International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 02 | Feb 2020 www.irjet.net p-ISSN: 2395-0072

A Survey on Human Facial Expression Recognition Techniques

Raj Baldania[1], Kaushal Patel[2], Barkha Bhavsar[3]

[1]MTECH Student, Dept. of Computer Engineering, LDRP Institute of Technology and Research, Gandhinagar, Gujarat, India 382015
[2]MTECH Student, Dept. of Computer Engineering, LDRP Institute of Technology and Research, Gandhinagar, Gujarat, India 382015
[3]Professor, Dept. of Computer Engineering, LDRP Institute of Technology and Research, Gandhinagar, Gujarat, India 382015

Abstract - *Human Facial Expression Recognition plays an important role in the field of computer vision and human computer interaction. It has become interesting due to its wide application in industries. Detecting facial expressions and extracting them helps to recognize the emotions of a human. Various applications of human facial expression recognition in real world such as Driver Fatigue Monitoring [1], Security systems, Facial Emotion Detection in Interview [2], Feedback extraction of customer [3], etc. This paper will give a short explanation and comparison of some of the Human Facial Expression Recognition techniques or methods are introduced which were proposed by some researchers. The first thing which we are also going to present in this paper is and overview or introduction of Human Facial Expression Recognition System, the basic emotions which are identified by the methods and the steps of expression extraction.*

***Key Words*: Facial Expression Recognition, Feature Extraction, Support Vector Machine, Principal Component Analysis, Convolution Neural Network, Back Propagation Neural Network, AdaBoost, Histograms of Oriented Gradients.**

1. INTRODUCTION

Facial expression basically occurs because of variations in human emotions or due to constriction of facial muscle groups. In terms of communicating and expressing, the human face has first-rate significance. Emotions like happiness, wonder, sadness, fear, anger, and disgust are the six simple feelings that everyone makes use of, for his or her day to day conversation. Current recognition systems analyze six basic expressions (joy, fear, anger, disgust, sadness and surprise) and classify them into emotion categories, rather than facial action.

In fact, human beings are capable of producing thousands of expressions varying in complexity which means that are not fully captured with a restrained quantity of expressions and emotion classes.

Face recognition is an interesting and successful application of Pattern recognition and Image analysis [4]. Machine recognition of faces is gradually becoming very important due to its wide range of commercial and law enforcement applications, which include forensic identification, access control, border surveillance and human interactions and availability of low cost recording devices [5]. Face recognition has a high identification or recognition rate greater than 90% for huge face databases with well-controlled pose and illumination conditions [6].

This paper will give a literature survey of different methods for human facial expression recognition. Algorithms such as Viola and Jones face detection algorithm, support vector machine (SVM), Ad-boost, back propagation, linear discriminant analysis (LDA) algorithm, etc. were used by the researchers in the given methods which are introduced in this paper.

Databases such as Cohn Kanade (CK), Cohn Kanade Extended (CK+), Yale, Jaffe and Multimedia Understanding Group (MUG) are used by the methods which are introduced in this paper.

The remainder of this paper is organized as follows: Section 2 presents the basic expressions of human face while Section 2.1 presents the steps involved in facial expression recognition system. The different methods or techniques of facial expression recognition are given in Section 3. Result as comparison between methods introduced in this paper is presented in Section 4, and the conclusions are given in Section 5.

2. BASIC EMOTIONS

Human facial expressions are one of the most important non-verbal ways we communicate. Psychologist Paul Ekman, in 1971, who invented six types of emotions that are accepted by universal: anger, disgust, fear, happiness, sadness, and surprise. However, humans make use of a much fuller range of facial expressions for everyday communication than these six, some are even Combinations of these basic ones.

The emotion categories can be classified into two groups. Compound emotions and basic emotions. Where compound emotions are combination of two or more basic emotions. Ekman investigated them at the beginning of his studies and proposed an expanded list of basic emotions.

In 2014, Jack et al from the University of Glasgow reported new research, published in the journal Current Biology, intimating that there are only four basic emotions. While the result obtained that happiness and sadness is distinct where, fear and surprise share some common signals, like wide open eyes. Same as anger and disgust share the wrinkled nose.

1. Happiness is an emotion that is often associated with a state of mind that reflects contentment, satisfaction, pleasure or joy. Happiness is one of the most popular emotions.
2. Sadness is another type of emotion often defined as a transient emotional state characterized by feelings of disappointment, grief, hopelessness, disinterest, and dampened mood.
3. Anger can be a particularly powerful emotion characterized by feelings of hostility, agitation, frustration, and antagonism towards others.
4. Surprise is usually quite brief and is characterized by a physiological startle response following something unexpected.
5. Fear is a powerful emotion that can also play an important role in survival. When you face some sort of danger and experience fear, you go through what is known as the fight or flight response.
6. Disgust sense of revulsion can originate from a number of things, including an unpleasant taste, sight, or smell.

Airports, military and security firms are researching the use of cameras and software to pair what we know about facial expressions and deception to aid with investigative interviewing and primary detection of potential suspects in large, busy areas.

2.1 Basic steps to extract emotions:

Steps which are involved in facial expression recognition system are:

1. Image pre-processing

In image pre-processing the images are cleaned as the image dataset can be noisy.

The aim of pre-processing is to improve the image data that contains distortions and also enhances the image features which are important for further processing, although. To do so there are different techniques and methods like geometric transformations (e.g. scaling, rotation, translation) and many more are used. At last the image is converted to grayscale for next phase.

2. Face detection

Facial detection is the first part of the system in which the faces of humans are detected.

The aim of face detection is to detect and locate the face present in the image regardless of illuminations, background, Occlusions, Facial pose, orientation and expressions. Face detection algorithms such as principal component analysis (PCA), Latent Dirichlet allocation (LDA), Viola and Jones algorithm and many more are used.

3. Feature extraction

In this phase of system, features such as eyes, eyebrows, nose and mouth will get extract. It is very important to extract only those features which are having highly contribution in identification of facial expression.

4. Classification of facial expressions

Classification is the final stage where the facial expressions are classified on the basis of features which are extracted from previous phase. There are various methods to classify the expression which also contains various algorithms and also mentioned in this paper. Two methods which are used by almost every system are: 1. Geometric based method, 2. Appearance based method. After getting the features, classification methods are applied to recognize facial expression.

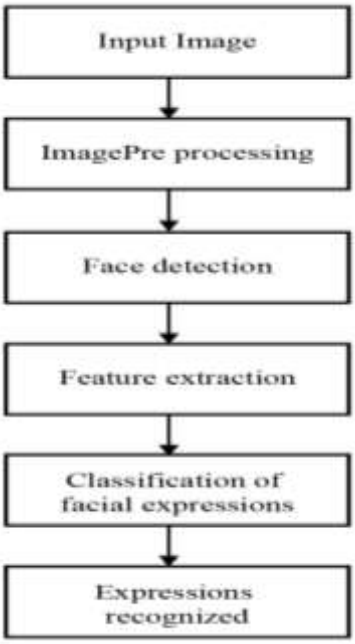

Figure 1 : Steps involved in facial expression recognition system

3. DIFFERENT TECHNIQUES OR METHODS FOR FACIAL EXPRESSIOIN RECOGNITION

3.1 Automatic Facial Expression Recognition Using Facial Animation Parameters and Multistream HMMs [7].

The method was proposed by Petar S. Aleksic and Aggelos K. Katsaggelos. In this method multistream Hidden Markov Model (MS-HMM) [8] is presented for automatic facial expression recognition system that utilizes MPEG-4 [9] compliant facial features. The Cohn-Kanade facial expression database [10] is used in this approach. Below is the block diagram of the proposed system

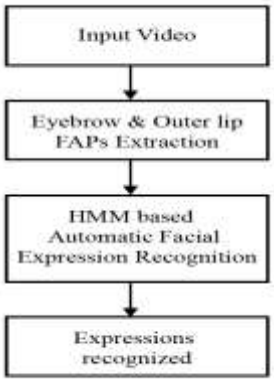

Figure 2 : Block diagram of Automatic Facial Expression Recognition Using Facial Animation Parameters and Multistream HMMs.

The outer-lip and eyebrow facial animation parameters (FAPs) are utilized as observations. The proposed approach introduces facial expression and FAP group dependent stream reliability weights. The stream weights are determined based on the automatic facial expression recognition results obtained when FAP streams are utilized individually. The proposed MS-HMM facial expression system, which utilizes stream reliability weights, achieves high recognition performance and significantly outperforms the single-stream HMM (SS-HMM) facial expression recognition system.

Outer-lip FAPs provided more information and produced better expression recognition results. Hence, the outer-lip stream weights were set in the experiments to be larger for the facial expressions for which eyebrow FAPs did not contain sufficient information. Stream weights were determined and utilizing the facial expression recognition results obtained when eyebrow and outer-lip FAPs were used individually. After the stream weights were chosen, MS-HMMs, described in the previous section, were utilized to perform facial expression recognition experiments and determine the effect of the introduction of stream reliability control to the recognition performance. The best recognition performance obtained in this method was 93.66%.

3.2 Real Time Facial Expression Recognition in Video using Support Vector Machines. [11]

This method was proposed by Philipp Michel and Rana El Kaliouby. They use a real time facial feature tracker to deal with the problems of face localization and feature extraction in spontaneous expressions. The tracker extracts the position of 22 facial features from the video stream. Displacements for each feature between a neutral and a representative frame of an expression was calculated and used together with the label of the facial expressions as input to the training stage of a support vector machine (SVM) classifier [12] [13]. The trained SVM model is subsequently used to classify unseen feature displacements in real time, either upon user request or continuously, for every frame in a video stream.

Below is the block diagram of the proposed method.

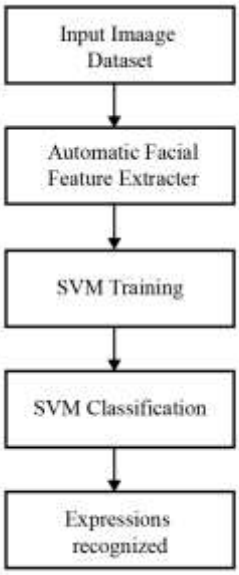

Figure 3 : Block diagram representing Facial expression recognition using SVM

This method is useful to detect the facial expression of a human in live video by using a commercial digital camcorder connected to a standard desktop computer. At lower frame rates and under different lighting conditions, the method performed well.

This method was applied first on datasets from CK facial expression database which results in high accuracy. After that they tested this method on a live video and got an accuracy rate of 87.5%

3.3 Neural-AdaBoost based facial expression recognition system. [14]

This method was based on AdaBoost (a boosting neural network) [15] and was proposed by Ebenezer Owusu, Yongzhao Zhan and Qi Rong Mao.

In this method, the data were reduced in dimensions by Bessel transform [16]and then the facial features are extracted by Gabor methods. This feature was further reduced via an AdaBoost-based feature reduction technique. The selected features represent samples of the facial deformation patterns of the expressive face. The datasets which were images from the Japanese Female Facial Expression (JAFFE) and Yale databases were partitioned into training and testing by leave-one-out cross validation. This selected features which represented the facial deformation patterns were then fed into a multilayer feed forward neural network (MFFNN) [17] which consists of 3 layers and is trained by a back-propagation algorithm.

The facial expression recognition was validated with the JAFFE and Yale facial expression databases.

The JAFFE database contains 213 images of 10 female Japanese persons. 140 images are taken for training and the rest was preserved for testing. The trial was performed using tenfold cross-validation to obtain the average recognition rate.

The Yale facial expression database contains 165 grayscale images in GIF format of 15 individuals. In this database 130 images were manually extracted. The datasets in this database were also partitioned into training and testing by using the same method described for the JAFFE.

The recorded average recognition rate was 96.83% in JAFFE and 92.22% in Yale.

Future improvements of recognition accuracies will depend on the increasing number of hidden neurons to expressions that recorded lower values.

3.4 Face Expression Recognition Using Histograms of Oriented Gradients with Reduced Features. [18]

This method utilizes histograms of oriented gradients (HOG) [19] descriptor to extract features from expressive facial images and was proposed by Nikunja Bihari Kar, Korra Sathya Babu and Sanjay Kumar Jena.

Below is the block diagram of the proposed method.

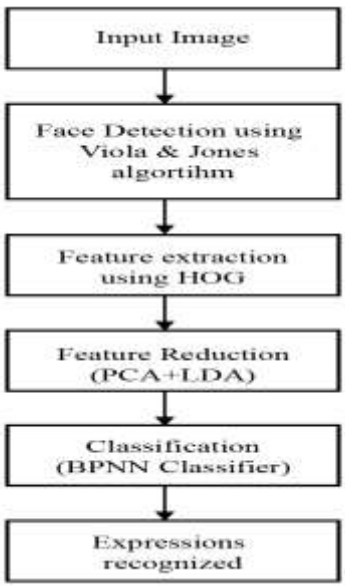

Figure 4 : Block diagram of Face Expression Recognition Using Histograms of Oriented Gradients with Reduced Features.

At first, the face images are converted to a grayscale image. This grayscale images are used by a popular Viola and Jones face detection algorithm [20] to detect the faces of humans. Each image is divided into a number of cells and a local 1-D histogram of gradient directions over the pixel are extracted for each cell. The output image is described by combining histograms of each cells. Local histograms are combined over a larger spatial region, called blocks. The normalized blocks are combined to represent HOG descriptor.

Now for feature reduction two linear methods are used: Linear Discriminant Analysis (LDA) and Principal Component Analysis (PCA). Classification of expressions is done by using a three-layer backpropagation neural network (BPNN) in which the input layer consists of 6 neurons as per the six features are selected after applying PCA+LDA standard. The learning performance was measured by root mean square error (RMSE).

Cohn-Kanade extended (CK+) was used to validate this method. The proposed method achieves recognition accuracy of 99.51 % with only six features.

3.5 Face expression recognition system based on ripplet transform type II and least square SVM. [21]

The method was proposed by Nikunja Bihari Kar, Korra Sathya Babu, Arun Kumar Sangaiah and Sambit Bakshi. It is based and efficient system consists of ripplet transform type II [22] and least square SVM (LS-SVM) [23].

Block diagram of this method is given below.

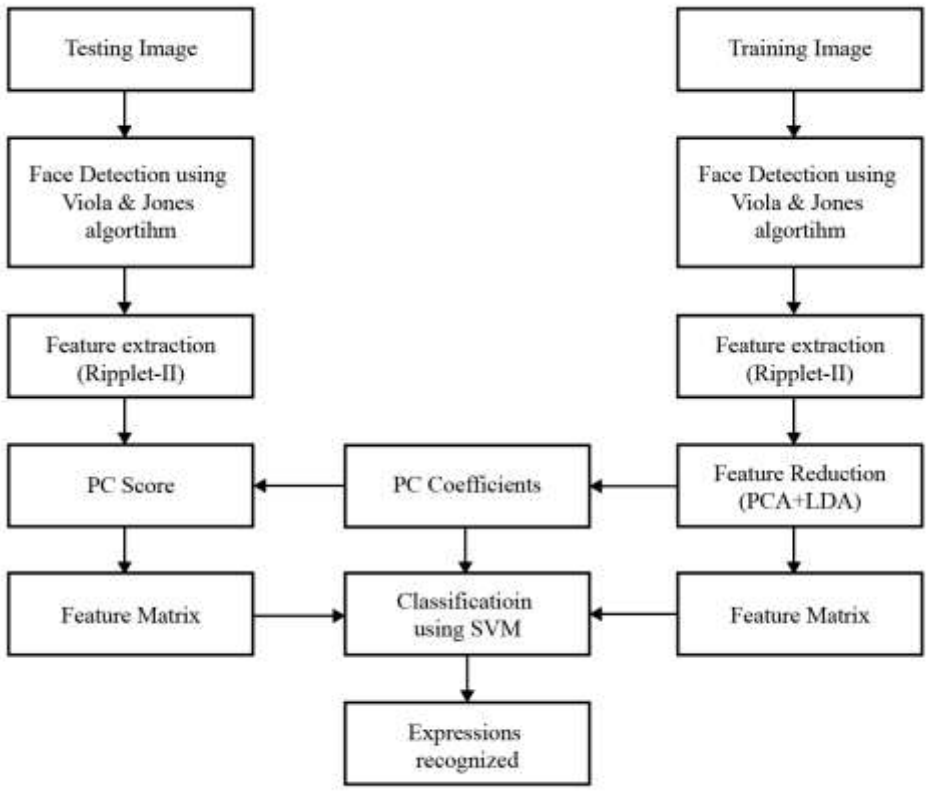

Figure 5 : Block diagram of Face expression recognition system based on ripplet transform type II and least square SVM.

At first the face is detected using Viola and Jones face detection algorithm from the input image. The region which was detected is then cropped and resized to maintain a uniform dimension of all input images. Now to extract features the method uses ripplet-II transform from the pre-processed image and carried out the discriminative features from the high dimensional ripple-II features by PCA+LDA approach.

Classification is performed using least-squares support-vector machine (LS-SVM). Datasets from JAFFE and CK+ databases were used to demonstrate the result as accuracy of detecting the facial expressions of humans. The proposed method achieves recognition accuracy of 98.97% and 99.46% on CK+ [24] and JAFFE [25] datasets respectively.

3.6 Automatic Facial Expression Recognition Using Combined Geometric Features. [26]

The present method proposed a geometrical feature based recognition system for automatic facial expression recognition and was proposed by Garima Sharma, Latika Singh and Sumanlata Gautam. The proposed methods use two datasets from JAFFE and Multimedia Understanding Group (MUG) [27].

At initial stage pre-processing of images is done to extract some useful details which are to be used for further processing. Removing noise from image, contrast enhancement and other geometric transformations are done first. After this, feature. extraction is done by landmark localization in this method which extracts the points from the facial regions like eyes, eyebrows, nose and mouth region. This method uses the iPar-CLR method to extract 49 fiducial points from the facial region. The spatial coordinates of the 49 landmark points extracted from the face image are utilized to calculate the relative distances using Euclidean distance measure among the facial points.

The final feature set spatial coordinates of landmarks points and distance features is constructed for images from JAFFE and MUG database. Three classification models (Multilayer perceptron neural network, SVM and Ensemble Neural Network) are trained using the feature sets.

Three experiments are performed on each classifier for both datasets. The first experiment is performed by training the classification models using only the facial landmark features, the second experiment is conducted by considering only the distance features and finally the third experiment uses the combined landmark and distance feature set for classification. The accuracy rate of this method for JAFFE database was 90.79% and for MUG database was 94.89.

3.7 Convolutional neural networks via a new face cropping and rotation strategy. [28]

The method was proposed by Kuan Li, Yi Jin, Muhammad Waqar Akram, Ruize Han and Jiongwei Chen which is based on convolution neural networks by cropping and rotating the input images. Below is the block diagram of the proposed method.

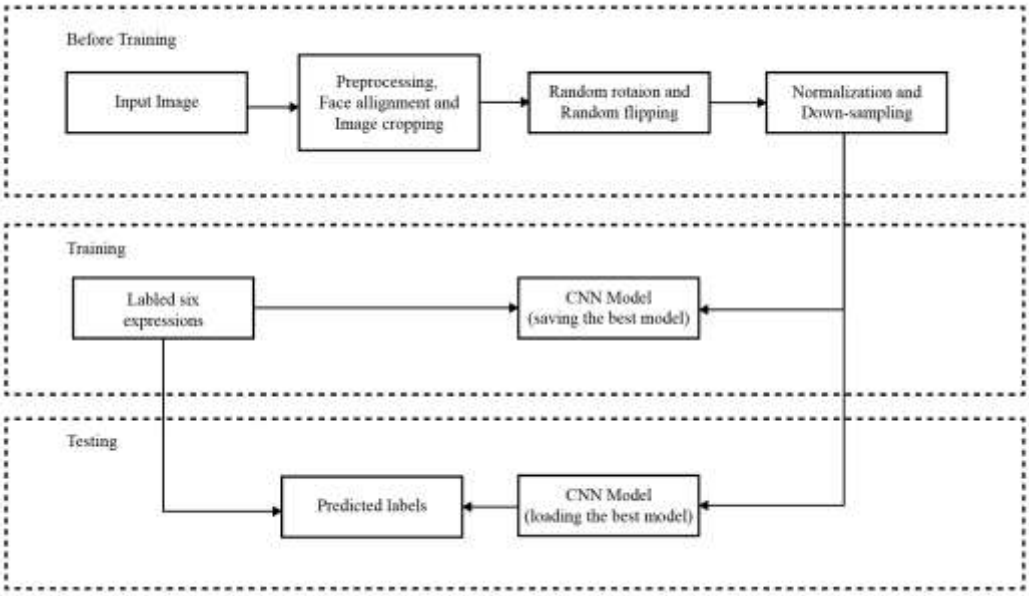

Figure 6 : Block diagram representing Convolutional neural networks via a new face cropping and rotation strategy.

At first stage the faces from images are aligned by rotation and then this rotated image is cropped in 3 ways (1. Face with background, 2. Face without background and 3. Face without forehead). In normalization phase, variations like brightness and contrast of the same face are reduced using histogram equalization.

After the pre-processing stages, to ensure the adaptability and abundance of data, random horizontal flip- ping and random rotation are adopted for training stage. In random horizontal flipping, an image can be flipped before training to address the problem of uneven cropping. Random rotation expands the original data by rotating an image by a random angle within an interval.

In present method convolution, neural networks (CNN) [29] was applied to extract features and categorize expressions. During the testing phase, the normalized testing images (without expansion) were sent to the CNN model from the training phase for prediction.

Here two widely used databases were used in the experiments: CK+ database and JAFFE database.

4. RESULTS

Human Facial Expression Recognition by using different methods is studied in this paper. The following Table I gives a short idea of comparison between these methods. The table also gives a brief idea of the databases used by each method. The conclusion of each technique is also provided in the table. Also, future scope if there are included in Table I.

Table 1 : A Survey on Human Facial Expression Recognition Techniques.

Sr.	Method/Technique(s)	Accuracy Rate	Conclusion	Future scope
1.	Automatic Facial Expression Recognition Using Facial Animation Parameters and Multistream HMMs.	Facial expression Recognition rate is 93.33% on Cohn-Kanade database.	It should result in more reliable training of the MS-HMMs and ultimately better recognition results then SS-HMMs.	Investigating possible applications of expression FAPs (high level FAPs) for facial expression analysis
2.	Real Time Facial Expression Recognition in Video using Support Vector Machines.	Facial expression Recognition rate is 86% on Cohn-Kanade database.	Information retrieved from more than one immediately succeeding frames of video sequences can increase the accuracy.	it can be made to work with a greater number of and more complex and train SVM for emotion categories.
3.	A neural-AdaBoost based facial expression recognition system.	Facial expression Recognition rate is 86% on JAFFE database.	Neural network is training with sample datasets from JAFFE and YALE facial expression databases and also automatic expression recognition are very accurate in surprise, disgust and happy; about 100%.	future improvements of recognition accuracies will look at the possibility of increasing the number of hidden neurons to expressions that recorded lower values.

4.	Face Expression Recognition Using Histograms of Oriented Gradients with Reduced Features.	Facial expression Recognition rate is 99.5% on Cohn-Kanade Extended database.	This scheme reaches recognition accuracy of 99.51 % with only six features.	In future, for augmented the performance of the system this scheme can be used.
5.	Face expression recognition system based on ripplet transform type II and least square SVM.	Facial expression Recognition rate is 98.97% on Cohn-Kanade Extended database and 99.46% on JAFFE database.	This method is tested on images captured in constrained environment with controlled lighting conditions, thus not hindering to the performance of the proposed system.	In future, other ground-breaking highlight extraction and portrayal devices, including profound learning, can likewise be used to improve the general execution of the FER framework.
6.	Automatic Facial Expression Recognition Using Combined Geometric Features.	Facial expression Recognition rate is 90.79% on JAFFE database and 94.89% on MUG database.	The results obtained by the proposed study are comparable with other methods and perform better in most of the cases.	The framework can be additionally improved by subsequently improves the general characterization precision of the incorporating appearance-based highlights with the expert framework.
7.	Convolutional neural networks via a new face cropping and rotation strategy.	Facial expression Recognition rate is 97.38% on Cohn-Kanade Extended database and 97.18% on JAFFE database.	The outcomes show that the proposed FER approach accomplishes focused outcomes as far as preparing time, testing time, and acknowledgment exactness.	In future, to improve high recognition accuracy and using different methods

5. CONCLUSION

The objective of this paper is to give a survey and compare different methods for human facial expression recognition system. In order to do this, the paper has looked into the details of every method introduced in section 3 and had given a short description with results. The basic expressions which are considered in this paper for recognition are happy, anger, sadness, disgust, fear and surprise. All the methods are having very good recognition rates when applied to widely used databases. Among all methods used, Face Expression Recognition Using Histograms of Oriented Gradients with Reduced Features and Face expression recognition system based on ripplet transform type II and least square SVM methods can recognize facial expressions of human with great accuracy in the presence of noise in the image.

REFERENCES

[1] N. P. S. Singh, "Monitoring driver fatigue using facial analysis techniques," in Proceedings 199 IEEE/IEEJ/JSAI International Conference on Intelligent Transportation Systems (Cat. No.99TH8383), Tokyo, Japan, 2002.

[2] M. Zetlin, "AI Is Now Analyzing Candidates' Facial Expressions During Video Job Interviews," 28 02 2018. [Online]. Available: https://www.inc.com/minda-zetlin/ai-is-now-analyzing-candidates-facial-expressions-during-video-job-interviews.html.

[3] Z. &. Y. S. &. I. Z. Kasiran, "Facial Expression Recognition as an Implicit Customers' Feedback.," Advances in Human Computer Interaction, p. 12, 2008.

[4] A. H. B. B. Pravin Dhulekar, "Identification Of Facial Features From The Real Time Video," International Journal of Advance Foundation and Research in Computer (IJAFRC) , vol. 2, no. 1, p. 5, 2015.

[5] N. C. I. Shivakumar Baragi, "Face Recognition using Fast Fourier Transform," in Research Advances in the Integration of Big Data and Smart Computing , IGI Global,, 2015.

[6] A. R. a. S. P. Anil K. Jain, IEEE Transactions on Circuits and Systems for Video Technology, vol. 14, 2004.

[7] A. K. P.S. Aleksic, "Automatic facial expression recognition using facial animation parameters and multistream HMMs," IEEE Transactions on Information Forensics and Security , vol. 1, no. 1, p. 8, 2006.

[8] B. H. J. Lawrence R Rabiner, Fundamentals of speech recognition, Englewood Cliffs, N.J. : PTR Prentice Hal, 1993.

[9] Facial Animation Engine (FAE), Genova: Tecnologia Automazione Uomo S.c.r.l.

[10] J. C. a. Y.-L. T. Takeo Kanade, "Comprehensive Database for Facial Expression Analysis," in 4th EEE International Conference on Automatic Face and Gesture Recognition, 2000.

[11] P. M. a. R. E. Kaliouby, "Real Time Facial Expression Recognition in Video using Support Vector Machines," Proceedings of the International Conference on Multimodal Interfaces, 2003.

[12] I. G. a. V. V. B. Boser, "A training algorithm for optimal margin classifiers," Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory, p. 144–152, 1992.

[13] N. C. a. J. Shawe-Taylor, An Introduction to Support Vector Machines and other Kernel-based Learning Methods., Cambridge: Cambridge University Press, 2000.

[14] Y. Z. a. Q. R. M. Ebenezer Owusu, "A neural-AdaBoost based facial expression recognition system," Expert Systems with Applications , vol. 41, no. 7, pp. 3383-3390, 2014.

[15] Y. F. E. Schapire, "A desicion-theoretic generalization of on-line learning and an application to boosting," Computational Learning Theory , pp. 23-37, 1995.

[16] M. P. P. C. G. S. V. Ganga, "Bessel transform for image resizing," Signals and Image Processing, pp. 1-4, 2011.

[17] A. B. N. Z. A. K. J. &. M. A. Bouzalmat, "Face recognition using neural network based Fourier Gabor filters & random projection," International Journal of Computer Science and Security, p. 376–386, 2011.

[18] K. S. B. a. S. K. J. Nikunja Bihari Kar, "Face Expression Recognition Using Histograms of Oriented Gradients with Reduced Features," in Proceedings of International Conference on Computer Vision and Image Processing, 2016.

[19] T. Gritti, "Localfeaturesbasedfacialexpressionrecog- nition with face registration errors.," in 8th IEEE International Conference on Automatic Face & Gesture Recognition, 2008.

[20] P. V. a. M. Jones, "Rapid Object Detection Using a Boosted Cascade of Simple Features," in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2001.

[21] K. S. B. A. K. S. &. S. B. Nikunja Bihari Kar, "Face expression recognition system based on ripplet transform type II and least square SVM," Multimedia Tools and Applications , pp. 4-24, 2017.

[22] X. J. &. W. D, "Ripplet transform type II transform for feature extraction," IET Image Processing, vol. 6, no. 4, pp. 374 - 385, 2012.

[23] J. S. &. J. Vandewalle, "Least Squares Support Vector Machine Classifiers," Neural Processing Letters , p. 293–300, 1999.

[24] C. J. K. T. S. J. A. Z. &. M. I. Lucey P, "The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression," in IEEE Computer society conference on computer vision and pattern recognition workshops (CVPRW), San Francisco, CA, USA, 2010.

[25] A. S. K. M. a. G. J. Lyons M, "Coding facial expressions with Gabor wavelets," in Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition, Nara, Japan, 1998.

[26] L. S. &. S. G. Garima Sharma, "Automatic Facial Expression Recognition Using Combined Geometric Features," 3D Research volume, 2019.

[27] C. P. a. A. D. Niki Aifanti, "The MUG facial expression database," in 11th International Workshop on Image Analysis for Multimedia Interactive Services WIAMIS 10 , Desenzano del Garda, Italy, 2010.

[28] Y. J. M. W. A. R. H. a. J. C. Kuan Li, "Facial expression recognition with convolutional neural networks via a new face cropping and rotation strategy," The Visual Computer, pp. 1 - 14, 2019.

[29] L. B. Y. B. a. P. H. Y. Lecun, "Gradient-based learning applied to document recognition," Proceedings of the IEEE , vol. 86, no. 11, pp. 2278 - 2324, 1998.