

基于有监督直接局部保持投影的人脸识别

李政仪¹, 朱益丹^{2,3}, 赵龙²

(1. 长沙民政学院软件学院, 长沙 410004; 2. 国防科技大学计算机学院, 长沙 410073; 3. 北京节能环保中心, 北京 100029)

摘要: 提出一种用于图像识别的有监督直接局部保持投影算法, 该算法结合样本类别信息, 通过同时对角化的方法求解局部保持投影问题, 避免矩阵的奇异性。在 ORL 人脸库上的测试结果表明, 该算法的识别率高于 PCA, PCA+LPP 等方法。

关键词: 局部保持投影; 图像识别; 主成分分析

Face Recognition Based on Supervised Direct Locality Preserving Projection

LI Zheng-yi¹, ZHU Yi-dan^{2,3}, ZHAO Long²

(1. Software School, Changsha Social Work College, Changsha 410004;

2. College of Computer, National University of Defense Technology, Changsha 410073;

3. Beijing Energy Conservation and Environment Protection Center, Beijing 100029)

【Abstract】 This paper proposes a supervised direct locality preserving projection algorithm for image recognition. This algorithm solves locality preserving projection problem via simultaneous diagonalization combined with samples' information, and can avoid the singularity of the matrices. Experimental results on the ORL face database show that this algorithm is more powerful than PCA and PCA+LPP for face recognition.

【Key words】 locality preserving projection; image recognition; Principal Component Analysis(PCA)

1 概述

人脸自动识别是一种用计算机分析人脸图像, 从中提取有效信息并进行自动鉴别的技术, 它在安全系统及人机交互等方面具有极大应用前景, 已成为计算机视觉和模式识别领域的重要研究课题之一。研究人员提出了许多人脸识别方法, 大致分为 3 类^[1]: 整体方法, 基于特征匹配的方法和混合方法。整体方法将整个人脸区域作为识别系统的输入, 不依赖人脸的先验知识和参考模型, 一直是人脸识别的研究热点之一, 基于子空间分析方法在整体方法中占重要地位。局部保持投影算法是一种最近提出的子空间方法, 它考虑了样本空间的图结构, 在降维过程中能保持样本空间的局部结构和本质几何特性。但在实际应用中, 局部保持投影算法在处理高维数据时面临着矩阵奇异问题, 因此, 在拉普拉斯方法中, 使用主成分分析进行降维。但上述 2 种算法的目标函数不同, 导致局部保持投影算法不能获得最优解。针对上述问题, 本文提出一种基于有监督的直接局部保持投影算法, 并在 ORL 人脸库上通过实验证明该方法的有效性。

2 局部保持投影算法

局部保持投影算法 LPP 是一种典型的流形学习算法——拉普拉斯特征映射^[2]的线性近似, 它可以用谱理论来阐述。假设 n 维数据集 $\{x_1, x_2, \dots, x_N\}$ 分布在一个低维子流形上, 希望找到一组 d 维 ($d \ll n$) 数据点 $\{y_1, y_2, \dots, y_N\}$, 该组数据点局部的邻域结构与数据 $\{x_1, x_2, \dots, x_N\}$ 一样。为了达到上述目的, 可以构建一个加权的图 $G=(X, W)$, 其中, X 表示数据点的集合; W 表示位置相似矩阵。在给定上述定义的基础上, 局部保持投影算法的目标函数定义为

$$\min \sum_{i,j} (y_i - y_j)^2 W_{ij}$$

其中, $y_i = \phi^T x_i$, 且 $\phi \in [\phi_1, \phi_2, \dots, \phi_{d-1}]$ 。经过一些简单的代数计算后, 可得

$$\frac{1}{2} \min \sum_{i,j} (y_i - y_j)^2 W_{ij} = \phi^T X(D-W)X^T \phi = \phi^T XLX^T \phi \quad (1)$$

其中, $X = [x_1, x_2, \dots, x_N]$; D 是一个对角矩阵, 其元素定义为 $D_{ii} = \sum_j W_{ij}$ 或 $D_{ii} = \sum_i W_{ij}$, W 是一个对称矩阵; $L = D - W$ 是拉普拉斯矩阵。可以将最小化问题简化为

$$\arg \min_{\phi} \phi^T XLX^T \phi \quad \text{Subject to: } \phi^T XDX^T \phi = 1 \quad (2)$$

该最小化目标函数的变换矩阵可以通过求解如下广义特征值问题得到:

$$XLX^T \phi = \lambda XDX^T \phi \quad (3)$$

设式(2)的解是列向量 $\phi_1, \phi_2, \dots, \phi_{d-1}$, 它们以相对应的特征值递增的顺序排列。最终的低维嵌入定义为

$$y_i = A^T X_i, A = (\phi_1, \phi_2, \dots, \phi_{d-1})$$

x_i 被降维成 d 维向量 y_i , y_i 保持了它和近邻点的局部结构关系。

3 有监督直接局部保持投影算法

3.1 算法主要步骤

本文提出一种可以有效结合样本类别信息的有监督直接局部保持投影算法。该算法以直接优化局部保持投影为准则, 采用同时对角化的方法进行求解。假设类别数为 C , 第 c 类 ($c=1, 2, \dots, C$) 的样本数为 n_c , 则有 $\sum_{c=1}^C n_c = N$ 。

作者简介: 李政仪(1979 -), 女, 讲师、硕士研究生, 主研方向: 模式识别, 虚拟现实技术; 朱益丹, 工程师、硕士研究生; 赵龙, 教授

收稿日期: 2008-11-29 **E-mail:** guofang80@163.com

本文算法的主要步骤如下：

Step1 建立带权图 G_1 和 G_2 。设带权图中有 N 个节点，对应 N 个样本点。若 2 个节点彼此是最近邻点，则构建带权图记为 G_1 。如果 2 个节点属于同一类别，则构建带权图记为 G_2 。

Step2 建立相似矩阵 W 。设 W^I 和 W^{II} 分别表示带权图 G_1 和 G_2 的相似矩阵，即

$$W_{ij}^I = \begin{cases} x_i^T x_j, & G_1 \text{中节点 } i \text{ 和 } j \text{ 相连} \\ 0, & \text{其他} \end{cases}$$

$$W_{ij}^{II} = \begin{cases} x_i^T x_j, & G_2 \text{中节点 } i \text{ 和 } j \text{ 相连} \\ 0, & \text{其他} \end{cases}$$

则最终的相似矩阵为

$$W = \theta \cdot W^I + (1 - \theta) \cdot W^{II}$$

其中， $0 < \theta < 1$ 是可调因子，用来表示流形局部结构信息和类别信息在构建带权图时各自所占的权重。

Step3 采用同时对角化的方法求矩阵 A 。

3.2 计算方法

在线性代数中，求解广义特征值问题的一种方法是采用同时对角化的方法。该方法的主要思想是通过角化 XDX^T 来去掉它的零空间^[3]，并通过投影和对角化 XLX^T 来寻找局部保持投影向量。

由式(2)和式(3)可得

$$g(\phi^*) = \arg \min \left(\frac{\phi^T XLX^T \phi}{\phi^T XDX^T \phi} \right) \quad (4)$$

利用式(4)，可以得到以下引理：

引理 对局部保持投影准则而言， XDX^T 的零空间不包含任何有用信息。

证明：在 XDX^T 的零空间内，存在 $XDX^T \phi = 0$ ，有 $\phi^T XDX^T \phi = 0$ 。在此情况下，不能得到最小值，且式(2)中的目标函数值将变得无穷大。如果 $XDX^T \phi \neq 0$ 且 $XLX^T \phi \neq 0$ ，则 $(\phi^T XLX^T \phi) / (\phi^T XDX^T \phi)$ 将取得最小值 0。

可见，把数据投影到 ϕ 方向时，它是 XLX^T 的零空间内的向量而不是 XDX^T 零空间的向量，将取得局部保持投影意义上的最优分类效果。考虑式(4)和引理，本文算法试图寻找一个投影矩阵，该投影矩阵同时对角化 XLX^T 和 XDX^T ，使得

$$WXDX^T W^T = I, \quad WXLX^T W^T = A \quad (5)$$

其中， A 是一个对角矩阵，它的对角元素以升序排列。为了降低维数到 $d (d \ll n)$ ，可以简单地挑选 W 的前 d 行，它们对应 A 中最小的 d 个对角元素。

本文算法的具体步骤如下：

(1) 用特征值分析的方法对角化 XDX^T

寻找矩阵 V 使 $VXDX^T V^T = A_1$ ，其中， $VV^T = I$ (I 代表单位阵)； A_1 是一个对角矩阵，且对角元素以降序排列。可以通过传统的特征值分解方法得到，即 V 的每一行是 XDX^T 的一个特征向量， A_1 包含了所有特征值。因为 XDX^T 奇异，所以可能一些特征值是零(或很接近零)。需要剔除此类特征值和特征向量，因为在这些方向的投影对局部保持投影而言没有任何有用的鉴别信息。仅保留 V 中特征值非零的主成分，即取 V 的前 m 行，令 Y (一个 $m \times n$ 的矩阵， n 为特征空间的维数) 为新的基，且 D_d 为对应于非零特征值的对角矩阵，则可以得到 $YXDX^T Y^T = D_d > 0$ ， D_d 是一个 $m \times m$ 的矩阵。 m 的个数可以通过检验非零特征值的个数自动确定。而 PCA+LDA 和

PCA+LPP 需要确定 PCA 步骤中主成分的个数。

(2) 投影和对角化 XLX^T

令 $Z = D_d^{-1/2} Y$ ，则 $ZXDX^T Z^T = D_d^{-1/2} YXDX^T Y^T D_d^{-1/2} = I$ ，使 XDX^T 白化。利用另一个特征值分析 $U(ZXLX^T Z^T)U^T = D_l$ 来因式分解 $ZXLX^T Z^T$ ，其中， $UU^T = I$ ； D_l 是一个对角矩阵。保留 $ZXLX^T Z^T$ 对应于最小特征值的特征向量，并将此类特征值按升序排列。

(3) 计算投影矩阵 A

令 $W = UZ$ ，则很容易验证 W 满足式(4)。可以选取 W 的前 d 行，记为矩阵 A ，它们对应 D_l 中 d 个最小的对角元素。对于一个给定的 n 维的输入 x ，它在特征空间的局部保持投影向量为 $y = Ax$ ，此时 y 的维数降为 d 维，且 $d \ll n$ 。

本文算法的计算流程如图 1 所示。



图 1 本文算法的计算流程

4 实验结果

对本文算法在 ORL 人脸库进行测试，图 2 给出了 ORL 人脸库中一个人的 6 幅图像。在人脸识别实验中，每个人任选 5 幅图像作为训练集，其他人脸图像作为测试集。在训练集和测试集中都有 200 幅人脸图像，所有人脸图像尺寸被压缩为 46×56 。首先分析本文算法在不同可调因子的识别率，结果如图 3 所示。然后采用文献[4]的做法，以抽取的样本维数为变量，分析该算法在不同分类器中的正确率，结果如图 4 所示。最后对已有的 PCA, PCA+LDA, PCA+LPP 和本文算法作特征提取，采用最小近邻分类器分类，比较它们的识别率。对 PCA+LDA 和 PCA+LPP，采用文献[5]的做法，对 PCA+LDA 方法，则在 PCA 步骤中保留 $(M-c)$ 个主成分， M 是训练样本的个数， c 是类别的个数。对于 PCA+LPP，在 PCA 步骤中保留 98% 的重建误差来确定主成分个数。



图 2 ORL 人脸库中一个人的 6 幅图像

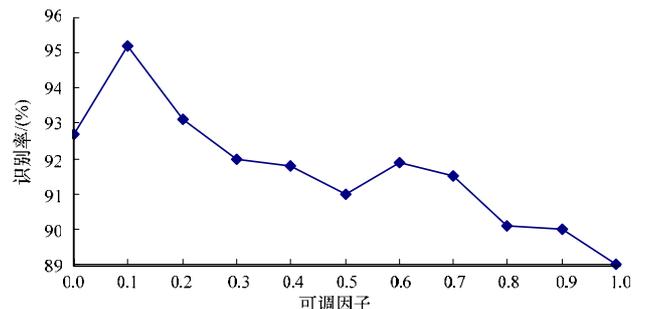


图 3 ORL 人脸库上不同可调因子下的识别率

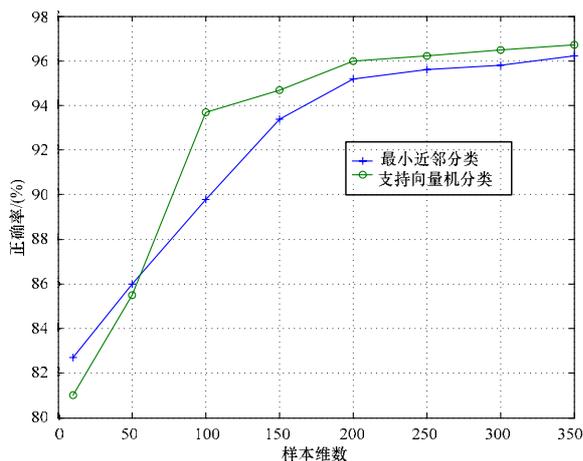


图 4 ORL 人脸库上不同分类器的性能比较

图 3 表示在 ORL 人脸库上,本文算法在不同可调因子 θ 值下的识别率。可以看出,当 $\theta=0.1$ 时,其识别率最高。图 4 给出了本文算法分别采用最小近邻分类器和支持向量机(SVM)分类器^[6]进行分类时的性能比较,可以看出,当样本维数大于 60 时,SVM 分类器的性能大于最小近邻分类器,且会随着样本维数的增加而增加。当维数大于 200 时,2 个分类器的性能基本趋向稳定。对此结果解释如下:特征信息基本集中分布在样本空间的一个的子空间中,如果在分类前对样本进行适当的特征抽取和降维,可以极大减少时间消耗,并能降低分类器性能的损失。

不同方法的性能比较如表 1 所示。

表 1 不同方法的性能比较

方法	训练时间/s	特征维数	单个样本识别时间/s	识别率(%)
PCA	3.11	82	0.015	90.1
PCA+LDA	5.12	40	0.052	90.8
PCA+LPP	78.85	76	0.425	92.4
本文算法($\theta=0.1$)	35.12	202	0.073	95.2

由表 1 可以看出,PCA 算法、PCA+LDA 算法、PCA+LPP 算法、本文算法的识别率分别为 90.1%、90.8%、92.4%、95.2%。从上述结果可以看出,由于本文算法引入了样本的类别信息,并采用同时对角化的方法进行计算,因此对高维的图像数据而言是一个较优的求解方案,取得了比其他 3 种算法更高的识别率。

5 结束语

本文提出一种新的线性子空间方法并将其应用于人脸识别中。该方法在传统 LPP 算法的基础上进行改进,与其他方法相比,取得了较高识别率。它适用于人脸识别,并能推广到其他图像识别问题。本文算法属于线性监督学习方法,为了进一步提高分类性能,可以将它发展为非线性方法。

参考文献

- [1] Zhao W, Chellappa R, Phillips P J, et al. Face Recognition: A Literature Survey[J]. ACM Computing Surveys, 2003, 35(4): 399-459.
- [2] 刘青山, 卢汉清, 马颂德. 综述人脸识别中的子空间方法[J]. 自动化学报, 2003, 29(6): 900-911.
- [3] Feng Guiyu, Hu Dewen, Zhou Zongtan. A Direct Locality Preserving Projections Algorithm for Image Recognition[J]. Neural Processing Letters, 2008, 27(3): 247-255.
- [4] Moghaddam B, Yang M H. Gender Classification with Support Vector Machines[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(5): 707-711.
- [5] He Xiaofei, Yan Shuicheng, Hu Yuxiao. Face Recognition Using Laplacian Faces[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(3): 328-340.
- [6] 鹿晓亮, 陈继荣, 黄戈祥. ICA 和改进的 SVM 在有限集字符识别中的应用[J]. 计算机工程, 2006, 32(23): 199-201.

编辑 陈 晖

(上接第 189 页)

从表 3 中可以得出以下结论:对于单字音说话人识别,3 种方法均能达到较高的识别率,本文方法在识别率方面有一定的优势,系统的识别时间更短。对于连续语音识别,单字特征矢量量化法采用了单字特征分别评分的方式,使识别过程能更好地适应自然发声方式带来的干扰,具有更强的自适应性和鲁棒性。

6 结束语

本文提出以单字音为特征提取单位的说话人特征提取方法,利用 DTW 算法和相邻帧矢量角度提取各单字帧的归整的特征矢量。最后利用这些矢量为每个说话人的每个单字音寻找码本中心。这种方法能更好地保留说话人发声过程中的片断信息。实验表明利用该特征实现的说话人识别系统具有较好的性能。但是这种方法需要对每个单字建立模板,只能运用于发音种类较少的语音环境中。采用分级搜索方式实现匹配的简化有望在一定时间范围内进一步提高发音种类的数量,使该识别系统的适用面更广。

参考文献

- [1] Kinnunen T. Spectral Features for Automatic Text-independent

Speaker Recognition[D]. [S. l.]: University of Joensuu, 2003.

- [2] 董志峰, 汪增福. 基于动态 MFCC 的说话人识别算法[J]. 模式识别与人工智能, 2005, 5(5): 596-601.
- [3] 李 波, 王成友, 杨 聪, 等. 基于语音频谱包络抽取的 MFCC 算法[J]. 国防科技大学学报, 2004, 26(4): 42-45.
- [4] Reynolds D A. An Overview of Automatic Speaker Recognition Technology[C]//Proc. of ICASSP'02. Orlando, Florida, USA: IEEE Press, 2002: 4072-4075.
- [5] Rabiner L, Juang Biinghwang. Fundamentals of Speech Recognition[M]. Beijing: Tsinghua University Press, 1999.
- [6] 刘敬伟, 徐美芝, 郑忠国, 等. 基于 DTW 的语音识别和说话人识别的特征选择[J]. 模式识别与人工智能, 2005, 18(1): 50-54.
- [7] Chen C T, Chen Chinta, Lung Shungyung. Efficient Genetic Algorithm of Codebook Design for Text-independent Speaker Recognition[J]. IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, 2002, 85(11): 2529-2531.
- [8] 于保华, 吴亚栋, 李治柱. 基于长时频谱统计参数的说话人识别[J]. 计算机工程, 2004, 30(2): 159-161.

编辑 任吉慧