Real-time Automatic Facial Expression Recognition in Video Sequence

Nivedita Singh[1] and Chandra Mani Sharma[2]

[1]Institute of Technology & Science (ITS)
Mohan Nagar, Ghaziabad-201007, (UP), India

[2]Institute of Technology & Science (ITS)
Mohan Nagar, Ghaziabad-201007,(UP), India

Abstract

It is an established fact that words have just 7% contribution in human communication and 93% part is contributed by non-verbal communication exhibited by body language. Facial expressions constitute a vital part when we talk about human body language. The representation and recognition of facial expressions by computers has been an area of vital interest since long. The recent time has witnessed increasing interest in this field because of its potential applications such as in marketing, human computer interaction, security, psychiatry analysis and visual surveillance. Though, various techniques for facial expression recognition have been proposed in available literature, yet they are inadequate for real-time video applications. This is due to the dynamic nature of video data and various other factors associated with it. This paper presents a simple overview of the facial expression recognition and describes a technique for facial expression recognition in video. The comparative analysis of various methods has also been presented in this paper.

Keywords: *Facial Expression Recognition, Pattern Recognition, Machine Vision, Algorithm, Support Vector Machine, Local Binary Pattern etc.*

1. Introduction

The ultimate evolution goal of computers is to create humanoids (computers with all human-like qualities). Creating programs to simulate and understand human emotions is the next big milestone in this direction. It is all with an intent to make human computer interaction as lively and natural as possible that computer no more remains a heartless machine but it becomes a emotional counterpart of humans. All the technological advancements are deeply impacting the society. Emotions play the role of a bonding fabric in society. If the research in the field of facial expression recognition yields something great then definitely it will have a positive impact on dealing various societal challenges and issues. Facial expressions are the signal of various emotions like happiness, sadness, anger, confusion, excitement, contempt, desire, disgust ,fear and surprise etc. Our smile can easily be perceived as approval or happiness while sad face can signal disapproval or sadness. This implies that our facial expression may showcase our true feelings for a particular circumstance to convey non-verbal information. A Facial expression may be defined as the set of motions or movements of our skin, muscles and tissues that involves brain for emotional experience. Facial expression is a form of non-verbal communication. A person expresses his/her sentiments by using facial expressions. There is a complex underlying mechanism behind the formation of facial expression.

To formulate this phenomenon mathematically and making it perfectly computer understandable, further research is required. There has not been ample research on this issue and most of the researchers have investigated various algorithms to handle expression variations at primitive level. Feature extraction is a general term for the mechanism of designing concatenation of the variables to get around these obstacles while still describing the input with accuracy.

This approach is beneficial when image sizes are too large and abstracted feature representation is needed to quickly tasks such as image matching and retrieval. There are various feature extraction mechanisms such as:Local Binary Patterns (LBP), Haar Wavelets, Color Histograms, Histogram of Oriented Gradients (HOG), Speeded Up Robust Features (SURF).

Figure 1 represents a general classification of distinctive expression recognition algorithms. Face recognition under changing facial expression algorithms is categorized as local or feature based approaches and holistic or appearance based approaches that operates the overall face for extracting facial features and give the brief information.

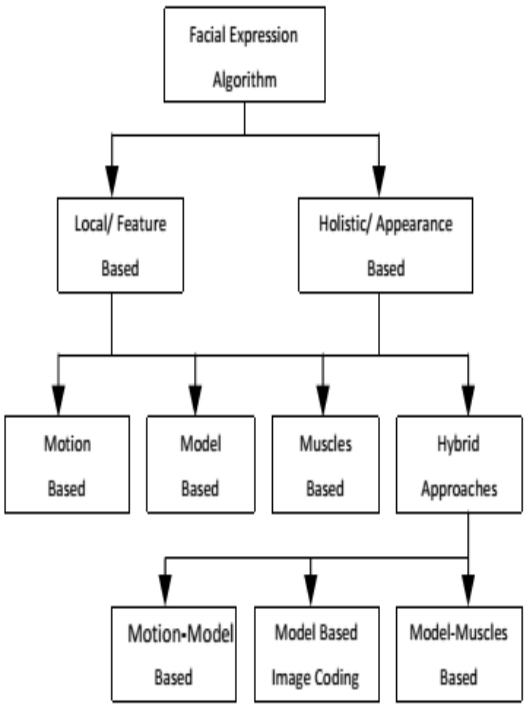

Fig. 1: Distinction of feature extraction and representation

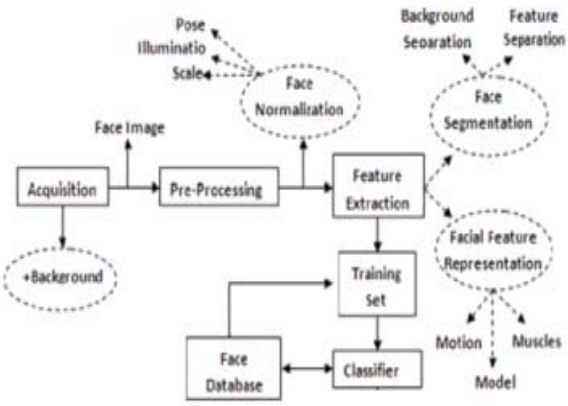

Fig. 2: General framework for facial expression recognition

Sometimes, all the information are not relevant because all facial features are not affected by the appearance of single emotion as, for example, the degree of smile doesn't changes all features but affects only the appearance of cheeks and lips etc. So, in contrast to holistic based approaches local based approaches offers a way to process only the affected facial features. Local graph is used in order to trace facial features while global graph is used to store the information of face texture. Feature and appearance based approaches are further categorized as motion, model based, muscles-based and hybrid approaches. There are further variants of motion-model based, motion based image coding and model-muscles based approaches. The general frame work for automatic facial Expression is shown in Figure 2. Primarily, the face images are obtained and standardized in order to exclude the complications like pose and radiance Factor during face consideration. It is maxim that feature extraction is a great milestone which uses numerous methods to portray facial features like motion, model, and muscles-based approaches. Finally these features are categorized and trained in manifold subspaces and then used for recognition.

Fasel et al.[1] present a survey of various methods of facial expression analysis. Tian et al.[2] also discuss the concept of facial expression recognition. Various techniques exist for facial expression recognition.

Moore et al.[3] discuss the use of Local Binary Patterns (LBP) for facial expression recognition. Bartlet et al.[4] delve into machine learning for facial expression recognition.

2. The Proposed Framework for Facial Expression Recognition

Figure 3 depicts the process of facial expression recognition. Here, a machine learning approach is used. A classifier is trained first using an annotated corpus of data containing different facial expressions of various genders, age groups and ethnic groups. Local binary patterns are extracted from training samples and then support vector machine is used for training a classifier.

This trained classifier is used to classify facial expressions in six primitive groups namely joy, sorrow, surprise, anger, disgust and fear. If a reasonable match is not found with any of these six expressions then it is classified as "other". In an image or video frame, before facial expression recognition, human face is recognized with the help of Viola Jones classifier [5]. The reason for using Viola Jones classifier is its real-time processing capability and accuracy in face detection.

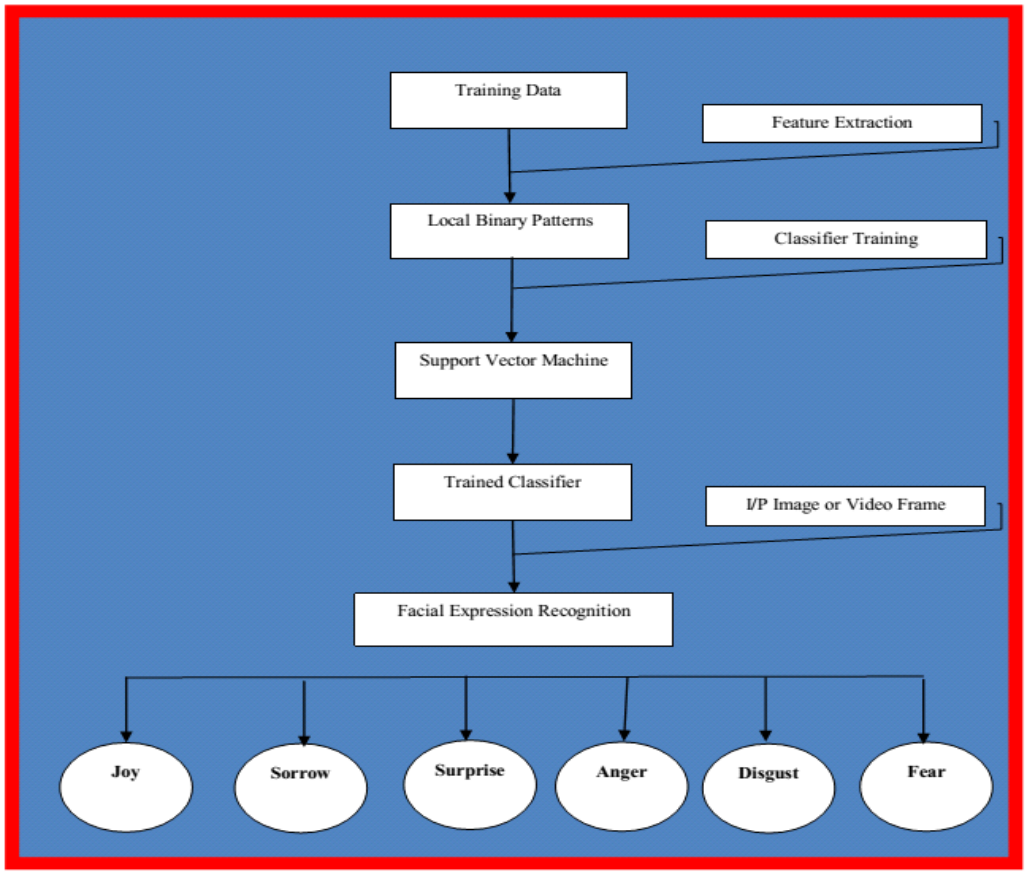

Fig. 3: The schematic diagram for proposed method of facial expression recognition

3. Experimental Results

The proposed technique successfully recognizes the six primitive facial expressions namely joy, sorrow, surprise, anger, disgust and fear. The accuracy matrix of the proposed method is depicted as given below:

Table 1: Facial expression recognition accuracy matrix for six different facial expressions

	Joy	Sorrow	Surprise	Anger	Disgust	Fear
Joy	**98.20%**	0.00%	1.80%	0.00%	0.00%	0.00%
Sorrow	0.00%	**97.46%**	0.00%	0.00%	1.27%	1.27%
Surprise	1.8%	0.00%	**96.00%**	0.00%	1.10%	1.10%
Anger	0.00%	0.00%	0.00%	**98.20%**	1.80%	0.00%
Disgust	0.00%	0.00%	0.00%	1.74%	**98.60%**	0.00%
Fear	0.00%	0.90	0.00%	0.00%	0.90%	**98.20%**

Accuracy matrix for facial expression recognition has beengiven in Table 1. It shows how accurately a facial expression is recognized, for instance, Joy is recognized 98.20% times accurately and only 1.80 % times it is misclassified as Surprise.

The average accuracy of proposed method = (98.20 + 97.46 + 96.00 + 98.20 + 98.60 + 98.20) / 6
= 586.66 / 6
= **97.77%**

Table 2: False Positive Rate (FPR) and False Negative Rate (FNR) for Facial Expression Recognition

Facial Expression	**FPR**	**FNR**
Joy	1.80%	1.80%
Sorrow	0.90%	2.54%
Surprise	1.80%	4.00%
Anger	1.74%	1.80%
Disgust	5.07%	1.74%
Fear	2.37%	0.90%

Table 2 depicts the False Positive Rate (FPR) and False Negative Rate (FNR) in facial expression recognition using proposed method. 'Disgust' has largest FPR (5.07%) and 'Sorrow' has least FPR (0.90%). 'Surprise' has largest FNR (4.00%) and 'Fear' has least FNR (0.90%).

The proposed method has been tested on a variety of facial expression datasets. Our dataset contains the image

sequences of several actors in with facial expressions namely anger, disgust, fear, joy, sorrow and surprise are exhibited. A sample snapshot has been taken from the dataset containing the six facial expressions of two female faces.

(a) Original input images, extraction of face and facial expressions respectively anger, disgust, fear, joy, sorrow, and surprise

(b) Original input images, extraction of face and facial expressions respectively anger, disgust, fear, joy, sorrow, and surprise

Fig. 4: Sample images used to test the proposed method for facial expression recognition

4. Facial Expression Recognition Databases

Humans have always had the innate ability to recognize and distinguish between faces, yet computers only recently have shown the same ability. In the mid-1960s, scientists started work on using the computer to identify human faces. Since then, facial recognition software has acquired a vast area. Now we will try to understand how the software works, Facial recognition examines the features of a person's face images data by a digital video camera. It computes the overall facial structure, comprising gap area between eyes, nose, jaw edges and mouth. These computations are retained in a database and used comparison when a user stands in front of the camera. Every face has numerous, differentiable marks, the different brows and chins that form up facial features. Each person face has around 80 nodal points. Some of these nodal points are measured by using:

- Distance between the eyes
- Width of the nose
- Depth of the eye sockets
- The shape of the cheekbones
- The length of the jaw line

These nodal points are measured by creating a numerical code; these numerical codes are called a faceprint, representing the face in the database.
The following four-stage process illustrates the way biometric systems operate:

Capture - a physical or behavioral data is trapped by the system during matriculation.

Extraction - vital data is mined from the data and a template is formed.

Comparison - the template created at the time extraction is then equated with a new sample.

Matching - the system then resolves if the features extracted from the new sample are identical or not

When the user comes to the camera, keeping distance around two feet from it. The system will trace the user's face and do matches against the requested identity or the facial database. It is probable that the user may tend to move and reattempt the confirmation based on his facial position. The system naturally shows output less than 5 seconds

5. Conclusion

This paper presents a discussion on facial expression recognition and describes a simple method for facial expression recognition using Local Binary Patterns and Support Vector Machine. The method is well suited for facial expression recognition in an image and in a video sequence both. The task of facial expression recognition is important because of its potential applications in various fields ranging from healthcare to security.

References

[1] B. Fasel, and J. Luettin "Automatic facial expression analysis: a survey" Pattern recognition, 36(1), 259-275, 2003.

[2] Tian, Y., Kanade, T., & Cohn, J. F. (2011). Facial expression recognition. In Handbook of face recognition (pp. 487-519). Springer London.

[3] Moore, S., & Bowden, R. (2011). Local binary patterns for multi-view facial expression recognition. Computer Vision and Image Understanding, 115(4), 541-558.

[4] Bartlett, M. S., Littlewort, G., Frank, M., Lainscsek, C., Fasel, I., & Movellan, J. (2005, June). Recognizing facial expression: machine learning and application to spontaneous behavior. In Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on (Vol. 2, pp. 568-573). IEEE

[5] Viola, P., & Jones, M. J. (2004). Robust real-time face detection. International journal of computer vision, 57(2), 137-154.

Nivedita Singh is a student of BCA at Institute of Technology & Science, Mohan Nagar Ghaziabad. She has interest in Computer Programming and User Interface Designing. She has won medals at various inter institutional competitions including Best Research Paper Award @ EPICTIS 2014, First Prize in Spardha 2014.

Chandra Mani Sharma is working as an Assistant Professor in Department of IT, ITS Mohan Nagar, Ghaziabad. He has more than 25 research papers published in various SCI Indexed and other refereed Journals, ACM/IEEE Conferences, Book Chapters and Edited Books. He received MHRD Scholarship during 2009-11 and has worked on UGC/DST funded research projects. His areas of interest include Computer Vision, Image Processing, Machine Learning, Software Testing and Web Application Development