Face Recognition from One Sample per Person

Swathi Chandra T R, Priya Thomas

***Abstract*— As one of the most visible applications in computer vision communication, face recognition (FR) has become significant role in the community. In the past decade, researchers have been devoting themselves to addressing the various problems emerging in practical FR applications in uncontrolled or less controlled environment. In many practical applications of FR (e.g., law enforcement, e-passport, ID card verification, etc.), there is only one sample per person. Face recognition (FR) with a One Sample per Person is a very challenging problem due to the lack of information to predict the variations in the query image. The number of training samples per person will greatly affect the performance of face recognition. In the case of One Sample per Person, existing manifold learning algorithms cannot be directly applied, since those algorithms need multiple training samples per person to represent the query face. To address the problem of One Sample per Person, we propose in this paper a novel multi manifold analysis method by learning discriminative features from each manifold. In this technique, each recorded face image is first detected. Detected face image is then partitioned into several non overlapping patches to form an image set where each set contains images belonging to the same subject. By modelling each image set as a manifold, we formulate the One Sample per Person face recognition problem as the computation of the distance between two manifolds and learn multiple feature spaces to maximize the manifold margins of different persons. Finally, we present a reconstruction-based manifold-manifold distance where identification is achieved by seeking the minimum manifold-manifold distance (MMD) from the probe to the gallery of image sets.**

***Index Terms*— Face recognition with image sets, single training sample per person, manifold–manifold distance**

I. Introduction

Within the past two decades, face recognition has received increasing attentions owing to its wide range of applications, such as society security, e-passport and ID card verification, human-computer interface etc. Numerous methods have been proposed to improve face recognition[2][4]. These methods can be divided into two categories: geometric-based methods and appearance-based methods. Geometric-based methods describe a face image by the relation of components such as eyes, mouth and nose. Appearance-based methods represent face by both shape and texture information. In recent years, appearance-based methods have become dominant methods for face recognition. Most appearance-based methods commonly require many training samples as possible for each person. So, performance of appearance-based methods might be heavily affected by the number of training samples enrolled for each person. More specifically, intra-personal and inter-personal variations would be estimated with large bias in case of only few samples for each person. Even worse, if only single sample per person is available, many methods, such as LDA [3] and PCA [4], will degrade seriously or even fail to work due to the lack of samples to compute the within-class scatter. This is the so-called One Sample per Person problem [1]. As a great challenge, the One Sample per Person problem has become a big obstacle to many real-world applications, such as e-passport, id card verification, since, in these scenarios, it is generally impossible to collect more than one sample per person. To solve this problem, many methods have been developed recently [4][5][11].

Given two face images, the similarity between the same semantic patches of these two images is usually higher than that of different semantic patches of the same image. Hence, when a face image is represented by a manifold there is high overlapping between these manifolds and the distances between patches at the same location of different images are usually smaller than those at different locations of the same image. Hence, a key challenge to the STSPP face recognition problem is to extract discriminative features such that the points in the same manifold become closer and those in different manifolds are far apart. To achieve this goal, we propose in this paper a novel multi manifold analysis method to learn the local features to maximize the manifold margins of different persons so that more information can be exploited for recognition. In the recognition phase, we propose a reconstruction-based manifold-manifold distance to identify the unlabeled subjects.

The remainder of this paper is organized as follows. Section II describes the proposed method with details about its basic idea, the formulation and the implementation. Section III discuss about the results and analysis of the method. Finally, conclusion is given in the last section.

II. PROPOSED APPROACH

The Facial recognition system mainly consists of two phases: Enrollment phase and the Identification phase. During the first phase features are extracted from the input face given by the user by a process called Feature extraction, and are stored in the database. The collection of all such enrolled models is called facial database.

The proposed method aims to improve the performance of

***Manuscript received March, 2014*.**

***Swathi Chandra T R**, Computer Science and Engineering, KMCT College of Engineering, Calicut, India.,*

***Priya Thomas**, Computer Science and Engineering, KMCT College of Engineering, Calicut, India.*

ISSN: 2278 – 7798

CBIR using wavelet decomposition by haar wavelet. After that features are extracted using F-norm theory. K-mean clustering is used to form the cluster of images and similarity matching is done using F-norm theory. CBIR is used to get the visual content of the image and search for the image that has highest similarity. The organization of the system is shown below diagrammatically through the figures

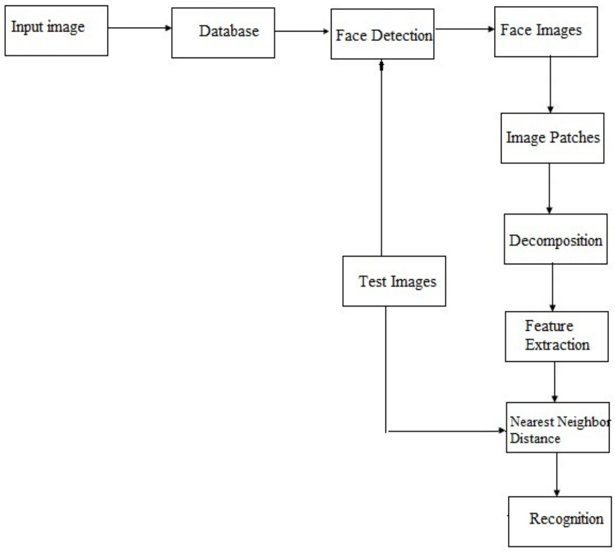

Fig.1: Structure of the Proposed approach

First step in the proposed approach is to input images to the system and save it to the database. Segmentation and representation of specific objects from input images by removing the background are the next step in the system. Face detection is the process that determines the locations and sizes of human faces in images. It detects facial features and ignores anything else. This is essentially a segmentation problem and in practical systems, most of the effort goes into solving this task. OpenCVs face detector [1] uses a method called the Viola-Jones method also known as the Haar cascade classifier.

The features that Viola and Jones used are based on Haar wavelets. These classifiers are single square waves (one high interval and one low interval). In two dimensions, a square wave is a pair of adjacent rectangles - one light and one dark. The presence of a Haar feature is determined by subtracting the average dark-region pixel value from the average light-region pixel value. If the difference is above a threshold (set during learning), that feature is said to be present.

To determine the presence or absence of hundreds of Haar features at every image location and at several scales efficiently, Viola and Jones used a technique called an Integral Image. In this case, the integral value for each pixel is the sum of all the pixels above it and to its left. Starting at the top left and traversing to the right and down, the entire image can be integrated with a few integer operations per pixel. To select the specific Haar features for use, and to set threshold levels, Viola and Jones used a machine-learning method called AdaBoost.

AdaBoost combines many weak classifiers to create one "strong" classifier. This combined series of classifiers as a filter chain is especially efficient for classifying image regions. Each filter is a separate AdaBoost classifier with a fairly small number of weak classifiers.

In the next step the images are divided into small parts known as patches. Next step is to read rgb values of the image. After creating red, green and blue matrices image is decomposed using haar wavelet transformation. Feature is extracted by using F-Norm theory. K-mean clustering is used to form the cluster of images and similarity.

Let $P = [P_1, P_2, .. P_n]$ be the training set, where P_j is the training image of the j^{th} person with a size of $a \times b$, $1 \leq j \leq N$, N is the number of face images in the training set. The first is to divide each image P_j into T non overlapping local patches with an equal size. Then store these patches to each manifold.

Haar wavelet transform [1] has been used as an earliest example for orthonormal wavelet transform with compact support. The Haar Wavelets are orthogonal, and their forward and inverse transforms require only additions and subtractions. It makes that it is easy to implement them on the computer. Haar Wavelets decomposes each signal into two components, one is called average (approximation) or trend and the other is known as difference. The Haar equations to calculate an average a_i and a wavelet coefficient c_i from an odd and even element in the data set S are:

$$a_i = \frac{s_i + s_{i+1}}{2} \qquad c_i = \frac{s_i - s_{i+1}}{2}$$

Feature extraction involves simplifying the amount of resources required to describe a large set of data accurately. When performing analysis of complex data one of the major problems stems from the number of variables involved. Analysis with a large number of variables generally requires a large amount of memory and computation power or a classification algorithm which over fits the training sample and generalizes poorly to new samples. Feature extraction is a
general term for methods of constructing combinations of the variables to get around these problems while still describing the data with sufficient accuracy. In the proposed approach feature vectors are extracted using F-Norm theory.

Next clustering process is done using K-means clustering. This non heirarchial method initially takes the number of components of the population equal to the final required number of clusters. In this step itself the final required number of clusters is chosen such that the points are mutually farthest apart. Next, it examines each component in the population and assigns it to one of the clusters depending on the minimum distance. The centroid's position is recalculated everytime a component is added to the cluster and this continues until all the components are grouped into the final required number of clusters.

In the recognition part, when a test image comes face image is detected first. Next the indexing process is done such as matrix decomposition and feature extraction. Finally the distance between the manifold of the test image and the trained images are calculated and recognize image with minimum manifold distance.

III. Results and Analysis

The application was tested during the research. It was important to have the same conditions with the testers to

ISSN: 2278 – 7798

prove how the program responded and to get conclusive results. Once their faces were added into the database the testers became database subjects. The application was tested during the research. It was important to have the same conditions with the testers to prove how the program responded and to get conclusive results. The results of the tests were favorable; around 8% of the tests were satisfactory. Some tests failed because of implementing problems. These tests were repeated, but the results of the investigation were not affected. The recognition rates dropped drastically by around 10% when trying to recognize the same subject again, after adding a new subject to the database.

The Face Detection System functioned with almost perfect accuracy, but many of the non-faces were erroneously detected as faces. This was acceptable because the false positive rate was only of 0.0001%. The application was able to detect, without problems, a number of faces at the same time from any source.

The face database was created without problems, so there was not needed to be conducted further work in this area. However, the application created was quite effective to detect and recognize faces in controlled environments.

IV. Conclusion

This paper presents relevant issues related to one sample per person problem in the area of face recognition. By using the STSPP face recognition method the face detection in the input image is done. In the next step, each enrolled face image is partitioned into several non-overlapping patches, and construct an image set for each sample per person. The next step is the haar wavelet transformation and then k means clustering is used. Finally test images are recognized by calculating manifold-manifold distance.

There are many factors that can improve the performance of the program. Good pre-processing of the images is very important for getting adequate results. It is important to have a lot of variation of conditions for each person so that the classifier will be able to recognize the person in different positions instead of looking for frontal images.

REFERENCES

[1] Felipe Ballester Toms, "Computer vision and Face Recognition".
[2] S. Kyungnam Kim,"Virtual Hypervisor: Face Recognition using Principle Component Analysis " in IEEE Transactions on image processing, Vol. 21, No. 10, October 2000
[3] Ruiping Wang et al, "Manifold-Manifold Distance and Its Application to Face Recognition With Image Sets ," in IEEE Transactions on image processing, Vol. 21, No. 10, October 2012 , 2003.
[4] Meina Kan, "Adaptive Discriminant Analysis for face recognition from Single Sample per Person ," in Department of Computer Science, University of Warwick, Coventry, CV4 7AL, UK ,
[5] Xingjie Wei et al., "Fixation and Saccade based Face Recognition from Single Image per Person with Various Occlusions and Expressions",2008.
[6] Matthew Turk, "Eigenfaces for Recognition," in Proc. of the International Computer Science and Engineering Department, Indian School of Mines , 2003.
[7] Suman Kumar Bhattacharya et al, "Face Recognition by Linear Discriminant Analysis," Lecture Notes in Computer Science , 2001.
[8] Jozef Ban et al.,"Enhancement of One Sample per Person Face Recognition Accuracy by Training Sets Extension ," in Proc. of OSDI , 2001.
[9] Yun Fu et al., "Classiffication and Feature Extraction by Simplexization ," in IEEE Trans-actions on Information Forensics and Security, Vol. 3, No. 1, March 2008 ,.
[10] De-An Huang, "With One Look: Robust Face Recognition Using Single Sample Per Person " in Proc. of the OSDI , 2002.
[11] Henry A. Rowley, "Human Face Detection in Visual Sceness," in Proc. Research Center for Information Technology Innovation, Academia Sinica, Taipei, Taiwan , 2005.

Swathi Chandra T R received the Bachelor of Information Technology degree with first class honors from Cochin University, India in 2010. She is currently pursuing her M.Tech degree in Computer Science Engineering from Calicut University, India. Her research interests include image processing.

Priya Thomas received the B.Tech from the Kannur University and M.Tech from the Calicut University in Computer Science and Engineering. She is currently working as an Assistant Professor in K.M.C.T College of Engineering, Calicut University. Her research interests include image processing.

ISSN: 2278 – 7798