A novel fusion method of improved adaptive LTP and two-directional two-dimensional PCA for face feature extraction*

LUO Yuan (罗元)[1], **WANG Bo-yu** (王薄宇)[1]**, **ZHANG Yi** (张毅)[2], **and ZHAO Li-ming** (赵立明)[2]

1. *Key Laboratory of Optoelectronic Information Sensing and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China*
2. *Engineering Research Center for Information Accessibility and Service Robots, Chongqing University of Posts and Telecommunications, Chongqing 400065, China*

(Received 20 October 2017; Revised 26 December 2017)

In this paper, under different illuminations and random noises, focusing on the local texture feature's defects of a face image that cannot be completely described because the threshold of local ternary pattern (LTP) cannot be calculated adaptively, a local three-value model of improved adaptive local ternary pattern (IALTP) is proposed. Firstly, the difference function between the center pixel and the neighborhood pixel weight is established to obtain the statistical characteristics of the central pixel and the neighborhood pixel. Secondly, the adaptively gradient descent iterative function is established to calculate the difference coefficient which is defined to be the threshold of the IALTP operator. Finally, the mean and standard deviation of the pixel weight of the local region are used as the coding mode of IALTP. In order to reflect the overall properties of the face and reduce the dimension of features, the two-directional two-dimensional PCA ($(2D)^2$PCA) is adopted. The IALTP is used to extract local texture features of eyes and mouth area. After combining the global features and local features, the fusion features (IALTP+) are obtained. The experimental results on the Extended Yale B and AR standard face databases indicate that under different illuminations and random noises, the algorithm proposed in this paper is more robust than others, and the feature's dimension is smaller. The shortest running time reaches 0.329 6 s, and the highest recognition rate reaches 97.39%.

Document code: A **Article ID:** 1673-1905(2018)02-0143-5

DOI https://doi.org/10.1007/s11801-018-7226-7

With the development of artificial intelligence technology, biometrics has become a very popular subject in the field of artificial intelligence and pattern recognition[1], where face recognition is the most representative in terms of practicality and wide applicability. Usually the face recognition process consists of two consecutive stages: The first stage is the face recognition used to extract feature information from a set of training images; The second stage is to extract the features by the classifier to identify the subject of the query image[2]. And facial feature extraction is the most critical step, which directly affects the effect and recognition rate of face feature classification recognition[3]. Since the local ternary pattern (LTP) operator uses the ternary mode coding to extract the facial feature, the dimension of LTP feature is too high. Secondly, the optimal threshold for LTP is obtained through a large number of experiments, and the threshold cannot adapt to complex environment. Besides, the above researches can only improve the recognition rate. The robustness and computational complexity under different illuminations and random noises need to be further studied.

In this paper, a novel fusion method of improved LTP and two-dimensional two-directional PCA for face feature extraction is proposed to improve the problem under different illuminations and random noises. Through the statistical characteristics of the central pixel and the neighborhood pixel we can get the description of each local neighborhood. The adaptively gradient descent iterative function can compute the new threshold automatically from one patch of pixels to another during the whole image based on the pixels' values in these patches. We use the new threshold in the definition of improved adaptive local ternary pattern (IALTP). In this way, a new way to calculate the threshold for LTP automatically with different illuminations and random noises is proposed. The two-directional two-dimensional PCA ($(2D)^2$PCA) can not only obtain the global feature but

* This work has been supported by the National Natural Science Foundation of China (No.51604056), and the Chongqing Science and Technology Commission (No.cstc2015jcyjBX0066).

** E-mail: 892105468@qq.com

also reduce the dimension of IALTP feature.

Obviously, there is no fixed threshold or the best threshold to adapt to different kinds of face images. So the best solution for this problem is to find a way to compute the threshold automatically based on different kinds of face images. Thus, we establish the statistical characteristics for each pixel in a neighborhood around a central one. The adaptive iterative gradient descent function is established to calculate the weight coefficient so that we can obtain the new threshold. This threshold changes automatically from one patch of pixels to another during the whole image based on the pixels' values in these patches.

Besides, the relationship between center pixel and neighborhood pixel can be regarded as a multivariate linear regression problem. Thus, we can use the gradient descent to solve the problem of multivariable linear regression.

Inspired by the cost function of multivariate linear regression model, the difference between the center pixel and the neighborhood pixel weights can be calculated as

$$J(\theta_s)=-\frac{1}{2}\left[P_c(i,j)-\sum_{s=1}^{M}\theta_s P_s(i,j)\right]^2 , \tag{1}$$

where $P_c(i, j)$ is the center pixel, $P_s(i, j)$ is the neighborhood pixel, θ_s is a weight coefficient of the arbitrary neighborhood pixel, and $\theta_s=[\theta_1, \theta_2,... \theta_M]$, $\sum_{s=1}^{M}\theta_s=1$. This equation minimizes the overall difference between the center pixel and the neighborhood pixel, if the above formula on the θ_s derivative can be obtained as

$$\frac{\partial J}{\partial \theta_s}=P_s(i,j)\left[P_c(i,j)-\sum_{s=1}^{M}\theta_s P_s(i,j)\right]=0 , \tag{2}$$

where M=1, 2,…8, because of the arbitrary neighborhood pixel.

The center pixel value in the above equation can be expressed in the sum of products of neighborhood pixel value and weights, and also can be represented as

$$P_c(i,j)=\theta_m P_m(i,j)+\sum_{\substack{s=1\\ s\neq M}}^{M}\theta_s P_s(i,j) . \tag{3}$$

According to the above equation, the weight of a neighborhood pixel value can be obtained as

$$\theta_m=\frac{P_c(i,j)-\sum_{\substack{s=1\\ s\neq M}}^{M}\theta_s P_s(i,j)}{P_m(i,j)} . \tag{4}$$

The gradient descent method is an error criterion function to minimize sample. Updating the learning parameters by giving the parameter values in advance, the structure of each updating can make the loss function smaller and finally reach the minimum. That is, we can get the gradient direction by the weight of neighborhood pixel value so that the direction of the decline in the most changes can be found. The weight of a neighborhood pixel value θ_m is as follows:

$$\theta_m:=\theta_m-\lambda\frac{\partial J}{\partial \theta_s} , \tag{5}$$

where λ is learning rate, and we can achieve the desired result when λ=1.

We can calculate the weights of different pixels according to the following steps:

Step1: Initialize $\theta_m=\frac{1}{M}$ for every m=1, 2,...M;

Step2: Use the update equation;

Step3: Loop for m=1 to M; $\theta_m:=\theta_m-\lambda\frac{\partial J}{\partial \theta_s}$;

Step4: Update θ_m with the new value. When computing the new value of θ_m for any pixel, use this value to compute θ_m for the next pixel.

When the number of local pixels is small, the optimization process is difficult to converge; when the number of local pixels is large, the resolution of the image is poor. In the experiment we find that when the number of local pixels is 8, although the resolution is poor, we can not only make the optimization process converge, but also obtain the optimal threshold by adjusting the parameters.

The mean and standard deviation of the pixel weight of the local region are used as the coding mode of the three modes of IALTP. Through the above steps we can get the weight coefficients of all neighborhood pixels, and define the formula of IALTP as

$$IALTP=\begin{cases}1 & \text{if} \quad \theta_m P_m(i,j)\geq\varepsilon+\omega\delta \\ 0 & \text{if}\ \varepsilon-\omega\delta<\theta_m P_m(i,j)<\varepsilon+\omega\delta \\ -1 & \text{if} \quad \theta_m P_m(i,j)\leq\varepsilon-\omega\delta\end{cases} , \tag{6}$$

where ε is the mean weight of local pixel, and δ is the standard deviation of local pixel's weight. According to the experiments, we can achieve the desired result when ω=0.6 and $\theta_m P_m(i,j)\geq\varepsilon+\omega\delta$.

The IALTP is applied to the processing of face image pixel information according to the following steps:

Step1: The face image is divided into non-overlapping sub-regions J_0, J_1, J_2,...J_{t-1}, where t is the number of non-overlapping regions;

Step2: Count the values of center pixel and neighborhood pixel within the pixel range, P is set as the center pixel and R is the neighborhood radius pixel (P, R);

Step3: Calculate θ_m for different pixels according to Eq.(5);

Step4: Substitute $\theta_m P_m(i, j)$ into formula IALTP, and finally we can get the IALTP operator.

After obtaining the adaptive weights of different pixels, we can define the contrast interval as [$\varepsilon-\omega\delta$, $\varepsilon+\omega\delta$], and the schematic diagram of IALTP operator is shown in Fig.1.

In order to simplify the calculation, each ternary pattern is further split into negative and positive parts. The two parts are processed as two separate channels of local binary pattern (LBP) descriptors (in Fig.2). The histograms of positive and negative channels are concatenated as the final feature description of the original image. Fig.3 shows the IALTP feature extraction process.

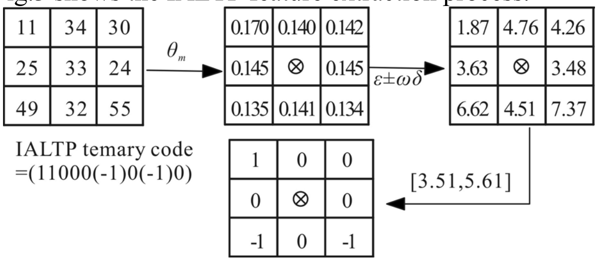

Fig.1 The schematic diagram of IALTP operator

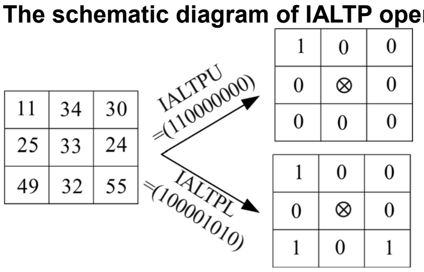

Fig.2 The example of splitting LTP code into positive and negative codes

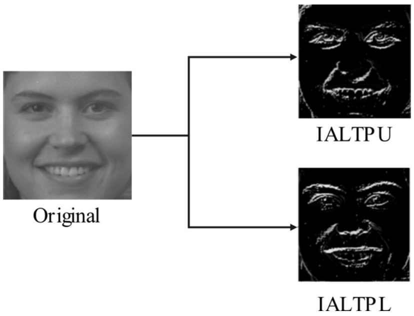

Fig.3 The IALTP feature extraction process

A novel fusion method of IALTP and $(2D)^2$PCA for face recognition is shown in Fig.4. First of all, the original face image is preprocessed. The experimental results show that if the histogram of the whole face image is extracted directly, the validity of the feature extraction will be affected, and the recognition rate will be reduced. Therefore, this paper will divide the face image into 8×8 blocks with small sizes, so that more effective feature information can be obtained.

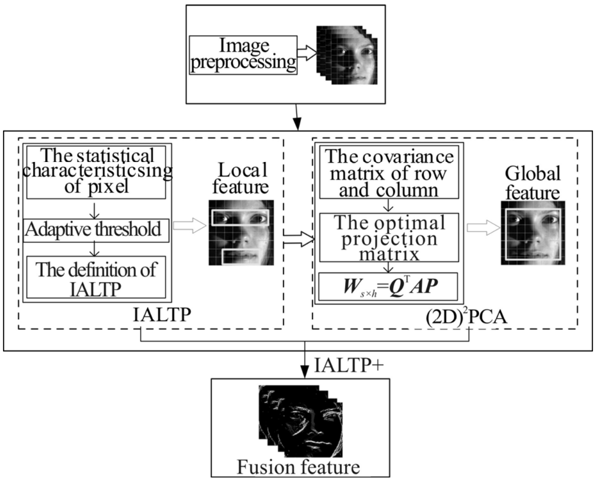

Fig.4 Fusion of IALTP and $(2D)^2$PCA

The IALTP algorithm is used to extract the local texture features of the face. Set the eye and mouth as the area-of-interest (ROI), then scan the ROI with IALTP operator to obtain the IALTP encoding image. Lastly, preserve the data as feature vectors from left to right, top to bottom. Combining global feature vectors with local feature vectors, the fusion feature vectors are obtained, and the support vector machine (SVM) is applied to classify the fusion features and recognize face image.

The experiment is designed based on the development tools: windows10+MATLAB 2016a, Intel Code i5, GPU: GTX960. The performance evaluation is carried out by comparing several best performance face characterization operators, including LBP[4], LTP[5], CS-LBP[6], relaxed local ternary pattern (RLTP)[7], optimized local ternary pattern (OLTP)[8], IALTP and $(2D)^2$PCA (IALTP+). In order to ensure the objectivity of the experiment, all the above algorithms are based on image blocks of 8×8, and the neighborhood (P, R) takes (8, 2). The Euclidean distance is used as the feature vector similarity measure.

The experiment is divided into three parts. The first part is comparing the calculation time of each algorithm and dimensions of extracting features. The second part is verifying the robustness of the proposed algorithm to the illumination. The third part is verifying the anti-noise performance of the algorithm.

The Extended Yale B face database comes from 64 individual face images captured by 38 persons under different lighting conditions. After processing, the face image size is 192×168. According to the illumination intensity changes, the image library is divided into five subsets, from set1 to set5. The illumination condition of set1~set5 is getting worse and worse. The AR face database comes from more than 4 000 face images of 126 persons, including 70 men and 56 women. After processing, the face image size is 100×100. The images have a positive view of different facial expressions, lighting conditions, and occlusion conditions (sunglasses and scarves).

In this experiment, the set1 of the Extended Yale B face database is used as the training group for the face image, and the remaining sets are used as the test groups. A human face image of each person in the AR face database is used as a training set, and the rest of the images are used as test sets.

The experimental comparison results of dimension and computation time are given in Tab.1. It's clear that the dimensions of the feature vectors for IALTP+ operator are smaller than those of others. The reason is that LBP, LTP and their derivative algorithms calculate the dimension by selecting the appropriate color range (the number of histogram bins), but the dimension of the feature is still so high. A simple strategy is to use $(2D)^2$PCA for further dimensional reduction. By projecting the high dimensional feature matrix into low dimensional feature matrix, we can obtain the low dimensional data.

Tab.1 Dimension and computation time of face image

Method	Extended Yale B		AR	
	Dimension	Time (s)	Dimension	Time(s)
LBP	59×64	0.451 2	59×24	0.946 3
CS-LBP	24×64	0.395 8	24×24	0.792 3
LTP	118×64	0.643 7	118×24	0.713 5
OLTP	86×64	0.398 8	86×24	0.648 3
RLTP	90×64	0.453 6	90×24	0.775 2
IALTP+	22×48	0.329 6	20×12	0.601 9

As shown, the computation time of IALTP+ operator for a complete face recognition is 329.6 ms which is less than that of CS-LBP by about 66 ms, and it is 1.93 times faster than LTP in the Extended Yale B database. At the same time, the computation time of IALTP operator for a complete face recognition is 46.4 ms less than that of CS-LBP. The experimental results show that the proposed IALTP+ not only greatly shortens the computation time of face feature extraction but also reduces the dimension of the eigenvector.

In order to verify the robustness of these algorithms to illumination, LBP, CS-LBP, LTP, OLTP, RLTP, IALTP+ operators are used for test in the extended Yale B and AR face databases, respectively. The first set of experiments uses the sufficient set of light sources in the Extended Yale B face database as the training set and the rest as the test set. The second set of experiments uses the set4 with poor light conditions as the training set and the remaining as the test set. The third set of experiments selects the set3 as the training set and the rest as the test set. The experimental results are shown in Fig.5, Fig.6 and Fig.7.

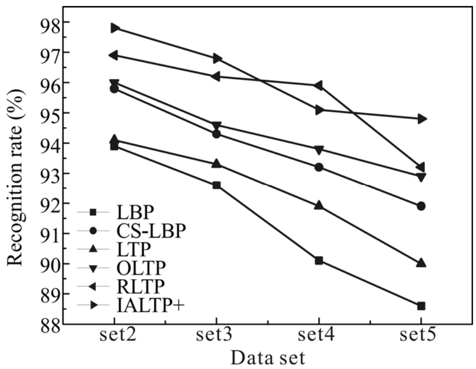

Fig.5 Recognition rates with set 1 from Extended Yale B database as the training samples

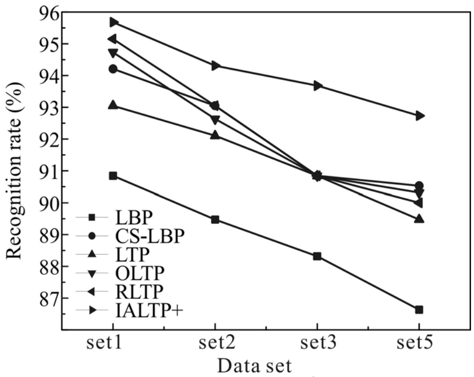

Fig.6 Recognition rates with set 4 from Extended Yale B database as the training samples

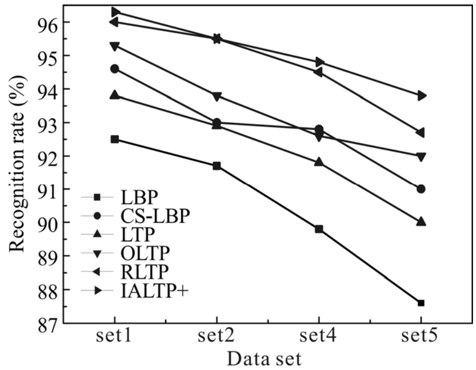

Fig.7 Recognition rates with set 3 from Extended Yale B database as the training samples

In the fourth set of experiments, a face image of each person in the AR human face database is selected as the training set, and the rest of the images are used as the test set. The experimental results are shown in Tab.2.

Tab.2 The face recognition rates in AR database

Number of training images	Number of testing images	Texture model	Recognition rate (%)
20	100	LBP	82.35
		CS-LBP	84.90
		LTP	86.91
		OLTP	89.34
		RLTP	89.98
		IALTP+	91.35

The results of the four sets of experiments show that the accuracy of IALTP+ for face recognition is higher than that of other algorithms. When the illumination is very sufficient, all operators can get high recognition rates and IALTP+ reaches the highest recognition rate. When the illumination is poor, the recognition rates of LBP, CS-LBP and LTP are decreased significantly, and the illumination robustness is poor. LBP cannot describe the face image of nonlinear illumination very well, and with poorer robustness. CS-LBP enhances the description of local texture features based on the LBP. Due to the adoption of the three-value model, LTP has a robustness to illumination changes with the appropriate threshold, but the recognition rate is not high. In contrast, IALTP+ is better than others. We can see that IALTP+ still has high recognition rate and robustness when the illumination is poor from Fig.3. The reason is that the IALTP+ operator considers the differences of illumination between neighborhood pixels, by the adaptive threshold to minimize the difference between center pixel and neighborhood pixel, so that it can solve the difference caused by the sample face, and the nonlinearity illumination which leads to the recognition rate reduction.

The experimental comparison results of the original images and the noisy images are given in Tab.3 and Tab.4. The Gaussian white noise is used as the noise model of this experiment, where λ is defined as the normalized standard deviation of noise. The noise intensity is simulated by adjusting the noise level, and then the anti-noise performance can be tested.

Tab.3 The face recognition rates in Extended Yale B database (%)

Method	Original image	Noise image		
		λ=0.02	λ=0.06	λ=0.1
LBP	91.59	80.53	75.56	47.63
CS-LBP	92.63	89.23	85.50	76.67
LTP	95.08	86.85	75.04	58.45
OLTP	96.07	89.69	87.35	78.89
RLTP	95.98	90.6	88.36	79.63
IALTP+	97.39	95.49	92.03	90.23

Tab.4 The face recognition rates in AR database (%)

Method	Original image	Noise image		
		λ=0.02	λ=0.06	λ=0.1
LBP	90.35	80.02	72.32	39.37
CS-LBP	91.88	85.35	80.47	70.02
LTP	91.36	84.32	72.30	47.39
OLTP	94.30	88.69	80.06	69.62
RLTP	93.87	91.35	82.67	70.39
IALTP+	95.38	93.12	90.83	88.64

Through the analysis of the above two sets of data, we can find that the recognition performance of the IALTP+ operator is better than that of other algorithms under noise environment. With the increase of the noise density, the recognition rate of the IALTP + operator is only reduced in a small range. When the noise level reaches λ=0.1, the recognition rate of the IALTP+ operator is higher than that of RLTP by 18.25%, and we can see that the IALTP+ has great anti-noise performance. The LBP operator has poor performance under noise, because it uses the contrast between the center pixel and the neighborhood pixel for coding. For the improved CS-LBP operator, the robustness to noise is enhanced by the contrast between pixels based on the center pixel symmetry. For IALTP+, the thresholds are obtained by finding the optimal quantization length of the center pixel and the neighborhood pixel, which greatly enhances the anti-noise performance. The algorithm proposed in this paper minimizes the difference between the pixel value of the sample and the center pixel value by adjusting threshold, and the $(2D)^2$PCA makes it have great robustness to illumination and noise.

In this paper, a novel fusion method of IALTP and $(2D)^2$PCA for face feature extraction is proposed. The results of comparison experiments in the Extended Yale B and AR face databases show that compared with current methods, under different illuminations and random noises, IALTP+ has higher face recognition, less computation time and low dimensional feature matrix. But we find that with the increasing number of face images, the learning rate of gradient descent will result in divergence. Thus, choosing the optimal learning rate to deal with a large number of face images will be further studied in the subsequent works to enhance the recognition rate.

References

[1] Hua G, Yang M H and Learnedmiller E, IEEE Transactions on Pattern Analysis & Machine Intelligence **33**, 1921 (2011).

[2] Mashhoori A and Jahromi M Z, Neurocomputing **108**, 111 (2013).

[3] Tasnim Tarannum and Anwesha Paul, Human Expression Recognition based on Facial Features, IEEE Computer Society Conference on Electronics and Vision, 990 (2016).

[4] Lei Z, Ahonen T and Pietikäinen M, Local Frequency Descriptor for Low-resolution Face Recognition, IEEE International Conference on Automatic Face & Gesture Recognition and Workshops, 161 (2011).

[5] Ibrahim M, Efat A and Kayesh H, Dynamic Local Ternary Pattern for Face Recognition and Verification, International Conference on Computer Engineering and Applications, 2014.

[6] Ruby A U and Chandran J G C, A Theoretical Approach on Face Recognition with Single Sample per Class using CS-LBP and Gabor Magnitude and Phase, 2016.

[7] Ren J, Jiang X and Yuan J, Relaxed Local Ternary Pattern for Face Recognition, IEEE International Conference on Image Processing, 3680 (2013).

[8] Raja G M and Sadasivam V, Journal of Electrical Engineering & Technology **12**, 402 (2017).