A Literature Review: Various Facial Expression

Neha
M.Tech (ECE), NCCE, Israna, Panipat, Haryana, India.

Deepti Jaglan
Assistant Professor, NCCE, Israna, Panipat, Haryana, India.

Abstract – The human face is main element to understand the individuals as well as gives the vital data, cutting-edge notion of user's behavior via their one of kind expressions. Therefore, in biometric vicinity of the studies the Automatic Facial Expression Recognition has been one of the latest research topic. The goal of facial expression recognition (FER) algorithms is to extract the discriminative and distinguishing features of a face. Multiple methods have been devised which helps in identifying face and facial expression. The various facial expressions are popularly identified on the basis of their geometric features, appearance features and hybrid features. This paper presents a literature overview over the various strategies used for facial expression reputation. Also, one-of-a-kind facial expression datasets available for the research or trying out of present techniques of facial expression popularity are mentioned. A comparative study is also carried out using various feature extraction techniques.

Index Terms – Facial Expression, Face Detection, Features Extraction, Feature categorization Expression Recognition.

1. INTRODUCTION

Facial expression is one of the most common non-verbal ways that humans use to convey their internal emotional states and, consequentially, plays a significant role in interpersonal interactions. A Facial Expression is a visible manifestation of the affective state, cognitive activity, intention, personality and psychopathology of a person [1]. Although there exists a wide range of possible facial expressions, psychologists have identified the six basic ones (happiness, sadness, fear, disgust, surprise, and anger) that are universally recognized (Izard 1971). It is straightforward that a system capable to perform an automatic recognition of the human emotions is a desirable task for a set of emerging applications. The wide range of application of Facial Expression includes image understanding, Psychological studies, facial nerve grading in medicine [4], face image compression and synthetic face animation [5]. Paul Ekman and Freisen have produced FACS- Facial Action Coding System for describing visually distinguishable Facial movements [2]. Also Mehrabian indicated that the verbal part of a message contributes for 7% to the effect of the message, the vocal part to 38% while facial expressions contribute for 55%. More specifically, the capability to automatically recognize the emotional state of a human being is a key factor to the challenging field of the human–robot interaction since on the one hand it allows to introduce behavioral metrics and on the other hand it could increase the level of technology acceptance.

In human-to-human communication, the articulation and belief of facial expressions form a conversation channel in addition to voice which includes essential facts about the mental, emotional, or even physical state of the people in communiqué. A character's facial expressions in its handiest form is a greater diffused satisfied or irritated mind, feelings or information of the speaker predicted or unexpected reaction from listeners, sympathy, or even what the speaker is announcing no sign can offer to computing historical past, brings our everyday human user to remain at the vanguard inside the fabric will flow to take in. This installation a usually prediction, pervasive computing and ambient intelligence inclusive of had to reap the following era of computing [12]. It is easy to evidently occurring multimodal human-human communiqué focused reaction to the person interface will want to be advanced to become aware of such interfaces and intentions and as expressed with the aid of feelings of social and emotional indicators will need to have the potential. This imaginative and prescient of the destiny motivates the studies for automatic popularity of nonverbal movements and expression. Facial expression reputation, pc vision, sample reputation and human-computer interaction studies has attracted growing attention in groups. Automatic reputation of facial expressions is so affective computing technologies, shrewd tutoring systems, including various bureaucracies the essence of the next generation computing device, affected person monitoring systems, and so forth. Non-public wellness profiled [13]. Human face different gender, one-of-a-kind age groups and different physical characteristics of a person varies.

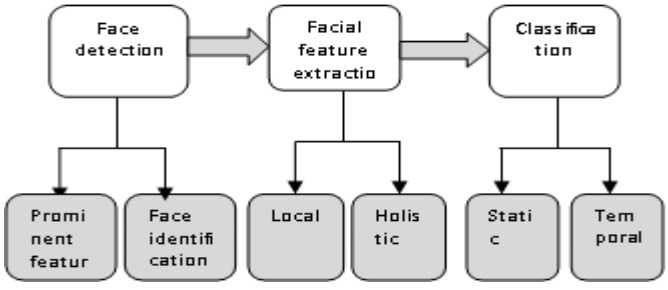

Fig-1: Generalized Framework for Automatic Facial Expression System

Overall framework of automated facial expression can be depicted from figure 1. From the discern, there are three important levels of typical face recognition system as shown in the figure 1.

(a) **Face detection**- It is an important part of face recognition and is the very first step in automatic face recognition. It is a computer technology that identifies human faces in digital images. It also refers to the psychological process by which humans locate and attend to faces in a visual scene.

(b) **Facial feature extraction**- This step transforms rich content of images into various content features. It is a process that generates features to be used in the classification task.

(c) **Classification**- 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, Intensity of Expression, Spontaneous expression[8]. In the next section Literature related to feature extraction is discussed in detail.

2. RELATED WORK

In this section, work related to previous techniques is discussed. Many researchers in their work have attempted to recognize the facial expression of an individual to the samples in a particular database of faces.

In the work of S. Kumar, *et.al.* the aim of facial expression recognition (FER) algorithms is to extract discriminative features of a face. However, discriminative features for FER can only be obtained from the informative regions of a face. Also, each of the facial subregions have different impacts on different facial expressions. Local binary pattern (LBP) based FER techniques extract texture features from all the regions of a face, and subsequently the features are stacked sequentially. This process generates the correlated features among different expressions, and hence affects the accuracy. The authors' approach entails extracting discriminative features from the informative regions of a face. In this view, he proposed an informative region extraction model, which models the importance of facial regions based on the projection of the expressive face images onto the neural face images. This feature extraction method reduces miss-classification among different classes of expressions. Experimental results on standard datasets show the efficiency of the proposed method.[1]

J. V. Patil, *et.al.* described facial expression as an essential part in communication. It is a challenging task in computer vision as well as in pattern recognition. Facial expression recognition has applications in many fields such as HCI, video games, virtual reality, and analyzing customer satisfaction etc. His proposed system focuses on emotion recognition using facial expressions which are captured by using Intel's Real Sense SR300 camera. This camera detects landmarks on the depth image of a face automatically using Software Development Kit (SDK) of Real Sense camera. Geometric feature based approach is used for feature extraction. The distance between landmarks is used as features and for selecting an optimal set of features brute force method is used. The Proposed system has used Multilayer Perceptron (MLP) neural network algorithm using back propagation method for classification. The proposed system recognizes three facial expressions namely neutral, happy, and surprised. The recognition rate achieved is 93.33%.[2]

J. d. A. Fernandes, *et.al.* in his article presented two facial geometric-based approaches for facial expression recognition using support vector machines. The first method performed an experimental research to identify the relevant geometric features for human point of view and achieved 85% of recognition rate. The second experiment employed the Correlation Feature Selection and achieved 96.11% of recognition rate. All experiments were carried out with Cohn-Kanade database and the results obtained are compatible with the state-of-the-art in this in this research area[3]

L. A. Jeni, *et.al.* worked on pose invariance in his article. He rendered 3D emotional database with different poses using BU 4DFE database, fitted 3D CLM, transformed the 3D shape to frontal pose and evaluated the outputs of our classifier. Results show that the high quality classification is robust against pose variations. The superior performance suggests that shape, which is typically neglected or used only as side information in facial expression categorization, could make a good benchmark for future studies.[4]

S. Jain, *et.al.* proposed that Conditional Random Fields (CRFs) can be used as a discriminative approach for simultaneous sequence segmentation and frame labeling. Latent-Dynamic Conditional Random Fields (LDCRFs) incorporates hidden state variables within CRFs which model sub-structure motion patterns and dynamics between labels. Motivated by the success of LDCRFs in gesture recognition, he proposed a framework for automatic facial expression recognition from continuous video sequence by modeling temporal variations within shapes using LDCRFs. The proposed approach outperforms CRFs for recognizing facial expressions. Using Principal Component Analysis (PCA) we study the reparability of various expression classes in lower dimension projected spaces. By comparing the performance of CRFs and LDCRFs against that of Support Vector Machines (SVMs). He demonstrated that temporal variations within shapes are crucial in classifying expressions especially for those with a small range of facial motion like anger and sadness.[5]

H. Li, *et.al.* in his work emphasized on Cross-age face recognition as it has remained a popular research topic because

the sophisticated facial change across age disables regular face recognition systems. Widely applied in age-related tasks, the hidden factor analysis (HFA) model decomposes face feature into independent age and identity factors. However, the hypothesis that the identity and age factors are independent is not in accordance with the fact that aging has different appearance changes on different people's faces. To address this problem, this letter presents a novel method for cross-age face recognition, called age-identity modified HFA, which exploits a new latent factor modeled as a linear combination with the age factor and the identity factor. Hence, the cross-age identity information can be extracted and separated preferably. A maximum likelihood strategy is proposed to judge which gallery face has the same identity with the probe image, while we do not need to know what the probe identity is. Extensive experiments are performed on the benchmark aging datasets MORPH and FG-Net, and the recognition rate of our method outperforms HFA by 10.4% and 1.15%, respectively.[6]

S. An and Q. Ruan, *et.al.* calculated the difference of gray value standard between neighboring pixels and the center pixel as a threshold to binary instead of the traditional LBP operation which only comparison of size between neighboring pixels and the center pixel. After he get the LTBP feature, fuse the LTBP and HOG (Histogram of Oriented Gradient) features to get multi-feature fusion for 3D facial expression recognition. His algorithm of 3D facial expression recognition comprises three steps: (1) extracting two sets of feature vectors and establishing the correlation criterion function between the two sets of feature vectors; (2) solving the two sets canonical projective vectors and extracting their canonical correlation features by the framework of canonical correlation analysis algorithm; (3) doing feature fusion for classification by using proposed strategy. He had performed comprehensive experiments on the BU-3DFE database which is presently the largest available 3D face database. He achieved verification rates of more than 90% for the 3D facial expression recognition.[7]

J. Kumari, *et.al.* in her work stated that Facial Expression is the easy way of telling/showing inner feelings. The Facial Expression Recognition system has many applications including human behavior understanding, detection of mental disorders, synthetic human expressions and many more. This paper presents a quick survey of facial expression recognition as well as a comparative study of various features on JAFFE and CK datasets. It mainly focuses on appearance based techniques. Recently, HOG has been widely used for feature extraction in image. It is found in our experiment that HOG feature gives comparable good recognition rate in facial expression recognition. Fusion of LBP with LGC and Fusion of HOG with other features like LDP and wavelets also improved their respective recognition rates.[8]

The work of T. W. Shen *et.al.* proposes that the Facial expressions recognition has gained a growing attention from industry and also academics, because it could be widely used in many field such as Human Computer Interface (HCI) and medical assessment. In this paper, we evaluate the strength of the Light Field Camera for facial expression recognition. The light filed camera can capture the directions of the incoming light rays which is not possible with a conventional 2D camera.. Firstly, a new facial expression dataset is collected by the light field camera. The depth map is estimated and applied on Histogram Oriented Gradient (HOG) to encode these facial components as features. Then, a linear SVM is trained to perform the facial expression classification. Performance of the proposed approach is evaluated using the new dataset with estimated depth map. Experimental results show that significant improvements on accuracy are achieved as compared to the traditional approach.[9]

S. M., *et.al.* in his article stated human detection in images as a fast growing and challenging area of research in computer vision with its main application in video surveillance, robotics, intelligent vehicle, image retrieval, defense, entertainment, behavior analysis, tracking, forensic science, medical science and intelligent transportation. This paper presents a robust multi-posture human detection system in images based on local feature descriptors such as HOG and BO (Block Orientation). The proposed system employs LLE method to achieve dimensionality reduction on the Hog feature descriptors and thus reduce time complexity. Performance of the proposed method is evaluated using feature and classifier based schemes with different datasets. By using classifier based schemes, fast-additive SVM outperforms other SVM classifiers. The combined feature vector can retain precision of HOG as well as improve the detection rate. The experiment results on INRIA person, SDL dataset, and TUDB russels dataset demonstrate that combined feature vector along with LLE and fast additive SVM significantly improves the performance.[10]

Whereas Yunsheng Jiang, *et.al.* in his paper presents an effective combination models with certain combination features for human detection. In the past several years, many existing features/models have achieved impressive progress, but their performances are still limited by the biases rooted in their self-structures, that is, a particular kind of feature/model may work well for some types of human bodies, but not for all the types. To tackle this difficult problem, combined certain complementary features/models together with effective organization/fusion methods. Specifically, the HOG features, color features and bar-shape features are combined together with a cell-based histogram structure to form the so-called HOG-III features. Moreover, the detections from different models are fused together with the new proposed weighted-NMS algorithm, which enhances the probable "true" activations as well as suppresses the overlapped detections. The experiments on PASCAL VOC datasets demonstrate that, both the HOG-III features and the weighted-NMS fusion algorithm are effective (obvious improvement for detection

performance) and efficient (relatively less computation cost).[11]

3. FACIAL EXPRESSION CATEGORIZATION

After a brief discussion of literature survey in the previous section expression categorization is presented in this section. 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.eg)Eyes, Eyebrows, Mouth, Tissue Textures, Nose. An example of geometric feature extraction is shown in Figure 2. The facial components or facial feature points are extracted to form a feature vector that represents the face geometry

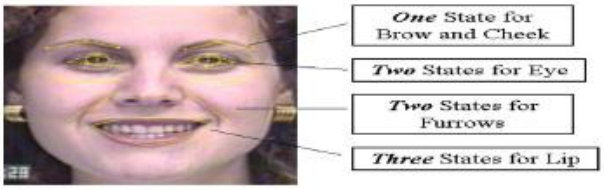

Fig-2: Geometric Feature Extraction

Appearance or transient Features: The features that appear temporarily in the face during any kind of Facial Expression. Eg. Different kinds of wrinkles, bulges, forefront, regions surrounding the mouth and eyes are shown in Figure-3. 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.[14]

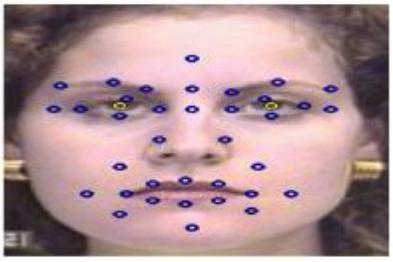

Fig-3 Appearance Feature Extraction

Fig-4: Prototypic facial expressions of six basic emotions

Figure 4 show the prototypic facial expressions of six basic emotions: disgust, happiness, sadness, anger, fear, and surprise.

4. OVERVIEW OF FEATURE EXTRACTION SYSTEM

This section provides a quick brief survey of few popular feature extraction techniques .The paper proceeds with brief idea of few well known methods such as Principal component analysis (PCA), Local binary pattern (LBP), Local directional pattern (LDP), Histogram of Oriented Gradients (HOG), Wavelet transformation, etc.

Principal component analysis- PCA is a basic and classical method. The PCA method computes the covariance matrix of the training images and calculates the eigen values and eigen vectors of this matrix Then the components that correspond to the top few largest eigen values (equal to the corresponding variances) are retained to achieve the purpose of data dimensionality reduction. Finally certain classifiers such as the nearest neighbor algorithm are used to recognize the query face images.[12]

Local binary pattern- LBP achieves salient results when applied to face images. After they detect and align the face image, they divide the whole face into small grids. From each grid, they compute a LBP histogram. And finally they concatenate all histograms together into a feature vector to represent the face image. They use weighted chi-square distance to compute the distance between faces. Details of LBP for face recognition can be found in [14]

Histogram of Oriented Gradients (HOG) is one of the gradient based powerful feature extraction mechanism. It is an easy method. It was developed for the human detection, later it is used to detect animals and also used in object detection **Histogram of Oriented Gradients (HOG)** is one of the gradient based powerful feature extraction mechanisms. For the computation of HOG feature vector, firstly the gradient magnitude and angular orientation is calculated for each pixel. Then, the gradient magnitude is histogrammed into specified number of bins depending upon their angular orientation. This binning is done with 50% overlapping cells. The feature vector obtained is normalized in order to have greater illumination variance.[13]

Wavelet Transform- Wavelet transformation is another approach for the feature extraction techniques. It is a powerful tool for signal representation. Today, it is widely used in image processing (feature extraction), pattern recognition, image compression (JPEG), signal processing etc. Haar wavelet is one of the oldest wavelet transform used. It was proposed in 1910. In the 1980's, lots of wavelets were proposed such as Morlet-wavelet concept (1981), Mallat and Meyer-multi-resolution (1988), Daubechies-compact support orthogonal wavelet (1988), Mallat-fast wavelet transforms (1989). On applying 2D discrete wavelet transform on an image, the image is divided

into 4 parts as an output, each part holding specific information.[8]

The above mentioned techniques along with other feature extaction techniques are summarized in Table 1.

Table 1

NAME	METHOD	PERFORMANCE	DISADVANTAGES
Low-Dimensional Procedure for Characterization of Human Faces	Principal Component Analysis	Recognition rate is low	Only single factor can be varied
Eigen faces vs. Fisher faces : recognition using class specific linear Projection	Fisher's Linear Discriminate	Recognition rate higher than PCA	Global feature vectors are generated
Recognizing Faces with PCA and ICA	Independent Component Analysis	Recognition rate is improved compared to PCA and FLD	Computationally expensive than PCA
Two-dimensional PCA: A new approach to appearance-based face representation and recognition	2-Dimensional Principal Component Analysis	Recognition rate is higher than PCA	Storage requirement is higher than PCA
The importance of the color information in face recognition	Global Eigen Approach using Color Images	YUV color space has highest recognition rate	RGB color space does not provide any improvement in recognition rate
A novel hybrid approach based on sub-pattern technique and E2DPCA for color face recognition	Sub pattern Extended 2-Dimensional Principal Component Analysis	Recognition rate higher than PCA,2DPCA	Variation in lighting, pose are not considered
Face Recognition using a Color Subspace LDA approach	Color Subspace Linear Discriminate Analysis	Recognition rate is higher than 2DPCA and LDA	Variation of performance in color spaces is not evaluated
Multilinear Image Analysis for Facial Recognition	Multilinear Image Analysis	Recognition rate higher than PCA	Less performance than Color Subspace LDA
Gabor Filter Based Face Recognition Technique	2-Dimensional Gabor Filter Bank	Higher recognition rate than PCA, LDA,2DPCA,Global Eigen Approach	Low and high frequency component attenuation
Local binary patterns for multi-view facial expression recognition	Local Gabor Binary Pattern	Better recognition rate than Gabor filter bank	Color information is not included

5. CONCLUSION

Facial expression plays an important role in communication and thus identifying the correct expression is as essential as knowing the exact matter of the communication. Therefore Facial Expression recognition has increasing application areas and requires more accurate and reliable FER system. A large number of methods has already been proposed and applied for the same; still the research is going for the betterment. This paper has presented a survey on facial expression recognition. Recent feature extraction techniques are covered along with comparison. The HOG features are better in comparison to other methods .The research is still going on (i) to increase the accuracy rate of predicting the expressions, (ii) to have applications based on dynamic images/sequence of images/videos, (iii) to handle the occlusion etc. by combining various techniques together.

REFERENCES

[1] S. Kumar, M. K. Bhuyan and B. K. Chakraborty, "Extraction of informative regions of a face for facial expression recognition," in *IET Computer Vision*, vol. 10, no. 6, pp. 567-576, 9 2016 doi: 10.1049/iet-cvi.2015.0273

[2] J. V. Patil and P. Bailke, "Real time facial expression recognition using RealSense camera and ANN," *2016 International Conference on Inventive Computation Technologies (ICICT)*, Coimbatore, 2016, pp. 1-6.

[3] J. d. A. Fernandes, L. N. Matos and M. G. d. S. Aragão, "Geometrical Approaches for Facial Expression Recognition Using Support Vector Machines," *2016 29th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, Sao Paulo, Brazil, 2016, pp. 347-354.

[4] L. A. Jeni, D. Takacs and A. Lorincz, "High quality facial expression recognition in video streams using shape related information only," *2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*, Barcelona, 2011, pp. 2168-2174.

[5] S. Jain, Changbo Hu and J. K. Aggarwal, "Facial expression recognition with temporal modeling of shapes," *2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*, Barcelona, 2011, pp. 1642-1649.

[6] H. Li; H. Zou; H. Hu, "Modified Hidden Factor Analysis for Cross-Age Face Recognition," in *IEEE Signal Processing Letters* , vol.PP, no.99, pp.1-1

[7] S. An and Q. Ruan, "3D facial expression recognition algorithm using local threshold binary pattern and histogram of oriented gradient," *2016 IEEE 13th International Conference on Signal Processing (ICSP)*, Chengdu, 2016, pp. 265-270

[8] J. Kumari, R. Rajesh and A. Kumar, "Fusion of features for the effective facial expression recognition," *2016 International Conference on Communication and Signal Processing (ICCSP)*, Melmaruvathur, 2016, pp. 0457-0461.

[9] T. W. Shen *et al*., "Facial expression recognition using depth map estimation of Light Field Camera," *2016 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*, Hong Kong, 2016, pp. 1-4.

[10] J. S. B. and S. M., "Multi-posture Human Detection Based on Hybrid HOG-BO Feature," *2015 Fifth International Conference on Advances in Computing and Communications(ICACC)*,Kochi,2015,pp.37-40. doi: 10.1109/ICACC.2015.99

[11] Yunsheng Jiang and Jinwen Ma, "Combination features and models for human detection," *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston,MA,2015,pp240-248. doi: 10.1109/CVPR.2015.7298620

[12] Shunfang Wang and Ping Liu "A New Feature Extraction Method Based on the Information Fusion of Entropy Matrix and Covariance Matrix and Its Application in Face Recognition" *Entropy* 2015, *17*, 4664-4683; doi:10.3390/e17074664

[13] Jyoti Kumaria, R.Rajesha, KM.Poojaa "Facial expression recognition: A survey" *in Second International Symposium on Computer Vision and the Internet.* Procedia 2015 Doi:10.1016/j.procs .2015.08.011

[14] C.P. Sumathi1, T. Santhanam2 and M.Mahadevi "Automatic Facial Expression Analysis A Survey" in *International Journal of Computer Science & Engineering Survey (IJCSES) Vol.3, No.6, December 2012* DOI : 10.5121/ijcses.2012.3604 47