Journal of Signal and Information Processing, **2013, 4, 173-175**
doi:10.4236/jsip.2013.43B031 Published Online August 2013 (http://www.scirp.org/journal/jsip)

Scientific Research

An Overview of Principal Component Analysis

Sasan Karamizadeh[1], Shahidan M. Abdullah[1], Azizah A. Manaf[1], Mazdak Zamani[1], Alireza Hooman[2]

[1]Advanced Informatics School(AIS), Universiti Teknologi Malayisa, Kuala Lumpur, Malaysia; [2]Faculty of Management (FOM), Multimedia University (MMU), Cyberjaya, Malaysia.
Email: Ksasan2@live.utm.my, mshahidan@ic.utm.my, azizah07@ic.utm.my, mazdak@utm.my, alireza_hooman@yahoo.com

Received May, 2013.

ABSTRACT

The principal component analysis (PCA) is a kind of algorithms in biometrics. It is a statistics technical and used orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables. PCA also is a tool to reduce multidimensional data to lower dimensions while retaining most of the information. It covers standard deviation, covariance, and eigenvectors. This background knowledge is meant to make the PCA section very straightforward, but can be skipped if the concepts are already familiar.

Keywords: Biometric; PCA; Eigenvector; Covariance; Standard Deviation

1. Introduction

Biometrics is derived from Greek .words "bio" meaning life and metrics meaning the term biometrics is derived from the Greek words bio meaning "life" and metrics meaning "to measure" [1]. Biometrics refers to the identification or verification of a person based on his/her physiological and/or behavioral characteristics [2]. Several verification and identification based biometrics have evolved based on various unique aspects of human body, ease of acquiring the biometric, public acceptance and the degree of security required [3].

Principal component analysis (PCA), also known as Karhunen-Loeve expansion, is a classical feature extraction and data representation technique widely used in the areas of pattern recognition and computer vision such as face recognition [4]. The strategy of the Eigenfaces method consists of extracting the characteristic features on the face and representing the face in question as a linear combination of the so called 'eigenfaces' obtained from the feature extraction process [5]. The principal components of the faces in the training set are calculated. Recognition is achieved using the projection of the face into the space formed by the eigenfaces [6]. A comparison on the basis of the Euclidian distance of the eigenvectors of the eigenfaces and the eigenface of the image under question is made [7]. If this distance is small enough, the person is identified [8]. On the other hand, if the distance is too large, the image is regarded as one that belongs to an individual for which the system has to be trained [9]. Principal component analysis is a statistics technical [10]. PCA used for reduce dimension vector to better recognize images [11]. PCA is a useful statistical technique that has found application in fields such as face recognition and image compression, and is a common technique for finding patterns in data of high dimension [12]. Before getting to a description of PCA, this tutorial first introduces mathematical concepts that will be used in PCA. It covers standard deviation, covariance, and eigenvectors [13]. This background knowledge is meant to make the PCA section very straightforward, but can be skipped if the concepts are already familiar [10, 14]. The basis of the eigenfaces method is the Principal Component Analysis (PCA). Eigenfaces and PCA have been used by Sirovich and Kirby to represent the face images efficiently [15, 16].

2. PCA Algorithm

Following are steps involve;

Step 1: Column or row vector of size N2 represents the set of M images (B1, B2, B3…BM) with size N*N

Step 2: The training set image average (μ) is described as

$$\mu = \frac{1}{m}\sum_{n=1}^{M} Bn \quad (1)$$

Step 3: the average image by vector (W) is different for each trainee image

$$Wi = Bi - \mu \quad (2)$$

Step 4: Total Scatter Matrix or Covariance Matrix is calculated from Φ as shown below:

$$C = \sum_{n=1}^{M} wnwnt = AAT , \quad (3)$$

where A= [W1W2W3…Wn]

Step 5: Measure the eigenvectors UL and eigenvalues λL of the covariance matrix C.

Step6: For image classification, this feature space can be utilized. Measure the vectors of weights

$$\Omega T = [w1, w2, \ldots, wM'], \quad (4)$$

whereby,

$$Hk = UkT (B - \mu), k = 1, 2, \ldots, M' \quad (5)$$

3. The Important of PCA in Face Recognition

The statistical information published in the area of facial recognition technology utilizing the PCA method reveals the significance of using this method for identifying and verifying facial features [8]. **Figure 1** below reveals the amount of publications that have used the words 'face recognition' and 'PCA' in their headings [17]

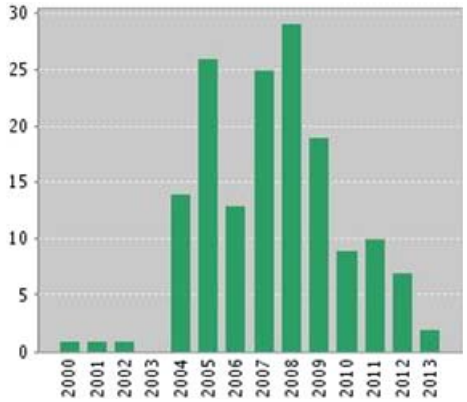

Figure 1. Number of publication utilizing [15].

Table 1 shows features about principal component analysis.

4. Advantage and disadvantage of PCA

PCA's key advantages are its low noise sensitivity, the decreased requirements for capacity and memory, and increased efficiency given the processes taking place in a smaller dimensions; the complete advantages of PCA are listed below:

1) Lack of redundancy of data given the orthogonal components [19, 20].

2) Reduced complexity in images' grouping with the use of PCA [19, 20]

3) Smaller database representation since only the trainee images are stored in the form of their projections on a reduced basis [19].

4) Reduction of noise since the maximum variation basis is chosen and so the small variations in the background are ignored automatically [19].

Table 1. The features of PCA are shown in the table below [18].

Feature	Principal component analysis
Discrimination between classes	PCA manages the entire data for the principal components analysis without taking into consideration the fundamental class structure.
Applications	PCA applications in the significant fields of criminal investigation are beneficial
Computation for large datasets	PCA does not require large computations
Direction of maximum discrimination	The directions of the maximum discrimination are not the same as the directions of maximum variance as it is not required to utilize the class information such as the within class scatter and between class scatter
Focus	PCA examines the directions that have widest variations
Supervised learning technique	PCA is an unsupervised technique.
Well distributed classes in small datasets	PCA is not as powerful as other methods.

Two key disadvantages of PCA are:

1) The covariance matrix is difficult to be evaluated in an accurate manner [19].

2) Even the simplest invariance could not be captured by the PCA unless the training data explicitly provides this information [4].

5. Conclusions

The PCA method is an unsupervised technique of learning that is mostly suitable for databases that contain images with no class labels. A detailed description of the PCA technique utilizing in face recognition has been provided. As mentioned above, the PCA method's advantages and disadvantages have also been explained in this study.

6. Acknowledgment

The work we presented in this paper has been supported by the University Technology Malaysia.

REFERENCES

[1] J. Ashok, V. Shivashankar and P. Mudiraj, "An Overview

of Biometrics," *International Journal,* Vol. 2, 2010.

[2] F. Ã–zen, "A Face Recognition System Based on Eigenfaces Method," *Procedia Technology*, Vol. 1, 2011, pp. 118-123.

[3] W. Miziolek and D. Sawicki, "Face Recognition: PCA or ICA," *Przeglad Elektrotechniczny*, Vol. 88, 2012, pp. 286-288.

[4] C. Li, Y. Diao, H. Ma and Y. Li, "A Statistical PCA Method for Face Recognition," *in Intelligent Information Technology Application*, 2008, pp. 376-380.

[5] H. Duan, R. Yan and K. Lin, "Research on Face Recognition Based on PCA," *in Future Information Technology and Management Engineering*, 2008, pp. 29-32.

[6] T. F. Karim, M. S. H. Lipu, M. L. Rahman and F. Sultana, "Face Recognition Using PCA-based Method," *in Advanced Management Science (ICAMS), 2010 IEEE International Conference on*, 2010, pp. 158-162.

[7] Z. Haiyang, "A Comparison of PCA and 2DPCA in Face Recognition," *in Electrical Power Systems and Computers: Springer*, Vol. 99, 2011, pp. 445-449. doi:10.1007/978-3-642-21747-0_55

[8] B. Poon, M. A. Amin and H. Yan, "PCA based Face Recognition and Testing Criteria," *in Machine Learning and Cybernetics, 2009 International Conference on,* 2009, pp. 2945-2949.

[9] Y. Wen and P. Shi, "Image PCA: A New Approach for Face Recognition," in Acoustics, Speech and Signal Processing, *IEEE International Conference on*, 2007, pp. 1241-1244.

[10] A. Boualleg, C. Bencheriet and H. Tebbikh, "Automatic Face Recognition Using Neural Network-PCA," *in Information and Communication Technologies*, 2006, pp. 1920-1925.

[11] Z. Wang and X. Li, "Face Recognition Based on Improved PCA Reconstruction," *in Intelligent Control and Automation (WCICA),* 2010 8th World Congress on, 2010, pp. 6272-6276.

[12] J. Meng and Y. Yang, "Symmetrical Two-Dimensional PCA with Image Measures in Face Recognition," *Int J Adv Robotic Sy,* Vol. 9, 2012.

[13] D. S. Pankaj and M. Wilscy, "Comparison of PCA, LDA and Gabor Features for Face Recognition Using Fuzzy Neural Network," *in Advances in Computing and Information Technology: Springer*, 2013, pp. 413-422.

[14] M. Safayani, M. T. Manzuri Shalmani and M. Khademi, "Extended Two-dimensional Pca for Efficient Face Representation and Recognition," *in Intelligent Computer Communication and Processing*, 2008. ICCP 2008. 4th International Conference on, 2008, pp. 295-298.

[15] L. Sirovich and M. Kirby, "Low-dimensional Procedure for the Characterization of Human Faces," Vol. 4, No. 3, 1987, pp. 519-524. doi:10.1364/JOSAA.4.000519

[16] M. Turk and A. Pentland, "Eigenfaces for Recognition," *Journal of Cognitive Neuroscience*, Vol. 3, No. 1, 1991, pp. 71-86. doi:10.1162/jocn.1991.3.1.71

[17] "Citation Report Title=(PCA face recognition)" Vol. 2013, 2013.

[18] K. S. Sodhi and M. Lal, "Face Recognition Using PCA, LDA and Various Distance Classifier," *Journal of Global Research in Computer Science*, Vol. 4, 2013, pp. 30-35.

[19] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min and W. Worek, "Overview of the Face Recognition Grand Challenge," *in Computer vision and pattern recognition,* 2005. CVPR 2005. IEEE Computer Society Conference on, 2005, pp. 947-954.

[20] D. Srinivasulu Asadi, Ch. DV Subba Rao and V. Saikrishna "A Comparative Study of Face Recognition with Principal Component Analysis and Cross-Correlation Technique," *International Journal of Computer Applications* Vol. 10, 2010.

