A Multimodal Biometric Authentication for Speech Controlled Automobile System

Akshay Ramesh[1], Sayantan Chakraborty[2], Jaykumar Chaudhary[3]

[1,2,3]Department of Electronics & Telecommunication Engineering
[1,2,3]Mukesh Patel School of Technology Management & Engineering,NMIMS(Deemed-to-be University),Mumbai,Maharashtra,India

---

Abstract - *Biometrics has been a subject of research by many scholars throughout these years. A biometric system consisting of only on a single biometric identifier in making a personal identification does not meet the desired performance requirements. Personal Identification Numbers (PIN) and key devices are not reliable and accurate techniques in secure environments. We introduce a multimodal biometric system on JAVA platform, which integrates face recognition, fingerprint verification, and speech recognition in making a personal identification. This system takes advantage of the capabilities of each individual biometrics.*

The processed information from all these three module is combined with the help of fusion algorithm in context switching in which feature vectors are made independently for query images and are then compared to the enrolment templates which are stored during database preparation for each biometric trait. Further, this paper is to propose the integration of Biometric technology with Automobile system and controlling the system using Speech input once the authentication is done through Android and Arduino, thus visualizing the future possibilities of Biometric Authentication and its applications in vehicular world.

***Key Words*: JAVA, Face Recognition, Speech Recognition, Fingerprint Recognition, Fusion, Bluetooth Technology, Arduino, Android, Automobile.**

1. INTRODUCTION

Biometric face, fingerprint and voice recognition are particularly attractive biometric approaches, since these three focuses on the same identifier that humans use primarily to distinguish one person from another: their "faces" and fingerprint is also around in criminal investigation since late 19th century[1]. One of the main goals of this system is to understand the complex human visual system as well as to use the knowledge of distinguishing one person's fingerprint from others and how humans represent faces in order to discriminate different identities with high accuracy with the use of the two most universally accepted biometric mechanisms.

1.1 Face Recognition

Image processing is a field that deals with manipulation of image with intent to carry out to enhance image and to extract some useful information from it.
In this work we used JAVA programming language in our aim to develop successful face recognition with a high recognition rate. The Haar cascade classifier is used to detect faces on a Java application. The classifier uses data stored in an XML file to decide how to classify each image location. Classification assumes a fixed scale for the face, say 50x50 pixels. The algorithm runs a series of feature detecting filters over the examples and generates data based on the shape of the object.
Once it has gathered this data, it can be used in the detection phase. In this phase a new object is introduced, and the same series of filters is passed over the object.

1.2 Speech Recognition

Speech recognition is also known as automatic speech recognition (ASR) or computer speech recognition which means understanding voice of the computer and performing any required task or the ability to match a voice against a provided or acquired vocabulary [2].
A speech recognition system consists of a microphone, for the person to speak into, and a speech recognition software. The software being described here uses Google voice and speech APIs. The voice command from the user is captured by microphone .this is then converted to text by using Google voice API. The text is then compared with the other previously defined commands inside the command configuration file. If it matches with any of them, then bash command associated with it will be executed.

1.3 Fingerprint Recognition

The fingerprint biometrics system is considered as one of the most efficient and trusted security system. The main reason for its reliability is that a fingerprint cannot have a positive match with someone else who is an unauthorized user. Each and every individual has a unique fingerprint and making it impossible to hack it [3]. We have used Fingerprint Module-R305, which is a serial fingerprint scanner that can be directly connected to the PC's com port, and to any controller via MAX232 IC. This Fingerprint scanner is capable of storing and comparing the fingerprint and accordingly giving the desired output. The scanner makes a copy of the fingerprint and compares its characteristics (such as branches and loops) to the ones stored beforehand.

Fingerprint processing includes two parts:

I. The system will process the two time finger images, generate a template of the finger based on processing results and store the template.
II. Fingerprint Matching: When matching, user enters the finger through optical sensor and system will generate a template of the finger and compare it with templates of the finger library.

For 1:1 matching, system will compare the live finger with specific template designated in the Module; for 1: N matching, or searching, system will search the whole finger library for the matching finger. In both circumstances, system will return the matching result, success or failure.

2. Literature survey

A Biometric system is a pattern recognition system which works in verification or identification mode. In verification mode system compares the query image against its stored template in the database (one to one). In identification mode the query image is compared against all the templates stored in the database (one-to-many).

Each character used depends on the application environment and has its own strengths and weakness.

Limitations imposed by unimodal biometric systems can be overcome using multiple sources of information of a person in establishing his identity. Integrating information from different traits (like face and fingerprint) or different samples of same trait (like multiple face images of person in different angles) results into more reliable, accurate biometric system [4].

In the research paper by Chowdhury S. "Implementation of Speech Recognition System for Bangla" [5], the author has used CMU Sphinx Tool which is a speech library and achieved an accuracy of 90.65% and by varying the speaker, an accuracy of 86.79% was obtained.

Similarly in the paper proposed by Bharti W. Gawali and Santosh Gaikwad, "Marathi Isolated Word Recognition System using MFCC and DTW Features"[6] , the classifiers MFCC (Mel-frequency cepstral coefficients) and DTW (Dynamic Time Warping) have been employed separately, and an accuracy of 84.65% was achieved for MFCC and 73.25% for DTW.

In M.A Abdou and M.H. Fathy's paper titled "An Efficient Hybrid Real Time Face Recognition Algorithm in Java Environment"[7], they use Hybrid skin color eigenface detection method which resulted in a better recognition rate and execution time.

Shervin Emami, Valentin Petruț Suciu proposed a face recognition technique titled "Facial Recognition using OpenCV" [8] where a set of detection algorithms were used that could be later packaged in an easily-portable framework. This technique gave a 95% successful recognition rate and less than 3% of the detected faces were false positives.

In fingerprint recognition, Le Hoang Thai and Ha Nhat Tam proposed a paper titled as "Fingerprint recognition using standardized fingerprint model" [9] in which they used pre-processing of fingerprint image, and found and adjusted the parameter sets. The results achieved after synthesizing fingerprint and applying post-processing for noise elimination gave an overall accuracy of 98.3-99.4%.The paper also compares this technique with the ones used by Xiping.

On the other hand, Madhuri and Richa Mishra [10] proposed a technique for fingerprint recognition and matching named "Local robust features SURF (Speeded up Robust Features)" which showed the presence of rotation and partial fingerprints and gave an accuracy of 99.46%.

3. Methodology

Developing a fully functional authentication system with three modalities: Face, Speech and finger print, and grant access to the person for an automobile system connected to the system via Arduino, which will be controlled using speech inputs.

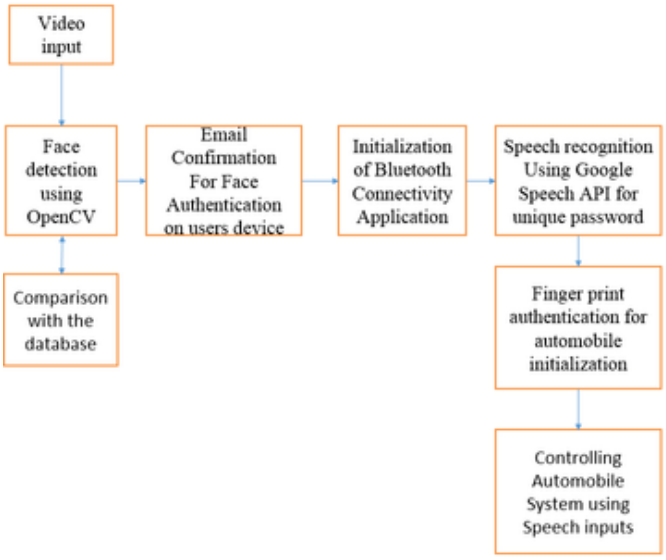

Fig-1: Block Diagram of System

3.1 Face recognition Module

Dynamic database for face detection is first created using MySQL database and then the face detection algorithm is carried out which gives a GUI (Graphic User Interface) output (using JAVA Swings) to Start, pause, Capture, Recognize, and register. Face Recognition is done using CBIR (Content Based Image Retrieval) technique. An email will be sent to the user's android phone will be notified to the Android app (Bluetooth SPP Pro).

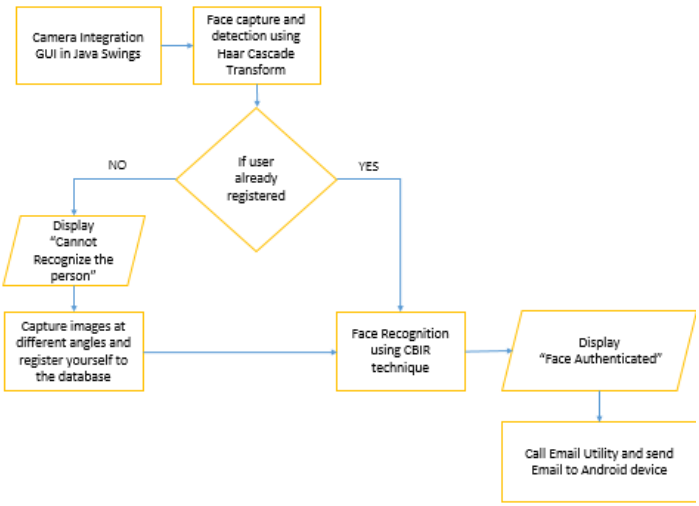

Fig-2: Flowchart of face module

3.2 Speech Recognition Module

Google API is integrated with our Android application and it takes a unique password from the user once the mail is received. The app will proceed to the next module only if the password matches the one in the database. Now running the Java with Google Speech API code will give us an output to record speech input and then display what has been said, by comparing the string input with the .gram file in the android code.

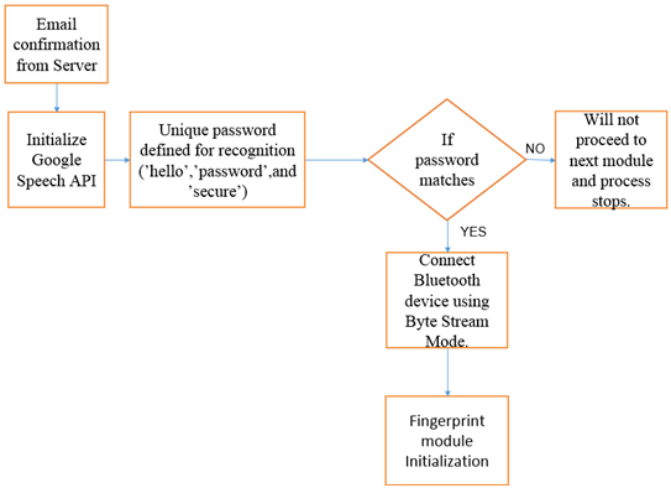

Fig-3: Flowchart of Speech module

3.3 Fingerprint Recognition Module

This will include digital Fingerprint image acquisition and feature extraction.

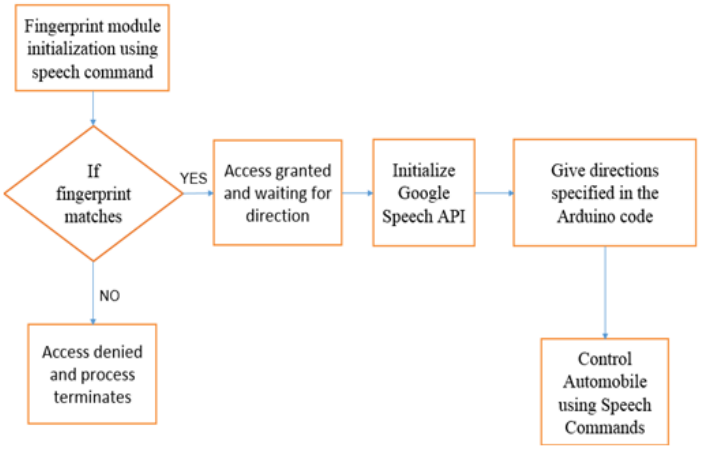

Fig-4: Flowchart of Fingerprint module

All the above modules will be fused together to form one multimodal system for authentication and if the above parameters are true, then it will start the automobile system through Arduino and the automobile will be further controlled using speech input from the App.

4. System Analysis

The system includes the following components which are all interconnected: Arduino Uno, Fingerprint Module(R 305), Bluetooth Module (HC-05), Android device, Motor Driver Circuit (L293D), and a Power Supply.

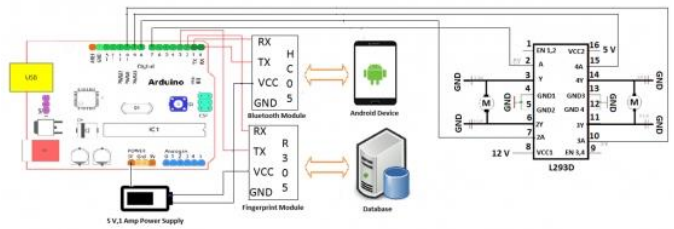

Fig-5: Detailed block diagram of system

The detailed block diagram of the system is as shown above and the explanation of each block is given further in this chapter:

4.1 Working of Face Recognition module

The first step is to detect the face using an external camera and recognizing the face by comparing it with the Database server, and then sending a mail to the user's android device if the face is valid.

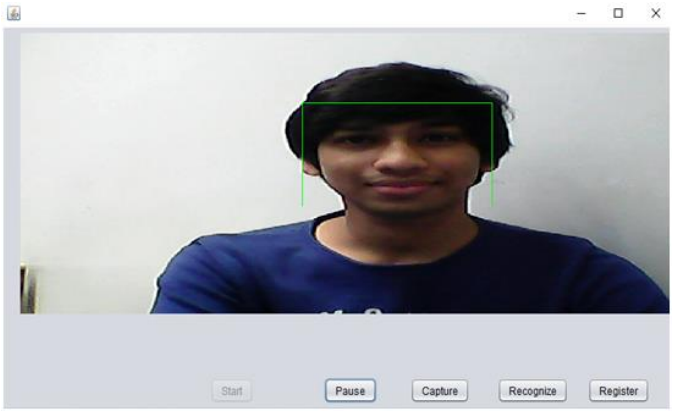

Fig -6: Starting the webcam and capturing images

After capturing the images at different angles and with different emotions, user can register himself to the database. The GUI also supports dynamic database creation of the user so that the person doesn't have to know the technical know-how of the system and directly register himself with the database.

The figure shows the registration form for the user.

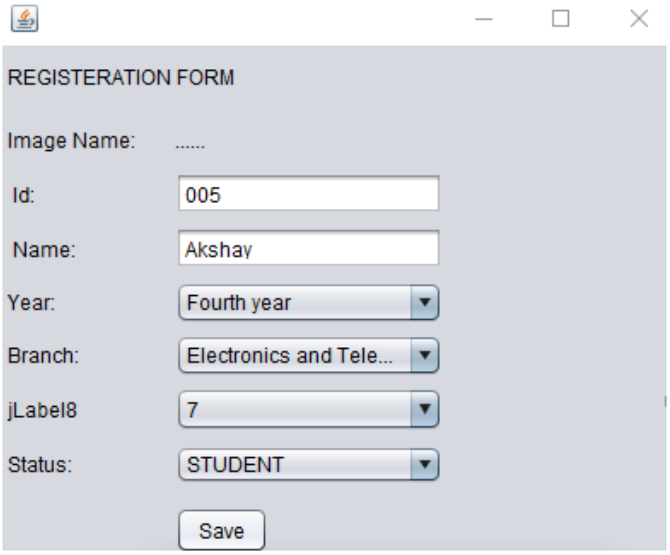

Fig -7: GUI to enter a person's registration details

After successful registration, user can check if his/her face is recognized by the system by re-running the code.

```
if(facesList!=null && facesList.size()>0)
{

    for(int i=0;i<facesList.size();i++)
    {
        Faces faces = facesList.get(i);

        if(faces.getPath().equals(iname))
        {

            JOptionPane.showMessageDialog(null, "Successfully Recognized");
            String emailId = "mukeshpatelnmims933@gmail.com";
            EmailUtility.sendEmail(emailId);
            System.out.println("Mail sent");


            check=true;
        }
    }
}
}
}catch(Exception e)
```

Fig-8: Java code for sending mail after Face Authentication

4.1 Working of Speech Recognition module

Google's current speech recognition system is speaker-independent and is built on deep neural networks together with hidden Markov models (DNN-HMM).

The strength of Google Speech lies in general purpose use, e.g. when making search requests on the World Wide Web. Google uses cloud-computing for speech recognition tasks.

Fig-9: Application asking speech input

After receiving password from user, it will compare with the passwords which are defined in the code. If password matches, only then it will initialize the application and application will search nearest Bluetooth connected automobile. If not, it will terminate the operation. After granting access to application-nearest Bluetooth connected device become visible.

4.3 Working of Finger Recognition module and Automobile System

After successful connection with device, it will ask for specific command to initialize the automobile. After initialization, it will ask user to put his finger onto

fingerprint module for verification as shown in Fig. 10.

After successful verification it will display a message "Access Granted" and it will wait for speech command for movement of the automobile.

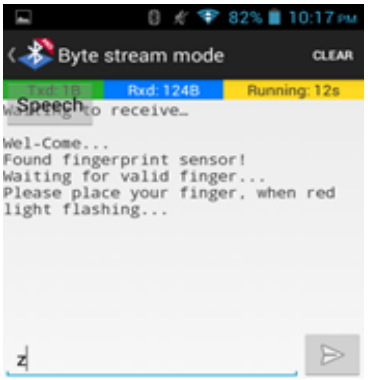

Fig -10: Waiting for valid command from user

The valid command for movement of automobile are defined in our Arduino code.

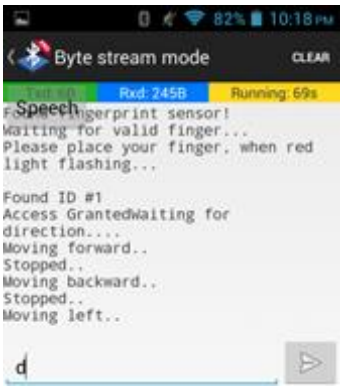

Fig-10: Command for vehicle left movement

5. Testing and Results

After successfully executing a face detection algorithm, an appropriately sized and scaled image of a face could be used for face recognition. The Haar cascade classifier algorithm was implemented by modifying sample code provided by OpenCV. A working Java application was created that could detect the user's face and place a green box around it and the face was successfully recognized.

The test results show that the detection method can accurately detect and trace human face in real time and the accuracy is represented in the below table:

Table-1: Confusion table for Face Recognition accuracy

Face Recognition Accuracy				
	User 1	User 2	User 3	Overall Accuracy
No. of times face was detected correctly (out of 10 attempts)	10	8	9	90%
No. of times face was detected correctly considering noisy environments (out of 10 attempts)	9	8	8	83.33%

Similarly for the Speech Module, Google Speech API was used and it was very much compatible with other software technologies (Java and Android) and the accuracy obtained is far better than the existing speech classifiers. The accuracy can be represented using a table as shown below:

Table-2: Confusion table for Speech Recognition accuracy

Speech Recognition Accuracy		
Password	No. of detection (out of 10 attempts)	Accuracy
'Hello'	10	100%
'Password'	9	90%
'Secure'	10	100%

The Fingerprint Module used in our project detects the user perfectly, and can be used in all environmental conditions. Its inbuilt DSP processor does all the processing step- by-step, and within milliseconds the digital image of the fingerprint is created and compared with the existing database and the user is authenticated if a match is found.

6. CONCLUSIONS AND FUTURE SCOPE

After some research was performed, it was discovered that OpenCV code can be executed on the Android platform, as well. One of the reasons that this implementation is desirable on the Android platform is the availability of a front facing camera on mobile Android

devices. A user interface can be developed with the touchscreen of the device to allow simple user commands. Also, we plan to make our system Speaker dependent rather than speech dependent system, and there will be a focus on development of the system automatically guess what the user intended to say, rather than what was actually said, to avoid mistakes and use many languages with translation according to user needs and microphone and sound system that will be designed to adapt more quickly to changing background noise level.

Further, the fingerprint sensors of Mobile devices can be used instead of the external fingerprint module and thus the whole system can be made portable for real time applications in the industries. But not all smartphones come with fingerprint sensors and thus we need to have a latest smartphone device for successful implementation.

REFERENCES

[1] Daniela Moctezuma, Cristina Conde, Isaac Martín de Diego and Enrique Cabello "Incremental Learning with soft-Biometric features for People Re-Identification in Multi-Camera Environments", *IEEE Conf. on Digital Image Computing: Techniques and Applications (DICTA)*, 2013, pp 1 – 7.

[2] M. Bacchiani, F. Beaufays, J. Schalkwyk, M. Schuster, and B. Strope, "Deploying goog-411: Early lessons in data, measurement, and testing," *in Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on, pp. 5260– 5263, IEEE*, 2008.

[3] Tiwalade O.Majekodunmi, Francis E.Idachaba "A Review Of The Fingerprint Recognition, Speaker Recognition, Face Recognition and Iris Recognition Based Biometric Identification Technology" *published in the proceeding of world congress on engineering 2011*, Vol.2, WCE 2011,July 6-8,2011, London, U.K.

[4]Dr. S. Ravi "Multimodal Biometric Approach Using Fingerprint, Face and Enhanced Iris Features Recognition", *IEEE Conf. on International Conference on Circuits, Power and Computing Technologies*, 2013, pp 1143-1150.

[5] Chowdhury, S., —"Implementation of Speech Recognition System for Bangla", *BRAC University, DHAKA, Bangladesh*, August 2010.

[6] R.K. Aggarwal and M. Dave, "Integration of multiple acoustic and language models for improved Hindi speech recognition", *Int. J. of Speech Technology*, Springer, DOI 10.1007/s10772-012-9131, 2012.

[7] Viraktamath SV, Mukund Katti, Aditya Khatawkar, Pavan Kulkarni, "Face Detection and Tracking using OpenCV," *The SIJ Transaction on Computer Networks & Communication Engineering (CNCE)*, 2013, 1(3).

[8] Shervin Emami, Introduction to face Detection and Face-Recognition http://www.shervinemami.info/faceRecognition.html.

[9] Le Hoang Thai and Ha Nhat Tam, "Fingerprint recognition using standardized fingerprint model", *IJCSI International Journal of Computer Science Issues*, Vol. 7, Issue 3, No 7, May 2010, Page No. 11-17.

[10] Madhuri and Richa Mishra, " Fingerprint Recognition using Robust Local Features", *in IJARCSSE*, 2012, Vol. 2, Issue 6, pp 1-5.