文章编号:1671 - 6833(2012)03 - 0125 - 04

基于类别多核局部判别嵌入的人脸识别

王永茂[1,2]，徐正光[1]，吴金霞[1]

(1. 北京科技大学 自动化学院,北京 100083; 2. 河南理工大学 计算机科学与技术学院,河南 焦作 454003)

摘　要: 在局部判别嵌入的基础上提出了一种有效的非线性子空间学习方法:类别多核局部判别嵌入. 首先针对给定数据的类别信息,定义基于每一个类别的局部核函数,形成多核,接着将不同的局部核函数进行线性组合作为最终的核函数引入到局部判别嵌入算法中,得到类别多核局部判别嵌入算法,在核空间内提取图像高阶非线性信息. 在 ORL 和 Yale 库上的人脸识别表明该方法是有效的.

关键词: 人脸识别;子空间学习;类别多核;降维

中图分类号: TP391.4　　**文献标志码**: A　　doi:10.3969/j.issn.1671 - 6833.2012.03.032

0　引言

为了缓解人脸识别中存在的维数灾难问题,需要进行有效的降维. 主元成分分析(PCA, Principle Component Analysis)[1]是最流行的降维方法之一,但其本质是线性的,不能很好地描述人脸图像中诸如光照、表情和姿态等复杂的非线性变化. 基于流形的降维方法是近年来兴起的一类降维方法,其中代表性的算法有等距映射(ISOMAP)[2]、局部线性嵌入(LLE, Locally Linear Embedding)[3]和拉普拉斯映射(LE, Laplacian Eigenmap)[4],它们能够很好学习到非线性流形结构,但是基于流形的降维方法的一大缺陷是不能直接映射新的测试点,为了解决这一问题,许多学者先后提出局部保形投影(LPP, Locality Preserving Projection)[5]、近邻保形嵌入(NPE, Neighborhood Preserving Embedding)[6]等线性降维算法,它们分别为 LE 算法与 LLE 算法的线性逼近.

上述提到的线性或非线性降维方法的设计初衷并不是应用于分类,而 Fisher 判别分析(FDA, Fisher Discriminant Anaysis)[7]得到的最佳映射子空间具有较好的鉴别能力,因此对于分类任务能够起到较好的效果. 为了保持数据的流形结构,研究者提出了一类既考虑类别信息又考虑邻域结构信息的降维方法[8-13],这些降维方法既具有较好的鉴别能力又能很好保持数据的近邻特性. 局部判别分析(LDE, Local Discrimiant Embedding)[8]是其典型代表. 然而,LDE 本质上仍然是线性方法,不能够很好揭示数据复杂的非线性结构,许多学者提出利用非线性核技巧在维数很高的空间内求解子空间的核局部判别嵌入(KLDE, Kernel Local Discriminant Embedding)降维方法[8,14]. 该方法将输入的数据空间转换到一个高维空间,其映射关系在全局范围内定义,在映射的过程中改变了数据分布,不能够很好保持数据的几何结构. 为解决这一问题,笔者提出一种基于类别多核局部判别嵌入算法(LMKLDE, Label Mulitple Kernel Local Discrimiant Embedding),首先针对给定样本数据的类别信息,定义类别局部核函数,形成多核,接着将不同的局部核函数进行线性组合,作为最终的核函数引入到 LDE 算法中,得到了 LMKLDE 算法.

1　LDE

假设有一数据集 $X = \{(x_1, l_1), (x_2, l_2), \cdots, (x_N, l_N)\}$, x_i 是一个 D 维向量, N 为样本的个数, $l_i \in L = \{1, 2, \cdots, m\}$ 为 x_i 的类别标签, L 是类别标签集, m 为类别总数. LDE 算法的目的就是寻找一投影矩阵 V, 得到 X 的低维嵌入 $Y = \{y_1, y_2, \cdots, y_N\}$, 使得 $Y = V^{\mathrm{T}}X$, y_i 是一个 d 维向量, 且 $d < D$.

根据图框架理论[9], LDE 算法构造两个近邻图:(1) 描述类内紧密性的近邻图 G_w(固有图),图中每个样本只与同类别的 k_1 个近邻点相连, $N_w(x_i)$ 表示样本 x_i 同类别的 k_1 个近邻点; 2) 描

收稿日期:2011 - 08 - 17;修订日期:2012 - 01 - 15

作者简介:王永茂(1976 -),男,河南焦作人,河南理工大学副教授,博士,现主要从事图像处理与模式识别方面的研究工作,E-mail:wymyjs2000@hpu.edu.cn.

述类间分离性的近邻图 G_p（惩罚图）：图中每个样本只与不同类别的 k_2 个近邻相连，$N_p(x_i)$ 表示样本 x_i 不同类别的 k_2 个近邻点.

计算近邻图 G_w 和 G_p 的权值矩阵 S_w 和 S_p，其值为

$$S_{w,ij}=\begin{cases}\exp(-\|x_i-x_j\|^2/t), & x_j\in N_w(x_i) \text{ 或 } x_i\in N_w(x_j)\\ 0 & \text{其他}\end{cases};\tag{1}$$

$$S_{p,ij}=\begin{cases}\exp(-\|x_i-x_j\|^2/t), & x_j\in N_p(x_i) \text{ 或 } x_i\in N_p(x_j)\\ 0 & \text{其他}\end{cases}.\tag{2}$$

为了保证 x_i 和 x_j 在相距很近且具有相同类别标记时，其低维嵌入 y_i 和 y_j 相距也很近；x_i 和 x_j 在相距很近但具有不同类别标记时，y_i 和 y_j 相距很远. LDE 算法的目标函数定义为

$$\begin{aligned}&\text{Maximize } J(V)=\sum_{i,j}\|V^{\mathrm T}x_i-V^{\mathrm T}x_j\|^2S_{w,ij}\\&\text{s.t. }\sum_{i,j}\|V^{\mathrm T}x_i-V^{\mathrm T}x_j\|^2S_{p,ij}=1.\end{aligned}\tag{3}$$

经过推导，LDE 目标函数变为

$$\begin{aligned}&\text{Maximize } J(V)=2\mathrm{tr}\{V^{\mathrm T}XL_wX^{\mathrm T}V\}\\&\text{s.t. }2\mathrm{tr}\{V^{\mathrm T}XL_pX^{\mathrm T}V\}=1.\end{aligned}\tag{4}$$

其中 $L_w=D_w-S_w$ 和 $L_p=D_p-S_p$ 为拉普拉斯矩阵，D_w 和 D_p 为对角矩阵，其值为权值矩阵 S_w 和 S_p 每一列（或行）数据之和，即 $D_{w,ii}=\sum_j S_{w,ij}$，$D_{p,ii}=\sum_j S_{p,ij}$.

式(5) 进行广义特征分解的前 d 个最大特征值对应的特征向量组成了 LDE 的基向量 $V=[v_1,v_2,\cdots,v_d]$.

$$XL_pX^{\mathrm T}v=\lambda XL_wX^{\mathrm T}v.\tag{5}$$

2 类别多核函数

核函数提供了一种数据相似度度量方法，对于两个样本 x 和 z，$\Phi(x)$ 和 $\Phi(z)$ 是样本 x 和 z 的高维非线性映射，则 $\Phi(x)$ 和 $\Phi(z)$ 的点积可用 x 和 z 的核函数表示，即

$$k(x,z)=\langle\Phi(x),\Phi(z)\rangle.\tag{6}$$

下面给出类别多核函数的构造.

对于样本 x，求其到类别 c 中心的矢量 x^c，其值为

$$x^c=x-\bar{x}^c,\ c=1,2,\cdots,m.\tag{7}$$

其中 $\bar{x}^c$ 为类别 c 的中心，其值为 $\bar{x}^c=\frac{1}{n_c}\sum_{i=1}^{n_c}x_i$，$n_c$ 为类别 c 中样本的个数. 样本 x 属于类别 c 的概率 p_x^c 定义为

$$p_x^c=f_x^c/\sum_{j=1}^m f_x^j,\tag{8}$$

其中 $f_x^c=\exp(-\|x-\bar{x}^c\|)$.

定义基于每一个类别的局部核，即

$$K^c(x,z)=\langle\Phi^c(x^c),\Phi^c(z^c)\rangle,\ c=1,2,\cdots,m.\tag{9}$$

将不同的局部核函数进行线性组合，得到最终的核函数，即

$$\begin{aligned}k(x,z)&=\langle\Phi(x),\Phi(z)\rangle\\&=\sum_{c=1}^m p_x^cp_z^cK^c(x,z).\end{aligned}\tag{10}$$

3 LMKLDE

将类别多核函数应用到 LDE 算法，以处理非线性结构数据，得到 LMKLDE 算法.

3.1 LMKLDE 的求解过程

假设 Φ 为一非线性映射函数，将样本 X 映射到希尔伯特空间内[15]，得到 $\Phi(X)=[\Phi(x_1),\Phi(x_2),\cdots,\Phi(x_N)]$. 于是，希尔伯特空间中的广义特征问题表示如下：

$$\Phi(X)L_p\Phi(X)^{\mathrm T}w=\lambda\Phi(X)L_w\Phi(X)^{\mathrm T}w.\tag{11}$$

由于上式的特征向量 w 为 $\Phi(x_1),\Phi(x_2),\cdots,\Phi(x_N)$ 的线性组合，故 w 可表示如下：

$$w=\sum_{i=1}^N\alpha_i\Phi(x_i)=\Phi(X)\alpha.\tag{12}$$

经过简单的变换，式(11) 变为

$$KL_pK\alpha=\lambda KL_wK\alpha.\tag{13}$$

其中：K 为核矩阵，其值由式(10) 定义的类别多核函数决定. 求解式(13) 对应的广义特征向量，得到 d 个特征向量 $\alpha^1,\alpha^2,\cdots,\alpha^d$，分别对应于 d 个最大的特征值. 那么对于一个测试样本 x，它在特征向量 w^k 上的投影为

$$(w^k\cdot\Phi(x))=\sum_{i=1}^N\alpha_i^k(\Phi(x_i)\cdot\Phi(x))=\sum_{i=1}^N\alpha_i^kk(x_i,x).\tag{14}$$

式中：α_i^k 是 α^k 的第 i 个分量. 因此，如果令 $\theta=[\alpha^1,\alpha^2,\cdots,\alpha^d]\in R^{N\times d}$，则新样本 x 的 d 维嵌入可以表示为 $y=\theta^{\mathrm T}K(\cdot,x)$，其中 $K(\cdot,x)=[k(x_1,x),\cdots,k(x_N,x)]$.

3.2 LMKLDE 的时间复杂度分析

KLDE 算法的时间复杂度主要有以下 3 个方

面决定:(1) 近邻点的搜索;(2) 核函数的构造;(3) 广义特征向量问题的求解.

近邻点的搜索的时间复杂度为 $O((D+k_1+k_2)N^2)$,DN^2 代表计算任意两个样本距离的时间复杂度,k_1N^2 代表寻找同类别的 k_1 个近邻点的时间复杂度,k_2N^2 代表寻找不同类别的 k_2 个近邻点的时间复杂度. 核函数的构造的时间复杂度为 $O(DN^2)$. 求解广义特征向量的时间复杂度为 $O(N^3)$.

LMKLDE 与 KLDE 区别仅在于核函数的构造,核函数构造的时间复杂度为 $O(mDn^2)$,大于 KLDE 的核函数构造的时间复杂度 $O(Dn^2)$,因此,LMKLDE 的执行时间要比 KLDE 长,当 $m \ll D$ 时,其执行时间差别不大.

4 实验设计和结果

为了验证算法的有效性,分别在 ORL 和 Yale 人脸库上进行实验对比. 实验中比较了 LMKLDE 与 PCA,FDA,LDE 以及 KLDE 的识别率, 在 KLDE,LMKLDE 方法中, 核函数采用线性核函数,$k(x,z)=(1+x^{T}z)^{d}$,$d=1$;以及高斯核函数,$k(x,z)=\exp(-\|x-z\|^2/2\sigma^2)$,$\sigma=1$,对于所有方法,均采用欧式距离度量下的最近邻分类器完成最终分类.

ORL 人脸库是由英国剑桥大学建立,共有 40 个人,每人 10 张图像,共有 400 张人脸图像,图像的面部表情和面部细节有着不同程度的变化,人脸姿势也有相当的程度变化,比较充分地反映了同一人不同人脸图像的变化和差异. 图 1 是 ORL 人脸库的部分样本. Yale 人脸库由美国耶鲁大学建立,包含 15 个人,每人 11 张图像,共有 165 张人脸图像,主要包括光照条件的变化、表情的变化及有无眼睛修饰等. 图 2 是 Yale 人脸库的部分样本. 实验使用的人脸图像经剪切后大小均为 32 × 32,然后将两个人脸库中的每个图像进行标准化.

在对 ORL 和 Yale 人脸库进行实验时,从每类人脸图像中随机选取 $i(i=2,3,4)$ 张图像作为训练集,剩余的图像作为测试集,重复进行 10 次,共获得 10 对不同的训练集和测试集,分别以 2Train,3Train 和 4Train 表示,最后取平均值作为识别结果. 表 1 和表 2 给出了 ORL 和 Yale 人脸库的平均最高识别率及其对应的子空间的维数,其中(L)为采用线性核函数的识别结果,(G)为采用高斯核函数的识别结果.

图 1 ORL 库中的部分人脸图像

Fig. 1 Some image samples of ORL faces database

图 2 Yale 库中的部分人脸图像

Fig. 2 Some image samples of Yale faces database

表 1 ORL 人脸数据库上的实验结果比较

Tab. 1 The experiment result on ORL database

算法	2Train	3Train	4Train
PCA	66.6% (79)	77.3% (119)	81.4% (159)
FDA	71.0% (28)	83.6% (39)	89.6% (39)
LDE	76.5% (39)	84.8% (39)	90.1% (39)
KLDE(L)	80.6% (39)	87.3% (39)	92.5% (39)
KLDE(G)	80.7% (39)	87.6% (39)	92.4% (39)
LMKLDE(L)	81.3% (39)	88.4% (39)	94.0% (39)
LMKLDE(G)	81.6% (39)	88.5% (39)	94.1% (39)

表 2 Yale 人脸数据库上的实验结果比较

Tab. 2 The experiment result on Yale database

算法	2Train	3Train	4Train
PCA	43.2% (29)	49.5% (44)	53.4% (58)
FDA	47.2% (10)	64.6% (14)	72.7% (14)
LDE	56.1% (14)	68.1% (14)	74.0% (14)
KLDE(L)	59.3% (14)	70.6% (14)	76.8% (14)
KLDE(G)	59.4% (14)	70.8% (14)	77.1 % (14)
LMKLDE(L)	60.5% (14)	71.5% (14)	77.9% (14)
LMKLDE(G)	60.6% (14)	71.5% (14)	77.8% (14)

从上面的实验可以看出:

(1)FDA 考虑样本的类别信息,识别率高于 PCA 方法;

(2)LDE 既考虑样本的类别信息又考虑样本之间的近邻关系,对于非线性流形结构有一定的保持作用,其识别率要高于 FDA;

(3)KLDE 以及 LMKLDE 在核空间提取图像的高阶非线性信息,其识别率高于其他方法,但 KLDE 方法的核函数在全局范围内定义,而 LMKLDE 方法的核函数是类别局部核函数的线性组合,考虑了样本的几何分布,因此 LMKLDE 的识别率高于 KLDE.

5 结论

笔者在 LDE 算法的基础上提出了一种基于类别多核局部判别嵌入的人脸识别算法. 算法通

过将不同类别的局部核函数进行线性组合所得到最终的核函数引入到LDE算法中,得到LMKLDE算法,能够较好的保持数据的几何结构.在ORL和Yale人脸库上的实验结果表明,在姿态、光照和表情等变化的情况下,该算法都具有良好的性能.

参考文献:

[1] TURK M, PENTLAND A. Eigenfaces for recognition [J]. Journal of Cognitive Neuroscience, 1991, 3(1): 72-86.

[2] TENENBAUM J B, SILVA V, LANGFORD J C. A global geometric framework for nonlinear dimensionality reduction[J]. Science, 2000, 290 (12): 2319-2323.

[3] ROWEIS S T, SAUL L K. Nonlinear dimensionality reduction by locally linear embedding [J]. Science, 2000, 290(12):2323-2326.

[4] BELKIN M, NIYOGI P. Laplacian eigenmaps for dimensionality reduction and data representation [J]. Neural Computation, 2003, 15(6):1373-1396.

[5] HE Xiao-fei, YAN Shui-cheng. Face recognition using laplacianfaces[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2005, 27(3):328-340.

[6] HE Xiao-fei, CAI Deng, YAN Shui-cheng, et al. Neighborhood preserving embedding[C]//Tenth IEEE International Conference on Computer Vision. Piscataway: Institute of Electrical and Electronics Engineers Inc., 2005:1208-1213.

[7] BELHUMEUR P N, HESPANHA J P, KRIEGMAN D J. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection [J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 1997, 19(7):711-720.

[8] CHEN HW, CHANG HW, LIU TL. Local Discriminant Embedding and Its Variants [C]//2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Pis cataway: Institute of Electrical and Electronics Engineers computer Society, 2005: 816-853.

[9] YAN Shui-cheng, XU Dang, ZHANG Ben-yu, et al. Graph Embedding and Extensions: A general framework for dimensionality reduction [J]. IEEE transaction on Pattern Analysis and Machine Intelligence, 2007, 29(1):40-51.

[10] 王超,王士同.最大间距准则与局部保持结合的特征提取方法[J].计算机工程,2009,35(14):209-211.

[11] 张召,业宁,业巧林.局部保持多投影向量Fisher判别分析算法[J].计算机学报,2010,33(5):385-396.

[12] 魏莱,王守觉,徐菲菲,等.近邻边界Fisher判别分析[J].电子与信息学报,2009,31(3):509-513.

[13] CAI Hong-ping, KRYSTIAN M, JIRI M. Learning linear discrimiant projection for dimensionality reduction of image descriptors [J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2011, 33(2):338-351.

[14] 王庆军,张汝波,潘海为.基于核正交局部判别嵌入的人脸识别[J].光电子激光,2010,21(9):1386-1389.

[15] 申中华,潘永惠,王士同.有监督的局部保留投影降维算法[J].模式识别与人工智能,2008,21(2):233-239.

Face Recognition Based on Label Multiple Kernel Local Discriminant Embedding

WANG Yong-mao[1,2], XU Zheng-guang[1], WU Jin-xia[1]

(1. School of Automation, University of Science and Technology of Beijing, Beijing 100083, China; 2. School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo 454003, China)

Abstract: Based on local discriminant embedding, an efficient nonlinear subspace learning method, Label Multiple Kernel Local Discriminant Embedding (LMKLDE), is developed. Firstly, according to the label information of given data set, local kernel function is defined and multiple kernel is gained. Then, different local kernel functions are merged by linear combination to form final kernel function. Finally, LMKLDE is developed by introducing label multiple kernel to LDE in order to deal with datasets of highly nonlinear structure. Experiments on ORL and Yale face database demonstrate the effectiveness of the proposed method.

Key words: face recognition; subspace learning; label multiple kernel; dimensionality reduction