Voice Recognition Security System With Face Recognition Using Matlab

Mohammed Mukhtar Latif
Shettigar Mokshan Devidas
Yadav Pankaj Sahablal
Sinha Amitkumar

Abstract: Biometric technology is fast gaining popularity as means of security measures to reduce cases of fraud and theft due to its use of physical characteristics and traits for the identification of individuals. The earliest methods of biometric identification included fingerprint and handwriting while more recent ones include iris/eye scan, face scan, voice print, and hand print. Security system using biometric technology: In this new technological era of new innovations and the obsolete technology getting compromised, we are to make striking balance for making our system foolproof. So in this project we have tried to implement an idea which uses multiple levels (stages) of authentication for a locker security System. Here we are using face and voice for security purpose.

Keywords: Registration and recognition, MATLAB, Basic locker mechanism.

I. INTRODUCTION

Biometrics has been used for identification or recognition purposes. The physical, behavioural, biological traits of an individual can verify a person’s identity. Physical traits include face, fingerprint, iris, and sclera. Behavioural traits are like gait, voice and biological include DNA. Each of these has its own advantages and disadvantages. Some traits could change over a period of time, cannot be used for recognition from a distance or can cause hygiene issues. A biometric may be more applicable in a particular scenario than the rest. No biometric is perfect or can be applied universally. Biometric systems compared with traditional authentication schemes are more reliable and it is difficult to copy, share or distribute the biometric feature [1]. The biometric characteristics have the following requirements due to which they can be used in authentication schemes.

- Universality: each person should have the characteristic.
- Distinctiveness: any two persons should be different.

FEATURES

- Two Modes of Operation (Registration and Recognition)
- Speaker Recognition in Matlab.
- Face Recognition in Matlab.
- Basic Locker Mechanism.

II. LITERATURE SURVEY

A. A MAXIMUM LIKELIHOOD APPROACH TO CONTINUOUS SPEECH RECOGNITION SYSTEM

Abstract- Speech recognition is formulated as a problem of maximum likelihood decoding. This formulation requires statistical models of the speech production process. In this paper, we describe a number of statistical models for use in speech recognition. We give special attention to determining the parameters for such models from sparse data. We also describe two decoding methods, one appropriate for constrained artificial languages and one appropriate for more realistic decoding tasks. To illustrate the usefulness of the methods described, we review a number of decoding results that have been obtained with them.

B. A SURVEY OF BIOMETRIC RECOGNITION METHODS

Biometric recognition refers to an automatic recognition of individuals based on a feature vector(s) derived from their physiological and/or behavioural characteristic. Biometric recognition systems should provide a reliable personal recognition schemes to either confirm or determine the identity of an individual. Applications of such a system include computer systems security, secure electronic banking, mobile phones, credit cards, secure access to buildings, health and social services.
By using biometrics a person could be identified based on "who she/he is" rather then "what she/he has" (card, token, key) or "what she/he knows" (password, PIN). In this paper, a brief overview of biometric methods, both uni-modal and multimodal, and their advantages and disadvantages, will be presented.

C. HMM CONTINUOUS SPEECH RECOGNITION USING PREDICTIVE LR PARSING

This paper proposes a new continuous speech recognition method using an efficient parsing mechanism, an LR parser, driving HMM modules directly without any intervening structures such as a phoneme lattice. Accurate and efficient speech parsing is achieved by combining HMM and LR parsing. This method is tested in Japanese phrase recognition experiments. Two grammars are prepared, a general Japanese grammar and a specific task with the general grammar is 72% for top candidates and 95% for the grammar, recognition rate is 80% and 99%, respectively.

D. TRAINING OF HMM WITH FILTERED SPEECH MATERIAL FOR HANDS-FREE RECOGNITION

This paper addresses the problem of hands-free speech recognition in a noisy office environment. An array of six omnidirectional microphones and a corresponding time delay compensation module are used to provide a beam formed signal as input to a HMM-based recognizer. Training of HMMs is performed either using a clean speech database or using a filtered version of the same database. Filtering consists in a convolution with the acoustic impulse response between speaker and microphone, to reproduce the reverberation effect. Background noise is summed to provide the desired SNR. The paper shows that the new models trained on these data perform better than the baseline ones. Furthermore, the paper investigates on MLLR adaptation of the new models. It is shown that a further performance improvement is obtained, allowing to reach a 98.7% WRR in a connected digit recognition task, when the talker is at 1.5 m distance from the array.

III. FLOW OF WORKING

Speech recognition, verification or identification systems work by matching patterns generated by the signal processing front end with patterns previously stored or learnt by the systems. Voice based security systems come in two flavours, Speaker Recognition and Speaker Verification. In Speaker recognition voice samples are obtained and features are extracted from them and stored in a database. These samples are compared with the various other stored ones and using and features increases this method becomes more taxing on the computer, as the voice sample needs to be compared with all other samples stored. Another drawback is that when number of users increase it becomes difficult to find unique features for each user, failure to do so may lead to wrong identification.

Viola-Jones algorithm for face detection: A face detector has to tell whether an image of arbitrary size contains a human face and if so, where it is. One natural framework for considering this problem is that of binary classification, in which a classifier is constructed to minimize the misclassification risk.

Since no objective distribution can describe the actual prior probability for a given image to have a face, the algorithm must minimize both the false negative and false positive rates in order to achieve an acceptable performance. This task requires an accurate numerical description of what sets human faces apart from other objects. It turns out that these characteristics can be extracted with a remarkable committee learning algorithm called Adaboost, which relies on a committee of weak classifiers to form a strong one through a voting mechanism. A classifier is weak if, in general, it cannot meet a predefined classification target in error terms. An operational algorithm must also work with a reasonable computational budget. Techniques such as integral image and attentional cascade make the Viola-Jones algorithm highly efficient: fed with a real time image sequence generated from a standard webcam, it performs well on a standard PC.

The figure 1 shown above is block diagram of the system.

A. COMPONENTS OF SYSTEM

WEBCAM: It is used to capture the image so as to give input in Matlab for sound output and convert into notepad. A webcam captures and recognises an object in view and tracks the user’s hand gestures using computer-vision based techniques. It sends the data to the computer. The camera, in a sense, acts as a digital eye, seeing what the user sees. It also tracks the movements of the thumbs and index fingers of both the user hands. The camera recognizes objects around you instantly.

MICROCONTROLLER: The AT89C2051 is a low-voltage, high-performance CMOS 8-bit microcomputer with 2K bytes of Flash programmable and erasable read-only memory (PEROM). The device is manufactured using Atmel’s high-density non-volatile memory technology and is compatible with the industry-standard MCS-51 instruction set. By combining a versatile 8-bit CPU with Flash on a
monolithic chip, the Atmel AT89C2051 is a powerful microcomputer which provides a highly-flexible and cost-effective solution to many embedded control applications.

**USB TO TTL:** The cable is easiest way ever to connect to your microcontroller/Raspberry Pi/WiFi router serial console port. Inside the big USB plug is a USB<->Serial conversion chip and at the end of the 36” cable are four wire - red power, black ground, white RX into USB port, and green TX out of the USB port. The power pin provides the 5V @ 500mA direct from the USB port and the RX/TX pins are 3.3V level for interfacing with the most common 3.3V logic level chipsets.

**Figure 2**

**HARDWARE REQUIREMENTS:**
- AT89C2051
- Motor Driver IC
- Power supply
- System: Pentium IV 2.4 GHz.
- Hard Disk: 40 GB.
- Floppy Drive: 1.44 Mb.
- Mouse: Optical Mouse.
- Ram: 2 GB.
- Keyboard: 101 Keyboards.
- Camera: 5 megapixel

**SOFTWARE REQUIREMENTS:**
- Diptrace for PCB layout
- Keil for Uc programing
- MATLAB for IP programing
- UcFlash to burn Uc code.

**CIRCUIT DIAGRAM:**

**IV. CONCLUSION**

Online Human Identification using face and voice recognition is still a very challenging topic after decades of exploration. A number of typical algorithms are presented separately, being categorized into appearance-based schemes.

**V. ACKNOWLEDGEMENTS**

Our thanks to M. H. Saboo Siddik College of Engineering, Department of Electronics Engineering, for giving us the initiative to do constructive work on such a great topic. We also thank anonymous reviewers for their constructive suggestions.

**REFERENCES**