ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 4, Issue 11, May 2015

Face and Expression Recognition Techniques: A Review

Rishin C. K, Aswani Pookkudi, A. Ranjith Ram
Advanced Communication & Signal Processing Laboratory, Department of Electronics & Communication engineering, Government College of Engineering Kannur, Kerala, India.

***Abstract*—in this review paper, we discuss various techniques for face and expression recognition. Finally, we propose a Hidden Markov Model (HMM)-based face recognition system and Principal Component Analysis (PCA) based expression recognition system. The person is identified using 7- state HMM as classifier and Singular Values Decomposition (SVD) coefficients as features for face recognition. Using quantized SVD as feature describing blocks of the image, each face is considered as a numerical sequence that can be easily modeled by HMM.**

***Index Terms*—Face Recognition, Hidden Markov Model, Singular Value Decomposition, Principal Component Analysis.**

.

I. INTRODUCTION

The task of facial expression recognition was addressed very anciently, in the era of the Greek philosophers (4th century BC) when they were trying to assess a stranger's character and personality based on their outlook and appearance, especially from their facial expression. A more recent scientific approach to this study was done by Paul Ekman [1] in the nineteen-sixties. He took photographs of men and women making a set of distinctive faces and travelled to Brazil, Argentina, Japan, jungles of Papua New Guinea, with those photographs. To his amazement, everywhere people agreed on what those expressions meant in those photographs. It was a ground breaking discovery that time, which establishes the study of facial expression analysis. Now a days face and expression recognition system is a computer application and it is used for automatically identifying or verifying a person and his expressions from a digital image or a video frame from a video source. This is done by comparing the selected facial features form the image and a facial database. In face analysis, a good face descriptor should have a high variance among classes (between different persons or expression), but little or no variation within classes (same person or expressions in different conditions). The face and expression recognition system deals with the human resources in terms of psycho-logical approach. In automatic face and expression analysis, higher accuracy with reasonable speed still remains a great challenge.

The face and expression recognition system is divided into four major steps as given in Fig. 1. They are

1) Face detection
2) Normalization
3) Feature extraction
4) Classification

Face detection and normalization phase detects the face and lighting effects are reduced to some extent. The next step is feature extraction which extracts the features and irrelevant features are eliminated in feature selection process. Final step is classification where the facial expressions are classified in to six basic emotions shown in Fig. 2. Generally, there are two techniques in the facial expression recognition process and the first technique is based on facial feature and the other technique considers the holistic view of the recognition problem.

A. Holistic Approach

In holistic approach or global feature, the whole face region is taken into account as input data into face detection system. Examples of holistic methods are eigenfaces (most widely used method for face recognition), probabilistic eigenfaces, fisher faces, support vector machines, nearest feature lines (NFL) and independent-component analysis approaches. They are all based on principal component-analysis (PCA) techniques that can be used to simplify a dataset into lower dimension while retaining the characteristics of dataset.

B. Feature-based Approach

A feature-based approach to face recognition in which the features are derived from the intensity data without assuming any information of the face structure is presented. The feature extraction is biologically provoked, and the locations of the facial appearance often correspond to salient facial features such as the eyes, nose etc. Then these features are segmented and then it can be used as the input data for structural classifier. The techniques like dynamic link architecture, pure

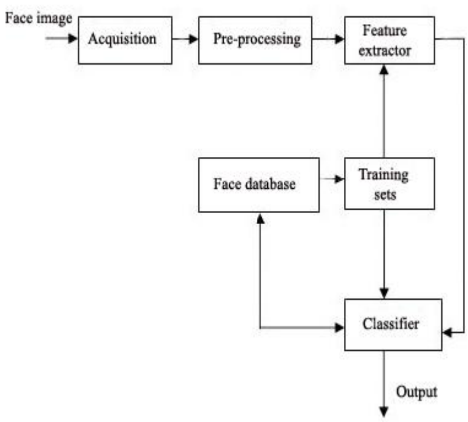

Fig. 1. Phases of face and expression recognition system

Geometry and hidden Markov model (HMM) are put under this category. During the survey it have been studied that eyes, mouth, and nose are the prominent features for face recognition.

C. Hybrid approach

Hybrid approach is a combination of above two mentioned approaches. The idea of this method comes from how human vision system perceives both local features and whole face. The methods like modular eigenface; hybrid local feature, shape normalized, and component based methods are used in hybrid approach. The human facial features have the great significance in the face recognition process. These facial features have the distinguishing characteristics which are not present in other facial components such as forehead, cheeks and chin. The local minima of a facial image are the eyes, the mouth, and the nostrils while the local maximum is the tip of the nose. The rest of the paper is organized as follows. In Section II, we present the current state of the art related to Hidden Markov Models, Singular Value Decomposition and Principal Component Analysis. In Section III, we discuss the various methods of Face and Expression Recognition. In Section IV we propose a system and in Section V we discuss its applications. Finally Section VI summarizes this paper with some concluding remarks.

II. CURRENT STATE OF THE ART

There are so many literature found in area of face and expression recognition. One of the methods developed and perhaps one of the most popular one to detect and track faces is the Kanade-Lucas-Tomasi tracker [2]. In earlier studies Kanade and Lucas developed a feature extraction algorithm [3] which matches two images for stereo matching and assumed that the second frame in a continuous frame of Images are a translation of the first one because of the small inter-frame motion.

Fig. 2. Facial expressions of six basic emotions.

Their implementation can successfully determine the distance to the object from camera and also can calculate brightness, contrast and five other camera parameters. In the presence of human supervision, the system worked really well, but their procedure also conjures errors and Tomasi et. al. updated and developed [2] the feature extraction algorithm on their own which iterates a few iterations over the basic solution that converges to a fast and simple solution. They define a feature as good, based on how well they can track that feature. Since by construction, their good feature selection criteria has become the optimal one. In 2004, Viola and Jones [4] developed a learning method based algorithm to detect frontal view faces. Their method is based on the Ada Boost learning algorithm and was found to be very fast and accurate. They computed an integral image from a source image using only a few operations per pixel based on a set of features they want to detect. After this computation, these features can be computed at any scale or location in constant time. After the detection of the features, they built a simple and efficient classifier by selecting a small number of important features from a huge library of potential features using AdaBoost, a learning program. But the problem they faced was that the feature space was rather large, far larger than the number of pixels. So in order to ensure a fast classification, they learning process must exclude a large majority of the available features. So they used AdaBoost to constraint each classifier to depend on a single feature. So the classification process actually becomes a visual feature detection process, which is also very fast and accurate. Recognizing facial expression is a complex task to com-plete and therefore several limitations are exists such as lightning condition, Age, similar expression type. As per Me Arabian et. al. [5], 55% communicative cues can be judge by facial expression; hence recognition of facial expressions became a major modality. For example, Smart Devices like computer/robots can sense/understand the human's intension from their expression then it will helpful to the system to assist them by giving suggestions or proposals as per their needs. The facial expressions under examination were defined by psychologists as a set of six basic facial expressions (anger, disgust, fear, happiness, sadness, and surprise). In order to make the recognition procedure more standardized, a set of muscle movements known as Facial Action Units (FAUs) that produce each facial expression, was created, thus forming the so called Facial Action Coding System (FACS) [6]. These FAUs are combined in order to create the rules responsible for the formation of facial expressions as proposed in [7].

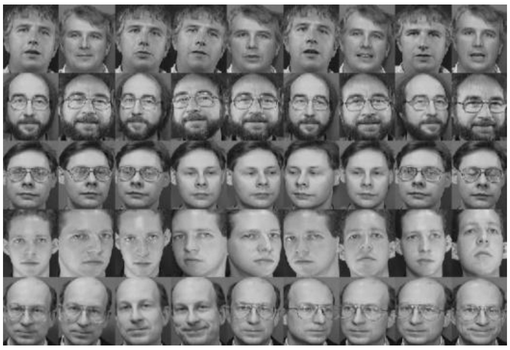

Fig. 3. Five examples of ORL face database

For the practical work over facial expression recognition, there is a need to have scientific research datasets for different subjects. The dataset which is having more uncontrollable conditions, it is a good choice. The uncontrollable conditions are like pose, occlusion, expression, illumination, expression variation, etc. There are many datasets presented to test methods of facial expression recognition. Some datasets are paid basis and some are publically available online. In some datasets preprocessed images are given for learners. In

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 4, Issue 11, May 2015

datasets, one person has different samples images. The datasets like FERET, Extended YaleB, CMU-PIE, AR, Cohn Kanade, ORL (as given in Fig. 3), and Indian Face database, Japanese Female Facial Expression JAFEE, etc. From this all, FERET face database and CMU (PIE) pose, illumination and expression face database is the one which are de-facto standard and are very courageous to handle different problem domain. In contrast to FERET database, there are some common expression databases which is openly available that are Cohn-Kanade database sometimes stated as CMU-Pittsburg AU coded database which has posed expressions [8] and is not fit for spontaneous expressions. Similar posed expression database are AR face database [9], Japnese Female Facial Expression Database (JAFFE) [10] etc.

III. METHODS OF FACE AND EXPRESSION RECOGNITION

There are different face and expression recognition techniques that apply mostly to the frontal faces. This section gives an overview of all these techniques with their advantages and disadvantages. The methods which are used for the face recognition are Eigen faces (Eigenfeatures), Neural Networks, Principal Component Analysis, Hidden Markov Model, and Geometrical Feature Matching.

A. Principal Component Analysis (PCA)

Principal Component Analysis (PCA), also known as the Eigen face approach is one of the popular methods for facial expression recognition [11]. It effectively characterizes geometry of face. Face can be easily reconstructed by only considering small amount of information that can be obtained by using Eigen faces; PCA faces difficulty if other factors like lighting and view pointing.

B. Fishers Linear Discriminant

Under severe variation in facial expression and illumination Fishers Linear Discriminant (FLD) is more suitable. FLD reduces the scattering of projected sample since it is a class specific method [12]. Error rate is reduced when compared to PCA.

C. Global Eigen Approach using Color Images

Conventional facial expression recognition techniques like PCA, LDA etc uses only the luminance information in face images. Global Eigen Approach uses the color information in face images [13]. RGB color space does not provide any improvement in recognition rate. In HSV color space, H component is removed since it reduces recognition rate. YUV color space provides high recognition rate.

D. Color Subspace Linear Discriminant Analysis

The 1DLDA AND 2DLDA are extended in color space to improve the face recognition accuracy. A 3-D color tensor is used to generate color LDA subspace [14]. Horizontal unfolding increases the recognition rate for 2DLDA while vertical unfolding improves recognition rate for 2DPCA. The performance evaluation of various color spaces is not done.

E. Local Gabor Binary Pattern

Appearance based features are being used for face recognition since it encodes specific details about human faces. In this facial image is divided into sub blocks and similarities between sub blocks is obtained [15]. An important advantage of Local Binary Pattern (LBP) is its illumination tolerance. In Local Gabor Binary Pattern (LGBP) method, LBP is extracted from gabor filters for feature vector generation. LGBP achieves better performance than gabor filter method [16]. All these techniques have improved the performance of the facial expression recognition system considerably by incorporating color component and tensor concepts. But none of the color spaces like RGB provide provision for head pose and lighting variations. Therefore perceptual color space like CIELab and CIELuv is used in this paper to improve the performance in varying conditions. The color images are represented using 3D matrix. But 2D filtering operation is needed to be applied to a 3D array, which is complex in nature. As a solution to this problem, 2D filters are to be applied three times over the three color components in images. Instead of doing this, tensor concept is used. The filtering operation is directly applied to the tensor i.e generated from the color image. In [16] author introduces a Tensor Perceptual Color Framework (TPCF) where color image components are horizontally unfolded to 2D tensors using multi-linear algebra and tensor concepts. Log-gabor filters are used for feature extraction since it overcome the limitations of gabor filter based method. For feature selection mutual information quotient method is utilized. Multiclass linear discriminant analysis classifier is used for classifying the selected features. TPCF can effectively recognize the facial expressions under different illumination conditions and thus performance can be improved.

F. Hidden Markov model

Hidden Markov models (HMM) are another promising method that works well for images with variation in different lighting, facial expression, and orientation. HMM is a set of statistical models used to characterize properties of signals. It has very good performance in speech recognition and character recognition, where the data is 1-dimentional. The system being modeled is assumed to be a Markov process with unknown parameters, and the goal is to find hidden parameters from the observable parameters. The each state in HMM has a probability distribution over the possible output whereas each state in a regular Markov model is observable. In Nefians paper [19], the authors use HMM approach for face recognition based on the extraction of 2- dimensional discrete cosine transformation (DCT) feature vectors. The author takes advantage of DCT compression property for feature extraction. An image is divided by blocks of a

sub-image associated with observation vector. More details about HMM method are provided in the following sections. In HHM, there are unobservable Markov chain with limited number of status in the model, the observation symbol probability matrix B, a state transition probability matrix A, initial state distribution , and set of probability density functions (PDF). A HMM is defined as the triplets
λ= (A; B; π).

For frontal human face images, the important facial components appear in top to bottom order such as hair, forehead, eyebrows, eyes, nose, mouth, and chin. This still holds although the image rotates slightly in the image plane. Each of the facial region is assigned to one state in 1-D continuous HMM. The transition probability a_{ij} and structure of face model is illustrated in Fig. 4.

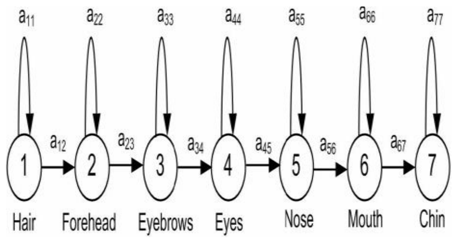

Fig. 4. A one dimensional HMM model with seven states for a face image with seven regions

G. Singular Value Decomposition

The Singular Value Decomposition (SVD) has been an important tool in signal processing and statistical data analysis. Singular values of given data matrix contain information about the noise level, the energy, the rank of the matrix, etc. As singular vectors of a matrix are the span bases of the matrix, and orthonormal, they can exhibit some features of the patterns embedded in the signal. SVD provides a new way for extracting algebraic features from an image. A singular value decomposition of a m n matrix X is any function of the form:

$$X = U \sum V^T$$

Where U (m×m) and V (m×m) are orthogonal matrix, and $\sum$ is a m×n diagonal matrix of singular values with components σ_{ij} = 0; i ≠ j and σ_{ij} > 0. Furthermore, it can be shown that there exist non-unique matrices U and V such that $\sigma_{1\geq}\sigma_{2\ldots\ldots} \geq 0$. The columns of the orthogonal matrices U and V are called the left and right singular vectors respectively; an important property of U and V is that they are mutually orthogonal [17]. The main theoretical property of SVD relevant to face image recognition is its stability on face image. Singular values represent algebraic properties of an image [18]. So because of these reasons and some experimental results, we find out that SVD is a robust feature extraction technique for face images.

H. Template Matching

We can exploit other face templates from different prospects in order to characterize single face. The complex-ity arises only during the extraction of template.

IV. PROPOSED SYSTEM

In the previous section we have discussed the current state of the art. The proposed method will be based on 7-state HMM as classifier and Singular Values Decomposition (SVD) coefficients as features for face recognition and Principal Component Analysis (PCA) for expression recognition. In this work the face database used is that of the PG batch of the Department of Electronics & Communication Engineering of Government College of Engineering Kannur. The data acquisition started using camera Nikon D300S in a controlled illumination at ACSP lab.

V. APPLICATIONS

With recent advancement in robotics and automated soft-wares, the requirement of a robust expression recognition system is more evident. As humans are in general responsive to each other’s emotional states, computers or automated systems must also gain this ability. With the advancement of human computer interaction study, researchers are bridging this gap between human and computer sensors. Video games consoles such as Kinect sensor or Wii can detect human movement and act accordingly connecting the physical world with the virtual world. Sleep detection sensors in automobiles can identify when a driver is sleepy and act accordingly to reduce the risk of accidents. Smart robots are being developed which can give company to human beings. Facial expression analysis will be very useful in all of these applications. The other applications include

- Information and forensics
- Multimedia applications
- Security applications
- Entry and exit verification
- Election accuracy

VI. CONCLUSION

We have discussed about the face and expression recognition systems. Face feature extraction is a major step for the efficient and effective representation of facial features.

The goal of the facial feature extraction is to find an efficient and effective representation of the facial images which would provide robustness during recognition process. This term paper propose an HMM based face recognition system and PCA for expression recognition. The database which is not having more uncontrollable conditions like pose, occlusion, expression, illumination etc. is a good choice.

REFERENCES

[1] M. Gladwell, “The naked face,” The New Yorker, pp. 38-49, August 2002.

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 4, Issue 11, May 2015

[2] C. Tomasi and T. Kanade, "Detection and tracking of point features," School of Computer Science, Carnegie Mellon Univ, April 1991.

[3] B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," Proc. of the 7th international joint conference on Artificial intelligence, 1981.

[4] P. Viola and M. Jones, "Robust real-time face detection," Inter-national journal of computer vision, pp. 137-154 2004.

[5] A. Mehrabian, "Communication without Words," Psychology Today, pp. 53-56, 1968.

[6] T. Kanade, J. Cohn and Y. Tian, "Comprehensive database for facial expression analysis," Proc. IEEE International Confer-ence on Face and Gesture Recognition, pp. 46-53, March 2000.

[7] M. Pantic and L. J. M. Rothkrantz, "Expert system for automatic analysis of facial expressions," Image Vis. Comput, pp. 881-905, August 2000.

[8] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," IEEE Transactions on Image Process, pp. 1635-1650, June 2010.

[9] T. Jabid, M. H. Kabir and O. Chae, "Local directional pattern (LDP) for face recognition," Proc. of the IEEE International Conference on Electronics, pp. 329-330, March 2010.

[10] T. Jabid, M. H. Kabir and O. Chae, "Robust facial expression recognition based on local directional pattern," ETRI J, pp. 784-794, 2010.

[11] L.Sirovich and M. Kirby, "Low Dimensional Procedure for Characterization of Human Faces," J. Optical Soc. Am, pp. 519-524, 1987.

[12] P. N. Belhumeur "Eigen faces vs. Fisher faces: recognition using class specific linear projection," IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 711-720, July 1997.

[13] L. Torres, J. Reutter, and L. Lorente, "The importance of the color information in face recognition," Proc. of the International Conference on Image Processing, pp. 627-631, 1999.

[14] M. A. O. Vasilescu and D. Terzopoulos, "Multilinear image analysis for facial recognition," in Proc. Int. Conf. Pattern Recognition., Quebec City, QC, Canada, pp. 511-514 August 2002.

[15] S. Moore and R. Bowden, "Local binary patterns for multi-view facial expression recognition," Comput. Vis. Image Understand, pp. 541-558, 2011.

[16] S. M. Lajevardi and H. R. Wu, "Facial Expression Recognition in Perceptual Color Space," IEEE Transactions on Image Processing, pp. 3721-3732, 2012

[17] Z. Hong, "Algebraic feature extraction of image for recognition", Pattern Recognition, pp. 211-219, 1991.

[18] V. C. Klema and A. J. Laub, "The Singular Value Decomposition: Its Computation and Some Applications," IEEE Transactions on Automatic Control, pp. 00189286, April 1980.

[19] A. V. Nefian,"Hidden markov models for face recognition," In proceedings, International Conference on Acoustics, Speech and Signal Proceeding, pp. 2721-2724, 1998.

AUTHOR BIOGRAPHY

Rishin. C. K. received the B-Tech degree in electronics and communication engineering in 2012 from Cochin university of Science and Technology kerala,. Currently doing M-tech degree in signal processing and embedded system at Kannur University Kerala

Awani Pookkudi received the B-Tech degree in electronics and communication engineering in 2012 from Kannur University. Currently doing M-tech degree in signal processing and embedded system at Kannur University Kerala

Dr. A. Ranjith Ram received the B- tech degree in electronics and communication at Calicut university**,** Graduated in Engineering at NSSCE Palakkad. Completed M-Tech from CET, Thiruvananthapuram. Recently got Doctorate Degree from IIT Bombay and currently work as Assistant Professor in ECE Dept. at GCE Kannur.