**ENGR 518 PROJECT REPORT School of Engineering**

**(5 pages max excluding references) Faculty of Applied Science**

**University of British Columbia**

**Project Title:** Speaker Recognition - Classification

**Group No.:** 14

**Members:** Juliya Johnson, Shubham Mohapatra, Kaimeng Du

**Date:** Oct 13, 2023

## Introduction

In this project we are going to collect data and train a classifier to identify speakers from different voice sources.

The project requirement asking to train a 2-class classifier to identify a certain speaker from the other. We decided to train a 3-class classifier to identify speech from each group member. Since this 3-class classifier includes 3 one-over-all 2-class classifiers, what we will do will fulfill the project requirement.

In today’s technology savvy market, machines are efficient in understanding human behaviors through training and in the emulation of these real time aspects. One such example is the speaker recognition project described in this report where the machine/system will be trained to identify and classify different voices. As the main objective of the project is to distinguish between different audio notes followed by outrightly identifying and recognizing the dedicated voice as the actual speaker, a time domain- based approach isn’t the best fit for dealing with this plan of action. Numerous audio samples at various frequencies and a wide range of sampling rates must be considered.

## Theory

Consider a beautiful painting with various colors but it is usually hard to figure out the individual colors of the painting. Similarly, getting to the individual pieces of a complicated and completed jigsaw puzzle is a tough job. In such cases, Fourier transform is the most efficient tool in breaking down the painting or puzzle into individual colors or pieces respectively. In a similar manner signals can be broken down and analysis can be performed.

A conversion technology involving the time domain and frequency domain must be adopted in order to achieve this. Fourier transform is the pick of the bunch when considering the conversion of time domain signals to frequency domain. The frequency elements can be easily examined and analyzed by using Fourier transforms. As sequence of values will be analyzed in this case, samples from the recorded continuous time signals will be broken down into discrete components. The best approach to understand these discrete values is by a using class of Fourier transforms called the discrete Fourier Transform (DFT). DFT finds its applications mainly in digital signal processing, numerical analysis and digital systems data. The Fast Fourier Transform (FFT) is the most accurate algorithm for the scrutinization and calculation of the DFT of a particular dataset. The FFT algorithm aims at limiting the extent of mathematical complexity encountered during the direct implementation of DFT for larger sets of data.

## Algorithm

3 subjects were recruited in this test. From each subject, around 10-minute stereo voice data was recorded with a voice recorder (R3312, Aigo) using 16kHz sampling rate and 16-bit data width in wav format. The wav files were then processed in MATLAB. Left channel data was segmented with 2048-point moving window with 75% overlap, then underwent 2048-point FFT. The spectrum was converted into log scale for further analysis.

## Results and Discussion

A screenshot of a graph

Description automatically generated

Figure 1, the frequency spectrum of the voice data collected from 3 subjects. Fundamentals and harmonics can be observed, pointed with arrows.

The fundamental, first and second harmonics of the voice data from Subject 1 and 2 can be clearly observed (green arrows), while only the fundamental and first harmonic can be observed for Subject 3 (black arrows).

## Conclusions

As shown in Figure 1, by comparing the shape of frequency spectrum below fundamental, as well as the position of the harmonic peaks, data from 3 subjects can be clearly identified. Therefore, we believe this data would be adequate for training a classifier to conduct this task. We may also utilize PCA or other data processing tools to down sample the spectrum or extract features before feeding the data into classifier in project.

## Appendices

Include in the appendices any of the following:

a. Calculations and Derivations.

b. Computer Program Listings / Information, etc

## References

[1] L. Auslander and R. Tolimieri, Is computing with the finite Fourier transform pure or applied mathematics? Bull. Amer. Math. Soc. (N.S.), 1 (1979), pp. 847–897.

[2] K. G. Beauchamp, Transforms for Engineers. A Guide to Signal Processing, Clarendon Press, Oxford, UK, 1987.

[3] W. L. Briggs and V. E. Henson, The DFT: An Owners’ Manual for the Discrete Fourier Transform, SIAM, Philadelphia, 1995.

[4] P. Burgisser, M. Clausen, and M. A. Shokrollahi ¨, Algebraic Complexity Theory, Springer, Berlin, 1997.