IJRESM
International Journal of Research in Engineering, Science and Management (IJRESM)
www.ijresm.com
Volume-1, Issue-5, May 2018

Face Parts Detection Through HMM and SVD Coefficients

J. K. Poonam

M.Tech. Student, Department of Computer Science Engineering, UBDT, Devangere, India

Abstract**—Images containing faces are essential to intelligent vision-based human computer interaction, and research efforts in face processing include face recognition, face tracking, pose estimation, and expression recognition. However, many reported methods assume that the faces in an image or an image sequence have been identified and localized. To build fully automated systems that analyze the information contained in face images, robust and efficient face detection algorithms are required. Given a single image, the goal of face detection is to identify all image regions which contain a face regardless of its three-dimensional position, orientation, and lighting conditions. Such a problem is challenging because faces are nonrigid and have a high degree of variability in size, shape, color, and texture. The features are extracted using Singular value decomposition method (SVD). A HMM (Hidden Markov Model) is used to train these features of the images from the given set of data. These resultant HMM's then compared with the HMM of the available face image to identify the person. This method gives good results compared to other 2D face recognition systems which are affected by pose variations and intensity of the 2D image.**

Index Terms—Adaboost method, Hidden markov model, Independent component analysis, Principal component analysis, Singular value decomposition

I. Introduction

Human face detection is gaining interest as an important research area with many applications. The applications are video conferencing, human-computer interaction, content based image retrieval, and automatic authorization etc. Face detection problem can be stated as, determining whether there are human faces in the image, and if here are, returning the location of each human face in the image, regardless of its position and lighting condition. In recent years, face recognition has concerned much attention. It has numerous applications in computer vision communication and regular access control system. Face detection is the elementary yet an important step towards automatic face recognition. However, face detection is not clear-cut because it has lots of variations of image look, such as pose variation (front, non-front), occlusion, image orientation, illuminating situation and facial appearance. Face detection is the middle of all facial analysis, e.g., face localization, facial feature detection, face recognition, face verification and facial expression recognition. Moreover, it is a fundamental technique for other applications such as content-based image retrieval, video conferencing, and intelligent human computer interaction. The objective of face detection is to find out whether or not there are any faces in the image and, if present, return the location and the extent of each face. While face detection is a trivial task for human vision, it is a challenge for computer vision due to the variations in scale, location, orientation, pose, facial expression, light condition, and various appearance features (e.g., presence of glasses, facial hair, makeup, etc.). With the ubiquity of new information technology and media, more effective and friendly methods for human computer interaction are being developed which do not rely on traditional devices such as keyboards, mice, and displays. Furthermore, the ever decreasing price, performance ratio of computing coupled with recent decreases in video image acquisition cost imply that computer vision systems can be deployed in desktop and embedded systems. 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. In the last five years, face and facial expression recognition have attracted much attention though they have been studied for more than 20 years by psychophysicists, neuroscientists, and engineers. Many research demonstrations and commercial applications have been developed from these efforts. 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). Facial expression, occlusion, and lighting conditions also change the overall appearance of faces. The human face is a dynamic structure with characteristics that can quickly and radically change with time. Face recognition is useful in many areas such as medical records, online banking, Passports, driver licenses , video surveillances , investigation , biometrics, access control, law enforcement, surveillance system, security systems, identification of criminals, verification of credit cards and so on. Unfortunately, many face features make development of facial recognition systems difficult. This difficulty is solved by the method called Principal Component Analysis. Face recognition systems have been conducted now for almost 50 years. Face recognition is one of the researches in area pattern recognition & computer vision due to its numerous practical applications in the area of biometrics, Information security, access control, law enforcement, smart cards and surveillance system. In order to design and develop a helpful and appropriate face recognition system several factors need to be take in hand. Is many different technique to achieve face detection? The categories are not exact, it can be quite different in other sources, and the aim of this description is to be pretty simple and well traceable. They are rule-based methods, they try to capture our facial features, and translate them into a set of rules. For example, a face usually has two symmetric eyes, and the eye area is darker than the cheeks. Facial features could be the distance between eyes or between the nose and mouth, and so on. The big problem of these methods is that it is difficult to build an appropriate set of rules. There could be many false positives if the rules were too general On the other hand, there could be many false negatives if the rules were too detailed. A solution is to build hierarchical knowledge based methods to overcome these problems. However, this approach alone is very limited. It's unable to find many faces in a complex image. Template matching methods is based on a hand-coded template. Not only face(s), but different features can be defined independently. For example, a face can be divided into eyes, face contour, nose and mouth. The advantage of this method is its simplicity, it's easy to implement. However, there are a lot of types of faces, so it is impossible to define a template, which can fit all of the faces. Also an image can be recorded in many angles, and the template can't tolerate it. The templates in appearance-based methods are learned from the examples in the images. In general, appearance based methods rely on techniques from statistical analysis and machine learning to find the relevant characteristics of face images.

II. Literature survey

Face recognition is an important research topic, this is because face recognition has numerous practical applications such as bankcard identification, access control, Mug shots searching, security monitoring, and surveillance system. Face recognition system can be formulated as: given an image, identify or verify one or more persons in the image by comparing with faces stored in a database. Some image processing techniques extract feature points such as eyes, nose, and mouth and use it as input data for the application. Various approaches have been proposed to extract these facial points from the images. This section gives an overview of various methods of face recognition.

1) Geometry Based Technique:

Geometrical feature matching techniques are based on the computation of a set of geometrical features from the face image. The total geometrical configuration is 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. In geometry based technique features are extracted using the size and the relative position of important components of images. First the direction and edges of important component is detected and then feature vectors from these edges and direction are prepared. In some methods the grayscales difference of unimportant components and important components are found, the important components of a face are eyes and eyebrows, nose and mouth. By using this data, Haar-like feature block is used in Adaboost method to change the grayscales distribution into the feature, these features are then further used for comparison. Geometrical feature matching technique used by Kanade in is based on the extraction of a set of geometrical features forming the picture of a face. Their system achieved 75 percent recognition rate on a database of twenty persons, using two images per person; one for training and the other for test. Mark Nixon presented a geometric measurement for eye spacing with the Hough transform technique to detect the instance of a circular shape and of an ellipsoidal shape. The result of this paper illustrate that it is possible to derive a measurement of the spacing by detection of the position of both the iris. All these techniques require image to be properly captured. Hence the image used for these methods must have the cooperation of the person whose image is to be captured, which in many cases is not possible. These techniques require a threshold for comparison, which may adversely affect the achieved performance.

2) Template Based Matching:

In template matching based techniques the test image is represented as a two-dimensional array of intensity values and is compared using a suitable metric, such as the Euclidean distance, with a single template representing the whole face. There are many other ways of template matching on face recognition. More than one face template from different viewpoints can also be used to represent an individual's face. A face from a single viewpoint can also be represented by a set of multiple distinctive smaller templates. Another technique described will extract facial feature based on the previously designed templates using appropriate energy function and the best match of template in facial image yield the minimum energy. Methods have been proposed by Yuille for detecting and describing features of faces using deformable templates. In deformable templates the feature of interest, for example an eye, is described by a parameterized template. These parameterized templates enable a priori knowledge about the expected shape of the features to guide the detection process. An energy function is defined to link peaks, edges and valleys in the image intensity with corresponding properties of the template. After that the template matching is done with the image, by altering its parameter values to minimize the energy function, thereby deforming itself to find the best fit. For the descriptor purpose final parameter value is used. In the template based eye and mouth detection, first an eye template is used to detect the eye from image. Then a correlation is found out between the eye templates with various overlapping regions of the face image. Eye region has a maximum correlation with the template. The Yuille, Fischler and Elschlager in, measured features automatically and described a linear embedding algorithm that used local feature template matching and a global measure of fit to find and measure facial features. These algorithms require a priori template modeling, in addition to their computational costs, which clearly affect their performance. Genetic algorithms have been proposed for more efficient searching times in template matching.

3) Appearance Based Technique:

This approach processes the image as two dimensional patterns. The concept of feature in this approach is different from simple facial features such as eyes and mouth. Any extracted characteristic from the image is referred to a feature. This method found the best performance in facial feature extraction because it keep the important information of image and reject the redundant information. Method such as principal component analysis (PCA), independent component analysis (ICA), DCT or Singular Value Decomposition (SVD) is used to extract the feature vector. The main purpose of doing this is to reduce the large dimensionality of observed variable to the smaller dimensionality of independent variable without losing much information. This technique would be later the foundation of the proposal of many new face recognition algorithms. In PCA analysis high order dependencies exist and this is the disadvantage of this method because much information may contain in the high order relationship. While other method ICA uses technique independent component analysis which not only uses second-order statistics but also uses high order statistics. In DCT the set of coefficients is very large and hence it becomes complex to analyze and store such data. Whereas compared to all these methods SVD has less number of important coefficients and gives best result with low complexity. Some problems with ICA method is that it requires image matrices to be transformed into vectors, which are usually of very high dimensionality and this causes high computational cost and complexity. This approach uses skin color to isolate the face area from the non-face area in an image. Any non-skin color region within the face is viewed as a candidate for eyes or mouth. The performance of such techniques on facial image databases is rather limited.

4) Knowledge-Based Top-Down Methods:

In this approach, face detection methods are developed based on the rules derived from the researcher's knowledge of human faces. It is easy to come up with simple rules to describe the features of a face and their relationships. For example, a face often appears in an image with two eyes that are symmetric to each other, a nose, and a mouth. The relationships between features can be represented by their relative distances and positions. Facial features in an input image are extracted first, and face candidates are identified based on the coded rules. A verification process is usually applied to reduce false detections. One problem with this approach is the difficulty in translating human knowledge into well-defined rules. If the rules are detailed (i.e., strict), they may fail to detect faces that do not pass all the rules. If the rules are too general, they may give many false positives. Moreover, it is difficult to extend this approach to detect faces in different poses since it is challenging to enumerate all possible cases. On the other hand, heuristics about faces work well in detecting frontal faces in uncluttered scenes. Yang and Huang used a hierarchical knowledge-based method to detect faces. Their system consists of three levels of rules. At the highest level, all possible face candidates are found by scanning a window over the input image and applying a set of rules at each location. The rules at a higher level are general descriptions of what a face looks like while the rules at lower levels rely on details of facial features. A multi-resolution hierarchy of images is created by averaging and subsampling.

III. Proposed Methodology

In the previous section we have discussed about various methods of face recognition, this section gives us the detail information about the Singular value decomposition and Hidden markov model which is used in the proposed method of face recognition.

1) Singular value decomposition:

Singular value decomposition (SVD) can be looked at from three compatible points of view. On one hand, one can see it as a method for transforming correlated variables into a set of uncorrelated ones that better expose the various relationships among the original data items. At the same time, SVD is a method for identifying and ordering the dimensions along which data points exhibit the most variation. The third way of viewing SVD is that once we have identified where the most variation is, it is possible to find the best approximation of the original data points using fewer dimensions. Hence, SVD can be seen as a method for data reduction. The basic idea behind SVD is taking a high dimensional, highly variable set of data points and reducing it to a lower dimensional space that exposes the substructure of the original data more clearly and orders it from most variation to the least. SVD is based on a theorem from linear algebra which says that a rectangular matrix 'A' can be broken down into the product of three matrices- an orthogonal matrix 'U', a diagonal matrix' S', and the transpose of an orthogonal matrix 'V' . The theorem is usually presented as given in equation.

$$A_{m\times n} = U_{m\times n}S_{m\times n}V^{T}{}_{m\times n} \quad (1)$$

Where, $U^T U=I$ & $V^T V=I$

The columns of U are orthonormal eigenvectors of AA^T, columns of V are orthonormal eigen vectors of A^TA and S is a diagonal matrix containing the square roots of eigen values from U or V in descending order.

2) Hidden Markov Model:

Hidden Markov Models (HMMs) are widely used in pattern recognition applications, most notably for speech recognition but can be extended to face recognition systems. A process in the HMM class can be described as a finite-state Markov Chain with a memory less output process which produces symbols in a finite alphabet. However, from the perspective of an observer who knows the parameters of some representation of the process and is able to observe the output symbols but not the internal states, things look different. For some processes there are infinitely many distinct states of such an observer's knowledge about the status of the process. This knowledge is defined in terms of conditional distributions on future symbols. This is the sense in which there can be infinitely many states. These states are more relevant than the original finite set of states. Hence to study the process, these states are selected since they allow for optimal prediction. Hidden Markov Model is a Markov Chain with an associated output mechanism which takes either states or transitions between states to either symbols or distributions on symbols. Hidden Markov Models are useful in modelling one dimensional data in face finding, object recognition and face recognition. HMM is associated with non-observable hidden states and an observable sequence generated by the hidden states individually.

3) Algorithm of Proposed Method:

1. The image from the database is obtained and converted to gray scale.
2. The image is resized to around 50 percent of its original size so as to reduce the computational complexity.
3. In order to compensate the flash effect and reduce the salt noise, a nonlinear minimum order static filter is

used. The filter has a smoothing role and reduces the image information.

4. After all the preprocessing on the image, the image is divided into blocks having an overlap of around 75 percent. The number of blocks extracted from each image is given by:

$$T = \left\{\frac{H-L}{L-P}\right\} + 1 \quad (2)$$

Where, H=height of the image
P=L-1; L= overlap size.

5. SVD for each block obtained in the previous step is calculated and the value for U_{11}, S_{22}, S_{11} are found.
6. Each U, S & V value is then given to the HMM model for training and an HMM model for a person is generated.
7. Similarly HMM model for each person is generated and saved in the database for further comparison.
8. While testing, all the steps are repeated and an HMM model for the query image is found. This model is then compared with the available database to find the probable match.

IV. Results

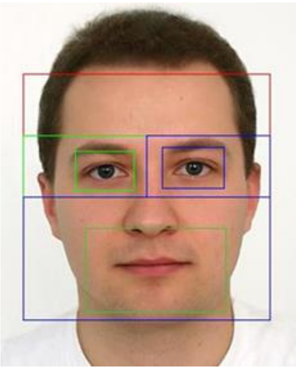

Fig. 1. Facial feature extraction

A successful face recognition system depends heavily on the feature extraction method. One major improvement of our system is the use of SVD coefficients as features instead of gray values of the pixels in the sampling windows (blocks). We use a sampling window of five pixels height and 64 pixels width, and an overlap of 80% in vertical direction (L=5, P=4). As we mentioned in previous section the observation vector is composed of large number of L×W (L=5 and W=64) blocks, each containing 320 pixels. The number of blocks of a 64×64 image is 60. Here one can simply see that how long the observation vectors will be. Using pixels value as features describing blocks, increases the processing time and leads to high computational complexity in training and recognition procedure. In this paper, we compute SVD coefficients of each block and use them as our features.

V. Conclusion and Future Scope

A fast and efficient system was presented. Images of each face were converted to a sequence of blocks. Each block was featured by a few number of its SVD parameters. Each class has been associated to hidden Markov model as its classifier. The evaluations and comparisons were performed on the two well-known face image data bases; ORL and YALE. In both data base, approximately having a recognition rate of 100%, the system was very fast. This was achieved by resizing the images to smaller size and using a small number of features describing blocks. In the future, we will focus on the use of larger and more complicated databases to test the system. For these complicated databases it is simply expected that all the previous methods will not repeat such efficiency reported in the papers. We will try to improve the feature extraction and the modeling of the faces. The use of 2-D HMM or more complicated models may improve the system performance.

References

[1] C. Zhang and Z. Zhang, “A survey of recent advances in face detection. Technical report,” Technical report, Microsoft Research, 2010.
[2] Ming-Hsuan Yang, D. J. Kriegman and N. Ahuja, "Detecting faces in images: a survey," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 1, pp. 34-58, Jan 2002.
[3] H. M. Naimi and P. Davari, “A New Fast and Efficient HMM-Based Face Recognition System Using a 7-State HMM Along With SVD Coefficients, *Iranian Journal of Electrical & Electronic Engineering*, vol. 4, no. 1 & 2, pp. 46-57, January 2008.
[4] N. Bhoi and M. N. Mohanty, “Template Matching based Eye Detection in Facial Image,” *International Journal of Computer Applications*, vol. 12, no.5, pp. 15-18, December 2010.
[5] K. Delac, M. Grgic and S. Grgic, “Independent Comparative Study of PCA, ICA, and LDA on the FERET Data Set,” *International Journal of Imaging Systems and Technology*, vol. 15, no. 5, pp. 252-260, 2006.