

Journal of Image
and Graphics

中国图象图形学报



ISSN1006-8961
CN11-3758/TB

2013 **3**
Vol.18 No.

中国科学院遥感应用研究所
中国图象图形学学会主办
北京应用物理与计算数学研究所

中国图象图形学报

Zhongguo Tuxiang Tuxing Xuebao

2013 年 3 月 第 18 卷 第 3 期(总第 203 期)

目次

图像处理和编码

沿边局部灰度差分椒盐噪声滤波 孔勇奇, 潘志庚(249)

图像分析和识别

核稀疏保留投影及在步态识别中的应用 王科俊, 阎涛, 吕卓纹, 唐墨(257)

局部 Gist 特征匹配核的场景分类 杨昭, 高隼, 谢昭, 吴克伟(264)

图像理解和计算机视觉

条件随机场模型的场景描述 赵龙, 郭立, 谢锦生(271)

基于 L_1 与空间点分类的 3 维场景重建 李静, 杨宜民, 付根平, 陈建平(277)

稀疏表示的 Lucas-Kanade 目标跟踪 徐如意, 陈靛影(283)

医学图像处理

管状特性和主动轮廓的 3 维血管自动提取 尚岩峰, 汪辉, 汪宁, 陈杰, 姚丽萍, 孙琨(290)

“第 17 届 CAD&CG 学术会议”专栏

多步攻击的规则树检测及可视化 胡亮, 赵剑明, 解男男, 努尔布力(299)

显著性边缘引导下的图像抽象化 王勋, 李红, 刘春晓, 严彩萍(305)

稀疏表达的运动数据压缩 齐天, 肖俊, 庄越挺(311)

基于语义的用户关注度计算 陈蕾英, 陈佳舟, 潘斌, 彭群生(318)

九宫格空间框架的图书图像检索 李宗民, 唐志辉(325)

光学相干层析视网膜体数据的 3 维分割 樊鲁杰, 孙延奎, 张田, 田小林(330)

混合软影绘制算法 刘浏, 周炜, 李华(336)

基于场的人群运动仿真 赵欣欣, 张勇, 孔德慧, 尹宝才(344)

利用大位移视图的自动可信图像修补 刘春晓, 金剑秋, 彭群生(351)

第八届国际数字地球会议征稿通知 (358)

Journal of Image and Graphics

(Monthly, Started in 1996)

Vol. 18 No. 3 March 2013

Contents

Image Processing and Coding

Salt and pepper noise filtering algorithm based on local border gray-scale differences Kong Yongqi, Pan Zhigeng(249)

Image Analysis and Recognition

Kernel sparsity preserving projections and its application to gait recognition Wang Kejun, Yan Tao, Lv Zhuowen, Tang Mo(257)

Scene categorization of local Gist feature match kernel Yang Zhao, Gao Jun, Xie Zhao, Wu Kewei(264)

Image Understanding and Computer Vision

Scene description based on the conditional random fields model Zhao Long, Guo Li, Xie Jinsheng(271)

Three dimensional scene reconstruction based on L_1 and scene point classification
..... Li Jing, Yang Yimin, Fu Genping Chen Jianping(277)

Lucas-Kanade tracking based on sparse representation Xu Ruyi, Chen Jingying(283)

Medical Image Processing

Three dimensional vessel extraction model based on tubular characters and active contour model
..... Shang Yanfeng, Wang Hui, Wang Ning, Chen Jie, Yao Liping, Sun Kun(290)

Issum of the CAD&CG'2012

Multi-step attacks detected by rules tree and visualization Hu Liang, Zhao Jianming, Xie Nannan, Nurbol(299)

Salient edge guided image abstraction optimization Wang Xun, Li Hong, Liu Chunxiao, Yan Caiping(305)

Motion data compression using sparse representation Qi Tian, Xiao Jun, Zhuang Yueting(311)

Semantic-based user interests computation Chen Leiying, Chen Jiazhou, Pan Bin, Peng Qunsheng(318)

Books image retrieval based on grid space frame Li Zongmin, Tang Zhihui(325)

Three dimensional segmentation to detect retinal boundary surfaces from OCT volume data
..... Fan Lujie, Sun Yankui, Zhang Tian, Tian Xiaolin(330)

Hybrid soft shadow rendering method Liu Liu, Zhou Wei, Li Hua(336)

Field-based crowd simulation Zhao Xinxin, Zhang Yong, Kong Dehui, Yin Baocai(344)

Automatic and accurate image completion from a large displacement view Liu Chunxiao, Jin Jianqiu, Peng Qunsheng(351)

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2013)03-0257-07

论文引用格式: 王科俊, 阎涛, 吕卓纹, 唐墨. 核稀疏保留投影及在步态识别中的应用[J]. 中国图象图形学报, 2013, 18(3): 257-263.

核稀疏保留投影及在步态识别中的应用

王科俊, 阎涛, 吕卓纹, 唐墨

哈尔滨工程大学自动化学院, 哈尔滨 150001

摘要: 针对已有基于线性变换的稀疏保留投影方法在解决实际问题时, 会遇到维数灾难和小样本问题。通过引入核方法, 提出一种核稀疏保留投影方法。首先采用非线性变换将原始数据映射到高维特征空间, 而后在这个高维空间进行稀疏重构, 并对得到的系数矩阵进行降维优化, 最终得到所需的投影矩阵。将其应用到步态识别中, 采用 CASIA(B) 步态数据库进行实验分析, 实验结果表明, 本文方法取得了令人满意的识别效果。

关键词: 稀疏保留投影; 核方法; 步态识别; 步态能量图

Kernel sparsity preserving projections and its application to gait recognition

Wang Kejun, Yan Tao, Lv Zhuowen, Tang Mo

College of Automation, Harbin Engineering University, Harbin 150001, China

Abstract: In order to solve the problem of the curse of dimensionality and the small sample problem, a kernel sparsity preserving projection is proposed. First, the nonlinear transformation is used to map the original data to a high-dimensional feature space. Then, the sparsity reconstruction in a high-dimensional space is used and, the coefficient matrix is reduced and optimized. Finally, the projection matrix is obtained. This method is evaluated on the CASIA (B) Gait database. The experimental results show that the proposed method can obtain stable classification and performs satisfactory recognition results.

Key words: sparsity preserving projections; kernel method; gait recognition; gait energy image

0 引言

在高维数据的分析和学习中, 数据降维起着重要作用。数据降维不仅可以简化数据以利于存储和计算, 而且通过降维能够去掉冗余信息, 找到高维数据的低维本质表示以利于分析。近年来, 已有多种数据降维方法被成功应用于模式识别、图像处理等领域。根据是否引入监督信息, 降维方法可以分为有监督的、无监督的方法。

有监督降维方法就是引入各种监督信息来指导投影变换过程, 最常见的是引入类别信息的线性判别分析(LDA)^[1], Xing 等人^[2]提出的基于 Side 信息(由一个相似、不相似对构成的集合)的降维算法。

无监督降维方法是直接按照某种规则来实现原始空间到目标空间的变换, 常见的无监督降维方法有主成分分析(PCA)^[3]、独立成分分析(ICA)^[4], 以及一些非线性算法, 如局部线性嵌入(LLE)^[5]、拉普拉斯特征映射(LE)^[6]。

对于无监督降维方法, 由于没有引入原始数据

收稿日期: 2012-05-31; 修回日期: 2012-09-12

基金项目: 国家高技术研究发展计划(863)基金项目(2008AA01Z148); 国家自然科学基金项目(60975022); 博士点专项科研基金项目(20102304110004)

第一作者简介: 王科俊(1962—), 男, 教授, 博士生导师, 1995 年于哈尔滨工程大学获船舶与海洋特辅装置与系统专业博士学位, 主要研究方向为模式识别、神经网络等。E-mail: wangkejun@hrbeu.edu.cn

所蕴含的类别信息,使得最终的识别效果往往低于有监督降维方法。因此为了提高无监督降维方法的识别效果,在投影变换时可以引入原始数据所蕴含的内在信息。如局部保留投影(LPP)^[7]通过保留原始数据的局部近邻信息得到了较好的识别效果。

受稀疏表示(SR)的启发,Qiao 等人^[8]提出了稀疏保留投影方法(SPP)。在投影变换过程中,SPP 首先通过稀疏重构处理,而后再对重构的稀疏表示系数矩阵进行降维投影。该算法所得到的系数矩阵在具有最大稀疏度的同时也保留了原始数据的局部近邻信息,使得最终获得的特征更加稳定。

然而该方法在使用时存在严重的维数灾难和小样本问题,一般需要先采用 PCA 方法降维,而后才能使用。因此使得算法的识别效果很大程度上取决于 PCA 的降维效果。而核方法^[9]是解决该类问题的有效方法,而且所得的特征矩阵的维数只与样本数有关。在实际应用中,样本数一般远远小于样本的维数,因此大大降低了运算量。

文献[10]将核方法引入到 PCA 方法中,提出了核主成分分析(KPCA)。借鉴该思想,将核方法引入到稀疏保留投影中,提出核稀疏保留投影(KSPP)。首先利用核函数将原始数据映射到高维特征空间中,而后再进行稀疏重构和降维投影。为了验证该方法的有效性,将其应用到步态识别中。实验结果表明,该方法取得了令人满意的识别效果。

1 稀疏表示

在模式识别、图像处理等领域中,寻找原始高维图像的低维表示一直是一个非常重要的研究方向,能否对图像建立有效的、更具区分性的表示方式直接影响到后续的分析处理。图像的稀疏表示是一种新兴的图像模型,其原理是首先通过稀疏重构,将原始数据重构成尽可能稀疏的表示方式,而后再进行投影达到降维的目的。重构的表示系数中较少的非零分量揭示了图像信号的主要结构与本质属性,同时神经生理学研究表明在哺乳动物的本原视觉皮层同样存在稀疏编码策略,因此稀疏表示模型能够很好地匹配人类的视觉感知特性。目前稀疏表示已在图像理解、数据压缩等领域得到了广泛应用。

稀疏表示的核心是对图像进行稀疏重构,即

$$\mathbf{x} = \mathbf{D}\boldsymbol{\alpha} \quad (1)$$

式中, $\mathbf{x} \in \mathbf{R}^m$ 为所要研究的图像样本所拉成的列向

量, $\mathbf{D} \in \mathbf{R}^{m \times l}$ 为数据字典,是 l 个 m 维向量的集合,每个向量被称为原子, $\boldsymbol{\alpha} \in \mathbf{R}^l$ 为图像的稀疏表示系数,为了使 $\boldsymbol{\alpha}$ 在理论上有所解,要求 $l \geq m$ 。

如果字典 $\mathbf{D}$ 中原子能够张成 m 维欧氏空间 $\mathbf{R}^m$,则字典 $\mathbf{D}$ 是完备的。当 $l > m$ 时,如果同时保证能够张成 m 维欧氏空间 $\mathbf{R}^m$,则原子是线性相关的,字典 $\mathbf{D}$ 是冗余的,此时称字典 $\mathbf{D}$ 是过完备的。图像在过完备字典下的稀疏表示系数 $\boldsymbol{\alpha}$ 并不唯一,使得可以根据不同的应用目的选择最适合的表示系数。因此在完备和过完备字典下的稀疏表示能得到较好的效果,同样说明了 $l \geq m$ 的必要。

由数据压缩领域可知,重构的表示系数越稀疏,则之后的降维效果越明显。因此重构的目标是使重构系数向量中大部分分量为零,只有少数的非零大系数。为了从众多的表示系数中选取最为稀疏的系数,采用 L_0 范数作为稀疏性度量函数,则图像的稀疏表示的目标函数为

$$\begin{aligned} \min_{\boldsymbol{\alpha}} \quad & \|\boldsymbol{\alpha}\|_0 \\ \text{s. t.} \quad & \mathbf{x} = \mathbf{D}\boldsymbol{\alpha} \end{aligned} \quad (2)$$

式中, $\|\cdot\|_0$ 是 L_0 范数,然而,当信号中含有噪声时, L_0 范数并不十分有效,以及在实际的信号表示方法中,很少表示能使许多系数严格地等于 0,会使得求解过程成为 NP 难问题。常见的方法是采用其他范数代替 L_0 范数,如 L_1 范数,则

$$\begin{aligned} \min_{\boldsymbol{\alpha}} \quad & \|\boldsymbol{\alpha}\|_1 \\ \text{s. t.} \quad & \mathbf{x} = \mathbf{D}\boldsymbol{\alpha} \end{aligned} \quad (3)$$

2 稀疏保留投影

受稀疏表示的启发,Qiao 等人^[8]提出稀疏保留投影方法。

首先将所要研究的图像拉成列向量,则样本集合为 $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]$, $\mathbf{x}_i \in \mathbf{R}^m$, $i = 1, 2, \dots, n$ 。根据稀疏表示理论,SPP 首先重构每个样本 $\mathbf{x}_i$ 并获取其对应的稀疏重构系数向量 $\mathbf{s}_i$ 。 $\mathbf{s}_i$ 的求取可转化为最小化问题

$$\begin{aligned} \min_{\mathbf{s}_i} \quad & \|\mathbf{s}_i\|_1 \\ \text{s. t.} \quad & \mathbf{x}_i = \mathbf{X}\mathbf{s}_i, \mathbf{1} = \mathbf{1}_n^T \mathbf{s}_i \end{aligned} \quad (4)$$

式中, $\mathbf{1}$ 为标量 1, $\mathbf{1}_n$ 是元素全为 1 的 n 维列向量; $\mathbf{s}_i = [s_{i,1}, \dots, s_{i,i-1}, 0, s_{i,i+1}, \dots, s_{i,n}]$ 是 n 维且第 i 个元素为 0 的列向量。

由此可见,与传统的稀疏表示理论相比,SPP

算法中,在为每一个样本进行稀疏重构时,所采用的数据字典 $\mathbf{D}$ 是整个样本集,其中被重构样本本身的重构系数强制为 0,即该样本在对应的字典中不起作用。而得到的稀疏重构系数向量中较大的系数所对应的是与被重构样本相似的样本,因此可以说该方法保留了原始样本的局部近邻信息。

在实际应用中,样本一般都有噪声信号,于是 $\mathbf{x}_i = \mathbf{X}\mathbf{s}_i$ 很难满足,为此将上式进行扩展,这里有两种扩展方式:

方式 1

$$\begin{aligned} & \min_{\mathbf{s}_i} \|\mathbf{s}_i\|_1 \\ & \text{s. t. } \|\mathbf{x}_i - \mathbf{X}\mathbf{s}_i\| < \varepsilon, 1 = \mathbf{1}_n^T \mathbf{s}_i \end{aligned} \quad (5)$$

式中, ε 为误差容忍度,是一足够小的数。

方式 2

$$\begin{aligned} & \min_{\mathbf{s}_i, \mathbf{t}_i} \|\mathbf{s}_i^T \mathbf{t}_i^T\|_1 \\ & \text{s. t. } \begin{bmatrix} \mathbf{x}_i \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{X} & \mathbf{I}_m \\ \mathbf{1}_n^T & \mathbf{0}_m^T \end{bmatrix} \begin{bmatrix} \mathbf{s}_i \\ \mathbf{t}_i \end{bmatrix} \Leftrightarrow \begin{cases} \mathbf{x}_i = \mathbf{X}\mathbf{s}_i + \mathbf{t}_i \\ 1 = \mathbf{1}_n^T \mathbf{s}_i \end{cases} \end{aligned} \quad (6)$$

式中, $\mathbf{t}_i$ 是一个 m 维的列向量, $\mathbf{0}_m$ 是一个元素全为 0 的 m 维列向量, $\mathbf{I}_m$ 是一个 m 阶的单位矩阵。

可以看出,两种方式的不同之处在于误差容忍度的不同,方式 1 的误差容忍度是一个固定的值,而方式 2 将 $\mathbf{s}_i$ 和 $\mathbf{t}_i$ 同时优化,可以说 $\mathbf{t}_i$ 就是根据每个样本所自动调节得到的误差容忍度。

式(5)(6)中的 L_1 范数最小化问题可以通过转换成标准线性规划问题来求解^[11],因此可对两种方式计算并优化得到每一个样本的系数向量,得到稀疏重构系数矩阵 $\mathbf{S}$ 定义为

$$\mathbf{S} = (\mathbf{s}_{ij})_{n \times n} = [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n]^T \quad (7)$$

由于系数向量 $\mathbf{s}_i$ 蕴含原始样本的局部近邻信息以及自身的稀疏性约束,使系数矩阵 $\mathbf{S}$ 既反映了原始样本的内在属性,也保留了一些潜在的判决性信息。

而后对系数矩阵进行降维优化,而从原始的高维空间投影到低维子空间时,总是希望保留所需要的特征信息。为探寻最优的特征向量,定义目标函数

$$\begin{aligned} & \min_{\mathbf{w}} \sum_i^n \|\mathbf{w}^T \mathbf{x}_i - \mathbf{w}^T \mathbf{X}\mathbf{s}_i\|^2 = \\ & \min_{\mathbf{w}} \left(\mathbf{w}^T \left(\sum_i^n (\mathbf{x}_i - \mathbf{X}\mathbf{s}_i)(\mathbf{x}_i - \mathbf{X}\mathbf{s}_i)^T \right) \mathbf{w} \right) \end{aligned} \quad (8)$$

令 $\mathbf{u}_i$ 为 n 维的单位列向量,式中第 i 个元素为 1,其他全为 0,则式(8)可变成

$$\min_{\mathbf{w}} \left(\mathbf{w}^T \left(\sum_i^n (\mathbf{X}\mathbf{u}_i - \mathbf{X}\mathbf{s}_i)(\mathbf{X}\mathbf{u}_i - \mathbf{X}\mathbf{s}_i)^T \right) \mathbf{w} \right) =$$

$$\begin{aligned} & \min_{\mathbf{w}} \left(\mathbf{w}^T \mathbf{X} \left(\sum_i^n (\mathbf{u}_i - \mathbf{s}_i)(\mathbf{u}_i - \mathbf{s}_i)^T \right) \mathbf{X}^T \mathbf{w} \right) = \\ & \min_{\mathbf{w}} \left(\mathbf{w}^T \mathbf{X} \left(\sum_i^n (\mathbf{u}_i \mathbf{u}_i^T - \mathbf{s}_i \mathbf{u}_i^T - \mathbf{u}_i \mathbf{s}_i^T + \mathbf{s}_i \mathbf{s}_i^T) \right) \mathbf{X}^T \mathbf{w} \right) = \\ & \min_{\mathbf{w}} \mathbf{w}^T \mathbf{X}(\mathbf{I}_n - \mathbf{S} - \mathbf{S}^T + \mathbf{S}^T \mathbf{S}) \mathbf{X}^T \mathbf{w} \end{aligned} \quad (9)$$

式中, $\mathbf{I}_n$ 是一个 n 阶的单位矩阵,为方便起见,将最小化问题转换为最大化问题

$$\max_{\mathbf{w}} \mathbf{w}^T \mathbf{X}(\mathbf{S} + \mathbf{S}^T - \mathbf{S}^T \mathbf{S} - \mathbf{I}_n) \mathbf{X}^T \mathbf{w} \quad (10)$$

式中,令 $\mathbf{S}_\alpha = \mathbf{S} + \mathbf{S}^T - \mathbf{S}^T \mathbf{S}$,并增加尺度约束 $\mathbf{w}^T \mathbf{X} \mathbf{X}^T \mathbf{w} = 1$,则式(10)的优化可转换为广义特征值问题的求解,即

$$\mathbf{X}(\mathbf{S}_\alpha - \mathbf{I}_n) \mathbf{X}^T \mathbf{w} = \lambda_0 \mathbf{X} \mathbf{X}^T \mathbf{w} \quad (11)$$

式(11)可变换成 $\mathbf{X}\mathbf{S}_\alpha \mathbf{X}^T \mathbf{w} = (\lambda_0 + 1) \mathbf{X} \mathbf{X}^T \mathbf{w}$,令 $\lambda = \lambda_0 + 1$,则

$$\mathbf{X}\mathbf{S}_\alpha \mathbf{X}^T \mathbf{w} = \lambda \mathbf{X} \mathbf{X}^T \mathbf{w} \quad (12)$$

求解该广义特征值问题,并取前 d 个最大的特征值所对应的特征向量构成投影矩阵 $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_d]$ 。

综上所述,SPP 首先通过优化式(5)或式(6)构建稀疏系数矩阵,再利用式(12)优化子空间得到投影矩阵 $\mathbf{W}$ 。

然而在实际应用中,样本的维数一般远大于样本数,会遇到维数灾难和小样本问题,使得不但无法很好地求解式(12),而且也不满足稀疏表示理论中所要求的字典中原子的个数应不小于样本的维数。因此,SPP 方法很难直接应用到实际的模式识别问题中,一般采用先对原始样本进行降维(如 PCA),而后再进行 SPP 的运算来进一步优化。

3 核稀疏保留投影(KSPP)

针对 SPP 在解决实际问题时会遇到的维数灾难和小样本问题,可以通过引入核方法来解决。

首先利用非线性映射 $\phi: \mathbf{R}^n \rightarrow \mathbf{F}, \mathbf{x} \rightarrow \phi(\mathbf{x}), \mathbf{X} \rightarrow \Phi(\mathbf{X})$,将输入样本映射到一个高维(甚至是无穷维)的 Hilbert 空间 $\mathbf{F}$ 中,则方式 1 可变成

$$\begin{aligned} & \min_{\mathbf{s}_i} \|\mathbf{s}_i\|_1 \\ & \text{s. t. } \|\phi(\mathbf{x}_i) - \Phi(\mathbf{X})\mathbf{s}_i\| < \varepsilon, 1 = \mathbf{1}_n^T \mathbf{s}_i \end{aligned} \quad (13)$$

因为特征空间 $\mathbf{F}$ 的维数很高,不可能进行通常的运算,而核方法可以利用内积核函数来隐含地进行运算。核函数是解决原样本空间中线性不可分问题时引入的一种非线性变换,通过构造内积来回避

计算非线性函数的复杂过程。定义核函数 $K_{ij} = (\phi(x_i) \cdot \phi(x_j))$, 则

$$\begin{aligned} & \|\phi(x_i) - \Phi(X)s_i\|^2 = \\ & K_{ii} - K_i^T s_i - s_i^T K_i + s_i^T K s_i \end{aligned} \quad (14)$$

式中, K_i 为核矩阵 K 的第 i 列, 将式 (14) 代入式 (13), 经分析可知, 该问题很难进行优化求解, 因此这里放弃方式 1。

在方式 2 中, t_i 的维数和 $\phi(x)$ 相同, 也是一个高维向量。由核空间理论可知, 高维空间中的矢量 t_i 可以由样本 $\{\phi(x_i); i = 1, 2, \dots, N\}$ 张成, $t_i = \Phi(X)\beta_i$, 则式 (6) 可变成

$$\begin{aligned} & \min_{s_i} \|\left[s_i^T \beta_i^T\right]^T\|_1 \\ & \text{s. t. } \begin{bmatrix} \phi(x_i) \\ 1 \end{bmatrix} = \begin{bmatrix} \Phi(X) & \Phi(X) \\ \mathbf{1}_n^T & \mathbf{0}_n^T \end{bmatrix} \begin{bmatrix} s_i \\ \beta_i \end{bmatrix} \end{aligned} \quad (15)$$

式中, β_i 是一个 n 维的列向量, $\mathbf{0}_n$ 是一个元素全为 0 的 n 维列向量。

为了引出高维空间里样本的内积, 式 (15) 中的约束等式两边同时左乘 $[\Phi(X)^T, \mathbf{1}_n]$, 则等式左边变成

$$[\Phi(X)^T, \mathbf{1}_n] \begin{bmatrix} \phi(x_i) \\ 1 \end{bmatrix} = K_i + \mathbf{1}_n \quad (16)$$

等式右边变成

$$\begin{aligned} & [\Phi(X)^T, \mathbf{1}_n] \begin{bmatrix} \Phi(X) & \Phi(X) \\ \mathbf{1}_n^T & \mathbf{0}_n^T \end{bmatrix} \begin{bmatrix} s_i \\ \beta_i \end{bmatrix} = \\ & [K + \mathbf{1}_{nn}, K] \begin{bmatrix} s_i \\ \beta_i \end{bmatrix} \end{aligned} \quad (17)$$

式中, $\mathbf{1}_{nn}$ 为元素全为 1 大小为 $n \times n$ 的矩阵, 式 (15) 最终变成

$$\begin{aligned} & \min_{s_i} \|\left[s_i^T \beta_i^T\right]^T\|_1 \\ & \text{s. t. } K_i + \mathbf{1}_n = [K + \mathbf{1}_{nn}, K] \begin{bmatrix} s_i \\ \beta_i \end{bmatrix} \end{aligned} \quad (18)$$

可以看出该式的形式与式 (6) 相似, 因此由式 (18) 可得系数矩阵 S , 定义目标函数

$$\min_w \sum_i \|\mathbf{w}^T \phi(x_i) - \mathbf{w}^T \Phi(X)s_i\|^2 \quad (19)$$

与 t_i 类似, 这里 w 也可以由样本 $\{\phi(x_i); i = 1, 2, \dots, N\}$ 张成, 则 $w = \Phi(X)\gamma$, 根据式 (8) — (10) 的推导, 另外由于核矩阵 K 是一个实对称阵, 式 (19) 可变成

$$\begin{aligned} & \max_w \mathbf{w}^T \Phi(X) (S + S^T - S^T S - I_n) \Phi(X)^T \mathbf{w} = \\ & \max_\gamma \gamma^T K (S + S^T - S^T S - I_n) K^T \gamma \end{aligned} \quad (20)$$

这里增加尺度约束 $\mathbf{w}^T \Phi(X) \Phi(X)^T \mathbf{w} = 1$, 又由于 $w = \Phi(X)\gamma$, 则

$$\begin{aligned} & \mathbf{w}^T \Phi(X) \Phi(X)^T \mathbf{w} = \\ & \gamma^T \Phi(X)^T \Phi(X) \Phi(X)^T \Phi(X) \gamma = \\ & \gamma^T K K^T \gamma \end{aligned} \quad (21)$$

与式 (11) (12) 类似, 最终可得

$$K S_\alpha K^T \gamma = \lambda K K^T \gamma \quad (22)$$

求解该广义特征值问题, 并取前 d 个最大的特征值所对应的特征向量构成投影矩阵 $r = [\gamma_1, \gamma_2, \dots, \gamma_d]$ 。

4 步态识别

步态识别^[12]是近年来计算机视觉和生物特征识别领域的一个备受关注的研究方向, 其目的是根据人们走路的姿势进行身份识别。Han 等人^[13]采用步态能量图 (GEI) 的形式表征步态, 取得了良好的效果。

首先对摄像机采集到的图像通过背景减除实现人体检测, 运用形态学操作和图形几何变换将图像标准中心化成统一的大小。而后进行周期检测提取一个周期的步态序列, 并生成 GEI。对于处理后的二值步态周期图像序列 $B_i(x, y), i = 1, 2, \dots, M$, 式中 M 为一个周期的步态序列的帧数, (x, y) 代表 2 维图像平面坐标, GEI 的定义为

$$G(x, y) = \frac{1}{M} \sum_{i=1}^M B_i(x, y) \quad (23)$$

在实际生活中, 人体的行走方向和行走状态完全是随机的, 因此会得到不同的能量图。图 1 为正常状态时不同视角 (人体行走方向与摄像机光轴方向夹角) 下所得到的 GEI。

图 2 为同一个人 90° 视角下的不同状态的 GEI。

在使用稀疏表示方法进行步态识别时, 为了增大步态能量图的稀疏程度, 提高最终的识别效果, 可以对训练样本进行零均值化, 去除冗余的相关信息。采用 SPP 时, 在使用 PCA 方法时就需要用到零均值化过程。而采用 KSPP 时, 需要在样本映射到高维空间后进行零均值化, 而在核空间这一过程可以通过对核矩阵 K 中心化来完成^[10]。

$$\tilde{K} = K - \frac{1}{n} K \mathbf{1}_{nn} - \frac{1}{n} \mathbf{1}_{nn} K + \frac{1}{n^2} \mathbf{1}_{nn} K \mathbf{1}_{nn} \quad (24)$$

利用式 (24) 对 K 中心化, 得到 $\tilde{K}$ 。根据式 (24) 的结构可以看出 $\tilde{K}$ 仍为实对称阵, 因此将 $\tilde{K}$ 代入式 (22) 即可。

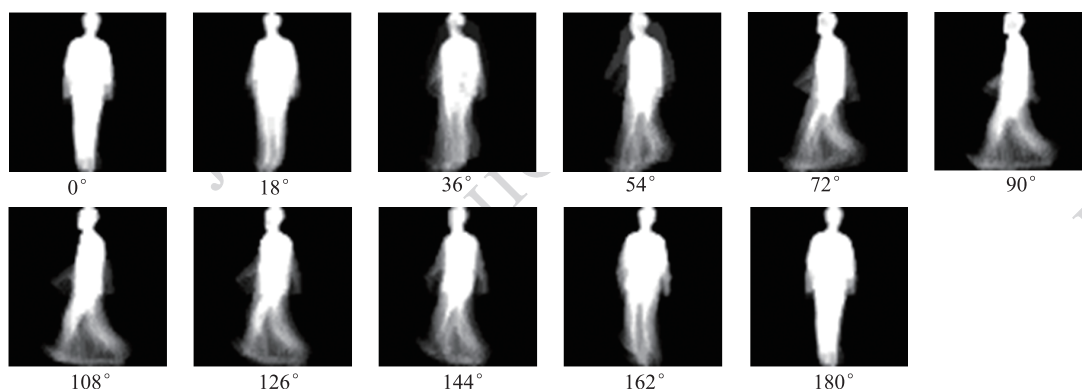


图1 正常状态下的不同视角的 GEI

Fig.1 The GEI of different views under walking normally

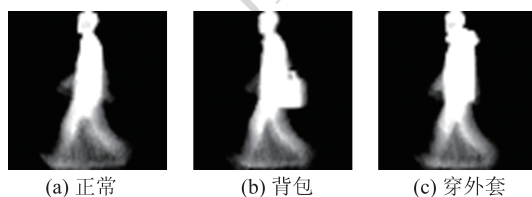


图2 同一个人 90°视角下的不同状态的 GEI

Fig.2 The GEI of different conditions under the view of 90°

维(所取特征的比重约为总特征的 99%),否则在线性规划过程中会遇到严重的维数灾难问题。而本文方法直接对 372×372 的矩阵进行线性规划,只需要比 PCA 略多的时间。在识别过程中,几种方法所需时间大体相当,均在 0.2 s 以内,能够很好地满足步态识别所需的实时性。

表1 实验1中几种方法的运算时间

Table 1 The computation time of several algorithms in experiment 1

方法	训练时间	识别时间
PCA	732.730 7	0.163 0
KPCA	16.713 5	0.169 5
PCA + SPP	1 147.453 3	0.171 9
KSPP	860.419 8	0.185 9

/s

5 实验分析

常见的核函数有多项式核函数、高斯核函数、Sigmoid 核函数等,而高斯核函数的特征空间维数是无穷大^[9],因此本文选择高斯核函数,高斯核函数的参数在 KPCA 中 $\sigma^2 = 0.1$,在 KSPP 中 $\sigma^2 = 0.02$ 。实验中 SPP 与 KSPP 方法均采用第 2 种优化方式。

实验采用 CASIA 步态数据库中的 B 库^[14],该库总共 124 人,在 11 个视角下采集,每人每个视角下 6 个视频为正常状态,2 个视频为背包状态,2 个视频为穿外套状态。实验中步态能量图的大小为 64×64 。

实验1 几种方法的运算时间

实验选取正常状态下 90°视角的前 3 组图像作为训练集,后 3 组图像作为测试集,注册集也采用前 3 组图像,采用最近邻分类器分类识别。

表 1 是在 CPU 2.93 GHz,内存为 2 GB, MATLAB 7.1 环境下几种方法的运算时间,可以看出 KPCA 方法具有最快的训练速度,这是由于该方法仅需对 372×372 的矩阵进行运算,PCA 需要对 $4 096 \times 4 096$ 的矩阵进行运算,而 PCA + SPP 和 KSPP 由于在稀疏重构过程中需要进行线性规划过程,因此需要较长的时间。实验中 PCA + SPP 首先采用 PCA 方法将维数从 $4 096 \times 372$ 降到 300×372

实验2 不同视角的步态图像之间的度量

实验选取正常状态下每个视角前 3 组图像作为训练集,后 3 组图像作为测试集,注册集也采用前 3 组图像,采用最近邻分类器分类识别。

从表 2 可以看出,相对于 PCA 方法,KPCA 方法的识别率在有的视角下得到了提升,而有的视角却降低了,这是由于在 KPCA 方法中核函数采用单一的参数设置不能适应不同视角下的数据。加入 SPP 方法后,大部分视角的识别率得到了提升,尤其是识别率较低的视角,识别率有了较大提升。这是由于加入 SPP 方法后,优化了原始特征的分布情况,提高了特征的可分性。而通过引入核方法,最终的识别率得到了进一步提升,说明 KSPP 不但能解决所存在的小样本问题,并且对不同数据具有很高的稳定性。

表 2 实验 2 中几种方法的识别率

Table 2 The identification rate of several algorithms in experiment 2

方法	视角											均值
	0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°	
PCA	96.24	67.47	61.56	79.30	94.09	93.28	92.74	89.52	83.33	93.55	96.24	86.12
KPCA	95.97	67.74	60.75	78.76	94.35	92.20	93.01	88.98	82.80	93.28	96.24	85.83
PCA + SPP	93.55	77.15	68.28	81.99	95.43	95.16	92.47	93.01	93.28	94.35	94.35	89.00
KSPP	95.16	79.84	70.70	84.68	96.24	96.77	94.09	94.35	92.74	93.55	95.16	90.30

实验 3 不同状态的步态图像之间的度量

实验分别选取 90° 视角下正常状态下前 3 组图像和背包、穿外套状态下的前一组图像作为训练集,测试集分别采用 90° 视角下正常状态下前 3 组图像,背包、穿外套状态下的后一组图像,注册集也采用前一半图像,最后采用最近邻分类器分类识别。

在表 3 的实验数据中,相对 PCA 方法,KPCA 方法未得到较高的识别率,可见在对 3 种步态一同训练时,KPCA 严重打乱了原有 3 种步态各自所具有的近邻关系。而加入 SPP 方法后,保留了原始样本所具有的局部近邻信息,使得识别率得到了提升,但提升并不是很大,而且识别率的提升主要在正常状态下的步态。这是由于这个实验中背包和穿外套状态下的训练集合是单样本集合,存在严重的小样本问题,而且 SPP 方法的重构系数矩阵是采用除本身外的所有样本作为数据字典得到的,理论上,较大的几个系数应该集中在被重构样本的同类样本上,因此单样本集合的重构系数会比较分散,使得最终得到的重构系数矩阵稀疏程度较差,以致于无法得到较高的识别效果。而采用 KSPP 时,通过解决小样本问题,使得最终在 3 种状态下均得到更高的识别率。

表 3 实验 3 中几种方法的识别率

Table 3 The identification rate of several algorithms in experiment 3

方法	测试集				/%
	正常	背包	穿外套	所有的测试集	
PCA	93.01	86.29	88.71	90.81	
KPCA	92.47	85.48	88.71	90.32	
PCA + SPP	95.16	84.68	89.52	91.94	
KSPP	96.51	87.90	90.32	93.55	

6 结 论

稀疏保留投影方法在投影时很好地保留了原始数据所蕴含的信息,在解决实际的模式识别和计算机视觉问题方面具有广泛的应用价值。而通过分析实验结果可以看出,基于核空间的稀疏保留投影取得了比基于线性映射的稀疏保留投影更好的识别效果,能更好的对模式进行表征。接下来的工作主要是引入监督信息,将稀疏保留投影方法扩展成有监督的降维方法。

志 谢

中国科学院自动化研究所为本文研究提供了 CASIA 步态数据库,在此表示感谢。

参考文献 (References)

- [1] Fisher R A. The use of multiple measurements in taxonomic problems[J]. Annals of Eugenics, 1936,48(7): 179-188.
- [2] Xing E, Ng A, Jordan M, et al. Distance metric learning with application to clustering with side-information [C]//Advances in Neural Information Processing Systems. Berkeley: University of California, 2003:521-528.
- [3] Hotelling H. Analysis of a complex of statistical variables into principal components[J]. Journal of Educational Psychology, 1933,24(6):417-441.
- [4] Comon P. Independent components analysis, a new concept[J]. Signal Processing, 1994,36(3):287-314.
- [5] Roweis S T, Saul L K. Nonlinear dimensionality reduction by locally linear embedding[J]. Science, 2000, 290(5500): 2323-2326.
- [6] Belkin M, Niyogi P. Laplacian eigenmaps and spectral techniques for embedding and clustering[C]//Advances in Neural Information Processing System. Chicago: University of Chicago, 2002:

- 585-591.
- [7] He X F, Niyogi P. Locality preserving projections [C]// Advances in Neural Information Processing System. Chicago: University of Chicago, 2004: 153-160.
- [8] Qiao L S, Chen S C, Tan X Y. Sparsity preserving projections with applications to face recognition [J]. Pattern Recognition, 2010, 43(1): 331-341.
- [9] Scholkopf B, Smola A J. Learning With Kernels [M]. Cambridge: MIT Press, 2002: 1-21.
- [10] Scholkopf B, Smola A, Muller K R. Nonlinear component analysis as a kernel eigenvalue problem [J]. Neural Computation, 1998, 10(5): 1299-1319.
- [11] Chen S, Donoho D, Saunders M. Atomic decomposition by basis pursuit [J]. SIAM Review, 2001, 43(1): 129-159.
- [12] Wang K J, Hou B B. A survey of gait recognition [J]. Journal of Image and Graphics, 2007, 12(7): 1152-1160. [王科俊, 侯本博. 步态识别综述 [J]. 中国图象图形学报, 2007, 12(7): 1152-1160.]
- [13] Han J, Bhanu B. Individual recognition using gait energy image [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(2): 316-322.
- [14] Yu S Q, Tan D L, Tan T N. A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition [C]// Proceedings of the 18th International Conference on Pattern Recognition. Hong Kong, China: IEEE Computer Society, 2006: 441-444.