Security Implementation using Biometric

Submitted in partial fulfillment of
the requirements for the award of the degree of

Bachelor of Technology
in
Computer Science and Engineering

Submitted by

B130253CS Shrimadhav U K

Under the guidance of
Dr. Vinod Pathari

Department of Computer Science and Engineering
National Institute of Technology Calicut
Calicut, Kerala, India – 673 601

Winter Semester 2017

Declaration

I, hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material which has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgement has been made in the text.

Place: NIT Calicut
Date: May 1, 2017
Name: Shrimadhav U K
Reg.No: B130253CS

Signature:

Department of Computer Science and Engineering

National Institute of Technology Calicut

Certificate

This is to certify that this is a bonafide record of the project presented by the students whose names are given below during Monsoon Semester 2016 in partial fulfilment of the requirements of the degree of Bachelor of Technology in Computer Science and Engineering.

B130253CS Shrimadhav U K

Dr. Vinod Pathari
(Project Guide)

Ms. Anu Mary Chacko
(Course Coordinator)

Acknowledgments

I would like to express my special thanks of gratitude to my guide **Dr. Vinod Pathari**, as well as to our course coordinator **Anu Maám**, who gave me the golden opportunity to do this project, which also helped me in doing a lot of research, helped me develop my own application program, and I came to know about many new things.
I am really thankful to them.
Secondly, I would also like to thank my parents and friends who helped me a lot in finishing this project in the limited time.
Last but not the least, I would also like to thank God for giving me the courage and strength to do this project in a feasible manner.

Shrimadhav U K
May 2017
National Institute of Technology Calicut

Abstract

The project aims to develop a biometric security system, which can protect the user's device(s) from unauthorized or unauthenticated access. The idea is inspired from Microsoft Windows Hello and Google Now, which allows us to speak our mind and the machine does it, through the profound advancement in machine learning and artificial intelligence. I plan to implement an Android application which can recognize the face and the voice of the user, and accordingly allow or deny access to the system.

Contents

List of Figures

Chapter 1

Introduction

Biometric Security is gaining more and more attention recently. This project attempts to implement an application which can take the voice input from a microphone, face input from a camera, and verify the authenticity of the user accessing the system.

1.1 Motivation

Human beings have reached a stage where it is no longer convenient to type the password when they want to be authenticated. This was the basic motivation of this project, i.e., to replace the password input using a keyboard, and instead ask the user to smile in front of their personal computer, and talk interactively to it. Then that personal computer unlocks, if it recognizes the integrity of the user.
Currently, no fool-proof solution exists which attempts to do both these tasks. There exists individual solutions for each of these individual tasks. But, these solutions are proprietary and requires specific licenses to use the offered services.

Chapter 2

Problem Statement

To design a security system for GNU/Linux operating system using biometric of the user, i.e., the face and the voice of the user, that would replace the traditional password input using a keyboard.

2.1 Related Works

1. Google Now This is an artificial inteligent personal assistant, made by Google, using their profound advancement in Natural Language Processing and Machine Learning. The assistant can do various actions, powered by their core serch technology. The most innovating function of this personal assistant is that in a smartphone with Android operating system, the phone can be locked, or unlocked by simply speaking to the smartphone using its microphone.
   `https://www.google.com/search/about/learn-more/now/`

2. Microsoft Windows Hello This is one of the authentication technologies used by Microsoft in their Windows operating system. The systems with supported cameras, and other hardware devices, includes improved support for biometric authentication. In this technology, the users show their face to the system camera, thereby allowing the user to be authenticated without the need to send their password.
   `https://support.microsoft.com/en-in/help/17215/windows-10-what-is-hello`

Chapter 3

Literature Survey

3.1 Face Recognition

3.1.1 Background and Related Work

Most of the work in recognition of faces was done by characterizing a face using a set of geometric parametrs , and then performing pattern recognition based on these values. One of the first system to do this is Kanade's face identification system.
In Kanade's face identification system [7], all steps of the recognition procedure were automated . This automation used a top-down control strategy with a generic model of expected feature characterstics. This system calculated a set of facial parameters of a single face image and used a pattern classification technique to match it from a known set. This depended primarily on local histogram analysis and absolute gray-scale values, since it was a statistical based approach.

3.2 Voice Recognition

3.2.1 Background and Related Work

Speaker recognition is the technique by which a person is identified using the characterstics of their voice biometrics.
Speech is a kind of complicated signal produced as a result of several transformations occuring at different levels: semantic, linguistic and acoustic. Differences in these transformations may lead to differences in the acoustic properties of signals. The recognizability of speaker can be affected not only

by the linguistic message but also the age, health, emotional state and effort level of the speaker.
Background noise and performance of recording device also interfere the classification process.
Speaker recognition is an important part of Human-Computer Interaction (HCI). As the trend of employing wearable computer reveals, Voice User Interface (VUI) has been a vital part of such computer. As these devices are particularly small, they are more likely to lose and be stolen. In these scenarios, speaker recognition is not only a good HCI, but also a combination of seamless interaction with computer and security guard when the device is lost. The need of personal identity validation will become more acute in the future. Telephone banking and Telephone reservation services will develop rapidly when secure means of authentication are available.

Chapter 4

Design

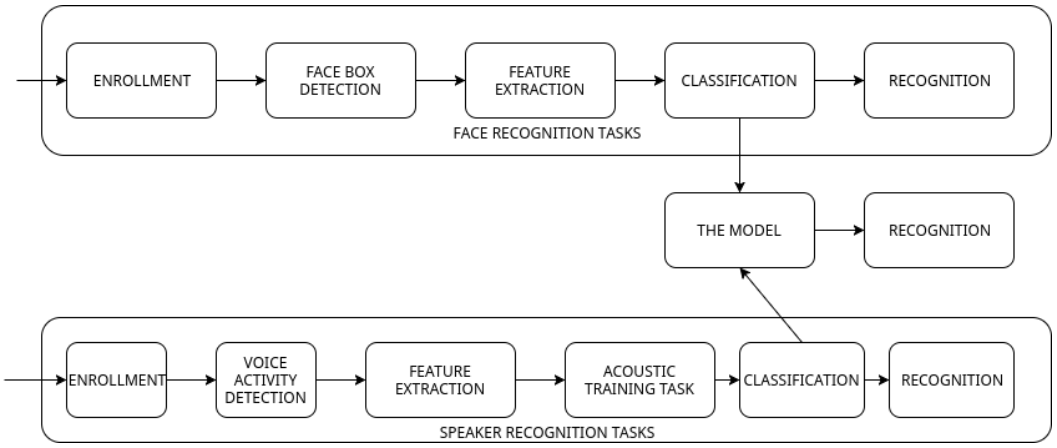

Figure 4.1: Design

Chapter 5

Implementation, Result, and Analysis

5.1 Face Recognition

5.1.1 The Local Binary Pattern Histogram Approach

The methods to improve the accuracy of the Face Recognizer have been made more stringent. The threshold used to control unknown faces, in the case of the EigenFaceRecognizer, from the calculated distance can be adjusted to allow better accuracy.

5.1.2 Principle Component Analysis

The EigenFaceRecognizer class applies PCA on each image, the results of which will be an array of Eigen values that a neural network can be trained to recognize.
The LBPHFaceRecognizer uses Local Binary Patterns (LBP) to create a feature vector using a Support Vector Machine or some other machine learning algorithm.

5.1.3 The Local Binary Pattern Histogram Classifier

The LBPH recognizer takes five variables:

radius The radius used for building the Circular Local Binary Pattern.

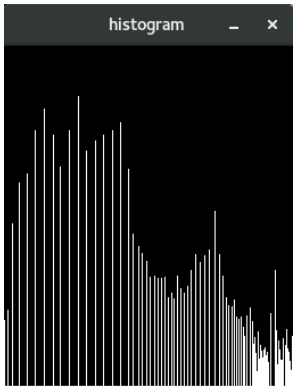

Figure 5.1: Local Binary Pattern

neighbors The number of sample points to build a Circular Local Binary Pattern from. OpenCV documentation suggests the value eight sample points. The more the number of sample points, higher will be the computational cost.

grid_x The number of cells in the horizontal direction. The common value used in publications is eight. The more cells, the finer the grid, the higher the dimensionality of the resulting feature vector.

grid_y The number of cells in the vertical direction. The common value used in publications is eight. The more cells, the finer the grid, the higher the dimensionality of the resulting feature vector.

threshold The threshold applied in the prediction. If the distance to the nearest neighbor is larger than the threshold, the method returns **-1**.

In the prototype implemented using the LBPH Approach (see Figure (5.1)), calculation of the LBP is done as part of the training process.
The recognition, using the LBPH approach, takes about 30 milliseconds implemented in Python on an Intel Core i5, using face images of size 132 x 132.

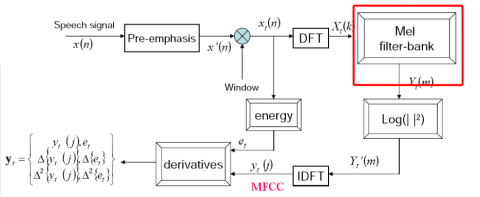

Figure 5.2: MFCC Feature Extraction Process

5.2 Voice Recognition

5.2.1 Algorithm Used

1. An utterance of a user is collected during the enrollment procedure.

2. Voice Activity Detection is performed: The silent part of the speech signals must be filtered out to decrease the bias that might occur during training.

3. Feature Extraction

   - Mel-Frequency Cepstral Coefficient (MFCC) is a representation of the short term power spectrum of a sound, based on a linear cosine transform of a log power spectrum on a non-linear mel-scale of frequency. This can be applied to speaker recognition task. The process to extract MFCC feature is demonstrated in Figure 5.2.

   - Linear Predictive Coding (LPC) is a tool used in audio signal processing and speech processing for representing the spectral envelope of a digital signal of speech in compressed form, using the information of a Linear predictive Model.
     The basic assumption that is used in LPC is that, the n th signal is a linear combination of the previous p signals. An optimization can be done by the Levinson-Durbin algorithm, to estimate the coefficients ai in the signal, the squared error should be minimized.

4. Gaussian Mixture Model (GMM) is used in acoustic learning task such as speaker recognition, since it describes the varied distribution of all the feature vectors. Therefore, GMM is merely a weighted combination of multivariate Gaussian distribution which assumes feature vectors are independent. We use diagonal covariance since the dimensions of the feature vector is independent to each other. GMM can describe the

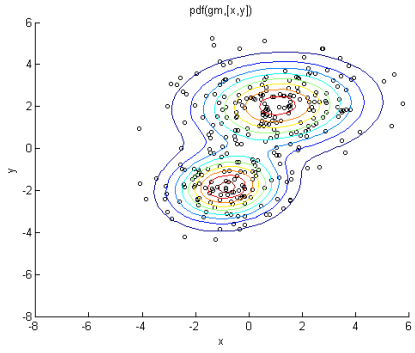

Figure 5.3: A Two Dimensional GMM with Two Components

distribution of feature vector with several clusters, as shown in Figure 5.3.

After training, the model can give the score of fitness for every input feature vector, measuring the probability that the vector belongs to this model. Therefore in the task of speaker recognition, we can train a GMM for every speaker. Then for a input signal, we extract lists of feature vectors for it, and calculate the overall likelihood that the vector belongs to each model. The speaker whose model fits the input best will be chosen as the answer.
Moreover, an enhancement has been done to the original GMM method. The training of GMM first requires a random initialization of the means of all the components. However, we can first use K-means algorithm to perform a clustering to all the vectors, then use the clustered centers to initialize the training of GMM. This enhancement can speed up the training and also give a better training result.

5. Joint Factor Analysis (JFA) is a typical method which behave very well in classification problems, due to its ability to account for different types of variability in training data. Within all the factor analysis methods, JFA was proved to outperform other methods in the task of speaker recognition.
   JFA models the user by supervector, i.e., a C x F dimension vector, where C is the number of components in the Universal Background Model, trained by GMM on all the training data, and F is the dimension of the acoustic feature vector. The supervector of an utterance is obtained by concatenating all the C means vectors in the trained GMM model.

5.3 The Biometric System

5.3.1 Advantages of Face Recognition

1. Face recognition systems are the least curious from a biometric sampling point of view. This is because they neither require contact nor requires the awareness of the user.

2. The face biometric can work with legacy photograph databases, video tape and other image sources.

3. It is a fairly good biometric identifier for small scale verification application.

5.3.2 Disadvantages of Face Recognition

1. A face needs to be well-lit by controled light sources in automated face authentication systems.

2. Face is a poor biometric for use in a pure identification protocol, it performs better in verification.

5.3.3 Advantages of Voice Recognition

1. Speech is a natural biometric. People use this to instinctively identify one another. Under certain circumstances, even machine decisions can be verified using Speech Recognition by relatively unskilled operators.

2. The voice biometric requires only inexpensive hardware and is easily deployable over existing, ubiquitous communications infrastructure. Voice is therefor very suitable for pervasive security management.

3. Voice recognition allows incremental authentication protocols. In an incremental authentication protocol, it waits for future voice data when the recognition confidence needs to be increased.

5.3.4 Disadvantages of Voice Recognition

1. Speech characterstics can drift away from models with age.

2. It becomes very easy to forge, or create non existent identities using machine synthesized voices, and hence it can create an automatic system that might be able to imitate a real human being.

3. Since the training of data depends to some extent on the quality of the audio signal captured, these systems are not immune to the background noise, channel noise, or other unknown channel or microphone characterstics.

Chapter 6

Conclusion

Face recognition is a fairly good biometric identifier for small-scale verification applications. Its systems are least least interrupting from a biometric sampling point of view. Another advantage of this is that, it uses low-power infrared illumination, to obtain rich images under poor lighting conditions.
Speaker recognition is also a fairly good biometric identifier. This type of biometric recognition is prevalent in every day communication. This also ensures higher accuracy and flexibility.
Hence, a combination of the above two biometric identification system is suitable for a good security implementation.
However, several points remain to be researched and extended.
For instance, since face to face meeting encompasses several modalities, such as speech and gesture, these capabilities need to researched and implemented.

References

[1] Microsoft Corporation (2014), Microsoft Windows Hello, Retrieved from https://support.microsoft.com/en-in/help/17215/windows-10-what-is-hello

[2] UCSD Computer Vision, Retrieved from http://vision.ucsd.edu/content/yale-face-database

[3] Google Inc. (2007), Google Now, Retrieved from https://www.google.com/search/about/learn-more/now/

[4] The CMU Audio Databases, Retrieved from http://www.speech.cs.cmu.edu/databases/

[5] P. Jonathon Phillips, A. Martin, C.l. Wilson, M. Przybocki (Febraury 2000). An Introduction to Evaluating Biometric Systems. Retrieved from http://www.face-rec.org/databases/151436.pdf

[6] Matthew Turk, Alex Pentland, Eigenfaces for Recognition

[7] Takeo Kanade, Computer recognition of human faces

[8] Arun Reddy Pothireddy, Spandana Sunkireddy, A Survey on Usage of Multiple Facial Recognition for Industrial Security