An ARM based Door Phone Embedded System for Voice and Face Identification and Verification by Open CV and Qt GUI Framework

K Uma Maheswari
M.Tech,Student,
DECS, Dept. of ECE
Vardhaman College of Engineering,
Hyderabad, India

J Krishna Chaithanya
Associate Professor,
Dept. of ECE
Vardhaman College of Engineering,
Hyderabad, India

ABSTRACT
Door phone embedded system for voice and face based user recognition is mainly used for security purpose. A natural way to identify a person is through their voice and face. Voice characteristics are different among individuals due to differences in their sound dynamics, vocal chords etc. A practical problem is voice can be trap by anyone. This proposed system is to recognize user password through voice and user recognize face at the door based on ARM processor. Speech technology is not 100% reliable but using an embedded microphone array increases speech in very noisy environments. Face Recognition is the process of matching the detected 'face' to one of many faces known to the file system. The system consists of three parts. First users pronounce a password through MIC which he enrolled and the MIC is present on ARM 9 board. Secondly, using an usb camera interface images is captured and the image is processed with help of Opencv and compared with existing database. And finally if the current image is match with any of existing image the system forwards user password to owner mobile through GSM module. The main aim of this project is to recognize the voice and face of the user at the door based on ARM processor. The algorithms were implemented in the system using C++. The system uses the Intel OpenCV library for image processing. However, when we use OpenCV library to detect a frontal face in an image using its Haar Cascade classifier face Detector, this will increase the human computer interaction by using real time face recognition. Phonon multimedia framework is used to display the image of the user. So that when user place face in front of camera it display recognize image of the user.

Keywords
Voice Recognition, Face Detection, Face Recognition, Qt GUI Framework, Open CV.

1. INTRODUCTION
Today, security is one of the utmost requirements of embedded systems. Now-a-days many security systems such as fire controlling systems, home security systems to prevent thieves from intruding are available in market. Due to advancement in technology, there is a wide development of these security systems to protect from security attacks.
A natural way to identify a person is through their voice and face. Voice characteristics are different among individuals due to differences in their sound dynamics, vocal chords, teeth etc. But, voice based security systems are accurate methods of identification, so in this paper we introduce face recognition system for automatic door access. In the last few decades, the importance of face recognition systems has been increased rapidly. Face recognition process can be implemented easily in security systems rather than other biometric systems such as fingerprint, signature, etc. Face recognition is the process where that detected and processed face is compared to a database of known faces, to decide who that person is.
According to T. Kinnunen and H. Li [1], the technology advancement in the recent years has addressed several technical challenges such as text/language dependency, channel effects and cross-talk speech, so they developed text independent telephone-based services with integrated speech recognition. F. T. H. den, M. Balm [2] states that, a residential gateway connects one or more access networks to one or more home networks and delivers services to the home environment. Saeed and M. K. Nammous [3] implemented speech and speaker identification method. It is based on spoken arabic digit recognition. The success rate of the speaker-identifying system obtained for individually uttered words is excellent and has reached about 98.8 %. S. J. Young [4] states that continuous speech from any speaker with average word error rates of between 15 and 30%. LVCSR systems are not robust to mismatched training and test conditions and cannot handle context. Face detection using artificial neural networks was done by Rowley [5]. Neural network is a nonlinear network adding features to the learning system. Hence, the features extraction step may be more efficient than the linear Karhunen- Loeve methods which chose a dimensionality reducing linear projection that maximizes the scatter of all projected samples [6]. This has classification time less than 0.5 seconds, but has training time more than hour or hours. Geometrical features image of a face is done by T Kanade [7]. The overall configuration can be described by a vector representing the position and size of the main facial features, such as eyes and eyebrows, nose, mouth, and the shape of face outline. Their system achieved a peak performance of 75% recognition rate. Graph matching is another method used to recognize face. M. Lades et al [8] presented a dynamic link structure for distortion invariant object recognition, which employed elastic graph matching to find the closest stored graph. But the matching process is complex and computationally expensive. I.J. Cox el [9] introduced a mixture-distance technique which achieved 95% recognition rate on a query database of 685 individuals. In this, each of the face was represented by 30 manually extracted distances.

2. SYSTEM OVERVIEW

The proposed system will use the OpenCV library from Intel for image processing. However, when we use OpenCV library to detect a frontal face in an image using its Haar Cascade classifier Detector, this will increase the human computer interaction by using real time face. Principal Component Analysis (PCA) is used to recognize a face in an image. PCA algorithm uses eigen faces nothing but a eigen values and eigen vector. Because we know that image is a multidimensional matrix in mathematics that can be represented by a matrix value. Image can be treated as a vector having magnitude and direction both. It is known as vector image or image vector. Phonon multimedia framework is used to display image of face. So that when user place face infront of camera the particular facial features of face is recognized.

This system supports feature of image/video processing by using various algorithms and features using image processing.

3. WORKING MECHANISM

In this concept, Password through voice and image of face should be enrolled by user before recognized by the system. First, users pronounce a password through mic and mic is present on board. If system recognize user password it goes to face recognition process. In this process, camera captures the image of user. Suppose if the system recognizes the user face, it display the user ID, buzzer beeps and user password sent to owner's mobile.

In our designed system users pronounce a password through MIC which he enrolled and the MIC is present on ARM 9 board. Secondly, using an usb camera interface images is captured and the image is processed with help of Opencv and compared with existing database. And finally if the current image is matched with any of existing image the system forwards user password to owner mobile through GSM module.

OpenCV is an open source computer vision library which runs under Linux, Windows, Mac OS X and it is written in C and C++. OpenCV was designed for computational efficiency and with a strong focus on real-time applications. OpenCV library provides us a greatly interesting demonstration for face recognition. Furthermore it provides us functions that they used to train classifiers for their face recognition system, called Haar-Training, so that we can create our own object classifiers using these functions.

ARM9 is an ARM architecture 32-bit RISC family. With this design generation, ARM moved from von Neumann architecture (Princeton architecture) to Harvard architecture with separate instruction and data busses, significantly increasing its potential speed. Most silicon chips integrating these cores will package them as modified Harvard architecture chips, combining the two address busses on the other side of separated CPU caches and tightly coupled memories.

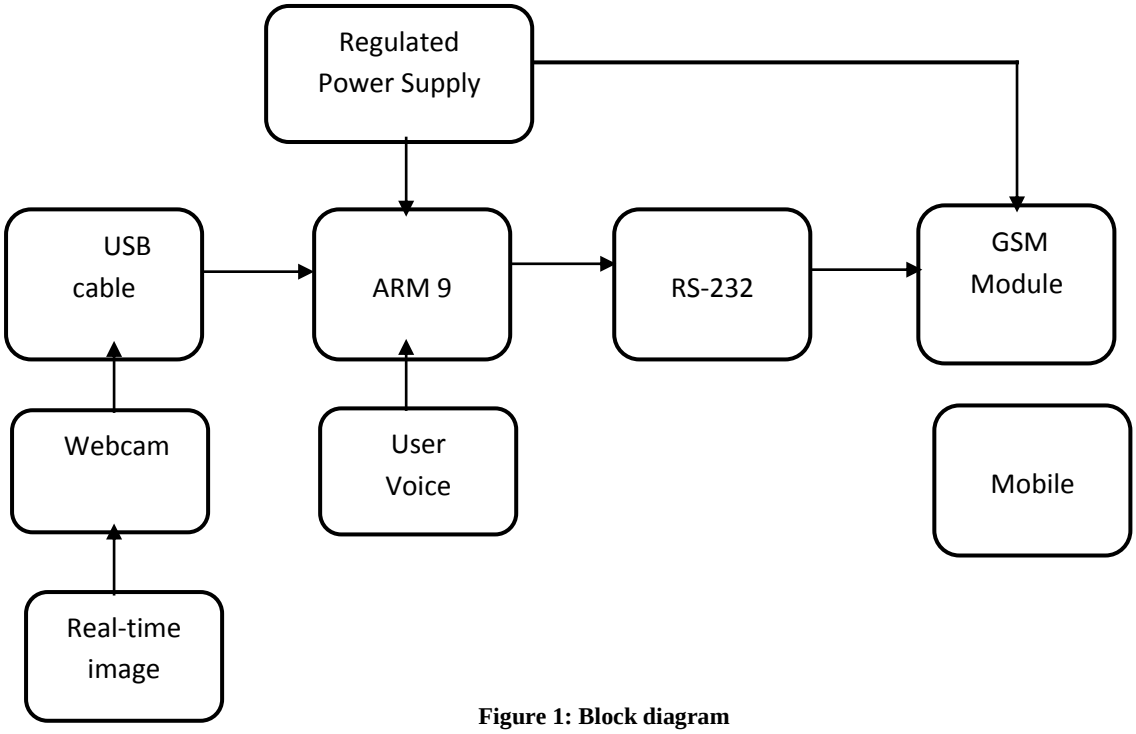

Figure 1: Block diagram

4. PROPOSED ALGORITHM

This paper proposes an algorithm for the real-time recognition of voice and face. This real-time algorithm is based on implementation of voice recognition, face detection, face recognition.This system uses linux operating system, opencv library and Haar classifier, Principal Component Analysis. The proposed system shown in Figure 2.

In this particular implementation first user should pronounce a password and next password verification will be processed. If it is verified it goes to face recognition. The live input real-time

image is captured through camera. The system detects face and next face recognition process takes place. System recognizes a face from previous stored database. Suppose user face is matched with existing database then the system display the user ID, simultaneously sends user password to owner's mobile and buzzer beeps which is present on system. Thus any potential gesture or movement made in front of camera from this point onwards will be processed for recognition.

For voice recognition we use pocketsphinx library, face detecion uses a Haar Cascade Classifier and for face recognition uses Principal Component analysis.

The Haarcascades training(haar training) is a quick tool to achieve face detection and recognition. The face, hand and body detection examples included in OpenCV installation example folder(/opencv/data/haarcascades/) demonstrate how fast haarcascade files help to do the job. In this proposed algorithm haarcascade training file is haarcascade_frontalface_alt2.xml is used for face detection.

Phonon is a cross-platform multimedia frame network that enables the use of audio and video content in Qt applications. Qt is a cross-platform application framework that is widely used for developing application software with a graphical user interface (GUI) (in which cases Qt is classified as a *widget toolkit*), and also used for developing non-GUI programs such as command-line tools and consoles for servers. Qt uses the phonon multimedia framework to provide functionality for display the captured image and also recognize the user from previous stored data.

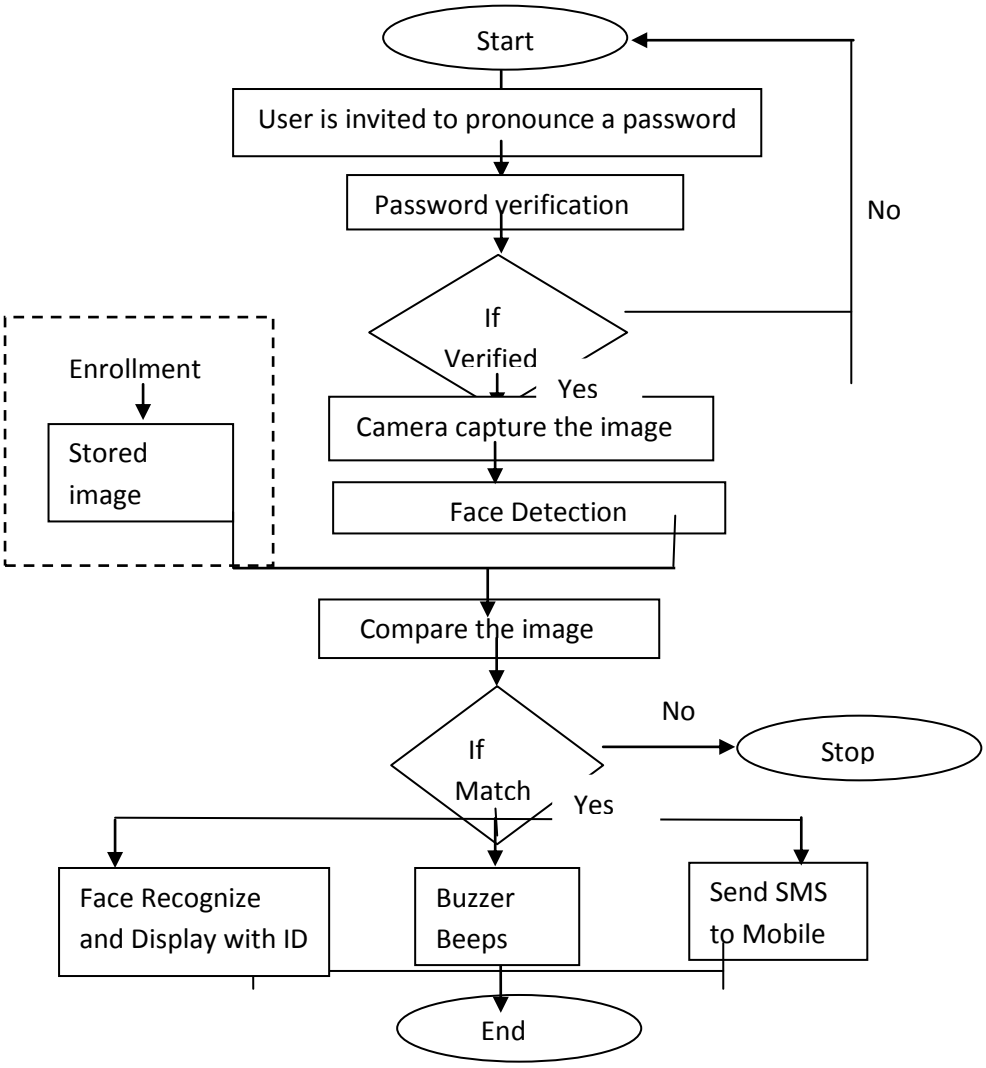

Figure 2: Proposed System

4. EXPERIMENTS AND RESULTS

Step 1:
Once all connections are provided for the proposed system as shown in Figure 3 then turn on the processor. LCD display starts with logo.

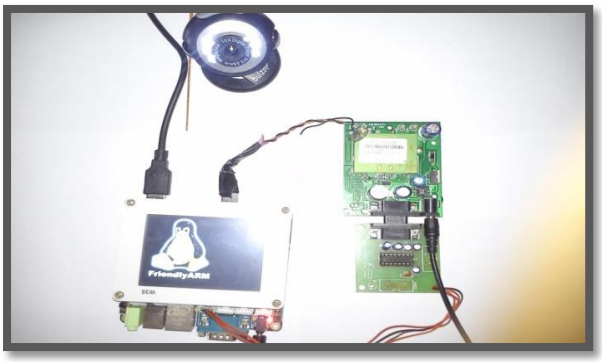

Figure 3: Proposed system starts with logo.

Step 2:
First a user start's with voice verification. Here P1 indicates person1 P2 indicates person2 and so on as shown in Figure 4. User should give the password through mic which he enrolled and the mic is present on board. And then click recognize button as shown in Figure 5. If it is recognized it will goes to the next step that is face recognition. Otherwise it stops there only.

Figure 4: User gives the password to the system.

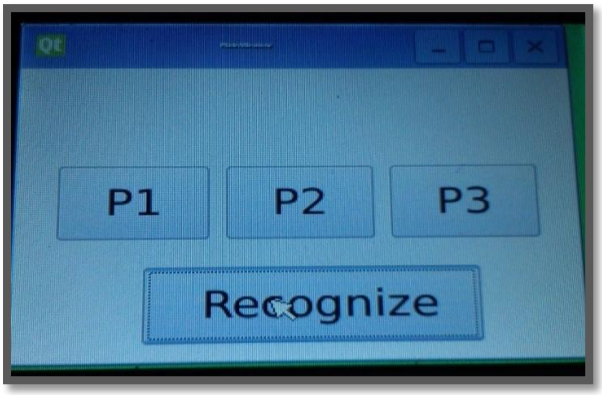

Figure 5: Users recognize the password.

Step 3:
In this proposed system capture button is used for enrolling the person images in database with ID's as shown in Figure 6. We've to capture our next images using add button as shown in Figure 7 until it display training images are enough. The image can be capture in different angles. But we've to capture only facial features eyes, nose and mouth. Image can be captured until the green rectangular box appears on face and this image is converting into gray scale image which is shown in Figure 7 top of the left corner.

Figure 6: Users enrol their image with ID

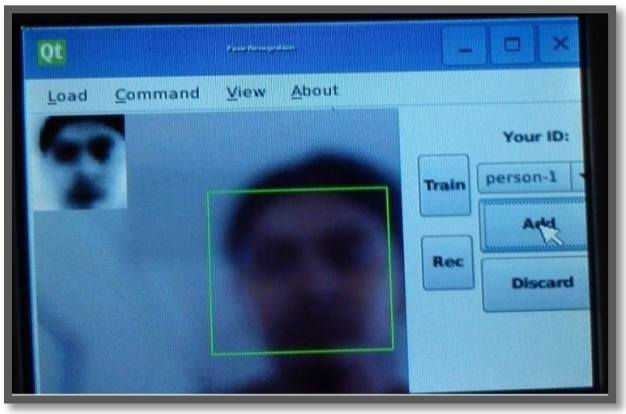

Figure 7: Users enrol their image in different angles.

Step 4:
By using train button check the image which we captured because the images should be clear with facial features is shown in Figure 8.

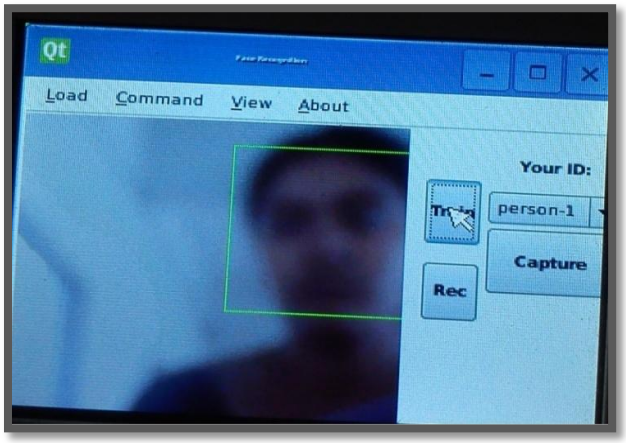

Figure 8: Users train their images.

Step 5:
The above steps are the process of enrolment and which is stored in database. Finally, when user wants to visit our home,

enter inside lab or wherever security needs user should pronounce a password. And after recognise a password user place a face in front of cam and click the Rec button nothing but recognition is shown in Figure 9. The system compare our image with enrol image which is stored in database. Suppose if it is recognized it display the person image with rectangular box, person ID and also simultaneously buzzer beeps which is present on processor and password will send to owner's mobile through GSM module shown in Figure 10. Buzzer beeps only after the verification process successfully completed.

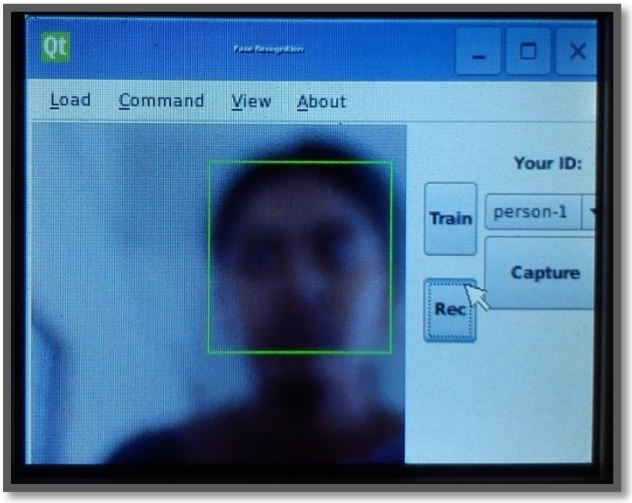

Figure 9: Recognizing a person using Rec button

Figure 10: Recognized person with ID and forward the user password to owner's mobile.

5. CONCLUSION

In this paper, we design an ARM based door phone embedded system for voice and face identification and verification by Open CV and Qt GUI framework. In present world there are many security systems available using different biometric applications such as fingerprint, signature, etc to provide security, but the voice and face recognition system is the best choice because it can be easily implemented toward high security, low cost and low power consumption.

6. REFERENCES

[1] T. Kinnunen and H. Li, "An overview of text-independent speaker recognition" Speech Communication, vol. 52, no. 1, pp. 12-40, Jan.2010.

[2] F. T. H. den Hartog, M. Balm, and C. M. de Jong, "Convergence of residential gateway technology," IEEE Communications Magazine, vol. 42, no.5, pp. 138-143, May 2004.

[3] Saeed, and M. K. Nammous, "A speech-and-speaker identification system: feature extraction, description, and classification of speech signal image," IEEE Trans. Industrial Electron., vol. 54, no. 2, pp. 887- 897, Apr. 2007.

[4] S. J. Young, "Large vocabulary continuous speech recognition: a review," *IEEE Trans. Signal Processing*, 1996.

[5] H. Rowley, S. Baluja, and T. Kanade, "Neural network-based face detection". *IEEE Trans. Pattern Analysis and Machine Intelligence*, 20(1):23–38, January 1998

[6] M. Kirby, and l. Sirovich, "Application of the Karhunen-Loeve procedure for the characterization of human faces". IEEE Trans. Patt. Anal. Mach. Intell. 12.

[7] T. Kanade, "Picture processing by computer complex and recognition of human faces," technical report, Dept. Information Science, Kyoto Univ., 1973.

[8] M. Lades, J.C. Vorbruggen, J. Buhmann, J.Lange, C. Von Der Malsburg, R.P. Wurtz, and M.Konen, "Distortion Invariant object recognition in the dynamic link architecture," IEEE Trans.Computers, vol. 42, pp. 300-311, 1993.

[9] I.J. Cox, J. Ghosn, and P.N. Yianios, "Feature-Based face recognition using mixturedistance" Computer Vision and Pattern Recognition, 1996.