International Journal of Computer Science and Engineering Communications,
Volume.5, Issue.4 (2017): Page.1702-1708
www.ijcsec.com

Facial Micro-Expression Recognition using Feature Extraction

[1]Anju P. SureshBabu, [2]Muralidharan K.B, [3]Arjun K. P
[1]PG Scholar, [2,3]Assistant Professor,
[1, 2, 3] Department of Computer Science and Engineering,
[1, 2] Cochin University of Science and Technology, Ernakulam, India
[3] Royal College of Engineering and Technology, Thrissur, India
anjupattali@cusat.ac.in, kbmuralidharan@cusat.ac.in, arjunkppc@gmail.com

Abstract: *Micro-expressions are rapid, involuntary facial expressions which reveal emotions that people do not intend to show. Studying small expressions is effective as recognizing them has several vital applications, notably in forensic science and psychotherapy. However, analyzing spontaneous Micro expressions is very challenging due to their short duration and low intensity. Nowadays, Video-based facial expression recognition has received significant attention because of its wide spread applications. One key issue for video-based facial expression analysis in practice is how to extract dynamic features. This work proposes a novel approach using histogram sequence of local binary patterns from three orthogonal planes (LBP-TOP). In this approach, every facial image sequence is firstly convolved with Gabor filters to extract the Gabor Magnitude Sequences (GMSs) that extract local information of magnitude, phase and orientation. The main Facial action units which are located at eye and mouth regions of face are extracted and apply feature extraction to increase the discrimination for micro-expressions. For classification nearest neighbour method is exploited. Our experimental results on CASME2 database demonstrate that proposed method achieved better results compared to other methods in past years.*

Keywords– *Micro-Expressions, Gabor filters, LBP-TOP, Nearest Neighbour method.*

I. Introduction

The human face is a medium that sends us a message. When we “read” a face, there is quite a lot of data to sift through. One part of the medium is its basic structure and muscle tone. We also perceive changes that have taken place such as scarring, weathering of skin or wrinkles. A facial expression is one or more motions or positions of the muscles beneath the skin of the face. Facial expressions are a form of nonverbal communication. Micro expressions are very brief facial expressions, lasting only a fraction of a second. They occur when a person either deliberately or unconsciously conceals a feeling. These expressions provide a full picture of the concealed emotion but so quickly that it is usually missed. Micro-expression was first discovered by Haggard and Isaacs in 1966 [1]. In an accidental opportunity, Ekman and Friesen also found the micro-expression in 1969[14] [2]. Dr. Ekman's research has revealed that seven emotions have universal signals such as anger, fear, sadness, disgust, contempt, surprise and happiness [6].

We combine the most effective feature extraction method LBP-TOP (Local Binary Patterns from Three Orthogonal Planes) with the most Useful filtering method Gabor filters [13] and direct classification method the nearest neighbour. We propose to extend the previously successful LBP method to the temporal domain, resulting in Local Binary Patterns from Three Orthogonal Planes (LBP-TOP). The structure of this paper is organized as follows, Section 2 we discuss the proposed methods. Experimental results are discussed in Section 3 and we conclude the paper in Section 4.

II. Proposed Methods

We present the framework to recognize micro-expressions in eye and mouth regions. At first, gray normalization performs to pre-process the image sequences as shown in figure 1. To recognize micro-expressions from eye and mouth region, it needs to locate and obtain the eye and mouth region from the pre-processed whole face images shown in figure 2. Then apply gabor filters to filtering the images and for feature extraction uses LBP-TOP, a powerful method to describe the texture features.

A) Eye and Mouth Region Extraction

To extract eye and mouth regions from the whole face we use Haar Feature-based Cascade Classifiers. Object Detection using Haar feature-based cascade classifiers is an efficient object detection method proposed by Paul Viola and Michael Jones [5].

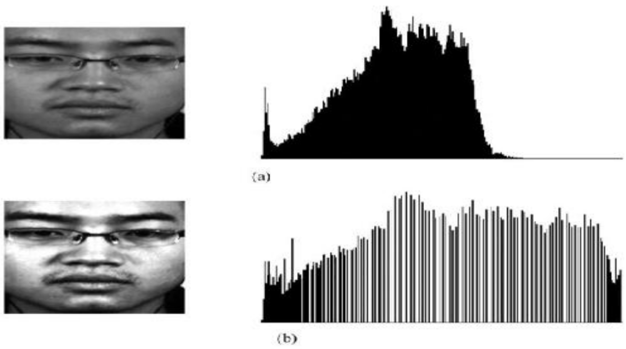


Figure 1: The effect of Gray Normalization

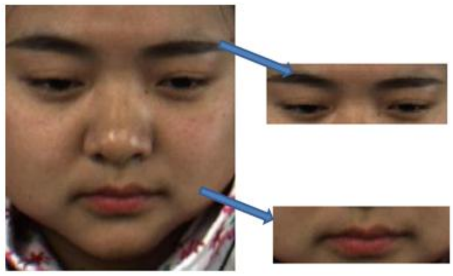


Figure 2: Illustration of extracting eye and mouth region.

B) Gabor Filter of image

In image processing, a Gabor filter [9], named after Dennis Gabor filter, is linear filter used for edge detection. Frequency and orientation representations of Gabor filters are similar to those of the human visual system, and they have used particularly appropriate for texture representation [10]. In the spatial domain, a 2D Gabor filter is a Gaussian kernel function modulated by a sinusoidal plane wave.

Figure 3: Gabor filters of images.

International Journal of Computer Science and Engineering Communications,
Volume.5, Issue.4 (2017): Page.1702-1708
www.ijcsec.com

An example of how an original image results in a set of Gabor Pictures after being convolved with a set of Gabor filters is shown in figure 3.

C) Local Binary Patterns from Three Orthogonal Planes(LBP-TOP)

In order to combine motion with appearance feature of a facial micro-expression sequence, we encode the Gabor feature representations with LBP-TOP [4] operators to further enhance the feature extraction. Local binary patterns from three orthogonal planes (LBP-TOP), proposed by Zhao, extends LBP to dynamic texture (DT) [7].

LBP-TOP combines the temporal and spatial features of an image sequence on the basis of LBP, and extracts the dynamic texture features of image sequences from three orthogonal planes. These dynamic texture features are used to express the spatial and temporal and motion characteristics of image sequences. The same as LBP operator to extract the features, LBP-TOP operator also see a pixel in the image sequence as the centre pixel and uniformly extract neighbourhood points from its three orthogonal planes.

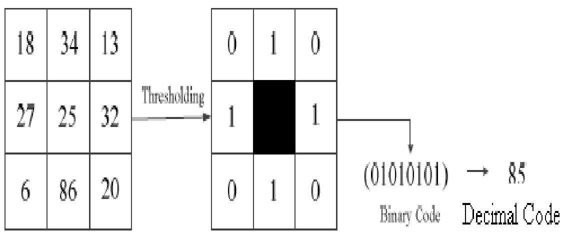


Figure 4: The calculation process of a basic LBP operator.

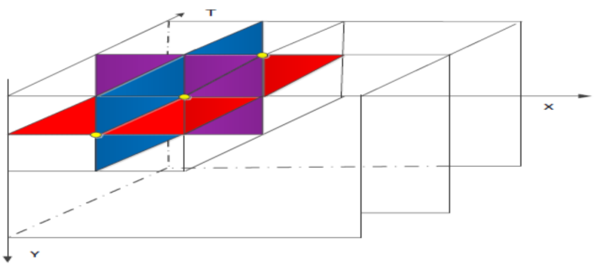


Figure 5: Extract neighbourhood points from three orthogonal planes.

D) Nearest Neighbor Method

The features extracted for each facial image sequences and eye and mouth region sequences in this paper are 3*59 matrixes. Then it transformed to 1*177 dimensional row vectors. Then we classify the features by calculating the spatial distance between each feature.

This paper selects the nearest neighbour method [3] as the classification method, because it can directly compare the distances between feature vectors. In all distance decisions, Euclidean distance is very best at measuring the similarity between samples. Here we choose Euclidean distance as the distance measurement. Euclidean distance is defined as follows. Suppose there is a two-dimensional space. Its coordinate axis are X and Y. The distance d_{12} between two points S_1 and S_2 in this space is (1),

$$d_{12} = \sqrt{(dx^2 + dy^2)} \quad (1)$$

Where: $$dx = x_2 - x_1, dy = y_2 - y_1 \quad (2)$$

The classification rule of the nearest neighbour method is to calculate the distance between each point in the feature space, and see the distances as the similarity measurement of samples. If the distance of two samples in the feature space is minimum and is very close, then the similarity of them is great.

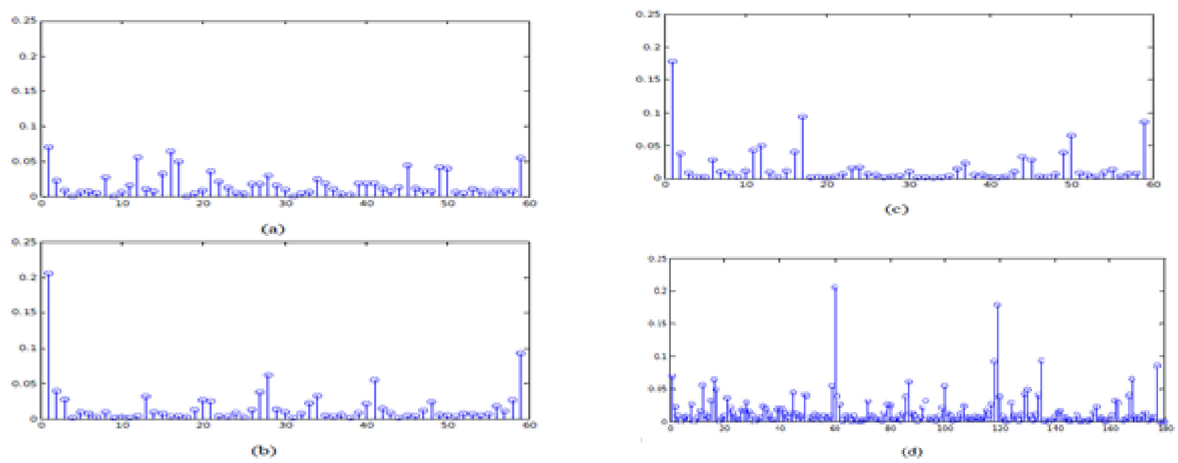
IV. Evaluation of Methods and Results

To understand the role of eye and mouth region in micro-expression recognition, we applied above proposed framework on the spontaneous database. The comparisons of performance between eye and mouth region and whole face regions are analyzed. Our aim is to ascertain that the spatio-temporal descriptor LBP-TOP with gabor filter results in more accurate in micro-expression spotting than both its static counterpart LBP with gabor filter and other dynamic appearance descriptors, in particular LBP-TOP, for this we uses a 195 micro-expressions filmed under 60fps of 35 participants from CASME2 Dataset [12].

A) LGBP-TOP

For each pixel in a whole face, eye and mouth regions, compare the pixel to each of its 8 neighbors (on its left-top, left- middle, left-bottom, right-top) [11]. Follow the pixels along a clockwise or anti-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 then converted to decimal for easy classification .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 feature vector [6]. Optionally normalize the histogram. Concatenate (normalized) histograms of all cells. This gives a feature vector for the entire window. The LBP responses, with only 59 patterns representing only strong edges and corners.

In LGBP, the Gabor [12] and LBP filtering follow one after the other. However, whereas in normal LBP the LBP filter operates only on the original image, in LGBP it operates on a number of filtered images, because a bank of different Gabor filters is used to generate response images. The final LGBP feature histogram of the image is then composed by a simple concatenation of the histograms composed for each GP, with histogram blocks in the same manner as for LBP. We have used 18 Gabor filters and apply LBP to a facial regions we would obtain a feature histogram descriptor of dimensionality 18*59=1062.



We use the LGBP-TOP method to extract dynamic features, then divide the sequences into pieces of x*y*t. So the changes of dynamic texture can be Easily seen. In this paper don't make the resolution of sequences in blocks. A facial micro-expression sequence is seen as a whole block. Then we extract the dynamic features of the sequences like that a center pixel of a frame in the sequence is *(x, y, t)*.Then extract its LGBP features from three orthogonal planes. Calculate the decimal value of the LBP features and record them as (3),

$$f_0(x_c, y_c, t_c), f_1(x_c, y_c, t_c), f_2(x_c, y_c, t_c) \quad (3)$$

Figure 7: LGBP features of three orthogonal planes. The horizontal axis represents the features after normalization. The vertical axis represents the corresponding proportions.(a)is the LGBP histogram of the XY plane, (b)is the LGBP histogram of the XT plane, (c)is the LGBP histogram of the YT plane, (d)is the LGBP histogram cascaded by three orthogonal planes, namely the LGBP-TOP feature of the sequence. Due to the differences between individual subjects, the sizes of cropping face regions and eye, mouth regions are not the same. Because the lengths of the sequences are different, the dimensions of extracted features are different. Therefore, we normalize the feature vector. In this paper, the features extracted are 3*59 matrixes for each filtered images .We normalize them to 1*177 dimensional row vectors, in order to calculate efficiently. One of the result histograms is as shown in Figure 7.

B) Use the Nearest Neighbor Method to Recognize Microexpression

In the feature extraction part we get a series of 1*177 dimensional row vectors for each set of gabor pictures, it is the final feature of the sequence. We randomly select some feature vectors as the training set, the remaining as the testing set. Then compare the Euclidean distance between each testing set and each training set and use the nearest neighbor method for classification [13]. Suppose there are m training set x and n testing set y. Define the distance between a testing vector y_j and a training vector x_i as in equation (4). Then, for each y_j, calculate its distances with the entire training sets. For each testing set, there is a nearest distance. The corresponding training set indicates the category of that testing set [8].

$$d_{ij} = \sqrt{\sum_{k=1}^{177} \left(x_{ik} - y_{jk} \right)^2} \quad (4)$$

The following formula (5) is used to calculate the nearest distance:

$$d_j = \min d_{ij} \quad (5)$$

In the experiment we select different expressions, In the feature extraction stage, we extract the LGBP-TOP feature of these sequences and save them. In the classification stage, we randomly select some feature vectors as training sets, the remaining as the testing sets [15]. All the samples are randomly divided into many groups of roughly equal size. For example, in the experiment when the ratio of the training set and testing set is 5:1, 1/6 of the total groups are selected as the testing samples. Some experimental results are shown in Table 1. Also compare the results with other feature extraction methods like LBP, LBP-TOP, LGBP on performing with CASMEII datasets and the accuracy results are shown in figure 8.

TABLE I. Recognition Results

Training set: Testing set	Recognition results	
	Face Region	*Eye and Mouth Region*
5:1	83%	91%
4:1	82.15%	89.12%
3:1	81.76%	88.54%
2:1	81.01%	86.53%
1:1	80.23%	84.30%

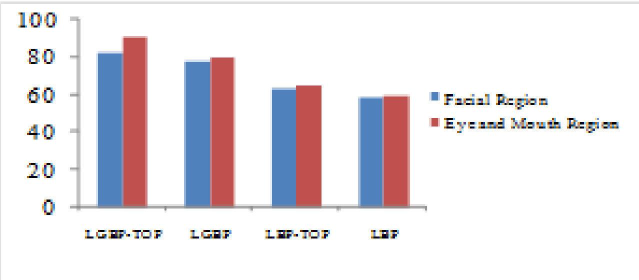


Figure 8: Accuracy of different feature extracti
methods performed on CASMEII dataset.

V. Conclusion

In this paper, we select a new combination of procedures and database, and recognize facial micro-expression successfully. Framework is presented to recognize micro-expressions from eye and mouth regions. Then all the image sequence is represented as 18 different Gabor filters. Then, LBP-TOP feature vector is extracted from an image sequence of eye and mouth regions corresponding to an original micro-expression sample. We compared the performances between the proposed frameworks with whole face region. The experimental results show that the proposed method performs better than whole face when recognizing micro-expressions. It validates that the eye and mouth region is more discriminative than whole face on these micro-expressions. And also compared the results of LGBP-TOP feature vector with previous methods and find out that proposed method gets high accuracy.

REFERENCES

[1]E. A. Haggard,et. al., “Micromomentary facial expressions as indicators of ego mechanisms in psychotherapy,” In: Auerback A, Ed: Methods of Research in Psychotherapy. New York: Appleton Century Crofts,pp.154-165,1966.

[2]P. Ekman and W. V. Friesen, “Nonverbal leakage and clues to deception,” Psychiatry, vol. 32, pp. 88–97, February 1969.

[3]Guo, Yanjun, et al. "Micro-expression recognition based on local binary patterns from three orthogonal planes and nearest neighbor method." Neural Networks (IJCNN), 2014 International Joint Conference on. IEEE, 2014.

[4]C. Shan,et. al., Facial expression recognition based on Local Binary Patterns: A comprehensive study, *Image and Vision Computing* , 27(6): 803–816, 2009.

[5]Duan, Xiaodong, et al. "Recognizing spontaneous micro-expression from eye region." Neurocomputing 217 (2016): 27-36.

[6]L.Wang,R.F.Li,K.Wang,etal.,”Feature representation for facial expression recognition based on FACS and LBP”, Int.J. Autom. Comput. 11(5) (2015) 459–468.

[7]Q.Zou,J.Zeng,L.Cao,etal.,”A novel features ranking metric with application to scalable visual and bioinformatics data classification”, Neurocomputing 173 (2015)346–354.

[8]C.Ding,H.Peng,”Minimum redundancy feature selection from microarray gene expression data”, J.Bioinform. Comput.Biol. 3(2)(2003)523–528.

[9]J. Zou, C. Liu, Y. Zhang and G. Lu, Object recognition using Gabor co-occurrence similarity, *Pattern Recognition*, 46(1):434–448, 2012.

[10] M. Yang and L. Zhang, Gabor feature based sparse representation for face recognition with gabor occlusion dictionary, In *European Conference on Computer Vision (ECCV)*, 2010: 448–461.

[11] S. Polikovsky, Y. Kameda and Y. Ohta, “Facial micro-expression recognition using high speed camera and 3D-gradient descriptor,” 3rd International Conference on Imaging for Crime Detection and Prevention (ICDP 2009), January 2009.

[12] K. Wang, B. Chen and G. Wu, Edge detection from high-resolution remotely sensed imagery based on Gabor filter in frequency domain, In *IEEE* International Conference on Geoinformatics, 2010: 1–6.

[13] Yan, Wen-Jing, et al. "CASME database: A dataset of spontaneous micro-expressions collected from neutralized faces." *Automatic Face and Gesture Recognition (FG), 2013 10th IEEE International Conference and Workshops on*. IEEE, 2013

[14] Ekman, Paul. *Telling lies: Clues to deceit in the marketplace, politics, and marriage (revised edition)*. WW Norton & Company, 2009.

[15] Almaev,et. al., "Local gabor binary patterns from three orthogonal planes for automatic facial expression recognition." *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, 2013.