Weighted Multimodal Biometric Recognition Algorithm Based on Histogram of Contourlet Oriented Gradient Feature Description

Xinman Zhang
School of Electronics and
Information Engineering
MOE Key Lab for Intelligent
Networks and Network Security
Xi'an Jiaotong University
No.28 xianning west road
Xi'an, P.R.China
zhangxinman@mail.xjtu.edu.cn

Dongxu Cheng
School of Electronics and
Information Engineering
MOE Key Lab for Intelligent
Networks and Network Security
Xi'an Jiaotong University
No.28 xianning west road
Xi'an, P.R.China
dxcheng@stu.xjtu.edu.cn

Xuebin Xu
Guangdong Xi'an Jiaotong
University Academy.
No. 3, Daliangdesheng East Road
Foshan, China
ccp9999@126.com.

Abstract—**Although the unimodal biometric recognition (such as face and palmprint) has higher convenience, its security is also relatively weak. The recognition accuracy is easy affected by many factors such as ambient light and recognition distance etc. To address this issue, we present a weighted multimodal biometric recognition algorithm with face and palmprint based on histogram of contourlet oriented gradient (HCOG) feature description. We employ the nonsubsampled contour transform (NSCT) to decompose the face and palmprint images, and the HOG method is adopted to extract the feature, which is named as HCOG feature. Then the dimension reduction process is applied on the HCOG feature and a novel weight value computation method is proposed to accomplish the multimodal biometric fusion recognition. Extensive experiments illustrate that our proposed weighted fusion recognition can achieve excellent recognition accuracy rates and outmatches the unimodal biometric recognition methods.**

I. INTRODUCTION

WITH the continuous progress and development of the artificial intelligence technology, the conventional recognition technology is unable to meet people's advanced needs. In recent years, biometric identity recognition technology [1] has been widely studied to accomplish the user's identity recognition quickly and accurately.

Since the face biometric recognition technology has higher security and accuracy, it has attracted a large number of scholars to carry out research [2-3]. Many feature extraction methods have been proposed, such as the local binary pattern (LBP) [2]，discrete wavelet transform (DWT) [3], nonsubsampled contour transform (NSCT) [4], Histogram of Oriented Gradient (HOG) [5] and Gabor transform [6] et al. These features have also been widely used in palmprint recognition and achieved good results [7-8]. However, when the biometric is affected by the environment factors, the recognition accuracy will decline. To address this issue, some scholars proposed the multimodal biometric recognition theory and carry out extensive research [9-10]. Since NSCT can capture the smoothness and continuities along the contour of the image, it possesses the properties of shift invariance and multi-scale and has been widely studied and employed to descript the features of the face and palmprint [11].

Inspired by this, we propose a novel weighted fusion algorithm based on NSCT and HOG to realize the multimodal biometric recognition with face and palmprint. Firstly, NSCT are utilized to decompose the face and palmprint images. Then, we propose a novel feature extraction method named as histogram of contourlet oriented gradient (HCOG). With full consideration of the intra-class similarity and the inter-class similarity of sample features, a weighted fusion strategy is proposed to realize multimodal biometric fusion recognition.

The main content structure of this paper is organized as follows. Section II discusses the proposed algorithm in detail. In section III, the proposed algorithm is verified by extensive experiments, and the experimental results are analyzed and discussed in depth. Finally, we summarize the conclusion in section IV.

II. PROPOSED ALGORITHM

A. Nonsubsampled Contourlet Transform

Contourlet transform [12] employs the Laplace pyramid (LP) transform to decompose the image, then combines the coefficient points with uniform direction into an image contour by using the dimensional filter bank (DFB). However, in the process of image decomposition and reconstruction, contourlet transform needs to adopt the down-sampling and up-sampling operations, which will give rise to frequency aliasing and cause image translation, larger distribution of decomposition coefficient, translation variability and some other problems. To address these issues, Cunha et al. [13] proposed the NSCT algorithm which adopted multi-resolution decomposition by nonsubsampled pyramid (NSP) and nonsubsampled directional filter bank (NSFB) decomposition. Figure 1 illustrates the principle of NSCT [13]. NSCT does not require down-sampling after filtering and up-sampling before band-pass filtering, it only up-sampling on the corresponding filter, and then using low-pass filter and band-pass filter to complete the low-frequency and a series of high-frequency sub-bands filtering.
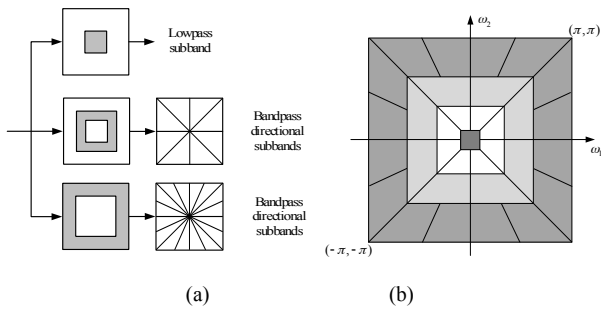
Fig. 1 Nonsubsampled contourlet transform. (a) NSFB structure that implements the NSCT. (b) Idealized frequency partitioning obtained with the proposed structure.

Since NSCT adopts the nonsubsampled operation, it has the stability with image shift, that is to say, it has the shift invariance.

B. Histogram of Oriented Gradient

Histogram of oriented gradient [5] is a classical feature description algorithm, which can describe the edge and shape information of an image well. The specific steps of HOG feature extraction algorithm can be summarized as follows, and figure 2 shows the schematic diagram.
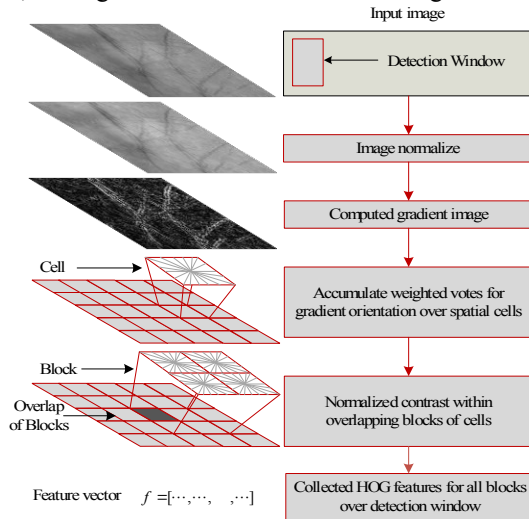

Fig. 2 Schematic diagram of HOG feature extraction.

(1) Normalize the image and calculate the gradient. (2) Divide the image into small cells with the same size and accumulate the weighted votes for gradient orientation with each cell to obtain the histogram of oriented gradient. (3) Group cells into large blocks and use the results to normalize all of the cells in the block. HOG descriptor is obtained by concatenating the normalized histogram vectors of all cells in each block. (4) Tile the detection window with a dense (overlapping is used) grid and cascade the feature descriptor of each block to generate HOG feature vector of the image.

The essence of HOG is the statistical information of image gradient, and the significant position of the gradient is mainly concentrated in the edge of the object. Therefore, it can achieve better recognition effect in the fields of pedestrian detection, and face recognition etc.

C. Histogram of Contourlet Oriented Gradient (HCOG) Feature Description

Although HOG has achieved stable and good effects in recognition, it lacks multi-scale adaptability due to the restriction of hierarchical rules. To address this issue, Bosch et al. [14] proposed a spatial pyramid representation to encode the object shapes. Inspired by this, we propose a novel HCOG feature descriptor by combining NSCT and HOG to improve the multi-scale performance of the conventional HOG feature.

HCOG feature extraction method is summarized as follows. (1) Implement the multi-scale decomposition of images based on NSCT. Figure 3 illustrates the NSCT decomposition results of a palmprint image, where the decomposition scale is 2, the corresponding bandpass directional subbands numbers are 2 and 4 respectively. (2) Extract HOG feature with the obtained lowpass subband and bandpass directional subbands respectively. (3) Concatenate these HOG descriptors with different subbands and scales to obtain the final HCOG feature vector.
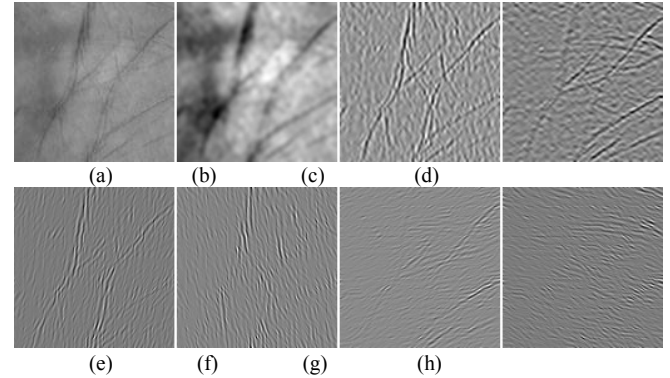

Fig. 3 The 2-level NSCT decomposition demonstration of a palmprint image. (a) Original image. (b) Lowpass subband. (c) and (d) are two different bandpass directional subbands of level 1. (e)-(h) are four different bandpass directional subbands of level 2.

D. Weighted multimodal biometric recognition algorithm

Although the feature extraction methods of face and palmprint images are same, the feature descriptions of different biometrics are quite different. In order to avoid the mutual interference between different biometrics, we present a novel weighted fusion recognition algorithm with face and palmprint. The specific algorithm steps are as follows.

Step 1. Image preprocessing. Preprocess the training face and palmprint images, including image normalization and ROI extraction etc.

Step 2. NSCT decomposition. NSCT decomposition is perform on the preprocessed face and palmprint images.

Step 3. HCOG feature extraction. Extract HOG feature from subbands obtained by NSCT decomposition with face and palmprint images respectively. Concatenate the HOG descriptors of the NSCT decomposition subbands with each face and palmprint sample to obtain the HCOG feature vector respectively. Then the feature matrix can be obtained, denoted as H^f and H^p for face and palmprint respectively.

Step 4. Feature dimension reduction. Employ the linear discriminant analysis (LDA) or principle component analysis (PCA) to reduce the dimensionality of the feature

matrices. Denote the dimension reduced training sample matrices as Tr^f and Tr^p for face and palmprint respectively.

Step 5. Weight value calculation. Denote the training matrices as $Tr^j = [Tr_1^j, Tr_2^j, ..., Tr_C^j]$, $j = \{f, p\}$, where C is the class number, suppose there are n samples for each class, then the training sample number can be calculated by $N = nC$. Denote $Tr_i^j = [tr_{(n-1)\times i+1}^j, ..., tr_{n\times i}^j]$, $j = \{f, p\}$ as the training subsets in the ith class. For each training sample x_k^f and x_k^p, calculate the similarity metric between different samples by using $S_{kl}^j = S(x_k^j, x_l^j)$, $j = \{f, p\}$, $(l = 1, ..., k-1, k+1, ...N)$. Then we calculate the intra-similarity as follows

$$intra\_S^j = \frac{1}{C}\sum_{i=1}^{C} intra\_S_i^j$$
$$= \frac{1}{n(n-1)C}\sum_{i=1}^{C}\sum_{\substack{k=(n-1)\times i+1}}^{n\times i}\sum_{\substack{l=(n-1)\times i+1 \\ l\neq k}}^{n\times i} S_{kl}^j, \quad j=\{f,p\}. \quad (1)$$

Calculate the inter-similarity as follows

$$inter\_S^j = \frac{1}{C}\sum_{i=1}^{C} inter\_S_i^j$$
$$= \frac{1}{C(C-1)n^2}\sum_{i=1}^{C}\sum_{k=(n-1)\times i+1}^{n\times i}\sum_{\substack{l=1 \\ l\notin I_i}}^{N} S_{kl}^j, \quad j=\{f,p\}. \quad (2)$$

where, $I_i = \{(n-1)\times i+1, (n-1)\times i+2, ..., n\times i\}, i=1,...,C$.

Define the discriminant function as follows

$$d^j = \frac{\left| intra\_s^j - inter\_s^j \right|}{intra\_s^j + inter\_s^j}, \quad j=\{f,p\}. \quad (3)$$

Then the weighted value is calculated by

$$\omega^j = \frac{d^j}{d^f + d^p}, \quad j=\{f,p\}. \quad (4)$$

Step 6. Fused recognition. For any given test images of face and palmprint, denoted as $y = [y^f, y^p]$, we employ the step 1 to step 4 to obtain the feature vectors, and denote it as $Tt = [tt^f, tt^p]$. Calculate the similarity between the test sample and each sample in the training set by $S_i^j = S(tt^j, tr_i^j), j=\{f,p\}, (i=1,2,...,N)$. Then we use the weighted value calculated by formula (4) to complete the fusion process

$$S_i = \omega^f S_i^f + \omega^p S_i^p, (i=1,2,...,N). \quad (5)$$

Finally, the nearest neighbor (NN) method is used to implement the recognition task.

III. EXPERIMENT AND ANALYSIS

In order to verify the effectiveness of the proposed algorithm, we carry out extensive simulation experiments on the multimodal database composed by the face and palmprint.

A. Experiment Database

Extended YaleB database contains 38 classes of face images, each class includes 64 different lighting conditions images. After manual cropping, the image size was adjusted to 192×168, and some face images are shown in figure 4.

Fig. 4 Face images of the Extend YaleB database.

Because the red blood cells in the human body has the characteristics of absorbing infrared light in specific bands. This can reduce the influence to the acquired palmprint images caused by palmprint desquamate, injury and sweating etc. We select the palmprint image acquired under near-infrared (NIR) illumination condition to fuse with the face biometric. The biometric research centre of Hong Kong polytechnic university (PolyU) has provided a multispectral palmprint database [1], which mainly includes four different spectral conditions (i.e. blue, red, green and NIR). It contains 500 different classes, and each class contains 12 palmprint images with four different spectral conditions. Figure 5 lists some palmprint images of this database.

Fig. 5 Some NIR palmprint images of the PolyU multispectral palmprint database.

B. Experiment Results and Analysis

Here we use the face and NIR palmprint images in Extend YaleB and PolyU database to constructed a multimodal database and conduct the fusion recognition experiment. According to the numbers of sample class and samples contained in each class, we construct a multimodal database with 38 classes and 12 samples for each class. We vary the training sample number from 2 to 7 from each class to demonstrate the experiments, the recognition rates of face and NIR palmprint are tested respectively to compare with the proposed fusion recognition algorithm. In this experiment, we used LDA method to conduct dimension reduction treatment. The specific experimental results are shown in table I.

TABLE I.
RECOGNITION RATES FOR DIFFERENT BIOMETRICS WITH THE TRAINING SAMPLE VARIES FROM 2 TO 7.

Biometrics	Recognition rates (%)					
	2	3	4	5	6	7
Face	84.74	89.47	93.75	94.74	95.18	97.89
Palm	98.95	98.83	100	100	100	100
Fusion case	99.74	99.71	100	100	100	100

From the table I, it is easy to see that the proposed algorithm outperforms the unimodal biometric recognition cases. Especially, when the number of training samples is small, (such as 2 and 3), our fusion recognition algorithm is 15% and 10.24% higher than the face recognition, and 0.79% and 0.88% higher than NIR palmprint recognition, respectively.

In order to verify the efficiency of our proposed HCOG feature, we implement experiment with different feature extraction methods, such as DWT, LBP, HOG, NSCT and Gabor feature etc. In addition, we use LDA and PCA methods to implement the dimension reduction and compare them with the case by using the original feature. In this experiment, the number of training samples is 3, and the rest are treated as test samples. Specific experimental results are shown in table II.

TABLE II.
RECOGNITION RATES WITH DIFFERENT FEATURE EXTRACTION METHODS

Feature extraction method	Recognition rates (%)		
	Original feature	LDA	PCA
DWT	97.66	98.54	97.08
LBP	97.95	98.54	97.37
HOG	92.40	91.81	94.74
NSCT	97.08	94.15	96.78
Gabor	98.25	98.54	99.12
Proposed method	99.74	99.74	99.71

From the data in table II, it is easy to see that the feature extraction algorithm proposed by us can effectively improve the recognition accuracy compared with the other feature extraction methods.

In order to verify the efficiency of the proposed algorithm, we draw the cumulative match characteristic (CMC) curve. As illustrated in the figure 6, the CMC curve takes Rank as the abscissa and the cumulative match score as the ordinate, which reflects the identity recognition ability of the biometric recognition system. It can be seen from the figure 6 that our proposed weighted fusion recognition algorithm can always achieve better fusion matching scores, which verifies that our recognition algorithm significantly outperforms the unimodal biometric recognition cases with face or NIR palmprint.
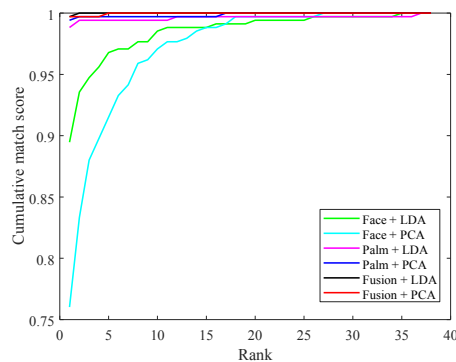

Fig. 6 Performance curves of our proposed weighted algorithms with the unimodal recognition in terms of Cumulative Match Characteristic.

IV. CONCLUSION

In this paper, a novel weighted fusion multimodal biometric recognition algorithm with face and palmprint is presented. By integrating the advantages of NSCT and HOG, a novel HCOG feature descriptor is proposed. At the same time, a fusion weight calculation strategy based on the sample similarity is presented to realize the biometric fusion recognition. Experimental results show that the proposed HCOG feature description and weighted fusion strategy can effectively improve the fusion recognition accuracy.

REFERENCES

[1] A. K. Jain, A. Ross, and S. Prabhakar. "An Introduction to Biometric Recognition," IEEE Trans. Circ. Syst. Vid. Tech., 2004, vol. 14 pp. 4-20, doi: 10.1109/TCSVT.2003.818349.

[2] J. Chen, V. Patel, and L. Liu, et al. "Robust Local Features for Remote Face Recognition," Image and Vision Computing, 2017, vol. 64, pp. 34-46, doi: 10.1016/j.imavis.2017.05.006.

[3] A. Ghasemzadeh, H. Demirel. "3D discrete wavelet transform-based feature extraction for hyperspectral face recognition," IET Biometrics, 2018, vol. 7, pp. 49-55, doi: 10.1049/iet-bmt.2017.0082.

[4] X. Xie, J. Lai, and W. Zheng. "Extraction of illumination invariant facial features from a single image using nonsubsampled contourlet transform," Pattern Recogn., 2010, vol. 43, pp. 4177-4189, DOI: 10.1016/j.patcog.2010.06.019.

[5] N. Dalal, B. Triggs. "Histograms of oriented gradients for human detection," International Conference on computer vision & Pattern Recognition, 2005, pp. 886-893, doi: 10.1109/CVPR.2005.177.

[6] C. Y. Low, A. B. Teoh, and C. J. Ng. "Multi-Fold Gabor, PCA and ICA Filter Convolution Descriptor for Face Recognition," IEEE Trans. Circ. Syst. Vid. Tech., 2019, vol.29, pp.115-128, doi: 10.1109/TCSVT.2017.2761829.

[7] A. Younesi, M.C. Amirani. "Gabor filter and texture based features for palmprint recognition," Procedia Comput. Sci. 2017, vol. 108, pp. 2488–2495, doi: 10.1016/j.procs.2017.05.157.

[8] S. W. Zhang, H. X. Wang, and W. Z. Huang, et al. "Combining modified LBP and weighted SRC for palmprint recognition," Signal Image Video Process. 2018, vol. 12, pp. 1035–1042, doi: 10.1007/s11760-018-1246-4.

[9] M. Haghighat, M. Abdel-Mottaleb, and W. Alhalabi. "Discriminant Correlation Analysis: Real-Time Feature Level Fusion for Multimodal Biometric Recognition," IEEE Trans. Inf. Foren. Sec., 2016, vol. 11, pp. 1984-1996, doi: 10.1109/TIFS.2016.2569061.

[10] N. Saini, A. Sinha. "Face and palmprint multimodal biometric systems using Gabor-Wigner transform as feature extraction," Pattern Anal. Appl., 2015, vol. 18, pp. 921-932, doi: 10.1007/s10044- 014-0414-6.

[11] W. F. Li, Y.C. Wang, and Z. Xu Z, et al. "Weighted contourlet binary patterns and image-based fisher linear discriminant for face recognition," Neurocomputing, 2017, vol. 267, 436-446, doi: 10.1016/j.neucom.2017.06.045.

[12] M. N. Do, M. Vetterli. "The Contourlet Transform: An Efficient Directional Multiresolution Image Representation," IEEE Transa. Image Process., 2006, vol. 14, pp.2091-2106, doi: 10.1109/ TIP.2005.859376.

[13] A. L. D. Cunha, J. P. Zhou, and M. N. Do. "The Nonsubsampled Contourlet Transform: Theory, Design, and Applications," IEEE Transa. Image Process., 2006, 15, pp. 3089-3101, doi: 10.1109/ TIP.2006.877507.

[14] A. Bosch, A. Zisserman, and X. Munoz, "Representing shape with a spatial pyramid kernel," International Conference on Image and Video Retrieval, 2007, pp. 401-408, doi:10.1145/1282280.1282340.

[15] D. Zhang, W. K. Kong, and J. You, et al. "Online palmprint identification," IEEE Trans. Pattern Anal. 2003, vol. 25, pp. 1041–1050, doi: 10.1109/TPAMI.2003.1227981.