A REVIEW ON FACIAL EXPRESSION RECOGNITION USING MACHINE LEARNING

Surbhi Singh[1], Dr. Vishal Shrivastav [2], Dr. Akhil Pandey [3]

[1] *M.Tech Scholar, Computer Science Engineering Department, Arya College of Engineering and IT, Jaipur, Rajasthan, India*
[2] *Professor, Computer Science Engineering Department, Arya College of Engineering and IT, Jaipur, Rajasthan, India*
[3] *Professor, Computer Science Engineering Department, Arya College of Engineering and IT, Jaipur, Rajasthan, India*

ABSTRACT

Human Face expression Recognition is one of the most powerful and challenging tasks in social communication. Generally, face expressions are natural and direct means for human beings to communicate their emotions and intentions. Face expressions are the key characteristics of non-verbal communication. This paper describes the survey of Face Expression Recognition (FER) techniques which include the three major stages such as preprocessing, feature extraction and classification. This survey explains the various types of FER techniques with its major contributions. The performance of various FER techniques is com- pared based on the number of expressions recognized and complexity of algorithms. Databases like JAFFE, CK, and some other variety of facial expression databases are discussed in this survey. The study on classifiers gather from recent papers reveals a more powerful and reliable understanding of the characteristics of classifiers for research fellows.

Keyword: - *expression, machine learning, deep learning, Databases - JAFFE, CK.*

1. INTRODUCTION

Facial expression recognition (FER) has a high impact in the field of pattern recognition, and a substantial effort is made by researchers to develop an FER system for human–computer interaction applications. The facial expression provides sensitive information cues to build an FER system and considered as the best tool for recognizing human emotions and intentions easily defined six distinct expressions (happy, sad, anger, surprise, fear, and disgust) as the basic emotions and each emotion is associated with a unique facial expression which readily recognized across different cultures. The psychologist proposed a study on information communication between humans. The study reveals 55% of information is conveyed by facial expression, 38% by supporting language like sound, speech and so on, and only 7% by oral language. Currently, an FER system plays a central part of artificial intelligence and serves as a potential real-world applications in different areas for psychological studies, driver fatigue monitoring, interactive game design, portable mobile application to automatically insert emotions in chat and assistance systems for autistic people, facial nerve grading in medical field , emotion detection system used by disabled to assist a caretaker, socially intelligent robot with emotional intelligence of the research work in FER system follows the framework of pattern recognition It consists of three phases: face detection, facial feature extraction, expression classification. It is quite substantial and noteworthy to research these phases. In this current survey, various phases of facial expression analysis are discussed with distinct algorithms to classify six basic expressions. Face detection is performed by algorithms such as Haar classifier, adaptive skin cooler algorithm and so on. Gabor feature, local binary patterns (LBPs), active appearance model, principal component analysis and other algorithms

exploited for feature extraction.[1] The classifiers used for expression classification are support vector machine (SVM), neural networks, nearest neighbor and on Facial expressions. Face expressions are the delicate signals of the larger communication. Non-verbal communication means communication between human and animals through eye contact, gesture, facial expressions, body language, and paralanguage. Eye contact is the important phase of communication which provides the mixture of ideas. Eye contact controls the contribution, discussions and creates a link with others. Face expressions include the smile, sad, anger, disgust, surprise, and fear. A smile on human face shows their happiness and it expresses eye with a curved shape. The sad expression is the feeling of looseness which is normally expressed as rising skewed eyebrows and frown. The anger on human face is related to unpleasant and irritating conditions. The expression of anger is expressed with squeezed eye- brows, slender and stretched eyelids. The disgust expressions are expressed with pull down eyebrows and creased nose. The surprise or shock expression is expressed when some unpredicted happens. This is expressed with eye-widening and mouth gaping and this expression is an easily identified one. The expression of fear is related with surprise expression which is expressed as growing skewed eyebrows. FER has the important stage is feature extraction and classification. Feature extraction includes two types and they are geometric based and appearance based. The classification is also one of the important processes in which the above-mentioned expressions such as smile, sad, anger, disgust, surprise, and fear are categorized. The geometrically based feature extraction comprises eye, mouth, nose, eyebrow, other facial components and the appearance based feature extraction comprises the exact section of the face this paper mainly focuses on various FER techniques with three major steps respectively preprocessing, feature extraction and classification. Also, this paper shows the advantages of different FER techniques and the performance analysis of different FER techniques. In this paper, only the image based FER techniques are chosen for the literature review and the video based FER techniques are not chosen. Mostly FER systems meet the problems of variation in illumination, pose variation, lighting variations, skin tone variations. Also this paper gives an essential research idea for future FER research. [2]

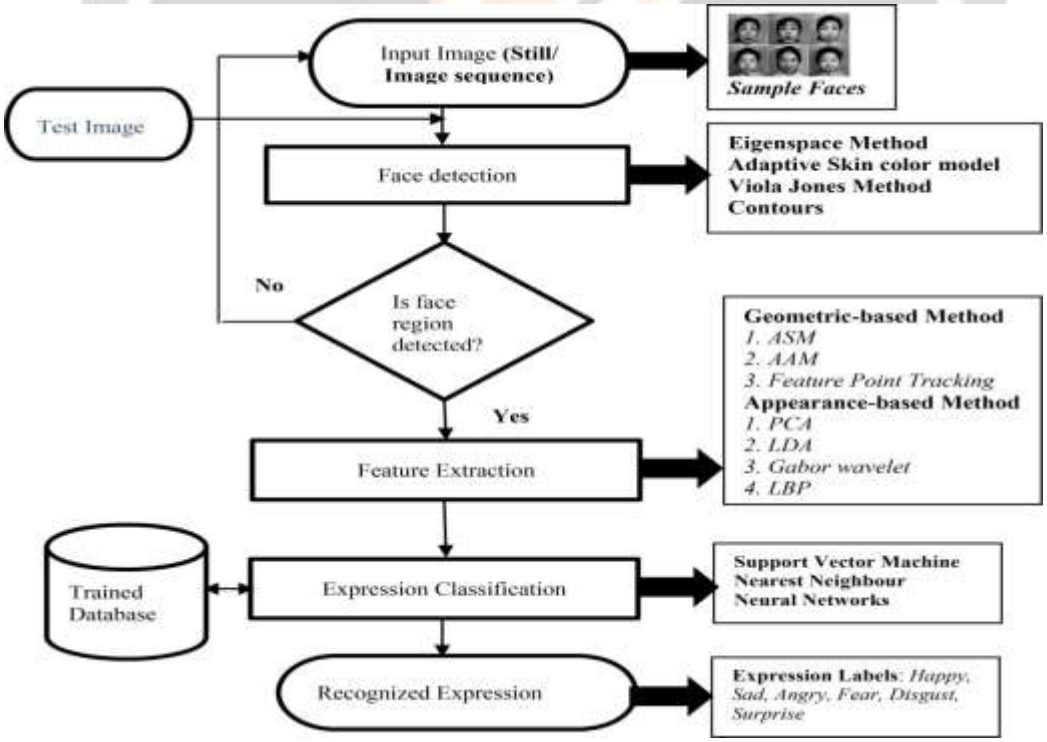

Fig -1: Face Expression Recognition Techniques

2. LITERATURE REVIEW

Many previous image processing methods discard low frequency components of images to extract illumination invariant for face recognition. However, this method may cause distortion of processed images and perform poorly under normal lighting. Although 3D face imaging is increasingly popular, many 3D facial imaging systems have significant noise components which need to be reduced by post-processing if meaningful recognition results are desired.[3] Biometric image recognition is the process of studying the closest match region in between the examining images. The study of the recognition is done about the spatial pixels (picture element) among the image. Recognition of two different biometric features, fingerprint and face images are attempted. One of the major challenges encountered by current Face Recognition (FR) techniques lies in the difficulties of handling varying poses and illuminations. In this paper we propose three novel techniques, viz. Face Recognition (FR) under varying lighting conditions and pose is very challenging. This paper proposes a novel approach for enhancing the performance of a FR system, employing a unique combination of Active Illumination Equalization (AIE), Image Sharpening (IS) the appearance of the face varies drastically when background, pose and illumination change. Variations in these conditions make Face Recognition (FR) [4]

Chao Qi, et,al (2018) .In this paper, a new expression recognition approaches presented based on cognition and mapped binary patterns. Atfirst, the approach is based on the LBP operator to extract the facial contours. Secondly, the establishment of pseudo 3D models used to segment face area into six facial expression sub regions. In this context, the sub-regions and the global facial expression images use the mapped LBP method for feature extraction, and then use two classifications which are the support vector machine and softmax with two kinds of emotion classification models the basic emotion model and the circumflex emotion model. At last, we perform a comparative experiment on the expansion of the Cohn Kanade (CK +) facial expression data set and the test datasets collected from ten volunteers. The experimental results showthat the method can effectively remove the confounding factors in the image. And the result of using the circumflex emotion model is obviously better than the traditional emotional model. By referring to relevant studies of human cognition, we verified that eyes and mouth express more emotion.[5]

He Jun et.al (2015) For facial expression recognition, the LBP feature is an important way of texture feature, but usually the whole of image is taken as extracting area, ignoring to extract the key areas of facial expression. In order to solve this problem, based on previous LBP feature extraction method, as well as the division of facial motion unit, we put forward a kind of expression recognition method using the fusion feature of key facial areas expression based on LBP, by dividing into several parts: eyes, eyebrows, between-eyebrow, nose, mouth, then we get the key areas of expression to extracted features independently, at the meaning time to hold global facial features, features of the whole facial is also extracted. After that we combine this two different features together and get a new feature which is called combine feature fused key expression Ares. The features combined then classified by SVM and NN to recognize different expressions. This article carries on the experiment in JAFFE database, the results show that the method of facial expression recognition rate obtained obvious ascension.[6]

Swapna Agarwal et.al (2019) For synthesis of realistic facial expressions displaying emotions, we need an efficient representation of pure (e.g., surprise) as well as mixed (e.g., happily surprised) emotional expressions. In this paper, we train an expression map (XM) that efficiently represents the emotional expressions. We propose an algorithm that utilizes the XM to synthesize emotional expressions, tailor-made for the facial structure of the target person. The proposed method can also control the proportions of different basic emotional expressions of those, when mixed together, to generate realistic emotional facial expressions. Unlike many existing methods, our expression synthesis model requires only one expression-neutral face image of the target person. Both qualitative and quantitative tests on four data sets show promising results. On average, we have achieved 92.4% correct validation of the expressions synthesized by our method. We also show that for both basic and mixed emotional expressions, our method generates finer expression details compared to existing state-of-the-art works.[7]

Yiming Wang et.al (2017) The main aim of face formalization is to synthesize the frontal facial appearances from non-frontal facial images. How to estimate the frontal face-shape is a crucial but very challenging problem in the formalization task. Most existing methods use a single shape template to fit in with frontal facial appearances, which will result in a loss of expression-related information. In this work, we present a novel facial expression-aware face formalization method which directly learns the pair-wise relations between non-frontal face-shape and its frontal counterpart. The support vector regression is explored to train the pair-wise regression model. Considered the pair-wise relationship is non-linear, an appropriate cascade manner is applied to iteratively adjust and optimize the model. With the estimated frontal shape, facial appearances are synthesized through a texture-fitting process formulated by solving a simple optimization problem. The proposed method has been evaluated on a in-the-wild facial expression database. The experimental results show an outstanding performance of both visual effects of expression recovery and facial expression recognition.[8]

Hari Prasad Mal et.al(2017) Facial expressions are the facial changes indicating internal state of human being, objectives or communal conversation. Subject to the change of emotions on the face, any persons face is the most important mode of conveying and deducing affective states of human ones. On the fly facial expression detection has become a major research area as it plays a key role in Human Computer Interaction. Facial expression detection has major application in areas of social interaction as well as social intelligence. This paper represents the various techniques used in facial expression detection along with system.[9]

Haifeng Zhang et.al (2019) Research shows that the facial expression recognition is strongly related to the person's identity. This paper presents an expression-identity fusion network to address the great inter-subject variations in facial expression recognition. The model is designed to jointly learn identity-related features and expression-related features via two branches with the same expression image input. A bilinear module is introduced to fuse two kinds of features and learn the relationship between them. Experimental results show that identity-related features can greatly boost the performance of facial expression recognition. Our method outperforms most of the state-of-the-art. On two popular facial expression databases (CK+ and Oulu-CASIA), our method achieves 96.02% and 85.21% recognition accuracy, respectively.[10]

3. FACE EXPRESSION RECOGNITION APPROACH

The overview of the FER system is illustrated in fig.2 The FER system includes the major stages such as face image preprocessing, feature extraction and classification.

Preprocessing: Preprocessing is a process which can be used to improve the performance of the FER system and it can be carried out before feature extraction process Image preprocessing includes different types of processes such as image clarity and scaling, contrast adjustment, and additional enhancement processes to improve the expression frames .

The cropping and scaling processes were performed on the face image in which the nose of the face is taken as midpoint and the other important facial components are included physically Bessel down sampling is used for face image size reduction but it protects the aspects and also the perceptual worth of the original image (The Gaussian filter is used for resizing the input images which provides the smoothness to the images Face alignment is also the preprocessing method which can be performed by using the SIFT (Scale Invariant Feature Transform) flow algorithm. For this, first calculate reference image for each face expressions. After that all the images are aligned through related reference images ROI Region of Interest) segmentation is one of the important type of preprocessing method which includes three important functions such as regulating the face dimensions by dividing the color components and of face image, eye or forehead and mouth regions segmentation In FER, ROI [11]

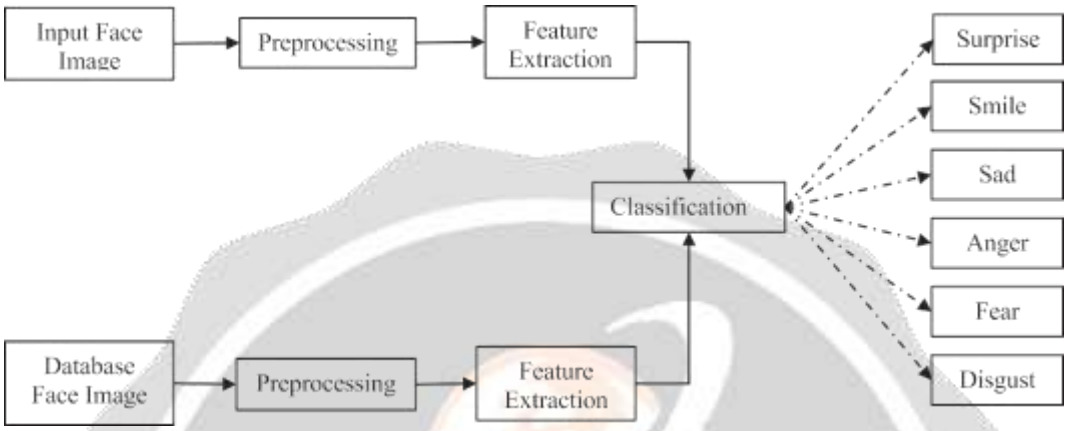

Fig -2: Architecture of Face Expression Recognition System

Segmentation is most popular because for convenient segmentation of face organs from the face images. The histogram equalization method is used to conquer the illumination variations this method is mainly used for enhancing the contrast of the face images and for exact lighting also used to improve the distinction between the intensities.[12] In FER, more preprocessing methods are used but the ROI segmentation process is more suitable because it detects the face organs accurately which organs are is mainly used for expression recognition. Next the histogram equalization is also another one important preprocessing technique for FER because it improves the image distinction.

Feature extraction: Feature extraction process is the next stage of FER system. Feature extraction is finding and depicting of positive features of concern within an image for further processing. In image processing computer vision feature extraction is a significant stage, whereas it spots the move from graphic to implicit data depiction. Then these data depiction can be used as an input to the classification. The feature extraction methods are categorized into five types such as texture feature-based method, edge based method, global and local feature-based method, geometric feature-based method and patch-based method.

The descriptors which extract the features based on the edge based methods are described as follows. Line Edge Map (LEM) descriptor is a facial expression descriptor which improves the geometrical structural features by using the dynamic two strip algorithm (Dyn2S)). Based on the motion analysis two types of facial features are extracted such as non discriminative and discriminative facial features Graphics-processing unit based Active Shape Model (GASM) is the feature extraction method which can be performed with edge detection, enhancement, tone mapping and local appearance model matching. After that the image ratio features are extracted from the expressed face images (Song et al., 2010). Histogram of Oriented Gradients (HOG) is a window supported feature descriptor which uses the gradient filter. The extracted features are based on the edge information of the registered face images. It extracts the visual features, for example a smile expression means curvature shaped eyes.[13]

The descriptors which extract the features based on the global and local feature-based methods are described as follows. Principal Component Analysis (PCA) method is used for feature extraction. It extracts the global and low dimensional features. Independent Component Analysis (ICA) is also a feature extraction method which extracts the local features using the multichannel observations Stepwise Linear Discriminated Analysis (SWLDA) is the feature extraction technique which extracts the localized features with backward and forward regression models. Depends

on the class labels the F-test values are estimated for both regression models.

The descriptors which extract the features based on the geo- metric feature-based methods are described as follows. Local Curve let Transform (LCT) is a feature descriptor which extracts the geometric features which depends on wrapping mechanism. The extracted geometric features are mean, entropy and standard deviation Addition to these geometrical features energy, kurtosis are extracted by using three stage steerable pyramid representation.
The descriptors which extract the features based on patch- based methods are described as follows. Facial movement features are extracted as patches depending upon the distance characteristics. These are performed by using two processes such as extracting the patches and patch matching. The patch matching is performed by translating extracted patches into distance characteristics. [14]

The texture feature based descriptors are more useful feature extraction method than the others because it extracts the texture features like related to the appearance which provides the important feature vectors for FER. Also Local Directional Number (LDN) pattern, Local Directional Ternary Pat- tern (LDTP) KL-transform Extended LBP (K-ELBP) and Discrete Wavelet Transform (DWT) texture feature based descriptors are used as feature descriptors in recent years FER.

Dataset: The database contains either static images or a sequence of images. The CK database extended CK (CK+) database, JAFFE database and so on, contains potential information to recognize different facial expressions under constrained environment. However, posed expressions may occur in the real- time environment due to different illumination conditions, pose variations, and occlusions. Databases like MMI database Oulu-CASIA database, AFEW & SFEW database and so on, were considered to handle not only frontal-view but also dual- view with varying lighting conditions. Here we discuss a few facial expression databases elaborately. JAFFE database: The JAFFE database is extensively used dataset for facial expression analysis. It includes static images with six basic expressions (happy, sad, fear, angry, disgust and surprise) along the neutral face that were posed under a controlled environment. Each subject has posed three or four samples for every single expression. The image resolution is 256×256 pixels, and the semantic ratings of the expressions were calculated on seven emotion categories by 60 female subjects as ground truth. The expression images are annotated based on the predominant expression of a particular image. CK database: The CK database is a popular benchmark dataset. The main purpose of this database is to evaluate face recognition and FER. It contains 486 video sequences from 97 subjects with neutral to apex displays. [15]

Fig -3: Sample images from the CK+ database

The frames have a resolution of 640×480 or 640×490 pixels and are fully FACS coded. Annotation of six basic expressions is also provided. The extended version (CK+) includes 593 posed expression sequence from 122 spontaneous smile sequences of 66 subjects. CK and CK + do not include occluded faces. Some sample images of CK+ database are shown in Fig. 3

4. CONCLUSIONS

Face detection is not an easy task since it is influenced by various factors such as non-uniform illumination, pose variation, occlusion, and complex background and so on. Adaptive skin color model is suitable to detect the face in the case of complex background and also shows high accuracy. In most of the research, algorithms such as Haar classifier and Gadabouts are used to detect faces in an image or video sequence due to its high accuracy and low-computational cost. Under the unconstrained environment, Haar classifier detects face efficiently and quickly than skin color model For real-time face tracking and motion detection, LK optical flow algorithm and MRASM methods are used in real time as it provides better efficiency and has increased stability in the extraction of facial landmarks. Geometric-based feature method describes the shape and structure of face components such as eyebrows, eyes, nose, and mouth. The feature extraction algorithms of these methods such as ASM and AAM are well suited in real- time applications due to efficient tracking. Appearance-based feature method defines the texture that appears on the face during expression and it has high discriminative power compared to geometric-based method but not suitable for real-time due to high computational cost and memory storage. The LBP, HOG, Gabor filter-based texture information, local directional ternary pattern provide the appearance features were employed to detect the facial expression of a static image and video.

5. REFERENCES

[1]. Arman Savran, B.S., 2017. Non-rigid registration based model-free 3D facial expression recognition. Comput. Vis. Image Underst. 162, 146–165. https:// doi.org/10.1016/j.cviu.2017.07.005.

[2]. B, K.S., Rameshan, R., 2017. Dictionary Based Approach for Facial Expression Recognition from Static Images. Int. Conf. Comput. Vision, Graph. Image Process. pp. 39–49.

[3]. Bashyal, S., Venayagamoorthy, G.K.V., 2008. Recognition of facial expressions using Gabor wavelets and learning vector quantization. Eng. Appl. Artif. Intell. 21, 1056–1064. https://doi.org/10.1016/j.engappai.2007.11.010.

[4]. Biswas, S., 2015. An Efficient Expression Recognition Method using Contourlet Transform. Int. Conf. Percept. Mach. Intell. pp. 167–174.

[5]. Chao Qi, Min Li, Qiushi Wang, Huiquan Zhang, Jinling Xing,ZhifanGao, Member, IEEE, and Huailing Zhang Facial Expressions Recognition Based on Cognition and Mapped Binary Patterns DOI 10.1109/ACCESS.2018.2816044, IEEE Access 2169-3536 (c) 2018 IEEE

[6]. He Jun CaiJian-feng Fang Ling-zhi He Zhong-wen A method of facial expression recognition based on LBP fusion of key expressions areas The 27th Chinese Control and Decision Conference (2015 CCDC) Year: 2015 ISBN: 978-1-4799-7017-9 DOI: 10.1109/IEEE Qingdao, China

[7]. Swapna Agarwal Dipti Prasad Mukherjee Synthesis of Realistic Facial Expressions Using Expression Map IEEE Transactions on Multimedia Year: 2019 DOI: 10.1109/IEEE

[8]. Yiming Wang Hui Yu JunyuDongMuweiJianHonghai Liu Cascade support vector regression-based facial expression-aware face frontalization2017 IEEE International Conference on Image Processing (ICIP) Year: 2017 ISBN: 978-1-5090-2175-8 DOI: 10.1109/ IEEE Beijing, China

[9]. Hari Prasad Mal P. SwarnalathaFacial expression detection using facial expression model 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS) Year: 2017 ISBN: 978-1-5386-1887-5 DOI: 10.1109/ IEEE Chennai, India

[10]. Haifeng Zhang Wen Su Zengfu Wang Expression-identity Fusion Network for Facial Expression Recognition ICASSP 2019 - 2019 IEEE International Conference on

[11]. Boqing Gong, YuemingWang, Jianzhuang Liu, X.T., 2009. Automatic Facial Expression Recognition on a Single 3D Face by Exploring Shape Deformation. Proc. 17th ACM Int. Conf. Multimed. pp. 569–572.

[12].Chang, H.T.Y., 2017. Facial expression recognition using a combination of multiple facial features and support vector machine. Soft Comput. 22, 4389–4405. https://doi.org/10.1007/s00500-017-2634-3.

[13]. Clawson, K., Delicato, L.S., Bowerman, C., 2018. Human Centric Facial Expression Recognition. Proc. Br. HCI pp. 1–12.

[14].Cossetin, M.J., Nievola, J.C., Koerich, A.L., 2016. Facial expression recognition using a pairwise feature selection and classification approach. IEEE Int. Jt. Conf. Neural Networks, pp. 5149–5155.

[15].Cui, R., Liu, M., Liu, M., 2016. Facial expression recognition based on ensemble of mulitple cNNs. Chinese

Conf. Biometric Recognit. 511–518. https://doi.org/ 10.1007/978-3-319-46654-5.

[16].Dahmane, M., Meunier, J., 2014. Prototype-based modeling for facial expression analysis. IEEE Trans. Multimed. 16, 1574–1584.

[17].Demir, Y., 2014. A new facial expression recognition based on curvelet transform and online sequential extreme learning machine initialized with spherical clustering. Neural Comput. Appl. 27, 131–142. https://doi.org/10.1007/s00521- 014-1569-1.