جامعة بيرزيت

BIRZEIT UNIVERSITY

Faculty of Engineering & Technology
Electrical & Computer Engineering Department

DIGITAL SIGNAL PROCESSING – ENCS4310

Matlab Assignment

Gender Recognition and Classification of Speech Signal System

Prepared by:

Hatem Hussein 1200894

Instructor: Dr. Ashraf Al-Rimawi

Section: 1

Date: 24-01-2024

Abstract/Objectives

This assignment focuses on the development and demonstration of a gender recognition and classification system for speech signals using time and frequency domains. The project involves utilizing a diverse dataset comprising at least 20 recordings for both male and female speakers, considering variations in age and environment. All recordings must feature the word "Zero" and share a common sampling frequency of 44,100 Hz. The tasks include splitting recordings into testing and training files, evaluating correlation, energy, zero-crossing count, and power spectral density. The system's accuracy is assessed by applying these techniques on different testing files.

Table of Contents

List of figures

Performance

Data Used

I used multiple voices for each training file (male/female) from different ages, countries, and in different environment. Moreover, other different voices have been chosen as test files for the system to see the accuracy of the system. The sampling frequency was set to be 44100Hz for all of the training files and testing files.

Notice that I chose multiple male training voices to be soft, so whenever a soft male test voice file applied to the system the recognition and classification be correct or got a high value of accuracy. On the other hand, I chose multiple female training voices to be sharp, so whenever a sharp female test voice file applied to the system the recognition and classification be correct or got a high value of accuracy.

The below is the system features of the males training-data used:

```
******************************************************************************

        THE RESULTS OF TRAINING MALES FILES

******************************************************************************
The ZCR of male is: 1210.3500   287.7250    509.3000
The CORRELATION of male is: 0.0003  0.0000  0.0002
The ENERGY of male is:  0.4348
The PSD of male is:      2.078120798887335e-05


******************************************************************************
```

Figure 1: Males training-data extracted features

The below is the system features of the females training-data used:

```
******************************************************************************

        THE RESULTS OF TRAINING FEMALES FILES

******************************************************************************
The ZCR of female is:   2384.9000   306.2000    723.8000
The CORRELATION of female is:   0.0000  0.0000  0.0033
The ENERGY of female is:    0.5269
The PSD of female is:       3.802493153816850e-05


******************************************************************************
```

Figure 2: Females training-data extracted features

Males Testing

The system was good in dealing with the time-domain features extracted from the males testing files. On the other hand, the system was not very accurate in dealing with the PSD. The system got an accuracy of 60% for the males testing files applied on the system with the extracted features got from the trained data shown previously. The percentage is good due to the very different voices files I used and I trained my system on. The below is the output of the full system applied on the five males testing voices:

```
******************************************************************************

        NOW LETS SEE TESTING RESULTS FOR MALES

******************************************************************************

______________________________________________________________________
Test file [male] #1 classified as male based on ZCR
Test file [male] #1 classified as female based on CORR
Test file [male] #1 classified as male based on ENERGY
Test file [male] #1 classified as male based on ZCR with ENERGY
Test file [male] #1 classified as female based on PSD

______________________________________________________________________
Test file [male] #2 classified as male based on ZCR
Test file [male] #2 classified as female based on CORR
Test file [male] #2 classified as female based on ENERGY
Test file [male] #2 classified as male based on ZCR with ENERGY
Test file [male] #2 classified as female based on PSD

______________________________________________________________________
Test file [male] #3 classified as male based on ZCR
Test file [male] #3 classified as male based on CORR
Test file [male] #3 classified as male based on ENERGY
Test file [male] #3 classified as male based on ZCR with ENERGY
Test file [male] #3 classified as female based on PSD
```

Figure 3: Males testing files results – part one

```
______________________________________________________________________
Test file [male] #4 classified as male based on ZCR
Test file [male] #4 classified as male based on CORR
Test file [male] #4 classified as male based on ENERGY
Test file [male] #4 classified as male based on ZCR with ENERGY
Test file [male] #4 classified as female based on PSD

______________________________________________________________________
Test file [male] #5 classified as female based on ZCR
Test file [male] #5 classified as male based on CORR
Test file [male] #5 classified as male based on ENERGY
Test file [male] #5 classified as female based on ZCR with ENERGY
Test file [male] #5 classified as female based on PSD

################################################################################

        THE ACCURACY OF TESTING MALE FILES
[============>       ] 60.00%

################################################################################
```

Figure 4: Males testing files results – part two

Females Testing

The system got better results than the males in dealing with the time-domain features extracted from the females testing files. Moreover, the system was very accurate in dealing with the PSD and got a highly percentage of accuracy. The system got an accuracy of 76% for the females testing files applied on the system with the extracted features got from the trained data shown previously. The percentage is good due to the very different voices files I used and I trained my system on. The below is the output of the full system applied on the five females testing voices:

```
********************************************************************************

        NOW LETS SEE TESTING RESULTS FOR FEMALES

********************************************************************************

________________________________________________________________________
Test file [female] #1 classified as female based on ZCR
Test file [female] #1 classified as female based on CORR
Test file [female] #1 classified as female based on ENERGY
Test file [female] #1 classified as female based on ZCR with ENERGY
Test file [female] #1 classified as female based on PSD

________________________________________________________________________
Test file [female] #2 classified as male based on ZCR
Test file [female] #2 classified as female based on CORR
Test file [female] #2 classified as male based on ENERGY
Test file [female] #2 classified as male based on ZCR with ENERGY
Test file [female] #2 classified as female based on PSD

________________________________________________________________________
Test file [female] #3 classified as male based on ZCR
Test file [female] #3 classified as female based on CORR
Test file [female] #3 classified as female based on ENERGY
Test file [female] #3 classified as female based on ZCR with ENERGY
Test file [female] #3 classified as female based on PSD
```

Figure 5: Females testing files results – part one

```
________________________________________________________________________
Test file [female] #4 classified as female based on ZCR
Test file [female] #4 classified as female based on CORR
Test file [female] #4 classified as male based on ENERGY
Test file [female] #4 classified as female based on ZCR with ENERGY
Test file [female] #4 classified as female based on PSD

________________________________________________________________________
Test file [female] #5 classified as female based on ZCR
Test file [female] #5 classified as female based on CORR
Test file [female] #5 classified as female based on ENERGY
Test file [female] #5 classified as male based on ZCR with ENERGY
Test file [female] #5 classified as female based on PSD

################################################################################

        THE ACCURACY OF TESTING FEMALE FILES
[===============>    ] 76.00%

################################################################################
```

Figure 6: Females testing files results – part two

GUI application

I built the below GUI application that enables the user to record a voice and then show the voice signal in time and frequency domains. Moreover, the application enables the user to test the recorded voice (Male/Female). The results shown are based on the system that has been built already and based on the training data used in the full-system tested.

I recorded my voice using the GUI application. I got the below Time-domain waveform:

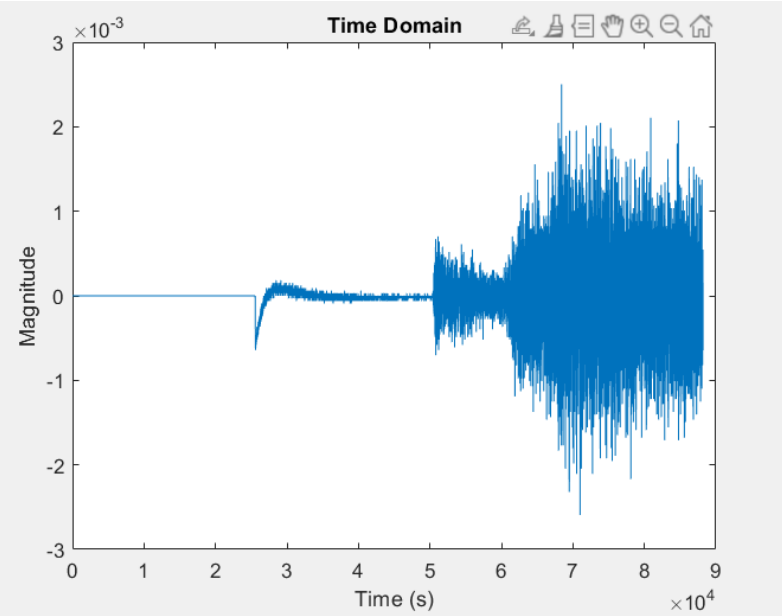

Figure 7: GUI application - Time domain tested voice signal

The frequency-domain waveform was as shown below:

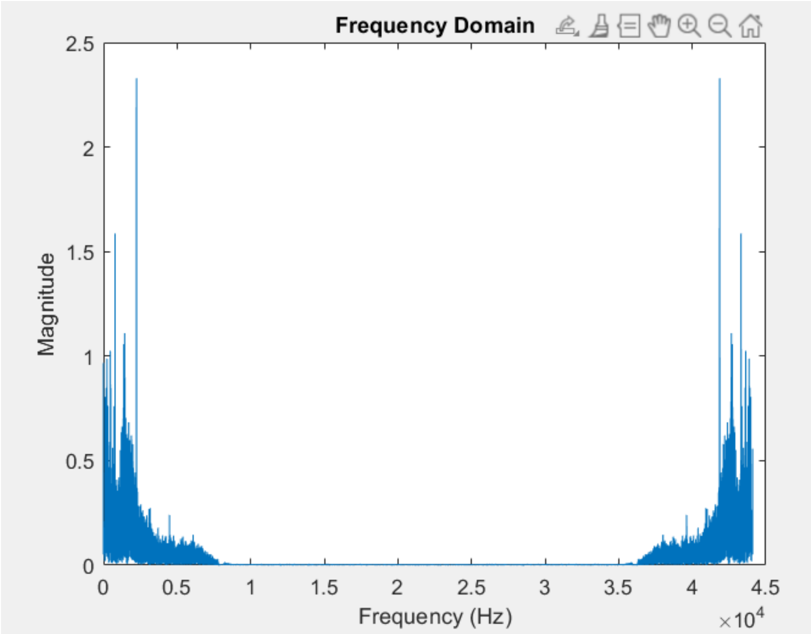

Figure 8: GUI application - Frequency domain tested voice signal

When I tested the voice recorded, the below results has been shown on the text areas illustrating what each extracted feature has classified the voice to:

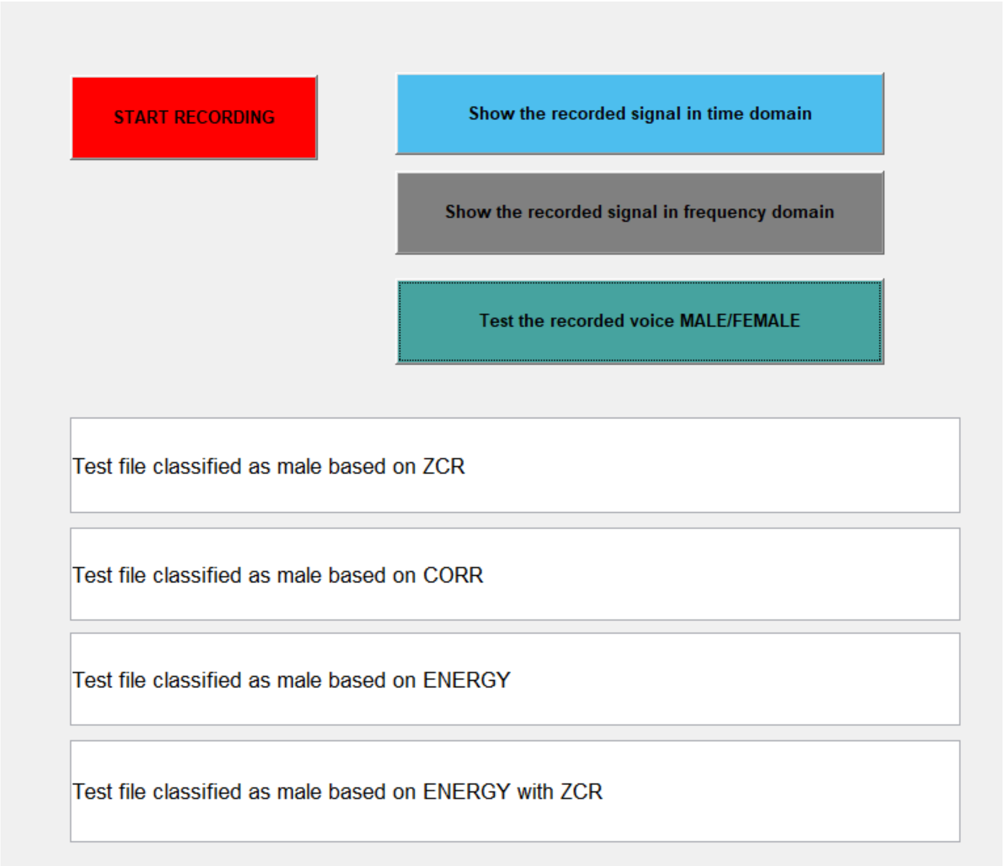

Figure 9: GUI application - tested voice results

NOTICE the below link is the screen recording for the results of my voice being tested:

https://drive.google.com/file/d/1aKNya4OYUQKsAvPQ9Ippe9Bwkide1y0s/view?usp=sharing

Improvements

The gender recognition and classification system can be improved by incorporating additional speech features such as formants and expanding the dataset to include recordings in different languages, ages, and environments. This comprehensive dataset will enhance the system's ability to recognize and classify gender accurately. Integration of machine learning techniques will improve adaptability, while the implementation of dynamic thresholds based on individual speaker characteristics will further refine the system. Additionally, contextual analysis by considering the word's context or emotional tone can provide a more nuanced understanding. Enhancements in noise robustness and the incorporation of user feedback will contribute to continuous improvement. Extending the system to recognize words beyond "Zero" and refining the graphical user interface will make the tool more versatile and user-friendly for speech signal analysis across various words and languages.

Conclusion

In conclusion, a gender recognition and classification system for speech signals was constructed using a dataset comprising a minimum of 20 recordings each for female and male speakers. The recordings considered various factors such as different ages and environments to ensure diversity. The chosen sentence for recording was "Zero," and all audio files maintained a consistent sampling frequency of 44100 Hz. The tasks included splitting the recording files into testing and training sets, followed by the evaluation of correlation, energy, and zero-crossing count. Additionally, the power spectral density was assessed. The system's accuracy was determined by applying these techniques, providing valuable insights into its performance. Moreover, when presented with a speech signal saying "Zero," the system plotted the recorded signal in both time and frequency domains. The final step involved matching the record with the training data to classify the speaker as either male or female. Through these tasks, the project successfully established a foundation for gender recognition in speech signals, employing a comprehensive dataset and employing various signal processing techniques to assess system accuracy.