

Evaluation of Statistical Feature Encoding Techniques on Iris Images

Chowhan S.S.¹, G.N. Shinde²

Summary

Feature selection, often used as a pre-processing step to machine learning, is designed to reduce dimensionality, eliminate irrelevant data and improve accuracy. Iris Basis is our first attempt to reduce the dimensionality of the problem while focusing only on parts of the scene that effectively identify the individual. Independent Component Analysis (ICA) is to extract iris feature to recognize iris pattern. Principal Component Analysis (PCA) is a dimension-reduction tool that can be used to reduce a large set of variables to a small set that still contains most of the information in the large set. Image quality is very important in biometric authentication techniques. We have assessed the collision of various factors on performance of ICA and PCA as well as evaluated which factors can be plausibly compensated on iris patterns.

keywords: Biometrics, Component analysis, feature Extraction, Iris recognition

Introduction

High dimensional problems are becoming increasingly common. With high dimensional data, it is difficult to understand the underlying structure [1]: Additionally, the storage, transmission and processing of high dimensional data places great demands on systems. All these are aspects of one of the most interesting computational and data analysis problems. Iris feature extraction is the crucial stage of the whole iris recognition process for personal identification [1]. A brief survey is made firstly in this paper on the methods that feasibly implemented in iris feature extraction. Because the iris capture devices in use are mostly exposed to the natural scene, so the natural illumination or other variant conditions sometimes can greatly influence the iris Images captured and further impact the recognition result.

Iris Representation

In order to acquire the image of an eye with clear iris pattern, a Sony Video camera was connected to the build-in video capture card in our IBM PC machine. Image Analysis requires a lot of complex mathematics calculations, which is why Matlab, with its powerful mathematics, signal and image processing capabilities, made itself the first choice of platform for developing algorithm [2].

¹Corresponding author. COCSIT, Ambajogai Road, Latur-413512, (M.S.) India, Email: csantu_149@rediffmail.com

²Indira Gandhi College, CIDCO, Nanded-431602, (M.S.) India, Email: shindegn@yahoo.co.in, Tel: +91-2462-263788

Iris Localization

Iris localization is an especially important step in the whole iris recognition system. Only when we segment iris correctly from the original iris image, we can obtain an accurate matching result. Iris localization, by definition, means to detect the location of iris' inner and outer boundaries as indicated in fig 1.

The aim of iris localization is to locate iris accurately and fast. In practice, if an algorithm performs accurately, it always needs a long time to locate iris. Daugman proposed a coarse to fine localization method. First, it finds out the coarse centre and radius of pupil and iris [3]; Another classic iris localization method was proposed by Wildes[4]. Wildes' system performed its contour fitting in two steps. First, transform the original iris image to a binary edge-map, which is recovered via gradient-based edge detection. Second, use the Hough transform to detect the circular curve through the edge points. The algorithms proposed by Wildes and Daugman can localize iris correctly. However, the two methods cost a long time[5][6].

Defining Pupillary Iris Boundary

The first procedure in the whole image analysis is to define the pupillary iris boundary. The inner boundary of the iris, forming the pupil, can be accurately determined by exploiting the fact that the boundary of the pupil essentially a circular edge. The pupil is generally dark while the iris is lighter, with varied pigmentation [3]. When an image is imported into Matlab the image is firstly transformed into a 256 level grayscale image between 0 being black, and 255 being white Pupillary boundary is detected as an abrupt and sudden change in brightness when summed along a circle whose radius steadily increasing. This sudden change will be maximum when the circle has its center near the true center of the pupil, and when its radius matches the true radius of the pupil. Thus, the image processing problem of finding the pupil can be formulated as an optimization problem, in which a series of "exploding circles" with the center coordinates located at each one of the number of trail points on a grid [6].

The image is filtered using histogram equalization method then thresholding, by setting Intensity which is greater than intensity threshold a to 255, white, and intensity less than b to black, 0. Then a circle mask of arbitrary area of r^2 is applied only the dark mass centroid in the pupil are left. Then an estimated center is found on the final filtered image [6].

The system searches for the maximum rate of change in this quantity as the radius expand. For the candidate circle that best describes the papillary boundary, there will be a sudden "spike" in the rate -of-change of luminance summed around its perimeter, when its radius just matches that of the Papillary boundary.

Iris Normalization

The purpose of iris normalization is to get the same region of iris to do matching regardless pupil dilation and the different iris size caused by the different distance between the eye and video zoom factor.

Wildes used an image-registration technique to eliminate dilation, scaling and rotation. This approach wraps a new iris image into alignment with a normal image according to a mapping function so that the difference between the two iris images is minimal [4]. Daugman map image coordinates to polar image coordinates. The angular coordinate ranges between $0 \sim 2\pi$ and the radial coordinate ranges from the iris [7,10]. Li Ma et al gave a method to eliminate other problems in normalization part to improve the performance of match. Enhancement is adopted to eliminate the non-uniform brightness problem and the low intensity contrast problem. There are two steps in this part. One is the background illumination estimation and the other is histogram equalization [8][21].

The normalization algorithm always depends on the algorithm of feature vector extraction and match. Moreover, it should be made the texture on iris become clearer and eliminate the factors that will lead to error of match in iris normalization operation.

Principal Component Analysis (PCA)

Principle component analysis has been widely used for analyzing the image data. It is applied to biometric systems as a classification design. PCA enables us to measure the difference between two images while allowing expression changes. In the vector space, PCA identifies the major directions and corresponding strengths, of variation in the data. PCA performs these by computing the eigenvectors and eigenvalues of the covariance matrix of the image data. Keeping only a few eigenvectors corresponding to the largest eigenvalues, PCA can be also used as to reduce the size of the data while hold the major variation of data. For example, we may reduce the database size by using only first twenty eigenspace while our whole eigenvectors are 250[1] [2].

Principle Component Analysis (PCA), also known as Eigen-XY analysis is a standard statistical technique for finding directions of maximum variations in data. These directions, called the principle components, can be used to reconstruct all of the information within the set and can be tested to which level a test image couples with an image of the training set. Principle components with smaller associated magnitudes can often be omitted, as they contribute less to the overall reconstruction of each data element. This allows sufficient representation of the original data set with a reduced set of principle components. Principle components are found by computing the eigenvalues and eigenvectors of the covariance matrix associated with the data. Eigenvectors with largest associated eigenvalues are

the principle components that describe the most variation in the data set. Principle component coefficients of data elements are found by projecting each datum onto the eigenspace of the covariance matrix [3].

Linear Principal Components Analysis

Using the example of projecting data from two dimensions to one, a linear projection requires the optimum choice of projection to be a minimization of the sum-of-squares error. This is obtained first by subtracting the mean \bar{x} of the data set. The covariance matrix is calculated and its eigenvectors and eigen values are found. The eigen vectors corresponding to the M largest eigen values are retained, and the input vectors x^n are subsequently projected onto the eigenvectors to give components of the transformed.

Vectors z^n in the M -dimensional space. Retaining a Subset $M < d$ the basis vectors μ_i so that M coefficients z_i are used allows for replacement of the remaining coefficients by constants b_i this allows each x vector to be approximated by an expression of the form:

$$\tilde{x} = \sum_{i=1}^M z_i \mu_i + \sum_{i=M+1}^d b_i \mu_i \quad (1)$$

Where represents a linear combination of d orthonormal vectors.[4][5].

Simple PCA

Simple PCA (SPCA) is not strictly speaking a Hebbian algorithm, although it does have similarities, it is iterative, local and gets the principal components in order (largest eigenvalues first). Most importantly, it does not necessarily add a Hebbian term. Instead, SPCA considers other forms of ϕ . In particular, considers two different functions ϕ_1 and ϕ_2

$$\phi_1(y_i, x_k) = \begin{cases} x_k & \text{if } y_i \geq 0 \\ 0 & \text{other wise} \end{cases}$$

$$\phi_2(y_i, x_k) = y_i x_k$$

Working of PCA on Iris image

In order to use PCA for image processing firstly it is required to convert RGB picture gray. Eye image is two dimensional array of $x * y$, where x and y are width and height of image. Each image can be represented as vector dimension [6].

P be an image from the database, $P_1, P_2, P_3, \dots P_n$ indicates the vectors from each image. The average of images can defined as

$$X = \frac{1}{m} \sum_{n=1}^M P_n \quad (2)$$

After getting average column vector matrix it is required to find the difference matrix of each image which should differ from the average vector

$$\phi = \begin{bmatrix} \phi_1 & \phi_2 & \dots & \phi_n \\ \vdots & \vdots & & \vdots \end{bmatrix} \quad \phi = P_i - X \quad (3)$$

The Difference Matrix is obtained. Covariance is always measured *between 2* dimensions The Covariance Matrix can be defined as

$$CM = \frac{1}{M} \sum_{n=1}^M \phi_n \phi_n^T = v v^T \quad (4)$$

Where

$$v = [\phi_1 \ \phi_{21} \dots \phi_M] \quad (5)$$

To provide a background for the matrix algebra required in PCA. Specifically we use eigenvectors and eigenvalues of a given matrix. After getting Covariance Matrix multiply two matrices together, provided they are compatible sizes. Eigenvectors are a special case of this. It is obtained by

$$v v^T v_i = \mu_i v_i \quad (6)$$

Hence we construct a matrix $M \times M$

$$L = v^T v \quad (7)$$

The above equation (3.3) and (3.4) can be defined as

$$U_l = \sum_{k=1}^M V_{lk} \phi_k \quad (8)$$

Theses eigenvectors are called eigenfaces. Among those eigen faces are with higher values are most useful for. Therefore $M < M'$ eigen faces that are most significant for building the eye subspace for image projections. Which are used for iris identification classification, recognition. [9,20]

Principle Component Analysis (PCA) and Independent Component Analysis (ICA). PCA has been used as a preprocessing step that reduces dimensions for obtaining ICA components for iris.

Principal Component Analysis (PCA): Simple to implement and fast performance. PCA is based on linear mapping and eigenvectors. It's one of the most traditional methods around, also known as Karhunen-Loeve expansion. Independent Component Analysis: Blind source separation method, very efficient de-mixing

non-Gaussian distributed features [11][12]. It is also an iterative, linear mapping technique [13,16]. Although Independent Component Analysis (ICA) and its many flavors were originally conceived for blind source separation (BSS), many studies were conducted on the application of ICA to feature extraction from time series and images [13]. Redundant distributed data from natural signals such as natural sounds or images have a high degree of redundancy [17,19].

Conclusion

This paper describes a procedure for using PCA as feature extractor in the context of Iris recognition. The fast tracking capability is enabled by the recursive nature of the complete eigenvector matrix updates. Results on experiments performed on a standard database show increased performances with respect to the use of PCA only as feature extractor. Performance comparisons with traditional algorithms, as well as a structurally similar perturbation-based approach demonstrated the advantages of the recursive PCA algorithm in terms of convergence speed and accuracy. There are many successful algorithms for Preprocessing.

References

1. Seung-in-Noh, Kwanghyuk B, Kang Ryoung P, Jaihie Kim, A New Iris Recognition Method Using Independent Component Analysis Volume E88-D , Issue 11 (November 2005), Pages 2573-2581, Year of Publication: 2005.
2. Boaz Lerner, Hugo Guterman, Mayer Aladjem and Its'hak Dinstein, "A comparative study of neural network based feature extraction paradigms" University of Cambridge Computer Laboratory, New Museums Site, Cambridge 84105, Israel Received 16 October 1997; revised 9 September 1998. Available online 16 April 2003.
3. Daugman J G. High Confidence Recognition of Person by Rapid Video Analysis of Iris Texture[C]. European Convebtion on Security Detection and Brighton, Brighton, UK, 1995: 244-251.
4. Wildes R P. Iris Rognition: An Emerging Biometric Technology[C]. In Proceeding of the IEEE, Piscataway, NJ, USA , 1997, 85: 1 348-1 363.
5. J. G. Daugman, Two-dimensional spectral analysis of cortical receptive field profiles, Vision Res. vol. 20, pp. 847-856
6. Iris Recognition by Frankin Cheung Department of Computer Science and Electrical Engineering The University of Queensland
7. Daugman J G. Recognition persons by their pattern[J]. Information Security Technical report, 1998, 13(1): 33-39.
8. Cui J L, Wang Y H, Tan T N, et al. A Fast and Robust Iris Localization

- Method Based on Texture Segmentation[C]. In Proceeding of SPIE, Bellingham, WA, 2004, 5404: 401-408.
9. XIE Mei, Iris Recognition Technique, Journal of Electronic Science and Technology of China, vol .4. No.3 School of Electronic Engineering, University of Electronic Science and Technology of China Chengdu 610054 China.
 10. Daugman, J. "How Iris Recognition Works", available http://www.ncits.org/tc{_}home/m1htm/docs/m1020044.pdf
 11. M. Mudrov'a, A. Proch'azka Principal Component Analysis In Image Processing
 12. I. T. Jolliffe. *Principal component analysis*.Springer-Verlag, New York, 1986. <http://www.eyejudge.com/pca.asp>
 13. Rafael A. Calvo Matthew Partridge Marwan A. Jabri A *Comparative Study of Principal Component Analysis Techniques*, University of Sydney NSW, 2006, Australia.
 14. Bishop, C. M 1995 *Neural networks for pattern recognition*, Oxford University Press, pp 310-313.
 15. Press, W.H, S. A. Teukolsky, W.T Vetterling, and B.P Flannery 1992 Numerical recipes in C: The art of scientific computing (second edition), Cambridge University press.
 16. Dorairaj V, Schmid N A, Fahmy G. Performance Evaluation of Iris Based Recognition System Implementing PCA and ICA Encoding Techniques[C]. Biometric Technology for Human Identification II, Proceeding of SPIE, Bellingham, WA, 2004: 51-58.
 17. R. Swiniarski, *Independent Component Analysis*, CS553 Class Notes, San Diego State University, 2004.
 18. Gonzalez, R.C., Woods, R.E, *Digital Image Processing*, 2rd ed., Prentice Hall (2002).
 19. Chowhan.S.S and G.N.Shinde *Iris Biometric Recognition Application in Security Management*, CISP/2008/Vol 1, pp-661-665.
 20. Vijaykumar Bhagavatula *Spatial Frequency domain methods for face and iris recognition*, Dept. of Electrical and Computer Engineering, Pittsburg.

