`A Review On Various Techniques To Recognize Gesture Based Facial Expressions

[1]Rohini Mahajan, [2]Devanand Padha
[1]Research Scholar, [2]Professor (HOD)
[1]Department of Computer Science, Central University of Jammu, Jammu
[2]Department Computer Science, Central University of Jammu, Jammu
[1]mahajanrohini10@gmail.com, [2]devanandpadha@gmail.com

ABSTRACT

Nowadays, the Facial Expression Recognition (FER) system is a very important research topic in the fields of pattern recognition, Human-Computer Interaction (HCI) and Artificial Intelligence. The most essential steps in FER system are facial feature extraction and classification of expressions which improves the system performance for human computer interaction. This article reviews the feature extraction approaches such as Local Binary Pattern (LBP) (Divide the examined window into cells (e.g. 16x16 pixels for each cell).For each pixel in a cell, compare the pixel to each of its 8 neighbors (on its left-top, left-middle, left-bottom, right-top, etc.). Follow the pixels along a circle, i.e. clockwise or counter-clockwise. Where the center pixel's value is greater than the neighbor's value, write "0". Otherwise, write "1". This gives an 8-digit binary number (which is usually converted to decimal for convenience).Compute the histogram, over the cell, of the frequency of each "number" occurring (i.e., each combination of which pixels are smaller and which are greater than the center). This histogram can be seen as a 256-dimensional feature vector. Optionally normalize the histogram. Concatenate (normalized) histograms of all cells. This gives a feature vector for the entire window.), Spatio-Temporal feature using Local Zernike moment and Local Directional Position Pattern (LDPP). Also, the various combinations of Convolutional Neural Network (CNN) and Deep Neural Network (DNN) classifiers are reviewed. In addition, this paper makes a comparative study of various techniques used for analysing gesture based facial expressions and their corresponding performance.
Keywords- Facial Expression Recognition (FER), Convolutional Neural Network (CNN), Deep Neural Network (DNN), Local binary pattern (LBP), feature extraction.

INTRODUCTION

In the recent years, various technologies referring to Human Computer Interaction (HCI) are enduring based on our essential requirements and we highly depend on these technologies. One such recent trend in technology is Facial Expression Recognition (FER) [1]. Face expressions play a vital role in social communication, and it does not only expose the feelings of human beings but also be used for recognizing their mentality. Thus, FER is a technique to recognize seven expressions of human beings such as anger, disgust, fear, happiness, neutral, sadness and surprise[2]. Recently, automatic FER has attracted growing attention in many fields such as Computer Vision, Pattern Recognition, and Artificial Intelligence.

FER aims to develop an accurate, efficient and automatic system to distinguish facial expression of human beings so that human emotions can be understood through their facial expression [4]. A fundamental automatic FER system comprise of three major steps namely pre-processing, facial feature extraction and facial expression classification[3]. At first, the pre-processing step detects the face and reduces light intensity effect. Next step is feature extraction that

extracts the required features and unnecessary features are removed. Final step is facial expression classification in which the expressions are classified in to six common emotions. Fig.1 shows that the basic process of facial expression recognition system.

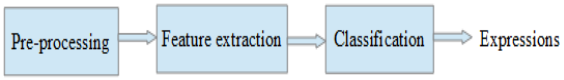

Fig.1 Fundamental process of FER system.

The present study conducts a review of various techniques used for facial feature extraction and facial expression classification in the FER system. The paper comprises of literature review which includes various FER techniques proposed by different authors, presents a comparative analysis of different feature extraction as well as classification methods of FER and finally concludes with the recommendations for future research.

LITERATURE REVIEW

The literature review is presented in two different categories namely facial feature extraction and expression classification.

A. Facial Feature extraction

Facial Expression Recognition (FER) system deals with extracting discriminative features from the human face. In FER techniques, texture features are extracted using Local Binary Pattern (LBP). This process creates the correlated features between different expressions and hence affects the accuracy. To address this issue, Kumar et al., (2016) proposed an informative region extraction model in which the facial regions based on the projection of the expressive face images is compared to common reference neutral image using Procrustes analysis. The weighted-projection based LBP feature is extracted from the informative regions of the face. This feature extraction technique minimizes the miss-classification between various classes of expressions.[9]

Uddin et al., (2017) proposed a novel feature extraction algorithm to extract salient features from the face. A new robust feature extraction method is employed namely Local Directional Position Pattern (LDPP). In LDPP, extracting local directional strengths for every pixel included in Local Directional Pattern (LDP) where in top directional strength positions are considered in binary along with their strength sign bits. Besides, the LDPP features are computed by Principal Component Analysis (PCA) and Generalized Discriminant Analysis (GDA) to provide more robust feature. These features are employed with Deep Belief Network (DBN) for expression training and recognition. [10]

Kamarol et al., (2016) proposed an appearance-based facial feature extraction to perform the facial expression recognition. The Spatio-Temporal Texture Map (STTM) has the ability of achieving the subtle spatial and temporal variations of facial expressions with less computational complexity. In this, the face is detected based on Viola–Jones Face Detector and irrelevant portions are removed. The feature extraction is achieved using STTM that uses the spatio-temporal information extracted from the three-dimensional Harris corner function. A block-based technique is adjusted to extract the dynamic features and these features are then classified into various facial expressions by the Support Vector Machine (SVM) classifier. The simulation results show that the STTM method achieves the highest performance rate compared with other feature extraction approaches. [11]

Fan & Tjahjadi, (2017) presented a Spatio-Temporal Feature using Local Zernike moment in the spatial domain. The dynamic feature consists of motion history image and entropy. A weighting scheme based on the feature and sub-division of the image frame is used to improve the dynamic information of facial expression. These extracted feature are used for expression classification using SVM. [13]

Yan, (2017) presented a Collaborative Discriminative Multi-Metric Learning (CDMML) based feature extraction for facial expression recognition. It possibly develops complementary and discriminative information for recognition in two parts: a) For every face frame, the feature descriptors are estimated to define both, the appearance in the facial image and the motion information of the subject in frame. b) Distance parameters with these extracted multiple features are approximated to analyze the most prominent expression. [14]

B. Expression classification

Zhang et al., (2017) proposed a Part-based Hierarchical Bidirectional Recurrent Neural Network (PHRNN) for analyzing the facial expression information. The temporal features are extracted based on facial morphological changes and dynamical evolution of expressions using facial landmarks. In addition, a Multi-Signal Convolutional Neural Network (MSCNN) is used to extract spatial features from face image. Both recognition and verification signals are generated to estimate the various loss functions thereby maximizing the variations of different expressions and minimizing the differences between identical expressions. The Deep Evolutional Spatial-Temporal Network is used to extract the geometry-appearance and dynamic-still information boosting the performance of facial expression recognition. [8]

A novel Deep Neural Network (DNN) based feature learning algorithm for FER has been proposed by Zhang et al., (2016). Scale Invariant Feature Transform (SIFT) is used to extract a set of landmark points from each facial image. The DNN algorithm take these extracted SIFT feature vectors as input, applies various hidden layers to define the relationship between the SIFT features and their corresponding high-level semantic information and finally obtain a set of appropriate optimal features by training the model. This model classifies the various facial expressions through the simulation process. The results of the process illustrate that the DNN model outperforms the other traditional algorithms. [7]

Xie & Hu, (2017) proposed a new classification algorithm namely Feature Redundancy Reduced with Convolutional Neural Network (FRR-CNN) for facial expression recognition. The convolutional kernel of FRR-CNN offer a more discriminative mutual difference between feature maps of the same layer. This results in producing less redundant features and earns a more compact illustration of an image. Besides, the transformation-invariant combing scheme is used to extract representative features cross-transformations. The simulation results show the FRR-CNN algorithm has more efficiency compared to other conventional classification algorithms. [6]

Du & Hu, (2017) proposed a Modified Classification and Regression Tree (M-CRT) algorithm based FER system. The facial expression details are attained by estimating the difference between the neutral expression and other basic expressions of image that illustrate the information irrelevant to identities. The global and local features are obtained using Local Binary Patterns (LBP) and Supervised Descent Method. M-CRT algorithm is used to detect the best classification decision based on the extracted global and local features. The simulation results demonstrates that M-CRT algorithm is simultaneously maximize the intra-class purity and the distance between classes that improves the classification accuracy. [5]

Hsieh et al., (2016) proposed a semantic features based facial expression recognition using Support Vector Machine (SVM). It locates the facial components by active shape model to extract seven dynamic face regions. Then, this semantic facial features is obtained using directional gradient operators such as Gabor filters and Laplacian of Gaussian. A multi-class support vector machine (SVM) is trained to classify six facial expressions. The experimental showed that the semantic facial features is effectively represent the changes between facial expressions. Also, the time complexity is lower than the other SVM based methods due to the less number of deployed features. [12]

COMPARISON OF RECENT TECHNIQUES

The comparative analysis of various classification techniques in Facial Expression Recognition (FER) is shown in Table1.

S. No	Paper Title	Technique	Database	Outcome
1	Facial expression recognition based on deep evolutional spatial-temporal networks (Zhang et al., 2017)	Part-based Hierarchical Bidirectional Recurrent Neural Network (PHRNN) and Multi-Signal Convolutional Neural Network (MSCNN)	CK+, Oulu-CASIA and MMI databases	Achieved accuracy of 71.18 %.
2	A Deep Neural Network-Driven Feature Learning Method for Multi-view Facial Expression Recognition (Zhang et al., 2016)	Deep Neural Network (DNN)	BU-3DFE and Multi-PIE databases	Achieved 70.1% of recognition accuracy.
3	Facial expression recognition with FRR-CNN (Xie & Hu, 2017)	Feature Redundancy Reduced with Convolutional Neural Network (FRR-CNN)	two public facial expression databases	Achieved 73.96% of recognition accuracy.
4	Modified classification and regression tree for facial expression recognition with using difference expression images (Du & Hu, 2017)	Modified Classification and Regression Tree (M-CRT)	Japanese Female Facial Expression and CK+ database	obtained 79.45% of recognition accuracy.
5	Effective semantic features for facial expressions recognition using svm (Hsieh et al., 2016)	Support Vector Machine (SVM)	Cohn–Kanade database	Attained 84.7 average recognition rate.

Table1: Comparative analysis of various glitch reduction techniques.

CONCLUSION

In this paper, various facial feature extraction and facial expression classification techniques for Facial Expression Recognition (FER) have been discussed. In classification, Convolutional Neural Network (CNN) and Deep Neural Network (DNN) have achieved the better performance than the Artificial Neural Networks (ANN) and support vector machine (SVM). Furthermore, a comparative analysis of various classification techniques is studied based on their performance. The design of new facial expression recognition (FER) system with high discriminative feature extraction algorithm and high classification accuracy rate remains a challenge.

REFERENCES

[1] Hsu, S. C., Huang, H. H., & Huang, C. L. (2017, April). "Facial Expression Recognition for Human-Robot Interaction." In Robotic Computing (IRC), IEEE International Conference on (pp. 1-7). IEEE.

[2] Căleanu, C. D. (2013, May). "Face expression recognition: A brief overview of the last decade." In Applied Computational Intelligence and Informatics (SACI), 2013 IEEE 8th International Symposium on (pp. 157-161). IEEE.

[3] Zhao, X., & Zhang, S. (2016). A Review on Facial Expression Recognition: Feature Extraction and Classification. IETE Technical Review, 33(5), 505-517.

[4] Bernin, A., Müller, L., Ghose, S., von Luck, K., Grecos, C., Wang, Q., & Vogt, F. (2017, June). "Towards More Robust Automatic Facial Expression Recognition in Smart Environments." In Proceedings of the 10th International Conference on PErvasive Technologies Related to Assistive Environments (pp. 37-44). ACM.

[5] Du, L., & Hu, H. (2017). "Modified classification and regression tree for facial expression recognition with using difference expression images." Electronics Letters, 53(9), 590-592.

[6] Xie, S., & Hu, H. (2017). "Facial expression recognition with FRR-CNN". Electronics Letters, 53(4), 235-237.

[7] Zhang, T., Zheng, W., Cui, Z., Zong, Y., Yan, J., & Yan, K. (2016). "A Deep Neural Network-Driven Feature Learning Method for Multi-view Facial Expression Recognition." IEEE Transactions on Multimedia, 18(12), 2528-2536.

[8] Zhang, K., Huang, Y., Du, Y., & Wang, L. (2017). "Facial expression recognition based on deep evolutional spatial-temporal networks". IEEE Transactions on Image Processing, 26(9), 4193-4203.

[9] Kumar, S., Bhuyan, M. K., & Chakraborty, B. K. (2016). "Extraction of informative regions of a face for facial expression recognition". IET Computer Vision, 10(6), 567-576.

[10] Uddin, M. Z., Hassan, M. M., Almogren, A., Alamri, A., Alrubaian, M., & Fortino, G. (2017). "Facial Expression Recognition Utilizing Local Direction-Based Robust Features and Deep Belief Network." IEEE Access, 5, 4525-4536.

[11] Kamarol, S. K. A., Jaward, M. H., Parkkinen, J., & Parthiban, R. (2016). "Spatiotemporal feature extraction for facial expression recognition. IET Image Processing", 10(7), 534-541.

[12] Hsieh, C. C., Hsih, M. H., Jiang, M. K., Cheng, Y. M., & Liang, E. H. (2016). "Effective semantic features for facial expressions recognition using svm. Multimedia Tools and Applications", 75(11), 6663-6682.

[13] Fan, X., & Tjahjadi, T. (2017). "A dynamic framework based on local Zernike moment and motion history image for facial expression recognition." Pattern Recognition, 64, 399-406.

[14] Yan, H. (2017). "Collaborative discriminative multi-metric learning for facial expression recognition in video". Pattern Recognition.