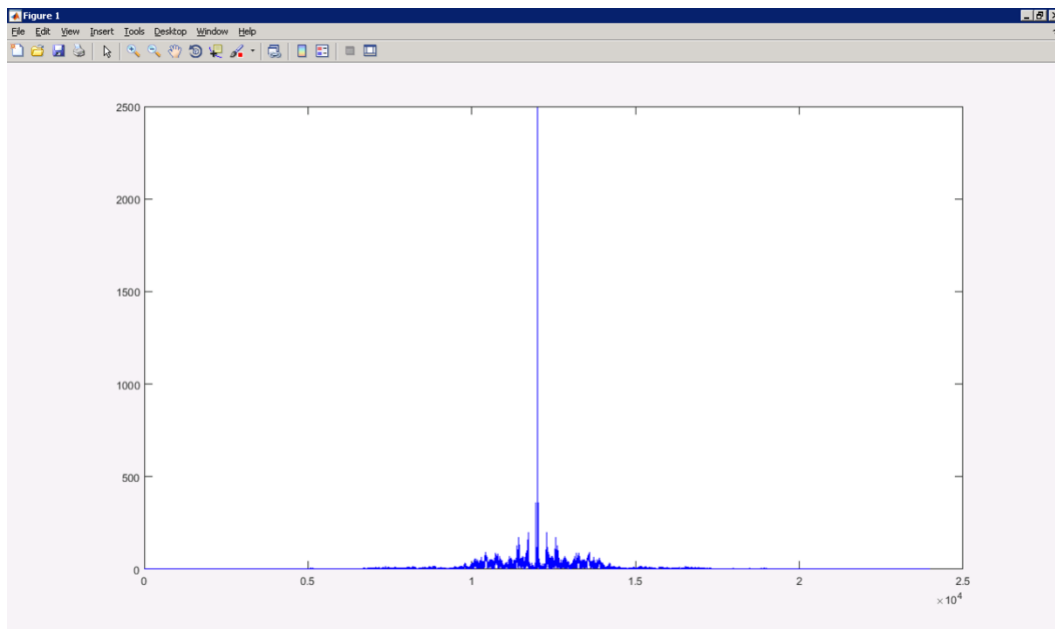
Voice 2
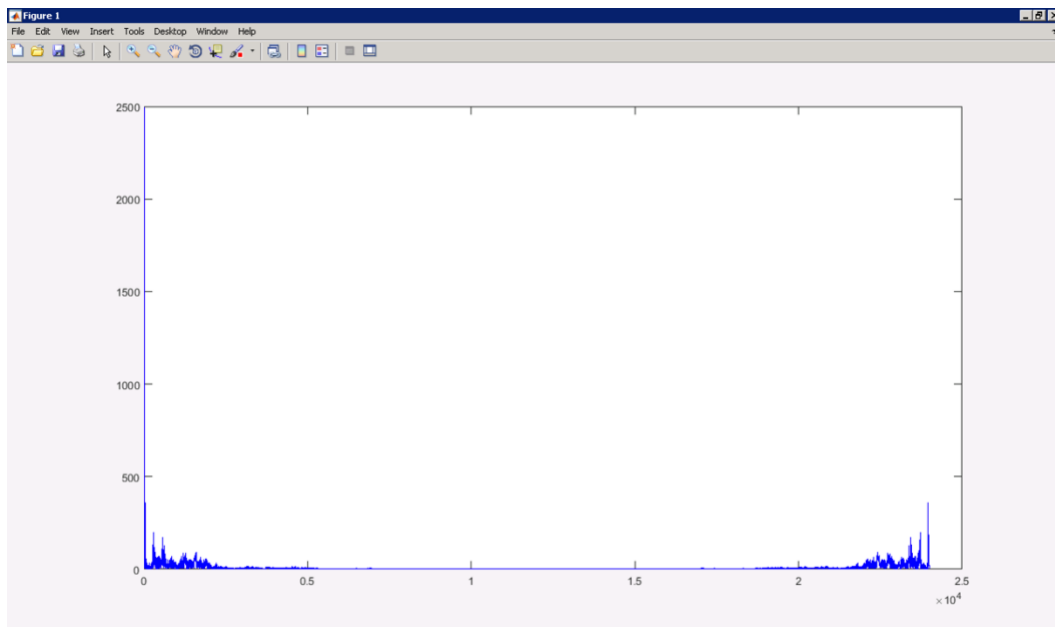
Voice 1 appears to have a higher frequency magnitude on average compared to Voice 2. This would make sense, since we know that Voice 1 is a female voice and Voice 2 is a male voice, and female voices tend to be a higher pitch than male voices. The voices appear symmetric about t = 12,000 (the midpoint of the function).

b)   Voice 1

Voice 2

Task 4:

Using the power spectrum, I obtained a percent accuracy of 99.23%. This is a 42.53% increase from the original SVM tested with the raw audio data.

Task 5:

For this task, I decided to use the average frequency of the voice to determine gender. This is because, on average female voices tend to have a higher pitch than male voices, so I thought this would be a good indicator. I took the Fourier transform of each audio file and calculated the average frequency of each file, and used that to train my model. My model performed with an accuracy of 99.43%, which is 0.20% better than using the power spectrum to determine gender.