International Journal of Information & Computation Technology.
ISSN 0974-2239 Volume 4, Number 10 (2014), pp. 959-964

http://www. irphouse.com

A secure face tracking system

Ankita Gupta, Preeti Gandhi and Udit Jagga

IT, GGSIPU,
Paschim Vihar East, Delhi, INDIA

Abstract

In this paper, a new application of face recognition system using Eigen vector, Fisher based on PCA and LBPH (Local binary pattern histogram) based on LBP is proposed. Combination of Fisher and LBPH improves the capabilities of Eigen Vector approach. Fisher provides more accuracy and LBPH gives more uniform image that is more resilient to changes in lighting. The application focuses on three modules:
(i)Face Recognition: This phase includes face detection and recognizing part. Using different approach and algorithms, the face is detected on camera and matched with the already stored images in the database. The compatibility goes further with multiple camera image processing.
(ii) Authentication: This part of the study includes an authentication and check on the user before logging him into the system by using speech to text and text to speech feature.
(iii) Location Tracking: This is the most important module of the study where the location of a person is tracked by recognizing the face of the person in various CCTV cameras installed at various locations in the premises. The last seen feature can also be added to impart better functionality of the system.
Having all these three modules together in an application would help in developing a system which would be of great help to the society.

Keywords-Threshold; open cv; speech recognition; navigator

1. Introduction

With the ubiquity of new information technology and media, more effective and friendly methods for human computer interaction (HCI) are being developed which do not rely on traditional devices such as keyboards, mice, and displays. The rapidly

expanding research in face processing is based on the premise that information about a user's identity, state, and intent can be extracted from images, and that computers can then react accordingly, e.g., by observing a person's facial expression. A first step of any face processing system is detecting the locations in images where faces are present. However, face detection from a single image is a challenging task because of variability in scale, location, orientation (up-right, rotated), and pose (frontal, profile). In this paper we analyze and evaluate a range of face recognition systems, each utilizing a different image processing technique, in an attempt to identify and isolate the advantages offered by each system.

In section 1 we begin with a brief explanation of the different face recognition methods. This is followed by Speech to Text module for authenticity in section 2. The proposed paper provides the additional application of tracking the location of a person in section 3.

2. The Eigen Face and Fisher Face Methods

PCA is one of the oldest and best known techniques in multivariate analysis. Over the past few years, several face recognition systems have been proposed based on PCA. Given a set of m centered (zero mean, unit variance) samples $\mathbf{x}_{k.}$ $\mathbf{x}_{k.}$ =[x_k1…..x_kn]^t belongs to R^n.PCA aims to find the projection directions that maximize the variance of a subspace. During preprocessing, they register a gallery of m trained images to each other and unroll each image into a vector of n pixel values. Next, the mean image for the gallery is subtracted from each and the resulting "centered" images are placed in a gallery matrix. Element [i,j] of M is the i^{th} pixel from the j^{th} image. A covariance matrix $\Omega=MM^T$ characterizes the distribution of the images in $\Re^n$.

2.1 Eigen Face

Eigenface is a practical approach for face recognition. Due to the simplicity of its algorithm, we could implement an Eigenface recognition system easily. Besides, it is efficient in processing time and storage. PCA reduces the dimension size of an image greatly in a short period of time. The accuracy of Eigenface is also satisfactory (over 90 %) with frontal faces. However, as there has a high correlation between the training data and the recognition data. Preprocessing of image is required in order to achieve satisfactory result. The drawback is that it is very sensitive for lightening conditions, lightening conditions and aging effects.

2.2 Fisher Face

Fisherface is similar to Eigenface but with improvement in better classification of different class images. With Fisher Linear Discriminating approach, we could classify the training set to deal with different people and different facial expression. We could have better accuracy in facial expression than Eigen face approach. Besides, Fisherface removes the first three principal components which is responsible for light intensity changes, it is more invariant to light intensity. Fisherface is more complex than Eigenface in finding the projection of face space.

2.3 Local Binary Pattern Histogram

The LBP operator labels the pixels of an image by thresholding a 3*3 neighborhood of each pixel with the center value and considering the results as a binary number. Formally, given a pixel at(x_c; y_c), the resulting LBP can be expressed in the decimal form as

LBP(x_c; y_c)= Σ s(i_n-i_c)2^n(1) for n=0 to n=7

Where n runs over the 8 neighbors of the central pixel, i_c and $i_n$are the gray-level values of the central pixel and the surrounding pixel, and s9x0=1 if x>=0 and 0 otherwise.

3. Speech to Text and Text to Speech

If a person finds it difficult or is not capable of handling the mouse ports and the keyboard and if the keyboard or mouse is faulty, there have to be other ways to handle the operating system. “Speech” may act as one of them. There is a growing demand for systems capable of handling Operating System using only the voice commands given by a person. And this paper represents a way how to control the OS by using voice command.

3.1 Module Basics:

SAPI (Speech Application Programming Interface)

It is an interface between our application platform and Microsoft Speech Engine. The speech recognition engine that is utilized by this voice controlled system is the Microsoft's speech recognition engine and the associated development kit 5.1 (Microsoft Speech SDK 5.1). The recognition rate of Microsoft's speech recognition engine is not high in continuous speech mode but extremely high under the command control mode. We use (SAPI) to implement voice function. SAPI provides a high level interface between applications and speech engine. Controlling and management of various speech engines need real-time operation technology. However, SAPI realizes and hides the underlying technical detail.

There are two basic types of SAPI engines: text-to-speech (TTS) systems and speech recognizers. The TTS systems can synthesize text strings and files into spoken audio using synthetic voices, whereas Speech recognizers can convert human spoken audio into readable text strings and files. Speech engine communicates with SAPI by the device driver interface (DDI) layer and SAPI communicates with applications by API. So by the use of these application interfaces, voice recognition and speech synthesis software can be developed.

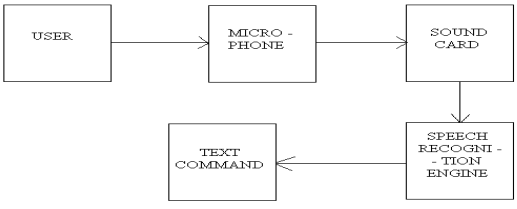

Figure 1: Working of speech to text module

As shown in Fig. 1, the user provides voice commands through microphone. The given command is then converted into electrical pulse by the microphone. The sound card converts electrical pulse into digital signal. The Speech Recognition Engine then converts digital signals into phonemes and finally we get text command. The respective operation is thus performed. This procedure repeats for every voice command.

4. Navigator

This project implements a small scale surveillance system with video analytics. Surveillance CCTV with required video analytics along with comprehensive Command and Control Centre is planned to be built to provide demonstration for public safety. The Scope of the work is limited to demonstrate a basic video analytics like facial recognition. Care is taken in such a way that this project will outline the overall solution considering advance video analytics for better vision and better solution.

4.1 Our Proposed Architecture

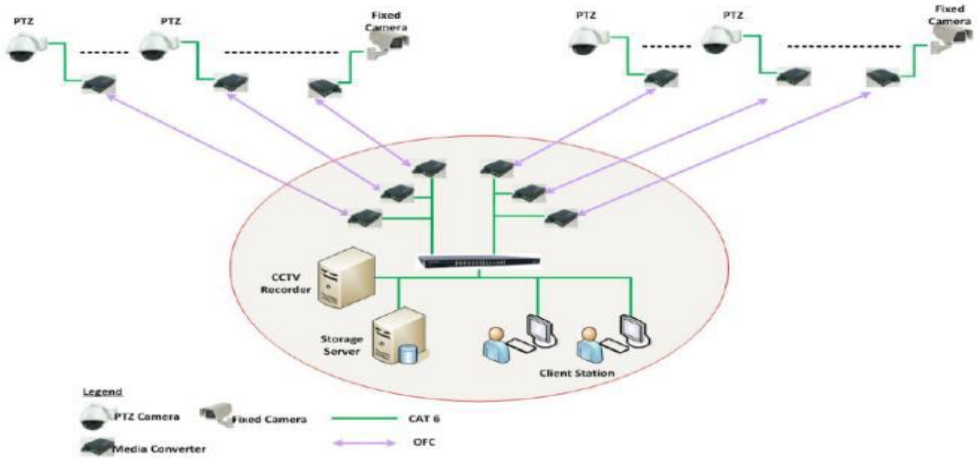

Figure 2: Here we are using only webcams in order to recognize faces.

4.2 Methodology

This system is based on using webcams in order to recognize faces. After the person is recognized further authentication is done by using speech to text module. Firstly the person has to speak out his/her ID (A unique code ID is allocated to authorized user), this ID is matched in the database and the text to speech module is used. A set of questions are being stored in the grammar and "what is your password?" question is asked to the user. After the ID and Password authentication, an authorized user message is being displayed.

The study further continues for navigation, which is used to track the location of a person by recognizing the face in various CCTV cameras installed at various locations

in the premises. Here, we are building a client server environment, where simultaneous CCTV images being stored at the server and sent to the client. The location may be analyzed and added to confirm and to make a trajectory of the person according to his/her movement within the premises.
To add better functionality to the system, a last seen feature is added which in case of no match found , the client can ask for last one-minute recording of the multiple CCTV cameras. The study would then result for any head and chief of the institute to track students and unknown people in the premises continuously to have a much safer environment around.

5. Conclusion

We have highlighted the importance of using image processing as a pre-processing step to well-known methods of face recognition and discussed the possibilities of combining face space dimensions of multiple systems in an attempt to utilize the advantages offered by numerous image processing techniques. This paper made a clear and simple overview of working of speech to text system (STT) in step by step process. There are many speech to text systems (STT) available in the market and also much improvisation is going on in the research area to make the speech more effective, natural with stress and emotions. This proposed system is getting ready to be implemented in many large scale sectors for their enhanced security. It can also be implemented in areas of maximum threat. This system is used in restricted areas where only the authenticated persons are allowed. The concept of the proposal is the intrusion detection by facial recognition. Since this project is very expensive and difficult to implement in a small scale by students we have showed the simulation results using C#.

References

[1] T. Heseltine, N. Pears, J. Austin, “Face Recognition: A Comparison of Appearance-based Approaches”, Proc. Of the 7th Int. Conf. on Digital Image Computing: Techniques and Applications, Vol. 1, pp. 59-69, 2003.
[2] International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-2, Issue-1, March 2012.
[3] International Conference on Intelligent Computational Systems (ICICS'2012) Jan. 7-8, 2012 Dubai.
[4] Mahalakshmi. R,Manjula. M, Saranya. S, Vaishnavi. P, Shalini. J., Arasa Kumar. R, “Intrusion Detection by Facial Recognition using CCTV Cameras with Video Management System“.
[5] Nishant Allawadi, Parteek Kumar, “A Speech-to-Text System”.