

基于矩阵完备投影的快速主分量分析算法

郭志波^{1),2)} 杨静宇¹⁾ 刘华军¹⁾ 严云洋¹⁾

¹⁾(南京理工大学计算机科学与技术学院, 南京 210094) ²⁾(扬州大学信息工程学院, 扬州 225009)

摘要 主分量分析是模式识别中经常采用的一种方法,但是由于经典的主分量分析在处理图像矩阵需要将图像展开成向量形式,因而造成其协方差矩阵维数和计算量太大,同时由于没有注意到图像矩阵中像素之间空间相关性,使得抽取的图像特征并不是优秀的,为此提出了一种基于矩阵完备投影的快速主分量分析算法(FMPCA),该算法不仅大大降低了分析过程中的计算量,而且发挥了图像矩阵行和列之间的空间特性,从而提高了整体性能。通过对 NUST603、Yale 和 ORL 图像库进行的实验证明,该算法不仅具有快速提取图像特征的能力,而且综合性能优于相应的一些主分量分析方法。

关键词 主分量分析 矩阵完备投影 特征抽取 街区距离

中图法分类号: TP391.41 **文献标识码**: A **文章编号**: 1006-8961(2007)04-0628-05

A New Method of Fast-complete Matrix-projection Principal Component Analysis

GUO Zhi-bo^{1),2)}, YANG Jing-yu¹⁾, LIU Hua-jun¹⁾, YAN Yuan-yang¹⁾

¹⁾(The College of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing 210094)

²⁾(The College of Information Engineer, Yangzhou University, Yangzhou 225009)

Abstract Principal component analysis (PCA) is a well-known method in pattern recognition. But the classical PCA transforms original image matrices into same dimensional vectors which will result in very large dimension of covariance matrix and very high computational complexity when processing image matrices. Moreover, extracted feature of the images are not excellent due to the fact that the pixel's spatial relativity based on the classical PCA was neglected. This paper presents a fast-complete matrix-projection principal component analysis (FMPCA) that decreases the computational complexity and utilizes the spatial relativity between rows and columns. The experiments conducted on NUST603, Yale and ORL face database demonstrate that the proposed algorithm can not only extract image feature efficiently but also maintain more powerful and excellent performance than some other principal component analysis methods.

Keywords principal component analysis (PCA), fast-complete matrix-projection PCA (FMPCA), feature extraction, block distance

1 引言

主分量分析(principal component analysis, PCA 或 K-L 变换)是模式识别中经常采用的一种方法,通过该方法不仅可以将样本投影到相应的一组正交向量上,而且通过设计正交投影向量,还可以将样本的主要分量提取到前面,也就是说,可利用少数的主

分量(鉴别特征)来代替整个样本,这样就可大大降低样本的维数,从而为后面的分类工作打下基础,所以是否能够快速提取有效的鉴别特征是解决问题的关键。如今许多在人脸识别方面的 PCA 方法^[1-10]都是基于经典 PCA 方法发展起来,经典 PCA 方法在性能上虽有着不错的表现,但是用它来处理图像矩阵时,则需要先将图像矩阵展开成向量形式,再用得到的协方差矩阵进行降维分析。可见经典 PCA

基金项目:国家自然科学基金项目(60632050,60472060,60473039)

收稿日期:2005-11-18;改回日期:2006-04-29

第一作者简介:郭志波(1975~),男,讲师。2003年获南京工业大学工学硕士学位,现为南京理工大学计算机系博士研究生。研究方向为模式识别、机器视觉等。E-mail:zhibo_guo@163.com

方法的弱点也很明显:首先,将图像矩阵转化为图像向量后,会造成图像向量的维数变得非常大,如需要处理分辨率为 $N \times N$ 大小的图像,其得到的协方差矩阵的维数就是 $N^2 \times N^2$, 所以对于高维的图像来说,完成 PCA 的速度就非常慢,同时将图像展开成向量只是将图像当成空间的一个点,并没有考虑行和列上的像素之间的分布关系,也就是一个像素和其四周像素之间的空间关系,这就造成抽取的鉴别特征并非原图像本身的,而是图像按行或列展开向量的鉴别特征。Yang 从欧氏距离角度提出了一种基于矩阵列向投影的主分量分析算法^[9,10] (简称 Yang PCA 算法),其实质是将图像的整行作为一个元素进行主分量分析,它是基于块结构的主分量分析的一个特例。由于该方法将图像整行作为一个元素,所以在降维计算中,计算量明显减小,在人脸识别中,该方法的识别效果优于经典 PCA 方法,但是由于该方法的一个明显的缺点就是抽取的鉴别特征向量的维数太大,并且没有考虑到行和列之间的混合关系,所以投影空间不完备,降维不彻底。从图像的 2 维正交变换(比如说 2 维离散余弦变换)来看,对于图像的降维分析,应该是同时进行行和列的双侧投影,其投影空间才是完备的,并且可以克服相应这些主分量分析的弱点。本文提出了一种基于图像矩阵完备投影的快速主分量分析算法(fast-complete matrix-projection PCA, FMPCA),不仅大大降低了过程中的计算量,并发挥了图像矩阵行和列之间的空间特性。实验证明,本文算法与相应的主分量分析相比,不仅具有快速抽取图像特征的能力,而且综合性能优于经典的 PCA 和 Yang PCA 方法,并弥补了它们的缺点。

2 FMPCA 算法设计

基于矩阵完备投影的快速主分量分析算法的思想是将 $s \times t$ 维的一组图像矩阵 $X^{(k)}$ 通过在行和列方向进行投影得到一组 $m \times n$ 维矩阵 $Y^{(k)}$, $Y^{(k)}$ 称为图像矩阵 $X^{(k)}$ 的鉴别特征矩阵($m < s$), ($n < t$), 其目的是为了便于识别和提高运算速度,所以应通过变换来尽量降低图像矩阵 $X^{(k)}$ 的维数。

定义 1 基于矩阵完备投影的主分量分析:将总数为 M 的 $s \times t$ 维图像矩阵 $X^{(k)}$, $k = 1, 2, \dots, M$, 通过在行和列方向上同时进行双侧投影来得到 M 个 $m \times n$ 维鉴别特征矩阵 $Y^{(k)} = U^T X^{(k)} V$ (U 和 V 分别是 $s \times m$ 、 $t \times n$ 维标准正交向量组)。

首先从欧氏距离角度出发,令 $\bar{Y} = \frac{1}{M} \sum_{k=1}^M Y^{(k)}$ 表

示所有鉴别特征矩阵的均值, $\bar{X} = \frac{1}{M} \sum_{k=1}^M X^{(k)}$ 表示所有样本矩阵的均值,则所有鉴别特征矩阵 $Y^{(k)}$ 到 \bar{Y} 的欧氏距离的平方为: $\|Y^{(k)} - \bar{Y}\|_F^2$, 其中 $\|\cdot\|_F$ 为 F-范数,根据定义 1 可得到:

$$\begin{aligned} \|Y^{(k)} - \bar{Y}\|_F^2 &= \|U^T (X^{(k)} - \bar{X}) V\|_F^2 \\ &= U^T (X^{(k)} - \bar{X}) V V^T (X^{(k)} - \bar{X})^T U \\ &= \sum_{i=1}^m \sum_{j=1}^n u_i^T (X^{(k)} - \bar{X}) v_j v_j^T (X^{(k)} - \bar{X})^T u_i \end{aligned} \quad (1)$$

根据式(1)可以得到所有鉴别特征矩阵 $Y^{(k)}$ 到 \bar{Y} 的距离平方的均值:

$$\begin{aligned} E \|Y^{(k)} - \bar{Y}\|_F^2 &= \sum_{i=1}^m \sum_{j=1}^n u_i^T \frac{1}{M} \sum_{k=1}^M (X^{(k)} - \bar{X}) v_j v_j^T (X^{(k)} - \bar{X})^T u_i \\ &= \sum_{i=1}^m \sum_{j=1}^n u_i^T E[(X^{(k)} - \bar{X}) v_j v_j^T (X^{(k)} - \bar{X})] u_i \\ &= \sum_{i=1}^m u_i^T E[(X^{(k)} - \bar{X}) V V^T (X^{(k)} - \bar{X})] u_i \end{aligned} \quad (2)$$

如果 v_j 是 $E[(X^{(k)} - \bar{X})^T (X^{(k)} - \bar{X})]$ 的本征向量,且 u_i 取 $E[(X^{(k)} - \bar{X}) V V^T (X^{(k)} - \bar{X})^T]$ 的本征向量,则式(1)可以表示为本征值之和,但此时本征值是 $E[(X^{(k)} - \bar{X}) V V^T (X^{(k)} - \bar{X})^T]$ 的本征值,而不是 $E[(X^{(k)} - \bar{X}) (X^{(k)} - \bar{X})^T]$ 的本征值,只有当图像样本的个数 M 等于 2 的时候,式(1)才可以表示为 $E[(X^{(k)} - \bar{X}) (X^{(k)} - \bar{X})^T]$ 的本征值之和,所以严格意义上讲,当样本个数 M 大于 2 时不能同时进行双侧投影,如果采用该方法,则需首先采用聚类分析获得样本的两个聚类中心,然后只是针对两个聚类中心,使用上述方法来获得正交投影向量组 U 、 V 。另外,如果在样本类别数较少的情况下,也可以近似使用上述方法,其性能优于行或列向投影。

如果 V 中的投影向量个数 $n = t$,且使用 $Y^{(k)}$ 的行向量进行计算:

$$\begin{aligned} \|Y^{(k)} - \bar{Y}\|_F^2 &= \sum_{i=1}^m (u_i^T X^{(k)} V - u_i^T \bar{X} V) (u_i^T X^{(k)} V - u_i^T \bar{X} V)^T \\ &= \sum_{i=1}^m u_i^T (X^{(k)} - \bar{X}) (X^{(k)} - \bar{X})^T u_i \end{aligned} \quad (3)$$

根据式(3)可得到:

$$\|E(Y^{(k)} - \bar{Y})\|_F^2 = \sum_{i=1}^m u_i^T E[(X^{(k)} - \bar{X}) (X^{(k)} - \bar{X})^T] u_i \quad (4)$$

由式(4)可看出,当且仅当 u_i 是 $E(X^{(k)} - \bar{X})(X^{(k)} - \bar{X})^T$ 的本征向量时, $\|E(Y^{(k)} - \bar{Y})\|_F^2 = \sum_{i=1}^m \lambda_i$ (λ_i 是 $E(X^{(k)} - \bar{X})(X^{(k)} - \bar{X})^T$ 的本征值),这说明投影变换与 V 无关, $Y^{(k)}$ 是 $m \times t$ 维矩阵,可见矩阵降维不彻底,同理可取 U 中投影向量个数 $m = s$,而且使用 $Y^{(k)}$ 的列向量进行计算也是如此。

下面从街区距离的角度出发,由鉴别特征矩阵 $Y^{(k)}$ 恢复的图像 $UY^{(k)}V^T$ 到某个参考矩阵鉴别特征 \hat{Y} 的恢复图像 $U\hat{Y}V^T$ 的街区距离为: $\|U(Y^{(k)} - \hat{Y})V^T\|_{m_1}$ ($\|\cdot\|_{m_1}$ 表示矩阵的 m_1 -范数),根据定义 1 可得:

$$\|U(Y^{(k)} - \hat{Y})V^T\|_{m_1} = \|UU^T(X^{(k)} - \hat{X})VV^T\|_{m_1}$$

此时,所有 $Y^{(k)}$ 恢复的图像 $UY^{(k)}V^T$ 到 \hat{Y} 的恢复图像 $U\hat{Y}V^T$ 的街区距离的均值为 $E\|U(Y^{(k)} - \hat{Y})V^T\|_{m_1} = E\|UU^T(X^{(k)} - \hat{X})VV^T\|_{m_1}$

$$= \|E|UU^T(X^{(k)} - \hat{X})VV^T|\|_{m_1}$$

其中, $|\cdot|$ 表示矩阵中的元素全部取绝对值后的矩阵,为便于计算,这里取参考矩阵 $\hat{X} = \{\hat{x}_{i,j}\}$ 中各元素值小于或大于所有图像矩阵 $X^{(k)}$ 中对应元素值(后面实验中取 $\hat{x}_{i,j} = \max(x_{i,j}^{(k)})$),可以认为

$$|UU^T(X^{(k)} - \hat{X})VV^T| = UU^T|X^{(k)} - \hat{X}|VV^T \quad (5)$$

根据式(5)可得到:

$$E\|U(Y^{(k)} - \hat{Y})V^T\|_{m_1} = \|UU^TE|X^{(k)} - \hat{X}|VV^T\|_{m_1}$$

如果 V 中的正交向量 v_j 取 $E|(X^{(k)} - \hat{X})^T| \times E|X^{(k)} - \hat{X}|$ 的本征向量,并且正交向量组 U 中的正交向量 u_i 取 $E|X^{(k)} - \hat{X}|E|(X^{(k)} - \hat{X})^T|$ 的本征向量,则得到

$$UU^TE|X^{(k)} - \hat{X}|VV^T = \sum_{i=1}^q \delta_i u_i v_i^T$$

其中, δ_i 是 $E|(X^{(k)} - \hat{X})^T|E|X^{(k)} - \hat{X}|$ 的本征值,也是 $E|X^{(k)} - \hat{X}|E|(X^{(k)} - \hat{X})^T|$ 的本征值,而且 $E|X^{(k)} - \hat{X}| = \sum_{i=1}^r \delta_i u_i v_i^T$, $q < r$, $q = \min(m, n)$, $r = \text{rank}(E|X^{(k)} - \hat{X}|)$, $\delta_1 > \delta_2 > \dots > \delta_r$ 。

所以,可以得出结论:当 V 中的正交向量 v_j 取 $E|(X^{(k)} - \hat{X})^T|E|X^{(k)} - \hat{X}|$ 的本征向量,且 U 中的正交向量 u_i 取 $E|X^{(k)} - \hat{X}|E|(X^{(k)} - \hat{X})^T|$ 的本征向量时, $E\|U(Y^{(k)} - \hat{Y})V^T\|_{m_1}$ 保留了 $\|E|X^{(k)} - \hat{X}|\|_{m_1}$,也就是 $E\|X^{(k)} - \hat{X}\|_{m_1}$ 的主要分量。

在使用本文 PCA 方法时, m 和 n 不一定要相同,试验中可以发现,对于正面人脸样本而言,列向投影特征比行向投影特征强烈,因为在维数相同的

情况下,选用的鉴别特征越强烈,保留原样本中的分量就越多,所以可通过适当增加 m ($m = a \times n, a > 1$) 来进一步提高识别效果。

3 实验结果

由于经典 PCA 方法(K_L 变换:按列向或行向将矩阵展开成向量进行降维)和 Yang PCA 方法(基于欧氏距离的矩阵列向投影)在文献[9,10]中已经进行过比较,所以在这里的实验中,只是采用本文方法(基于街区距离的矩阵完备投影)和 Yang PCA 方法进行速度、降维效果和识别效果的对比,为了检查算法对采用不同距离度量方式的实验效果,对两种方法都采用了欧氏距离和街区距离度量方式。本文实验选用的图像人脸库是 NUST603 人脸标准库、Yale 人脸库以及 ORL 人脸库,其中,NUST603 人脸标准库包含 96 个人,每人 10 幅不同照明及表情的 32×32 大小的标准图像(如图 1 所示);Yale 库包含 15 个人,每人 11 幅不同照明、不同面部表情、不同面部姿态的 112×92 大小的图像(如图 2 所示);ORL 库包含 40 个人,每人 10 幅 112×92 大小的图像,在 10 幅图像中,人的脸部表情和脸部细节有着不同程度的变化,人脸姿态也有相当程度的变化(如图 3 所示)。实验条件为:处理器 Celeron D-2.66M、内存 256M、Matlab7.0。



图 1 NUST603 人脸图像标准库

Fig. 1 NUST603 human face standard database



图 2 Yale 人脸图像库

Fig. 2 Yale human face database

实验时,首先进行 NUST603 人脸标准库的识别,同时采用每人的前 5 幅图像作为训练样本,后 5 幅作为测试样本,这样训练样本和测试样本总数均为 480,实验结果如表 1 所示。使用最近邻测量方法进行分类测试,在使用街区距离测量时,本文方法



图 3 ORL 人脸图像库

Fig.3 ORL human face database

使用了 108 个鉴别特征,识别率就达到 96.25%,而 Yang PCA 方法用了 224 个鉴别特征,识别率才达到 95.21%,使用欧氏距离测量时,本文方法使用了 108 个鉴别特征,识别率就达到 94.79%,而 Yang PCA 方法用了 320 个鉴别特征,识别率才达到 93.54%,从实验结果中还可以看出,本文 PCA 方法

在速度上也明显快于 Yang PCA 方法。

为了进一步证实这一方法的可靠性,本文还选用了 Yale 人脸库进行了 Yang PCA 和本文 PCA 方法的对比实验,实验中采用每人的前 6 幅图像作为训练样本,后 5 幅作为测试样本,并使用最小距离测量方法进行分类测试。实验结果如表 2 所示。在使用街区距离测量时,本文方法使用了 147 个鉴别特征,识别率就达到 98.67%,而 Yang PCA 方法用了 784 个鉴别特征,识别率才达到 85.33%;使用欧氏距离测量时,本文方法使用了 147 个鉴别特征,识别率就达到 93.33%,而 Yang PCA 用了 784 个鉴别特征,识别率才达到 88%,由此可以看出,对于较复杂条件下的小类别样本,本文方法综合性能优于 Yang PCA 方法。

表 1 NUST603 人脸标准库上测试结果对照

Tab.1 The comparison of experiment results on NUST603 face standard database

	识别率 (%)									
	本文 PCA 方法 ($a=3$)					Yang PCA 方法				
	$4a \times 4$	$5a \times 5$	$6a \times 6$	$7a \times 7$	$8a \times 8$	32×6	32×7	32×8	32×9	32×10
最近邻(街区距离)	95.21	95.63	96.25	95.42	95.00	94.37	95.21	94.37	94.58	94.79
最近邻(欧氏距离)	92.29	94.37	94.79	94.58	93.33	93.33	93.33	93.13	93.33	93.54
总时间(s)	8.4390	9.4700	11.4390	14.1410	16.3440	18.1400	19.6710	21.0620	23.1090	24.2030

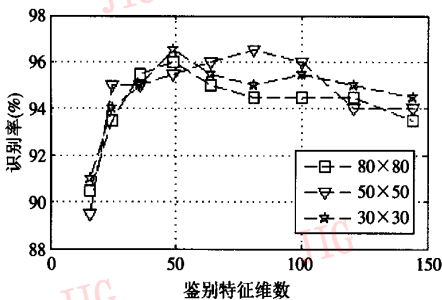
表 2 Yale 人脸标准库上测试结果对照

Tab.2 The comparison of experiment results on Yale face database

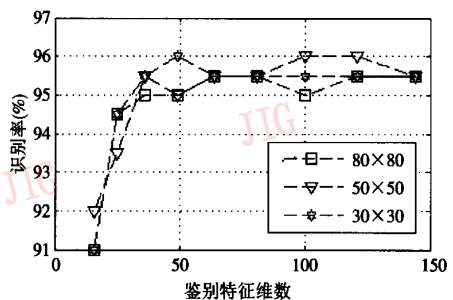
	识别率 (%)									
	本文 PCA 方法 ($a=3$)					Yang PCA 方法				
	$5a \times 5$	$7a \times 7$	$9a \times 9$	$11a \times 11$	$13a \times 13$	112×5	112×7	112×9	112×11	112×13
最小街区距离	96.00	98.67	97.33	97.33	96.00	84.00	85.33	82.67	81.33	82.67
最小欧氏距离	90.67	93.33	93.33	93.33	92.00	84.00	88.00	88.00	88.00	85.33
总时间(s)	0.5150	0.5620	0.5930	0.6710	0.7650	0.6570	0.7810	0.8910	1.0940	1.0310

最后为了验证算法对低分辨率图像也具有较好的识别性能,将 ORL 库中的人脸图像通过双线性插值变成 80×80 、 50×50 、 30×30 3 种分辨率的图像

分别进行测试比较,测试时选择 5 幅训练、5 幅作为测试样本,使用最近邻度量方式进行测量。图 4 的纵坐标是识别精度,横坐标是鉴别特征的维数(取



(a) 街区距离



(b) 欧氏距离

图 4 ORL 人脸库实验结果

Fig.4 The experiment results on ORL face database

的是 $4 \times 4 \sim 12 \times 12$), 由图 4 可见, 无论使用街区距离还是欧氏距离来进行测量, 3 种分辨率人脸图像的识别效果都比较相近, 低分辨图像的识别性能也是比较好的, 所以该算法对图像分辨率的适应范围较宽。

4 结 论

本文针对目前相关的一些 PCA 方法的不足, 提出了一类矩阵完备投影快速主分量分析算法并将其应用到人脸识别方面, 实验证明, 该方法不仅计算速度快、降维尺度大, 且识别性能也优于相关的一些方法。该方法不仅可以用在图像识别上, 还可以用在图像检测、目标跟踪等方面。

参考文献 (References)

- 1 Kirby M, Sirovich L. Application of the Karhunen-Loeve procedure for the characterization of human faces [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1990, 12 (1): 103 ~ 108.
- 2 Peng Hui, Zhang Chang-shui, Rong Gang, *et al.* Research of automated face recognition based on K-L transform [J]. *Journal of Tsinghua University (Science and Technology)*, 1997, 37(3): 67 ~ 70. [彭辉, 张长水, 荣钢等. 基于 K-L 变换的人脸自动识别 [J]. *清华大学学报 (自然科学版)*, 1997, 37(3): 67 ~ 70.]
- 3 Turk M A, Pentland A P. Face recognition using eigenfaces [A]. In: *Proceedings of the IEEE Conference in Computer Vision and Pattern Recognition [C]*, Maui, Hawaii, USA, 1991: 586 ~ 591.
- 4 Valentin D, Abdi H, O' Toole A J. Principal component analysis and neural network of face images: What can be generalized in gender classification [J]. *Journal of Mathematical Psychology*, 1997, 41(4): 398 ~ 413.
- 5 Swets D, Weng J. Using discriminant eigenfeatures for image retrieval [R]. Technical Report CPS-96-16, Department. of Computer Science, MSU, 1996.
- 6 Belhumeur P N, Hespanha J P, Kriegman D J. Eigenfaces vs. Fisherfaces: recognition using class special linear projection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997, 19(7): 711 ~ 720.
- 7 Scholkopf B, Smola B, Muller K R. Nonlinear component analysis as a kernel eigenvalue problem [J]. *Neural Compute*, 1998, 10(5): 1299 ~ 1319.
- 8 Yang M H. Kernel Eigenface vs. kernel Fisherface: face recognition using kernel methods [A]. In: *IEEE International Conference on Face and Gesture Recognition [C]*, Washington, DC, USA, 2002.
- 9 Yang J, Yang J Y. From image vector to matrix: a straightforward image projection technique—IMPCA vs. PCA [J]. *Pattern Recognition*, 2002, 35(9): 1997 ~ 1999.
- 10 Yang J, Zhang D, Frangi, A F, Yang J Y. Two-dimensional PCA: a new approach to appearance-based Face representation and recognition [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, 26(1): 131 ~ 137.