



JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象图形学报

2013
08
VOL.18

ISSN1006-8961
CN11-3758/TB



北魏



隋



初唐

数字文化遗产研究 P968

综述

尺度不变特征变换算子综述

刘立, 詹茵茵, 罗扬, 刘朝晖, 彭复员 885

图像处理和编码

改进的H.264帧间模式选择算法

任克强, 张旭光, 罗会兰 893

均质雾天下摄像机动态标定的车速估计

宋洪军, 陈阳舟, 郜园园 901

基于人脸五官特征的空域差错掩盖算法

张江鑫, 谢晋, 邝万坤 913

采用圆谐-傅里叶矩的图像区域复制粘贴篡改检测

秦娟, 李峰, 向凌云, 殷蓓蓓 919

图像分析和识别

基于视频感知哈希的视频篡改检测与多粒度定位

朱映映, 文振焜, 杜以华, 邓良太 924

视觉注意模型的道路监控视频关键帧提取

刘云鹏, 张三元, 王仁芳, 张引 933

格拉斯曼流形上的半监督判别分析

姜伟, 陆瑶, 杨炳儒 944

精神疲劳实时监测中多面部特征时序分类模型

陈云华, 张灵, 丁伍洋, 严明玉 953

图像分类中的概率乘积核函数

杨赛, 赵春霞 961

轮廓整体结构约束的壁画图像相似性度量

唐大伟, 鲁东明, 杨冰, 许端清 968

应用色彩空间聚类方法实现道路建模

向宸薇, 王拓, 于舰 976

图像理解和计算机视觉

杂乱背景和摄像机移动下的时空兴趣点检测

刘长红, 陈勇, 王明文 982

形状和颜色混合不变量在图像检索中的应用

公明, 曹伟国, 李华 990

计算机图形学

结合统计分类和边缘特征的最优视图提取

潘翔, 陈敖, 章国栋 1004

医学图像处理

局部熵驱动下的脑MR图像分割与偏移场恢复耦合模型

张建伟, 杨红, 陈允杰, 方林, 詹天明 1011

改进的模糊聚类亚实质肺结节3维分割

李翠芳, 聂生东, 王远军, 孙希文, 郑斌 1019

遥感图像处理

高光谱海岸带区域分割的活动轮廓模型

王相海, 金弋博 1031

改进的基于非局部均值的极化SAR相干斑抑制

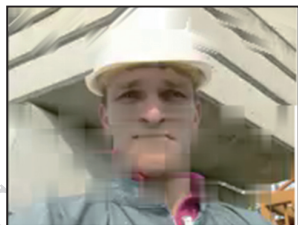
赵忠民, 赵拥军, 牛朝阳 1038

小波变换和稀疏表示相结合的遥感图像融合

刘婷, 程建 1045

《中国图象图形学报》论文摘要写作要求

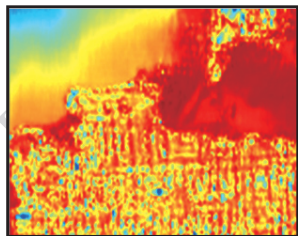
1054



基于人脸五官特征的空域差错掩盖算法 (第913页)



应用色彩空间聚类方法实现道路建模 (第976页)



改进的基于非局部均值的极化SAR相干斑抑制 (第1038页)

CONTENTS

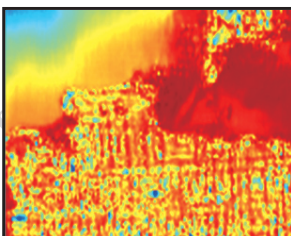
JOURNAL OF IMAGE AND GRAPHICS



Spatial error concealment algorithm based on the facial features (P913)



Road modeling using color spatial clustering (P976)



Improved polarimetric SAR speckle filter based on non-local means (P1038)

Review

Summarization of the scale invariant feature transform

Liu Li , Zhan Yinyin, Luo Yang, Liu Chaohui, Peng Fuyuan 885

Image Processing and Coding

Improved inter frame mode selection algorithm for H.264

Ren Keqiang, Zhang Xuguang, Luo Huilan 893

Vehicle velocity estimation based on camera calibration under homogenous fog

Song Hongjun, Chen Yangzhou, Gao Yuanyuan 901

Spatial error concealment algorithm based on the facial features

Zhang Jiangxin, Xie Jin, Kuang Wankun 913

Detection of image region copy-move forgery using radial harmonic Fourier moments

Qin Juan, Li Feng, Xiang Lingyun, Yin Changming 919

Image Analysis and Recognition

Video forgery detection and multi-granularity location based on video perceptual hashing

Zhu Yingying, Wen Zhenkun, Du Yihua, Deng Liangtai 924

Key frame extraction based on the visual attention model for lane surveillance video

Liu Yunpeng, Zhang Sanyuan, Wang Renfang, Zhang Yin 933

Semi-supervised discriminant analysis on Grassmannian manifold

Jiang Wei, Lu Yao, Yang Bingru 944

Time-series classification model based on multiple facial feature for real-time mental fatigue monitoring

Chen Yunhua, Zhang Ling, Ding Wuyang, Yan Mingyu 953

Probability product kernels in image classification

Yang Sai, Zhao Chunxia 961

Similarity metrics between mural images with constraints of the overall structure of contours

Tang Dawei, Lu Dongming, Yang Bing, Xu Duanqing 968

Road modeling using color spatial clustering

Xiang Chenwei, Wang Tuo, Yu Jian 976

Image Understanding and Computer Vision

Spatio-temporal interest point detection in cluttered backgrounds with camera movements

Liu Changhong, Chen Yong, Wang Mingwen 982

Application of combined shape and color invariants to image retrieval

Gong Ming, Cao Weiguo, Li Hua 990

Computer Graphics

Best view selection based on statistical classification and edge features

Pan Xiang, Chen Ao, Zhang Guodong 1004

Medical Image Processing

Brain MR image segmentation and bias correction model based on local entropy

Zhang Jianwei, Yang Hong, Chen Yunjie, Fang Ling, Zhan Tianming 1011

Segmentation of sub-solid pulmonary nodules based on improved fuzzy

C-means clustering

Li Cuifang, Nie Shengdong, Wang Yuanjun, Sun Xiwen, Zheng Bin 1019

Remote Sensing Image Processing

The active contour model for segmentation of coastal hyperspectral remote sensing image

Wang Xianghai, Yin Yibo 1031

Improved polarimetric SAR speckle filter based on non-local means

Zhao Zhongmin, Zhao Yongjun, Niu Chaoyang 1038

Remote sensing image fusion with wavelet transform and sparse representation

Liu Ting, Cheng Jian 1045

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2013)08-0944-09

论文引用格式: 姜伟, 陆瑶, 杨炳儒. 格拉斯曼流形上的半监督判别分析[J]. 中国图象图形学报, 2013, 18(8): 944-952. [DOI: 10.11843/jig.20130808]

格拉斯曼流形上的半监督判别分析

姜伟¹, 陆瑶¹, 杨炳儒²

1. 辽宁师范大学数学学院, 大连 116029; 2. 北京科技大学计算机与通信工程学院, 北京 100083

摘要: 将图像集合表示为格拉斯曼流形上的点能够获得更好的识别性能。传统格拉斯曼流形上的判别分析方法仅考虑了带标签样本的统计信息, 忽略了无标签样本。鉴于此, 基于流形正则化思想, 提出了一个新的格拉斯曼流形上的半监督判别分析方法(SDAGM), 将其应用于图像集合的识别问题。通过构建近邻图刻画格拉斯曼流形上的所有样本局部几何结构, 并使其作为正则化项添加到格拉斯曼流形上的判别分析目标函数中, 本文方法不但考虑标签信息, 而且利用了一致性假设。标准数据集上的实验结果表明了 SDAGM 的有效性。

关键词: 格拉斯曼流形; 图像集合; 判别分析; 半监督

Semi-supervised discriminant analysis on Grassmannian manifold

Jiang Wei¹, Lu Yao¹, Yang Bingru²

1. School of Mathematics, Liaoning Normal University, Dalian 116029, China;

2. School of Computer & Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China

Abstract: Recent research has shown that a better recognition performance can be attained through representing image sets as points on Grassmannian manifolds. However, the conventional discriminant analysis methods based on such manifolds take into account only the statistical information of labeled samples and suffer from ignoring unlabeled samples. To address this issue, a new method based on manifold regularization, called semi-supervised discriminant Analysis on Grassmannian Manifold (SDAGM), is presented and applied to the image sets recognition problem. In SDAGM, a nearest neighbor graph is constructed to capture the local geometrical structure of all samples on the Grassmannian manifold and incorporates them into the objective function of discriminant analysis on Grassmannian manifold as a regularization term. Not only does the proposed algorithm consider the label information, but it also uses a consistency assumption. The feasibility and effectiveness of SDAGM are verified on several standard data sets with promising results.

Key words: Grassmann Manifold; image sets; discriminant analysis; semi-supervised

0 引言

传统图像识别领域通常是以单幅静态图像作为研究目标, 在可控场景的情况下取得了良好的性能^[1-2]。然而, 基于单幅静态图像学习算法在很多实际应用场景中效果不佳^[3-4]。随着数字成像和海量存储设备等硬件设备快速发展以及互联网资源的触手可及, 人们能够获取的图像涵盖了物体的各种表现变化, 如姿态、表情、光照角度等, 从而为基于图

收稿日期: 2012-10-19; 修回日期: 2012-12-25

基金项目: 国家自然科学基金项目(61175048)

第一作者简介: 姜伟(1970—), 男, 副教授, 2011年于北京科技大学获计算机应用技术博士学位, 主要研究方向为机器学习, 模式识别。

E-mail: swxxjw@yahoo.com.cn

像集合^[4,6]学习问题提供了丰富的数据。在这类问题中,每个图像集合通常包含同一类别的多个图像样本,利用了图像样本综合信息提取特征,提升算法的性能,因此,基于图像集合的识别问题受到越来越广泛的关注。

研究表明,将图像集合建模为子空间^[4,7],可以达到更好的性能。Yamaguchi 等人^[8]提出了针对人脸识别问题的互子空间方法(MSM),开创了利用子空间对图像集合非参数建模方法研究的先河,引起各领域科学工作者的广泛关注,衍生出一系列改进算法。Fukui 等人^[9]提出约束互子空间方法(CMSM),通过把图像集合投影到一个对分类最有利的子空间上,然后计算投影后的图像集合之间的相似性,算法的不足之处在于需人为设置子空间的维数及其相应的参数,在无监督情况下进行参数的选择还没有可靠的理论依据。Kim^[10]对 Fukui^[9]的工作做了进一步的发展,提出了基于典型相关的图像集合判别(DCC)学习和识别算法,采用主角作为子空间相似性度量,并期望寻求投影方向,使得数据映射到低维空间后,类间样本尽可能“分开”。Chu 等人^[11]提出基于图像集的核判别变换(KDT),该算法是 DCC^[10]算法的非线性拓展。Li 等人^[12]把集成策略用于图像集合的分类,提出了提升的全局和局部主角方法(BGLPA)。Wang 等人^[13]提出了基于流形-流形距离框架的图像集合人脸识别方法,将流形表示为一组局部线性模型的组合,采用子空间刻画各局部模型。

监督学习和无监督学习是降维算法的两种重要形式。通常,对无标签数据进行大量的标定工作会耗费很大的代价。如果只使用少量的标记样本,那么所训练出的学习系统很难具有良好的泛化能力,如果只使用未标记样本则浪费了标记样本中所提供的有用信息。因此,无论是监督算法还是无督算法都不能获得令人满意的结果。半监督学习是介于监督学习和无监督学习之间,它利用少量带标签样本来指导降维,同时又利用大量的不带标签样本来刻画并保持样本的局部几何结构。

处理子空间问题的一种变通的方法是将子空间表示为格拉斯曼流形(Grassmann manifold)^[14-15]上的点。近年来,一些研究者探讨了在格拉斯曼流形上实现判别分析。这类算法的基本出发点是把格拉斯曼流形上松散分布的样本,通过非线性变换压缩到一个低维的格拉斯曼流形中,使样本在该低维格

拉斯曼流形中分布更紧凑,更利于分类,降低计算复杂度。Hamm 等人^[15]提出了格拉斯曼流形判别分析(GDA),该方法将子空间表示成格拉斯曼流形上的点,利用格拉斯曼核度量空间的相似度,实现格拉斯曼流形上数据非线性判别分析。尽管 GDA 表现出优越的性能,然而该方法并没有明确考虑数据分布的内蕴几何结构^[16-18],从而影响样本的判别性和模型的泛化能力。鉴于此,Harandi 等人^[19]在图嵌入框架的基础上,提出了一种基于格拉斯曼流形局部判别分析方法(GEDAGM),采用在格拉斯曼流形上构建类内图和类间图刻画类内紧致性与类间可分性,以此获取流形上数据的几何结构和局部信息。在全局方法无法奏效的情况下,局部学习方法提供了有效的解决方案。本文提出一种新颖的格拉斯曼流形上的半监督判别分析方法(SDAGM),通过在格拉斯曼流形上构建所有样本(有标签和无标签)的近邻图来估计样本分布的几何结构,然后将其作为正则化项整合到 GDA 目标函数中,在此正则化项的作用就是用来估计样本空间的内蕴几何结构,从而学习更为有效的投影矩阵。本文方法具有以下优点:1)以图像集合为输入单位,将子空间表示为格拉斯曼流形上的点,由此,图像集合的判别分析问题转化为格拉斯曼流形上数据的判别分析问题;2)引入投影核度量格拉斯曼流形之间距离,提升算法时间效率;3)在图嵌入框架基础上,通过引入“局部保持”正则化项,将格拉斯曼流形上的判别分析与正则化项整合于单个目标函数。

1 格拉斯曼流形及其上判别分析

1.1 格拉斯曼流形

格拉斯曼流形已成功应用于控制领域和信号处理,近年来才逐步用于机器学习和模式识别领域。以下为格拉斯曼流形的定义^[20]。

定义 假设样本均匀采样于 D 维欧氏空间中的 m 维流形上,格拉斯曼流形 $G_{D,m}$ 是 $D \times m$ 正交矩阵等价类集合,即 $G_{D,m} = \{Y | Y = XV, \forall V \in O_k\}$,式中, X, Y 是 $D \times m$ 正交矩阵, V 是 $m \times m$ 正交矩阵, O_k 为正交群的一个特例,即有 $O_k = \{Q | Q^T Q = I, \det(Q) = 1\}$ 。

依据定义,如果格拉斯曼流形上的两个点能够通过正交变换相互获得时,那么这两个点是等价的。即矩阵 $W_1 = W_2 Q$,式中, Q 是 $r \times r$ 正交矩阵,并且

列空间 $\text{span}(W_1) = \text{span}(W_2)$, 从而 W_1 和 W_2 等价。由此可见, 格拉斯曼流形 $G_{D,m}$ 直观解释为格拉斯曼流形是空间 $\mathbf{R}^D$ 中所有 m 维线性子空间的集合, $G_{D,m}$ 上的每个点是一个 $D \times m$ 维正交矩阵列空间所张成的子空间^[19]。

如果图像集合用线性子空间描述, 则它们之间的距离可以用两个子空间的主角来表示。事实上, 格拉斯曼流形上两点之间的距离可以由格拉斯曼核来表示。采用投影核^[19]来表示格拉斯曼流形上两点之间的距离。

1.2 格拉斯曼流形上的判别分析

格拉斯曼流形上的判别分析^[15]基本思想是通过非线性映射将格拉斯曼流形 M 上数据点映射到一个新的格拉斯曼流形 M' 上, 使各类内样本尽可能“紧凑”, 类间样本尽可能“分开”。所以使得最大化判别准则为

$$J = \arg \max \frac{S_B^{M'}}{S_W^{M'}} \quad (1)$$

假设样本集 X 来自格拉斯曼流形 M 上, M 是 $D \times m$ 正交矩阵等价类集合。设带标签的样本集为 $X^L = \{X_i\}_{i=1}^l \subset M$, 样本分别属 c 个类别; 不带标签样本集为 $X^U = \{X_i\}_{i=l+1}^{l+u} \subset M$, 式中, $l+u=n$ 。样本全体为 $X = X^L \cup X^U = \{X_i\}_{i=1}^n$, $X_i \in \mathbf{R}^{D \times m}$, $i=1, 2, \dots, n$ 。 $X_j^{(i)}$ 表示第 i 个类中第 j 个样本。 n_i 表示第 i 类中元素的个数。存在非线性映射 $\Phi: M \rightarrow M'$, $X_i \rightarrow Y_i$ 把样本映射到新的格拉斯曼流形 M' 上。定义流形 M' 上的类间散度矩阵 $S_B^{M'}$ 和类内散度 $S_W^{M'}$ 矩阵为

$$S_B^{M'} = \sum_{i=1}^c n_i (m_i^{M'} - m^{M'}) (m_i^{M'} - m^{M'})^T$$

$$S_W^{M'} = \sum_{i=1}^c \sum_{j=1}^{n_i} (Y_j^{(i)} - m_i^{M'}) (X_j^{(i)} - m_i^{M'})^T$$

$$m_i^{M'} = \frac{1}{n_i} \sum_{j=1}^{n_i} Y_j^{(i)}$$

$$m^{M'} = \frac{1}{c} \sum_{i=1}^c n_i m_i^{M'} = \frac{1}{l} \sum_{i=1}^c \sum_{j=1}^{n_i} Y_j^{(i)}$$

与核学习方法相类似, 假设流形 M 所嵌入的高维特征空间的特征向量为 α_i, α_i 一定位于所有的训练样本 X 所张成的子空间中, 即存在系数 a_{ij} ($j=1, \dots, l$) 使 $\alpha_i = \sum_{j=1}^l a_{ij} X_j$, 有

$$Y_i = [\langle \alpha_1, X_i \rangle, \langle \alpha_2, X_i \rangle, \dots, \langle \alpha_d, X_i \rangle]^T \quad (2)$$

式中, d 为 G' 上子空间的维数, 即 G' 为 $\mathbf{R}^D$ 上 d 维子

空间所在的格拉斯曼流形, 假设 d 个特征向量构成投影矩阵 $A = [\alpha_1, \dots, \alpha_d]$ 。由

$$\begin{aligned} \langle \alpha_s, X_i \rangle &= \left\langle \sum_{j=1}^l a_{sj} X_j, X_i \right\rangle = \\ &= \sum_{j=1}^l a_{sj} \langle X_j, X_i \rangle = \sum_{j=1}^l a_{sj} k_{ij}^L = A_s^T K_i^L \end{aligned} \quad (3)$$

式中, $A_s = [a_{s1}, a_{s2}, \dots, a_{sl}]^T$, $K_i^L = [k_{i1}^L, k_{i2}^L, \dots, k_{il}^L]^T$ 。把式(3)代入式(2)得

$$Y_i = [A_1^T K_i^L, A_2^T K_i^L, \dots, A_d^T K_i^L]^T = A^T K_i^L \quad (4)$$

式中, $A = [A_1 | A_2 | \dots | A_d]$ 。

将式(4)代入式(1), 经过简单数学推导, 得

$$J = \arg \max \frac{A^T K^L W (K^L)^T A}{A^T K^L (I - W) (K^L)^T A}$$

式中, 核矩阵 $K^L = [K_1^L | K_2^L | \dots | K_l^L]$, W 是分块对角矩阵。第 i 块的矩阵 $W^{(i)} = \mathbf{1}_{n_i} \mathbf{1}_{n_i}^T / n_i$ ($\mathbf{1}_{n_i}$ 表示 1 的个数是 n_i 个全 1 向量)。记所有带标签数据的样本矩阵为 $X^L = [X^{(1)}, X^{(2)}, \dots, X^{(c)}]$, 定义 $L \times L$ 矩阵

$$W = \begin{bmatrix} W^{(1)} & 0 & \dots & 0 \\ 0 & W^{(2)} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & W^{(c)} \end{bmatrix} \quad (5)$$

投影矩阵可通过广义特征值问题 $K^L W (K^L)^T A = \lambda K^L (I - W) (K^L)^T A$ 求得。

2 格拉斯曼流形上的半监督判别分析

格拉斯曼判别分析是一种监督判别分析方法, 其目的是降维之后不同类样本之间的可分性最大, 然而该方法并没有明确考虑对分类问题尤为重要的内蕴几何结构。SDAGM 包括一个正则化项, 能够更好地保持和优化数据点之间的几何结构和判别信息。

2.1 目标函数

基于图的半监督降维算法的关键之处是一致性先验假设。也就是说近邻图像集合在投影低维流形时有相似的嵌入。具体而言, 已知 n 个数据点 $X = \{X_1, X_2, \dots, X_n\}$ 分布在格拉斯曼流形上, 如果数据点 X_i 属于 X_j 的 k -近邻, 或者 X_j 属于 X_i 的 k -近邻, 则称 X_i 和 X_j 为近邻点, 并在图上以边相连, 并附上相应的权重 S_{ij} , 其定义如下:

$$S_{ij} = \begin{cases} \exp(-\|X_i - X_j\|^2 / 2\sigma^2) & X_i \in N_k(X_j) \text{ 且 } X_j \in N_k(X_i) \\ 0 & \text{其他} \end{cases} \quad (6)$$

式中, $N_k(\mathbf{X}_i)$ 和 $N_k(\mathbf{X}_j)$ 分别表示 $\mathbf{X}_i$ 和 $\mathbf{X}_j$ 的 k -最近邻, σ 和 k 为自由参数。依据流形假设, 如果数据点之间有边连接, 则被投影到低维流形时, 数据点依然保持近邻。定义局部保持正则化项

$$\begin{aligned} R &= \frac{1}{2} \sum_{i,j} (\mathbf{Y}_i - \mathbf{Y}_j)^2 S_{ij} = \\ &= \frac{1}{2} \sum_{i,j} (\mathbf{A}^T \mathbf{K}_i - \mathbf{A}^T \mathbf{K}_j)^2 S_{ij} = \\ &= \sum_i \mathbf{A}^T \mathbf{K}_i \mathbf{K}_i^T \mathbf{A} S_{ij} - \sum_{ij} \mathbf{A}^T \mathbf{K}_i \mathbf{K}_j^T \mathbf{A} S_{ij} = \\ &= \mathbf{A}^T \mathbf{K} \mathbf{D} \mathbf{K}^T \mathbf{A} - \mathbf{A}^T \mathbf{K} \mathbf{S} \mathbf{K}^T \mathbf{A} = \\ &= \mathbf{A}^T \mathbf{K} \mathbf{L} \mathbf{K}^T \mathbf{A} \end{aligned} \quad (7)$$

式中, $\mathbf{A} = [\mathbf{A}_1 | \mathbf{A}_2 | \cdots | \mathbf{A}_d]$ 为投影矩阵; $\mathbf{K} = [\mathbf{K}_1 | \mathbf{K}_2 | \cdots | \mathbf{K}_n]$ 为核矩阵; 矩阵 $\mathbf{D}$ 为一对角矩阵, 且 $D_{ii} = \sum_j S_{ij}$, $\mathbf{S}$ 为权矩阵, $\mathbf{L} = \mathbf{D} - \mathbf{S}$ 为图拉普拉斯矩阵。最小化正则化项式(7)意味着, 如果图中有权相连的两点 $\mathbf{X}_i$ 和 $\mathbf{X}_j$ 被映射后反而远离时, $\mathbf{S}$ 的损失函数会产生很大的补偿。因此, 最小化正则化项就是试图确保如果 $\mathbf{X}_i$ 和 $\mathbf{X}_j$ 近邻, 那么 $\mathbf{Y}_i$ 和 $\mathbf{Y}_j$ 同样近邻^[2]。

在格拉斯曼流形上的判别分析的基础上, 考虑当数据取样于嵌入在高维空间低维流形数据的几何结构, 最大化目标函数

$$J = \arg \max \frac{\mathbf{A}^T \mathbf{K}^L \mathbf{W} (\mathbf{K}^L)^T \mathbf{A}}{\mathbf{A}^T \mathbf{K}^L (\mathbf{I} - \mathbf{W}) (\mathbf{K}^L)^T \mathbf{A} + \beta \mathbf{A}^T \mathbf{K} \mathbf{L} \mathbf{K}^T \mathbf{A}} \quad (8)$$

式中, β 为正则化参数, 用于折中目标函数中两项的重要程度; 易知, 若 $\beta = 0$, 则 SDAGM 就退化为 GDA。

最大化准则函数可以通过求解广义特征值问题, 即

$$\mathbf{K}^L \mathbf{W} (\mathbf{K}^L)^T \mathbf{A} = \lambda (\mathbf{K}^L (\mathbf{I} - \mathbf{W}) (\mathbf{K}^L)^T + \beta \mathbf{K} \mathbf{L} \mathbf{K}^T) \mathbf{A} \quad (9)$$

如果 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_d$ 为式(9)广义特征方程的前 d 个最大特征值所对应的特征向量, 定义投影矩阵 $\mathbf{A} = [\mathbf{A}_1 | \mathbf{A}_2 | \cdots | \mathbf{A}_d]$ 。对于一个新的图像集 $\mathbf{X}_q$ 在 α_k 上的投影为

$$\begin{aligned} V_q^k &= \langle \alpha_k, \mathbf{X}_q \rangle = \left\langle \sum_{j=1}^n a_{kj} \mathbf{X}_j, \mathbf{X}_q \right\rangle = \\ &= \sum_{j=1}^n a_{kj} \langle \mathbf{X}_j, \mathbf{X}_q \rangle = \mathbf{A}_k^T \mathbf{K}_q \end{aligned} \quad (10)$$

式中, $\mathbf{K}_q = [\langle \mathbf{X}_1, \mathbf{X}_q \rangle, \cdots, \langle \mathbf{X}_n, \mathbf{X}_q \rangle]^T$ 。

获取降维后样本的特征向量之后, 采取最近邻

分类器进行分类。设有 n 训练图像集合 $\{\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_n\}$, 这些图像样本集是由 N 个单图像组成的, 共有 c 类, 经过 SDAGM 算法提取的特征向量为 $\{V_1, V_2, \cdots, V_n\}$, 每一个图像集合都属于一个类别 $\omega_k, k=1, \cdots, c$ 。测试图像集合 $\mathbf{X}_{\text{test}}$ 的经变换后的特征向量为 $\mathbf{V}_{\text{test}}$, 如果满足

$$d(\mathbf{V}_{\text{test}}, V_i) = \min_i d(\mathbf{V}_{\text{test}}, V_i), V_i \in \omega_k$$

$$i=1, \cdots, N; k=1, \cdots, c$$

则测试图像集合 $\mathbf{X}_{\text{test}} \in \omega_k$ 。

2.2 算法描述

为了不失一般性, 假设带标签样本集 $\mathbf{X}^L = \{\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_l\}$ 中的样本点是按照它们的标签顺序排列的。SDAGM 算法描述如下:

数据 $\mathbf{X} = \{\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_l, \mathbf{X}_{l+1}, \cdots, \mathbf{X}_n\} \subset M$, 式中, l 表示带标签的样本的个数, n 表示 $\mathbf{X}$ 中全部样本的个数。带标签的样本 $\mathbf{X}_i |_{i=1}^l \in \mathbf{R}^{D \times m}$, 无标签样本 $\mathbf{X}_i |_{i=l+1}^n \in \mathbf{R}^{D \times m}$, 降维后的维数为 d 。

1) 在 $\mathbf{X}^L$ 上, 按照式(5)构造 $\mathbf{W}_{l \times l}$ 。定义矩阵 $\mathbf{W} \in \mathbf{R}^{n \times n}$ 如下

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_{l \times l} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

对 $\mathbf{X}_i \in \mathbf{X}^L$ 和 $\mathbf{X}_j \in \mathbf{X}^L$, 计算核矩阵 $\mathbf{K}_{l \times l}$ 。定义核矩阵

$$\mathbf{K}^L = \begin{bmatrix} \mathbf{K}_{l \times l} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \mathbf{K}^L \in \mathbf{R}^{n \times n}$$

定义矩阵 $\tilde{\mathbf{I}}$ 为

$$\tilde{\mathbf{I}} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

式中, $\mathbf{I}$ 为 $l \times l$ 阶单位矩阵。

2) 在 $\mathbf{X}$ 上依据近邻数, 构建近邻图, 按照式(6)计算权值矩阵 $\mathbf{S}$, 并求出拉普拉斯矩阵 $\mathbf{L} = \mathbf{D} - \mathbf{S}$ 。对所有 $\mathbf{X}_i \in \mathbf{X}$ 和 $\mathbf{X}_j \in \mathbf{X}$, 计算核矩阵 $\mathbf{K}$ 。

3) 求解广义特征值问题

$$\mathbf{K}^L \mathbf{W} (\mathbf{K}^L)^T \mathbf{A} = \lambda (\mathbf{K}^L (\tilde{\mathbf{I}} - \mathbf{W}) (\mathbf{K}^L)^T + \beta \mathbf{K} \mathbf{L} \mathbf{K}^T) \mathbf{A}$$

得到最优投影矩阵 $\mathbf{A} = [\mathbf{A}_1 | \mathbf{A}_2 | \cdots | \mathbf{A}_d]$, 式中, $\mathbf{A}_i$ 为前 d 个最大特征值对应的特征向量。

3 实验与结果

为了全面验证 SDAGM 有效性, 针对本文算法与其他相关算法的关系, 做了一系列对比实验。对

比的算法包括边界费舍尔分析(MFA)^[2]、线性判别分析(LDA)^[21]、核费舍尔判别分析(KFDA)^[22]、GEDAGM^[19]、黎曼局部保持映射^[5](RLPP)、GDA^[15]。其中,MFA、LDA和KFDA是经典的基于单图像降维方法,而GEDAGM、RLPP和GDA是近年提出的基于图像集合的方法。为了保证对比实验的充分性与广泛性,选用了4个具有不同特点的数据集,包括3个基准人脸数据集CMU PIE、Extended YaleB和Honda/UCSD与一个物体识别ETH-80数据集。

3.1 描述

CMU PIE数据集包括68个人的41 368幅人脸图像。这些图像具有变化的13种姿态、43种不同光照和4种面部表情,平均每人约为170张图像。在实验中,仅选取接近正面视角的人脸图像构成的子集,该子集包括不同光照和表情变化下的5个poses(C05、C07、C09、C27、C29)人脸图像,图1给出了某人的部分人脸图像。把图像调整为 32×32 大小,并将像素灰度值归一化处理(零均值、单位方差)。



图1 CMU PIE中某个人的部分人脸图像

Fig. 1 Some face images of one person from CMU PIE data set

Extended Yale B数据集包括38个人的2 414幅正面人脸图像,其中每个人大约64张图像。这些图像均在实验室控制光照的条件下拍摄。把图像调整为 32×32 像素大小,并将像素灰度值做归一化处理。图2显示了来自某人的部分人脸图像。



图2 Extended Yale B中某个人的部分人脸图像

Fig. 2 Some face images of one person from Extended Yale B data set

实验选取Honda/UCSD数据集的第一个子集,其中包括20个对象的59段视频,每人有2~5段视频,每段视频大约含有300~500帧图像,图像大小为 486×640 ,人脸发生2维或3维的姿态变化、表情变化、光照变化和尺度变化。把图像调整为 32×32 像素大小,并将像素灰度值归一化到单位区间。部分实验人脸图像如图3所示。



图3 Honda/UCSD中部分人脸图像

Fig. 3 Some face images from Honda/UCSD data set

ETH-80数据集包括8类不同对象,每类物体包括10组不同的子类,如图4所示,每个子类包含该子类不同视角下的41幅图像。实验中把图像调整为 32×32 像素大小,并将像素灰度值作归一化处理。

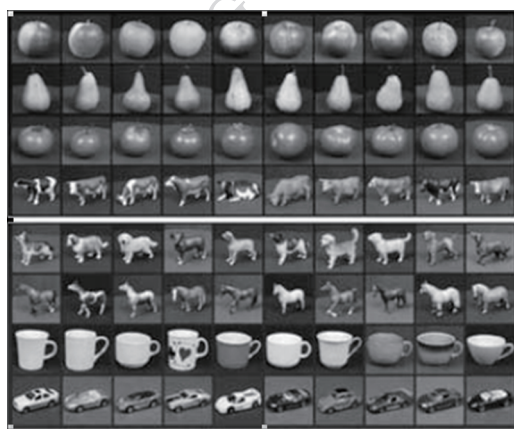


图4 ETH-80的部分物体

Fig. 4 Some objects from ETH-80 data set

3.2 实验环境设置

3.2.1 单图像方法

随机地从每类中选择一半图像作为训练集,剩余部分用于测试集。首先基于训练样本,并分别利用MFA、LDA、KFDA算法获取从高维空间到低维特征空间显式线性映射,然后将训练和测试样本通过线性映射投影至各自的低维特征空间,最终基于测试样本,采用最近邻分类器计算识别率。实验重复执行20遍,平均识别率作为最终性能的评价指标。为了避免小样本问题,在执行算法之前,采用了主成分分析(PCA)算法对数据进行预处理,在降维过程中保留95%的数据能量。KFDA中的核函数选择通用的高斯核,根据经验将 σ 固定为训练样本的平均范数 σ_0 ,近邻参数 k 设置搜索范围为 $\{1, 2, \dots, l-1\}$, l 为每类中训练样本个数。

3.2.2 图像集合方法

对于CMU PIE,取20个人做实验,每次采用无放回随机地选择每个人的10幅人脸图像组成一个图像集合,这样每个人可构成17个图像集合,其中10个图像集合作为训练集,7个图像集合作为测试

集,训练集中 4 个图像集合无标签。对于 Extended Yale B,取 20 人做实验,同样采用无放回随机地选择每个人的 6 幅人脸图像组成一个图像集合,这样每个人可构成 10 个图像集,其中 6 个图像集作为训练集,4 个图像集作为测试集,训练集中 2 个图像集合无标签。对于 Honda/UCSD,选择一个视频段中 100 幅图像构成训练集,每个图像集合是由 10 幅图像构成,这样每个人有 10 个训练图像集合,其中 2 个是无标签的图像集合;选择另外一个视频 100 幅图像构成测试集,同样每个测试图像集合是由 10 幅图像构成的。对于 ETH-80 数据集,采用每

个子类不同视角下的 41 幅图像组成一个图像集,这样每个物体可构成 10 个图像集合,其中 3 个图像集合作为训练集,7 个图像集合作为测试集,在训练集中 1 个图像集未带标签。近邻参数 k 设置搜索范围为 $\{1, 2, \dots, l-1\}$, l 为每类中训练样本个数;核函数均使用投影核。实验重复执行 20 遍,平均识别率作为最终性能的评价指标。

3.3 识别率

通过最近邻分类器来评价和比较不同降维算法的性能。表 1 给出了各种方法对应的最好结果及其对应的维数。

表 1 不同算法在 4 种数据集上的识别率

Table 1 Face recognition accuracies of different algorithms on four data sets

算法	识别率(维数)			
	CMU PIE	Extended Yale B	Honda/UCSD	ETH-80
MFA	62.5% (60)	64.7% (90)	61.2% (40)	65.8% (14)
LDA	59.9% (50)	62.5% (80)	70.8% (50)	63.4% (12)
KFDA	68.3% (50)	68.9% (50)	65.3% (120)	62.2% (18)
GEDAGM	88.6% (120)	86.7% (40)	80.3% (50)	67.5% (12)
RLPP	72.1% (150)	70.2% (110)	73.6% (130)	91.8% (18)
GDA	90.8% (70)	91.7% (40)	82.7% (50)	68.2% (18)
SDAGM	100% (20)	94.3% (90)	87.2% (20)	79.3% (12)

针对图像集合方法,图 5 给出了各种降维算法的平均识别率对特征维数的变化曲线。

3.4 参数的敏感性

针对不同的数据集,配置恰当的参数,能够提升算法的识别性能。然而,为参数选择合适的值是不容易的。通常采用交叉验证的方式进行参数选择,但非常耗时,并需要大量的训练样本,在实际中难以实现。参数选择对于带标号数据很少的机器学习算法而言,目前还没有非常有效的方法。图 6 显示了对应的不同近邻参数 k 所构建的近邻图,SDAGM 在 4 个数据集上的识别率对比。具体地,固定低维子空间维数 $d = 20$ 和正则化参数 $\beta = 20$,遍历不同近邻数 k 对应的识别率,参数 β 用于折中类内方差和样本的正则化项。如果 β 选取较小,则意味着正则化项在目标函数中所占的比重少,这样,对样本内蕴几何结构的影响也较小,极端情形是 β 取 0 时,算法退化为 GDA 方法;相反,如果参数 β 选取较大,则正则化项在目标函数中所占的比重较大,突出了数据

的几蕴几何结构作用,弱化了样本的类内方差。为了提高算法的时间效率, β 大多采用经验选取。图 7 显示固定低维子空间维数取 $d = 20$ 和近邻数 $k = 5$,不同正则化参数 β ,SDAGM 在 4 个数据集上的识别率对比。从图 7 中不难看出,SDAGM 对正则化参数不敏感,很大程度上得益于在格拉斯曼流形上所构建的正则化项,有效地获得了样本的类内几何结构。最小化正则化项,在某种程度上迎合了使得样本映射到低维空间后,各类样本尽可能“紧凑”。

3.5 实验结果讨论

总体而言,SDAGM 在绝大多数情况下显示出最优的性能。根据上面的图表可以得到以下结论:

1) 通过对比相同数据集下的单样本方法和集合方法识别率,集合方法均一致地超过单样本方法。这表明由集合方法获得的投影方向较单样本方法能够充分利用图像集合提供的整体信息,这样,更有效地保留样本的判别信息,规避单样本算法的弊端。

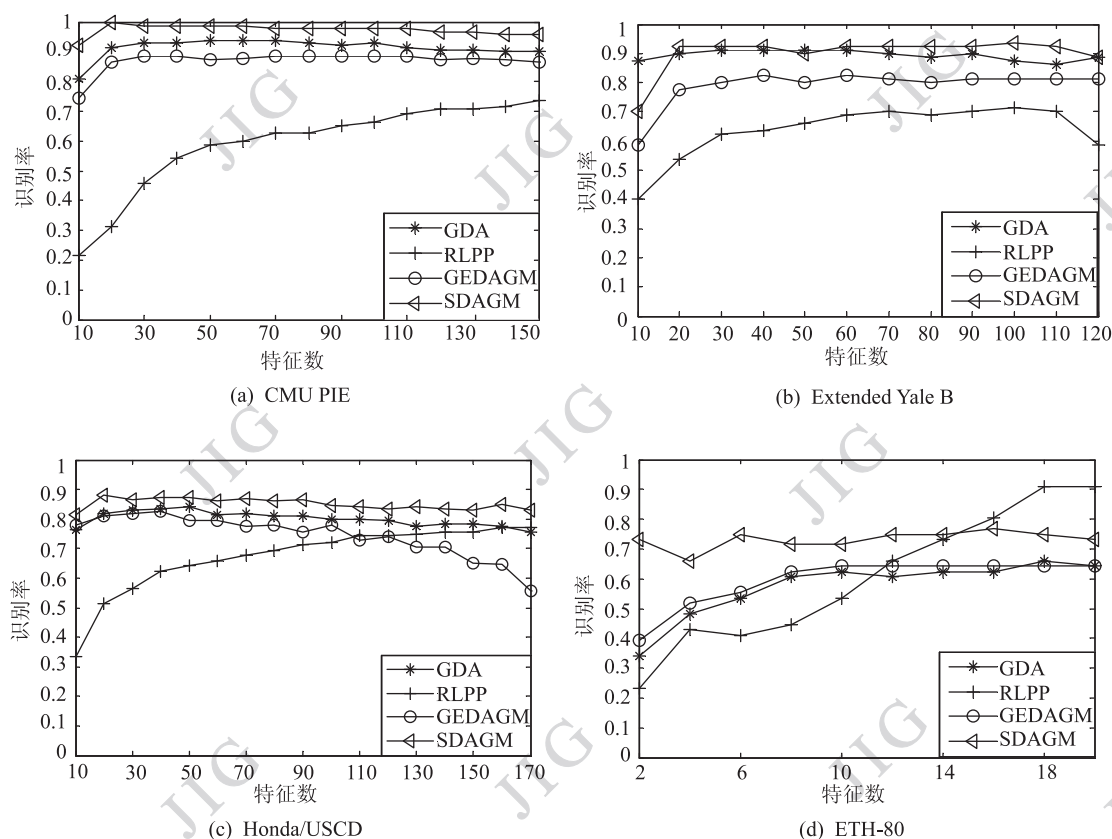
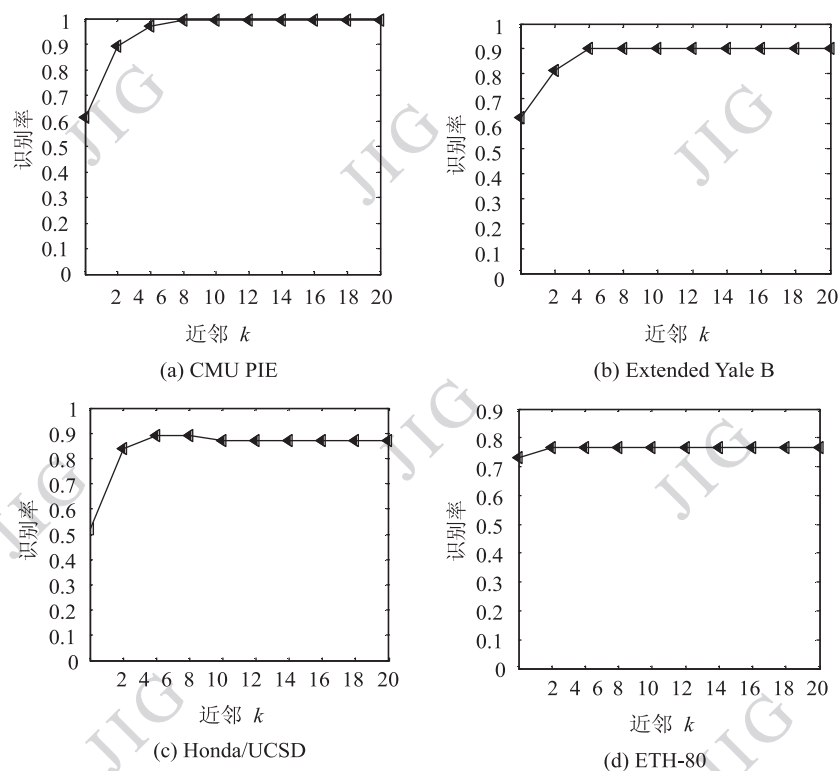
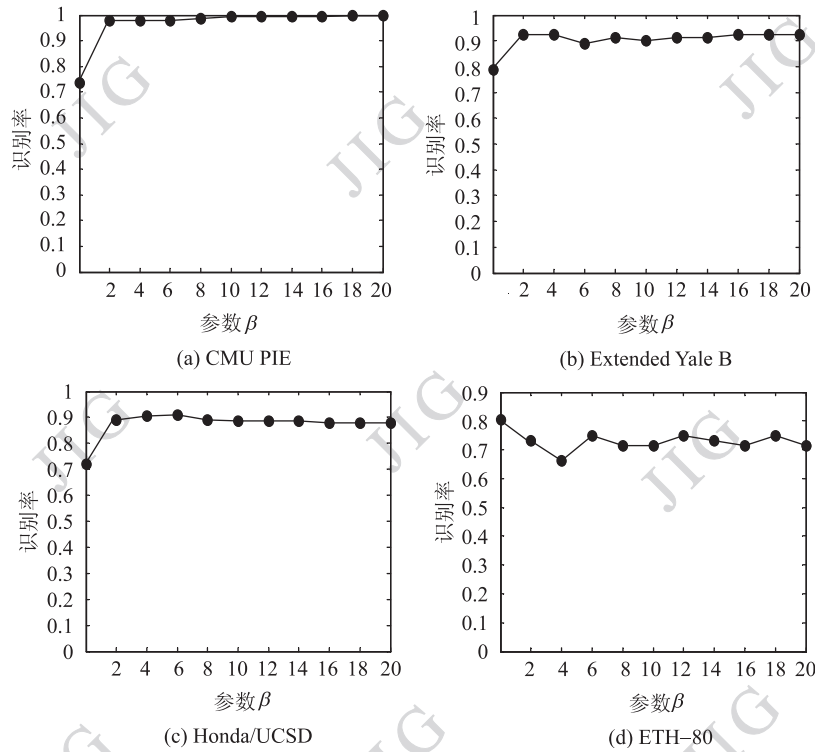


图5 不同算法在4个数据集中识别率曲线图

Fig. 5 Recognition rate curves of different algorithms versus dimension of reduced space on four databases

图6 不同邻域大小 k 对应识别率Fig. 6 The classification accuracies using different neighborhood size k

图 7 不同的正则化参数 β 对应的识别率Fig. 7 The classification accuracies using different regularization parameter β

2) 针对集合算法, 只要低维子空间的维数超过某个阈值, SDAGM 大多能够超过 GEDAGM、RLPP 和 GDA 3 种方法, 这说明具有全局判别结构和局部保持正则化方法在一定程度上提升后续分类器的识别率。

3) 对于 4 个数据集综合比较, ETH-80 的效果是最差的, 原因可能是算法中参数选择不当造成的。而 Honda/UCSD 的识别率也较其他数据集低, 原因可能是视频图像背景复杂、姿态变化较大, 加大识别的难度。

4) 值得指出的是, RLPP 在 Honda/UCSD 上给出了最高的识别率, 而 SDAGM 识别率则略低, 但 SDAGM 的识别率更稳定。

5) SDAGM 对近邻数和正则化参数不敏感。

4 结 论

提出一种新颖的格拉斯曼流形上的半监督判别分析方法, 在流形正则化的半监督学习框架下, 利用在格拉斯曼流形上所有样本 (带标签和不带标签) 估计样本子空间的流形结构来帮助提取有效的判别特征。在大多数数据集上的实验结果表明, SDAGM

表现最好, 主要是因为 SDAGM 同时考虑了数据的局部几何结构、判别信息, 并且全面继承了 GDA 方法的优点, 从而得到最优的识别性能。下一步的研究方向是如何有效地将图像集合方法应用到更为普遍的黎曼流形上。

参考文献 (References)

- [1] Belhumeur P N, Hespanha J, Kriegman D J. Eigenfaces vs. fisherfaces: recognition using class specific linear projection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(7): 711-720. [DOI:10.1109/34.598228]
- [2] Yan S C, Xu D, Zhang B Y, et al. Graph embedding and extensions: a general framework for dimensionality reduction[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(1): 40-51. [DOI:10.1109/TPAMI. 2007. 12]
- [3] Kim M, Kumar S, Pavlovic V, et al. Face tracking and recognition with visual constraints in real-world videos[C] //Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Anchorage, Alaska, USA: IEEE Computer Society, 2008:1-8. [DOI:10.1109/CVPR. 2008. 4587572]
- [4] Arandjelovic O., Shakhnarovich G, Fisher J. Face recognition with image sets using Manifold Density divergence[C] //Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. San Diego, CA, USA: IEEE Computer Society, 2005:

- 581-588. [DOI: 10.1109/CVPR.2005.151]
- [5] Harandi M T, Sanderson C, Wiliem A, et al. Kernel analysis over Riemannian Manifolds for visual recognition of actions, pedestrians and textures [C]//Proceedings of IEEE Workshop on Applications of Computer Vision. Breckenridge, CO, USA; IEEE Computer Society, 2012, 433-439. [DOI: 10.1109/WACV.2012.6163005]
- [6] Sanin A, Sanderson C, Harandi M, et al. K-tangent spaces on Riemannian manifolds for improved pedestrian detection [C]//Proceedings of IEEE International Conference on Image Processing. Orlando, Florida, USA; IEEE Computer Society, 2012:7-10.
- [7] Fan W, Yeung D Y. Locally linear models on face appearance manifolds with application to dual-subspace based classification [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. New York, USA; IEEE Computer Society, 2006:1384-1390. [DOI:10.1109/CVPR.2006.178]
- [8] Yamaguchi O, Fukui K, Maeda K. Face recognition using temporal image sequence [C]//Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition. Nara, Japan; IEEE Computer Society, 1998: 318-323. [DOI: 10.1109/AFGR.1998.670968]
- [9] Fukui K, Yamaguchi O. Face recognition using multiview point patterns for robot vision [C]//International Symposium on Robotics Research. Siena, Italy; Springer Tracts in Advanced Robotics, 2003:192-201. [DOI: 10.1007/11008941_21]
- [10] Kim T K, Kittler J, Cipolla R. Discriminative learning and recognition of image set classes using canonical correlations [J]. IEEE Transactions Pattern Analysis and Machine Intelligence, 2007, 29 (6): 1005-1018. [DOI: 10.1109/TPAMI.2007.1037]
- [11] Chu W S, Chen J C, Lien J J. Kernel discriminant transformation for image set-based face recognition [J]. Pattern Recognition, 2011, 44 (8): 1567-1580. [DOI: 10.1016/j.patcog.2011.02.011]
- [12] Li X, Fukui K, Zheng N N. Image-set based face recognition using boosted global and local principal angles [C]//Proceedings of Asian Conference on Computer Vision. Queenstown. New Zealand; the Asian Federation of Computer Vision Societies, 2010: 323-332. [DOI:10.1007/978-3-642-12307-8_3]
- [13] Wang R, Shan S, Chen X, et al. Manifold-manifold distance with application to face recognition based on image set [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Anchorage, Alaska, USA; IEEE Computer Society, 2008:1-8. [DOI:10.1109/CVPR.2008.4587719]
- [14] Shirazi S, Harandi M T, Sanderson C, et al. Clustering on Grassmann manifolds via kernel embedding with application to action analysis [C]//Proceedings of IEEE International Conference on Image Processing. Orlando, Florida, USA; IEEE Signal Processing Society, 2012.
- [15] Hamm J, Lee D D. Grassmann discriminant analysis: a unifying view on subspace-based learning [C]//Proceedings of International Conference on Machine Learning. Helsinki, Finland; ACM, 2008:376-383. [DOI: 10.1145/1390156.1390204]
- [16] Chen J, Ye J, Li Q. Integrating global and local structures: a least squares framework for dimensionality reduction [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis, Minnesota, USA; IEEE Computer Society, 2007:1-8. [DOI: 10.1109/CVPR.2007.383040]
- [17] Yin X S, Hu E L. Semi-supervised locality dimensionality reduction [J]. Journal of Image and Graphics, 2011, 16 (9): 1615-1624. [尹学松, 胡恩良. 半监督局部维数约减 [J]. 中国图象图形学报, 2011, 16 (9): 1615-1624.] [DOI: 10.11834/jig.20110901]
- [18] Harandi M, Nili M, Ahmatabadi, et al. Optimal local basis: a reinforcement learning approach for face recognition [J]. International Journal of Computer Vision, 2009, 81 (2): 191-204. [DOI: 10.1007/s11263-008-0161-5]
- [19] Harandi M T, Sanderson C, Shirazi S, et al. Graph embedding discriminant analysis on Grassmannian manifolds for improved image set matching [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Colorado Springs, CO, USA; IEEE Computer Society, 2011: 2705-2712. [DOI: 10.1109/CVPR.2011.5995564]
- [20] Edelman A, Arias T A, Smith S T. The geometry of algorithms with orthogonality constraints [J]. SIAM Journal on Matrix Analysis and Applications, 1998, 20 (2): 303-353. [DOI: 10.1137/S0895479895290954]
- [21] Belhumeur P N, Hespanha J P, Kriegman D J. Eigenfaces vs. fisherfaces: recognition using class specific linear projection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19 (7): 711-720. [DOI: 10.1109/34.598228]
- [22] Mika S, Ratsch G, Weston J, et al. Mullers, fisher discriminant analysis with kernels [C]//Proceedings of IEEE Signal Processing Society Workshop on Neural Networks for Signal Processing IX. Wisconsin, USA; IEEE Signal Processing Society, 1999. [DOI: 10.1109/NNSP.1999.788121.]