

Research and Application on Face Recognition Algorithm Based on FWLD Method and Deep Belief Nets

Yihong Zhang^{1, 2, a}, Xiang Li^{1, b}, Zhijie Wang^{1, 2, c}

¹ College of Information Science & Technology, DongHua University, Shanghai 201620, China;

² Engineering Research Center of Digital textile and garment technology, Shanghai 201620, China.

^a zhangyh@dhu.edu.cn, ^b lix930@163.com, ^c zjwang@dhu.edu.cn

Keywords: Weber local descriptor; deep belief nets; fuzzy logic; local binary pattern.

Abstract. Aiming at the problem of insufficient feature extraction and sensitive to noise for traditional face recognition algorithms, a face recognition algorithm based on improved Weber local descriptor and deep belief nets is proposed. First analyze the shortcomings of Weber local descriptor, and based on the fuzzy logic, an improved FWLD face description method is proposed. To make full use of the domain pixels in WLD direction component, based on introduction of LBP coding, the fuzzy logic is used to optimize the LBP coding, and then the improved WLD algorithm is used as the input of the deep belief network. And the network parameters are obtained through the layer-by-layer greedy pre-training network. Finally, the BP neural network is used to fine tune and optimize the DBN network, and the test samples are predicted by the trained network. In the ORL data set, the correct rate is 95%. The simulation results show that the face recognition algorithm proposed in this paper is higher in recognition rate and more robust than the traditional recognition.

1. Introduction

Face recognition technology, as an important part of pattern recognition and computer vision, is concerned by more and more researchers. At the beginning of the study, the researchers proposed some methods, including the geometric feature based method, the model-based method and the statistics-based method, which is better than other methods, such as Principal Component Analysis (PCA) [1], and Linear Discriminant Analysis (LDA) [2], etc. Under the limiting conditions, there are always better recognition effects, but under the non-limiting conditions, such as changes in light, attitude changes, and occlusion, the recognition rate drops sharply. In response to this problem, people are increasingly studying local features. In the literature [3], local Binary Pattern (LBP) is proposed. The LBP is featured by simple calculation, invariant rotation and constant gray scale, but the description ability is insufficient and is susceptible to noise interference. In literature [4], (Weber Local Descriptor, WLD) is proposed, unlike the LBP method, the WLD feature consists of the differential excitation component and the direction vector. The WLD differential excitation component counts the difference between the neighborhood pixel and the central pixel, and then divides the center pixel, which further improves the WLD robustness. In literature [5], WLD is proved to be insensitive to light and convenient to calculate. In literature [8], WLD is used for face recognition, and good experimental results are achieved. In literature [9], a multi-scale WLD algorithm is proposed to further increase the robustness of WLD. In literature [10], a rotation-invariant WLD algorithm is proposed to enhance its rotation invariance by introducing angle calculation on the direction component. In literature [11], Gabor and WLD combined algorithm is proposed, and successfully applied in the iris recognition.

Different from the current improved method, the algorithm in this paper introduces the fuzzy logic, and fuzzy LBP coding is conducted on the WLD direction component. This not only makes full use of the domain pixels on the direction component and improves the description ability of WLD, but also improves the robustness of WLD by fuzzy logic. Experiments show that the algorithm proposed in this paper has stronger description ability and robustness than traditional WLD.

2. Weber Local Descriptor

The Weber Law shows that the ratio of the amount of change ΔI of the perceived signal to the size of the original signal value I is a constant, that is $\Delta I / I$, which is called the Weber Rate, indicating that the ratio of the stimulus intensity changes reflects the perceived stimulus amplitude. Inspired by this law, Chen Jie et al. [4] proposed Weber Local Descriptor (WLD) to describe the texture information of the image. He believed that the features extracted by this method are more in line with the human perception system, and have a better robustness to the light and noise.

The Weber local descriptor introduces two components to describe the texture features of the image, which respectively are the differential excitation component and the direction component.

2.1 Differential Excitation Component

Human perception is simulated to find significant texture changes in the image, and the differential excitation component is the ratio of the sum of the difference between the current pixel and the domain pixel to this pixel. The differential excitation is described as follows:

X_0	X_1	X_2
X_7	X_c	X_3
X_6	X_5	X_4

Fig. 1 3*3 image block

In Fig.1, X_c is the gray value of the current pixel, and $X_i (i = 0,1,2, \dots, n-1)$ is the gray value of n field pixels, where $n = 8$. The differential excitation component $\xi(X_c)$ can be calculated by the following formula:

$$\xi(X_c) = \arctan(v_s^{00} / v_s^{01}) \tag{1}$$

$$v_s^{00} = \sum_{i=0}^{n-1} (X_i - X_c) \tag{2}$$

$$v_s^{01} = X_c \tag{3}$$

The differential excitation component $\xi(X_c)$ is valued as $(-\pi/2, \pi/2)$. When $\xi(X_c)$ is a positive, it can be considered that the gray value of the domain pixel is larger than that of the central pixel, that is, the center pixel is darker. When $\xi(X_c)$ is negative, the gray value of the domain pixel is smaller than that of the center pixel, that is, the center pixel is brighter.

2.2 Direction Component.

The so-called direction component is the gradient direction of a certain pixel in the image relative to the neighborhood pixel, which can be expressed by the formula (5).

$$\theta(X_c) = \arctan\left(\frac{v_s^{11}}{v_s^{10}}\right) \tag{4}$$

$$v_s^{11} = X_7 - X_3, v_s^{10} = X_5 - X_1 \tag{5}$$

Where $\theta(X_c)$ is the direction value, and the valuing range is $(-\pi/2, \pi/2)$. To simplify and obtain more texture information, $\theta(X_c)$ can be mapped to $\theta'(X_c)$, ie.

$$\theta'(X_c) = \arctan(v_s^{11}, v_s^{10}) + \pi \tag{6}$$

$$\theta'(X_c) = \begin{cases} \theta(X_c) & v_s^{11} > 0, v_s^{10} > 0 \\ \pi + \theta(X_c) & v_s^{11} > 0, v_s^{10} < 0 \\ \theta(X_c) - \pi & v_s^{11} < 0, v_s^{10} > 0 \\ \theta(X_c) & v_s^{11} < 0, v_s^{10} < 0 \end{cases} \tag{7}$$

The valuing range of the mapped $\theta'(X_c)$ is $[0, 2\pi]$, which makes the direction component more intuitive and easy to calculate. The quantized equation is as follows:

$$\Phi_t = f_q \theta'(X_c) = \frac{2t}{T} \pi \tag{8}$$

3. The Method Proposed in This Paper

According to WLD direction component calculation method, it is only simple and linear addition and subtraction of the upper and lower, left and right pixels, which are susceptible to noise interference and do not take into account the 8 pixels of the neighborhood. In order to improve the description ability of texture features, combined the features of WLD algorithm, this paper introduces the fuzzy logic theory and proposes FWLD algorithm.

To enhance the description ability of traditional WLD, considering the 8 neighborhood pixels of the central pixel, the LBP algorithm is introduced to encode the image. In the pixel i_c and its domain i_p , the LBP is encoded as $z_p = i_p - i_c$, and each LBP pixel can be expressed as:

$$b_p = \begin{cases} 1 & \text{if } z_p \geq 0, \\ 0 & \text{if } z_p < 0. \end{cases} \tag{9}$$

LBP is only insensitive to the gray scale of the image, but some shadows or some particularly small noise can cause the pixel coding to be completely different. As shown in Fig. 1. Ideally, the area shown in Fig. 1 shall be coded as "11111111", but as the subtle noise, the code has become "01010111".

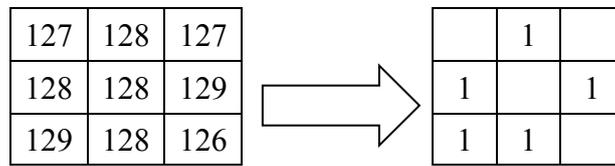


Fig. 2 LBP calculation diagram

On the above basis, fuzzy logic is introduced, and the problem of sensitive to noise is solved by fuzzification of LBP coding process. The calculation method of the Formula (9) is replaced by establishing a membership function to denote the probability that z_p is encoded as "0" or "1". We choose the most commonly used linear membership function:

$$f_{1,d}(z_p) = \begin{cases} 0 & \text{if } z_p < -d, \\ 0.5(1 + \frac{z_p}{d}) & \text{if } -d \leq z_p \leq d, \\ 1 & \text{if } z_p > d. \end{cases} \tag{10}$$

Where $f_{1,d}(z_p)$ and $f_{0,d}(z_p)$ respectively represent the probability that the pixels are encoded as "1" and "0". The parameter d represents the fuzzy quantity.

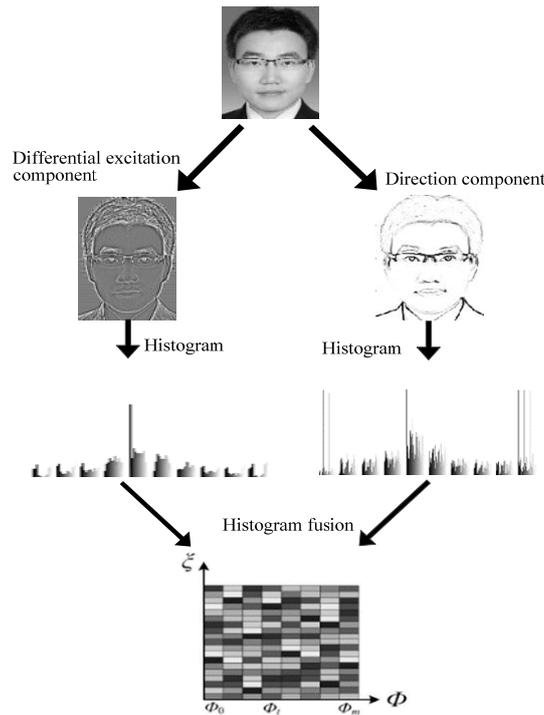


Fig. 3 FWLD feature calculation process

Fig.3 illustrated the establishment process of FWLD feature calculation and experiment results.

4. Establishment of Recognition Model

Deep Belief Nets (DBNs) uses an unsupervised greedy layer-by-layer training approach to solve difficult problems in deep neural networks. DBNs is a generalization of BP neural network. The hidden layer is composed of many non-linear structural units, and its structural elements generally are Restricted Boltzmann Machine.

The Restricted Boltzmann Machine (RBM) consists of visible and implicit elements, and both the visible and implicit elements are binary variables, and the structure is shown in Fig. 4.

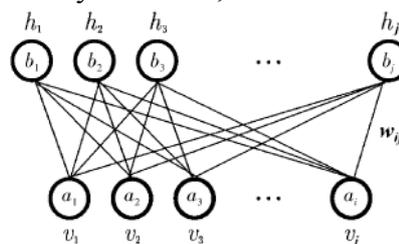


Fig. 4 Structure chart of Restricted Boltzmann Machine

RBM is a model based on the energy function, each layer of RBM neurons are only in two states, activated and suppressed. For any given set of (v, h) , define $E(v, h)$ as a whole energy function of a layer of RBM.

$$E(v, h) = -\sum_i a_i v_i - \sum_i b_i h_i - \sum_i \sum_j v_i w_{i,j} h_j \tag{11}$$

Where v_i is the state of the i -th neuron of the visible layer, h_j is the state of the j -th neuron of the hidden layer, a_i is the bias of the i -th neuron in the visible layer, b_j is the bias of the i -th neuron in the visible layer, $w_{i,j}$ is the weight between the i -th neuron of the visible layer and the j -th neuron of the hidden layer.

RBM is an unsupervised learning method, and pre-training can obtain a good initial network value, so the last layer of the DBN model needs to use BP neural network and marked data to fine tune the

overall network parameters. The pre-training process can be regarded as the initialization process of parameters obtained in RBM training to the BP network. It can overcome the shortcomings of long training time and easy to fall into the local optimum caused as random initialization of BP network.

5. Experimental Results

To validate the algorithm proposed in this paper, the ORL face database is used for face recognition experiments. At the same time, to further evaluate the algorithm, experiments are carried out on the face database Yale and Yale-B under the limiting condition. The experiment in this paper is carried out under the following hardware configuration environment: windows7 64-bit system, CPU Core i5-4590 @ 3.30GHz, and 8G of memory.

ORL face database contains 400 pictures of 40 people in a variety of lights, expressions and gestures, and the size of each picture is 112*92 pixels. We take each people's first 7 pictures as training samples, and the remaining are used as test samples. Because the calculated amount of DBN model is relatively large, if the face image is directly used as the input training model, it will consume a lot of time. In order to reduce the data dimension, the images are reduced in dimension and the FWLD features are extracted to test the recognition performance of the algorithm in this paper under different resolutions. The images are respectively down-sampled in 2*2, 4*4, 6*6 and 8*8 to constitute the face data in four different resolutions, 56*46 , 28*23 , 19*16 and 14*12 . In the algorithm of this paper, DBN model of three-layer RBM is adopted, and the model structure is 2576 – 500 – 300 – 200 – 40 . The last layer of the model is 200-40 BP neural network, and 3000 times of loop iterations are done for each layer. The learning rate is 0.01.

First of all, conduct training and recognition intensively in different sizes of dimensions, and repeat it for 10 times to take the average recognition rate. It could be seen from the experimental results shown in the figure 6, when the image dimension is small, the recognition rate of each algorithm is not good. But when the dimension increases, we can see that the effect in FWLD+DBN method is better.

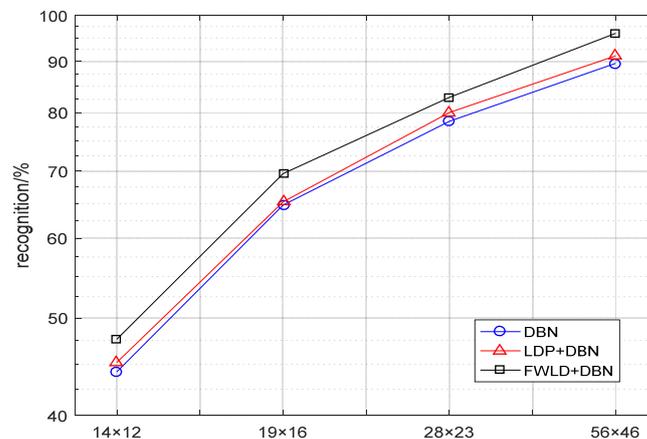


Fig. 6 The highest recognition rate of different algorithms on the ORL face database

6. Summary

In this paper, a face recognition method based on improved WLD feature and DBN is proposed. After analyzing the shortcomings of WLD feature, fuzzy LBP coding id introduced on the direction component to further enhance its description ability, which improves the description ability and robustness of WLD to local feature. Combined with the deep learning model DBN, DBN can learn more local shallow feature on the basis of learning overall deep features. The experiments on ORL and YALE-B databases show that the algorithms proposed in this paper have a good recognition rate in expression, illumination change and occlusion. The experimental results show that the improved FWLD feature and the deep belief nets can extract more texture features, have higher accuracy and

stronger robustness. However, combined with the deep learning model, the model parameters must be manually set and training time is too long, so how to quickly find the appropriate parameters and reduce training time will be the focus of the further study.

Acknowledgements

This research was financially supported by the ShangHai Science Foundation-“The Research and Application of Intelligent Garment Defect Detecting based on Cloud Computing”, Funding Number: 14ZR1400400. This research was financially supported by the China central university foundation. Project number [15D110406].

References

- [1]. Belhumeur P N, Hespanha J P, Vs K M E. Fisherfaces: Recognition using class specific linear projection[M]// Computer Vision — ECCV '96. Springer Berlin Heidelberg, 1996:43-58.
- [2]. Ling Y, Yin X, Bhandarkar S M. Sirface vs. Fisherface: Recognition using class specific linear projection[J]. 2003, 3:III-885-8 vol.2.
- [3]. Pietikäinen M, Ojala T, Xu Z. Rotation-invariant texture classification using feature distributions[J]. Pattern Recognition, 2000, 33(1):43-52.
- [4]. Heikkilä M, Pietikäinen M, Schmid C. Description of interest regions with local binary patterns[J]. Pattern Recognition, 2009, 42(3):425-436.
- [5]. Bui H N, Na I S, Kim S H. De-Noising Model for Weberface-Based and Max-Filter-Based Illumination Invariant Face Recognition[M]// Ubiquitous Information Technologies and Applications. Springer Berlin Heidelberg, 2014:373-380.
- [6]. Li S, Gong D, Yuan Y. Face recognition using Weber local descriptors[J]. Neurocomputing, 2013, 122:272-283.
- [7]. Muhammad G, Hussain M, Alenezzy F, et al. Race recognition using local descriptors[C]// IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 2012:1525-1528.
- [8]. Hussain M, Muhammad G, Saleh S Q, et al. Copy-Move Image Forgery Detection Using Multi-Resolution Weber Descriptors[C]// Eighth International Conference on Signal Image Technology and Internet Based Systems. IEEE Computer Society, 2012:395-401.
- [9]. Banerjee A, Das N, Nasipuri M. Texture Classification Using Deep Neural Network Based on Rotation Invariant Weber Local Descriptor[C]// International Conference on Recent Trends in Image Processing and Pattern Recognition. Springer, Singapore, 2016:277-292.
- [10]. Li J, Sang N, Gao C. Log-Gabor Weber Descriptor for Face Recognition[J]. Journal of Electronic Imaging, 2014, 24(5):053014.
- [11]. Zhang Z, Wang L, Zhu Q, et al. Pose-invariant face recognition using facial landmarks and Weber local descriptor[J]. Knowledge-Based Systems, 2015, 84(C):78-88.