The Thirty-Third International
FLAIRS Conference (FLAIRS-33)

Variance Linear Discriminant Analysis for IRIS Biometrics

Sung-Hyuk Cha
Computer Science department, Pace University
1 Pace plaza, New York, NY 10038
scha@pace.edu

Teryn Cha
Essex County College
303 University Ave, Newark, NJ 07102
yan@essex.edu

Abstract

Dichotomy transformation in biometric authentication problem creates a two class ("within" or "between") classification problem in multivariate distance space. Linear discriminant analysis, which is a linear classifier, results in good performance in IRIS biometric authentication problem. However, it assumes that the distributions of two classes are normal, whereas they are closely related to the log-normal distributions. Here a modified variance linear discriminant analysis algorithm is proposed and its superior experimental results on the IRIS biometric database are reported.

Introduction

Establishing the individuality of biometrics is of great importance in cybersecurity and computational forensics. The *dichotomy transformation model*, which is a statistically inferable methodology, was first introduced to establish the individuality of handwriting in (Cha and Srihari 2000b). Studies on establishing the Individuality of Fingerprints and IRIS were conducted in (Pankanti, Prabhakar, and Jain 2002) and (Yoon et al. 2005), respectively.

The problem of reducing dimensionality is one of the fundamental problems in machine learning and pattern recognition. It is also considered in biometrics such as in (Tantawi et al. 2013). If the dimension is reduced, the better results are often derived because a very high dimensional space often leads to a curse of dimensionality. The dichotomy transformation model often involves high dimensional multivariate distance space and thus, the dimensionality reduction is necessary.

Two of the most popular techniques for dimensionality reduction are Principal Components Analysis (PCA) and Linear Discriminant Analysis (LDA). The main objective of both PCA and LDA is to reduce dimensionality. While PCA ranks new axes by maximizing variance of data, LCA ranks new axes by maximizing the ratio of the between-class variance to the within-class variance. Instead of the covariance matrix in PCA, LCA utilizes a scatter matrix that maximizes class separability. PCA was utilized in face recognition (Belhumeur, Hespanha, and Kriegman 1997) and LDA has been also utilized in face recognition in (Belhumeur, Hespanha, and Kriegman 1997; Chen et al. 2000).

LDA which assumes that samples in classes are normally distributed. Within and between distance classes in a dichotomy model do not follow the Gaussian distributions. Hence, new variation of LDA that is suitable for the dichotomy transformation model is proposed. A new scatter distance matrix is utilized. In this paper, IRIS biometirc data is used to demonstrate the effectiveness of the proposed system, but it can be also generalized to many other biometric authentication systems.

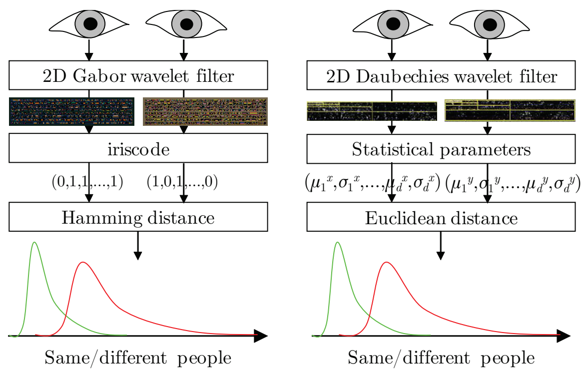
The rest of this paper is organized as follows. Section II reviews the dichotomy transformation model and presents how the model can be effectively applied to the IRIS biometrics. Next, section III reviews the linear discriminant analysis. Section IV introduces a new scatter matrix for the linear discrimination analysis for the dichotomy transformation model. Experimental results on the iris biometric authentication are also reported in section V. Finally, Section VI concludes this work.

Dichotomy transformation model

This section reviews the dichotomy transformation model, which was first introduced in (Cha and Srihari 2000b). The IRIS biometric authentication system is used to illustrate the model. In order to visualize the decision boundary, hypothetical two dimensional data samples are used as well.

There are two fundamental models in biometrics: identification and verification (Cha and Srihari 2000b). While the identification model is a many class classification (polychotomy) problem, the verification model is a two class (dichotomy) problem. The identification model involves the feature space domain. Consider the many-class problem where the number of classes (individuals) is too large to be completely observed, such as the population of a country. Most biometric identification problems fall under the aegis of the many-class problem. For this reason, it was shown that the verification model is clearly more suitable than the identification model for establishing the individuality of a biometric modality (Cha and Srihari 2000b).

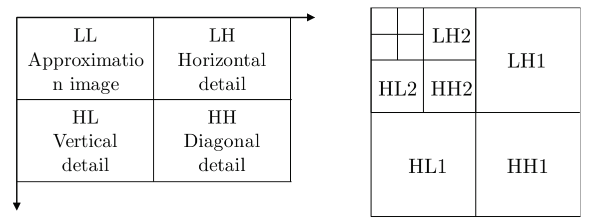
The biometric verification or authentication problem is whether two biometric samples are from the same person or two different people and is a two (either *'within'* or *'between'*) class classification problem. Let $s(x)$ denote the subject identity of the biometric sample x. If two randomly selected biometric samples are from the same subject, the



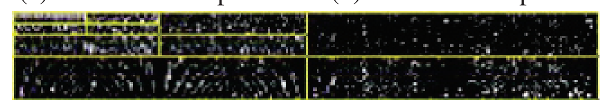
(a) Iriscode

(b) 3 level wavelet statistical parametersl

Figure 1: Two simple match models for IRIS biometrics.



(a) wavelet decomposition (b) 3 level decomposition



(c) 3 level wavelet decomposition

Figure 2: The 2D Daubechies wavelet transformation.

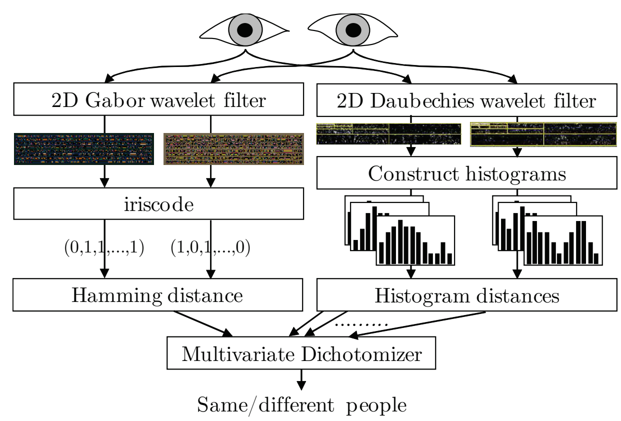


Figure 3: Dichotomy model for IRIS biometrics

scalar distance value between them belongs to the *within* class (intra-person), W, as defined in (1). If they are from two different subjects, it belongs to the *between* class (inter-person), B, given in (2).

$$W = \{d(x,y)|s(x) = s(y)\} \quad (1)$$

$$B = \{d(x,y)|s(x) \neq s(y)\} \quad (2)$$

A simple distance based biometric match model utilizes a certain distance measure between two biometric data and the scalar distance value is classified based on the threshold value t as defined in (3) on the belief that the within-class distance tends to be smaller than the between-class distance.

$$c(x,y) = \begin{cases} w & \text{if } d(x,y) \leq t \\ b & \text{otherwise} \end{cases} \quad (3)$$

A typical conventional biometric verification model is the distance-based simple match (SM). A certain proximity measure is applied to generate two scalar valued distance distributions.

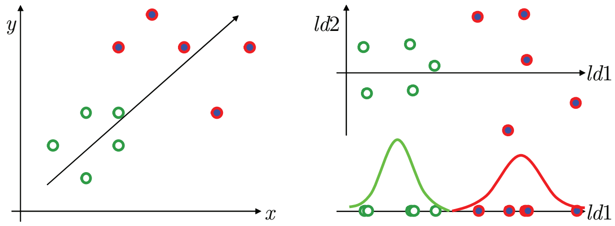
The dichotomy transformation model involves the multivariate feature distance space. The original feature space is transformed to a feature distance space. For example, an intra-person distance, W (within), and an inter-person distance, B (between) correspond to the points W and B in the feature distance space, respectively. Thus, there are only two categories: intra-person distance and inter-person distance in the feature distance space. When artificial neural networks or support vector machines are used as dichotomizers, much higher accuracy than simple match models were reported in handwriting (Cha and Srihari 2000b) and iris (Yoon et al. 2005).

There are three main steps in the SM procedure: feature extraction, applying a proximity measure, and a statistical performance evaluation on decision. In IRIS biometric simple match models, 2D Gabor wavelet filter was used to extract iriscode, which is a 256 binary feature vector in (Daugman 1993). Hamming distance was used to get the scalar distance value and this model is illustrated in part of Figure 1 (a).

The multi-level 2D wavelet decomposition technique was introduced in (Mallat 1989) and it has been widely applied to extract statistical parameters in IRIS biometrics in (Kee et al. 2001; Ma et al. 2003), as illustrated in Figure 1 (b).

The hierarchical wavelet transform decomposes the original iris image into a set of frequency windows having narrower bandwidths in the lower frequency region. Decomposing images with the wavelet transform yields a multi-resolution from detailed images to approximation images in each level. LH, HL, and HH represent detailed images for horizontal, vertical, and diagonal orientation, respectively, as shown in Figure 2 (a). Sub-image LL corresponds to an approximation image that is further decomposed, resulting in a two-level wavelet decomposition. The result of a three-level decomposition is shown in the lower-left portion of Figure 2 (b) and (c).

In (Yoon et al. 2005), histograms, instead of extracting statistical parameters, are constructed and histogram distance measure in (Cha and Srihari 2002) is directly used to transform into the distance space, as illustrated in Figure 3. Figure 3 is given so that the dichotomy transformation model is excellent to combine multiple classifiers and features can be heterogeneous as stated in (Cha and Srihari 2000a). Any IRIS biometric authentication system can be augmented easily to the dichotomy model.



(a) Original spacel (b) $ld1$ and $ld2$ space

Figure 4: LDA illustration.

Linear Discriminant Analysis

In this section, Linear Discriminant Analysis, or simply LDA, is briefly reviewed. LDA with only two classes is considered, albeit it can be easily generalized to the multiple class case. Consider a sample two dimensional data with two classes in Figure 4 (a). LDA finds and ranks new axes by rotating the data such the first new axis ($ld1$) has the highest separation between two classes. Figure 4 (b) shows the rotated space by LDA and when projected into the $ld1$ axis, two classes are best separated.

LDA requires defining the scatter $(d \times d)$ square matrix, Z, that determines the separability between two classes. The typical scatter matrix is the ratio between the *between-class variance* (S_B) and the *within-class variance* (S_W). Let C_1 and C_2 be sets of samples in classes 1 and 2, respectively. Let n_1 and n_2 be the sizes of each sets: $n_1 = |C_1|$ and $n_2 = |C_2|$. Let $\mu(C_1)$ and $\mu(C_2)$ be the d-dimensional means for all samples in the classes, C_1 and C_2, respectively. The *between-class variance* (S_B) is $(d \times d)$ square matrix and defined as follows:

$$S_B = |\mu(C_1) - \mu(C_2)|^T |\mu(C_1) - \mu(C_2)| \tag{4}$$

The further apart between two class means, the better separation between two classes.

Let cov(C_x) be the the covariance matrix for the class C_x, which is a $(d \times d)$ square matrix. The *within-class variance* (S_W) is $(d \times d)$ square matrix and defined as follows:

$$S_W = \frac{\text{cov}(C_1)}{n_1} + \frac{\text{cov}(C_2)}{n_2} \tag{5}$$

The lower variance within each class, the better separation between two classes. Let Z be the scatter matrix of the LDA.

$$Z = S_W^{-1} S_B \tag{6}$$

The eigenvalues, λ, and eigenvectors, V, of Z can be computed if S_W is non-singular. Eigenvectors in V are sorted by the eigen values in descending order. Original data are transformed into the linear discriminant axes when they are multiplied by V', the sorted eigenvectors.

If the original data is in d-dimensional space and first k LDA axes are selected, where $k < d$, LDA lowers dimensionality while the class separability is maximized. When data are projected onto a single axis, Z fisher's criteria is closely related to the z-test in hypothesis testing in eqn (7).

$$z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \tag{7}$$

LDA for dichotomy transformation model

In this section, a variation of LDA is introduced for the dichotomy transformation model. The dichotomy transformation model of a biometric verification problem is multivariate (d-dimensional) two-class classification problem, where two classes are within (or intra)-distance and between (or inter)-distance. Since the dimension of the distance space is quite large, the reduction is often required.

LDA requires several assumptions (see (Büyüköztürk and Çokluk Bökeoğlu 2008) for the list of assumptions). One of them is that samples in classes are normally distributed However, within and between distance classes in a dichotomy model do not follow the Gaussian distributions, but seem to follow log-normal distributions. Directly applying the standard LDA to this two class classification problem does not yield good performance. A new criteria for these distributions is needed.

The intra-person distance distribution is clustered toward the origin, whereas the inter-person distance distribution is scattered and away from the origin. Utilizing the fact that the intra-person distance is smaller than inter-person distance, a dichotomizer can be designed to establish the decision boundary between the intra and inter-person distances. The goal of the new scatter matrix in equation (8), instead of Z matrix in the ordinary LDA, is to minimize the intra-person distance while maximizing the inter-person distance.

$$F = (W^T W)^{-1} (B^T B) \tag{8}$$

As before, the eigenvalues, λ, and eigenvectors, V, of F can be computed if $(W^T W$ is non-singular. Eigenvectors in V are sorted by the eigen values in descending order. Original data are transformed into the linear discriminant axes when they are multiplied by V', the sorted eigenvectors. This variation of LDA shall be referred to vLDA to distinguish it from the ordinary LDA.

The $(d \times d)$ square scatter matrix, F can be realized as a ratio between covariances of two classes. Let $W' = W \cup -W$ and $B' = W \cup -B$. Negated data are added to the within and between distance sets. Means for both W' and B' are the same, the origin. The scatter matrix, F can be redefined as follows:

$$F = \text{cov}(W')^{-1} \text{cov}(B') \tag{9}$$

When data are projected onto a single axis, F criteria is closely related to the F-test of the equality of two variances in hypothesis testing in eqn (7).

$$F = \frac{\sigma_B^2}{\sigma_W^2} \tag{10}$$

The letter, F if F-test of the equality of two variances is coined in honour of Sir Ronald A. Fisher who studied the variance ratio in the 1920s.

vLDA for IRIS biometrics

In this section, experimental results on the IRIS biometrics using the vLDA are reported. Figure 5 depicts the dichotomy transformation model with the dimension reduction by a

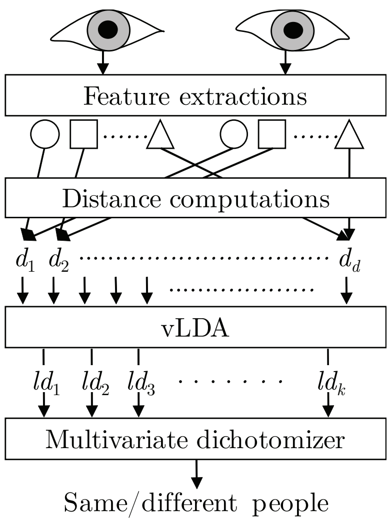


Figure 5: Dichotomy transformation model for IRIS biometrics with vLDA

		Accuracy	FRR	FAR
vLDA	Training	95.4%	5.4%	3.0%
	Test 1	93.8%	9.0%	3.4%
	Test 2	93.7%	10.6%	2.0%
	Test 3	94.5%	7.8%	3.2%
ANN 12 full dimension	Training	99.2%	1.0%	0.4%
	Test 1	96.1%	5.0%	2.8%
	Test 2	96.1%	5.2%	2.6%
	Test 3	96.7%	4.8%	1.8%
ANN + vLDA 8 dimension	Training	99.1%	1.2%	0.6%
	Test 1	97.7%	4.4%	0.2%
	Test 2	97.0%	5.4%	0.6%
	Test 3	97.7%	3.4%	1.2%

Table 1: Experimental results on IRIS biometrics

proposed vLDA. The result of the dichotomy transformation is multivariate distance vector and its dimensionality can be very high. Hence, vLDA is used to reduce the dimension.

The IRIS biometric database of twelve histogram distances is extracted from the 3 level wavelet decomposition, which was studied in (Yoon et al. 2005), is used in this experiment. Cardinalities of W and B are 500 in all training and three independent testing sets. When the Bayesian decision classifier is used in the two distributions projected on the best vLDA axis, the results are given in Table 1 under vLDA rows. Although low, the performance is quite good considering that only single feature is used.

As the multivariate dichotomizer, an *Artificial Neural Network*, or simply *ANN*, is used. ANN 12 full dimension in Table 1 is the case when an ANN is trained using the original full twelve distance data. ANN + vLDA 8 dimension in Table 1 is the case when an ANN is trained using the first eight vLDA's. When first eight vLDA's are used, the performances on the testing sets are better than those without the dimensionality reduction on average.

Conclusion

This paper proposes a new scatter matrix when the LDA is used in the dichotomy transformation model in biometric authentication problems. The effectiveness of the newly proposed model, vLDA, is demonstrated using the IRIS biometric authentication. The performance was better when the dimensionality was reduced.

References

Belhumeur, P. N.; Hespanha, J. P.; and Kriegman, D. 1997. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Tran. on Pattern Analysis and Machine Intelligence* 19(7):711–720.

Büyüköztürk, S., and Çokluk Bökeoğlu, O. 2008. Discriminant function analysis: Concept and application. *Egitim Arastirmalari - Eurasian Journal of Educational Research* 8:73–92.

Cha, S.-H., and Srihari, S. N. 2000a. Multiple feature integration for writer verification. In *7th Int'l Workshop on Frontiers in Handwriting Recognition*, 333–342.

Cha, S.-H., and Srihari, S. N. 2000b. Writer identification: Statistical analysis and dichotomizer. *LNCS - Advances in Pattern Recognition* 1876:123–132.

Cha, S.-H., and Srihari, S. N. 2002. On measuring the distance between histograms. *Pattern Recognition* 35(6):1355–1370.

Chen, L.-F.; Liao, H.-Y. M.; Ko, M.-T.; Lin, J.-C.; and Yu, G.-J. 2000. A new lda-based face recognition system which can solve the small sample size problem. *Pattern Recognition* 33(10):1713–1726.

Daugman, J. G. 1993. High confidence visual recognition of persons by a test of statistical independence. *IEEE Tran. on Pattern Analysis and Machine Intelligence* 15(11):1148–1161.

Kee, G.; Byun, Y.; Lee, K.; and Lee, Y. 2001. Improved techniques for an iris recognition system with high performance. In *Advances in Artificial Intelligence*, 177–188. Springer Berlin Heidelberg.

Ma, L.; Tan, T.; Wang, Y.; and Zhang, D. 2003. Personal identification based on iris texture analysis. *IEEE Tran. on Pattern Analysis and Machine Intelligence* 25(12):1519–1533.

Mallat, S. G. 1989. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Tran. on Pattern Analysis and Machine Intelligence* 11(7):674–693.

Pankanti, S.; Prabhakar, S.; and Jain, A. K. 2002. On the individuality of fingerprints. *IEEE Tran. on Pattern Analysis and Machine Intelligence* 24(8):1010–1025.

Tantawi, M. M.; Revett, K.; Salem, A.; and Tolba, M. F. 2013. Fiducial feature reduction analysis for electrocardiogram (ecg) based biometric recognition. *Journal of Intelligent Information Systems* 40(1):17–39.

Yoon, S.; Choi, S.-S.; Cha, S.-H.; Lee, Y.; and Tappert, C. C. 2005. On the individuality of the iris biometric. In *Image Analysis and Recognition: Second International Conference, ICIAR*, 1118–1124. Springer Berlin Heidelberg.