

基于 QR 分解与 2DLDA 的单样本人脸识别

覃 磊^{1,2}, 李德华¹, 周 康²

(1 华中科技大学 自动化学院, 湖北 武汉 430074; 2 武汉轻工大学 数学与计算机学院, 湖北 武汉 430023)

摘 要: 提出了一种新的基于矩阵的 QR 分解与 2DLDA 的单样本人脸识别算法 (QR decomposition + 2DLDA)。利用矩阵的 QR 分解, 将单样本人脸图像进行 QR 分解后提取有效的部分信息构成虚拟图像, 结合原训练图像生成新的训练样本集, 应用 2DLDA 进行特征提取和识别。在 ORL 人脸数据库上对算法进行了实验, 实验结果表明此算法的识别效果不仅优于 PCA、SPCA、 $(PC)^2A$ 、 $E(PC)^2A$ 算法, 而且对于光照、表情等因素具有良好的鲁棒性。

关键词: 虚拟图像; 单样本; 二维线性判别分析; QR 分解

中图分类号: TN402

文献标识码: A

文章编号: 1000-7180(2015)02-0065-04

Single Sample Face Recognition Based on QR Decomposition and 2DLDA

QIN Lei^{1,2}, LI De-hua¹, ZHOU Kang²

(1 School of Automation, Huazhong University of Science & Technology, Wuhan 430074, China;

2 School of mathematics and computer, Wuhan Polytechnic University, Wuhan 430023, China)

Abstract: This paper proposes a new single sample face recognition algorithm based on matrix QR decomposition and 2DLDA. The algorithm uses QR decomposition of a matrix on the single sample, extracts the effective information to constitute a virtual image, combines the original training images and generates a new training sample set, then the application of 2DLDA algorithm for feature extraction and recognition. The algorithm is tested on ORL face database, experimental results show that this algorithm is not only superior to the traditional algorithms on recognition rate, such as PCA, SPCA, $(PC)^2A$, $E(PC)^2A$ algorithms, but also has good robustness for factors such as illumination, expression etc.

Key words: virtual image; single sample; 2DLDA; QR decomposition

1 引言

近几十年来, 人脸识别问题已成为计算机视觉、模式识别等领域的研究热点之一, 研究者们提出了如主成分分析^[1]、线性判别分析、局部二值模式、隐马尔可夫模型等方法。这些方法往往需要每种模式类有多个样本才能达到较好的识别效果, 但在实际应用中, 人们往往面临只有少量甚至单样本的特殊场景, 如护照、身份证验证、法律实施等, 每个人只有一副图像。这时传统方法的性能往往大幅下降, 如

2DLDA 方法, 因为是单样本, 无法计算类内散度, 所以方法失效。对于单样本人脸识别问题, 近来研究者们提出了不少方法。如 Wu Jianxin 和 Zhou Zhi-Hua 提出的 $(PC)^2A$ 算法^[2], 通过计算原图像的水平 and 垂直投影, 结合原图像构建投影结合的图像, 然后将 PCA 应用于投影结合的图像。此方法实质是一种预处理的 PCA 方法, 即将原图像进行了投影滤波。在此基础上, Chen Songcan 等提出了 $E(PC)^2A$ 算法^[3], 不仅考虑了图像的一阶投影, 而且考虑了其二阶甚至 n 阶投影, 同样应用 PCA 于所得

收稿日期: 2014-05-20; 修回日期: 2014-06-10

基金项目: 国家自然科学基金项目 (61179032, 61072143); 湖北省自然科学基金项目 (2011CDB229); 湖北省教育厅科技计划项目 (Q20131703)

到的投影图像、投影结合的图像. 相对于 $(PC)^2A$ 算法, 此时单样本识别问题可转化为一般多样本识别问题, FERET 数据库上的实验结果表明其识别率相对于 $(PC)^2A$ 算法有所提高. Zhang Daoqiang 等提出了一种基于奇异值扰动的单样本人脸识别算法^[4], 在文中先将图像矩阵做奇异值分解, 然后扰动分解后所得到的对角矩阵, 生成虚拟图像后应用 PCA 算法, 相对于 $(PC)^2A$ 算法能有效地提高识别率, 降低存储要求. Gao Quan-xue 等提出了一种新的 FLDA 算法^[5], 通过图像矩阵的奇异值分解构造一个虚拟图像, 将单样本问题转化为多样本问题, 并应用 2DLDA 算法^[6]. 本文继承上述文献的思想, 通过矩阵的 QR 分解^[7]构造虚拟样本, 并应用 2DLDA 算法, 寻求更有区分力的投影方向.

2 2DLDA

2.1 投影轴的建立

设 A 表示 $m \times n$ 的图像, x 表示一个 n 维的列向量, A 通过 $y = Ax$ 投影到 x 上. 设实验中有 L 个模式类, 第 i 个模式类用集合 T_i 表示, 共有 M 个训练图像, 第 T_i 类中第 j 个图像用 A_i^j 表示, $\bar{A}$ 表示 M 个训练样本的平均值, N_i 表示第 T_i 类图像的个数, $\bar{A}_i$ 表示第 T_i 类图像的平均值. 最优投影向量 x 通过最大化类间散度和最小化类内散度来达到, 即使得

$$J(x) = \frac{tr(TS_B)}{tr(TS_W)} \quad (1)$$

达到最大, 其中

$$\begin{aligned} TS_B &= \sum_{i=1}^L N_i (\bar{y}_i - \bar{y})(\bar{y}_i - \bar{y})^T \\ &= \sum_{i=1}^L N_i [(\bar{A}_i - \bar{A})x][(\bar{A}_i - \bar{A})x]^T \end{aligned} \quad (2)$$

$$\begin{aligned} TS_W &= \sum_{i=1}^L \sum_{j \in T_i} (y_i^j - \bar{y}_i)(y_i^j - \bar{y}_i)^T \\ &= \sum_{i=1}^L \sum_{j \in T_i} [(A_i^j - \bar{A}_i)x][(A_i^j - \bar{A}_i)x]^T \end{aligned} \quad (3)$$

所以

$$\begin{aligned} tr(TS_B) &= x^T \left(\sum_{i=1}^L N_i (\bar{A}_i - \bar{A})^T (\bar{A}_i - \bar{A}) \right) x \\ &= x^T S_B x \end{aligned} \quad (4)$$

$$\begin{aligned} tr(TS_W) &= x^T \left(\sum_{i=1}^L \sum_{j \in T_i} (A_i^j - \bar{A}_i)^T (A_i^j - \bar{A}_i) \right) x \\ &= x^T S_W x \end{aligned} \quad (5)$$

求解广义特征值问题 $S_B x = \lambda S_W x$, 所得广义特

征向量 $x_1, x_2, \dots, x_d$ 即为最佳投影矩阵, 即 $X = (x_1, x_2, \dots, x_d)$.

2.2 特征提取

对于任意一个给定的图像 A , 我们将其投影到前文所得到的 X 上, 求出 $Y = (y_1, y_2, \dots, y_d) = AX = A(x_1, x_2, \dots, x_d)$, Y 即为图像 A 的二维 Fisher 特征矩阵.

2.3 分类

训练时, 将所有的训练样本向投影矩阵做投影, 得到所有训练样本的特征矩阵. 测试时, 将每一幅测试样本图像向投影矩阵做投影, 得到其特征矩阵. 比较此特征矩阵和所有训练样本图像的特征矩阵的距离, 利用最小距离分类器做出判决. 设两幅图像特征矩阵分别为 $Y_1 = (y_1^1, y_1^2, \dots, y_1^d)$, $Y_2 = (y_2^1, y_2^2, \dots, y_2^d)$, 则 $d(Y_1, Y_2) = \sum_{k=1}^d ||y_k^1 - y_k^2||_2$, 对于给定的测试图像 Y , 若 $d(Y, Y_k) = \min_j d(Y, Y_j)$ 且 $Y_k \in T_l$, 则 $Y \in T_l$.

3 基于 QR 分解与 2DLDA 的单样本识别

因为 2DLDA 算法要求每类样本至少有两幅甚至多幅图像, 这样才能有效计算每类样本的类内散度. 但在如身份验证、法律实施等一些特定场景, 每个人往往只能提供一副图像, 计算出样本的类内散度为零. 为了突破此限制, 本文从图像的 QR 分解入手, 将每个图像矩阵做 QR 分解, 选取分解后 R 矩阵中部分对角元所对应的行和矩阵 Q 生成虚拟图像, 结合原图像合并为新的训练样本集. 这样每个人将有两幅训练样本图像, 2DLDA 算法就可在此训练集中实施.

3.1 图像的 QR 分解和虚拟图像的生成

假定一副图像 $A \in R^{m \times n}$, $m \geq n$, 则可对 A 做如下分解

$$AE = QR \quad (6)$$

式中, $Q \in m \times m$, $R \in m \times n$, Q 为正交矩阵, R 为上三角矩阵, E 为置换矩阵, R 的对角元呈降序排列, 即

$$R = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix},$$

$$\text{且 } |r_{11}| \geq |r_{22}| \geq \cdots \geq |r_{nn}| \quad (7)$$

记 $Q = (Q_1, Q_2, \dots, Q_m)$, 其中 $Q_i (i = 1, 2, \dots, m)$ 代表 Q 的第 i 列, $R = (R_1, R_2, \dots, R_n, 0, \dots, 0)^T$, 其中 $R_j^T (j = 1, 2, \dots, n)$ 代表 R 的第 j 行. 所以式 (6) 可简写为

$$AE = QR = (Q_1, Q_2, \dots, Q_m)(R_1, R_2, \dots, R_n, 0, \dots, 0)^T$$

$$= Q_1 R_1^T + Q_2 R_2^T + \dots + Q_n R_n^T = \sum_{i=1}^n Q_i R_i^T \quad (8)$$

式 (8) 表明 AE 是由一系列基矩阵 $F_i = Q_i R_i^T$, $(i = 1, 2, \dots, n)$ 所组成, 因为 R 的对角元呈降序排列, 所以 A 的大部分能量应集中于前面的 F_i , 选取前面的 k 个基矩阵 F_i 之和作为 AE 的近似矩阵. 选取 $k = 40$, 即

$$AE \approx \sum_{i=1}^k Q_i R_i^T \quad (9)$$

$$A \approx \left(\sum_{i=1}^k Q_i R_i^T \right) E^{-1} \quad (10)$$

记 $\hat{A} = \left(\sum_{i=1}^k Q_i R_i^T \right) E^{-1}$, 即将 $\hat{A}$ 作为原始图像 A

通过 QR 分解所生成的虚拟图像.

3.2 特征的提取和分类

对于每一幅训练图像, 通过 QR 分解对应产生一幅虚拟图像, 所以若有 M 幅训练图像, 将生成 M 幅虚拟图像, 把两者合并为新的训练样本集, 应用 2DLDA 算法进行特征提取, 然后分类.

4 实验与分析

4.1 实验数据库

本文选用 ORL 人脸数据库进行算法的检测, ORL 数据库由剑桥大学的 Olivetti Research Laboratory 于 1992 年 4 月到 1994 年 4 月拍摄. 数据库包括 40 个人, 每人 10 幅图像, 一共 400 幅人脸图像, 照片拍摄于不同的时间. 在一定范围内存在光照变化、面部表情变化 (如睁眼或闭眼、微笑或不微笑) 以及脸部细节变化 (如是否戴眼镜). 所有照片都是正面直立的, 甚至还有一定程度的头部倾斜和转动, 转动角度不超过 20 度, 图像为 112×92 的灰度图像. 实验中每人任选一副图像, 一共 40 幅图像作为训练样本集, 其余的 360 幅图像作为测试样本.

为更好地检测算法的准确性和鲁棒性, 每次实验重复 10 次, 且重复的每次实验都重选不同的训练样本和测试样本.

4.2 实验的结果与分析

本文首先比较本文 QR decomposition +

2DLDA 算法和 PCA、SPCA、 $(PC)^2A$ 、 $E(PC)^2A$ 算法随投影轴数变化的识别率, 实验结果如图 1 和表 1 所示.

从图 1 可以看出, 当投影轴数偏小时, QR decomposition + 2DLDA 算法的识别率大大超出了 SPCA 等其他算法. 而投影轴数较小意味着将产生较小的特征矩阵, 这对于特征矩阵后续的存储、计算等处理无疑是有利的. 随着特征轴数的增加, 虽然其他算法的识别率在逐步增加, 但具体分析发现 SPCA 算法的最高识别率只有 59.03%, 对比算法的最低识别率 65.50% 甚至还要低 6.47 个百分点. $E(PC)^2A$ 算法的最高识别率也只有 65.97%, 与本文算法的最低识别率相差无几. 虽然 PCA、 $(PC)^2A$ 的最高识别率分别为 69.20%、69.50%, 但离 QR decomposition + 2DLDA 算法的最高识别率 71.72% 仍分别相差 2.52%、2.22%, 具体数据见表 1.

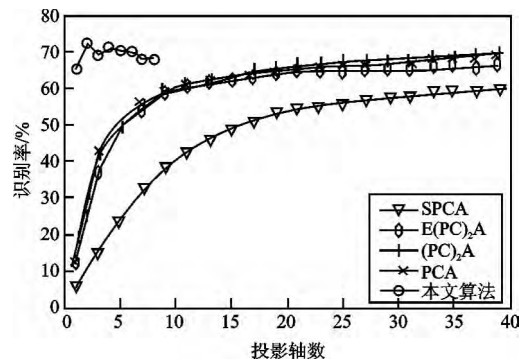


图 1 随投影轴数变化的各算法的识别率

表 1 各算法的最高识别率

算法	PCA	SPCA	$(PC)^2A$	$E(PC)^2A$	本文算法
最优投影轴数	39	34	39	39	2
特征矩阵维数	39×1	34×1	39×1	39×1	112×2
识别率/%	69.20	59.03	69.50	65.97	71.72

QRdecomposition + 2DLDA 算法明显超出其他算法的原因在于 PCA 算法仅考虑了样本间的全局散度, 而对于类内散度、类间散度却不加以区分, 在单样本条件下更是难以区分各类的本质属性. 而 $(PC)^2A$ 、 $E(PC)^2A$ 只是先对图像进行了垂直和水平滤波, 然后再采取 PCA 算法, 其本质并未脱离 PCA 算法. 对于 SPCA 算法, 虽然其原理和本文算法有些相似之处, 但奇异值扰动可能过多地改变了样本图像本身的拓扑结构, 反而未达到好的识别效果. 另外从表 2 可以看出 QR decomposition + 2DLDA 算法和其他几种算法的特征提取时间相差很小, 本文算法的特征提取时间为 3.21 s, 而特征提

取时间最少的(PC)²A 的特征提取时间为2.37 s.

表2 各算法的特征提取时间

算法	(PC) ² A	PCA	E(PC) ² A	SPCA	本文算法
时间/s	2.37	2.87	3.16	4.04	3.21

式(2)中还有一个参数 k ,它决定了生成怎样的虚拟图像.为了研究此参数对算法的影响,我们对 k 的取值进行了实验.从ORL数据库随机选取了一组训练样本图像,每人一幅图像.对 $10 \leq k \leq 70$ 的 k 值进行了均匀的取值,每隔10作为一个步长,记录最高识别率.实验结果如表3所示,可看出当 k 偏小(接近于10)或偏大(接近于70)时,算法的识别率并不高,而当 k 取为中间值,如 k 为40时达到最好的识别效果,所以前文中我们将 k 值取为40.原因在于若 k 过小,生成的虚拟图像包含原图像的信息过少,与原图像的差异过大,这样客观上扩大了类内散度.而当 k 过大时,虽然较丰富地反映了原图像的信息,但因为过拟合使得过于强调同类样本的细节信息,反而影响识别效果.对于如何更精确地选取合适的 k 值,这需要在以后通过大量的实验研究甚至理论推导,这也可以作为后续的研究主题之一.

表3 随 k 值变化的本文算法的识别率

k	10	20	30	40	50	60	70
识别率/%	70.56	74.17	73.61	75.00	72.50	68.89	61.67

5 结束语

本文在生成虚拟样本思想的启发下,应用QR分解生成虚拟图像,在后续的特征提取中应用2DLDA算法.从实验结果上来看,QR decomposition+2DLDA算法超出了PCA、SPCA等算法;其次,对于2DLDA算法无法解决单样本问题的限制,因为虚拟样本算法的产生,2DLDA算法得以实施.而且应用于ORL数据库的实验结果也表明对于光照、表情、姿势等因素本文算法体现出相当的鲁棒性.

应用QR decomposition+2DLDA识别单样本问题不仅适用于人脸识别,而且对于其他的小样本、单样本问题都可以利用本文思想,可应用矩阵分解生成一幅甚至多幅图像扩充样本库,更好的发现样本的内部属性,而这正是我们下一步需要探讨的主题.

参考文献:

- [1] Turk M, Pentland A. Eigenfaces for recognition[J]. Journal of Cognitive Neuroscience, 1991, 3(1):71-86.
- [2] Wu Jianxin, Zhou Zhihua. Face recognition with one training image per person[J]. Pattern Recognition Letters, 2002, 23(14):1711-1719.
- [3] Chen Songcan, Zhang Daoqiang, Zhou Zhihua. Enhanced (PC)²A for face recognition with one training image per person[J]. Pattern Recognition Letters, 2004, 25(10):1173-1181.
- [4] Zhang Daoqiang, Chen Songcan, Zhou Zhihua. A new face recognition method based on SVD perturbation for single example image per person[J]. Applied Mathematics and Computation, 2005, 163(2):895-907.
- [5] Gao Quanyue, Zhang Lei, David Zhang. Face recognition using FLDA with single training image per person[J]. Applied Mathematics and Computation, 2008, 205(12):726-734.
- [6] Li Ming, Yuan Baozong. 2D-LDA: a statistical linear discriminant analysis for image matrix[J]. Pattern Recognition Letters, 2005, 26(5):527-532.
- [7] 程云鹏, 张凯院, 徐仲. 矩阵论[M]. 西安:西北工业大学出版社, 2006: 178-233.

作者简介:

覃磊 男, (1979-), 博士研究生, 讲师. 研究方向为模式识别. E-mail: 53039943@qq.com

李德华 男, (1946-), 博士生导师, 教授. 研究方向为模式识别与人工智能.

周康 男, (1965-), 博士, 教授. 研究方向为模式识别.