Voice Authentication System

Group Members: Saad Zahid Qureshi (saadzahid133204@gmail.com)
Bilal Jabbar (bilaljabbar381@gmail.com)
Supervisor: Dr. Shehzad Faisal (Shahzad_faisal@ciit-attock.edu.pk)
Co-Supervisor: Mr. Najam Dar (najam.dar@ciit-attock.edu.pk)
Department of Computer Science

Introduction

- To prevent breeches, into any system we should need to implement a secure and reliable system.
- Voice authentication system is kind of system that provides a voice based authentication that is quit different from the other commonly used biometric systems like fingerprint, face detection etc.
- This authentication process will be accomplished by using voice features matching with already trained model.

Fig. 1 Human voice Activities

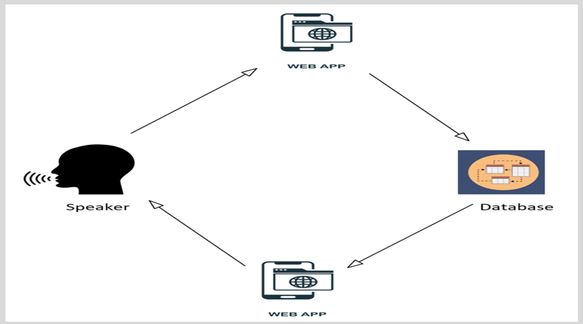

Fig. 2 Control flow of system

Objectives

The objectives of this project are to

- To design and implement a more inventive and secure mechanism.
- Create a system that validate user's voice regardless of gender and age.
- To investigate and use Artificial Intelligence (AI) approached to speaker recognition.
- Extends the capabilities of the computer system with minimum keyboard input.

Methodology

Human Voice Recognition involves the following steps:

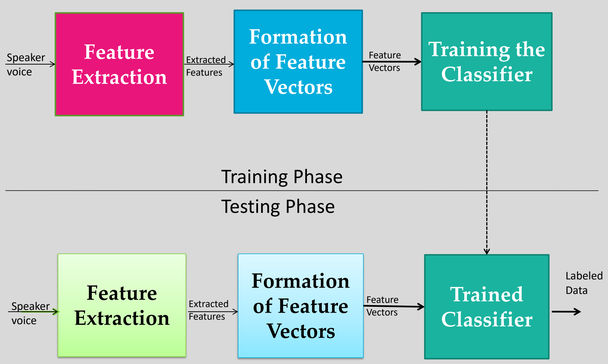

Fig. 3 voice Recognition Process

Architecture

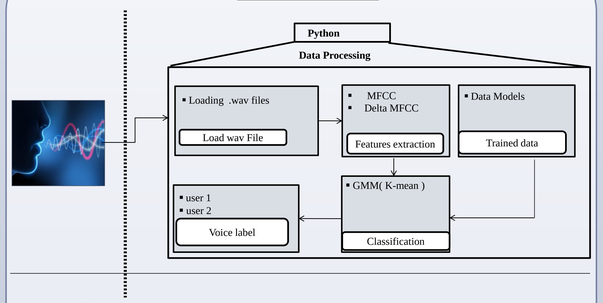

Fig. 4 Proposed Architecture of voice Recognition System

Flow chart

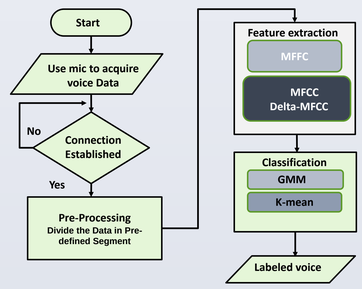

Fig. 5 Flow Chart of proposed voice recognition Systems

Voice Sampling For Data Acquisition

- Take user voice sample using microphone
- Then split his/her voice into segments.
- Save data to the specific user.
- Create a class as per user

Features Extraction

MFCC

- It stands for Mel Frequency Cepstral Coefficients.
- It is very useful in speaker recognition process because it works upon the human peripheral auditory system.
- MFCC coefficients are found out by taking the log magnitude of the windowed.
- S (n) = s(n) – a*s(n-1)

Delta-MFCC

- By taking first derivative of MFCC features we can extract the Delta MFCC features.
- Delta features are used to represent the related Delta features to the change in the cepstral features with time.
- This technique adds an approximation of the first time derivatives of basic features (for example MFCCs) to the feature vector.

$$C_n = \sum_{k=1}^{k} (\log D_k) \cos\left[m\left(k - \frac{1}{2}\right)\frac{\pi}{k}\right]$$

- where Cn represents the MFCC and m is the number of the coefficients here m=20 so, total number of coefficients extracted from each frame is 20.

$$d_t = \frac{\sum_{n=1}^{N} n(c_{t+n} - c_{t-n})}{2\sum_{n=1}^{N} n^2}$$

- Where dt is a delta coefficient, from frame t computed in terms of static coefficients ct+N to ct-N.
- A typical value for N is 2.

GMM Likelihood Algorithm

- Maximum likelihood estimation (MLE) is an estimation method that allows us to use a sample to estimate the parameters of the probability distribution that generated the sample.
- A sample that we use to make statements about the probability distribution that generated the sample
- The GMM probability density function can be presented by the
- covariance matrix (∑), Mean vector and parameter set of the mixture weight (Ki). The multidimensional
- Gaussian mixture model is given by equation

$$K_i = \frac{1}{(2\pi)^{\frac{d}{2}}|\sum_i|^{\frac{1}{2}}} \exp\{-\frac{1}{2}(x-\mu_i)^T \sum_i^{-1}(x-\mu_i)\}$$

Fig. 6 GMM likelihood Algorithm

Results

Table 1. Dataset of Acquired voice Samples

User	Labels	Time Duration (sec)
User 1	0	6
User 2	1	6
User 3	2	6
.	.	.
.	.	.
.	.	.
User n	n-1	6

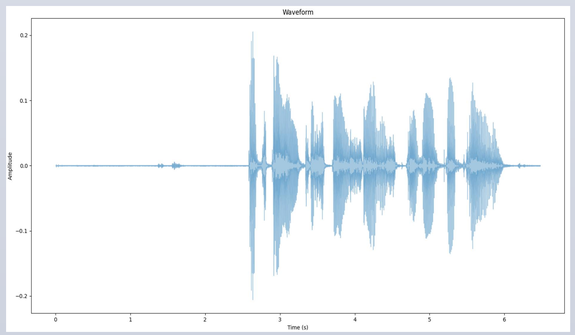

Fig. 7 Time and Amplitude of a wav

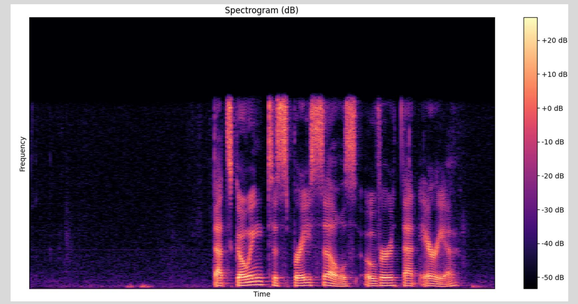

Fig. 8 spectrogram of wav file

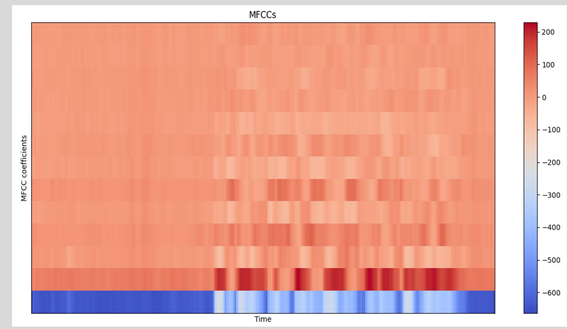

Fig. 9 Time and MFCC of wav file

Conclusion

- The voice-based security system can play vital role in the security of personal or organizational
- The given system is developed to provide the authentication login system to the employees of the organization (i.e., Bank).
- To provide the touchless experience to the employees which can be very useful and hygienic to due to current diseases (i.e., Corona).
- In in normal situations this can be use as security layer with traditional login procedures..

References

[1] n. kodaagoda, "voislock-Human voice authentication using Hidden Markove Model," dec, 2008.
[2] s. b. w. a. frederick t.sheldon, "voice biometriuc identity authentication for Iot devices," 2020.
[3] R. Tanwar, "an approach to ensure secuirty using voice authentication system," 2019.
[4] P. Dhakal, "A Near Realtime Automatic Speaker Recognition Architecture for voice based user interface," 2019.
[5] A. Subramanian, "Building a Speaker Recognition Model," 17 July 2021. [Online]. Available: https://blog.dominodatalab.com/building-a-speaker-recognition-model.
[6] b. p. a. k. Atul anand, "Speaker Identification over Call Record," 2018.
[7] V. Bhapkar, "Speaker Identification using Machine Learning," medium, [Online]. Available: https://medium.com/analytics-vidhya/speaker-identification-using-machine-learning-3080ee202920. [Accessed 04 April 2022].