Cascaded Feature Reduction in Palm Vein Recognition based on Variance Order

Prasti Eko Yunanto
School of Computing, Telkom University
Telekomunikasi No.1, Bandung,
Indonesia, 40257
prastieko@students.telkomuniversity.ac.id

Hertog Nugroho
Electric Engineering Department
Bandung State of Polytechnic, Gegerkalong Hilir, ds. Ciwaruga, Bandung, Indonesia, 40012
hertog@polban.ac.id

Tjokorda Agung Budi W.
School of Computing, Telkom University
Telekomunikasi No.1, Bandung,
Indonesia, 40257
cokagung@telkomuniversity.ac.id

Abstract

Feature or dimensionality reduction has become one of fundamental problem in the field of pattern recognition such as biometrics. The choosing number of feature or dimension has become one challenge. Instead choosing number of feature manually, in this paper, we proposed an automatic feature reduction by using a cascaded feature reduction schemes based on variance order of the DCT feature space and eigenvalue of k-PCA in the palm vein recognition. Based on experiment results, our proposed scheme can achieve recognition rate above 0.92 accuracy which uses fewer features and can reduce time process significantly until 99.5% comparing with traditional manual feature reduction method.

1 Introduction

In recent years, biometrics technology was widely studied and used for alternative method in personal identification and authentication [1][2]. The reason is the weakness of conventional system such as personal identification number (PIN), password, and ID card, which were easily lost, stolen, can be duplicated and forgotten[1]. Biometrics technology was expected to be an alternative solution for the problem of conventional system, where the people don't need to remember, memorize, and bring the identifier to anywhere, because the identifiers are not from third party but from human itself (e.g. finger print, eyes, face, hand geometrics, signature, and etc.), so that, people automatically bring their identifier every time and everywhere[1].

In biometrics system, the identification algorithms must be accurate and fast [3]. Based on survey of palm print recognition that has been done by Adam Kong et al.[3], there are five objectives of biometrics system which need to be considered: cost, user acceptance, accuracy, speed and security. Palm vein is one kind of biometrics trait that has high accurate rate authentication like iris recognition and also serve as a convenient form as fingerprint recognition[2]. Palm vein are widely developed for biometrics, because of their advantages, such as: (1) palm vein is more secure than other biometrics trait, because a vein is located in human body that only can be used while the person still alive; (2) high user convenience; and (3) not affected by external environment[2][4].

Dimensionality or feature reduction has become one of fundamental problem, since in the real world application there are large volumes of high dimensional data[5]. Features reduction can reduce time process of the system, also can remove irrelevant features. One popular approach that has been used in biometrics is principle component analysis[1][3][6]. But the process to find the number of feature was still done manually. Although the effort was done only on training process, if there are large data, it will became a problem.

2 Proposed Method

In this paper we proposed a feature reduction based on variance order of discrete cosine transform (DCT) coefficient and eigenvalue of kernel principle component analysis (k-PCA) in palm vein recognition as shown in figure 1.

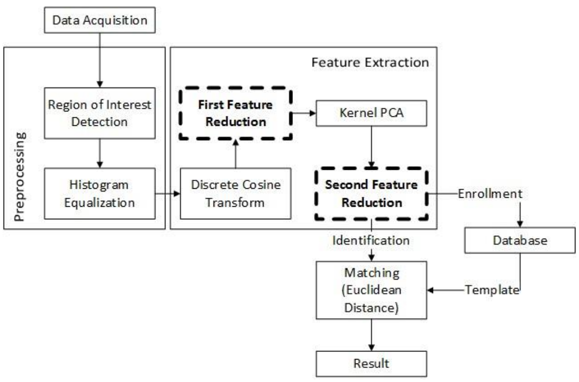

Figure 1. Block diagram of proposed biometrics system

Firstly, the palm vein images are preprocessed by competitive hand valley detection (CHVD)[7] to get region of interest and histogram equalization to enhance the image quality[8]. Then, the spatial feature of the images are transformed into frequency domain using DCT [9]. Here we introduce the first feature reduction process, based on variance of the DCT coefficient. After that, the features were extracted by k-PCA[10] to provide the eigenvector and eigenvalue. In this step, we also introduce the second feature reduction based on curve of eigenvalue, where the high of eigenvalue correspond with the eigenvector which contains high discriminant features[6]. Lastly, the training data features is compared with testing features based on the euclidean distance.

2.1 Basic Idea

The basic idea of feature reduction which is proposed in this paper, was adopted from principle component analysis (PCA). The features can be reduced by choosing eigenvector corresponding with the largest eigenvalue[6]. Eigenvalue represents how significant the feature can discriminate classes/clusters. Figure 2 shows a curve of the eigenvalue. We assume that eigenvalue which have high discriminant capability (on left side of boundary line) give significant contribution to the identification accuracy. The problem is to decide the boundary line. The previous method[6][3][1][9] used manual selection and observe it's impact on the identification result. These try and error approaches requires high cost to get optimal solution. In this paper, we proposed an automatic feature reduction scheme with comparable result.

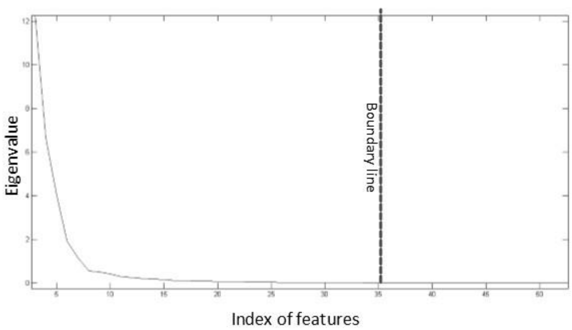

Figure 2. example of distribution value on eigenvalue

Based on that idea, we treat the eigenvalue as variance of the data. If the variance of data small, it means that data distribution is very close, and if the variance is large the distribution spreads out.

$$\sigma^2 = \frac{\sum_{i=1}^{n} (x_i - \mu)^2}{n} \quad (1)$$

where σ^2 is variance, x is data, μ is mean of data and n is number of data. Feature vector $X = (x_1, x_2, x_3, ..., x_m)$ and $Y = (y_1, y_2, y_3, ..., y_m)$ consist of m features. For every features of x_i and y_i, i=1,2,..,m, feature i can be said as close if only if ${\sigma^2}_i$ has small variance, zero means that the features are similar or identical.

In biometrics, image data can be transformed into matrix of feature vector, then by calculating the variance for each features, we can calculate the discriminant capability of those features. By choosing the larger variance feature, we can reduce number of feature and process time.

Here we proposed a scheme to reduce the number feature automatically. Let's λ is array of variance that has been sorted in descending order $\lambda = (\lambda_1, \lambda_2, \lambda_3, ..., \lambda_n)$, where $\lambda_i > \lambda_{i+1}$. Delta between λ_i and λ_{i+1} can be calculated by equation 2.

$$\Delta\lambda_i = \lambda_i - \lambda_{i+1}, \; (i = 1, 2, 3, ..., n) \quad (2)$$

if $\Delta\lambda_i$ ¡ threshold, then i can be chosen as the index of boundary for feature reduction. In this case, threshold can be calculated by equation 3.

$$\text{threshold} = \frac{\text{mean}(\lambda)}{\text{maximum}(\lambda)} \quad (3)$$

The algorithm is written bellow.

Algorithm 1 Function autoReduction

```
% * I.S. array of variance λ
% * F.S. number dimension for reduction
threshold = mean(λ)/ max(λ);
n = length(λ);
for i = 1 to n − 1 do
  Δλi = (λ[i] − λ[i + 1])/λ[i];
  if Δλi < threshold then
    break;
  end if
end for
return i;
```

In this paper we proposed cascaded feature reduction, where there are two kind of feature reduction based on above ideas. The first reduction were done on DCT coefficient and the second one on eigenvalue of k-PCA.

2.2 Preprocessing

Region of Interest: In this step, Competitive Hand Valley Detection (CHVD) to extract region of interest in hand palm region proposed by Michael Goh, et al.[7] was adopted. Firstly, the images are segmented to separate the palm hand region from the background[7]. Then, to locate the ROI in palm hand, it is necessary to trace contour of hand, where a valley point is a point between two fingers. Detail description of this method can be consulted in ref[7].

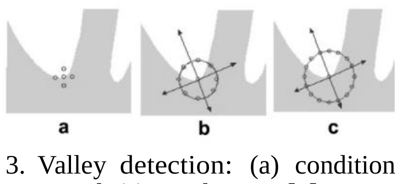

Figure 3. Valley detection: (a) condition 1; (b) condition 2; and (c) condition 3[7]

After the valley points were acquired, then the region can be drawn (see figure 4).

Histogram Equalization: Histogram equalization aims to change a picture in such a way as to produce a picture with flatter histogram[8]. The example of image can be shown by figure 5. In our scheme, this step is useful to spread out the distribution of feature.

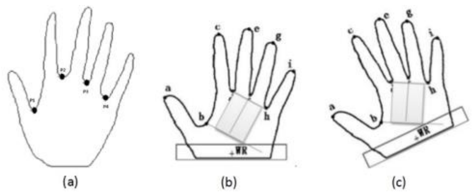

Figure 4. ROI detection: (a) valley points; (b) before rotation; (c) after rotation

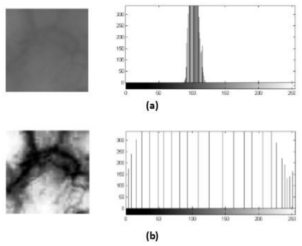

Figure 5. Example of histogram equalization: (a) original palm vein images and histogram; (b) image and histogram after equalization

2.3 Feature Extraction

First Feature Reduction: In this work, instead of using original palm vein image, we used DCT matrix from that image. We hypothesizes here that frequency feature can be more discriminative as has been proven from our experiments. To acquire feature vector from DCT matrix, DCT coefficients were scanned by zigzag scan (usually use for compression). If we look the position of the features as depicted in figure 6, the discriminant feature will be stacked on the left side. So, only

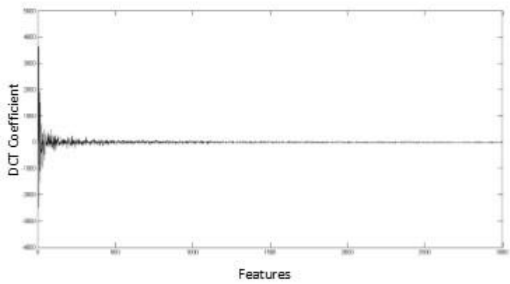

Figure 6. Example of DCT coefficient an image the

lower frequencies of DCT can provide discriminant features. If the variance for each feature of DCT matrices was calculated then almost all of variance will be ordered descending. So, the features which have lower variance can be discarded.

Based on figure 6, for feature reduction a half features were chosen. After that, the DCT features should be grouped into classes (by averaging data from same class), then the variance of those feature can be used for feature reduction by using the proposed approach described in subsection 2.1.

Second Feature Reduction: Based on the concept of principle component analysis, it states that it does not need to choose all eigen image for discrimination, just those eigenvector that corresponding with the largest eigenvalue[6]. So, for second reduction we assume the eigenvalue as variance of the data, and then we implement our proposed approach onto that eigenvalue.

2.4 Matching

In the matching process, euclidean distance were used for finding distance between testing sets feature and template feature. And then, the result will be decided by minimum distance $d_{Euc} = \sqrt{\sum_{i=1}^{d} |P_i - Q_i|^2}$ [12].

3 Experiment Result

Palm vein images which is used in this experiment are palm print images from CASIA Multi-Spectral Palm print Image Database V1.0 (http://www.cbsr.ia.ac.cn/english/index.asp). Those dataset consist of palm print images which contain vein pattern were captured by near-infrared light (NIR) (850nm and 940nm). Vein patterns can only be captured under the NIR with 800 until 1000nm wavelength[1]. Therefore, only palm print images with 850nm were selected. The selected data consists 600 of palm vein images (100 classes, 6 images for each class), so the total is 1,200 images (left and right hand) with size of 768 x 578 pixels (after preprocessing, the images were resize into 100 x 100 pixels). In this experiment, the data were divided into two groups: training set (4 images); and testing set (2 images) for each class. The experiment of left and right hand were done separately.

In the experiment, we compare our proposed method with manual one. For manual feature reduction, we trace all possible feature number from 1 until maximum features (400 features, after reduction by k-PCA) to find the optimal features that can provide the highest accuracy. The result, shown in figure 7, achieved the highest recognition rate at 0.9450 or 94.5% with 66 features for left hand and 0.940 or 94.0% with 64 features for right hand. The result of proposed method

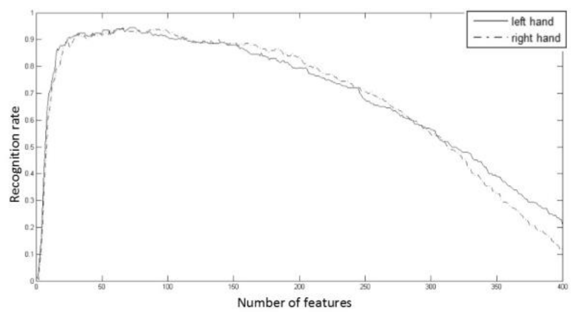

Figure 7. recognition rate both of hand based on number of features

is shown by table 1, which can reduce process time until 99%, and achieve recognition rate above 92%. A further analysis for experiment was made by calculating the False Rejection Rate (FRR); and False Acceptance Rate (FAR)[1]. In this schemes, we reduced training sample into 50 classes (200 images), with the testing sample similar with previous one, so there are

Table 1. Result of recognition rate

Method	Features*	Recognition rate	Time
N_{left}	Full & 66	94.5%	2548.8 s
N_{right}	Full & 64	94.0%	2516.4 s
P_{left}	1869 & 78	93.5%	7.5 s
P_{right}	5050 & 48	92.0%	11 s

Note:

- features* n & m : n features of DCT, and m feature of k-PCA
- N uses full DCT feature and manual selected k-PCA features at maximal accuracy.
- P uses our proposed method

50 classes as unknown in testing sample. The objective is to find the threshold that make FRR and FAR balance (known as Equal Error Rate(ERR)). The results were be shown in figure 8 and 9. So, the threshold that

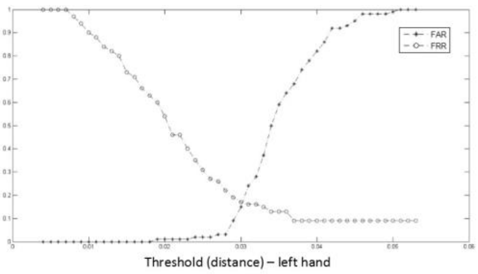

Figure 8. The FAR and FRR of the proposed method(left hand)

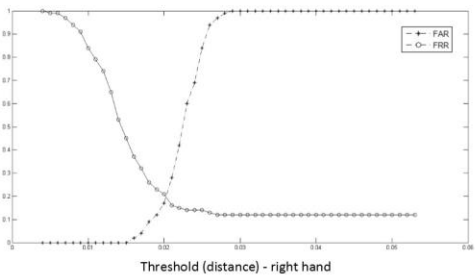

Figure 9. The FAR and FRR of the proposed method(right hand)

can be used to minimize recognition rate of impostor or unknown as known class are 0.0302 with ERR=0.18 for left hand and 0.0202 with ERR=0.195 for right hand.

4 Conclusion

In this paper, cascaded feature reduction based on variance order has been described. Experiment result shown that the method can reduce feature effectively on left and right hand, without reducing the recognition rate, up to 92%. Beside that, the process time also can be reduced significantly (until 99%) compared with the manual approach.

References

[1] A. K. Jain, P. Flynn, and A. A. Ross, Handbook of Biometrics. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2007.

[2] L. Xueyan and G. Shuxu, “The fourth biometric - vein recognition,” InTech, Pattern Recognition Techniques, Technology and Applications, 2008.

[3] A. Kong, D. Zhang, and M. Kamel, “A survey of palmprint recognition,” Pattern Recogn., vol. 42, pp. 1408 –1418, July 2009.

[4] Y. Zhou and A. Kumar, “Contactless palm vein identification using multiple representations,” in 2010 Fourth IEEE International Conference on Washington DC, ,, pp. 1–6, IEEE, Biometrics: Theory Applications and Systems (BTAS), September 2010.

[5] W. Yang, C. Sun, and L. Zhang, “A multi-manifold discriminant analysis method for image feature extraction,” Pattern Recogn., vol. 44, pp. 1649–1657, Aug. 2011.

[6] G. Lu, D. Zhang, and K. Wang, “Palmprint recognition using eigenpalms features,” Pattern Recogn. Lett., vol. 24, pp. 1463–1467, June 2003.

[7] G. K. Ong Michael, T. Connie, and A. B. Jin Teoh, “Review article: Touch-less palm print biometrics: Novel design and implementation,” Image Vision Comput., vol. 26, pp. 1551–1560, Dec. 2008.

[8] M. Nixon and A. Aguado, Feature Extraction and Image Processing. Elsevier, 2008.

[9] A. Chadha, P. P. Vaidya, and M. M. Roja, “Face recognition using discrete cosine transform for global and local features,” CoRR, vol. abs/1111.1423, 2011.

[10] B. Scholkopf, A. J. Smola, and K.-R. Müller, “Advances in kernel methods,” ch. Kernel Principal Component Analysis, pp. 327–352, Cambridge, MA, USA: MIT Press, 1999.

[11] K. Hon, An Introduction to Statistics. CreateSpace Independent Publishing Platform, 2014.

[12] S.-H. Cha, “Comprehensive survey on distance/similarity measures between probability density functions,” International Journal of Mathematical Models and Methods in Applied Sciences, vol. 1, no. 4, pp. 300–307, 2007.