Advances in Computer Science Research, volume 71
4th International Conference on Machinery, Materials and Information Technology Applications (ICMMITA 2016)

Hand Vein Recognition with Single-layer Feature Learning Model

Haoyu Wang [a*], Xiaomin Liu [b], Bingguang Chen [c]

School of Information and Electrical Engineering, China University of Mining and Technology, Xuzhou, 221000, China

[a]email: wzr78998@126.com, [b]email: xiaominliu@cumt.edu.cn, [c]email: cbgcumt@126.com

Keywords: Vein Recognition; Single-layer; Feature Learning; K-Means; SVM

Abstract. Performance of feature extraction and representation, sticking point of image recognition task, will directly influence the accuracy of final recognition. The traditional feature extraction algorithm of vein recognition is based on the sufficient prior knowledge of analysis on vein information characteristics, the shortcoming of which reflects in long time consumption spent on tuning parameters and special selection about later classifier to guarantee the final recognition rate as high as possible. The paper makes the attempt to introduce the K-means model, single-layer feature representation architecture, to the vein recognition task with some targeted modification, and adopts the SVM at the link of classifiers design. Finally, the proposed approach is rigorously evaluated on the self-built database and achieves the state-of-the-art RR (Recognition Rate) of 98.34%, which demonstrates the effectiveness of the proposed model.

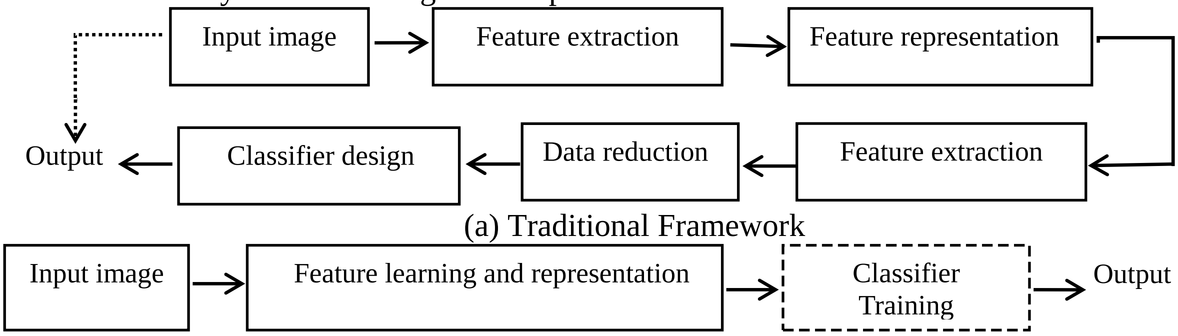
Introduction

Vein recognition, which is not introduced until 1990 by MacGregor and Welford [1], has become one of the most popular biometric identification methods with the advantage of unique, portable and inherent properties. The main characteristics we stress when putting vein recognition into practice not only because the vein patterns of individual is different even between identical twins [2], but also lies in the fact that it is easy-acceptable, anti-counterfeit and also with high recognition rate.

Meeting all the requirements of biometric identification like other personal traits [3], vein recognition is also characterized with high convenience in image acquisition and feature representation which results in wide investigation on vein information research in hand covering palm vein [4-6], dorsal vein [7] [8] [9] and finger vein [10]. And the entire framework for the three kinds of recognition is more or less same, which covers vein image preprocessing, vein image feature extraction and representation, classifier design and vein recognition [11]. However the hand-crafted feature extraction methods of the traditional framework cannot ensure the high recognition accuracy, well robust performance, shorter time consumption to the same level. The analytical route cause is that the feature representation methods design are based on the single or mixture observation about the part of vein image information features, and it's a common sense that the proposed methods cannot achieve the stage covering all the vein information representation based on the hand-crafted feature analytical result. The major goal in machine learning is to learn deep hierarchies of features for other tasks. A typical approach taken in the literature is to use an unsupervised learning algorithm to train a model of the unlabeled data and then use the results to extract interesting features from the data [12] [13] [14]. Depending on the choice of unsupervised learning scheme, it is sometimes difficult to make these systems work well. K-means has already been identified as a successful method to learn features from images by computer vision researchers. The popular "bag of features" model [15] [16] from the computer vision community is very similar to the pipeline that we will use in this chapter, and many conclusions here are similar to those identified by vision researchers [17] [18]. So the paper introduces the K-means to realize the vein feature representation system so as to find the feature distribution without adopting hand-crafted feature as the prior knowledge, and the experimental results demonstrate the efficiency of the proposed single-layer feature learning model in solving the feature learning problem in hand vein recognition task.

Overview of Feature Learning Algorithms for Hand Vein Recognition

To well get rid of the existing problem in traditional hand vein recognition framework (as shown in Fig .1(a)), which has the disadvantage of great redundancy and lower recognition rate according to the application demand, the paper proposes the learning feature model (as shown in Fig. 1(b)) and designs the recognition experiment to demonstrate the necessity of introducing the learning representation theory into vein recognition experiment.



(a) Traditional Framework

(b) Feature Learning Framework

Fig.1. Different Hand Vein Recognition Framework

Feature learning, the product of deep learning and also called representation learning, is a new and rising concept [12]. The concrete step for feature learning is firstly trying to design the suitable network structure and training it from the original image pixel data with unsupervised or supervised training, and then effective representation of original pixel could be obtained in the hidden layer and the classification result could also be predicted with the structure of importing classifying module in the output layer. The main representation learning methods make up with two different kinds of structures, the first of which is single-layer based structure including K-means learning structure[13], the SRBM(Sparse restricted Boltzmann machine) [14]etc. This kind of structure is characterized by simplicity in layer composition and has been applied into many different sorts of vision tasks [15] [16] [17], while the other one is the deep learning structure.

Due to the data dimensionality and dispersion of vein images are both at the low level; the paper introduces the single-layer representation learning structure for the recognition experiment design. The traditional single-layer structure involves the network training and the feature extraction procedures, for the part of training process, training images are firstly processed into patches (at least 100000) and sent for the especially designed pre-process methods, and then is the selection and modification of single-layer structure for training on the image patches to learn the network weight matrix (filters), following the training step is the feature extraction part. When applied into vein image processing experiment, patches are extracted by the pre-set step rule (the size should be kept the same with the training part), and then every single patch is projected onto the format of feature vector by the learned weight matrix, and finally the vector are processed with the max pooling to get the final representation expression for the input vein images.

Patching and pre-process of vein images

The database of the paper is from [19]. Randomly select 400 images from the self-built database(a total of 500) and conduct the patching process to get the result of $w \times w$ for each image(the concrete result of this paper is respectively about 8×8,12×12,16×16) with the expressions set as $P=\{p^{(1)},p^{(2)},\cdots,p^{(N)}\}, p^{(i)}\in R^{w\times w}$, and then before the training the single-layer network with unsupervised method, the contrast normalization and white process[13] are conducted on the patches to decrease the correlation.

(1) Contrast normalization

With the theory that structural feature is the most important and considerable feature adopted at the link of feature extraction part, so it is necessary to design the contrast normalization method to decrease the influence of the brightness feature on the feature learning set, the corresponding formula is set as follows:

$$x^{i}=\frac{p^{(i)}-mean(p^{(i)})}{\sqrt{var(p^{(i)})+\varepsilon}} \quad (1)$$

The ε is set to avoid the situation that there exists 0 in the denominator, and the specific value is set as 0.01 to realize the noise suppression.

(2) PCA whitening

Whitening, the process to decrease the redundancy, process would rich the features of the images including low correlation and same variance.

The PCA projection is introduced firstly to remove the relevancy of the data and then the scaling process on the eigenvalues factor is adopted to ensure the unit variance. Suppose the data collection of the input image patches to be $X_{m\times N}$(each row consists of the image patches vector), the corresponding covariance matrix could be expressed as $C=\mathrm{cov}(X_{m\times N})$, and then the Eigen decomposed result is that$[V,D]=eig(C)$, the PCA whitening can be expressed as follow:

$$X_{PCAwhite}=\frac{(X_{m\times N}-mean(X_{m\times N}))\cdot V}{\sqrt{diag(D)+\varepsilon}} \quad (2)$$

Weight Matrix Learning Based on the K-means Algorithm

In this part, the paper introduces the improved K-means algorithm for learning the weight matrix to realize the mapping function: $\mathrm{f}: R^{w\times w}\to R^{K}$ after the proposed pre-process part, the function could realize the mapping procedure from the input vein image patches $x^{(i)}$ to the K-dimensional feature vector $f^{(i)}$. The paper defines the input vein image patches as $x^{(i)}\in R^{n}, i=1,2,\cdots,m$ and then is the vector quantization process to learn the dictionary with the center $\mathrm{D}\in R^{n\times k}$, the criterion of learning is shown as follows:

$$minimize\sum_{i}\left\|D\cdot s^{(i)}-x^{(i)}\right\|_{2}^{2} \quad (3)$$

$$\mathrm{s.t.}\left\|s^{(i)}\right\|_{0}\leq 1,\forall i \ and \ \left\|D^{(j)}\right\|_{2}=1,\forall j \quad (4)$$

The expression of $s^{(i)}$ stands for the code vector of the input vector $x^{(i)}$.The first constrain is set to guarantee the sparsity with the condition that there exists one nonzero item at most, while the second criterion is set to avoid that there exists some unexpected data with large value in the feature vectors, which is the set to realize the constrain on the dictionary, the specific solving process is according to the methods in [18]. The final process is coding on the new input vectors to get the features after getting the training dictionary:

$$s_{j}^{(i)}=\begin{cases}D(j)^{T}x^{(i)} & if\ j==argmax\left|D^{(l)T}x^{(i)}\right| \\ 0 & otherwise\end{cases} \quad (5)$$

Convolved Feature Extraction

After getting the weight matrix W by training process, the feature extraction on the new input image is conducted by the convolved operation. The specific procedure involves the following steps: Conducting the convolution operation between the new input vein image and the weight matrix W to get the feature map with the size of$(n_{H}-w+1)\times(n_{w}-w+1)$, and an analytical but knotty result is that the feature map, each pixel of it is represented as a K-dimensional vector, cannot be adopted directly for final feature representation with the problem of high dimensionality resulting in higher computational complexity. To decrease the dimensionality of the feature map and obtain the feature representation characterized with anti-translation, the paper proposes the theory that cutting the feature map into four same-sized patches, and then conduct the pooling operation on each patch to get the final vein image feature vector with 4K dimensionality.

Results and Discussion

During the classification procedure, the SVM classifier is firstly trained with the convolved feature vectors of the 350 training samples, and the realization of SVM is based on the Liblinear Software [20]. After obtaining the trained model, the classification experiment with the remaining samples is conducted and the specific recognition result is shown in Table 1.

Table 1. Recognition Accuracy with Single-layer Feature Learning and Representation Model(%)

Method	Size of patches (W)	Dimensionality of learning feature(K)		
		400	800	1000
Improved K-MEANS	8	87.19	92.76	97.21
	12	87.58	93.91	97.63
	16	87.86	96.57	98.34

It could be inferred from the results in Table 1 that the recognition result is directly proportional to the dimensionality of the learning feature while the size of patches has few influence on the final result. On the other hand, the classification result of feature learning-based vein recognition method is directly and greatly influenced by the size of the training and testing database, and it is illustrated [18] that the more sample of the database, the better recognition result, and the recognition result of this model and the analytical result fully demonstrates that it is necessary to try the single-layer feature learning and representation method in vein recognition experiment design.

Conclusion

To improve the recognition result from the perspective of feature learning, the paper innovatively proposes the relatively simple but effective single-layer based feature learning and representation model into the recognition experiment, and to demonstrate the effectiveness. The proposed model firstly conduct contrast normalization and PCA whitening to obtain the pre-processing input samples, followed by the feature transformation matrix learning based on K-Means strategy, and then the learned matrix dictionary is adopted to the new input vein images for feature extraction and then the SVM is realized to obtain the final state-of-the-art classification result. What's more, the paper also making detailed analysis on how to improve the recognition accuracy with the proposed feature learning model.

For future extension of our work, we are interested in trying to enrich the self-built vein database under the premise of guaranteeing the effectiveness of the captured vein images and try them on the proposed single-layer feature learning and representation model to observe the change of the recognition result, on the other hand, we are intended to introduce the deep learning architecture into the vein recognition task by making some necessary modification to it.

References

[1] M. J. Morgan, "Features and the primal sketch," Vis. Res., vol. 51, no. 7, pp. 738-753, 2011.

[2] A. Kumar, M. Hanmandlu, and H. M. Gupta, "Online biometric authentication using hand vein patterns," in Proc. IEEE Symp. Comput. Intell. Secur. Defence Appl.,
Ottawa, ON, Canada, 2009, pp. 1-7.

[3] Moreno. P, Marín-Jiménez. M. J, Bernardino. A, Santos-Victor. J and Blanca. N. P. D. L, "A comparative study of local descriptors for object category recognition: sift vs hmax," Lecture Notes in Computer Science, vol. 4477, pp. 515-522.

[4] C. Lin and K. Fan, "Biometric verification using thermal images of palm-dorsa vein patterns," IEEE Trans. Circuits Syst. Video Technol., vol. 14, no. 2, pp. 199-213, Feb, 2004.

[5] S. Malki and L. Spaanenburg, "Hand veins feature extraction suing DT-CNNs," Proc. SPIE, vol. 6590, May. 2007.

[6] P. Ladoux, C. Rosenberger, and B. Dorizzi, “Palm vein verification system based onSIFT matching,” in PROC. IEEE Int. Conf. Biometrics, Alghero, Italy, 2009, pp. 1290-1298.

[7] S. Zhao, Y. Wang, and Y. Wang, “Biometric identification based on low quality hand vein pattern images,” in PROC. IEEE Int. Conf. Mach. Learn. Cybern., Kunming, China, 2008, pp. 1172-1177.

[8] A. Kumar and K. Prathyusha, “Personal authentication using hand vein triangulation and knuckle shape,” IEEE Trans. Image Process., vol. 18, no. 9, pp. 2127-2136, Sep. 2009.

[9] Y. Wang, K. Li, j. Cui, L. Shark, and M.Varley, “Study of hand-dorsa vein recognition,” in Proc. Int. Conf. Intell. Comput., Changsha, China, 2010, pp. 490-498.

[10] N. Miura, A. Nagasaka, and T. Miyatake, “Feature extraction of finger-vein pattern based on repeated line tracking and its application to personal identification,” Mach. Vis. Appl., vol. 15, no. 4, pp. 194-203, 2004.

[11] Mikolajczyk, K. Schmid, and C. Schmid, “A performance evaluation of local descriptors,” IEEE Trans. Pattern Anal and Mach Int., vol. 27, no. 10, pp. 1615-1630, Oct. 2005.

[12] Y. Bengio, A. Courville, P. Vincent, “Representation learning: a review and new perspectives” [J], IEEE Transactions on Pattern Analysis and Machine Intelligence,35(8): 1798-1828, 2013.

[13] A. Coates, H. Lee, A. Y. Ng, “An analysis of single-layer networks in unsupervised feature learning” [C], in International Conference on Artificial Intelligence and Statistics, 2011.

[14] J. Ngiam, P. W. Koh, Z. Chen, et al, “Sparse filtering” [C], in Neural Information Processing Systems, 2012.

[15] A. Coates, B. Carpenter, C. Case, et al. “Text detection and character recognition in scene images with unsupervised feature learning” [C], in International Conference on Document Analysis and Recognition, 2011.

[16] K. Kavukcuoglu, M. Ranzato, R. Fergus, “Learning invariant features through topographic filter maps” [C], in IEEE Conference on Computer Vision and Pattern Recognition, 2009.

[17] Y. Netzer, T. Wang, A. Coates, et al, “Reading digits in natural images with unsupervised feature learning” [C], in NIPS Workshop on Deep Learning and Unsupervised Feature, 2011.

[18] A. Coates, A. Y. Ng, “Learning feature representation with K-means” [J], Neural Networks: Tricks of the Trade, 7700: 561-580, 2012.

[19] Jun Wang, Guoqing Wang, et al, “ Hand vein recognition based on PCET” [J], OPTIK, 127(19): 7663-7669, 2016.

[20] R. Fan, K. Chang, C. Hsieh, et al, “LIBLINEAR: A library for large linear classification”[J], Journal of Machine Learning Research, 9:1871-1874, 2008.