

# 主成分分析法在掌纹图像识别中的应用<sup>①</sup>

蔡平胜<sup>1</sup> 闫乐林<sup>1,2</sup> (1.山东教育学院 计算机科学与技术系 山东 济南 250013;

2.北京邮电大学 信息与通信工程学院 北京 100876)

**摘要:** 掌纹识别技术是生物特征识别领域的又一新兴技术,在网络安全、身份鉴别等方面有广阔的应用前景。将主成分分析法应用于掌纹图像的特征提取,阐释了传统主成分分析与加权主成分分析在处理掌纹图像时的差异,并在不同数据库上对两种方法进行了实验,结果表明传统主成分分析比加权主成分分析有更高的识别率以及加权主成分分析能够削弱光照对识别结果的影响。

**关键词:** 生物特征识别;掌纹识别;主成分分析;加权主成分分析

## Application of Principal Component Analysis to Palmprint Images Recognition

CAI Ping-Sheng<sup>1</sup>, YAN Le-Lin<sup>1,2</sup> (1.Department of Computer Science and Technique, Shandong Institute of Education, Jinan 250013, China; 2.Information & Telecommunication Engineering School, Beijing University of Posts and Telecommunications, Beijing 100876, China )

**Abstract:** The palmprint recognition is a new biometric technology, which has a good prospect of applications in the areas of network security, identity authentication etc. In this paper, the principal component analysis method is applied to palmprint image feature extraction, and the differences between the traditional principal component analysis(PCA) and the weighted principal component analysis(WPCA) in addressing the palmprint image are explained. According to the experimental results of two methods on two databases, PCA has a higher precision of palmprint recognition than WPCA, and the affection of light condition is weakened by WPCA.

**Keywords:** biometric recognition; parmprint recognition; PCA; WPCA

## 1 引言

计算机辅助身份识别在当今信息社会变得越来越重要,生物特征识别技术是其中一种可靠且有效的身份识别方法。与其它生物特征相比,掌纹识别有以下优点: 1)图像分辨率低; 2)冒犯性低; 3)线特征稳定; 4)用户可接受程度高。基于以上几点,掌纹识别在近年得到了较快发展。

主成份分析(PCA)又称 K-L(karhunen-loeve),其变换本质是利用低维的子空间去近似一个向量或图

像的方法<sup>[1]</sup>。该方法通常是利用最小均方误差准则(MSE)来获取最优的子空间,它的优点是在充分保留有用信息的基础上,能有效的降低原有特征向量的维数,因此已经被广泛应用到生物特征识别技术领域。

在文献[2]中,提出了将 PCA 用于掌纹识别的方法,并取得较不错的识别率。在文献[3]中,将 PCA 用于了人脸识别,但文献[4]却指出去掉三个最大的特征值会取得更好的识别效果。区别与此,本文提出了一种对掌纹图像的特征向量进行加权处理的方法,

<sup>①</sup> 基金项目:国家自然科学基金(60743007);山东省教育厅自然科学基金(J07WJ16)

收稿时间:2010-03-15;收到修改稿时间:2010-05-01

不去掉某些特征向量，而对其进行平衡处理。实验结果表明该方法的有效性和局限性。

## 2 掌纹图像的传统PCA算法

掌纹图像在计算机中逻辑上常以矩阵方式存储和描述，针对掌纹图像的主成分分析首要的任务的是将图像矩阵转化为一个图像向量，通常的做法是将图像矩阵的各列依次相接，即第  $N$  列的头接第  $N-1$  列的尾，这样便将一个图像矩阵转变成一个图像向量<sup>[5]</sup>。设变量  $x$  有  $N$  个样本  $x_k (k=1, 2, \dots, N)$ ， $x_k (k=1, 2, \dots, N)$  即假设训练库中  $k$  幅掌纹图像，将每幅图像按上述方法转换为列向量，其均值矢量记为  $\bar{x}$ ，则总体散布矩阵可表示为：

$$S_T = \frac{1}{N} \sum_{k=1}^M (x_k - \bar{x})(x_k - \bar{x})^T \quad (1)$$

令  $Z = [x_1 - \bar{x}, x_2 - \bar{x}, \dots, x_N - \bar{x}]$ ，则有：

$$S_T = \frac{1}{N} ZZ^T \quad (2)$$

令矩阵  $W = [w_1, w_2, \dots, w_n]$ ，其中  $w_1, w_2, \dots, w_n$  为  $S_T$  的特征向量。则用矩阵  $W$  对样本矢量进行的变换，称为  $K-L$  变换，其中  $w_1, w_2, \dots, w_n$  称为  $K-L$  变换轴。若要求数据变换后的维数为  $d$  维，则选用前  $d$  个最大特征值对应的特征向量  $w_1, w_2, \dots, w_d$  作为变换轴得出的矢量  $X$  即是  $x$  的所有  $d$  维描述中均方误差最小者，其中  $X$  由下式定义

$$X = W_{pca}^T x W_{pca} \quad (3)$$

其中， $W_{pca} = [w_1, w_2, \dots, w_d]$

因此，PCA 方法是一种最优维数压缩技术。可以证明，上述变换得出的数据中各分量之间是互不相关的，因此，也称 PCA 方法是一种去相关变换方法。

利用转换矩阵  $W_{pca}$ ，可将掌纹图像形成的列向量变换至 PCA 空间中，然后在该空间中利用简单的欧氏距离匹配方式获取最为相似的掌纹图像对，达到掌纹识别的目的。

在 PCA 空间中，一个特征向量和其均值向量之间的欧氏距离平方的统计平均可以由下面的式子得出：

$$E\{D_E^2(X, \bar{x})\} = E\{(X - \bar{x})^T (X - \bar{x})\} = Tr(C) = \sum_i I_i \quad (4)$$

式中  $D_E$  为欧氏距离， $I_i$  为协方差矩阵  $C$  的特征值。

可以看出，特征向量中每个分量对距离平方的贡献就是相应的特征值。因此，少数几个对应于大的特征值但对于分类并不重要的分量可能会主导 PCA 空间的相似性测度的计算，从而减小或彻底淹没那些对应小的特征值但对于分类更为重要的分量。

## 3 掌纹图像的传统PCA算法

将掌纹图像通过转换矩阵投影到 PCA 空间，为了平等看待该空间中的每一个特征分量，即那些互不相关的对分辨掌纹图像起作用的特征向量，可以对传统的使用原有转换投影矩阵张成的 PCA 空间进行加权处理，令：

$$Y_i = \text{diag}(I_1^{-1/2}, I_2^{-1/2}, \dots, I_m^{-1/2}) X_i = \Lambda^{-1/2} X_i \quad (i=1, 2, \dots, N) \quad (5)$$

这样就将传统 PCA 空间变换到了加权 PCA (WPCA) 空间，且其空间中特征向量的协方差矩阵  $C'$  为：

$$\begin{aligned} C' &= \frac{1}{N} \sum_{i=1}^N Y_i Y_i^T = \frac{1}{N} \sum_{i=1}^N \Lambda^{-1/2} X_i X_i^T \Lambda^{-1/2} \\ &= \frac{1}{N} \sum_{i=1}^N \Lambda^{-1/2} W^T (x_i - \bar{x})(x_i - \bar{x})^T W \Lambda^{-1/2} \\ &= \Lambda^{-1/2} W^T C W \Lambda^{-1/2} = I \end{aligned} \quad (6)$$

因此，WPCA 空间中的特征向量不仅相互正交，而且具有单位方差。这样，可以选取 WPCA 空间的变换矩阵为：

$$V = \Lambda^{-1/2} W = \text{diag}(I_1^{-1/2}, I_2^{-1/2}, \dots, I_m^{-1/2}) W \quad (7)$$

这里假设取前  $m$  个较大的特征向量，同时，也可以看出，WPCA 与具有传统 PCA 相同的计算复杂度<sup>[6]</sup>。

原掌纹图像可以从 WPCA 空间中进行如下估计：

$$\hat{X}_i = W \Lambda^{-1/2} X_i + \bar{x}, (i=1, 2, \dots, m) \quad (8)$$

从而得出估计误差为：

$$\begin{aligned} \Delta X_i &= X_i - \hat{X}_i = W(n) \Lambda^{1/2}(n) X_i(n) - W(m) \Lambda^{1/2}(m) X_i(m) \\ &= \sum_{k=m+1}^n X_i(k) I_k^{1/2} j_k \quad (i=1, 2, \dots, N) \end{aligned} \quad (9)$$

其中  $j_k$  为  $I_k$  所对应特征向量，则其均方误差为：

$$e^2 = E\{\|\Delta X_i\|^2\} = \sum_{k=m+1}^n I_k \quad (10)$$

由此可见, WPCA 具有与传统 PCA 相同的重建误差, 同时, 也应选择  $m$  个最大的特征值对应的  $m$  个特征向量构成变换矩阵。

假设  $\bar{Y} \in R^m$  是 WPCA 空间中的均值向量, 那么任一特征向量与该均值向量的欧氏距离的平方为:

$$D_E^2(Y, \bar{Y}) = (Y - \bar{Y})^T (Y - \bar{Y}) = (Y - \bar{Y})^T C^{-1} (Y - \bar{Y}) = D_M^2(Y, \bar{Y}) \quad (11)$$

(式中  $D_E$  和  $D_M$  分别表示欧氏和马氏距离)

因此, 在 WPCA 空间中, 马氏距离和欧氏距离是相等的。同时, 还有:

$$\begin{aligned} D_E^2(Y, \bar{Y}) &= (X - \bar{X})^T \Lambda^{-1/2} \Lambda^{-1/2} (X - \bar{X}) \\ &= (X - \bar{X})^T C^{-1} (X - \bar{X}) \\ &= D_M^2(X, \bar{X}) \end{aligned} \quad (12)$$

即在 WPCA 空间中按最小欧氏距离分类, 如同在传统 PCA 空间按最小马氏距离分类, 而后者效果优于前者的。

通过对传统 PCA 空间进行加权处理, 得到 WPCA 空间, 使用新转换投影矩阵将掌纹图像变换至 WPCA 空间, 然后用欧氏距离进行分类, 这样能取得按马氏距离分类的效果。

#### 4 仿真实验

首先选用一较小的掌纹库(掌纹库 I)进行验证, 该库由 50 个人掌纹图像构成, 每个人有 5 张图像, 且为同为右手掌, 切割后形成大小为  $128 \times 128$  像素的感兴趣区域(256 灰度级)<sup>[7]</sup>。分别选取每个人的幅图像做训练样本, 其余的  $5 - n$  幅图像用做识别样本, 识别结果如表 1 所示。

表 1 掌纹库 I 中 PCA 与 WPCA 识别结果

n	1	2	3	4
训练样本数量(幅)	50	100	150	200
样本识别数量(幅)	200	150	100	50
PCA 识别率(%)	60	60.67	74	92
WPCA 识别率(%)	64	62	75	94

为验证算法的普遍性, 本文采用来自香港理工大学的 PolyU Palmprint Database 为掌纹库 II。与上一掌纹库相比, 该库样本数量更多, 且由于手掌是放在仪器的固定位置处, 这样避免了掌纹选取时因角度不同和离摄像头远近而引起的比例不一致等问题。实验结果如图 1 - 图 3 所示。

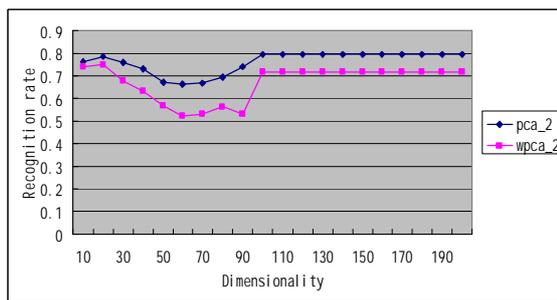


图 1 掌纹库 II 中 PCA 和 WPCA 的识别结果(训练样本数为 2)

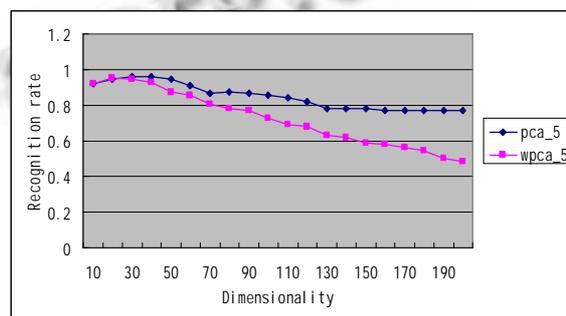


图 2 库 II 中 PCA 和 WPCA 的识别结果(训练样本数为 5)

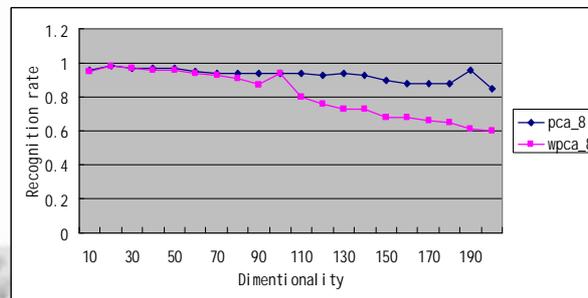


图 3 库 II 中 PCA 和 WPCA 的识别结果(训练样本数为 8)

从上述图表中, 我们可以看出, 在掌纹库 II 中, 无论是在小样本(样本数为 2), 还是在大样本(样本数为 8)情况下, WPCA 的识别率都不如 PCA 的识别率高。并且随着选取的主成份分量数目的增加, PCA 以及 WPCA 的识别率均不再增加或有所下降。

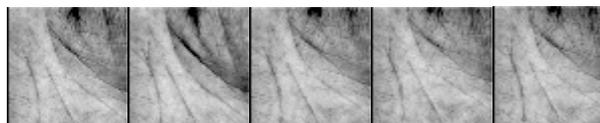


图 4 掌纹库 I 中的手掌图像(同一人)



图5 掌纹库 II 中的手掌图像(同一人)

比较图4和图5的掌纹库 I 和 II 手掌图像,可以看出掌纹库 I 比掌纹库 II 模糊,其原因是由于掌纹库 I 的采集条件比掌纹库 II 要差,从而导致了库 I 中掌纹图像受光照影响较大。也正是基于此原因, WPCA 在掌纹库 I 中的表现性能要强于 PCA,因为它平均了各个特征值及特征向量,使得光线对识别结果的影响大大降低了,从而取得了较好的识别率。而在掌纹库 II 中,由于其固定采集掌纹图像,采集条件较好,抑制了光线对图像的影响,而此时使用 WPCA,则相当于平均了各主成份分量,削减了那些本应对分类起主要作用的分量所产生的作用,从而取得了比 PCA 较低的识别率。

## 5 结论

从以上的分析可以看出,在采集条件较好的情况下,光照对识别结果并不会产生较大的影响。而 WPCA 由于平均了特征向量,削弱了原本对识别结果起主要作用的分量所产生的影响,因而取得了比 PCA 低的识别率。

同时,可以看到 WPCA 正是 ICA 预处理中的白化过程,它具有单位方差,比去相关稍微好点,正是基于这种原因, WPCA 在削减光照对识别结果的影响上起一定作用,和人脸识别中的去除前三个最大特征值所对应的特征向量有着相似的作用。但这种白化过程与 ICA 的正交是有着本质区别的,对识别结果的影响也是极其巨大的。

## 参考文献

- 1 边肇祺,张学工.模式识别.北京:清华大学出版社,2000.
- 2 Lu G, Zhang D, Wang K. Palmprint recognition using eigenpalms features. Pattern Recognition Letters, 2003, 24(9):1463 - 1467.
- 3 Turk M, Pentland A.. Eigenfaces for Recognition, Journal of Cognitive Neuroscience,1991,13(1):71 - 86.
- 4 Chen S, Zhao H, Kong M, Luo B. 2D-LPP: a two-dimensional extension of locality preserving projections. Neurocomputing, 2007,(70):912 - 921.
- 5 Duta N, Jain AK, Mardia KV. Matching of palmprint. Pattern Recognition Letters, 2001,23(4):477 - 485.
- 6 Kaplan NR. Palmprint verification:an implementation of biometric technology. ICPR, 1998,35(4):847 - 859.
- 7 李强,裘正定,孙冬梅等.基于改进的二维主成分分析的在线掌纹识别.电子学报,2005,33(2):7 - 9.