

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 3, March 2019

Facial Expressions Recognition Based on Cognition and Mapped Binary Patterns

Shweta K Gaikwad¹, Pravin P Kalyankar²

M. E Scholar, Dept. of CSE, TPCT's COE Osmanabad, Maharashtra, India¹

Asst. Professor, Dept. of CSE, TPCT's COE Osmanabad, Maharashtra, India²

ABSTRACT: In this paper, a new expression recognition approach is presented based on cognition and mapped binary patterns. At first, the approach is based on the LBP operator to extract the facial contours. Secondly, the establishment of pseudo-3-D model is used to segment the 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 circumplex emotion model. At last, we perform a comparative experiment on the expansion of the Cohn-Kanade(CK+) facial expression dataset and the test data sets collected from ten volunteers. The experimental results show that the method can effectively remove the confounding factors in the image. And the result of using the circumplex 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.

KEYWORDS: Facial expressions recognition, local binary patterns, human cognition, emotion model, machine learning.

I. INTRODUCTION

In human society, facial expressions play a significant role. As the carrier of information, facial expressions are able to contain numerous information that can not be conveyed by voice in the process of interpersonal communication [1]. Therefore, automatic facial expressions recognition has been becoming a growing topic in computer vision research in recent years. The automatic classification of facial expressions is a great significant for human-computer interaction technology, and also has potential application value in many fields such as computer-aided training, medical field and distance education. The traditional electronic intelligent teaching has the drawback that the interaction between students and teachers about teaching feedback can only be delivered via voice information. Emotional calculation based on facial expressions can solve the shortcomings of the traditional electronic teaching. In the medical field, the automatic recognition of facial expressions can help the doctor know the patient's feelings in time. In addition, the application of facial expressions recognition technology has also offered new developments in the fields of animation film and face image synthesis. Facial expressions recognition (FER) systems consist of face detection, facial feature extraction and expression classification. American psychologists Ekman and Friesen defined six basic human facial expressions, like happy, angry, surprise, disgust, etc. They also developed facial action coding system (FACS) based on action units (AUs) to describe facial expressions [2]. In recent years, many methods of automatic facial expression recognition have been proposed, such as principal component analysis (PCA) [4], linear criterion analysis (LDA) [5], [6], elastic bunch graph matching (EBGM) [7], independent component analysis (ICA) [8], two-dimensional principal component analysis (2D-PCA) [9], artificial neural networks, embedded hidden Markov models (EHMM) [10] and others. Although these algorithms improved the accuracy and speed of automatic facial expressions recognition, they are also affected by changes in light, pose, aging, alignment, etc [12]. In the process of facial expression recognition, the facial expression features extracted from the facial images are the highly abstract expression of the facial expressions. There are two main approaches: the overall template matching system and the system based on geometric features. In the overall system, the template can be a

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 3, March 2019

pixel or a vector. In the geometric feature system, the principal component analysis and the multi-layer neural network are widely used to obtain the low-dimensional representation of the face, and to detect the main feature points and the main parts in the picture. On the one hand, feature-based approaches are computationally larger than template-based methods, but are not sensitive to scale, size, head orientation and face position [11]. On the other hand, a large number of physiologic and psychological studies have shown that the human visual system in the Pre attentive stage detects and deals with the pop-out feature areas of the scene. These feature areas in the scene are different from most of the background of the regions and make the visual cells more stimulated. However, the existing methods are mostly feature extraction based on locating facial feature points. In order to verify and compare with the research results of relevant human cognitive disciplines, we used the method of feature areas.

II.RELATED WORK

With the development of database acquisition equipment and acquisition technology, the collected image pixels are getting higher and higher, and at the same time, the information that the pictures can express is also getting more and more delicate. Human facial expression pictures can explain and express a more delicate state of mind. And human facial expression images have been able to detect human emotions by analyzing the sequence of emotions images or videos. Since the amount of information contained in an image or video is too large, the general method is to extract the most important information from the static facial expressions images or dynamic facial expressions videos, and to reduce the number of image dimension without losing the most essential information. That means the features of the images are extracted. In recent years, by integrating emotional computing, computer vision and machine learning contents, the automatic expressions recognition has become a popular direction in the field of artificial intelligence. However, most of the previous works were based on experiments in laboratory controlled conditions, which ignored the individual differences of the experimental object and the impact on the experimental results. Generally speaking, automatic expressions recognition can be divided into two categories: one category is based on the traditional feature classification method that includes selecting different features and matching different classification algorithms. The other category is based on the deep learning methods that focus on how to construct and train the different neural network models. For the former, most of methods can be divided into three types: Principal Component Analysis(PCA),Local Binary Pattern (LBP) and Gabor wavelet. After decades of development, automatic recognition of facial expressions has been summarized as a complete system. In facial expression recognition, researchers hope to extract the facial feature information as complete as possible, so the original image needs to be pre-processed before the feature extraction, and the image should be normalized by pre-processing and remove the complex background. Object detection technology can solve well the problem of removing the complex background of the original image, and it usually depends on category-based training to handle intra-class variations. When the positive sample of the target detection technology or the target sample to be detected is a human face, it can well solve the problem of removing the complicated background of the image. Our proposed a classification method based on structured sparse representation for robust face recognition. This method also works better than the sparse representation face recognition method when the face is occluded or the illumination changes. Due to the different ways of image capture, the format and size of the image will be different. In order to facilitate subsequent feature extraction and image classification, the original image usually needs to be normalized. Normalization includes gray normalization and geometric normalization or position calibration. The former mainly solves the problem of large differences in gray caused by uneven lighting. The latter is mainly used to solve the face scale changes and face rotation problems caused by imaging distance or face pose that changes the size difference and tilt angle.

Face feature extraction is a very important part in automatic facial expression recognition. Although the network structure proposed in some deep learning models can self-learn features, the initialization and adjust men to parameters are too dependent on the trick strategy heavily. Many excellentfeatureextractionmethodscanremovefacialexpressionindependent interference information such as light, background and noise while the most essential features of facial expressions are preserved. It can reduce the image dimension and select the appropriate data structure, so as to achieve the purpose for correct classification. Feature extraction can be divided into geometric feature-based extraction methods and appearance-based methods. Among them, the geometric feature extraction method needs to be combined with the facial feature points detection technique.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 3, March 2019

Feature points detection technology can accurately mark the face of each organ feature points, but this approach likely ignore the changes in facial texture. The appearance-based methods are able to extract pixel-level underlying information in face images through specific operations and can be sensitive to local texture changes. The LBP operator calculates the gray scale order relationship between each pixel in the image and its local neighborhood pixels, and then encodes the binary sequence to form a local binary pattern. The LBP texture feature map can be shown in Figure.1. Finally, the multi-region histogram which worked out various non overlapping blocks is used as the feature description of the image. Due to LBP has many significant advantages, such as invariance on gray-level and rotation, it is used widely in research work [13]. Although LBP has made good progress in the texture feature extraction of images, the interference of noise also hinders the accuracy of image classification. Making texture analysis is not only robust against noise disturbances but also the direction of many methodological studies. However, many existing methods used the LBP texture feature, overlooked the complexity of its inherent law and did not reflect the association between the local facial features and the different facial expressions.

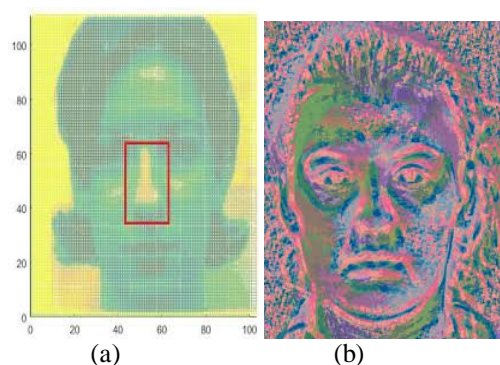


FIGURE 1. (a) A facial image and (b) its LBP texture feature map. Owing to the LBP features are robust to illumination changes, it is widely used in face recognition and target detection.

III.METHOD

In this section, the various parts of the work will be introduced. To begin with, we extract the facial contours by seeking the largest island method based on local binary pattern (LBP) and segment the facial regions by establishing pseudo 3D model which makes the gray values of the images as its z-axis. Moreover, we use Multi Dimensional Scaling (MDS) and the approximation of Earth Mover's Distance (EMD) to reduce the dimension of data. Once more, we use the feature images reduced dimension to train the convolutional neural network model to predict the expressions, and also choose the different feature regions or combinations as the reference to find out the best distinguished region or combination and compare with the conclusions in cognitive neurology. Finally, we have a comparative experiment between traditional basic emotion model and dimension space model.

IV.EXPERIMENTS AND RESULTS

Facial expression recognition can be divided into two steps: feature extraction and facial expression classification. Among them, facial expression feature extraction can be divided into feature training and feature testing. In this experiment, we take 150 images from each facial expression and use Extended Cohn-Kanade Dataset [3] as the experimental dataset. Firstly, we divide this dataset into two parts with 50 images as the test set, and the other as the training set. In the facial expressions section, we use the mapped LBP method, on the other hand, to reduce the dimensionality of the image features and then use the convolution neural network for facial expression classification. And on the other hand, we use two classifiers: the support vector machine (SVM) and softmax for classification and recognition. In the experiment, the image should be pre-processed firstly, and the Cohn-Kanade expression database is

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 3, March 2019

composed of a series of action images from the beginning of the expression to the expression of the mature. In order to get more accurate results, only the facial images that are forward and have no occlusion or large sloshing are chooesd. The images we collected from ten volunteers should be pre-processed, too. The Figure2 is a pre-processed facial expression image whose unrelated information, such as hair, ear, background, etc., has been removed. The purpose is to make the follow-up work more pertinent. Only the features of the face retained, and the unnecessary interferences are excluded.



FIGURE 2. The facial expressions images were collected from the volunteer and were preliminarily pre-processed.

It summarizes the results base done using two classifiers SVM and softmax on the test set, specifically for each class(SeeFigure3). In Figure3, we can see that the performance of the softmax is better than SVM.

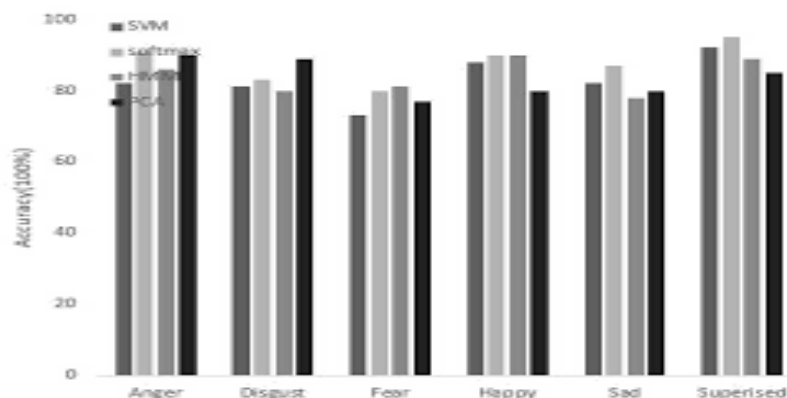


FIGURE 3. Comparison of SVM, softmax, HMM and PCA method.

Moreover, we compare our method with existing HMM method and PCA method. In the experiment, we average the results by using three LBP representations because the different LBP representation hasn't much difference and influence in recognition accuracy. The experimental results show that this method can measure and express the distance between LBP codes more accurately in binary space. And this method is able to, make the features more salient and the data dimension smaller. And this method is capable of obviating the double interference of the light and noise. Moreover, this method is in a position to make the calculation more convenient to. According to the work of Shokoohietal. if the movements of the eyes and the mouth area can be accurately identified, the recognition rate of the expression will be improved. So we divide the face into six regions by the way of the method of division mentioned in the third section. Compared with single block method and multi layer block method, our approach is more reasonable

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 3, March 2019

and avoids the irrationality of dividing a whole organ into small pieces in a fixed-size block. Our approach also solves the problem of small size block method which will ignore the information of the local expression. Then, in order to discuss the correspondence between the local region and the facial features, we select different expression are as for the experiment. Under the condition of only retaining eyes and mouth regions and covering other regions, the specific situations of identification are shown in the following Table 1.

TABLE 1. Confusion matrix and automatic facial expression recognition performance the mouth and eyes regions.

Anger	39	4	0	4	3	0
Disgust	4	41	0	2	3	0
Fear	1	8	33	3	5	0
Happy	6	0	3	37	4	0
Sad	9	6	4	0	28	0
Surprise	0	0	8	5	4	33

The results show that the eyes and mouth are better able to distinguish the expression than other regions, especially on anger and disgust.

TABLE 2. Confusion matrix and automatic facial expression recognition performance the mouth, eyes and left-cheek regions

Anger	40	4	0	4	3	0
Disgust	3	42	0	2	3	0
Fear	0	9	33	4	4	0
Happy	6	0	1	40	3	0

As shown in Table 2, it describes the number of classifications of each facial expression in detail after the left cheek region is added.

TABLE 3. Confusion matrix and automatic facial expression recognition performance the mouth, eyes and forehead regions.

Anger	39	4	0	4	3	0
Disgust	4	41	0	1	3	0
Fear	0	9	33	3	5	0
Happy	5	0	2	38	5	0
Sad	8	7	4	0	30	1
Surprise	0	0	7	6	3	34

Similarly, Table 3 show that the forehead region is different from cheeks region on expressing emotions. This result also supports the conclusion that human beings in the cognitive neurology mainly through the eyes and mouth to recognize expressions. Unlike the general discrete model, the circumplex model of expressions has two dimensions: pleasure and activeness, so it can express more expressive content.

Finally, we did a comparative test based on the traditionally basic emotional model and the circumplex emotion model in the experimental data collected from the volunteers. We use P1,P2,P3,P4 and P5 to represent the five ranges of pleasure, P1 means very unpleasant and P5 means very pleasant. Similarly, I1,I2,I3,I4 and I5 are used to represent the five intervals of activity, I1 is very inactive and I5 is very active. According to the work of Shietal.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 3, March 2019

Pleasure and intensity have a clear positive correlation. We apply the circumplex emotion model in facial expression recognition. On the basis of the traditional basic emotional model of the classification, the process and the results according to the traditional basic emotional model for the classification in the previous part has been discussed clearly. This section introduce the experimental process and results in the light of the circumplex emotion model and compares it with the basic emotional model. We find that the circumplex emotion model performed well in dealing with some expressions that can't be described and expressed. This part of the expressions are not in the seven discrete expressions, and the volunteers said that these are the complicated results composed of various emotions. Because the amount of the expressional data we collected is small, there is little difference in performance between the SVM and the softmax. The accuracy of the experiment is shown in Figure 4.

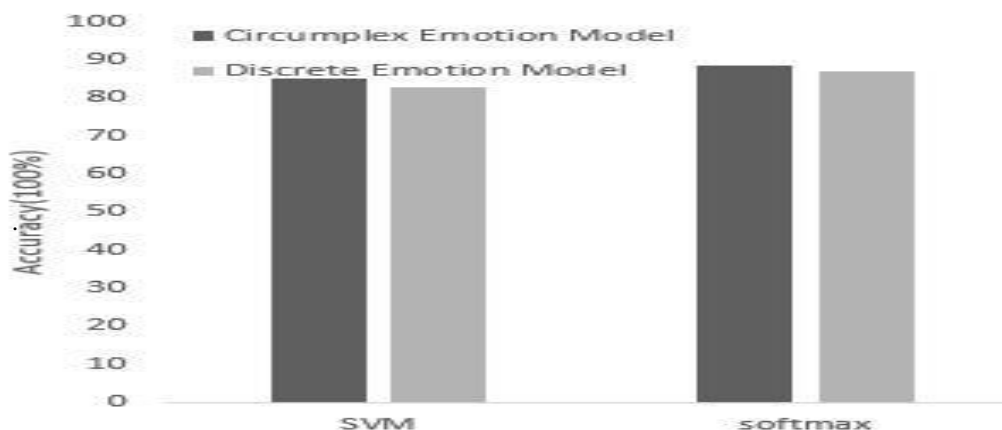


FIGURE 4. Comparison of two emotion models: the traditionally basic emotional model and the circumplex emotion model.

V. CONCLUSION

In this paper, we propose a model mainly improve the current work and find that the local facial features and expressions are correlated. We divide the facial area on the basis of the cognitive neurological theory to recognize and compare with the theory. It is conducive to the future expressions recognition with the existence of occlusion do their complex conditions. We use LBP codes to describe the images and take into account the distances between the points of the image distribution in binary space that can't be measured directly in European distance. Finally, the basic emotion classification model and the the circumplex emotion model based on dimension space theory are discussed and contrasted. And the recognition accuracy of the circumplex emotion model is higher, which will help us in the next step on the complex expression and micro-expression recognition progress.

REFERENCES

- [1] A. Mehrabian and J. A. Russell, An Approach to Environmental Psychology. Cambridge, MA, USA: MIT Press, 1974.
- [2] P. Ekman and W. V. Friesen, Facial Action Coding System: A Technique for the Measurement of Facial Movement. Palo Alto, CA, USA: Consulting Psychologists Press, 1978, pp. 271–302.
- [3] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, "The extended Cohn-Kanade dataset (CK+): A completed dataset for action unit and emotion-specified expression," in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops, Jun. 2010, pp. 94–101.
- [4] M. Turk and A. Pentland, "Eigenfaces for recognition," J. Cognit. Neurosci., vol. 3, no. 1, pp. 71–86, 1991.
- [5] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection," IEEE Trans. Pattern Anal. Mach. Intell., vol. 19, no. 7, pp. 711–720, Jul. 1997.
- [6] K. Etemad and R. Chellappa, "Discriminant analysis for recognition of human face images," J. Opt. Soc. Amer. A, Opt. Image Sci., vol. 14, no. 8, pp. 1724–1733, 1997.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 3, March 2019

- [7] L. Wiskott, J.-M. Fellous, N. Kuiger, and C. von der Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 775–779, Jul. 1997.
- [8] M. S. Bartlett, J. R. Movellan, and T. J. Sejnowski, "Face recognition by independent component analysis," *IEEE Trans. Neural Netw.*, vol. 13, no. 6, pp. 1450–1464, Nov. 2002.
- [9] J. Yang, D. Zhang, A. F. Frangi, and J. Yang, "Two-dimensional PCA: A new approach to appearance-based face representation and recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 1, pp. 131–137, Jan. 2004.
- [10] D.-J. Kim, K.-W. Chung, and K.-S. Hong, "Person authentication using face, teeth and voice modalities for mobile device security," *IEEE Trans. Consum. Electron.*, vol. 56, no. 4, pp. 2678–2685, Nov. 2010.
- [11] M. S. Bartlett, G. Littlewort, I. Fasel, and J. R. Movellan, "Real time face detection and facial expression recognition: Development and applications to human computer interaction," in *Proc. Comput. Vis. Pattern Recognit. Workshop*, Jun. 2003, p. 53.
- [12] M. Yeasin, B. Bulot, and R. Sharma, "From facial expression to level of interest: A spatio-temporal approach," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2004, pp. 922–927.
- [13] C. Shan, S. Gong, and P. W. McOwan, "Robust facial expression recognition using local binary patterns," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2005, pp. 370–373.
- [14] P. Ekman, "Universals and cultural differences in facial expressions of emotion," in *Proc. Nebraska Symp. Motiv.*, 1987, p. 712.
- [15] S.S. Tomkins, *Affect, Imagery, Consciousness: Cognition Duplication and Transformation of Information*. New York, NY, USA: Springer, 1963.
- [16] J. A. Russell, "A circumplex model of affect," *J. Pers. Social Psychol.*, vol. 39, no. 6, pp. 1161–1178, Dec. 1980.
- [17] L. F. Barrett and J. A. Russell, "Independence and bipolarity in the structure of current affect," *J. Pers. Social Psychol.*, vol. 74, no. 4, pp. 967–984, 1998.