AUTOMATIC FACIAL EXPRESSION ANALYSIS
A SURVEY

C.P. Sumathi¹, T. Santhanam² and M. Mahadevi³

1 Department of Computer Science, SDNB Vaishnav College for Women, Chennai, India
santsum@hotmail.com
2 Department of Computer Application, DG Vaishnav College for Men, Chennai, India
santhanam_dgvc@yahoo.com
3 Department of Computer Science, SDNB Vaishnav College for Women, Chennai, India
maxsaran@gmail.com

ABSTRACT:

The Automatic Facial Expression Recognition has been one of the latest research topics since 1990's. There have been recent advances in detecting face, facial expression recognition and classification. There are multiple methods devised for facial feature extraction which helps in identifying face and facial expressions. This paper surveys some of the published work since 2003 till date. Various methods are analysed to identify the Facial expression. The paper also discusses about the facial parameterization using Facial Action Coding System (FACS) action units and the methods which recognizes the action units parameters using facial expression data that are extracted. Various kinds of facial expressions are present in human face which can be identified based on their geometric features, appearance features and hybrid features. The two basic concepts of extracting features are based on facial deformation and facial motion. This article also identifies the techniques based on the characteristics of expressions and classifies the suitable methods that can be implemented.

KEYWORDS:
Facial Expression, FACS, Geometric Features, Appearance Features, Deformation, Facial Motion.

1. INTRODUCTION:

A Facial Expression is a visible manifestation of the affective state, cognitive activity, intention, personality and psychopathology of a person [1]. Facial Expression convey non-verbal communication cues in face-to-face interactions. Paul Ekman and Freisen have produced FACS –Facial Action Coding System for describing visually distinguishable Facial movements [2]. Using the FACS, Action Parameters are designated to each of the Expressions which classifies the Human Emotions [2]. Also Mehrabian indicated that the verbal part of a message contributes for 7 percent to the effect of the message, the vocal part contributes 38 percent while facial expression contributes for 55 percent [3]. There are wide range of Applications of Facial Expressions, some of which include image understanding, Psychological studies, facial nerve grading in medicine [4], face image compression and synthetic face animation [5]. Previous survey on the automatic analysis of Facial Expressions have discussed the classification methods and the extraction methods [6][7][8]. This Paper classifies different...
methods of the published work after 2003 in extracting those expressions which helps in recognizing the non-verbal communication in Human.

1.1 MEASURING FACIAL ACTIONS:

Prior to the compilation of the FACS in 1977, most of the facial behaviour researchers were relying on the human observers who would observe the face of the subject and give their analysis. But such visual observations cannot be considered as an exact science since the observers may not be reliable and accurate. Ekman et al. questioned the validity of such observations by pointing out that the observer may be influenced by context [9,10]. They may give more prominence to the voice rather than the face and furthermore, the observations made may not be the same across cultures; different cultural groups may have different interpretations [9,10].

The limitations that the observers pose can be overcome by representing expressions and facial behaviours in terms of a fixed set of facial parameters. With such a framework in place, only these individual parameters have to be observed without considering the facial behaviour as a whole. Even though, since the early 1920s researchers were trying to measure facial expressions and develop a parameterized system, no consensus had emerged and the efforts were very desperate [9,10]. To solve these problems, Ekman and Friesen developed the comprehensive FACS system which has since then become the de-facto standard.

Facial Action Coding is a muscle-based approach. It involves identifying the various facial muscles that individually or in groups cause changes in facial behaviours. These changes in the face and the underlying (one or more) muscles that caused these changes are called Action Units (AU). The FACS is made up of several such action units. For example:

- AU 1 is the action of raising the Inner Brow. It is caused by the *Frontalis* and *Pars Medialis* muscles,
- AU 2 is the action of raising the Outer Brow. It is caused by the *Frontalis* and *Pars Lateralis* muscles,
- AU 26 is the action of dropping the Jaw. It is caused by the *Masetter, Temporal* and *Internal Pterygoid* muscles, and so on [10]. However not all of the AUs are caused by facial muscles. Some of such examples are:
  - AU 19 is the action of ‘Tongue Out’,
  - AU 33 is the action of ‘Cheek Blow’,
  - AU 66 is the action of ‘Cross-Eye’, and so on [9,10]. The Face can be divided into Upper face[11] and Lower Face Action units[12] and the subsequent expressions are also identified. The Figures shows some of the combined action units.
2. BASIC STRUCTURE OF FACIAL EXPRESSION ANALYSIS:

The facial expression analysis system estimates the measurement of actions and also classifies the actions according to the manual of FACS[9]. The steps involved in Facial Expression analysis are 1. Face Acquisition 2. Facial Expression Extraction 3. Expression Recognition.

Face acquisition is the first step in which faces are detected from the input images or image sequences. The Face acquisition may detect faces in input images or detect face in the first frame and track the face in the remaining frames in case of image sequences. Based on the type of input the faces can be classified into 2D or 3D faces. After the face is located then the
Facial features are extracted to identify the facial expression. Facial expression can be classified into two types namely Geometric or InTransient features and Appearance Features or Transient Features.

**Geometric or Intransient Features:** The features that are always present in the face but may be deformed due to any kind of facial expression. For example, eyes, eyebrows, mouth, tissue textures, nose. The facial components or facial feature points are extracted to form a feature vector that represents the face geometry.

**Appearance or Transient Features:** The features that appear temporarily in the face during any kind of Facial Expression. For example, different kinds of wrinkles, bulges, forehead, regions surrounding the mouth and eyes. With appearance-based methods, image filters, such as Gabor wavelets, are applied to either the whole-face or specific regions in a face image to extract a feature vector.

Facial Expression Recognition is the last step in facial expression analysis where the extracted features are recognized based on the action units. The Recognizer identifies not only the basic emotions like anger, happy, surprise, sad [13] but also identifies the expression caused due to pain [14], temporal dynamics [15], Intensity of Expression [16], Spontaneous Expression [17].

![Figure 3: Basic Structure of Facial Expression analysis](image)

### 2.1 FACE ACQUISITION:

The two basic face acquisition methods are to detect faces both in frontal view images and near frontal view images. To detect the faces, two methods are used namely face detection and head pose estimation.

#### 2.1.1 FACE DETECTION:

Many detection methods have been employed to detect faces [18][19][20][21][22]. Some of the previous face detection methods since 2003 are summarized here. Mohammed Yeasin had used Robust and automated face detection [23] to segment the face region which was based on the work of Rowley [19]. Marian Stewart Bartlett had designed boosting techniques for face detection in a generative framework [17] based on their own work [18]. Y. Tong had drawn geometry of face by encoding the edges of the face in the graph [24] based on the work of L. Wiskott [25]. Irene Kotsia had employed convolutional neural network for detecting the face and the classification is performed using a rule-based algorithm [19].
2.1.2 HEAD POSE ESTIMATION:

To handle the out-of-plane head motion, head pose estimation can be employed. The methods for estimating head pose can be classified as 3D model-based methods [26,27] and 2D image-based methods [28]. In 3D Model based method Bartlett used a canonical wire-mesh face model to estimate face geometry and 3D pose from hand-labelled feature points. In 2D image based method To handle the full range of head motion for expression analysis, Tian et al. [28] detected the head instead of the face. The head is identified using the smoothed silhouette of the foreground object as a segment using background subtraction and computing the negative curvature minima (NCM) points of the silhouette.

2.2 FACIAL EXPRESSION EXTRACTION:

The extraction is basically based on the type of features, Geometric Features and Appearance Features. The two basic concepts employed for extracting features are based on identifying facial deformation and facial motion. The deformation based features recognize the Action Units, and the classifier is trained to differentiate human emotional states based on identified Action Units. The Deformation kind of extraction is applied to images [29] and to image sequences [30]. The motion based features exploit the temporal correlation of facial expressions to identify variations within a probabilistic framework[15]. Image based models extract features from images, or reduced dimensional facial components[31]. Model based features are usually shape or texture models that fit human faces. The output of the feature extractor stage must contain separable and classifiable vectors. Active appearance models[32] and point distribution models[33] are used to fit on the shapes of interest. These shapes constitute the feature vectors. The expression extraction methods are widely classified under two kinds namely deformation extraction and motion extraction. As for motion extraction techniques, some commonly used methods are dense optical flow[34], feature point tracking[35], and difference images[36].

The Various techniques under facial expression extraction methods are tabulated in the table.

<table>
<thead>
<tr>
<th>Deformation Extraction</th>
<th>Geometric Features</th>
<th>Appearance Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image Based</td>
<td>Gabor Filter[29]</td>
<td>Local Gabor Filter Bank[31][16], Fisher’s Linear Decomposition, Singular Value decomposition[37][38]</td>
</tr>
<tr>
<td>Model Based</td>
<td>Point Distribution Model[33]</td>
<td>Feature point Tracking[35]</td>
</tr>
<tr>
<td>Motion Extraction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frame Based</td>
<td>Active contour(snake)[39]</td>
<td>Gabor Filter Bank[17]</td>
</tr>
<tr>
<td>Sequence Based</td>
<td>PCA[23][44], Gabor Filter Bank &amp; AdaBoost[40]</td>
<td>Haar like feature[41], Multimodal facial feature tracking[24], Candid Grid Node[30].</td>
</tr>
</tbody>
</table>

Table 1: Facial Expression Extraction Methods.
2.2.1 Geometric Feature Extraction:

Geometric Extraction is to detect and track changes of facial components in near frontal face images. Tian et al. develop multi-state models to extract the geometric facial features. A three-state lip model describes the lip state: open, closed, tightly closed. A two-state model (open or closed) is used for each of the eyes. Each brow and cheek has a one-state model. Some appearance features, such as nasolabial furrows and crow’s feet wrinkles (Fig. 5), are represented explicitly by using two states: present and absent.

Model Based:

Automatic Active Appearance Model (AAM) mapping can be employed to reduce the manual preprocessing of the geometric feature initialization. Xiao et al. [43] performed the 3D head tracking to handle large out-of-plane head motion and track nonrigid features. Once the head pose is recovered, the face region is stabilized by transforming the image to a common orientation for expression recognition [42].

Image Sequence:

Given an image sequence, the region of the face and approximate location of individual face features are detected automatically in the initial frame. The contours of the face features and components then are adjusted manually in the initial frame. After the initialization, all face feature changes are automatically detected and tracked in the image sequence. The system groups 15 parameters for the upper face [11] and 9 parameters for the lower face [12], which describe shape, motion, and state of face components and furrows. To remove the effects of variation in planar head motion and scale between image sequences in face size, all parameters are computed as ratios of their current values to that in the reference frame.

2.2.2 Appearance Feature Extraction:

Gabor wavelets [31] are widely used to extract the facial appearance changes as a set of multiscale and multiorientation coefficients. The Gabor filter may be applied to specific locations on a face [31] or to the whole face image [16]. There are two types of features to recognize expressions, the geometric positions of 34 fiducial points on a face and 612 Gabor wavelet coefficients extracted from the face image at these 34 fiducial points.
The Preprocessing procedure steps for the Gabor Filters are 1). detecting facial feature points manually including eyes, nose and mouth; 2). rotating to line up the eye coordinates; 3) locating and cropping the face region using a rectangle according to face model as shown in Figure 6.

**Image Sequence:**

Techniques like Haar-like Feature, Facial Feature tracking are used to identify the facial features that produces the expressions. A multi-modal tracking [24] approach is required to enable the state switching of facial components during the feature tracking process. Twenty-six fiducial points and 56 contour points are used in the facial model. Using the Facial model the fiducial points are marked for an image sequence using feature tracking method. The marked features in an image sequence is shown in the figure 8.

**2.3 Expression Recognition:**

Recognizing the facial expressions is the last step in Facial Expression Analysis. This is classified into basic categories namely Frame Based Expression recognition and Sequence Based Expression recognition.

**2.3.1 Frame-Based Expression Recognition:**

Frame-based expression recognition does not use temporal information for the input images. It uses the information of current input image with/without a reference frame. The input image can be a static image or a frame of a sequence that is treated independently. Several methods can be found in the literature for facial expression recognition such as neural networks [35,36], support vector machines [30], linear discriminant analysis [31], and rule-based classifiers [46].
2.3.2 THE SEQUENCE-BASED RECOGNITION:

This method uses the temporal information of the sequences to recognize the expressions of one or more frames. To use the temporal information, the techniques such as HMM [15], recurrent neural networks, and rule-based classifier [47] were employed in facial expression analysis. The article by Yunfeng Zhu, Fernando De la Torre, Jeffrey F. Cohn have made comparative studies for FACS AU recognition in spontaneously occurring behaviour by using the same RU-FACS database. Several systems had tried to recognize AUs or expression in spontaneously occurring behaviour [17].

The Expression recognition classifiers are listed below.

<table>
<thead>
<tr>
<th>Technique</th>
<th>Title</th>
<th>Year</th>
<th>Author</th>
<th>Method</th>
<th>Database</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frame based</td>
<td>A New Facial Expression Recognition Method based on local Gabor filter bank and PCA plus LDA</td>
<td>2005</td>
<td>Hong-Bo Deng, Lian-Wen Jin, Li-Xin Zhen, Jian – Cheng Huang</td>
<td>PCA plus LDA</td>
<td>JAFFE</td>
</tr>
<tr>
<td>Automatic Facial Expression Recognition using facial animation parameters and multistream HMMS</td>
<td>2006</td>
<td>Petar S. Aleksic, Member, IEEE, and Aggelos K. Katsaggelos, Fellow, IEEE</td>
<td>Multistream Hidden Markov Models</td>
<td>Cohn Kanade</td>
<td></td>
</tr>
<tr>
<td>A Region Based methodology for facial expression recognition</td>
<td>2006</td>
<td>Anastasios C. Koutlas, Dimitrios I. Fotiadis</td>
<td>Neural Networks</td>
<td>JAFFE</td>
<td></td>
</tr>
<tr>
<td>A Facial Expression Classification System Integrating Canny, Principal Component Analysis and Artificial Neural Network</td>
<td>2011</td>
<td>Le Hoang Thai, Nguyen Do Thai Nguyen and Tran Son Hai, Member, IACSIT</td>
<td>PCA &amp; ANN</td>
<td>JAFFE</td>
<td></td>
</tr>
<tr>
<td>Sequence Based</td>
<td>Recognition of Facial Expressions and Measurement of Levels of Interest From Video</td>
<td>2006</td>
<td>Mohammed Yeasin, Senior Member, IEEE, Baptiste Bullot, and Rajeev Sharma, Member, IEEE</td>
<td>Hidden Markov Model</td>
<td>Cohn-Kanade</td>
</tr>
</tbody>
</table>
3. Discussion:

In this survey on automatic facial expression analysis, facial expression analysis with regard to different motion and deformation-based extraction methods, model and image-based representation techniques as well as recognition and interpretation-based classification approaches are discussed. While trying to classify the facial expression, two approaches had been handled namely, classifying expressions based on facial action coding system [1][2], and direct and indirect interpretation of facial expressions. Few recent articles are discussed for its recognition rate and action unit detection. The various techniques and their classification rates are listed below.
<table>
<thead>
<tr>
<th>Authors</th>
<th>Year</th>
<th>Action Element detected</th>
<th>Emotion Classified</th>
<th>Extraction methods</th>
<th>Classification methods</th>
<th>Classification Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amir Jamshidinzhad, Md jan Nordin</td>
<td>2011</td>
<td>8</td>
<td>Geometric measurement</td>
<td>Fuzzy logic</td>
<td>90.33%</td>
<td></td>
</tr>
<tr>
<td>P.Geetha, Dr. Vasumathi Narayan</td>
<td>2010</td>
<td>24</td>
<td>2D-Principal Component Analysis</td>
<td>Dynamic 2D Cellular Automata video frames</td>
<td>94.13%</td>
<td></td>
</tr>
<tr>
<td>Sander Koelstra, student member IEEE, Maja Pantic, Senior Member IEEE, Ioannis Patras, member IEEE</td>
<td>2010</td>
<td>27</td>
<td>Quadtree Decomposition</td>
<td>Hidden Markov model</td>
<td>94.3%</td>
<td></td>
</tr>
<tr>
<td>Peng Yang, Qingshan Liu, Dimitris N.Mexatras</td>
<td>2009</td>
<td>8</td>
<td>Haar like feature</td>
<td>Adaboost</td>
<td>96.6%</td>
<td></td>
</tr>
<tr>
<td>Le Hoang Thai, Nguyen Do Thai, Nguyen Tran son hai, Member IACSIT</td>
<td>2011</td>
<td>6</td>
<td>PCA</td>
<td>Neural Network</td>
<td>85.7%</td>
<td></td>
</tr>
<tr>
<td>Guoying Zhao, Matti Pretikainen</td>
<td>2009</td>
<td>6</td>
<td>Adaboost</td>
<td>Boosted Multi resolution spatio temporal descriptors</td>
<td>93.85%</td>
<td></td>
</tr>
<tr>
<td>Pooja Sharma</td>
<td>2011</td>
<td>6</td>
<td>Pattern Tracking</td>
<td>Optical flow based analysis</td>
<td>83.33%</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: List of recent work for recognition rate.

In the above table the first four system performs facial expression classification based on facial action elements that is detected. The remaining systems perform direct or indirect interpretations of facial expression. Furthermore, the facial expression intensities are studied in some systems for smile detection [16], pain detection[14],identification of posed and Genuine pain[49]. Much more algorithms are focussing on extraction algorithms and in classification for an optimal recognition rate.

4. APPLICATIONS OF FACIAL EXPRESSION ANALYSIS:

Automatic Facial expression analysis find applications in different interesting areas. Mostly used in human computer interaction interface. Expression recognition acts as an intelligent system in identifying the genuine pain and posed pain, hence thereby finds its place in medical field, identifies the expression of student’s understanding capability and finding its place in Education environment, identifies the expression of a person during any sort of interrogation and brings out its necessity in crime and security division. Recognition of
expression intensities helps in identifying the faces of pain like chronic or acute. The Expression system is used in many domains like Telecommunications, Behavioural Science, Video Games, Animations, Psychiatry, Automobile Safety, Affect sensitive music juke boxes and televisions, Educational Software, etc.

5. CONCLUSION:

The objective of this paper is to show a clean survey on the structure of analysing the facial expression. The steps involved in expression analysis like face acquisition, feature extraction and expression classification had been discussed. Each step is discussed with the approaches and methods that can be applied to attain the required goal. The expression recognition based on FACS and direct or indirect interpretation are also discussed with some of the recent research work. Although many researchers have been investigating facial expressions, basic expressions like happy, sad, disgust, surprise had been the interesting topic that is been widely discussed. Topics like Expressions recognition during spontaneous movement, intensity of expressions, combination of facial action elements detection, temporal segmentation, pain analysis are still some topics of interest that are under the cover which needs to be unwrapped.

REFERENCES:


[36] Sander Koelstra, Student Member, IEEE, Maja Pantic, Senior Member, IEEE, and Ioannis (Yiannis) Patras, Member, IEEE, “A Dynamic Texture-Based Approach to Recognition of Facial Actions and Their Temporal Models”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 32, No. 11, November 2010


[48] Yunfeng Zhu, Fernando De la Torre, Jeffrey F. Cohn, Associate Member, IEEE, and Yu-Jin Zhang, Senior Member, IEEE"Dynamic Cascades with Bidirectional Bootstrapping for Action Unit Detection in Spontaneous Facial Behavior", Journal of LATEX Class Files, October 2010 .