

监督式正交迹比判别投影在图像集人脸识别中的应用^①

张 强^{②*} 蔡云泽^{**}

(* 淄博职业学院电子电气工程学院 淄博 255314)

(** 上海交通大学电子信息与电气工程学院 上海 200240)

摘 要 研究、分析了人脸识别中提取原始数据特征的已有方法,在此基础上给出了一种应用监督式正交迹比判别投影(SOTRDP)的新型特征提取方法,即 SOTRDP 方法。不同于现有的非监督判别投影(UDP)方法,SOTRDP 方法能够同时利用局部信息和类别信息建立相似性矩阵。在利用改进局部切空间对齐(ILTSA)非线性降维的基础上,利用聚类中心或最靠近它的样本作为输入,拓展 SOTRDP 用于图像集人脸识别。在 PIE 和 Honda/UCSD 人脸数据库上的实验结果验证了所提方法的有效性。

关键词 非监督判别投影(UDP), 监督式正交迹比判别投影(SOTRDP), 改进局部切空间对齐(ILTSA), 图像集人脸识别

0 引 言

在人脸识别任务中,从原始数据中提取特征是一个重要的步骤,目的是将输入数据投影到一个能够反映原始人脸数据内在结构的特征空间。众所周知的主成分分析(PCA)^[1]和线性判别分析(LDA)^[2]考虑的都是数据的全局欧氏结构,没有考虑数据的流形结构。如果数据位于能够反映数据空间内在结构的子流形上,PCA 和 LDA 都难以发现数据中的潜在流形。新的特征提取技术,比如局部线性嵌入(LLE)^[3]、等距映射(ISOMAP)^[4]、局部保持投影(LPP)^[5]和它们的一些拓展方法^[6,7],常用于表征数据的子流形结构。其中,LPP 是从对稠密黎曼流形上拉普拉斯 Beltrami 算子的离散逼近导出的,通过建立数据点集之间的邻接图来包含流形上局部邻域内的邻接信息。Yang 等人在局部散度和非局部散度概念的基础上提出了非监督判别投影(unsupervised discriminant projection, UDP)^[8],其准则函数通过非局部散度与局部散度之比的最大化来描述,目的是使得投影变换后,相互邻近的样本仍然相互邻近,同时相互疏远的样本依然相互远离。

基于 UDP 的特征提取算法是一种线性算法,它能同时保持局部散度最小化和全面散度最大化,但没有考虑数据的类别信息。针对这种情况,本文提出了一种基于监督式正交迹比例判别投影(supervised orthogonal trace ratio discriminant projection, SOTRDP)的新型的特征提取算法。在人脸识别中,训练样本的类别标签常常已经给出,类标签信息能够帮助提取更具判别力的特征。本文所提出的算法,在设计相似性矩阵时既考虑了局部信息也考虑了数据的类别信息,通过设计一个迹比型目标函数和迭代计算步骤,得到一组正交判别基向量。最近基于图像集的人脸识别算法研究日益受到重视^[9-11]。同基于每人少量训练图像的识别方法相比,在图像分辨率低,外部照明条件、表情或姿态变化大的情况下,基于图像集的人脸识别算法能获得更好的性能。基于 SOTRDP,本文提出了一种新的图像集人脸识别方法。该方法根据对图像集的改进的局部切空间对齐(improved local tangent space alignment, ILTSA)^[12]非线性降维结果,选择分级聚类中心或最靠近它的图像用作 SOTRDP 的训练样本。在 SOTRDP 特征提取的基础上,被测图像集的分类,采用对每一帧进行多数投票方法求得。

① 国家自然科学基金(61004088),上海市基础研究重点(09JC1408000)资助项目。

② 男,1974 年生,博士,讲师;研究方向:模式识别,机器学习,图像处理;联系人,E-mail: zhangqiang741002@sina.com (收稿日期:2011-11-29)

1 监督式正交迹比判别投影(SOTRDP)

本节提出用一种新颖的 SOTRDP 方法进行特征提取,并通过考虑数据的类别信息,设计基于迹比目标函数的迭代学习算法。

1.1 判别相似性矩阵

假设数据集 $Y = [y_1, y_2, \dots, y_N]$ 表示数据集 $[x_1, x_2, \dots, x_N]$ 投影到 $d(d < n)$ 维子空间后的点集,那么投影可以表示为 $y_i = A^T x_i$, 其中 A 为线性变换矩阵。为此我们可以构造一个加权图 $G = (V, E, W)$, 其中 V 是所有点构成的集合, E 是连接点的边构成的集合, W 是表征两点之间相似性的权值构成的相似性矩阵。在分类问题里面,类别标签信息可用来指导特征提取。假设给定的观察为 (x_i, τ_i) ($i = 1, \dots, N$), 其中 $x_i \in R^n$ 并且 τ_i 是类标签, $\tau_i \in \{1, 2, \dots, c\}$, 其中 c 是类别总数。假设 N_i 是包含 x_i 的 k 个最近邻的集合,那么它可以进一步分为只包含同类最近邻样本的类内集合 N_i^w 和只包含异类最近邻样本的类间集合 N_i^b 。进而我们可以定义对应每个样本的判别因子,表示每个样本对分类的贡献。事实上,更加有效的分类应该是让处于边界的点向各自的类别中心移动并远离其它不同类别的点。换句话说,样本越靠近中心,对分类的贡献率越小;反之,越远离中心并越靠近边界的样本,对分类的贡献率越大。因此,定义如下判别因子来表示每个样本的分类贡献率大小:

$$u_i = \frac{\sum_{j, x_j \in N_i^b} \exp(-\|x_i - x_j\|^2 / \delta^2)}{\sum_{j, x_j \in N_i} \exp(-\|x_i - x_j\|^2 / \delta^2)} \quad (1)$$

其中, δ 为一自定义参数。按本文方法,两个点 x_i 和 x_j 之间的判别相似性定义为

$$W_{ij} = \begin{cases} \exp(-\frac{\|x_i - x_j\|^2}{\beta})(1 + u_i u_j) & (x_i \in N_j \text{ 或 } x_j \in N_i, \text{ 且 } \tau_i = \tau_j) \\ \exp(-\frac{\|x_i - x_j\|^2}{\beta})(1 - u_i u_j) & (x_i \in N_j \text{ 或 } x_j \in N_i, \text{ 且 } \tau_i \neq \tau_j) \\ 0 & (\text{其他}) \end{cases} \quad (2)$$

其中欧氏距离 $\|x_i - x_j\|$ 在指数上,参数 β 用作调整参数。我们选择 β 为所有数据点对间平均欧氏距离的平方。这里 $\exp(-\|x_i - x_j\|^2 / \beta)$ 为局部权值, $(1 + u_i u_j)$ 和 $(1 - u_i u_j)$ 分别为类内和类间判别

权值,判别相似性既能反映局部邻域结构又能反映数据集的类别信息。

1.2 SOTRDP 算法

有了如式(2)定义的判别相似性矩阵 W , 我们定义如下的优化问题:

$$A^* = \arg \max_{A^T A = I} \frac{\text{tr}(A^T X L_b X^T A)}{\text{tr}(A^T X L_w X^T A)} \quad (3)$$

其中, $H_{ij} = 1 - W_{ij}$, Q 是对角矩阵, $Q_{ii} = \sum_{j=1}^N H_{ij}$, $L_b = Q - H$ 为拉普拉斯矩阵。 D 是对角矩阵, $D_{ii} = \sum_{j=1}^N W_{ij}$, $L_w = D - W$ 为拉普拉斯矩阵。

通过采用正交迹比目标函数,期望能够获得比求解迹比问题更好的解。为了直接求解迹比问题,本文不加证明地引用如下定理:

定理 1[文献[13], Th. 2] 为了求解迹比问题(式(3)),可以等效地求解相应的迹差问题。也就是说,首先可以找到迹差函数

$$f(\lambda) = \max_{A^T A = I} \text{tr}[A^T (X L_b X^T - \lambda X L_w X^T) A] \quad (4)$$

的零点 λ^* , 即求解一个迹差方程 $f(\lambda) = 0$, 然后可以求解 A^* , 得出:

$$A^* = \arg \max_{A^T A = I} \text{tr}[A^T (X L_b X^T - \lambda^* X L_w X^T) A] \quad (5)$$

为了求解提出的监督式正交迹比判别投影(SOTRDP)问题,所设计的基于 DNMI 迭代方法^[14]的监督式正交迹比判别投影算法步骤总结如下:

步骤 1: 对于每个数据点 x_i , 通过 k 最近邻(k -nearest neighbor, KNN)算法决定它的 k 个最近邻,利用式(1)计算每个样本 x_i 的判别因子 u_i 。

步骤 2: 依据式(2)构造判别相似性矩阵 W , 计算拉普拉斯矩阵 L_w 和 L_b 。

步骤 3: 初始化 $\lambda_0 = 0$, $t = 0$ 。

步骤 4: 特征值分解 $X L_b X^T - \lambda_t X L_w X^T$ 。

步骤 5: 应用一阶泰勒展开逼近 λ_t 附近的特征值:

$$\hat{\beta}_k(\lambda) = \beta_k(\lambda_t) - \omega_k^T(\lambda_t) X L_w X^T \omega_k(\lambda_t) (\lambda - \lambda_t) \quad (6)$$

其中, $\beta_k(\lambda_t)$ 是 $X L_b X^T - \lambda_t X L_w X^T$ 的第 k 个特征值, $\omega_k(\lambda_t)$ 是其对应的单位特征向量。

步骤 6: 逼近迹差函数 $f(\lambda)$ 为 $\{\hat{\beta}_1(\lambda), \dots, \hat{\beta}_m(\lambda)\}$ 中最大的 d 个值的和:

$$\hat{f}_D(\lambda) = \max_i \sum_{k=1}^d \hat{\beta}_{i_k}(\lambda) \quad (7)$$

步骤 7: 通过求解 $\hat{f}_D(\lambda) = 0$ 求得 λ_{t+1} 。

步骤 8: 如果 $|\lambda_{t+1} - \lambda_t| < \varepsilon$, 转步骤 9, 否则, $t = t + 1$, 转步骤 4。

步骤 9: 输出 $\lambda^* = \lambda_t$, 求解 A^* 为对应 $XL_b X^T - \lambda^* XL_w X^T$ 最大 d 个特征值的特征向量。

2 实验结果与讨论

基于图像集的人脸识别可以这样表述: 给出一个训练人脸图像序列 $G = \{g_1, g_2, \dots, g_N\}$, 从中选出最有代表性的样本

$$E = \{e_1, e_2, \dots, e_K\}, \quad K \ll N \quad (8)$$

本研究希望选出的样本能够包含照明、姿态、表情和其它类内人脸变化因素。不同于以前的方法^[11], 为了能够揭示人脸图像集的内在非线性结构, 本文首先将训练人脸图像序列投影到改进的局部切空间对齐 (ILTSA) 空间, 然后采用分级聚类方法选择代表图像。其中, 分级聚类方法同时在多个尺度上, 通过生成一个聚类树, 把数据划分为许多嵌

套的数据集合。在实验中, 分级聚类的步骤包括:

(1) 初始化聚集合 $C_i, i = 1, 2, \dots, c$ 。

(2) 找到最近的两个聚, 比如 C_i 和 C_j , 并把它们融合成一个新的聚。不断重复这一步骤, 直到达到指定的聚数目。

选择聚类中心或者最靠近每个聚类中心的图像帧作为代表样本, 实施 SOTRDP 特征提取算法。为了判定测试图像序列的所属类别, 采用多数投票方法组合测试集中每一帧的判别结果。

2.1 在 PIE 数据库上的实验结果

CMU PIE 数据库^[15] 包含 68 个人的总共 41368 张人脸图像。其中, 人脸图像是用 13 个同步照相机和 21 个闪光灯, 在不同的姿态、照明和表情变化条件下拍摄的。从 PIE 数据库中, 选择 5 个接近正面的姿态 (C05, C07, C09, C27, C29), 其中每人 170 张图像的子数据库。每张图像以双眼位置为基准对齐, 并裁剪、缩放到 32×32 尺寸大小。图 1 给出了一个人的 40 张图像样本。

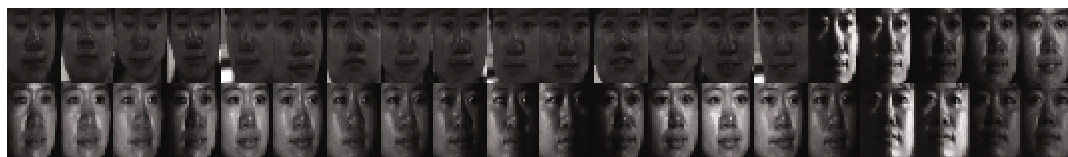


图 1 PIE 实验数据库中一个人的例子图像

从每人 170 张图像中, 随机抽取 80 张作为训练图像集, 剩余的 90 张用于测试。从每人 90 张测试图像中, 随机抽取长度为 40 的图像集 20 次。形成每人 20 个测试图像集。在实验中, 首先固定每类选

出的代表样本数目为 5, ILTSA 算法和 SOTRDP 算法的最近邻数分别选为 5 和 8。图 2 给出了从训练序列中提取出的一个人的聚类中心代表图像。



图 2 从 PIE 训练序列中提取出的代表图像

本实验比较了 SOTRDP 方法和其它几个最近提出的典型特征提取方法, 20 次实验平均测试结果如表 1 所示, 其中 CMVM 为约束最大差异映射 (constrained maximum variance mapping), LLDE 为局部线性判别嵌入 (locally linear discriminant embedding)。最优约简维数的选择根据实验设定。从表 1

可以看出, 由于能够提取非线性人脸数据流形中的有用信息, 基于局部学习的特征提取方法在性能上优于全局方法。更重要的是, 本文所提出的 SOTRDP 方法取得了比其它方法更好的正确识别率。

表 1 在 PIE 数据库上的实验结果

样本选择方法	PCA	UDP	LDA	CMVM ^[6]	LLDE ^[7]	SOTRDP
随机选择	38.87	69.96	70.89	70.52	71.64	72.35
LLE + k 均值[11]	61.58	84.52	86.56	87.69	87.89	88.51
ILTSA + 分级聚类	64.67	86.54	88.04	88.13	88.36	89.28

图 3 给出了改变选出的代表样本数目(聚类数)时三种局部学习方法的识别率与聚类数的关系

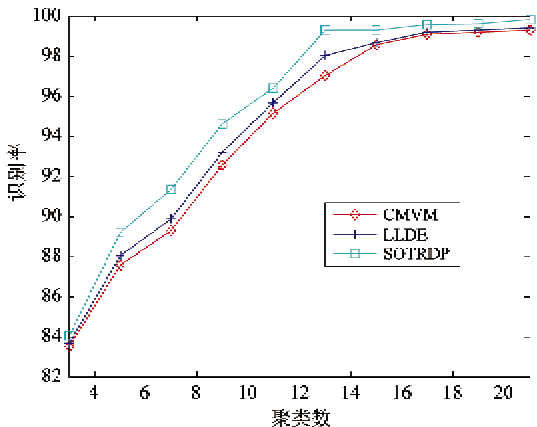


图 3 PIE 数据库中识别率与样本数关系

曲线。从图中可以看出,随着聚类数的增加,三种方法的识别率相应也增加,但 SOTRDP 方法能取得更好的结果。

2.2 在 Honda/UCSD 图像集数据库上的实验结果

本文进行了将 SOTRDP 算法用于 Honda/UCSD 图像集人脸识别的实验。Honda/UCSD 数据库^[16]是一个常用的视频人脸识别数据集。在本实验中,应用它的第一个数据集。该数据集由 20 个人的 59 个视频序列构成。每个视频包含大约 300-500 帧图像,图像变化包括大的姿态和尺度变化,头部的平面内、外旋转和表情变化。图 4 给出其中一个人的例子图像。其中的图像是通过 Viola-Jones 人脸检测方法检测到的每帧图像,图像尺寸为 32×32 。为了补偿照明变化的影响,我们对图像进行了简单的直方图均衡化预处理。



图 4 从 Honda/UCSD 视频数据库中提取出的一个人的例子图像

对于每个人,其中的一个视频用于训练,其余的视频用于测试。从每个人的测试序列中,我们随机抽取长度为 200 的图像集 20 次。这样形成每人 20 个测试图像集,总共 $20 \times 19 = 380$ 个被测试图像集。在实验中,我们固定每类选出的代表样本数目为 7, ILTSA 算法和 SOTRDP 算法的最近邻数分别选为 7

和 12。图 5 给出了从训练视频序列中提取出的三个人的代表图像。本实验比较了 SOTRDP 算法和其它几个最近提出的典型特征提取方法,20 次实验平均测试结果如表 2 所示,其中最优约简维数的选择根据实验设定。从表中可以看出,本文所提出的 SOTRDP 取得了比其它方法更好的正确识别率。



图 5 从训练视频序列中提取出的代表图像

表 2 在 Honda/UCSD 图像集数据库上的实验结果

样本选择方法	PCA	UDP	LDA	CMVM ^[6]	LLDE ^[7]	SOTRDP
随机选择	31.56	54.64	61.33	64.26	65.01	69.77
LLE + k 均值 ^[12]	60.35	68.57	76.14	79.97	79.36	82.54
ILTSA + 分级聚类	62.06	71.31	79.44	80.23	80.79	83.78

图 6 给出了改变选出的代表样本数目(聚类数)时三种局部学习方法的识别率与聚类数的关系曲线。从图中可以看出,随着聚类数的增加,三种方法的识别率相应也增加,但 SOTRDP 算法能取得更好的结果。

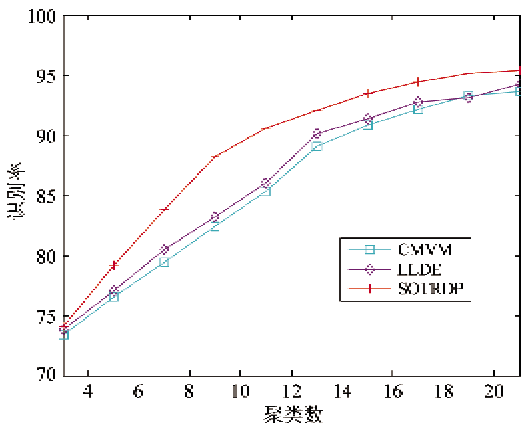


图 6 Honda 数据库识别率与样本数关系

3 结论

从原始图像数据中提取特征是实现人脸识别的重要步骤。本研究提出了一种新颖的应用监督式正交迹比判别投影(SOTRDP)的特征提取方法,另外提出了一种基于改进的局部切空间对齐(ILTSA)典

型样本选择和 SOTRDP 法特征提取的图像集人脸识别方法。基于 PIE 和 Honda/UCSD 人脸数据库的人脸识别实验结果验证了该算法设计和理论分析的合理性,以及算法的有效性。将来的工作是在其它复杂人脸数据库上的算法验证,以及将所提方法用于其它生物特征识别。

参考文献

[1] Turk M A, Pentland A P. Face Recognition Using Eigenfaces. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Maui, USA, 1991. 586-591

[2] Belhumeur P N, Hespanha J P, Kriegman D J. et al. Fisherfaces: Recognition Using Class-Specific Linear Projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997, 19(7): 711-720

[3] Roweis S T, Saul L K. Nonlinear Dimensionality Reduction by Locally Linear Embedding. *Science*, 2000, 290(5500):2323-2326

[4] Tenenbaum J B, Silva V, Langford J C. A Global Geometric Framework for Nonlinear Dimensionality Reduction. *Science*, 2000, 290(5500):2319-2323

[5] He X F, Yan S C, Hu Y X, et al. Face Recognition Using Laplacianfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, 27(3): 328-340

- [6] Li B, Huang D S, Wang C, et al. Feature extraction using constrained maximum variance mapping, *Pattern Recognition*, 2008, 41(11): 3287-3294
- [7] Li B, Zheng C, Huang D. Locally linear discriminant embedding: an efficient method for face recognition, *Pattern Recognition*, 2008, 41(12): 3813-3821
- [8] Yang J, Zhang D, Yang J Y, et al. Globally maximizing, locally minimizing: unsupervised discriminant projection with application to face and palm biometrics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, 29(4): 650-664
- [9] Wang T, Shi P. Kernel grassmannian distances and discriminant analysis for face recognition from image sets. *Pattern Recognition Letters*, 2009, 30(13): 1161-1165
- [10] Mehrtash T H, Conrad S, Sareh S, et al. Graph embedding discriminant analysis on grassmannian manifolds for improved image set matching. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Providence, USA, 2011. 2705-2712
- [11] Fan W, Yeung D. Face recognition with image sets using hierarchically extracted exemplars from appearance manifolds. In: Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition, Southampton, UK, 2006. 177-182
- [12] Zhang P, Qiao H, Zhang B. An improved local tangent space alignment method for manifold learning. *Pattern Recognition Letters*, 2011, 32: 181-189
- [13] Guo Y, Li S, Yang J, et al. A generalized foley-sammon transform based on generalized fisher discriminant criterion and its application to face recognition. *Pattern Recognition Letters*, 2003, 24(1-3): 147-158
- [14] Jia Y, Nie F, Zhang C. Trace ratio problem revisited. *IEEE Transactions on Neural Networks*, 2009, 20(4): 729-735
- [15] T. Sim, S. Baker, M. Bsat. The CMU pose, illumination, and expression database. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, 25(12): 1615-1618
- [16] Lee K C, Ho J, Yang M H, et al. Visual tracking and recognition using probabilistic appearance manifolds. *Computer Vision and Image Understanding*, 2005, 99(3): 303-331

Application of supervised orthogonal trace ratio discriminant projection to image set face recognition

Zhang Qiang*, Cai Yunze**

(* College of Electronic and Electrical Engineering, Zibo Vocational Institute, Shandong, 255314)

(** School of Electronic, Information and Electrical Engineering, Shanghai Jiaotong University, Shanghai 200240)

Abstract

The existing feature extraction methods for face recognition were studied and analyzed, and based on this, a novel feature extraction method using supervised orthogonal trace ratio discriminant projection (SOTRDP), called the SOTRDP method for short, was proposed. Unlike the unsupervised discriminant projection (UDP) method, this new method can simultaneously use both local information and class information to model the similarity of the data. Based on the improved local tangent space alignment (ILTSA) nonlinear dimensionality reduction, the SOTRDP was extended to image set face recognition application by using the clustering centers as input. The results of the experiments on the PIE and Honda/UCSD image set face databases demonstrate the effectiveness of the proposed face recognition method.

Key words: unsupervised discriminant projection (UDP), supervised orthogonal trace ratio discriminant projection (SOTRDP), improved local tangent space alignment (ILTSA), image set face recognition