Tariq Odeh
1190699
Birzeit University
1190699@student.birzeit.edu,

Qays Safa
1190880
Birzeit University
1190880@student.birzeit.edu,

Mahmoud Samara
1191602
Birzeit University
1191602@student.birzeit.edu

Abstract

In this report, we will discuss our Digital Signal Processing Course project. The aim in this project is to recongnize the gender of any speaker depend on the frequency of the input audio file that is .wav files. Using main method that is the fundamental frequency which can be estimated from the short frames (20-30ms) using auto-correlation. The report is simply discusses the Data, Evaluation Criteria, approach techniques we used then discusses the result of our work.

1. Introduction

The main idea of this project is to implement, test and recongnize the gender of any speaker depend on the frequency of the input audio file that is .wav files. Our main testing will depend on the autocorrelation result since Threshold frequency (Fth) of a male voice lies between 85-155 Hz whereas for a female, it lies between 165 to 255 Hz. Moreover, we will Find the first max peak of the autocorrelation values and implement and test our full system. Simply, your system should take an audio file (.wav) as an input and recognizes the speaker gender.

2. Problem Specification

In this project the basic problem was how we can divide the samples into short frames because without this step we can't continue our work and find the auto correlation. Also, we need to Find autocorrelation values of each short frame, for all k. After that, we came to the most important part in our problem that is: Convert the estimated pitch period P into seconds and then find the corresponding pitch frequency F0 because this value of F0 will lead us to know the gender of our speaker.

2.1. Equations

In our project we will have main and important equation that is how to find the autocorrelation since I said before our road to the goal and to recognize the gender of the speaker is to know the value of F0 that cannot happen if we will not find the autocorrelation for each frame.

The autocorrelation equation is:

2.1.1. Autocorrelation equation:

$$R(k) = \frac{1}{N}\sum_{n=0}^{N-1} s[n]s[n-k]$$

Where, R(k) is the autocorrelation function at k. s[n] is the short frame samples with length N samples. Clearly, R(k=0) would be equal to the average energy of the signal s[n] over the N sample frame.If s[n] is perfectly periodic with a period of P samples, then s[n+P] = s[n]. Therfore, R(k=P)=R(k=0) = average Energy.

3. Data

The project was designed to recognize the gender of the speaker sound plus plot the audio signal. This project take input audio file and use the autocorrelation value to get the results we need. Moreover, in our project the user can enter the name of the input file but of course it must be in the same file of the matlab folder.

The data is taken from user as string so he has to enter the name of the input audio file (.wav) and depend on the input file the output will appear. The data values that will be printed are: sampling frequency (Fs) that will be calculated using the audioread function in matlab, also we will show full information of the input file. Finally, we will find the corresponding pitch frequency F0 and use it as our references to recognize the gender.

4. Evaluation Criteria

To test the program if it is working right or not and finding the accuracy of the expected output, various strings were used and to get correct output with full details and plots. For example, the name of the input file and what full information of it were used to find and solve our problem. Project performance was measured based on the accuracy score so after testing a large set of audios by what user will enter so the project must have an error rate of 0% and an accuracy of 100%.

5. Approach

In the beginning, it is important to build the project smoothly and appropriately with the user, so we started by asking the user to enter the name of the file he wants to read and identification, with an emphasis on it to be in the form shown formula ["data#.wav"] where # is the number of files. After entering the file name, we go to the second step, which is reading the file entered by the user, where we will first print some details about the entered file such as location, SamplesPerFrame, ReadRange, OutputDataType, PlayCount this inforamtions will appear after we used dsp.AudioFileReader function in Matlab . After that, we will read the sound and find the sampling frequency (Fs) for the audio, and then we will work on plot the shape of the sound to show how the nature of the sound and what we're going to do with it.

```
%-----------------------------------------------------------------
%Print the Menu

    fprintf(' +----------------------------------------------------------------------------------------+\n');
    fprintf(' |        [Welcome in this system to identification of males and females from their speech analysis]        |\n');
    fprintf(' +----------------------------------------------------------------------------------------+\n');
    fprintf(' |  [Please use the next formula [''data#.wav''] where # is the number of file when enter the name of file]  |\n');
    fprintf(' +----------------------------------------------------------------------------------------+\n');

%-----------------------------------------------------------------
```

Figure 1: The menu that appear to user

```
%-----------------------------------------------------------------
%Read an audio file.
%Find its sampling frequency (Fs).

    file = input('       #-->Please use name of file:   ');
    fileReader = dsp.AudioFileReader(file);
    [A,Fs] = audioread(file);
    figure;
    plot(A)          %plot audio file
    title('Audio');
    xlabel('Time');
    ylabel('Amplitude');
    grid

%-----------------------------------------------------------------
```

Figure 2: Read and plot the audio file

In the next step, we will work on calculating the important values that we need to reach the identification of the voice, male or female, so we will find the values of each maximum speech, frames, overlap, estimates of the fundamental frequency.

```
%-----------------------------------------------------------------
%Find the values.

    m500 = Fs/500;                          %maximum speech Fx at 500Hz
    m50 = Fs/50;                            %maximum speech Fx at 50Hz
    windowLength = round(0.03*Fs);          %short frames 30ms
    overlapLength = round(0.01*Fs);         %overlap 10ms
    F0 = pitch(A,Fs);                       %estimates of the fundamental frequency

%-----------------------------------------------------------------
```

Figure 3: Values needed for the recognition operation

In the next stage, we will work on the most important part of our work so the following steps our main road to solve the project, so we will work on divided the samples into short frames and then find autocorrelation values of each short frame. Next, finding the max peak of the autocorrelation values and find Fx.

```
%-----------------------------------------------------------------
%Autocorrelation.

    AUCO = xcorr(A,m50);
    Delay = (-m50:m50)/Fs;
    figure;
    plot(Delay,AUCO);              %plot autocorrelation
    title('Autocorrelation');
    xlabel('Delay(s)');
    ylabel('Correlation Coefficient');
    grid
    AUCO = AUCO(m50 + 1 : 2 * m50 + 1);     %region corresponding to positive delays
    [AUCO_max, Tx] = max(AUCO(m500:m50));
    Fx = Fs/(m500 + Tx - 1);

%-----------------------------------------------------------------
```

Figure 4: Performing autocorrelation

Here we found the Fth then we checked If Fx is above a threshold then the voice is of a male, otherwise it is of a female and the threshold frequency (Fth) of a male voice lies between 85-155 Hz whereas for a female, it lies between 165 to 255 Hz.

```
%-----------------------------------------------------------------
%Threshold frequency (Fth) of a male voice lies between 85-155 Hz
%Threshold frequency (Fth) of a female voice lies between 165 to 255 Hz

    Fth = ((155 + 165) / 2);        %Threshold frequency (Fth)
    Fmax = 255;
    Fmin = 85;

    if( Fx >= Fmin && Fx < Fth )
          fprintf('--->[It is a male voice!]\n');

    elseif( Fx <= Fmax && Fx > Fth )
          fprintf('--->[It is a female voice!]\n');

    else
          fprintf('Unable to identification, please make sure the sound is clear!!\n');
    end

%-----------------------------------------------------------------
```

Figure 5: If statements to recognize the gender of the speaker

6. Result and Analysis

After test a large group of input audio files by using pitch function and readaudio function then run the code using the main methods, the result is completely identical and the same as the theoretical part and what the gender supposed to be. In addition, the accuracy is very excellent, so thereis no error in this project.

Our result will show the following things:

First the full information of the input audio file as the following figure.

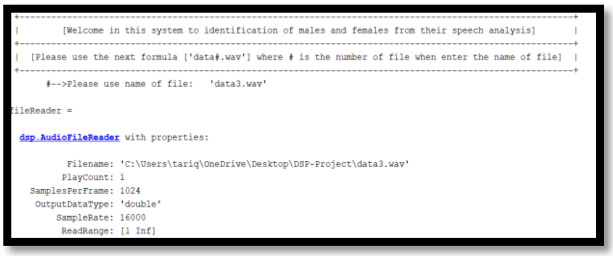


Figure 6: full information of the input audio file

After that, the Audio signal plot of the input audio file will appear so the X axis is: Time and the Y axis is: Amplitude.

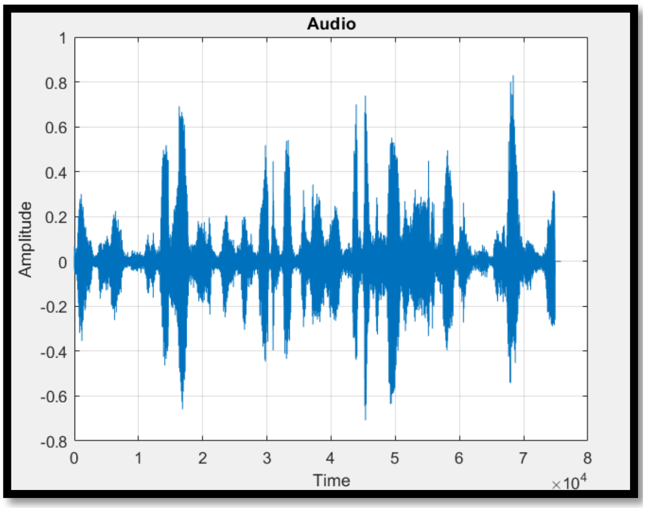


Figure 7: Audio signal plot for male

After that, the autocorrelation plot will appear so the X axis is: Delay(s) and the Y axis is: Correlation Coefficient.

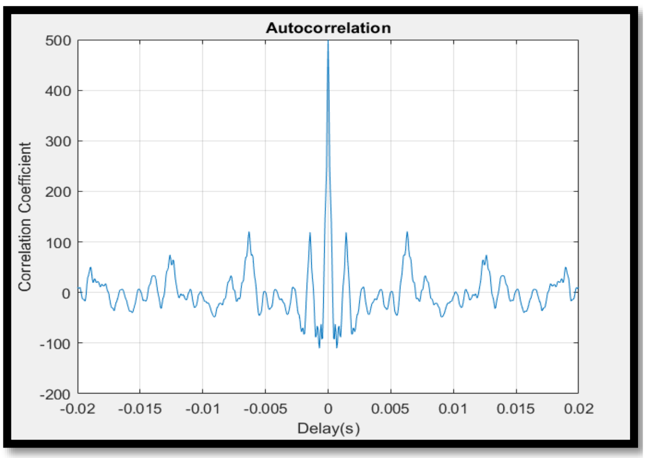


Figure 8: Autocorrelation plot for male

Finally, the gender of the speaker will be printed and in the workspace part the values of our variables will be calculated.

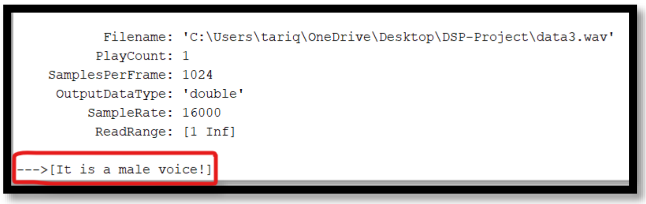
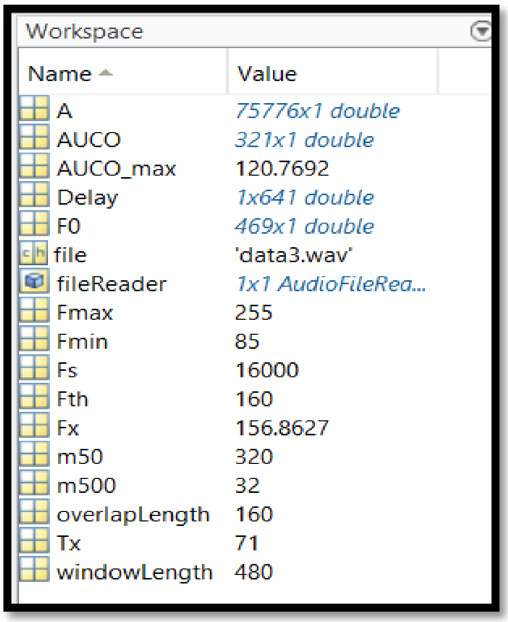


Figure 9: Male input audio file result



Workspace

Name	Value
A	75776x1 double
AUCO	321x1 double
AUCO_max	120.7692
Delay	1x641 double
F0	469x1 double
file	'data3.wav'
fileReader	1x1 AudioFileRea...
Fmax	255
Fmin	85
Fs	16000
Fth	160
Fx	156.8627
m50	320
m500	32
overlapLength	160
Tx	71
windowLength	480

Figure 10: Results of the different variables

Also, if we repeat the steps for another file for example for female the following result will appear

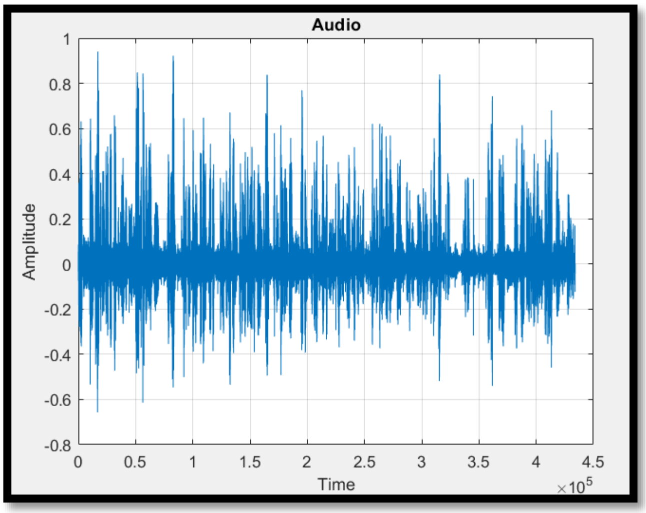


Figure 11: Audio signal plot for female

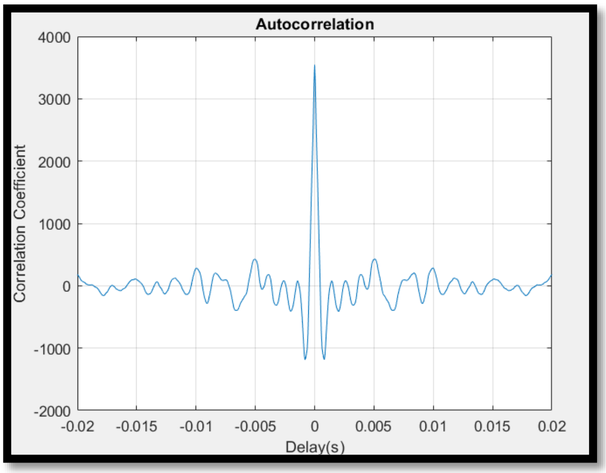


Figure 12: Autocorrelation plot for female

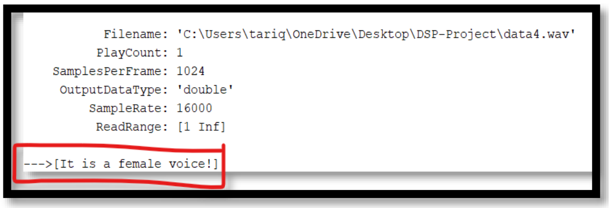


Figure 13: Female input audio file result

7. Development

In our case, we get a lot of noise in the audio (.wav) file when we read it that appears in the plot of it, but we can improve our approach and the ideas we used in our project. So for that first of all we made the project general that the user he has the freedom to test any input file he want this will help the user to feel comfortable to make many tests. By the way, we used the (audioread) function in matlab and got the sound without any noise and it calculated the sample frequency immediately that help us to keep going in the project faster.

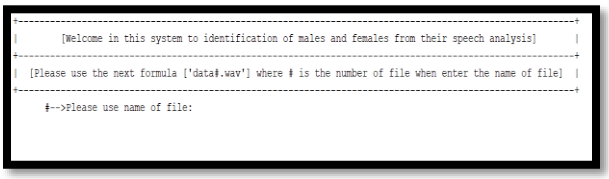


Figure 14: Freedom to user to test any input audio file

8. Conclusion

In this project we learned how to use Matlab program and knew new functions in it, these functions helped us in our project. As we want clearly to know the gender of the speaker and clearly without exposing the transmitted signals to noise or distortion and then get an ideal and clear result, so this project can be used to send a message as a sound wave file through any network and know the gender of the speaker of it. According to what was noted when searching for the importance of this method (pitch and autocorrelation) we found that Autocorrelation gives information about the trend of a set of historical data so that it can be useful in the technical analysis for the equity market and for the Pitch detection it is crucial task in singing voice separation also. Pitch detection also play important role in Musical information retrieval, Identification of the singer and in lyric recognition. Pitch can identify gender of singing voice. It means it can identify the gender of the singer. Finally, we want to say that it was a wonderful project to introduce us to how to reach a result from encrypting these messages, as well as how the messages reach the other side Without any distortion on it and the method of its arrival, and this is due to the importance of communications between the world.

9. References

1. Digital Signal Processing textbook
2. https://www.mathworks.com/help/matlab/ref/audioread.html
3. https://github.com/DAGsHub/audio-datasets
4. https://xpertsvision.wordpress.com/2015/12/04/gender-recognition-by-voice-analysis/
5. https://wiki.aalto.fi/pages/viewpage.action?pageId=149890776

10. Appendix

```
clc
clear
close all
%---------------------------------------------------------------
%Print the Menu
   fprintf(' +-----------------------------------------------------------------------------------------------------------+\n');
   fprintf(' |       [Welcome in this system to identification of males and females from their speech analysis]       |\n');
   fprintf(' +-----------------------------------------------------------------------------------------------------------+\n');
   fprintf(' |  [Please use the next formula ["data#.wav"] where # is the number of file when enter the name of file]  |\n');
   fprintf(' +-----------------------------------------------------------------------------------------------------------+\n');

%---------------------------------------------------------------
%Read an audio file.
%Find its sampling frequency (Fs).

   file = input('      #-->Please use name of file:   ');
   fileReader = dsp.AudioFileReader(file)
   [A,Fs] = audioread(file);
   figure;
   plot(A)       %plot audio file
```

```
    title('Audio');
    xlabel('Time');
    ylabel('Amplitude');
    grid
%--------------------------------------------------------------
----
%Find the values.

    m500 = Fs/500;                    %maximum speech
Fx at 500Hz
    m50 = Fs/50;                      %maximum speech
Fx at 50Hz
    windowLength = round(0.03*Fs);      %short
frames 30ms
    overlapLength = round(0.01*Fs);     %overlap
10ms
    F0 = pitch(A,Fs);                 %estimates of the
fundamental frequency

%--------------------------------------------------------------
----
%Autocorrelation.

    AUCO = xcorr(A,m50);
    Delay = (-m50:m50)/Fs;
    figure;
    plot(Delay,AUCO);           %plot autocorrelation
    title('Autocorrelation');
    xlabel('Delay(s)');
    ylabel('Correlation Coefficient');
    grid
    AUCO = AUCO(m50 + 1 : 2 * m50 + 1);
%region corresponding to positive delays
    [AUCO_max, Tx] = max(AUCO(m500:m50));
    Fx = Fs/(m500 + Tx - 1);

%--------------------------------------------------------------
----
%Threshold frequency (Fth) of a male voice lies
between 85-155 Hz
%Threshold frequency (Fth) of a female voice lies
between 165 to 255 Hz

    Fth = ((155 + 165) / 2);        %Threshold
frequency (Fth)
    Fmax = 255;
    Fmin = 85;

    if( Fx >= Fmin && Fx < Fth )
          fprintf('--->[It is a male voice!]\n');

    elseif( Fx <= Fmax && Fx > Fth )
          fprintf('--->[It is a female voice!]\n');

    else
          fprintf('Unable to identification, please make
sure the sound is clear!!\n');
    end
```