

# An Improved Algorithm based on Palm Print Recognition

Tong Zhang<sup>1</sup> Xinghui Zhang Tao Li Xiaoqin Liu

**Abstract.** This paper aims at the deficiency of classical PCA algorithm on only extracting linear palm print features, sufficiently uses the double matching thought of recognition technology to propose a new algorithm which relatively efficient in extracting the classified identification information. The new proposed algorithm based on PCA algorithm and by using transformation of kernel function to convert the input samples into high-dimensional space through which efficiently solving the problems of small samples and sharply rising computation quantities, while the two aforementioned problems were caused by the number of training samples after matched to the high-dimensional space that smaller than the dimensions themselves. Experimental results show that the proposed algorithm is effective and the accuracy is improved obviously.

**Keywords:** PCA. transformation of kernel function. non-linear transformation. palm print recognition. transformation of dimension

## 1.1 Introduction

Palm print, provided with rich amount of information, low-cost gathering device and the possession of non-invasive feature, etc, obtains more and more attention from many researchers in recent years and palm print identification technology also becomes a hot research area in area of biological characteristics recognition.

How to extract palm print features occupies a decisive position in palm print recognition system and the results greatly determines the effect of matching classification. The existing palm print features extracting algorithms can be classified as follows: the feature extraction method based on structure, the feature extraction method based on time-frequency transform and the feature extraction method based on statistics and subspace, etc. thereinto, the most classical is the Principal Components Analysis(PCA) method and which gained extensive attention from many researchers for the good results that had achieved in facial recognition area.

---

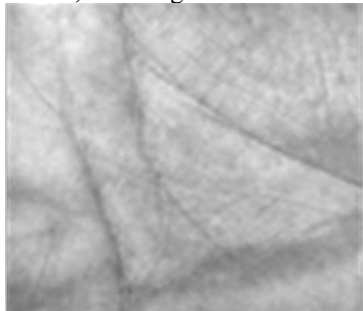
<sup>1</sup> Tong Zhang(✉)

School of Information Science and Engineering, Yunnan University. 650091 Kunming, China  
e-mail: lunwenlz@sina.cn

The PCA technique is widely used in image compression and pattern recognition area for the extracted object ROI image can be projected a shadow on low-dimensional space from high-dimensional space through PCA transformation to meet the goal of reducing training as well as recognition process, but its shortcoming is leaving out of consideration of the different training samples among different classes in transform process to reduce dimension, so PCA transformation is a ideal dimensionality reduction method from this perspective but not a ideal pattern classification method, what is more, PCA feature extracted method belongs to linear transformation and only linear features can be extracted, whereas, the amount of non-linear pattern features existing in application in fact. In order to overcome the classification defect in PCA transformation, this paper proposes a new palm print recognition method which uses kernel function to convert input sample space to another high-dimensional space where is more conducive to recognize and extract stage classification information, moreover, the utilization of kernel function efficiently avoids so called “curse of dimensionality” and huge computation problem after the non-linear feature projected a shadow on high-dimensional space and obtains better experimental results under laboratory conditions.

## 1.2 The Introduction of Typical PCA Transformation

Suppose sample set has  $l$  palm print ROI images in total, the size of which is  $w \times h$ , put each of ROI image pixels in order by left to right and top to bottom way and then convert  $r$ -th dimension column vector in sequence:  $x_1, x_2, \dots, x_l$ ,  $1 \leq i \leq l$ ,  $r = w \times h$ , thereinto,  $x_i$  represents column vector formal training sample obtained after the  $i$ -th palm print ROI image expanded. The first step was to extract palm print ROI image before PCA transformation and then the image enhancement was performed, the design sketch are shown in Fig.1.1 and Fig.1.2



**Fig.1.1** The extracted palm print ROI image



**Fig.1.2** The palm print ROI after image enhancement

### 1.21 The Main Steps of Typical PCA Transformation

First, calculate the mean vector value of all training samples:

$$\mu = \frac{1}{l} \sum_{i=1}^l x_i \quad (1.2.1)$$

Then, every training sample minus mean value  $\mu$  :

$$\Phi_i = x_i - \mu \quad (1.2.2)$$

Then, calculate covariance matrix C:

$$C = \frac{1}{l} \sum_{i=1}^l (x_i - \mu)(x_i - \mu)^T \quad (1.2.3)$$

Then, calculate the feature vector and eigenvalue of C, take out the maximal k eigenvalues  $\lambda_1, \lambda_2, \dots, \lambda_i, \dots, \lambda_k$ , and their corresponding feature vectors  $\mu_1, \mu_2, \dots, \mu_i, \dots, \mu_k$ ,  $1 \leq i \leq k$ , thereinto, k is far less than r.

Then, the feature vector  $\mu_i$  is normalized in order to obtain orthonormalization feature vector  $w_i$ , as follows:

$$w_i = \frac{u_i}{\|u_i\|} \quad (1.2.4)$$

Thereinto,  $\|\bullet\|$  is euclidean distance in feature space and  $W=[w_1, w_2, \dots, w_k]$  represents the basis of feature subspace of palm print ROI image. When input palm print  $x_s$  to recognize, the recognition steps are as follows:

1. The unknown palm print minus mean value vector

$$\Phi_s = x_s - \mu \quad (1.2.5)$$

2. Let  $\Phi_s$  project to feature subspace which was obtained by training

$$Y_s = W^T \Phi_s \quad (1.2.6)$$

3. Let  $\Phi_j$  ( $1 \leq j \leq l$ ) project to feature subspace which was obtained by training

$$Y_j = W^T \Phi_j \quad (1.2.7)$$

4. Calculate euclidean distance  $d_E$  between  $Y_s$  and  $Y_j$ , and preset a threshold, if the euclidean distance  $d_E$  is smaller than the preset threshold that shows unknown sample  $x_s$  and j-th training sample come from the same one user.

### 1.3 The improved new algorithm proposed in this paper

The new algorithm aims at the deficiency of classical PCA algorithm on only extracting linear palm print features but whose shortcoming is missing a lot of non-linear features, sufficiently using the double matching thought of recognition technology and based on the foundation of PCA algorithm rough location, making training samples project a shadow on high-dimensional space, where processes LDA transformation to feature data (see details in reference 1), then through kernel function thought solves “curse of dimensionality” and huge computation problem, and finally reaches the target for fine matching by extracting non-linear features in palm print image.

#### 1.3.1 The Related Fundamental Principle of the New Algorithm

##### 1.3.1.1 The Operating Principle of the New Algorithm

Suppose training set T has m people palm print ROI images in total,  $T=\{X_1, X_2, \dots, X_m\}$  and every people has n palm print images, so set  $x_{ij}$  represent the column vector transformed by i-th people j-th palm print image ROI, then the training set has the number of images is  $l=m \times n$ . For simplicity's sake, we can represent training set as  $\{x_r, r=1, 2, \dots, l\}$ , the basic idea of the new algorithm proposed in this paper is performed by the following non-linear mapping function to let training set project to a high-dimensional space F

$$\varphi: R^l \rightarrow F, x_r \mapsto \varphi(x_r) \quad (1.3.1)$$

Then processes LDA transformation to the feature data which obtained by mapping the  $x_r$  in high-dimensional space (see details in reference 1), in formula 8,  $\varphi(\cdot)$  represents a non-linear mapping function who can be a radial basis function or an exponential function, etc, the aim of the new algorithm is maximum following Fisher criterion

$$\lambda = \frac{v^T S_B^{\varphi} v}{v^T S_T^{\varphi} v}, v \neq 0 \quad (1.3.2)$$

In the above formula,  $S_B^{\varphi}$  is the scatter matrix between classes,  $S_T^{\varphi}$  is the total scatter matrix, and the two matrixes can be defined as follows:

$$S_B^{\varphi} = \sum_{i=1}^m n(\bar{\varphi}(x_i) - \bar{\varphi})(\bar{\varphi}(x_i) - \bar{\varphi})^T \quad (1.3.3)$$

$$S_T^\varphi = \sum_{r=1}^l (\varphi(x_r) - \bar{\varphi})(\varphi(x_r) - \bar{\varphi})^T \quad (1.3.4)$$

Thereinto,  $\bar{\varphi}$  is all the feature data's  $\varphi(x_r)$  mean vector obtained by palm print training set through  $\varphi(\cdot)$  mapped,  $\bar{\varphi}(x_i)$  is all the feature data's mean vector obtained by i-th people's all palm print samples through  $\varphi(\cdot)$  mapped.

### 1.3.1.2 The Fundamental Principle of Kernel Function Transformation

To random  $\varphi(x_r) \in F$ , suppose there exist a non-linear mapping  $\varphi(\cdot)$  and a kernel function  $k(x_r, x_s)$  that satisfy  $k(x_r, x_s) = \varphi(x_r) * \varphi(x_s)$ , through defining the kernel function  $k(x_r, x_s)$  in low-dimensional space, we can calculate the inner product of the two vector  $\varphi(x_r)$  and  $\varphi(x_s)$  which obtained by non-linear mapping to high-dimensional space. The kernel function can accept radial basis function, exponential function and Sigmoid function, etc, thereinto, radial basis function is most common and its expression is as follows:

$$k(x_r, x_s) = \exp\left(\frac{-\|x_r - x_s\|^2}{2\delta^2}\right) \quad (1.3.5)$$

In the above expression,  $\delta$  represents standard deviation of the radial basis function.

For the all solutions of  $v$  in formula 1.3.2 were obtained by the expansion of feature space  $F$ , so we can view  $v$  as the linear combination of  $\varphi(x_r)$ :

$$v = \sum_{i=1}^l \alpha_i \varphi(x_i) \quad (1.3.6)$$

Thereinto,  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_l)^T$  and simplified by formula 1.3.2 and formula 1.3.6 as shown follows:

$$\lambda = \frac{\alpha^T K W K \alpha}{\alpha^T K K \alpha} \quad (1.3.7)$$

Thereinto,  $K$  is a  $l \times l$  kernel matrix:

$$K = (k(x_r, x_s))_{r=1, \dots, l; s=1, \dots, l}$$

$W$  is a  $l \times l$  block diagonal matrix:

$$W = (W_i)_{i=1, \dots, l}$$

$W_i$  is a  $n \times n$  square matrix.

### 1.3.2 The Fundamental Steps of the New Algorithm

1. Based on the obtained kernel matrix  $K$  and block diagonal matrix  $W$  and processes eigenvalue decomposition of  $K$ , to get the feature vector matrix  $P$  and feature diagonal matrix  $T$  through normalization, so we can get follows:

$$K = P\Gamma P^T \quad (1.3.8)$$

2. Use  $P$  and  $\Gamma$  represent formula 1.3.7 so we can get  $\lambda = \alpha^T P\Gamma P^T W P\Gamma P^T \alpha / \alpha^T P\Gamma P^T P\Gamma P^T \alpha$ , set  $\beta = \Gamma P^T \alpha$ , so we can get as follows:

$$\lambda \beta = P^T W P \beta \quad (1.3.9)$$

Then solves feature vector  $\beta$  and eigenvalue to obtain the normalized  $\alpha$ .

3. To input sample  $x_z$ , its projection  $y_z$  in feature space is as follows:

$$y_z = v^T \varphi(x_z) = \sum_{i=1}^l \alpha_i \cdot k(x_i, x_z) \quad (1.3.10)$$

4. Calculate euclidean distance between  $y_z$  and  $y_j$ , if the euclidean distance smaller than the preset threshold that shows unknown sample  $x_z$  and  $j$ -th training sample come from the same one user.

## 1.4 Experimental Results and Analysis

In order to testify the proposed algorithm, research team have gathered 50 people's left hand and right hand palm print images and every hand were gathered 10 images, according to research of biological characteristics, every people's left hand and right hand has the different palm print, so we can regard the left hand and right hand belong to the same people as two different hands that come from two different people, then the palm print data base used for experiment can regard as 1000 images and the experimental environment are intel (R) 3.30GHz, 1G RAM, Matlab R2010b. In order to testify the recognition rate of the proposed algorithm, the palm print ROI image size of  $80 \times 80$ ,  $100 \times 100$ ,  $120 \times 120$  and  $200 \times 200$  were used to train and recognize, the recognition rate of the algorithm before and after improved are shown in the following table:

The image size for algorithm	80x80	100x100	120x120	200x200
Original algorithm(PCA)	96.12%	96.62%	97.41%	96.33%
The improved algorithm	99.39%	99.67%	99.92%	99.61%

**Table 1.1**

We can see from the above table, in different size of images, the PCA algorithm mean recognition rate is 96.62% and the improved algorithm mean recognition rate is 99.65%, when the size of image is  $120 \times 120$ , namely the feature dimension is 14400, the recognition rate is the highest of all. No matter how big the image is, the improved algorithm always possesses higher recognition rate than original PCA and in the same algorithm, with the bigger as the image is, the recognition rate is better improved, but this do not prove that the bigger the image is ,the higher recognition rate would have, on the contrary, when the size of the image is  $200 \times 200$ , the recognition rate does not rise but decline.

Through analysis of the wrongly classified images, we can find that illumination conditions is still a key factor to influence recognition rate, this is possible when gathering palm print images, as time went on, illumination conditions had changed so as to influence the quality of the gathered palm print images.

## 1.5 Conclusion

This paper, aiming at the deficiency of classical PCA algorithm on only extracting linear features but whose shortcoming is missing a lot of non-linear features, proposes a new algorithm to extract non-linear features which based on the linear feature foundation of PCA algorithm and making input samples project a shadow on high-dimensional space in order to recognize and extract under the transformation of LDA and kernel function. The proposed new algorithm sufficiently uses the double matching thought of recognition technology and based on the foundation of PCA algorithm rough location (linear feature) to achieve fine matching, unlike descending dimension algorithm like PCA algorithm discards image information, the mapping recognition rate is highly enhanced. The experiment shows that the new algorithm is efficient and the recognition rate is better improved than the original one.

## 1.6 Reference

- [1] Shi Jing, Fan Meng, Li Wan-Long. Topic analysis based on LDA model. *Acta Automatica Sinica*, 2009, 35(12): 1586-1592
- [2] X. Wu, D. Zhang, K. Wang, and B. Huang. Palmprint Classification Using Principal Lines[J]. *Pattern Recognition*, 2004, 37(10): 1987-1998.
- [3] X. Wu, D. Zhang, and K. Wang. Palm-Line Extraction and Matching for Personal Authentication[J]. *IEEE Transactions on Systems, Man, and Cybernetics, Part A*, 2006, 36(5): 978-987.
- [4] D.-S. Huang, W. Jia, and D. Zhang. Palmprint Verification Based on Principal Lines[J]. *Pattern Recognition*, 2008, 41(4): 1316-1328.
- [5] A. K. Jain and J. Feng. Latent Palmprint Matching[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, 31(6): 1032-1047.
- [6] J. You, W.-K. Kong, D. Zhang, and K. H. Cheung. On Hierarchical Palmprint Coding with Multiple Features for Personal Identification in Large Databases[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2004, 14(2): 234-243.
- [7] W. Li, J. You, and D. Zhang. Texture-Based Palmprint Retrieval Using a Layered Search Scheme for Personal Identification[J]. *IEEE Transactions on Multimedia*, 2005, 7(5): 891-898.
- [8] D. Zhang, W. K. Kong, J. You, and M. Wong. Online Palmprint Identification[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, 25(9): 1041-1050.
- [9] W. K. Kong, D. Zhang, and W. Li. Palmprint Feature Extraction Using 2-D Gabor Filters[J]. *Pattern Recognition*, 2003, 36(10): 2339-2347.
- [10] A. W.-K. Kong, D. Zhang, and G. Lu. A Study of Identical Twins' palmprints for Personal Verification[J]. *Pattern Recognition*, 2006, 39(11): 2149-2156.
- [11] X. Pan and Q.-Q. Ruan. Palmprint Recognition Using Gabor-Based Local Invariant Features[J]. *Neurocomputing*, 2009, 72(7-9): 2040-2045.