Scientific Journal of Impact Factor(SJIF): 3.134
e-ISSN(O): 2348-4470
p-ISSN(P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 2,Issue 4, April -2015

Various Techniques for Facial Detection and Expression Classification

Desai Ghata Sanjeevkumar[1], Pankaj Kumar Gautam[2]

[12]CSE Department, Parul Institute of Technology, Limda, Vadodara, India.

Abstract — *The facial expression is important in social interaction and social intelligence. Automatic recognition of facial expression is most important component in human-machine interface. This field contains various applications like user authentication, person identification, video surveillance, information security, data privacy etc. This paper includes the survey of various facial detection and expression recognition techniques.*

KEYWORDS- FACE DETECTION, FEATURE EXTRACTION, CLASSIFICATION, EXPRESSION RECOGNITION, FACIAL DATABASES

I. INTRODUCTION

In designing computerized environments recent challenge is to keep the human user at the core of the system. The computer needs to be capable of interacting with the user in a natural way. To recognize an emotional state, system should interpret non-verbal behaviour like: voice, body gesture and facial expressions. Among three, facial expression is the most natural means of communicating human emotions, intentions and opinions to each other. [1] Showed that 55% of the emotional information is conveyed by facial expression alone. Remaining voice tone and spoken words conveys 38% and 7% of the information respectively. [2] did a psychological research on facial expression and they concluded that there are six basic facial expressions which are universal. The six basic expressions include happiness, sadness, disgust, anger, surprise and fear.

II. SYSTEM STRUCTURE

The general architecture of face recognition system consists of four components as depicted in Figure 1. Face detection finds the face areas in the input image. If the input is a video, face detection is only performed on key frames and a tracking algorithm is applied on interval frames. This will be more efficient and also achieve better robustness. Face alignment is very similar to detection, but it is aimed at achieving a more accurate localization. In this step, a set of facial landmarks (facial components), such as eyes, brows and nose, or the facial contour are located; based on that, the face image is rotated, chopped, resized and even warped, this is called geometrical normalization. Usually the face is further normalized with respect to photometrical properties such as illumination and gray scale. Feature extraction is performed on a normalized face to provide effective information that should be useful for recognizing and classifying labels in which there is interest, such as identity, gender, or expression. The extracted feature vector is sent to a classifier and compared with the training data to produce a recognition output.

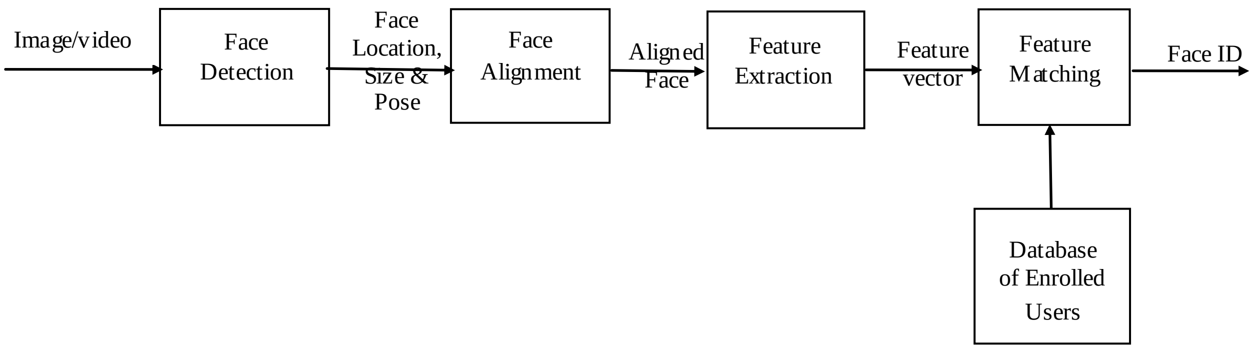

Figure 1: Face Recognition Processing Flow **[3]**

III. Face Detection

The first step in face detection is pre-processing. The reason is to obtain pure facial images with normalized intensity, uniform size and shape. The steps involved in converting a image to a normalized pure facial image for feature extraction is detecting feature points, rotating to line up, locating and cropping the face region using a rectangle, according to the face model. Detecting faces in a single image involves four methods:

2.1 Knowledge based

In this approach face detection methods are developed based on the rules derived from our knowledge of human faces. It is easy to come up with simple rules that describe the features of a face and their relationships. For example, a face usually appears with two eyes that are symmetric to each other, a nose and a mouth. The relationships between these features can be represented by their relative distances and positions. Faces are then detected by applying these codes rules to find features. Face candidates are detected after applying all the rules. A verification process is optionally used to increase avoid false detections. Problem with this approach is that it is difficult to translate our knowledge about faces into rules effectively. If the rules are detailed, it may fail to detected faces that do not pass all the rules. If the rules are too general, it may give many false positives.

2.2 Feature Invariant

The underlying assumption is based on the observation that human can detect faces effortless because facial features always exist in the image , regardless of the poses , viewpoints or lighting conditions. The assumption, if hold, are also the advantages of this approach since faces of different poses can be detected. On the other hand, the major problem with feature based algorithm is that the image feature can be badly corrupted due to illumination, noise or occlusion. Feature boundaries can be weakened by illumination, and shadows can cause numerous strong edges that render perceptual grouping algorithms useless.

2.3 Template Matching

In template matching, a standard pattern of a human face is stored first. Given an input image, the correlation values with several sizes of the standard pattern are calculated for the face contour, eyes, nose and mouth independently. From these correlation values of the face portions, the final determination is done for the existence of a face. This approach has the advantage of being simplistic. However, it has proved to be inadequate for face detection since it cannot deal with variation of faces in scale, pose and shape. Multiresolution, multiscale and deformable templates have been proposed achieve scale and shape invariance.

2.1.4 Appearance Based

Provide several key advantages and are widely used in face detection. Especially, as they allow one to learn the models from training data, the large amount of intra-class variation, expression, and pose can be accounted for in training by using a large training set.

IV. FEATURE EXTRACTION

There are two common types of facial features extraction techniques: geometric based methods and appearance based methods. Geometric based approach mainly depends on how we perceive an expression; i.e., based on shrinking/ widening of eyes, eyebrows, lips, nose etc. This approach mainly focuses on finding out the displacement/ angular changes of the feature points with respect to a neutral expression. But, the main drawback with this technique is its sensitivity noise and accumulation of errors while tracking those features which makes the accurate detection. Whereas, appearance features are less sensitive to noise and can encode micro patterns present in the skin texture of face which are very important information for expression recognition.
Various methods for feature extraction are [4]:

4.1 Principal Component Analysis

Principal Component Analysis (PCA), also known as the Eigen face approach is one of the popular methods for facial expression recognition. The major goal of PCA is to reduce the dimensionality for effective face indexing and retrieval. Also, PCA uses linear projection, which maximizes the projected sample scattering. In this, the identity of the person is the only varying factor.PCA faces difficulty if other factors like viewpoint, lighting are varied.

4.2 Fisher's Linear Discriminant

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

4.3 Independent Component Analysis

Both PCA and LDA generate spatially global feature vectors. But for effective facial expression recognition spatially localized feature vectors is needed. Therefore Independent Component Analysis (ICA) generates statistically independent basis vector. The average recognition rate is improved. But ICA is computationally expensive than PCA.

4.4 2Dimensional Principal Component Analysis

In PCA, feature extraction is done based on 1D vector. Therefore the image matrixes need to be transformed into vector. 2dimensional Principal Component Analysis (2DPCA) uses 2D matrix instead of 1D vector. The recognition rate of 2DPCA is higher than PCA. But the storage requirement for 2DPCA is higher than PCA since 2DPCA needs more coefficients for image representation.

4.5 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.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 colorspace provides high recognition rate.

4.6 Sub pattern Extended 2-Dimensional PCA

The recognition rate of PCA is low and has small sample size problem. For gray facial expression recognition, 2DPCA is extended to Extended 2DPCA. But E2DPCA is not applicable for color images. Therefore Sub pattern extended 2-Dimensional PCA (SpE2DPCA) is introduced for color face recognition. The recognition rate is higher than PCA, 2DPCA, E2DPCA and problem of small sample size in PCA is also eliminated.

4.7 Multilinear Image Analysis

Facial expression recognition needs different factors like pose, lighting, expressions to be considered. But the conventional PCA addresses only variations in single factor. Multilinear image analysis use multilinear algebra. In this, the concept of „Tensor faces‟ is used, which separates different factors underlying the formation of an image. Recognition rate is greater when compared to PCA approach. Color information is not incorporated in multilinear image analysis.

4.8 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. 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.

4.9 2D Gabor filter bank

The Gabor filtering is considered as one of the most important feature extraction technique in facial expression recognition. Gabor filter bank performs better in terms of recognition rate than the other methods like PCA, LDA etc. The major limitation of Gabor filter is its bandwidth limitation i.e. Maximum bandwidth is limited to one octave. Gabor filters loss high and low frequency information since it is band pass in nature.

4.10 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. 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.

V. CLASSIFICATION

Different classifiers are explained below:

5.1 Hidden Markov Model (HMM)

Hidden markov model is developed to classify the higher level emotions states like interested, unsure, disagreeing, encouraging and discouraging from the lower level emotions like neutral, joy, surprise and sad. An emotional indexing is modelled is used to understand the states of emotions thus it functions as database hence there is an one to one mapping is done between an facial emotions and expressions. The indexer receives the symbol and matches them against stored in the mind and then chooses symbols to represent concepts in the index. Hidden markov expert rule (HMER) is used for segmentation and to recognize emotional states from a set of video sequences. A classification framework is used for every incoming video frame and hence the facial expression recognizer identifies the head and actions which combine to form displays and thus HMER represents dynamic display and classification framework. HMER topology is constructed for four state emotional states like N represents neutral state and SU for surprise etc from which it is possible to transit the higher level emotion states. Here the probabilistic framework for modelling, time varying sequences and convergence of recognition computation runs in real time. The performance of recognizing the emotions unsure is 87% and disagreeing 78%.

5.2 Support Vector Machine

An SVM is a method of supervised learning for classification by using a hyperplane to divide the space into two regions, each region classifying one type of element. For training, from the negative training samples and positive training samples, a margin sep arator is defined by mathematical functions. Using this margin separator, the testing sample is considered as positive case or negative case. The SVM is an effective machine learning method when using a relatively small number of training data.

5.3 AdaBoost

AdaBoost is a short for Adaptive Boosting. It is a boosting learning algorithm to combine many weak classifiers into a single powerful classifier. It adaptively changes the weight for weak classifiers from the misclassified data sample. It is mainly used to fuse many binary classifiers into a strong classifier.
There are many other classifiers that can be used for the classification purpose.

VI. DATABASES

When benchmarking an algorithm it is recommendable to use a standard test data set for researchers to be able to directly compare the results. While there are many databases in use currently, the choice of an appropriate database to be used should be made based on the task given (aging, expressions, lighting etc). Another way is to choose the data set specific to the property to be tested (e.g. how algorithm behaves when given images with lighting changes or images with different facial expressions).

6.1 The Color FERET Database

The FERET program set out to establish a large database of facial images that was gathered independently from the algorithm developers. Dr. Harry Wechsler at George Mason University was selected to direct the collection of this database. The database collection was a collaborative effort between Dr. Wechsler and Dr. Phillips. The images were collected in a semi-controlled environment. To maintain a degree of consistency throughout the database, the same physical setup was used in each photography session. Because the equipment had to be reassembled for each session, there was some minor variation in images collected on different dates. The FERET database was collected in 15 sessions between August 1993 and July 1996. The database contains 1564 sets of images for a total of 14,126 images that includes 1199 individuals and 365 duplicate sets of images. A duplicate set is a second set of images of a person already in the database and was usually taken on a different day. For some individuals, over two years had elapsed between their first and last sittings, with some subjects being photographed multiple times. This time lapse was important because it enabled researchers to study, for the first time, changes in a subject's appearance that occur over a year.

6.2 SCface - Surveillance Cameras Face Database

SCface is a database of static images of human faces. Images were taken in uncontrolled indoor environment using five video surveillance cameras of various qualities. Database contains 4160 static images (in visible and infrared spectrum) of 130 subjects. Images from different quality cameras mimic the real-world conditions and enable robust face recognition algorithms testing, emphasizing different law enforcement and surveillance use case scenarios. SCface database is freely available to research community.

6.3 The Yale Face Database

Contains 165 grayscale images in GIF format of 15 individuals. There are 11 images per subject, one per different facial expression or configuration: center-light, w/glasses, happy, left-light, w/no glasses, normal, right-light, sad, and sleepy, surprised, and wink.

6.4 Cohn-Kanade AU Coded Facial Expression Database

Subjects in the released portion of the Cohn-Kanade AU-Coded Facial Expression Database are 100 university students. They ranged in age from 18 to 30 years. Sixty-five percent were female, 15 percent were African-American, and three percent were Asian or Latino. Subjects were instructed by an experimenter to perform a series of 23 facial displays that included single action units and combinations of action units. Image sequences from neutral to target display were digitized into 640 by 480 or 490 pixel arrays with 8-bit precision for grayscale values. Included with the image files are "sequence" files; these are short text files that describe the order in which images should be read.

6.5 Japanese Female Facial Expression (JAFFE) Database

The database contains 213 images of 7 facial expressions (6 basic facial expressions + 1 neutral) posed by 10 Japanese female models. Each image has been rated on 6 emotion adjectives by 60 Japanese subjects.

VII. CONCLUSION

Although humans recognize facial expressions virtually without effort or delay, reliable expression recognition by machine is still a challenge, in achieving optimal pre-processing, feature extraction or selection, and classification, particularly under conditions of input data variability. To attain successful recognition performance, most current expression recognition approaches require some control over the imaging conditions because many real-world applications require operational flexibility. In particular, research into automatic expression recognition systems capable of adapting their knowledge periodically or continuously has not received much attention. Robustness of expression recognition, against the variability of facial characteristics, would be difficult to achieve without incorporating adaptation in the recognition framework in future.

REFERENCES

[1] A. Mehrabian, Nonverbal communication, Aldine, 2007.

[2] P. Ekman, W. V. Friesen, and J. C. Hager, “Facial action coding system,” A Human Face, Salt Lake City,2002.

[3] Yuan Ren,” Facial Expression Recognition System,” Waterloo, Ontario, Canada, 2008.

[4] Ms.Aswathy.R, “A Literature review on Facial Expression Recognition Techniques”, IOSR Journal of Computer Engineering (IOSR-JCE), Volume 11, Issue 1, May. - Jun. 2013, PP 61-6.

[5] Anima Majumder, Laxmidhar Behera and Venkatesh K. Subramanian, "Local Binary Pattern Based Facial Expression Recognition Using Self-Organizing Map," International Joint Conference on Neural Networks (IJCNN), 2014.

[6] Brian C. Lovell, Shaokang Chen, Ting Shan,”Real-time Face Detection and Classification for ICCTV,”2007.

[7] G.Hemalatha, C.P. Sumathi, “A Study of Techniques for Facial Detection and Expression Classification”, International Journal of Computer Science & Engineering Survey (IJCSES) Vol.5, No.2, April 2014.

