Journal of Advances in Computer Research
Quarterly ISSN: 2008-6148
Sari Branch, Islamic Azad University, Sari, I.R.Iran
(Vol. 11, No. 4, November 2020), Pages: 107-115
www.jacr.iausari.ac.ir

Facial Expression Recognition Using Texture and Edge Descriptors

Davar Giveki✉1, **Nastaran Mirzaei**2
1) Department of Computer Engineering, Malayer University, P. O. Box 65719-95863, Malayer, Iran
2) Faculty of Engineering, Lorestan University/ Pol-e Dokhtar , Iran

davood.giveki@gmail.com; nastaranmirzaei3@gmail.com

Received: 2020/11/24; Accepted: 2021/04/22

Abstract

Facial expression recognition is one of the most important computer vision issues that has many applications. In this paper, an efficient method for facial expression recognition using texture and edge descriptors is proposed. Facial expression recognition generally consists of three steps: preprocessing, feature extraction and classification. In this paper, histogram equalization has been used as pre-processing method to increase the quality of the input face images. In this paper, our focus is on the feature extraction by a combination of LDP[1] and HOG[2] descriptors. After feature extraction, the support vector machine was used to classify the face images. In our experiments we used JAFFE database. The database contains 213 images of seven facial expressions (happy, sad, angry, fear, disgust, surprised and natural) taken from 10 Japanese female models. The results showed that the proposed method achieved 99.04% accuracy on JAFFE dataset which is a higher performance compared to other methods.

Keywords: *facial expression recognition, histogram equalization, texture and edge descriptors, LDP, HOG, Support Vector Machine.*

1. Introduction

In recent years, one of the active fields of image processing was the facial expressions recognition. Face plays an important role in emotions recognition and in social relations and non-verbal communication [1]. The human face shows many information and characteristics such as facial expressions, identity, age, gender, as well as information about the state of fatigue, interest, confusion and stress[2]. The seven main facial expressions that are often considered for diagnosis are: happy, sad, angry, fear, surprised, disgust and neutral [3]. Facial expressions are created by changes in different parts of the face. Non-verbal communication or body language is a set of body movements and gesture that take place to exchange information. Face shape is one of the effective factors in communication [4]. Hence, face analysis can reveal valuable information. Mouth, lips, eyebrows, eyes, etc are the main parts of a face that are informative in facial expression

[1]Local Derivative Pattern (LDP)
[2]Histogram of Oriented Gradient (HOG)

recognitions [5]. There are some challenges in facial expressions recognition. For instance, the similarities between the different poses make it more difficult to distinguish the facial expressions. Another complication we face in real life is the presence of occlusion on the face. This occlusion is due to the use of glasses, scarves, hats, scarves and placing hands on the face. In real situations, a part of the face is often covered by beards and mustaches and etc. These issues emphasize that the facial expression recognition is still a challenging issue in image processing and computer vision [6]. In this article, we try to present a new method using a combination of different descriptors to detect facial expressions. The rest of the paper is organized as followings. Section 2 deals with the related works. In section 3, the JAFFE database is explained. Section 4 elaborates the LDP and the HOG descriptors. The proposed method is discussed on section 5. The results of the experiments are given in section 6. A comparison between the proposed method and other methods is reported in section 7. Finally, the paper is concluded in section 8.

2. Related Works

Facial expression is one of the most powerful, natural and universal signals for human beings to convey their emotional states and intentions [7]. Gholami used LGBP method to extract features and K-nn method to classify different facial expressions and reached 96% accuracy [8]. Hamster et al. used convolutional network to detect facial expressions and achieved 94.1% accuracy [9]. Stanosh and Kumar proposed a combination of PCA and LBP descriptors to represent image content. They employed a neural network for classification, and achieved an accuracy of 64.28% [10]. Happy and Rotary investigated PHOG and LBP methods to extract facial features. Also, support vector machine was used for classification. The authors achieved 83.86% accuracy [11]. Zhang et al. adopted Gabor wavelet coefficients and geometric positions to construct the feature vector for each image and applied a two-layer perceptron to classify seven different facial expressions. Their recognition rate is 90.1% [12]. In this paper, we apply the histogram equalization to preprocess our images. Then, we fuse features extracted by LDP and HOG descriptors. Finally, a support vector machine classifier with Gaussian kernel is used to classify different facial expressions.

3. Database Used

The database used in this article is the JAFFE database. The images in this database were taken in the psychology department of Kyushu University. It contains 213 images of seven facial expressions (happy, sad, surprised, fear, disgust, angry and natural) taken by 10 Japanese female models. Each image is categorized into 60 emotional states by 60 Japanese [13]. Figure 1 shows an example of the images in this database.

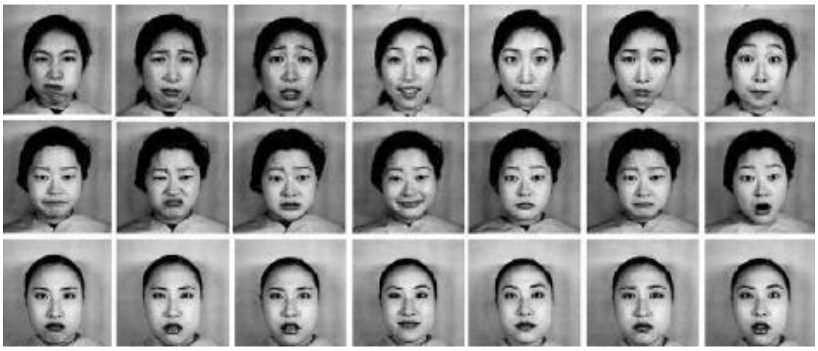

Figure 1. An example of a JAFFE database

4. LDP descriptor

The LDP descriptor was proposed by Buchang Zhang for face recognition as a high-end descriptor [14]. The LDP operator calculates the derivative in different directions of order $(n-1)^{th}$ based on the binary function. The LBP operator calculates the first-order derivative in all directions, while the LDP computes the higher-order derivative, which contains more properties. The LBP operator cannot display more definite features, while the LDP extracts more detail from the image. Considering the image $I(z)$, the first-order derivatives of the image $I(z)$ in the directions 0, 45, 90, 135 degrees are determined by $I'_{\alpha}(z)$ that α is= 0°, 45°, 90°, 135°. As shown in Figure 2, assuming that z_0 is a point in the image $I(z)$ and z_i (i = 1,…, 8) is the neighboring point around z_0. The four first-order derivatives at $z = z_0$ can be written as

$$I'_{0^{\circ}}(z_0) = I(z_0) - I(z_4) \quad (1)$$

$$I'_{45^{\circ}}(z_0) = I(z_0) - I(z_4) \quad (2)$$

$$I'_{90^{\circ}}(z_0) = I(z_0) - I(z_4) \quad (3)$$

$$I'_{135^{\circ}}(z_0) = I(z_0) - I(z_4) \quad (4)$$

Z_1	Z_2	Z_3
Z_8	Z_0	Z_4
Z_7	Z_6	Z_5

Figure 2. An example of eight neighbors around the point z_0

The second-order directional LDP in α direction, in z=z_0 point is defined as.

$$LDP_{\alpha}^{2} = \{f\ (I_{\alpha}^{'}(z_0), I_{\alpha}^{'}(z_1)), f\ (I_{\alpha}^{'}(z_0), I_{\alpha}^{'}(z_2)..., f\ (I_{\alpha}^{'}(z_0), (I_{\alpha}^{'}(z_8))\} \quad (5)$$

In a general formulation, the nth -order LDP is a binary string (n-1)th describing gradient as trend changes in a local region of directional -order derivative images

$$LDP_{\alpha}^{n} = \{f\ (I_{\alpha}^{n-1}(z_0), I_{\alpha}^{n-1}(z_1)), f\ (I_{\alpha}^{n-1}(z_0), I_{\alpha}^{n-1}(z_2)..., f\ (I_{\alpha}^{n-1}(z_0), (I_{\alpha}^{n-1}(z_8))\}$$

(6)

In this formula, $I_{\alpha}^{n-1}(z_0)$, is a derivative of the order $(n-1)^{th}$ in the direction ɑ at the point $z = z_0$. In this paper, the focus is on the feature extraction step and to improve the existing methods, a combination of LDP and HOG descriptors is used to extract the image features.

5. HOG descriptor

A HOG descriptor is an image descriptor used in computer vision and image processing to identify humans. Dalal and Triggs first used the HOG method in 2005 for pedestrian detection [15]. Recently, the issue of facial recognition has been investigated using the HOG descriptor [16]. The basic idea in the HOG descriptor is that a shape can be well described by the distribution of edge directions or gradients. In this method, the image is first divided into several small areas, called cells, and the image is divided into blocks with 50% overlap. The gradients in the x and y directions are then calculated per pixel. The next step is to obtain the histograms of each cell. To do this, at first, the distance between 0 and 180 degrees or 0 to 360 degrees, depending on whether the gradients are signified or not, are divided into several equal parts, each part forming a histogram bin. Then for each cell, a histogram of the edge directions is calculated. In this way, each pixel in the cell votes for a histogram bin based on the direction of the edge and the amount of gradient in it. Finally, histograms are normalized to compensate for light intensity. To do this, it considers several neighboring cells as a block and normalizes the histogram of these cells. The combination of these histograms ultimately represents a HOG descriptor. As shown in Figure 3, a 64 * 64 image is divided into the some cells. These cells are categorized and each of the four cells is called a block. The number of cells in each block is 2 * 2. On the other hand, the size of each cell is 8 * 8. As a result, seven blocks are seen in the horizontal direction and seven blocks in the vertical direction. Therefore, in total there are 49 blocks in each image. To improve detection, blocks are considered so that each block has 50% overlap with the adjacent block. It is necessary to use this technique to increase the accuracy of performance [16]. In this paper, to evaluate the effect of changing HOG descriptor parameters such as cell and block sizes, 2 * 2 cell size has been used.

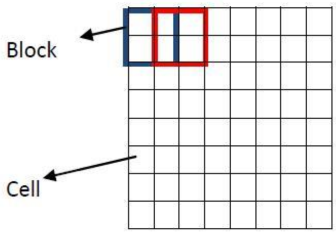

Figure 3. An example of blocks and cells in an image with a size of 64 * 64

6. Proposed Method

To reach a high performance facial expression recognition method, a new combination of image descriptors has been used to extract facial features. HOG descriptor's basic idea is that a shape can be well described by distributing edge directions or gradients. The LDP descriptor calculates the higher-order derivative, which contains more accurate and more textural features. Therefore, high accuracy can be achieved by combining these two complementary descriptors. The flow diagram of the proposed method is shown in Figure 4.

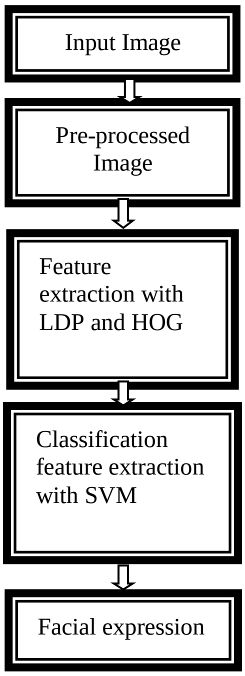

Figure 4. The flow diagram for proposed method

6.1 Pre-processed Image

In the preprocessing stage, the data is preprocessed using the histogram equalization. When we say low image contrast, we mean that the difference between the minimum and maximum brightness of the image is small. Through this setting, the intensity can be better distributed in the histogram. Figure 5 shows an example of a database image and a pre-processed image.

Figure 5. An example of a database image (left) and a pre-processed image (right)

6.2 Feature extraction

In this paper, the focus is on the feature extraction step and to improve the existing methods. LDP and HOG descriptors, and a combination of two LDP and HOG descriptors have been used separately. In the proposed method, after preprocessing, LDP and HOG descriptors are used to extract the image features.

6.3 Classification

In the final step, the database is divided into two parts, training and test, and then the facial expressions are classified using the nonlinear support vector machine with Gaussian kernel. In this paper, 50% of the data was used for training and 50% of the data for test. To avoid overfitting, 5-fold cross validation is investigated.

7. Experimental Results

As can be seen in Table1, the proposed method is compared with the methods proposed by other researchers for facial expression recognition on JAFFE database.

Table 1. Comparison of the proposed method with other methods on the JAFFE database

Row	Researchers	Feature extraction method	Classification method	Accuracy
1	[7]	LGBP	K-nn	%96
2	[9]	Multi-channel convolution network	Multi-channel convolution network	%94.1
3	[10]	PHOG+LBP	Support Vector Machin	%83.86
4	[12]	PCA+LBP	Neural Network	%64.28
5	Proposed method	LDP+HOG	Support Vector Machin	%99.04

Table 1 compares our proposed method with other methods on the JAFFE database. As shown in this table, using a combination of LDP and HOG descriptors, the proposed method has the best performance compared to the methods presented by other researchers. To evaluate the performance of the proposed method in this paper, the confusion matrix and the mean Average Precision (mAP) are reported. Figures 6 and 7 show the confusion matrix and the mean Average Precision diagrams of the proposed method, respectively.

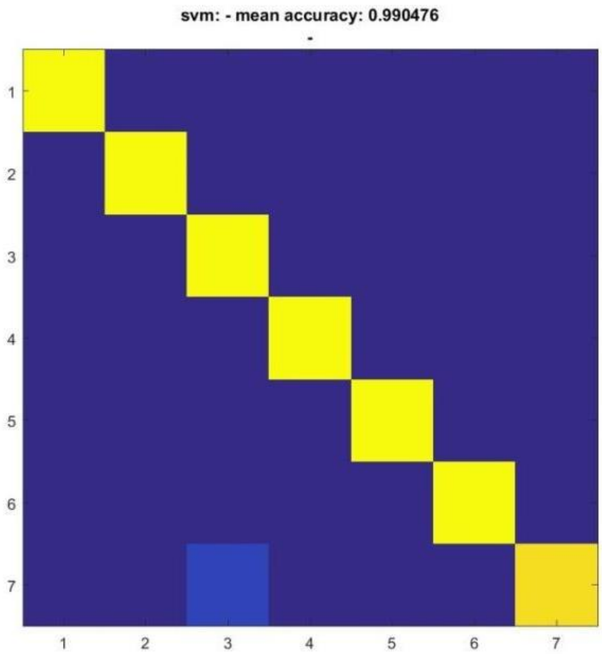

Figure 6. Confusion matrix of obtained by the proposed

As can be seen in Figure 6, class seven (surprise), is confused with class three (fear). The rest of the classes are correctly identified. Figure 7 shows the mean Average Precision obtained by our proposed method.

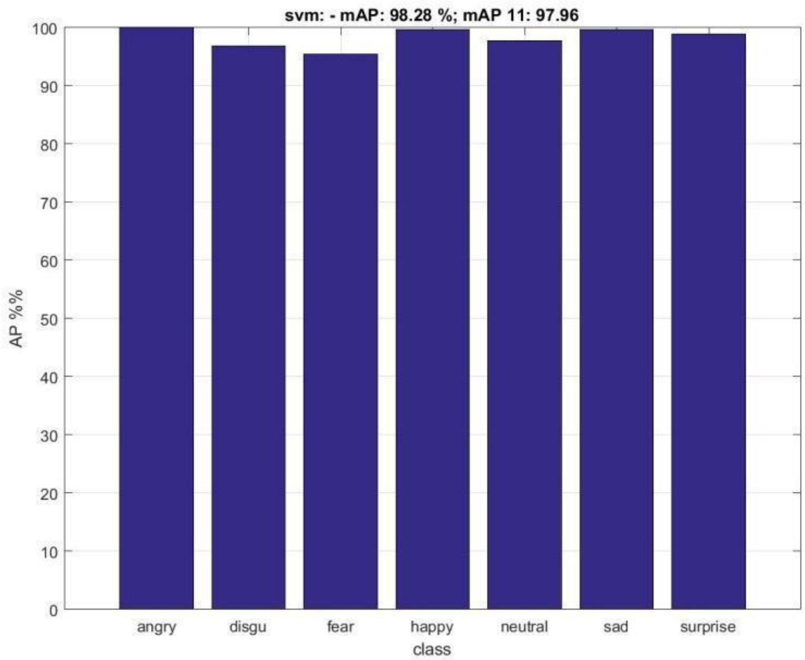

Figure 7. Graph of the mean Average Precision obtained for 2 * 2 cell size using LDP and HOG

As can be seen in Figure 7, class one (anger), is more easily identified than the other classes, and class two (hate) is more difficult to be distinguished compared the other classes. The mean Average Precision using LDP and HOG is 98.28%.

8. Conclusion

For many years, facial expressions recognition, regardless of a person's identity, has received a great deal of attention in image processing and machine vision applications. There are different challenges in facial expression recognition. For example, occlusion is usually happening for faces in crowd. In real life, part of the face is often covered. This concealment is due to the use of glasses, hats, scarves and placing hands on the face. These issues have made face recognition to be still a challenging issue in image processing. One of the most important steps in recognizing facial expressions is feature extraction. In this paper, the feature extraction method is investigated by fusing two descriptors, LDP and HOG. Nonlinear support vector machine classifier with Gaussian kernel was used to classify each facial expression. Our proposed method reached 99.04% accuracy on JAFFE dataset.

References

[1] Al-Qaisi, A., AlTarawneh, M. S., ElSaid, A., & Alqadi, Z. (2020). A Hybrid Method of Face Feature Extraction, Classification Based on MLBP and Layered-Recurrent Network. *Traitement du Signal*, *37*(4), 555-561.

[2] Eliasi, M., Manzuri, M., Yaghoubi, Z., Eliasi, A. (2014). Novel Texture Description and Face Identification Methods by Defining Bridle Paths and Using Gabor Phases. *Journal of Advances in Computer Research*, 5(3), 35-53.

[3] Goyal, S. J., Upadhyay, A. K., Jadon, R. S., & Goyal, R. (2018). Real-Life Facial Expression Recognition Systems: A Review. In *Smart Computing and Informatics* (pp. 311-331). Springer, Singapore.
[4] Motameni, H. (2019). Face recognition using sparce reprasentations and p-laplacian. *Journal of Advances in Computer Research*, 10(4), 37-49.
[5] Qayyum, H., Majid, M., Anwar, S. M., & Khan, B. (2017). Facial expression recognition using stationary wavelet transform features. *Mathematical Problems inEngineering*, *2017*.
[6] Yang, S., Luo, P., Loy, C. C., & Tang, X. (2016). Wider face: A face detection benchmark. In *Proceedings of the IEEE Conference on Computer Vision and PatternRecognition* (pp. 5525-5533).
[7] Krumhuber, E. G., Skora, L., Küster, D., & Fou, L. (2017). A review of dynamic datasets for facial expression research. *Emotion Review*, *9*(3), 280-292.
[8] Gholami, F. (2017). Facial Expressions Recognition Based on a Combination of the Basic Facial Expression Using Weighted the Local Gabor Binary Pattern (LGBP). *International Journal of Computer Science and Network Security (IJCSNS)*, *17*(4), 352.
[9] Hamester, D., Barros, P., & Wermter, S. (2015, July). Face expression recognition with a 2-channel convolutional neural network. In *Neural Networks (IJCNN), 2015International Joint Conference on* (pp. 1-8). IEEE
[10] Santosh, M., & Kumar, A. (2017). Superiority of PCA algorithm for Facial Expression Recognition. *International Journal*, *8*(3).
[11] Happy, S. L., & Routray, A. (2015). Automatic facial expression recognition using features of salient facial patches. *IEEE transactions on Affective Computing*, *6*(1), 1-12.
[12] Yang, S., Luo, P., Loy, C. C., & Tang, X. (2016). Wider face: A face detection benchmark. In *Proceedings of the IEEE Conference on Computer Vision and PatternRecognition* (pp. 5525-5533).
[13] Santosh, M., & Kumar, A. (2017). Superiority of PCA algorithm for Facial Expression Recognition. *International Journal*, *8*(3).
[14] Parameswaran, S., Parameswaran, M., Joseph, S., & Viswanathan, D. M. (2020). Exploring Various Aspects of Gabor Filter in Classifying Facial Expression. In *Advances in Communication Systems and Networks* (pp. 487-500). Springer, Singapore.
[15] Gowda, S. N. (2017). Human activity recognition using combinatorial Deep Belief Networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (pp. 1-6).
[16] Allagwail, S., & Elbkosh, A. (2020). Face Recognition with Symmetrical Face Training Samples Based on Histograms of Oriented Gradients.
[17] Alsubari, A., & Ramteke, R. J. (2017). Extraction of Face and Palmprint Features based on LBP, HOG and Zernike Moments. *Extraction*, *172*(5).