International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2017 IJSRCSEIT | Volume 2 | Issue 1 | ISSN : 2456-3307

To Recognize Human Emotions Based on Facial Expression Recognition : A Literature Survey

Komal D. Khawale*, D. R. Dhotre

Department of Computer Engineering, Amravati University, SSGMCE, Shegaon, Maharashtra, India

ABSTRACT

Facial Emotion Recognition has been a very significant issue and an advanced area of research in the field of Human- Machine Interaction and Image Processing. Human-Machine relation is a major field for that different approaches have been proposed for developing methods for recognition of automated facial emotion analysis using not only facial expressions, also speech recognition. Facial expression detection the multiple varieties of human faces like texture, color, shape, expressions etc. are considered. Firstly, to detect a facial emotions of the human with variations in the facial movements including mouth, eyes, and nose are to be determined and after that considering those features using a very good classifier to recognize the human emotions. This paper gives a brief summary of emotion recognition methods like Feature Fusion, Deep Auto-Encoder, Sigma Pi-Neural Network, Genetic Algorithm, PHOG and Hierarchical Expression Model etc. which are used to recognize human emotions are presented.

Keywords: Human Emotion Recognition, PHOG, Emotion Recognition, Expression Recognition.

I. INTRODUCTION

The expressions of a face are studied as the most crucial ways to commiserate human emotional state of a particular individual to social communications. It is also referred as a type of non-verbal communication. The most significant issue in the field of research on Human-Machine Interaction and intelligent control systems is facial expressions. In Man-Machine intercourse community, face expression detection does not compulsorily used for the recognition emotions but it relates to the division of face features into six emotional states as shown in Fig.3. The proposed system of expression recognition is accurately used in the field of security, military, biometric authentication, and medical science, and manufacturing and design units. Man Emotion Acknowledgment could have numerous significant usages and accomplishments. The human computer interaction, patient monitoring and social intelligence Etc. may use the areas of emotion recognition. In a hypothecated form of a facial expression detection system, a camera device which acts as a input device and capture the query image (input image) and after that it interacts with the machine. Then findings of the features that come out from the face region is carried out and after that, the classifiers like LDA or KNN is used to classify them into six basic emotion type like happiness, sadness, disgust, anger, and fear etc. Various types of face recognition or detection techniques and classification algorithms [2] that can be used to detect and classify the emotions of the face using a significant concept of the face i.e. face expression detection. The proposed paper contains various algorithms for face expression recognition is reviewed. Both facial expression approaches and speech recognition approaches were considered for the recognition of emotions are viewed and equated. In this paper techniques that are dealing with the facial expressions are revisited. Some customized algorithm for this is considered in this paper. Feature Fusion [3], Deep Auto-Encoder [4], PHOG and Hierarchical Expression Model [5], Sigma Pi- Neural Network [6], Kernel methods and Statistical Model [7], and Genetic Algorithm [8] are considered for study in this paper. These algorithms are customized in the sense that they are designed to

conquer the expression variations on the same subject. JAFFE database, Cohn-Kanade database, BAUM-2 database, RaFD database, etc. were some of the databases recommend by the authors of these papers for experimentation. The paper is structured as follows. Chapter II gives a concise description of the methodology for Facial Emotion Recognition. The next chapter III gives the conclusion on different methods used in the proposed paper is discussed.

II. METHODS AND MATERIAL

Literature Survey

This section presents a summary of all these techniques with their advantages and disadvantages. There are various approaches of recognition of emotions from face expressions have been improved and accomplished over past few years.

A. To Recognize facial expressions with emotion-based feature fusion

In 2015, [3] proposed innovative approach facial expression recognition with emotion-based feature fusion using the Discriminant-Analysis of Canonical Correlations (DCC) [9]. In the proposed method, four efficient descriptors are considered like (Weber Local Descriptor) WLD [12], Local Binary Pattern (LBP) [10], (Local Phase Quantization) LPQ [11], and (Pyramid of Histogram Oriented Gradient) PHOG [13]. For dimensionality reduction and manifold learning the Supervised Locality Preserving Projection (SLPP) is considered [14]. The Fig. 1 below shows the Emotion-based feature fusion scheme for facial expression recognition with training and testing databases. In the proposed scheme the first stage is input test image then the four descriptors are used since

- Mostly used for facial expression recognition and
- They represent facial expressions in terms of various aspects like intensity, shape, and phase.

WLD consists of two components –

1) Differential excitation
2) Orientation.

LBP is the first descriptor used in our approach and it also referred as a texture descriptor. The advantage of LBP is that it is insensible to monotonic variations caused by illumination changes. LPQ is the second descriptor considered, also proposed as a texture descriptor. It uses the blur invariance property of Fourier phase information with the assumption that blur is centrally symmetric.

PHOG is an extension of HOG descriptor [15], which is commonly used for object recognition. It represents an image using its local shape of different scales. The canny edge detector is used which divides the image into spatial cells which are based on the number of levels. The next stage is manifold learning preserves the locality information with the help of class information then DCC is borrowed to fuse the best two features set by planning them into a coherent subspace. After that, the proposed classification method uses the adaptive descriptor selection algorithm to further enhance the interpretation of a facial expression recognition system.

The experiments were conducted on three databases like JAFFE[16], BAUM-2 database [17] and a combination of above two databases. JAFFE consists of images from 10 Japanese females that express 6 basic emotions and the neutral BAUM-2 database consists of expression videos extracted from movies.

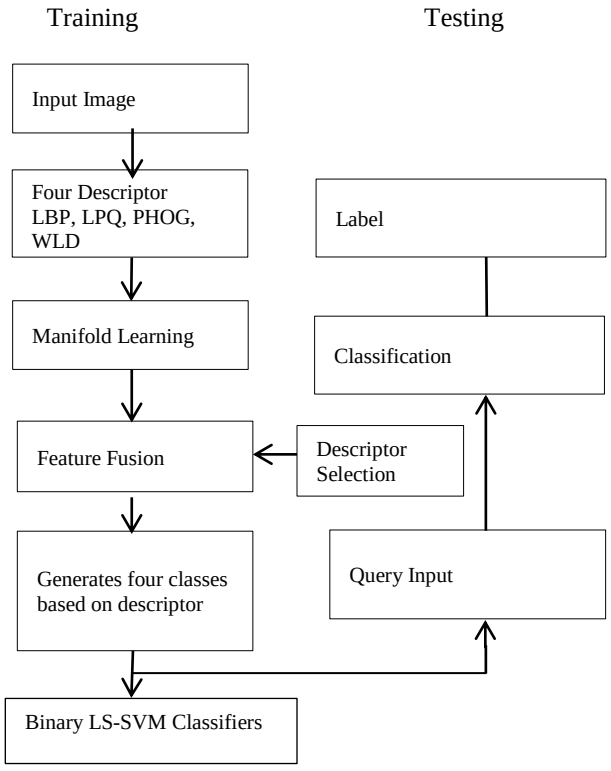

Figure 1. Stages of Emotion –Based feature

Recognition rates are greater for fused features than the individual descriptors and the adaptive descriptor selection algorithm enhances the accuracy up to 2% for JAFFE, BAUM-2, and BAUM-2+JAFFE databases. The results show that the BAUM-2 database (68.71%) is lower than for JAFFE (92.13%) as the images are extracted from a movie that contains the pose, illuminations and resolution variations.

B. Research on speech emotion recognition based on deep auto-encoder

In 2016, [4] proposed a new approach based on deep autoencoder for speech emotion recognition. The method of deep auto-encoder (DAE) were divided into five secret layers and used to extract the phonetic features which are based on the theory of deep learning. Deep auto-encoder (DAE) is a special type of deep neural network and is used to extract the speech emotion feature. DAE is an unsupervised learning method which handles a huge amount of unlabelled data. There are two uses of DAE as follows.

- ✓ It can learn the efficient encoding mechanism, and,
- ✓ It can extract the representation of original data in the hidden layers, namely, feature extraction.

In the proposed method, layers are divided into an input layer, more than one secret layers, and a production layer. The first step, the query input data, which is divided the audio into short frames where every frame of speech emotion signal was the breakdown with wavelet and then estimated the Fourier Transform. There are some high-level features and traditional features are extracted then the SVM was used for classification and recognition.

The experiments were carried out on Chinese Academy of Sciences Institute of Automation of speech emotion database (CASIA) [21] which is created by four actors. There are four traditional speech features for extraction of emotion features from speech signals that includes Mel Frequency Mel cepstral coefficients (MFCC) [18], Perceptual Linear Prediction cepstral coefficient (PLP) [20], Linear Predictive Cepstral Coefficients (LPCC) [19], after comparing with the traditional features, the highest accuracy rate of 86.41% for the speech emotion feature DAE as compared to other traditional voice features.

C. PHOG and a Hierarchical Expression Mode based Emotion Recognition

In 2013, [5] proposed an innovative approach for the recognition of facial emotions using PHOG and a hierarchical expression model. To detect the human facial emotions the proposed approach used the PHOG feature descriptors and a hierarchical multiple stage scheme anterior images of human faces, also the algorithm to accomplish an adequate detection delicacy in real time with lesser time and space complexities. In contrast with the other traditional approaches, it makes it applicable to low cost hardware like the mobile device. The researchers were conducted on the Cohn-Kanade (CK) database to verify the extendibility of the proposed framework. The two classes of experiments were proposed for emotion detection and gender classification. The Cohn-Kanade (CK) database [22] contains the 90 university Students from 19 to 32 years old. Through them 70 percent were female, 10 percent were Indian faces, and 3 percent were Asian. Total of 394 images were used in the experiments. In the meanwhile, the experiments demonstrated that the suggested approach has good robustness and extensibility.

D. Research on Emotion Recognition using Sigma-Pi Neural Network

In 2016, [6] suggested a new approach based on the problem of emotion state recognition using sigma-pi ANN. In the proposed method, the two different types of activation functions were used such as Sigmoid and bell-shaped. An algorithm is also proposed for the high approximation accuracy particularly for non-linear processes in real time. The experiment on the emotions recognition and learning algorithm was accomplished on the images from the faces database [23] and the public images.

E. Genetic Algorithm for emotion recognition based on feature selection

In 2016, [7] proposed a new approach to feature selection, which is an important aspect in facial emotion classification based on Genetic Algorithm. In the proposed system, facial emotion recognition consists of three stages- 1) Extraction of features 2)Selection of features and 3)Classification of features.

The diagrammatic representation of stages for facial emotion recognition as shown in Fig.2-

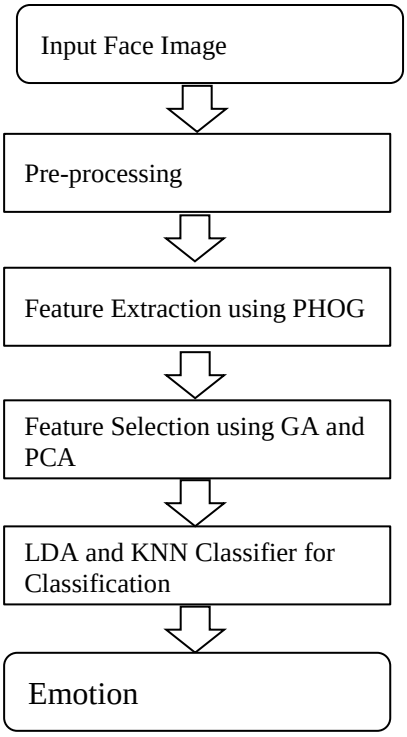

Figure 2. Phases for Facial Emotion Recognition

- ✓ Input face image and Pre-processing - The first stage is a pre-processing stage with the input face image having identical shape and size. The detection of the face is carried out with the used method of [24] based on the haar-like property and AdaBoost learning algorithm [24]. After that, the edge detection using canny edge detection operator and images are adjusted to 110*110 pixels.
- ✓ Extraction of feature using PHOG- The next stage is feature extraction. PHOG is used to extract the feature vector. It is an extension of HOG and uses a similar process but having a different structure like a pyramid. In this stage the feature extraction using PHOG extracting features in a hierarchical way [25].
- ✓ Selection of Features using GA and PCA - GA algorithm is used to select the features from the set of features by elimination certain values that seem insignificant knowledge around face emotions. The proposed framework of GA feature selection consists of various stages includes generation of the initial population, the compatibility value is calculated for each individual solution, after that the selection of best compatibility value and then undergo crossover and mutation. The whole process is repeated until we get the better solution.
- ✓ Classification of features- The last stage is classification using LDA and KNN classifiers [26, 27]. LDA classifier intends to perform a maximum class separation by finding the maximum distance between classes and minimum distance within classes. The KNN is also one of the most applied classification algorithms for classifying unlabeled samples. The experiments were carried out on the Radboud face database (RaFD) [28]. The total of 76 samples having male, female, and children's are included. The images are scaled to 681*1024 pixels with colourful jpg file format. The two dataset of training having 305 images of size 110*110 pixel and 149 images for testing out of total 454 images. Here, the used of six basic emotions such as fear, anger, surprise, happiness, disgust, and sadness.

The results show that the GA as feature selector expressively enhanced the accuracy as compared to other approaches such as PCA which is well known as dimensionality reduction techniques. The recognition accuracy of 99.33% for the proposed GA+PHOG+LDA and 88% for PCA+PHOG+KNN [29]. From the experimental analysis, the GA feature selection uses the feedback from the classifier to select a best performance feature subset. On the other hand, PCA involves only a single iteration and do not enclose such feedback. Another benefit of GA as the feature selection is that it truthfully decreases the number of features in the pattern vectors.

F. Kernel Methods and Statistical Models based on facial expression analysis for recognition of emotions

In 2014, [8] proposed a new approach based on facial expression analysis using Kernel methods and Statistical Models for the recognition of human emotions. In the proposed method, an AAM [30] is developed. For facial expression recognition the eyes, eyebrows, and mouth are considered through this the detection and tracking of features are done using the Active Appearance Model (AAM) [30]. It not only observes the characterization methodology from a parametric model but also quantitatively estimated the accuracy for feature detection and evaluation of the parameters associated with facial expressions, analyzing its roughness to adaptation in pose. Again, the procedure of emotion depiction is recommended to achieve the acknowledgment. Additionally, for emotion recognition, the kernel method using deluge classifiers are accomplished.

The experiments were carried out on Cohn-Kanade [31]. The proposed database contains the face images

having facial expressions and emotions with six basic emotions types such as surprise, sadness, Disgust, Anger, Joy, and Fear[1]. The CK database contains the total of 486 sequences from 97 posers.

Figure 3. Shows an example of this database.

The results present that the suggested model can efficiently distinguish the various facial expressions. The model adopts and depiction approaches occurrences effective to expose the emotion like in 92.4% of the cases. Table I shown in below provides the summary of efficient approaches to facial expression recognition techniques.

Table I. Literature Survey on Facial Expression Recognition

Sr. No	Title	Year	Accuracy/Method	Conclusion
1	To Recognize facial expressions with emotion-based feature fusion	2015	- Recognition rates are greater for fused features. - Experiments were carried out on JAFFE, BAUM-2 database.	This paper introduced Feature Fusion method for the Emotion-Based system.
2	Research on speech emotion recognition based on deep auto-encoder	2016	- Highest accuracy of 86.41% for speech emotion feature recognition. - CASIA.	Accuracy rates are higher for speech emotion recognition as compared to other traditional methods.
3	PHOG and a Hierarchical Expression Mode based Emotion Recognition	2013	- Experiments were carried out on Cohn-Kanade database. - The proposed approach having greater robustness and extensibility.	It has adequate detection accuracy in lesser time and space complexities.
4	Research on Emotion Recognition using Sigma-Pi Neural Network	2016	- Proposed approach having higher accuracy for non-linear processes. - Experiments carried out on Caltech face database.	This paper introduced sigma-pi neural network for addressing an emotion recognition that enhances the recognition accuracy.
5	Genetic Algorithm for emotion recognition based on feature selection	2016	- The proposed method has recognition accuracy of 99.33%. - RaFD database.	Higher accuracy for GA+PHOG+LDA.
6	Kernel Methods and Statistical Models based on facial expression analysis for recognition of emotions	2014	- Characterization methodology shows detection of emotions in 93.4% of cases. - Cohn-Kanade AU-Coded Facial Expression database.	The proposed method uses Kernel methods and statistical models for emotion recognition.

III. CONCLUSION

In this paper, we reviewed some of the facial emotion recognition and also give the brief idea about the algorithm used by these techniques based on both facial expressions and speech features. The expressions that are considered for recognition are happy, sad, anger, disgust, fear, and surprise the techniques here are considered give a good recognition rate with JAFFE, Cohn- Kanade, and RaFD database. The maximum recognition accuracy for RaFD face database is almost

99.33% for GA+PHOG+LDA, while the accuracy of the BAUM-2 database is lower than JAFFE as it contains video with variations in pose and illuminations. There are some techniques which are considered so that recognition accuracy increases and achieves better performance. We also focused and estimated different algorithms and ideas, used to solve some real life applications.

IV. REFERENCES

[1]. P. Ekman and Friesen, "Constants across cultures in the face and emotion," Journal of Personality and Social Psychology, vol. 17, no. 2, pp.124–129, 1971.

[2]. Ashim Saha, Anurag De, Dr. M .C. Pal and Nirmalya Kar, "Different techniques of automatic facial expression recognition: a survey," Proc. of the Second Intl. Conf. on Advances in Computing, Communication and Information Technology- CCIT 2014.

[3]. Cigdem Turan, Kin-Man Lam, Xiangjian He,"Facial Emotion Recognition with Emotion-Based feature Fusion", Proceedings of APSIPA Annual Summit and Conference, 2015.

[4]. Fei Wang, Xiaofeng Ye, Zhaoyu Sun, Yujia Huang, Xing Zhang, Shengxing Shang,"Research on speech emotion recognition based on deep auto-encoder",Proc. of the 6th Annual IEEE International Conference on Cyber Technology in Automation, Control ,and Intelligent Systems June 19-22, 2016, Chengdu, China.

[5]. Zhao Zhong, Gang Shen, Wenhu Chen," Facial Emotion Recognition Using PHOG and a Hierarchical Expression Model", 5th International Conference on Intelligent Networking and Collaborative Systems,2013.

[6]. Nonna Kulishova," Emotion Recognition Using Sigma-Pi Neural Network", IEEE First International Conference on Data Stream Mining, Lviv, Ukraine, August 2016.

[7]. Hadjer Boubenna , Dohoon Lee," Feature Selection for Facial Emotion recognition Based on Genetic Algorithm", 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), 2016.

[8]. Hern´an F. Garcia,Cristian A. Torres, and Jorge Ivan Marin Hurtado,"Facial Expression Analysis for Emotion Recognition Using Kernel Methods and Statistical Models", 2014.

[9]. T.-K. Kim, J. Kittler, and R. Cipolla, Discriminative learning and recognition of image set classes using canonical correlations, Pattern Analysis and Machine Intelligence, IEEE Transactions on, 29(6): p. 1005-1018, 2007.

[10]. T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," Pattern Analysis and Machine Intelligence, IEEE Transactions on, 24(7): p. 971-987, 2002.

[11]. V. Ojansivu and J. Heikkilä, Blur insensitive texture classification using Local phase quantization, in Image and signal processing, Springer.pg. 236-243, 2008.

[12]. J. Chen et al., WLD: A robust local image descriptor, Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2010.

[13]. A. Bosch, A. Zisserman, and X. Munoz, Representing shape with a spatial pyramid kernel, in Proceedings of the 6th ACM international conference on Image and video retrieval, 2007.

[14]. X. Niyogi, Locality preserving projections. In Neural information Processing systems, 2004.

[15]. N. Dalal and B. Triggs, Histograms of oriented gradients for human detection, in Computer Vision and Pattern Recognition, 2005, CVPR 2005, IEEE Computer Society Conference on, 2005.

[16]. M.J. Lyons et al., The Japanese female facial expression (JAFFE) Database, 1998.

[17]. C.E. Erdem, C. Turan, and Z. Aydin, BAUM-2: a multilingual audiovisual affective face database, Multimedia Tools and Applications, pg. 1-31, 2014.

[18]. Maorong Wang, Ping Zhou, Xinxing Jjing. Mixing parameters of MFCC and short TEO energy used in speaker recognition Microelectronics and Computer, 2016.

[19]. Wenjing Han, Haifeng Li, Huabin Ruan, Lin Ma. Review of the progress in the research of speech emotion recognition. Journal of Software, 25(1):37-50, 2014.

[20]. Nirmesh J. Shah, Bhavik B. Vachhani, Hardik B. Sailor and Hemant A. Patil, "Effectiveness of PLP-Based Phonetic Segmentation for speech synthesis", IEEE International Conference on

Acoustic, Speech and Signal Processing (ICASSP), 2014.
[21]. Kohei Arai and Rosa Andrie, "Gait Recognition Method Based on Wavelet Transformation and Its Evaluation with Chinese Academy of Sciences (CASIA) Gait Database as a Human Gait Recognition Dataset", Ninth International Conference on Information Technology- New Generations, 2012.
[22]. T. Kanade, J. F. Cohn and Y. Tian, "Comprehensive database for facial expression analysis,"In Automatic Face and Gesture Recognition, Proceedings. Fourth IEEE International Conference on, pp. 46-53, 2000.
[23]. Ch.-Yi Lee, and Li-Ch. Liao, "Recognition of Facial Expression by Using Neural-Network System with Fuzzified Characteristic Distances Weights", in Proceedings of IEEE Int. Conf. Fuzzy Systems FUZZ-IEEE 2008.
[24]. P.Viola and M.Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," IEEE CVPR, 2001.
[25]. A. Bosch, A. Zisserman, and X. Munoz, "Representing shape with a spatial pyramid kernel," ACM International Conference on Image and Video Retrieval, 2007.
[26]. P.N. Belhumeur, J. Hespanha, and D.J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection" , 1996.
[27]. B. V. Dasarathy (Editor), "Nearest neighbor (NN) norms: NN pattern classification techniques", IEEE Computer Society Press, Los Alamitos, California, 1990.
[28]. O. Langner, R. Dotsch, G. Bijlstra, D. Wigboldus, S. Hawk, and A. Knippenberg, "Presentation and validation of the Radboud Faces Database," Cognition and Emotion 24 (8), 2010.
[29]. Shirin Shokrani, Payman Moallem, and Mehdi Habibi, "Facial emotion recognition method based on Pyramid Histogram of Oriented Gradient over three direction of head," IEEE Computer and Knowledge Engineering (ICCKE), 2014.
[30]. I. Matthews and S. Baker, "Active appearance models revisited," Int. J. Comput. Vision, vol. 60, pp. 135–164, November 2004.
[31]. P. Lucey, J. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, "The extended Cohn-Kanade dataset (CK+): A complete dataset for action unit and emotion-specified expression," in Computer Vision and Pattern Recognition Workshops (CVPRW), IEEE Computer Society Conference on, June 2010.