International Journal of Engineering & Technology, 7 (4.5) (2018) 95-97

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Speaker Recognition System for Home Security using Raspberry Pi and Python

Dr.Bageshree Pathak[1], Shriyanti Kulkarni[2*], Aditi Karandikar[3*], Anaya Kharwadkar[4]

[1]M.K.S.S.S Cummins College of Engineering for Women, Associate Professor
[2,3,4]Student
**Corresponding author E-mail shriyanti.kulkarni@cumminscollege.in, [*]aditi.karandikar@cumminscollege.in*

Abstract

The transfer of manual controls to machine controls is automation. Automation is the need of the hour. Home automation is automation of home systems to create smart homes. It includes security systems, appliance control and environment control. The increasing need for safety and security has brought biometric security systems to the forefront. Speech being unique and individualistic can be used for biometric identification. The proposed system is a prototype which can be fitted for speaker recognition for home security. The system will identify the registered speakers and will allow access to the recognized speaker. The system is implemented on Raspberry pi platform using Python language.

Keywords: *Automation, Python, Raspberry pi, Security, Speech*

1. Introduction

There are several techniques that can be used for home automation. Out of these techniques the proposed system will be using speaker recognition. The system checks if the audio input provided by a speaker matches a registered sample of the system's database and will then allow or deny access. The proposed system finds use where the other biometric scans like facial recognition, iris or fingerprint identification is not possible. For real time recognition the system is implemented on Raspberry Pi platform[2], a compact and affordable single board computer developed by Raspberry Pi Foundation. We are using the Raspberry Pi 3 B model which is faster and more powerful than its predecessors. The Raspberry Pi uses Raspbian which is a Linux kernel-based operating system. Python is an open source programming language which is used to implement the algorithm of the system[5].

There are two major techniques of speaker recognition namely: speaker verification or authentication and speaker identification. Verification is done when the speaker claims to be of a certain identity and the speaker's voice is used to verify this claim. On the other hand, identification is used to determine an unknown speaker's identity. Our system will be using speaker verification.

Speaker recognition systems have two phases training and testing. Training phase includes extracting a template with the help of feature extraction from a recorded speaker's voice. Testing phase consists of a comparison of an utterance with the existing voice templates in the system's database.

2. Need of the System

Due to the increasing crime rate, there is a growing need for reliable and accurate security systems. The existing systems which are already in use, are implemented using wired networks or sensors. Our proposed system is a substitute for these heavily wired systems. By using speaker recognition, we aim to reduce the effort and inconvenience caused to the physically impaired as the system reduces the need for interaction between the user and the system. Our proposed system finds its use in environments where other biometrics face limitations.

3. Description

The block diagram of the proposed system is given in the following Fig. 1.

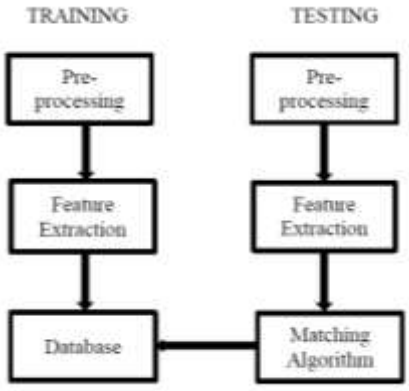


Fig. 1: Block Diagram

The system consists of two phases: training phase and testing phase. The training phase includes pre-processing, feature extraction and database generation. The testing phase includes the same pre-processing and feature extraction techniques and will be followed by a matching algorithm to compare the input audio signal of the testing phase with the templates in the database.

The preprocessing block (shown in Fig. 2.) includes sampling, windowing and frame formation and denoising. Preprocessing

Copyright © 2018 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

helps to increase the efficiency and accuracy of subsequent feature extraction and matching stages. This improves theoverall system performance.

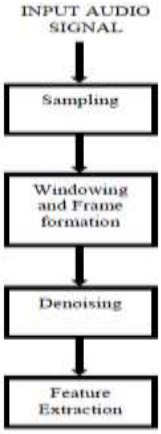


Fig. 2: Preprocessing Block

The input audio signal provided as input to the system is shown in Fig. 3.

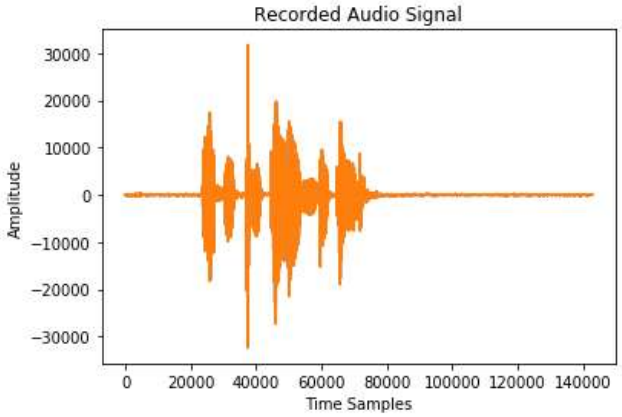


Fig. 3: Recorded Audio Signal

The continuous time input audio signal will be sampled and quantized to make it discrete in nature. Windowing function (as shown in Fig. 4.) will scale the signal. After dividing the input into a sequence of partial signals it is multiplied with a hamming window (as shown in Fig. 5.) to packetize it into frames.

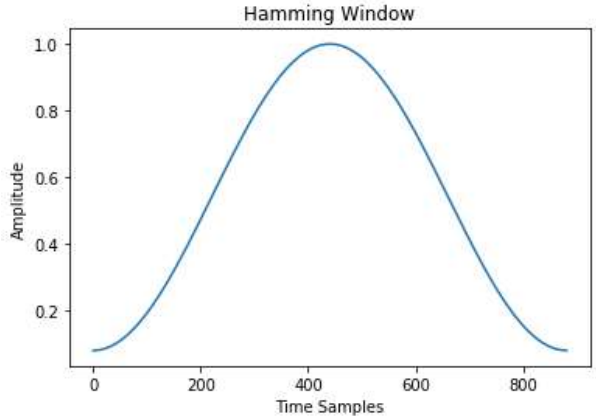


Fig. 4: Hamming Window

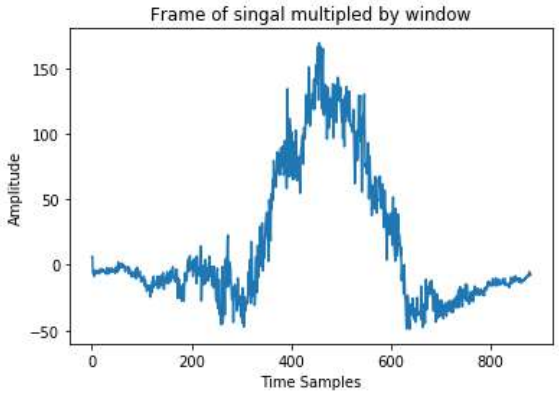


Fig. 5: Frame of signal multiplied by window

The input signal maybe corrupted by microphone related noise, environmental and electrical noise. Denoising and speech enhancement is done to improve the quality of speech signal by reducing the noise. The challenge of denoising is to eliminate external noise without tampering with the low intensity components of the speech input. These denoised samples are given to feature extraction block.

Feature extraction is deriving descriptive features from windowed and enhanced signals. Since speech is unique and individualistic, the features of each voice sample will be distinctive using which the speaker can be recognized. Ideally in this process the significant components will be emphasized and all other information will be suppressed. Various feature extraction techniques like Mel-frequency cepstral coefficients (MFCC), linear predictive coding (LPC), formants, pitch[1] can be used. Out of these techniques, our design uses MFCC analysis[4]. The templates of extracted features are stored in a database. The MFCC coefficients of one segment of an audio input are given in Fig. 6.

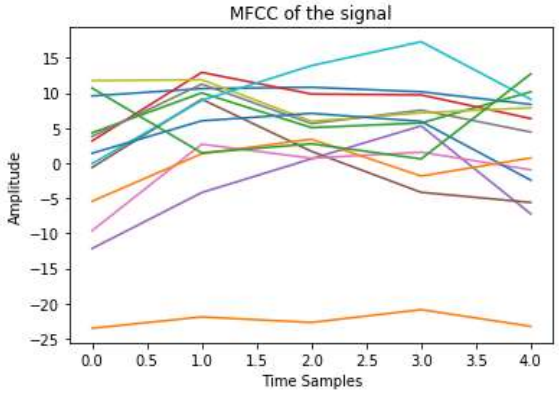


Fig. 6: MFCC of signal

Database of 8 speakers is created considering the average size of an Indian family. The input audio signal is taken using a Phillips microphone.

In the testing phase, the preprocessing and feature extraction routine is followed by a matching algorithm. Various matching algorithms like K nearest neighbor classifier (KNN), Naïve Bayes classifier or Euclidean distance can be used. The system is using Euclidean distance method as a matching algorithm[4].

3.1. Raspberry Pi

Raspberry Pi 3B model consists of various functional blocks as shown in Fig. 7. Out of which the system will use the power supply, SD card slot, USB ports and LCD interface[2].

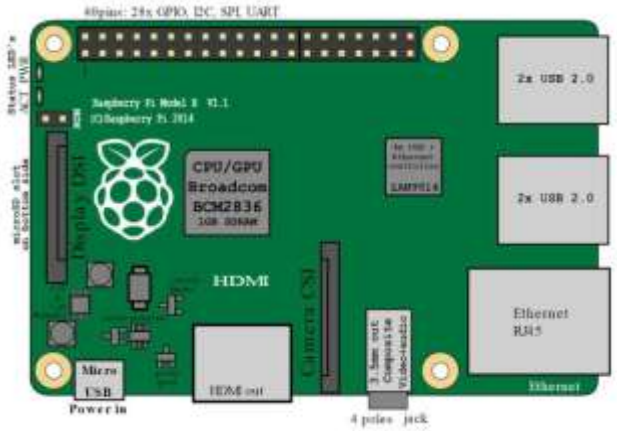


Fig. 7: Raspberry Pi

4. System Development

The constructional details of the system are as given in the Fig. 8.

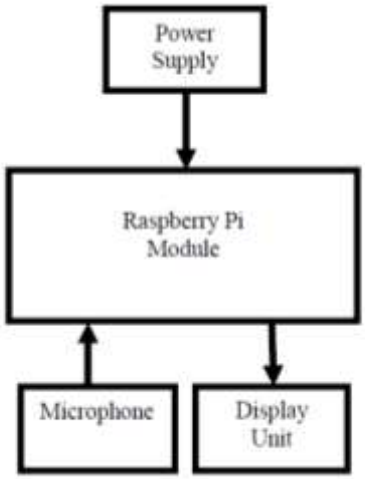


Fig. 8: System Layout

The microphone provides audio input to the system. The on-chip computer implements the algorithm[3]. The SD card acts as the data and instruction storage of the system. The output section consisting of a LCD displays if the access to the user is allowed or denied.

5. Conclusion

The system is designed for speaker recognition using Raspberry Pi. It is a prototype for home security system. Speaker recognition based home security systems are not implemented as much as the other biometric security systems. This system will not only find its use for physically impaired but also ease the basic functionality of our lives. The accuracy and efficiency of the system is better due to the MFCC analysis used. This system will be low cost, easy to handle and user friendly. By making use of other advanced technologies like the Internet of Things (IoT) the system can adapt to other dynamic environments.

Acknowledgement

We would like to express our sincere gratitude towards our project guide, Dr. Bageshree Pathak for her constant encouragement and valuable guidance. We take this opportunity to express our sincere thanks to all the staff members of M.K.S.S.S Cummins College of Engineering for Women for their help whenever required.

Finally, we would like to thank all those who directly or indirectly helped us during the project work.

References

[1] Juan A. Morales-Cordovilla, Antonio M. Peinado, Victoria Sanchez, Jose A. Gonzalez (2011) “Feature Extraction based on Pitch Synchronous Averaging for Robust Speech Recognition”.

[2] Patchava Vamsikrishna, Sonti Dinesh Kumar, Dinesh Bommisetty, Akshat Tyagi, (2016) “Raspberry Pi voting system, a reliable technology”.

[3] Radoslaw Weychan, Tomasz Marciniak, Adam Dabrowski, (2015) “Implementation aspects of speaker recognition using Python language and Raspberry Pi platform”.

[4] Umrani J Suryavanshi, Prof. Dr S.R.Ganorkar, (2014) “Hardware Implementation of Speech Recognitioon using MFCC and Euclidean distance”.

[5] John Glover, Victor Lazzarini , Joseph Timoney, ”Python For Audio Signal Processing”.