ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 4, April 2018

A Literature Survey on Emotion Recognition System Using Facial Expressions

[1] Rachoori Keerthi, [2] A. Obulesh, [3] Pallam Ravi,[4]Deepika.S
[1] M.Tech. Student, [2][3][4] Associate Professor
[1][2][3][4] Dept. of CSE, Anurag Group of Institutions, Hyderabad, T.S., India.

Abstract: Facial Expression Recognition(FER) is progressively picking up significance in different Affective computing applications. FER contributes to various application domains like computer vision, image processing, human computer interaction, information security, affective computing. This paper describes various emotion recognition techniques like LBP, and their performance is listed. Two common methods of Facial Expression Recognition System are appearance based and geometry based. This paper aims at describing a general procedure of how to recognize various facial expressions and making comparative study of various procedures. Also, describes various problems in facial expression recognition systems and methods implemented in resolving those issues.

Index Terms— Expression Recognition, Affective computing, Emotion Recognition, Facial Expression Recognition.

1. INTRODUCTION

Affective computing deals with machines and emotions. Emotion recognition[6][7] is very essential to develop effective Human Computer Interaction[8]. Human emotions are recognized by various non-verbal cues like facial expressions, gestures, body posture or speech. Among them facial expressions is easy to obtain. Facial expressions can be used to obtain 7 categories of expressions like neutral, happiness, surprise, disgust, fear, anger and surprise. The general procedure of determining facial expressions has three important steps (fig.1) . An image is given as input. The first step is the detection of face in the image in which important features are extracted and then face is identified. The second step is to extract the expression features from the image. Then extracted features are given to the classifier to identify the expressions as output.

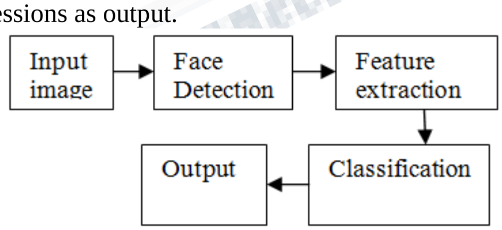

Fig.1. General procedure of Facial Expression Recognition.

2. EMOTION RECOGNITION SYSTEM

Emotion Recognition System[6][7] involves the process of acquiring the images, processing the images, detection of faces then extracting the expression features. The System consists of three main steps. First step is to identify the face region from the acquired image and then pre-processed so as to minimize the environmental and other variations in the image. The next step is to extract expression features which are then classified in the third step. The classifier provides the output of the expression which is recognized.

i) FACE DETECTION

Face detection process is to extract the face region from the background. Input images having various illumination conditions and complex backgrounds can be confusing to identify the expressions. It involves segmentation and extraction of facial features from uncontrolled background.

ii) PRE PROCESSING OF AN IMAGE

The image pre-processing step is important in facial expression recognition. Raw images may be corrupted with noise and unwanted effects. If the test image has different illumination conditions to that of the training set images the facial expression recognition may fail. To minimize the environmental and illumination changes in the image the image is pre-processed. The motive of image pre-processing step is to obtain images of normalized intensity so that illumination and environmental changes will have no effect.

iii) FEATURE EXTRACTION

The process of translating the features from the input image into some set of features is termed as feature extraction. Feature extraction helps in reducing the huge

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 4, April 2018

amount of data to small data which increases the computational efficiency. A set of points are selected which represent the characteristics of human face. The face parts like mouth corners, eyebrows are identified using features such as pixel intensities.
The two main categories of feature extraction methods used in Facial Expression Recognition system are: Geometric based features[5] and Appearance based features[5].

1) Geometric based : In this, geometric feature vectors are formed considering geometric relationships such as angles and positions between different facial parts like nose, eyes, ears. The facial expression is determined by the movement of the facial points. The effectiveness of this method is dependent on the accurate detection of facial components. The crucial task in this method is to determine and track the exact point in face region.

2)Appearance based : The features focus on facial appearances like bulges, wrinkles around mouth and eyes. This method applies image filters on the whole image to extract a feature vector. Appearance based algorithms are of wide range. These include Local binary Patterns(LBP)[4], Local directional Patterns(LDP)[5], Local Directional Ternary Patterns(LDTP)[9], Gabor wavelets[10].

iv) CLASSIFICATION
The features extracted from previous block tries to classify the features based on the similarities between the feature data. Classification is done by supervised learning. The classifier has to be trained first and then the input test data is given to recognize the expression of the image.

3. DATASETS

Datasets play a vital role in training, testing and validation of facial expression recognition systems. Expression databases are mainly of 2 kinds pose-based expressions and spontaneous-based expressions[5] [11] [12] . Pose-based [5] or deliberate expressions are those in which the actors are asked to display different emotional expressions while in spontaneous-based expressions, the expressions are natural. MMI dataset [11] [9] is a pose-based dataset, in which 205 expressions sequences are collected from 30 subjects. Each expression sequence has 6 basic expression classes (sad, neutral, fear, disgust, anger, happiness and surprise). CASME II dataset [11] contains 246 spontaneous micro-expression sequences by 26 subjects. The expression sequences have 5 expression classes (happiness, disgust, repression, surprise and anger). JAFFE dataset [9] [13] containing 213 pose-based facial expression images of basic 7 facial expressions. CK+ expression database [9] [13] of 123 subjects containing pose-based expressions and spontaneous-based smiles. Radboud Faces Database (RAFD) [12] contains 7 basic expressions (neutral, sad, disgust, fear, anger, surprise, happiness) is a multi-pose database. Karolinska Directed Emotional Faces Dataset (KDEF) [12] [14], a multi-posed dataset. Multi-PIE face dataset [12] contains 750,000 images under various illumination conditions. BU-3DFE [12] [9], a 3D facial Expression database. Static Facial Expressions in the Wild (SFEW) dataset [12] contains facial expression images with various poses, occlusions and illumination conditions. GEMEP-FERA dataset [9] contains 5 emotions (anger, joy, relief, fear, sadness).

4. FEATURE EXTRACTION

Facial expressions are instinctively dynamic in nature, while static FER systems are those which rely on static facial features. In order to capture the expression dynamics of facial expression sequences saptio-temporal features[11] are utilized. In [11] proposed a new spatio-temporal feature representation learning for FER system that sustains expression intensity variations. In this paper, they used representative expression states (e.g., onset, apex, offset expressions) that specifies the facial sequences regardless the expression intensity. As RGB image reveal person's identity leading to problems related to privacy, RGB images are replaced with depth images. Modified Local Directional Patterns (MLDP)[5] is used, which considers the top directional strengths with respect to the signs. Also, depth images solve the issue of pixel intensities.

In [14] proposed, special coefficients called as action units (AU) for seven facial expressions which were used as features. Softmax Regression deep sparse auto encoder network (SRDSAN) [13] is proposed to recognize facial expression. Region of interest (ROI) (mouth, eyes, eyebrows) are chosen as extracted areas of facial expression features. Auto encoder, which is able to reconstruct the data and change the representation of data which helps in increasing learning efficiency. Sparsity, is applied to auto encoder which helps in reducing the computational complexity. A pose-based Hierarchical bayesian theme model [12] is proposed to address challenges multi pose facial expression recognition (FER). Before recognizing an expression, an intermediate face

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 4, April 2018

representation is learned using local appearance features and geometric information.

Translation of facial image may deteriorate recognition performance and lack of robust classifiers are 2 limitations of emotion recognition systems. To resolve this, Stationary wavelet entropy [15] is used to extract features. Local directional ternary pattern (LDTP) [9] is used as face descriptor for facial expression recognition. LDTP encodes information of emotion-related features(i.e., eyes, eyebrows, upper-nose and mouth) using directional information and ternary patterns.

4.1 CLASSIFICATION

In [5] GDA (Generalized Discriminant Analysis)[5], supervised classification method is applied to obtained MLDP features and trained with DBN. Spatial [11] image characteristics of expression-state frames are learned via convolutional neural network and temporal features of spatial feature representation [11] obtained in the first part are learned via long short term memory(LSTM). Facial expression recognition is done either in view invariant manner or in multi-variant manner, a pose-LDA [12] classifier is used for classification. SVM[9] classifier with Radial Basis Function(RBF) [9] Kernel is used to classify expressions. Softmax Regression(SR) [13] is used to classify expression features obtained by Deep sparse autoencoder network (DSAN) [13]. k-NN classifier [14] and MLP neural network [14] are used to classify expressions. Feed forward neural network [15] is employed as the classifier. Jaya algorithm [15] is introduced to prevent the training of the classifier fall into optimum points.

4.2 PARAMETERS

Parameters η,α,β, θ and ξ are used in [12]. Dirichlet hyperparameters 'α' are estimated using maximum likelyhood estimation [16]. β, θ and ξ are estimated using Expectation Maximization (EM) [12]. Generalized linear model (GLM) parameters also used in [12]. Gradient Descent method is employed to obtain optimal sparse parameters [13].Common Controlling Parameters(CCP) are used in [15]. As, all algorithms have same Common Controlling Parameters(CCP) [15] with size of population 20. Parameters in [11]E1,E2,E3,E4,E5. Where 'E1' minimizing expression classification error, 'E2' minimizing intra-class variation, 'E3' minimizing expression-state classification error, 'E4' minimizing expression-state variation, 'E5' preserving expression-state continuity.

5. EXPERIMENTAL RESULTS AND DISCUSSIONS

S.No	Reference	Approach	Feature extraction	Classifier	Database	Performance
1	[5]	DBN	MLDP	GDA	RGB images Depth Images	93.33 96.25
2	[11]	CNN	spatio-temporal feature representation	LSTM	MMI CASME II	78.61 60.98
3	[12]	Hierarchical Bayesian Theme Model	SIFT	Pose-LDA	BU3DFE RAFD KDEF	79.11 74.91 74.91
4	[9]	Cross-Validation	LDP^4_{22}	SVM	JAFFEE MMI GEMEP-FERA BU-3DFE CMU-PIE	94.8 99.8 81.3 88.1 89.5
5	[13]	SRDSAN	DSAN	SR	JAFFE CK+	89.12 89.03
6	[14]	k-NN and MLP	Action Units	3-NN MLP	KDEF KDEF	90.00 96.00

7	[15]	Jaya Algorithm	Stationary Wavelet Entropy	Feed forward neural network	JAFFE	96.80

The analysis of various facial expression recognition systems is as follows. RGB images are replaced by Depth images in [5] to resolve privacy issue and noise due to illumination. As pixel intensities in depth images [5] remain unchanged even under noisy conditions. Though [5] resolves few issues, it is constrained only to pose-based expressions. To obtain, expression dynamics of facial expression sequences, saptio-temporal features [11] are utilized as static facial features could cause feature confusion between facial expression features and facial identity features. Inspite of resolving the issue of expression dynamics [11] has an average performance of about 78%. Multi-pose FER system is addressed in [12] as pose invariant FER systems are robust in nature. [12] has too many parameters and its performance is about 79%. [9] has optimum recognition rate in various kinds of databases ranging from microexpression dataset. Also [9] there are too many parameters hence increases the computational complexity. Sparsity [13] is applied to reduce the computational complexity, with recognition rate of about 89%.

6. CONCLUSION

The objective of this paper is to give a brief overview of Facial Expression Recognition system and to discuss various techniques adopted to implement FER system. A robust FER system, has to satisfy the performance in terms of accuracy, computational complexity, recognition rate. In addition, an FER system should satisfy pose-invariance, illumination variance.

REFERENCES

[1] L. Zhang and D. Tjondronegoro, "Facial Expression Recognition Using Facial Movement Features," IEEE Trans. Affect. Comput., vol. 2, no. 4, pp. 219–229, Oct. 2011.

[2] R. D. Ward and P. H. Marsden, "Affective computing: problems, reactions and intentions," Interact. Comput., vol. 16, no. 4, pp. 707–713, Aug. 2004.

[3] Institute of Electrical and Electronics Engineers., IEEE Computer Society., M. IEEE Systems, and IEEE Computational Intelligence Society., IEEE transactions on affective computing. Institute of Electrical and Electronics Engineers Computer Sociery, 2010.

[4] H. Zhang, Z. Qu, L. Yuan, and G. Li, "A face recognition method based on LBP feature for CNN," Proc. 2017 IEEE 2nd Adv. Inf. Technol. Electron. Autom. Control Conf. IAEAC 2017, pp. 544–547, 2017.

[5] M. Z. Uddin, M. M. Hassan, A. Almogren, M. Zuair, G. Fortino, and J. Torresen, "A facial expression recognition system using robust face features from depth videos and deep learning," Comput. Electr. Eng., vol. 63, pp. 114–125, 2017.

[6] J. Hakura, R. Domon, and H. Fujita, "Emotion recognition method using facial expressions and situation," in 2013 IEEE 12th International Conference on Intelligent Software Methodologies, Tools and Techniques (SoMeT), 2013, pp. 257–263.

[7] A. Kumar and A. Agarwal, "Emotion recognition using anatomical information in facial expressions," in 2014 9th International Conference on Industrial and Information Systems (ICIIS), 2014, pp. 1–6.

[8] P. SIGCHI (Group : U.S.), S. SIGART, J. G. IEEE Robotics and Automation Society, and Institute of Electrical and Electronics Engineers, HRI'16 : the Eleventh ACM/IEEE International Conference on Human Robot Interation : March 7-10, 2016, Christchurch, NZ. IEEE Press, 2016.

[9] B. Ryu, A. R. Rivera, J. Kim, and O. Chae, "Local Directional Ternary Pattern for Facial Expression Recognition," IEEE Trans. Image Process., vol. 26, no. 12, pp. 6006–6018, 2017.

[10] Borui Zhang, Guangyuan Liu, and Guoqiang Xie, "Facial expression recognition using LBP and LPQ based on Gabor wavelet transform," in 2016 2nd IEEE International Conference on Computer and Communications (ICCC), 2016, pp. 365–369.

[11] D. H. Kim, W. Baddar, J. Jang, and Y. M. Ro, "Multi-Objective based Spatio-Temporal Feature Representation Learning Robust to Expression Intensity

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 4, April 2018

Variations for Facial Expression Recognition," IEEE Trans. Affect. Comput., vol. 3045, no. c, pp. 1–1, 2017.

[12] Q. Mao, Q. Rao, Y. Yu, and M. Dong, "Hierarchical Bayesian Theme Models for Multipose Facial Expression Recognition," IEEE Trans. Multimed., vol. 19, no. 4, pp. 861–873, 2017.

[13] L. Chen, M. Zhou, W. Su, M. Wu, J. She, and K. Hirota, "Softmax regression based deep sparse autoencoder network for facial emotion recognition in human-robot interaction," Inf. Sci. (Ny)., vol. 428, pp. 49–61, 2018.

[14] P. Tarnowski, M. Kołodziej, A. Majkowski, and R. J. Rak, "Emotion recognition using facial expressions," Procedia Comput. Sci., vol. 108, pp. 1175–1184, 2017.

[15] S. H. Wang, P. Phillips, Z. C. Dong, and Y. D. Zhang, "Intelligent facial emotion recognition based on stationary wavelet entropy and Jaya algorithm," Neurocomputing, vol. 272, pp. 668–676, 2018.

[16] G. Ronning, "Maximum likelihood estimation of dirichlet distributions," J. Stat. Comput. Simul., vol. 32, no. 4, pp. 215–221, Jul. 1989.

