2019 IEEE 3rd Information Technology,Networking,Electronic and Automation Control Conference (ITNEC 2019)

Deep Learning for Face Recognition under Complex Illumination Conditions Based on Log-Gabor and LBP

Liyun Zhuang[1,2] ,Yepeng Guan[1,3]

1. School of Communication and Information Engineering, Shanghai University, Shanghai, China
2. Faculty of Electronic and Information Engineering, Huaiyin Institute of Technology, huaian, Jiangsu, China
3. Key Laboratory of Advanced Displays and System Application, Ministry of Education, Shanghai, China

zly418@126.com , ypguan@shu.edu.cn

Abstract—**Complex illumination condition is one of the most critical challenging problems for practical face recognition. In this paper, we propose a novel method based on deep learning to solve the adverse impact imposed by illumination variation in the face recognition process. Firstly, illumination preprocessing is applied to improve the adverse effects of intense illumination changes on face images. Secondly, the Log-Gabor filter is used to obtain the Log-Gabor feature images of different scales and directions, then, LBP (Local Binary Pattern) features of images subblock is extracted. Lastly, texture feature histograms are formed and input into the deep belief network (DBN) visual layer, then face classification and recognition are completed through deep learning in DBN. Experimental results show that superior performance can be obtained in the developed approach by comparisons with some state-of-the-arts.**

Keywords—complex illumination; Log-Gabor filter; LBP features; deep learning

I. INTRODUCTION

Face recognition has been a research hotspot in pattern recognition and image processing in the past few years due to its friendliness and convenience. In recent years, many algorithms have been proposed by scholars[1-4]. Most previous researches can obtain satisfying recognition performance under uniform illumination conditions with frontal face images. However, it is still a very challenging research area because even the images of same person seem different due to occlusion, illumination, expression and pose variation, which can cause sharp decline in recognition rate [5,6]. Improvement of face recognition performance in complex light environment is still a difficult problem in the field of artificial intelligence and computer vision.

At present, domestic and overseas researchers have proposed numerous illumination processing algorithms for face images under complex illumination conditions, and achieved relatively good experimental results. They can be roughly divided into two categories: 2-D and 3-D based models. Face recognition method based on 3-D model is very effective to overcome the influence of attitude and illumination in environmental factors. However, the 3-D model method is complex and has a long fitting time, so it is difficult to achieve real-time requirements. The algorithms based on 2-D model are the research hotspots. Local binary pattern (LBP) [7-10] is an effective local texture descriptor. It is widely used in face recognition because of its advantages in image texture description. Wolf et al. [11] optimize the description of LBP(local binary pattern) and combine it with Gabor wavelet to obtain the best representation of face image features under complex illumination conditions. Marsico et al. [12] proposed the FACE (face analysis for commercial entities) algorithm for face recognition under conditions. This algorithm mainly normalizes the attitude and light under complex illumination conditions, and it can obtain accurate recognition rate. On the basis of LBP, a modified central symmetric local binary pattern (CSLBP) descriptor is proposed [13-15]. The features extracted by the existing algorithms may not achieve high discrimination, and the expression of features is overly dependent on manual selection. For the face images with severe illumination effects, they still cannot achieve satisfying result.

Recently, more and more scholars pay attention to deep learning. Deep learning simulates the deep organizational structure of brain groups, and it forms more abstract and effective high-level representation by combining low-level features[16]. Deep belief network (DBN) is a typical deep learning method. DBN automatically learns the abstract features of different levels from bottom to top, and finally obtains the nonlinear description of features. An automatic feature extraction process without artificial choice is presented. DBN has been successfully applied to handwritten numeral recognition, dynamic human detection, and many other fields[17]. However, DBN may ignore the local structure of the image and it is difficult to learn the local features of the face image [18]. At the same time, the network will learn the unfavorable feature expression for the influence of illumination and other factors when the pixel-level face features are used as the input of DBN. Liang and Zhang etc. propose to use LBP feature as the input of the deep learning network [19-21]. It improves the performance of LBP and deep learning algorithm respectively. However, for severe illumination effect, its result still cannot meet the requirement of practical application.

This paper presents an effective method to extract the robust deep features of face images under severe illumination conditions. It is on the basis of combining the Log-Gabor filter and LBP with the Deep Belief Network (DBN), which is an effective deep learning network. Firstly, the image is preprocessed to effectively improve the adverse effects of intense illumination changes on the face image. Secondly, the Log-Gabor filter is used to obtain the Log-Gabor feature images of different scales and directions. Then, LBP features of images subblock is extracted. Finally, texture feature histograms are formed and input into the deep belief network (DBN) visual layer, then face classification and recognition are completed through deep learning in DBN.

II. TECHNICAL METHOD

A. Log-Gabor filter electing

Log-Gabor function proposed by Field is an alternative to the Gabor function. The Gauss transfer function can better describe natural images at logarithmic frequency scales. The definition of Log-Gabor filter in frequency domain is as follows.

$$G(f) = \exp\left\{ \frac{\left(-\log \frac{f}{f_0}\right)^2}{2\left(-\log \frac{k}{f_0}\right)^2} \right\} \tag{1}$$

where f_0 is the filter's centre frequency and k denotes control bandwidth. There are two important characteristics of Log-Gabor filter. Firstly, log-Gabor functions, by definition, always have no DC component, therefore, the influence of illumination conditions on image processing is relatively small, which can overcome the adverse effects of illumination on face recognition to a certain extent. Secondly, the transfer function of the log Gabor function has an extended tail at the high frequency end. Log -Gabor functions, which has extended tails, should be able to encode natural images more efficiently than ordinary Gabor functions that would over-represent the low frequency components and under-represent the high frequency components in any encoding.

B. LBP Operator

LBP has the advantages of gray translation invariance, rotation invariance and simple calculation. The texture features of LBP have been successfully applied in texture classification, face recognition, image analysis, background modeling and other fields. It shows superior performance. The texture features of a two-dimensional face image $F_{M \times N}(x, y)$ can be obtained by comparing each pixel in the image with its neighboring pixels, the encoding method is defined by (2).

$$F_l(x,y)_{P,\mathrm{R}} = \sum_{p=0}^{P-1} s(g_P - g_c)2^P \tag{2}$$

where $s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x<0 \end{cases}$ g_c is the central pixel of window with size m x n, g_P (P =0, 1, 2,……P-1) indicates the evenly distributed pixels with number of P on the circle. And g_c is the center while R stands for the radius. Then different weight coefficients are assigned according to different positions of pixels and weighted summation that is obtained. The LBP value of window can be calculated as follows.

$$\mathrm{LBP}_{g_c} = \sum_{i=0}^{mn-2} s(g_i - g_c)2^i \tag{3}$$

According to the above operation method, LBP operator is applied to every pixel of the image and a corresponding LBP value can be get to describe the texture information of the image.

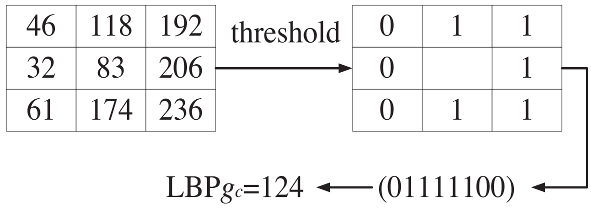

Fig. 1. Basic LBP operator.

C. DBN

Deep learning architecture is a non-supervised neural network consisting of multi-layers. The output of the former layer is usually set as the input of the latter layer. The learning aim is to make the original input information and the final output information as similar as possible by constructing the network architecture and training the parameters. Some typical deep architectures have been proposed, such as DBN [22, 23], Convolution Neural Network (CNN) [24], and so on. The DBN consists of a number of unsupervised Restricted Boltzmann Machines (RBM), as shown in Fig. 2.

For a DBN with L-layer hidden units, the joint distribution between visual units and hidden units can be expressed as.

$$\begin{aligned} &p(v,h^{(1)},h^{(2)},......,h^{(l)}) \\ &= P(v|h^{(1)})p(h^{(1)}|h^{(2)})...P(h^{(l-1)}|h^{(l)}) \end{aligned} \tag{4}$$

where v=$h^{(0)}$, v is visual unit of DBN, $h^{(k)}$(k=1,2,…,l) represents the k layer hidden unit. The hidden units in k and k+1 layers satisfy the following formulas.

$$P(h^{(k)}|h^{(k+1)}) = \prod_i P(h_i^{(k)}|h^{(k+1)}) \tag{5}$$

$$P(h_i^{(k)}=1|h^{(k+1)})=\sigma(b_i^{(k)} + \sum_j W_{ij}^{(k)} h_j^{(k+1)}) \tag{6}$$

where $\sigma(x) = 1/(1+\exp(-x))$, $b_i^{(k)}$ represents the offset of k-layer, $W_{ij}^{(k)}$ indicates the weight between the k-layer and (k+1)-layer. $P(h^{(l-1)}|h^{(l)})$ is regarded as a RBM model in DBN network.

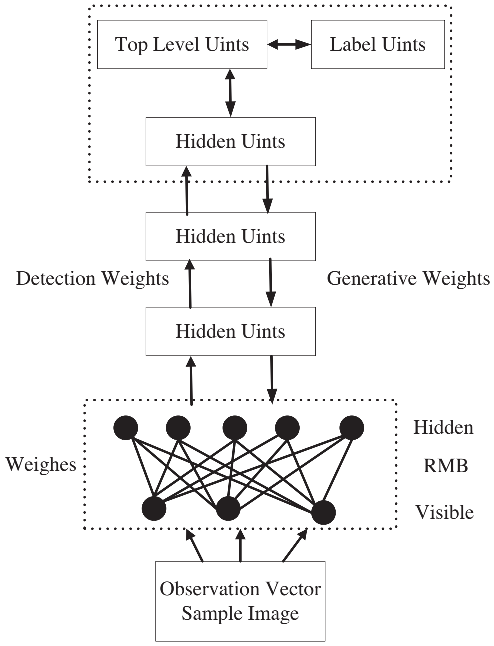

Fig. 2. DBN structure model.

DBN is a typical deep learning network. Its prototype is similar to human brain tissue structure. The DBN extracts the input data features from simple to complex and from low-level to high-level. The Soft-Max regression is applied to classify the features in the top level units, so that the class labels of input data can be obtained [24]. The DBN does not rely on the manual selection, and it learns the input data actively and digs out rich information hidden in the known data automatically.

III. PROPOSED METHOD BASED ON LOG-GABOR-LBP AND DBN (PBLGLD)

The PBLGLD is proposed in the paper and the flowchart of PBLGLD method is displayed in Fig. 3.

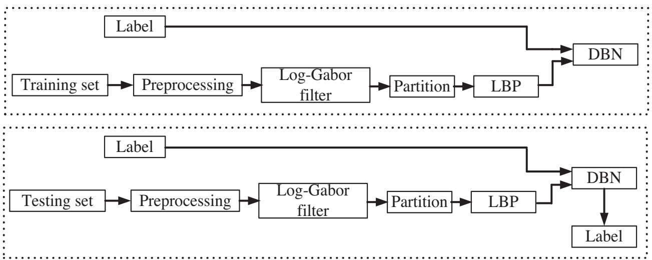

Fig. 3. Flowchart of proposed method.

In this paper, Log-Gabor-LBP and DBN are used to study face recognition under complex illumination conditions, as shown in Fig. 3. The steps of the proposed algorithm are as follows.

(1) Before using Log-Gabor filter to preprocess the image, we first improve the brightness uniformity of image. According to the model of retinal local adaptation [25], the adaptation factors are firstly computed for each pixel by performing a low-pass filter on the input image, as shown in (7).

$$H(p) = I_c(p) * G_F \tag{7}$$

$$G_H(x,y) = e^{-[(x^2+y^2)/2\sigma_F^2]} \tag{8}$$

where p is a pixel in the image; $H(p)$ is the adaptation factor at pixel p; I_c is the intensity of the input image, normalized between [0, 1]; * denotes the convolution operation; G_F is a two dimensional Gaussian filter with spatial constant and we use σ_F=3. The input image I_c is preprocessed by a local nonlinearity operation as follows.

$$I_{cl}(p) = (I_c(\max) + H(p))\frac{I_c(p)}{I_c(p) + H(p)} \tag{9}$$

The term $I_c(\max) + H(p)$ is a normalization factor that ensures that I_c is again scaled in the range of [0, 1]. To improve the preprocessed image with the maximum intensity range, I_{cl} is further normalized by (10).

$$I_{clf} = max_v \frac{I_{cl} - minI_{cl}}{maxI_{cl} - minI_{cl}} \tag{10}$$

where $maxI_{cl}$ and $minI_{cl}$ are the maximum and minimum value of I_{cl}, respectively. max_v is the maximum intensity and is empirically set to 255.

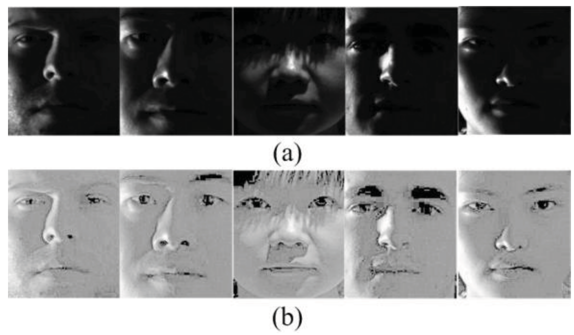

Fig. 4. (a) Original images. (b) Results of the preprocessed image.

(2) After preprocessing the input image, Log-Gabor filtering is applied to the transformed face image to obtain local feature images of different scales and directions.

(3) The database of face images is divided into training set and testing set; each image in training set and testing set is divided into subblocks, and the partition of each image is 4 × 4 in this paper. LBP texture features of each sub block are extracted. The feature of each sub block is connected to form the LBP texture feature of the sample, named as H.

(4) The texture feature vector (H) obtained in step (3) is input into the DBN visible layer. The joint distribution of the DBN visual units and hidden units is shown in

$$\begin{aligned} &p(H, h^{(1)}, h^{(2)}, \ldots\ldots, h^{(l)}) \\ &= P(H|h^{(1)})\,p(h^{(1)}|h^{(2)})\ldots P(h^{(l-1)}|h^{(l)}) \end{aligned} \tag{11}$$

In (11), $h(1)$, $h(2)$, . . . , $h(l)$ are higher level features that are learned layer by layer. The number of hidden layers is set to 2, and the hidden unit number is 200 in the paper. According to (5), the joint distribution of the visual layer and two hidden layers can be obtained by(12).

$$p(H,h^{(1)},h^{(2)})=P(H|h^{(1)})\,p(h^{(1)}|h^{(2)}) \qquad (12)$$

where H is visual layer. $h(1)$ is the first hidden layer and $h(2)$ is the second hidden layer. The active probability of hidden units in the first hidden layer is defined by

$$P(H_i=1|h)=\sigma(b_i+\sum_{i=1}^{num}W_{ij}H_i) \qquad (13)$$

where H_i indicates the visual unit; num is the number of visual units.

(5) In order to optimize the weights W_{ij} for the optimal training network, the DBN iterative algorithm is presented; the iteration number is m. The judgment of the optimal network is based on the fact that the maximum probability function value of the training set is the largest. The maximum probability function is achieved by

$$P=\arg\max_{w} E\left[\sum_{h\in H}\log p(h)\right] \qquad (14)$$

In (14), w is weight matrix; H is the LBP texture feature matrix of training set. The iteration number (m) is set to 3000. According to adjustment, the learning rate is set to be 0.001. After step (5), the optimal network is obtained. The category labels of the testing samples are obtained by classifier in the top of the DBN network.

IV. Results and Discussion

In this section, Yale B+ [26], CMU-PIE [27] databases are selected to evaluate the performance of the method introduced in this paper. Face recognition on those databases is still a challenging task owing to their complex lighting conditions. Yale B+ consists of 38 subjects under 64 lighting conditions. The CMU-PIE includes 68 subjects under 21 illumination conditions with a little hair and background in the face images.

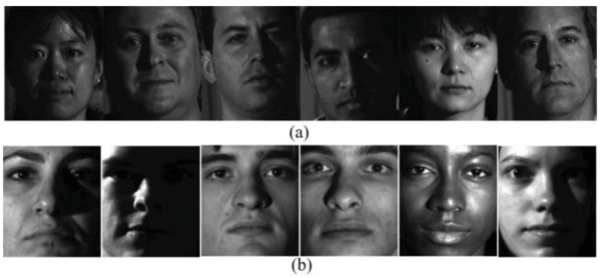

Fig. 5. Face image examples (a) CMU-PIE. (b) Yale B+.

The PBLGLD recognition rate is compared with the recognition rate of the methods based on LBP- DBN, DBN, LBP and SVM, PCA and SVM. The experimental selections of training sets and testing sets on Yale B+, and CMU-PIE are exactly the same as the previous experimental selections. The experiments of the PBLGLD and the LBP - DBN are under the situations with the best partitioning way and the best hidden unit numbers. For these experimental methods, entire face images need to be directly input. The face recognition experiments are performed 20 times by each face recognition method, and the average of recognition rates are shown in Table 1.

TABLE 1: EXPERIMENTAL RESULTS OF DIFFERENT METHODS(%)

Methods	CMU-PIE	Yale B+
PCA+SVM	72.53	85.78
LBP+SVM	80.52	92.16
DBN	84.36	93.61
LBP+DBN	90.12	95.17
PBLGLD	95.80	98.89

As shown in Table 1, for these experimental databases, Yale B+, and CMU-PIE, the proposed PBLGLD method can achieve the highest recognition rate on the face databases. The PBLGLD recognition rate is up to 95.80% on CMU-PIE and 98.89% on Yale B+ database. The proposed method has superior performance in face recognition by comparing with that of other methods.

V. Conclusions

A novel method based on deep learning to solve the adverse impact imposed by illumination variation in the face recognition is proposed in this paper. In order to improve the adverse effects of intense illumination changes on face images, illumination preprocessing is applied. Log-Gabor filter is used to obtain the Log-Gabor feature images of different scales and directions, and LBP features of images are extracted. Then, these texture features are learned by DBN network to complete the classification and the recognition. Experimental results have shown that the proposed method has superior performance in face recognition by comparing with some state-of-the-arts.

References

[1] W. Zhao, R. Chellappa and P. J. Phillips. "Face recognition: A literature survey. " ACM computing surveys (CSUR), Vol.35, No.4, pp. 399-458, 2003.

[2] M. Sharif, K. Ayub and D. Sattar. "Enhanced and fast face recognition by hashing algorithm." Journal of applied research and technology , Vol.10 , No.4 , pp. 607-617, 2012.

[3] L. Ramírez-Valdez and R. Hasimoto-Beltran."3D-facial expression synthesis and its application to face recognition systems. " Journal of applied research and technology, Vol.7 , No.3 , pp. 354-373, 2009.

[4] T. Zhang, B. Fang and Y. Y. Tang. "Topology preserving non-negative matrix factorization for face recognition." IEEE Transactions on Image Processing, Vol.17 , No.4 , pp. 574-584, 2008.

[5] X. Zhao, Z. He, S. Zhang, S. Kaneko, and Y. Satoh. "Robust face recognition using the GAP feature." Pattern Recognition, vol. 46, no. 10, pp. 2647-2657, 2013.

[6] K. C. Lin, X. Wang, and Y. Q. Tan. "Adaptive region enhancement based facial feature localization under complex illumination." Chinese Journal of Scientific Instrument, vol. 35, no. 2, pp. 292-298, 2014.

[7] L. Lei, D. H. Kim, W. J. Park, and S. J. Ko. "Face recognition using LBP Eigenfaces." IEICE Transactions on Information and Systems, vol. 97, no. 7, pp. 1930-1932, 2014.

[8] G. Zhang, J. S. Chen, and J. Wang. "Local information enhanced LBP."Journal of Central South University, vol.20, no. 11, pp. 3150-3155, 2013.

[9] W. S. Li, L. Chen. "The LBP integrating neighbor pixels on face recognition. "Advanced Materials Research, vol.756, pp. 3834-3840, 2013.

[10] W. Yuan, L. Huanhuan, W. Kefeng, and T. Hengqing. "Fusion with layered features of LBP and HOG for face recognition."Journal of Computer-Aided Design & Computer Graphics, vol. 27, no. 4, pp. 640-650, 2015.

[11] L. Wolf, H. T. Hassner, and Y. Taigman. "Effective unconstrained face recognition by combining multiple descriptors and learned background statistics." IEEE transactions on pattern analysis and machine intelligence ,vol. 33, no.10, pp.1978-1990,2011.

[12] M. De Marsico, M. Nappi, D. Riccio, et al. "Robust face recognition for uncontrolled pose and illumination changes. " IEEE transactions on systems, man, and cybernetics: systems, 2013, 43(1): 149-163.

[13] R. Davarzani, S. Mozaffari, and K. Yaghmaie. "Perceptual image hashing using center-symmetric local binary patterns." Multimedia Tools and Applications, vol. 75, no. 8, pp. 4639-4667, 2016.

[14] Y. Huixian, H. Dilong, L. Fan, L. Yang, and L. Zhao. "Face recognition based on bidirectional gradient centersymmetric local binary patterns." Journal of Computer-Aided Design & Computer Graphics, vol. 29, pp. 130-136, 2017.

[15] J. D. Li, Z. X. Chen, and C. Y. Liu. "Low-resolution face recognition based on blocking CS-LBP and weighted PCA. " International Journal of Pattern Recognition and Artificial Intelligence, vol. 30, no. 8, article ID. 1656005, 2016.

[16] Bengio, Yoshua, and O. Delalleau. "On the expressive power of deep architectures." International Conference on Algorithmic Learning Theory. Springer, Berlin, Heidelberg, 2011.

[17] Taylor, W. Graham, G. E. Hinton, and S.T. Roweis. "Modeling human motion using binary latent variables." Advances in neural information processing systems. 2007.

[18] Arel, Itamar, D. C. Rose, and T. P. Karnowski. "Deep machine learning-a new frontier in artificial intelligence research." IEEE computational intelligence magazine 5.4 (2010): 13-18.

[19] S. F. Liang, Y. H. Liu, and L. C. Li. "Face recognition under unconstrained based on LBP and deep learning. "Journal on Communications, vol. 35, no. 6, pp. 154-160, 2014.

[20] W. Zhang and W. Wang. "Face recognition based on local binary pattern and deep learning." Journal of Computer Applications, vol. 35, no. 5, pp. 1474-1478, 2015.

[21] X. Ma and Q. Sang. "Handwritten signature verification algorithm Based on LBP and deep learning. " Chinese Journal of Quantum Electronics, vol. 34, no. 1, pp. 23-31, 2017.

[22] P. Zhong, Z. Gong, S. Li, and C. B. Schonlieb. "Learning to diversify deep belief network for hyperspectral image classification. " IEEE Transactions on Geoscience and Remote Sensing, vol. 55, no. 6, pp. 3516-3530, 2017.

[23] Y. Chen, X. Zhao, and X. Jia. "Spectral-spatial classification of hyperspectral data based on deep belief network. " IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 8, no. 6, pp. 2381-2392, 2015.

[24] L. Chang, X. M. Deng, , et al. "Convolutional neural networks in image understanding. " Acta Automatica Sinica, vol. 42, no. 9, pp. 1300-1312, 2016.

[25] Meylan, Laurence, D. Alleysson, and S. Süsstrunk. "Model of retinal local adaptation for the tone mapping of color filter array images." JOSA A 24.9 (2007): 2807-2816.

[26] K.C. Lee. The Extended Yale Face Database B , http://vision.ucsd.edu/~iskwak/ExtYaleDatabase/ExtYaleB.html

[27] T. Sim, S. Baker, and M. Bsat. 'The CMU pose, illumination, and expression database, IEEE Trans', Pattern Anal. Mach. Intell, 2003,25, pp. 1615–1618.