Facial Emotion Recognition Techniques: A Survey

Namita Rathore¹ Rohit Miri²
¹²Department of Computer Science and Engineering
¹²Dr C V Raman Institute of Science and Technology

Abstract—Facial expression recognition system has turn into a most emphasizing research area since it plays a most important part in human-computer-interaction. The facial expression recognition system finds foremost application in areas like social interaction and social intelligence like in diverse surveillance systems, defense systems, substantiation or verification of individual like criminals etc...The face can articulate emotion sooner than people verbalize or even understand their posture. Thus there is apparent need of unfailing recognition and identification of facial expressions. In this paper a review of various techniques used in facial expression recognition like PCA (Principal component analysis), LDA (Linear discriminant analysis) etc has been carried out. We have formulated the Face expression recognition technique as different phases and the survey is represented accordingly in tabular form emphasizing on their performance.

Key words: PCA, LDA, DFAT, FACS

I. INTRODUCTION

Human communication has two main aspects; verbal (auditory) and non-verbal (visual). Facial expressions are an important component of interpersonal communication. Despite their non-verbal nature, they convey a lot of information about the person and the person’s affective state, intention and personality. Particularly for the recognition of the affective state, humans rely heavily on analyzing facial expressions [10, 18]. Facial expressions also support verbal communication due to their complementary nature to the acoustic side of the spoken words. Unlike humans, current computer systems can hardly recognize the affective state of a human user. The last decade has witnessed a trend towards an increasingly ubiquitous computing environment, where powerful and low-cost computing systems are being integrated into mobile phones, cars, medical instruments and almost every aspect of our lives. This has created an enormous interest in automatic processing of digital images and videos in a number of applications, including biometric authentication, surveillance, human-computer interaction, and multimedia management. Research and development in automatic face recognition follows naturally. Face recognition is a visual pattern recognition problem where a three-dimensional object is to be identified based on its two-dimensional image. In recent years, significant progress has been made in this area; owing to better face models and more powerful computers, face recognition system can achieve good results under constrained situations. However because face images are influenced by several factors: illumination, head pose, expression and so on, in general conditions, face recognition is still challenging. From a computer vision point of view, among all these “noises” facial expression may be the toughest one in the sense that expressions actually change the three-dimensional object while other factors, such as illumination and position, only affect imaging parameters. To get rid of expression noise², one first needs to estimate the expression of an image, this is called Facial Expression Recognition". Another, maybe more important motivation of facial expression recognition is that expression itself is an efficient way of communication; it’s natural, non-intrusive, and [12] has shown that, surprisingly, expression conveys more information than spoken words and voice tone. To build a friendlier Human Computer Interface, expression recognition is essential. The importance of facial expression system is widely recognized in social interaction and social intelligence. The system analysis has been an active research topic since 19th century. The facial expression recognition system was introduced in 1978 by Suwa et. al. The main issue of building a facial expression recognition system is face detection and alignment, image normalization, feature extraction, and classification. There are number of techniques which we use for recognizing the facial expression. Some of the researchers [1] introduced the system can recognize the different human gesture in color image. In this paper Viola and Jones describe the face detection technique using Add Boost Haar classifier. After performing the pre-processing operation the recognition is performed, the simplicity and robustness of the system is significant. Depending on threshold value the researchers system can recognize the facial expression. The approach of this system can be adapted to real time and it briefly describes the schemes of capturing the image and to recognize the gestures. Robert Axelrod has also shown the ability to recognize that they have met before and distinguish them from strangers is one of the bases for humans to form cooperation [3].

II. LITERATURE SURVEY

M. Pantic and L.J.M. Rothkrantz focused on a threefold, Modelling the facial motion and its intensity (i.e. dealing with face image sequences and AU intensity) will increase the overall performance of the system. Developing a Fuzzy Expert System for face action tracking and face action emotional classification will increase the quality of the system by allowing it to reason about the involved face actions according to the accuracy of the performed facial feature tracking. Designing and developing a learning facility, which will allow the user to define his/her own interpretation categories, will yield a broader and more realistic classification of the encountered expressions.

Philipp Michel and Rana El Kaliouby states that in their implementation correctly recognized expressions in 78% of trials, with subsequent improvements including selection of a kernel function customized to the training data boosting recognition accuracy up to 87.9%, as illustrated in Table 1. Incorporating further possible enhancements such as adjusting data to account for head motion or performing automatic SVM model selection is likely to yield even better performance and further increase the suitability of SVM-based expression recognition approaches in building affective and socially intelligent human-computer interfaces.
Pushpaja V. Saudagare, D.S. Chaudhari told that automatic facial expression recognition systems are overviewed. The neural network approach is based on face recognition, feature extraction and categorization. The approach of facial expression recognition method involve the optical flow method, active shape model technique, principle component analysis algorithm (PCA) and neural network technique. The approach does provide a practical solution to the problem of facial expression recognition and it can work well in constrained environment.

Akshat Garg, Vishakha Choudhary states that Principal Components Analysis is a method that reduces data dimensionality by performing a covariance analysis between factors. Thus Facial expression recognition or cognitive assessment can be done by comparing the principal components of default image or slice with the new/any respective subject.

Victor-emil neagoe, Adrian-dumitru ciotoc had dedicated to the challenging computer vision task of subject-independent emotion recognition from facial expressions. The original key idea of the proposed model is the increasing of the neural classifier training set size by adding virtual samples generated with a system of Concurrent Self-Organizing Maps (CSOM). The model consists of the following main processing cascade: (a) Gabor Wavelet Filtering (GVF); (b) dimensionality reduction using Principal Component Analysis (PCA); (c) Radial Basis Function (RBF) neural classifier trained with virtual samples generated by CSOM system (VSG-CSOM).

Alex Graves, Jürgen Schmidhuber, Christoph Mayer, Matthias Wimmer, Bernd Radig had presented a complete system for automatic facial expression recognition. The Candide-3 face model is used in conjunction with a learned objective function for face model fitting. The resulting sequence of model parameters is then presented to a recurrent neural network for classification. The advantage of using a recurrent network is that the temporal dependencies present in the image sequences can be taken into account during the classification. Since the entire process is automatic, and the recurrent networks can be used to make online predictions, the system would be ideal for real-time recognition. This would make it suitable for the CoTeSys ‘coffee break’ scenario, where guests must be recognized and served by robot waiters. Promising experimental results are presented on the Cohn-Kanade database.

Ashutosh Saxena, Ankit Anand, Prof. Amitabha Mukerjee explained an efficient, local image-based approach for extraction of intransient facial features and recognition of four facial expressions from 2D image sequences is presented. The algorithm uses edge projection analysis for feature extraction and creates a dynamic spatio-temporal representation of the face, followed by classification through a feed-forward net with one hidden layer. A novel transform for extracting lip region for color face images based on Gaussian modeling of skin and lip color is proposed. The proposed lip transform for colored images results in better extraction of lip region in the feature extraction stage. The algorithm achieves an accuracy of 90.0% for facial expression recognition from grayscale image sequences.

III. FACE EXPRESSION RECOGNITION SYSTEM

Face recognition system can be formulated as following phases:

Image Acquisition  
Pre-processing  
Feature Extraction  
Classification

Fig. 1: Face Expression Recognition System

Image Acquisition: Images used for facial expression recognition are static images or image sequences. An image sequence contains potentially more information than a still image, because the former also depicts the temporal characteristics of an expression. With respect to the spatial, chromatic, and temporal dimensionality of input images, 2-D monochrome (grey-scale) facial image sequences are the most popular type of pictures used for automatic expression recognition. However, colour images could become prevalent in future, owing to the increasing availability of low-cost colour image acquisition equipment, and the ability of colour images to convey emotional cues such as blushing.

Pre-processing: Image pre-processing often takes the form of signal conditioning (such as noise removal, and normalisation against the variation of pixel position or brightness), together with segmentation, location, or tracking of the face or its parts. Expression representation can be sensitive to translation, scaling, and rotation of the head in an image. To combat the effect of these unwanted transformations, the facial image may be geometrically standardised prior to classification. This normalisation is usually based on references provided by the eyes or nostrils. Segmentation is concerned with the demarcation of image portions conveying relevant facial information. Face segmentation is often anchored on the shape, motion, colour, texture, and spatial configuration of the face or its components [13]. The face location process yields the position and spatial extent of faces in an image; it is typically based on segmentation results. A variety of face detection techniques have been developed [13]. However, robust detection of faces or their constituents is difficult to attain in many real-world settings. Tracking is often implemented as location, of the face or its parts, within an image sequence, whereby previously determined location is typically used for estimating location in subsequent image frames.

Feature Extraction: Feature extraction converts pixel data into a higher-level representation of shape, motion,
Facial Emotion Recognition Techniques: A Survey

All rights reserved by www.ijsrd.com

Facial Emotion Recognition Techniques: A Survey

IJSRD/Vol. 3/Issue 03/2015/498

All rights reserved by www.ijsrd.com

Facial Emotion Recognition Techniques: A Survey

IJSRD/Vol. 3/Issue 03/2015/498

All rights reserved by www.ijsrd.com

Colour, texture, and spatial configuration of the face or its components. The extracted representation is used for subsequent expression categorisation. Feature extraction generally reduces the dimensionality of the input space. The reduction procedure should (ideally) retain essential information possessing high discrimination power and high stability. Such dimensionality reduction may mitigate the ‘curse of dimensionality’ [10]. Geometric, kinetic, and statistical- or spectral-transform-based features are often used as alternative representation of the facial expression prior to classification [14]. Figure 1: Facial Expression Processing Flow.

Classification: Expression categorisation is performed by a classifier, which often consists of models of pattern distribution, coupled to a decision procedure. A wide range of classifiers, covering parametric as well as non-parametric techniques, has been applied to the automatic expression recognition problem [14]. The two main types of classes used in facial expression recognition are action units (AUs) [6], and the prototypic facial expressions defined by Ekman [8]. The 6 prototypic expressions relate to the emotional states of happiness, sadness, surprise, anger, fear, and disgust [8]. However, it has been noted that the variation in complexity and meaning of expressions covers far more than these six expression categories [12]. Moreover, although many experimental expression recognition systems use prototypic expressions as output categories, such expressions occur infrequently, and fine changes in one or a few discrete face parts communicate emotions and intentions [6] [19]. An AU is one of 46 atomic elements of visible facial movement or its associated deformation; an expression typically results from the agglomeration of several AUs [6] [8]. AUs are described in the Facial Action Coding System (FACS) [7]. Sometimes, AU and prototypic expression classes are both used in a hierarchical recognition system – for example, categorisation into AUs can be used as a low-level of expression classification, followed by a high-level classification of AU combinations into basic expression prototypes [15].

IV. COMPARISON

<table>
<thead>
<tr>
<th>Reference</th>
<th>Pre-processing</th>
<th>Feature Extraction</th>
<th>Expression Classification</th>
<th>Recognition Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bourel et al.</td>
<td>Using a point tracker</td>
<td>Scalar quantisation of facial dynamics</td>
<td>rank-weighted k-nearest neighbor classifier.</td>
<td>Relatively little degradation in recognition under partial face occlusion or tracker noise.</td>
</tr>
<tr>
<td>Pantic and Rothkrant</td>
<td>Multiple detectors (e.g. snakes, neural networks, …).</td>
<td>Extraction of static geometric measurements</td>
<td>rule-based expert system</td>
<td>91% recognition of basic expression prototypes.</td>
</tr>
<tr>
<td>Essa and Pentland</td>
<td>View-based and Modular Eigenspace methods</td>
<td>(i) peak activation of each muscle. (ii) motion estimates</td>
<td>(i) maximum correlation with muscle activation template, (ii) minimum distance to motion energy template.</td>
<td>92% recognition of facial expressions.</td>
</tr>
<tr>
<td>Tian et al.</td>
<td>Gaussian mixture model, Lucas-Kanade tracking algorithm</td>
<td>Continuous and discrete representation of magnitude and direction for motion of face</td>
<td>multi-layer perceptron for upper-face and lower-face</td>
<td>96.7% recognition of lower-face AUs, and neutral expression.</td>
</tr>
<tr>
<td>Alex Graves</td>
<td>Multitude of features with a multi-stage fitting approach</td>
<td>Feature-point tracking by optical flow discriminates</td>
<td>Recurrent Neural Networks</td>
<td>85.4% AU recognition</td>
</tr>
<tr>
<td>Philipp Michel</td>
<td>Automatic Facial Feature Tracker</td>
<td>vector of displacements is calculated</td>
<td>Support Vector Machine</td>
<td>87.9% AU recognition</td>
</tr>
<tr>
<td>L. Sirovich</td>
<td>-</td>
<td>Low-Dimensional Procedure for Characterization of Human Faces</td>
<td>Principal Component Analysis</td>
<td>Recognition rate is low</td>
</tr>
<tr>
<td>Martin et al.</td>
<td>-</td>
<td>Using AAM based</td>
<td>AAM classifier set</td>
<td>Anger emotion with</td>
</tr>
</tbody>
</table>
Table 1

<table>
<thead>
<tr>
<th>Year</th>
<th>Model</th>
<th>Classification Method</th>
<th>Average Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>Appearance Based</td>
<td>SVM based classifier</td>
<td>94.9%</td>
</tr>
<tr>
<td>2009</td>
<td>Model vertices are determined using PCA</td>
<td>SVM based classifier</td>
<td>90%</td>
</tr>
</tbody>
</table>

V. DATABASE

The good choice of database under uncontrollable condition like occlusion and pose, illumination, expression variation is a very challenging task that deals with testing the novel approaches. Databases are used to test the proposed system on different images under varying condition like pose, illumination, occlusion, expression etc. Some databases are publically available for researchers. In some cases various databases stores the preprocessed data of images for learners. One subject or individual has number of samples in different varying conditions. Number of databases includes FERET, CMU-PIE, Extended YaleB, Cohn Kanade, AR, ORL, Japanese Female Facial Expression JAFFEE, Indian Face database etc. In 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 and is not fit for spontaneous expressions. Similar posed expression database are AR face database, Japanese Female Facial Expression Database (JAFFE) etc [12].

VI. CONCLUSION

Human beings naturally and intuitively use facial expression as an important and powerful modality to communicate their emotions and to interact socially [13]. There has been continued research interest in enabling computer systems to recognize expressions and to use the emotive information embedded in them in human-machine interfaces.

This paper presents a high-level overview of automatic expression recognition; it highlights the main system components and some research challenges. This work provided a framework for facial expression recognition that can effectively maximize information gathered about the emotion change and minimize the impact of person identity. Experiment results reveal that recognition performance varies for different neural network techniques for expression classification. So we can take best feature extraction parameter and neural network technique for expression classification method to increase Recognition accuracy.

REFERENCES


