Digital Signal Processing Project

1190207 Maysam Khatib , 1190069 Majd Abubaha , 1190507 Khalid Addabe

Birzeit University

I. INTRODUCTION

In this project we are going to implement a program using MATLAB to determine different voice recordings whether it belongs to a male or a female (adults) from their speech analysis, By examining multiple aspects of the voice sample, this technique seeks to determine the speaker's gender.

II. PROPLEM SPECIFICATION

The aim of this project, is to imply our theoretical knowledge which we took in the course, to implement the program, we used a dataset contains both male and female voices, Using the correlation approach; we take the voices of numerous persons (both female and male) and analyze their sound waves to identify their frequencies.

If the frequencies of the voices are between 85 and 155, the owner of the voice is deemed male, and if the frequencies are Between 165 - 255, the owner of the voice is regarded female. Children's voice frequencies vary from 250 to 400, and if the frequency is not present within these ranges, it produces a notice indicating that it is out of range.

III. DATA

We obtained people's voices in two ways: the first was by having a number of individuals record their voices for us, and the second was through an internet dataset, in which we acquired a number of voices from various websites and utilized them at random to determine the speaker's gender.

IV. EVALUATION CRITERIA

We examined all of the noises, whether we recorded them or got them from the Internet, and the results were right in every case: if the speaker is male, the voice is male, and if the speaker is female, the voice is female.

We also experimented with people voices of all ages and genders, since it was not enough with just one voice we wanted to try a variety of them.

V. APPROACH

At first, we store the sounds in a file, and then randomly select one, MATLAB reads it and save the sampling frequency (Fs), which in all cases was 44.1 KHz, which is the most common audio sample rate and represents the CD format sampling frequency.

```
clear all
close all
clc

% Read a random audio file from the PathName with format .wav
PathName = 'C:\Users\HP\OneDrive\Documents\121-2\DSP\Project';
List  = dir(fullfile(PathName, '*.wav'));
Index = randi(length(List), 1);
File  = fullfile(PathName, List(Index).name);
[audio,fs]= audioread(File);
fileInfo = audioinfo(File);

% Play the audio
sound(audio, fs);
```

Then we used MATLAB to play the voice and display the wave, which represents the voice signal, and then we get the Maximum Speech Frequency at 500 Hz and 50 Hz, which we'll need later in the calculations, we've represented the maximum value of the autocorrelation amplitude at time 0 which is equaled to 1, We used the function xcorr() with a brief frame of 20 milliseconds to compute the autocorrelation of the audio signal, and we estimated the right time for the period of the signal that occurred so that we could display it appropriately.

```
% plot the audio
t = 0:seconds(1/fs):seconds(fileInfo.Duration);
t = t(1:end-1);
figure
plot(t, audio, 'r')
title('Audio Signal Plot')
xlim(seconds([0 0.02]))
ylabel('Amplitude');
xlabel('Time');
```

```
% Calculate the autocorrelation using xcorr() function and plot it
twomsec = fs/500; % 2 msec frame = the maximum speach at 500 Hz
shortFrame = fs/50; % 20 msec short frame = the maximum speach at 50 Hz
correlationResult = xcorr(audio, shortFrame, 'normalized');
% determine the period of the plot to be the same as the correlation result
t2 = (-shortFrame:shortFrame)/fs;
figure
plot(t2, correlationResult,'k');
title('The Autocorrelation Plot');
xlabel('Delay (sec)');
ylabel('Correlation Coefficient');
```

To compute the frequency, first, we took the positive component of the autocorrelation, then took the maximum value of the correlation to calculate the Tx, and then got the frequency product using the equation:

Voice freq = fs / (2ms + Tx – 1).

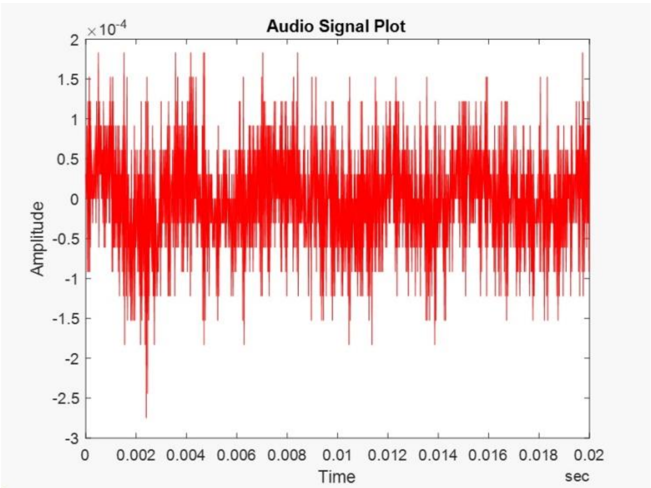
Finally, we matched the obtained frequency to the ranges to determine the speaker's gender, and we printed the result for the speaker's gender as well as the frequency.

```
% take the positive area of the correlation result
correlationResult = correlationResult(shortFrame + 1 : (2*shortFrame) + 1);
[correlationResultMax, tx] = max(correlationResult(twomsec:shortFrame));

% find the speaker voice frequency
voiceFreq = fs/(twomsec+tx-1);

% Determine the gender based on the speaker voice's frequency
if (voiceFreq >= 165) && (voiceFreq <= 255)
    disp('The Gender is: Female, with frequency:')
elseif (voiceFreq >= 85) && (voiceFreq <= 155 )
    disp('The Gender is: Male, with frequency:')
elseif (voiceFreq >= 255) && (voiceFreq <= 400)
    disp('The Gender Cannot be Recognized because the speaker is a child with frequency:')
else
    disp('The Gender Cannot be Recognized! The frequency is:')
end
disp(voiceFreq)
disp('First Max Peak:')
disp(correlationResultMax)
```

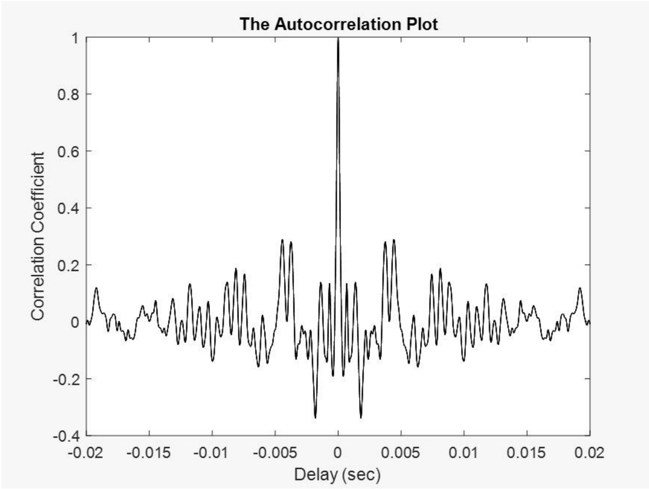
VI. Results and Analysis

When we run the code, the program chooses any voice randomly, going through the stages that we've explained previously in Approach, we've got these results

#1 Female Voice



The Previous plot represents the audio signal.

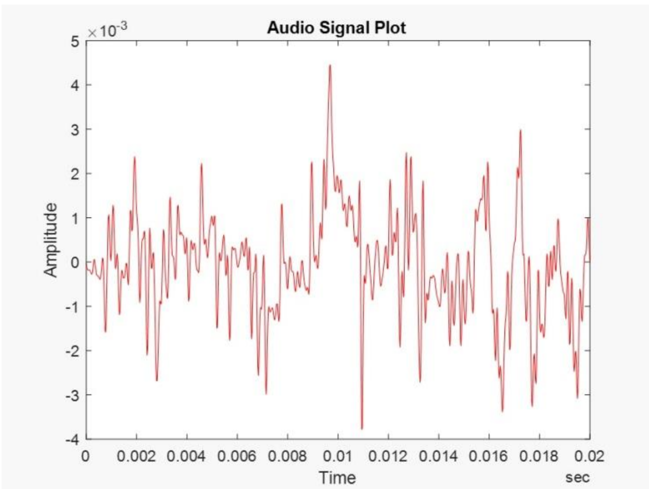


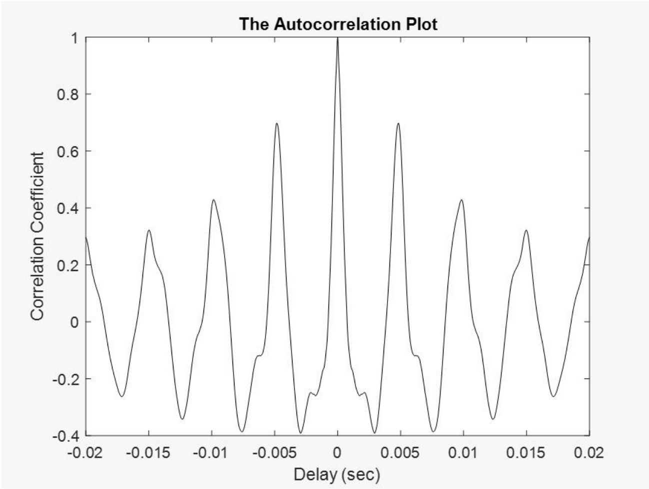
The previous plot represents the Autocorrelated audio signal.

```
The Gender is: Female, with frequency:
  224.7706

First Max Peak:
    0.2893

>>
```

The previous figure represents the result with voice frequency.

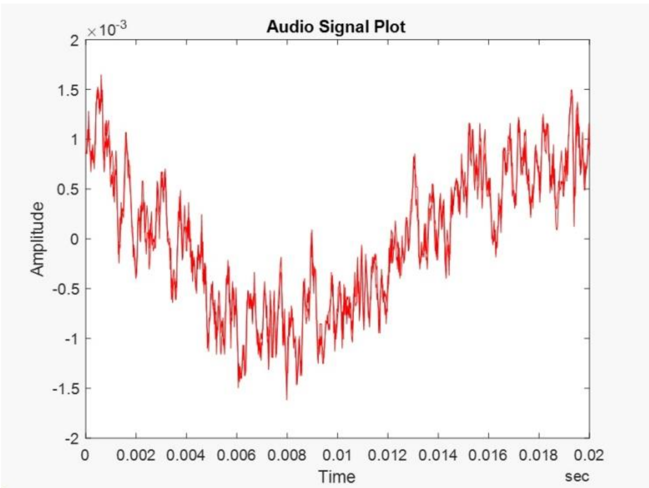
#2 Female Voice

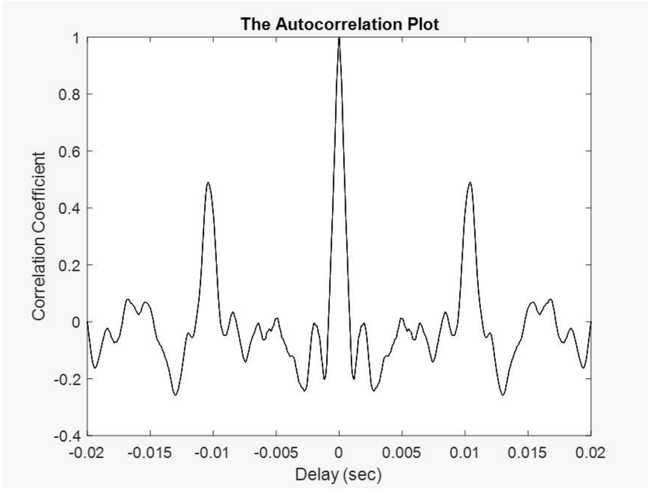





```
Command Window
The Gender is: Female, with frequency:
  206.0086

First Max Peak:
    0.6980

fx >>
```

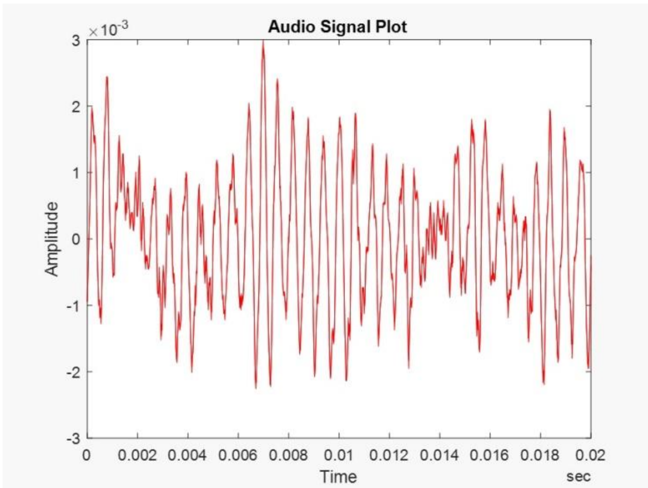
#3 Male Voice

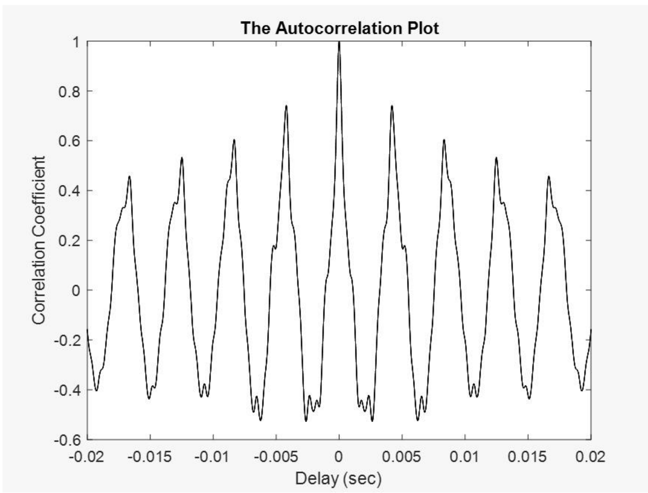


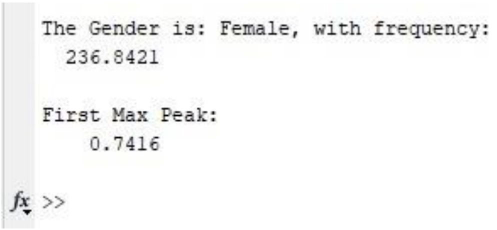



```
The Gender is: Male, with frequency:
   96.0366

First Max Peak:
    0.4898
```

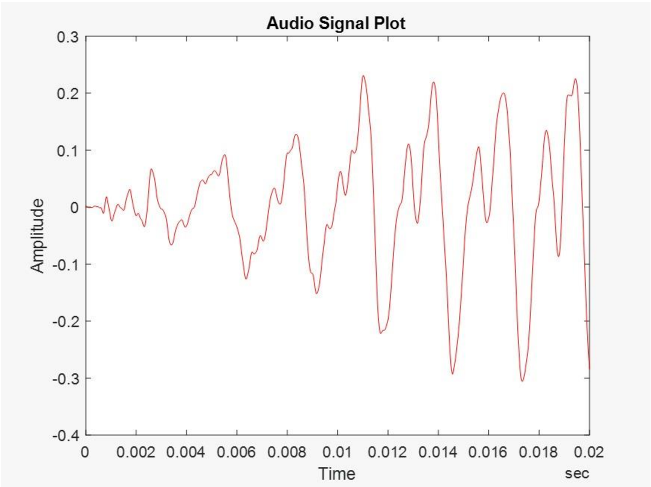
#4 Female Voice

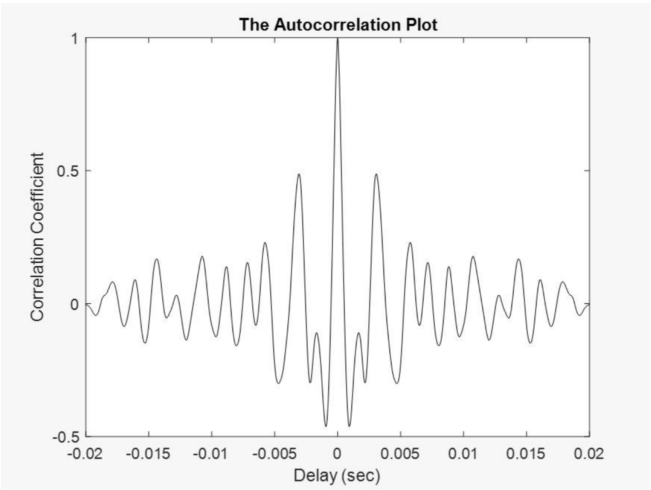


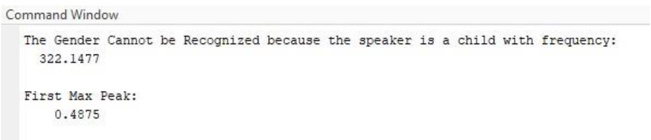




#4 Child Voice







VII. Development

The voices with some noise were occasionally incorrect; for example, if the speaker was male or female and there was some noise in their voice, the gender of the speaker would not be accurately detected. It is possible to alleviate this issue by removing the noise.

We can also utilize the DVD format to acquire greater quality and a sample frequency that is double that of the CD format (96 KHz), resulting in more accurate findings than utilizing the CD format.

VIII. Conclusions

We learned how to use Matlab to recognize sounds as male, female, or kid voice, and how to utilize the correlation function to achieve the above aim in this project.

We looked at a lot of voice samples and divided them into male, female, and kid sounds based on the frequencies that each gender has.

IX. Acknowledgements

We would like to thank our instructor Dr.Alhareth Zyoud (Birzeit University) for providing us enough theoretical information to achieve the level of understanding, which made it much easier for us to implement or knowledge in coding.

X. References

1. HTTPS://XPERTSVISION.WORDPRESS.COM/2015/12/04/GENDER-RECOGNITION-BY-VOICE-ANALYSIS/
2. HTTPS://TOWARDSDATASCIENCE.COM/40-OPEN-SOURCE-AUDIO-DATASETS-FOR-ML-59DC39D48F06
3. HTTPS://MATLABHELPER.COM/BLOG/MATLAB/PITCH-DETECTION-OF-SPEECH-SIGNAL-USING-AUTOCORRELATION-METHOD-IN-MATLAB/
4. HTTPS://WWW.MATHWORKS.COM/HELP/MATLAB/REF/AUDIOREAD.HTML
5. HTTPS://WWW.RESEARCHGATE.NET/PUBLICATION/306024204_A_TUTORIAL_TO_EXTRACT_THE_PITCH_IN_SPEECH_SIGNALS_USING_AUTOCORRELATION