

LSPSA: 基于局部结构保持的共享子空间分析

杜琳琳 朱振峰 段红帅 赵耀

(北京交通大学信息科学研究所 北京 100044)

(现代信息科学与网络技术北京市重点实验室 北京 100044)

摘要 多输出正则投影(MORP)算法将输入特征向量和由类标签形成的多输出特征向量经过因子分解方法映射到一个共享子空间,从而建立输入特征与类标签的关联。在 MORP 的基础上,通过引入图约束,提出了一种基于局部结构保持的共享子空间分析方法(LSPSA),该方法在获取共享信息的同时,保持原始多视角特征空间与共享子空间中的数据具有相近的局部几何结构关系,从而避免多视角数据在共享子空间的过拟合问题。此外,还提出了一种图模型逼近方法,实现了 LSPSA 的在线扩展,解决了在线获取新测试样本在共享子空间中表征的高复杂度问题。在 UCI 多特征手写体数据库上的分类及检索实验验证了所提出的算法的有效性。

关键词 多视角,共享子空间,局部结构保持,图模型,在线扩展

中图分类号 TP391 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2014.10.015

Local Structure Preserved Shared-subspace Analysis

DU Lin-lin ZHU Zhen-feng DUAN Hong-shuai ZHAO Yao

(Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China)

(Beijing Key Laboratory of Advanced Information Science and Network Technology, Beijing 100044, China)

Abstract With the rapid development of information technology, multi-view data has become increasingly common and how to obtain the shared information from the multi-view data has become one of the hottest research topics in the field of machine learning. As a shared subspace method for multi-view data, Multi-output regularized feature projection (MORP) has been proposed recently to build the correlation of multi-view data in the shared subspace by matrix factorization. Compared with the classical multi-view analysis method CCA, MORP has been proved to be more effective. On the basis of MORP, we proposed a local structure preserved shared-subspace analysis (LSPSA) method by imposing an extra graph constraint. While obtaining the shared information from multi-view data like MORP, the local geometrical structure of data in both shared subspace and original multi-view feature space can be well preserved. Thus, in the obtained shared subspace, the over-fitting problem of multi-view data can be avoided to some extent for MORP model. Meanwhile, we also proposed a graph approximating method to provide an online extension of LSPSA for the problem of out-of-sample. Without loss of performance, the computational complexity of online extension of LSPSA for seeking the representation of out-of-sample in the shared subspace can be reduced greatly, especially with the increasing size of dataset. The final experimental results on UCI multi-view hand-written digit dataset demonstrate that LSPSA achieves much better performance for classification and retrieval tasks.

Keywords Multi-view, Shared-subspace, Local structure preserving, Graph model, Online extension

1 引言

随着信息技术的迅猛发展,人们生产与收集的数字化数据越来越复杂。除了数据的维度越来越高,表示形式也越来越多样化,呈现出多态性、多源性和多描述性的特点,这些数据通常被统一称为多视角数据。如何通过多视角分析挖掘多视角数据所包含的共享信息已成为当前机器学习领域研究的

热点^[1-3]。

目前,分析多视角问题的代表性方法有:典型相关分析(Canonical Correlation Analysis, CCA)^[4]、偏最小二乘(Partial Least Squares, PLS)^[5]、核 CCA^[6]、半监督 CCA^[7]以及核 PLS^[8]。CCA 是 Hotelling 在 1936 年提出的一种研究同一对象的两个视角数据间相关性的有效算法,该算法通过将多视角数据进行合适的线性变换使两个视角之间具有最大的相关

到稿日期:2013-07-02 返修日期:2013-08-09 本文受国家“973”计划(2012CB316400),国家自然科学基金(61172129),中央高校基本科研业务费专项资金(2012JBZ012),北京市自然科学基金(4112043)和 PCSIRT(IRT201206)资助。

杜琳琳(1988-),女,硕士生,主要研究方向为模式识别;朱振峰(1974-),男,副教授,主要研究方向为图像与视频理解、计算机视觉、机器学习等, E-mail:zhfzhu@bjtu.edu.cn;段红帅(1987-),女,硕士生,主要研究方向为图像处理、机器学习;赵耀(1967-),男,教授,博士生导师,主要研究方向为图像与视频编码、数字水印与信息隐藏、基于内容的信息检索。

性。PLS是一种综合了PCA(主成分分析)和CCA的算法,该算法既具有典型相关分析的相关性分析能力,同时还具备主成分分析保持数据变化信息的能力^[9]。最近,Abhishek等人提出了广义多视角分析(Generalized Multiview Analysis, GMA)^[10],并指出PCA、CCA以及PLS等均可视作GMA框架下的特殊形式。此外,在多标签的情况下,可以结合标签和特征的子空间。Chen等人在CCA的基础上提出了用于多标签分类的特征感知的标签子空间降维方法^[11]。该方法相比于只利用标签的子空间分析方法效果有很大的提高。

近年来,Yu等人提出了多输出正则投影算法^[12](Multi-Output Regularized feature Projection, MORP),该算法将输入特征向量和由类标签形成的多输出特征向量映射到一个共享子空间,从而建立输入特征与类标签的关联。MORP虽然作为一种监督投影算法被提出,但如果把多个输出特征看作是输入特征共同描述同一对象的某个视角,该算法同样适用于多视角问题分析。研究结果表明,应用MORP算法提取多视角数据的关联特征进行分类识别,性能优于CCA。然而,在多输出正则投影模型中,由于没有考虑样本的局部结构特性,原始多视角特征经因子分解映射到共享子空间后,多视角数据的原始样本结构无法得到很好的保持,同时也易于产生过拟合问题。此外,针对新的测试样本(Out-of-Sample)问题,MORP没有给出一个有效的求解方法。

针对上述问题,本文在MORP的基础上,通过引入图约束来保持共享子空间与原始多视角数据的局部结构特性,以克服过拟合问题。此外,提出了一种逼近方法,以解决新测试样本(Out-of-Sample)的在线扩展问题。

2 多输出正则投影

多输出正则投影算法采用监督方式,通过因子分解将输入特征向量和由类标签形成的多输出特征向量映射到一个共享子空间,从而建立输入特征与类标签的关联。设有描述某一对象的 n 个观测样本 $\{x_i, y_i\} \in \mathbb{R}^{1 \times s} \times \mathbb{R}^{1 \times t}$,其中 $\{x_i\}_{i=1}^n$ 和 $\{y_i\}_{i=1}^n$ 分别来源于视角 X 和视角 Y , $X=[x_1, \dots, x_n]^T \in \mathbb{R}^{n \times s}$ 和 $Y=[y_1, \dots, y_n]^T \in \mathbb{R}^{n \times t}$,用 $\{x_i, y_i\}$ 表示样本集 $\{x_i, y_i\}_{i=1}^n$ 的第 i 对样本。则多输出正则投影对应的优化问题可表述为:

$$\begin{aligned} \min_{A, B, U} (1-\beta) \|X-UA\|_F^2 + \beta \|Y-UB\|_F^2 \\ \text{s. t. } U^T U = I \end{aligned} \quad (1)$$

其中, $U=[u^1, \dots, u^p] \in \mathbb{R}^{n \times p}$ 为视角 X 和视角 Y 在共享子空间中的 p 维投影描述, $A \in \mathbb{R}^{p \times s}$ 和 $B \in \mathbb{R}^{p \times t}$ 分别是视角 X 和视角 Y 的载入矩阵, $\|\cdot\|_F$ 表示Fresenius范数且 $\|A\|_F = \sqrt{\sum_{i,j} a_{ij}^2}$, $\beta(0 < \beta < 1)$ 作为一个权重系数来平衡两个视角的重建误差,模型(1)中的正交约束 $U^T U = I$ 的目的是保持 p 维共享子空间中的 p 个变量彼此独立。经过化简,模型(1)等价于:

$$\begin{aligned} \max_{U \in \mathbb{R}^{n \times p}} U^T ((1-\beta)XX^T + \beta YY^T)U \\ \text{s. t. } U^T U = I \end{aligned} \quad (2)$$

由最优化问题(2)可以看出,MORP模型的最优解 $U=[u^1, \dots, u^p] \in \mathbb{R}^{n \times p}$ 由矩阵 $(1-\beta)XX^T + \beta YY^T$ 的 p 个最大特征值对应的特征向量组成。

3 局部结构保持的共享子空间分析

为了解决MORP在保持局部结构特性时易产生多视角数据在共享子空间中过拟合的问题,本文在最优化问题(1)的基础上提出了如下优化模型:

$$\begin{aligned} \min_{A, B, U} (1-\beta) \|X-UA\|_F^2 + \beta \|Y-UB\|_F^2 + \\ \lambda \sum_i \sum_j \left\| \frac{u_i}{\sqrt{d_i}} - \frac{u_j}{\sqrt{d_j}} \right\|_F^2 w_{ij} \\ \text{s. t. } U^T U = I \end{aligned} \quad (3)$$

其中, $\sum_i \sum_j \left\| \frac{u_i}{\sqrt{d_i}} - \frac{u_j}{\sqrt{d_j}} \right\|_F^2 w_{ij}$ 为图约束项, w_{ij} 为反映图中第 i 个样本节点与第 j 个样本节点的关系权值,且有:

$$w_{ij} = \begin{cases} \exp[-d^2(x_i, y_j)/2\sigma^2], & x_i \in N_k(x_j), \text{ or } x_j \in N_k(x_i) \\ 0, & \text{else} \end{cases}$$

其中, $d(x_i, y_j)$ 为某种距离度量, $x_i \in N_k(x_j)$ 表示 x_i 属于 x_j 的 k 个近邻,本文中采用欧氏距离度量。 σ 为高斯核半径, $\sigma = \frac{\sum_{i=1}^n \sum_{j=1}^n d(x_i, y_j)/n^2}$ 。

化简式(3)中的约束项,容易得到:

$$\begin{aligned} \min_{A, B, U} (1-\beta) \|X-UA\|_F^2 + \beta \|Y-UB\|_F^2 + \\ \lambda \text{Tr}(U^T L_u U) \\ \text{s. t. } U^T U = I \end{aligned} \quad (4)$$

其中, $L_u = I - D^{-0.5} W D^{-0.5}$ 为拉普拉斯矩阵, $W=[w_{ij}]$, $D = \text{diag}(\sum_j w_{ij})$ 。在由式(4)给出的优化问题中,通过约束项 $\sum_i \sum_j \left\| \frac{u_i}{\sqrt{d_i}} - \frac{u_j}{\sqrt{d_j}} \right\|_F^2 w_{ij} = \text{Tr}(U^T L_u U)$,可以尽量地保证在高维空间中邻近的点投影后在低维空间中对应的映射值仍然是邻近的,从而保持了样本数据的局部结构特性。为此,本文称模型(4)为局部结构保持的共享子空间分析模型(Local Structure Preserved Shared-subspace Analysis, LSPSA)。

仿照文献[12],针对模型(4)的求解,有如下定理1和定理2。

定理1 若 U, A 和 B 是最优化问题(4)的最优解,则有: $A=U^T X, B=U^T Y$ 。

证明:设

$$J(A, B, U, L_u) = (1-\beta) \|X-UA\|_F^2 + \beta \|Y-UB\|_F^2 + \lambda \text{Tr}(U^T L_u U)$$

对 A, B 求导可得:

$$\begin{cases} \frac{\partial J}{\partial A} = 2(1-\beta)(U^T X - U^T U A) = 0 \Rightarrow A = U^T X \\ \frac{\partial J}{\partial B} = 2\beta(U^T Y - U^T U B) = 0 \Rightarrow B = U^T Y \end{cases}$$

定理1得证。

定理2 若 U, A 和 B 是最优化问题(4)的最优解,且 $G=(1-\beta)XX^T + \beta YY^T$,则最优化问题(4)可转化为:

$$\begin{aligned} \max_{u \in \mathbb{R}^{n \times 1}} u^T (G - \lambda L_u) u \\ \text{s. t. } u^T u = 1 \end{aligned} \quad (5)$$

且最优解 $U=[u^1, \dots, u^p]$ 由矩阵 $G - \lambda L_u$ 的 p 个最大特征值对应的特征向量组成。

证明:同定理1,令

$$J(A, B, U, L_u) = (1-\beta) \|X-UA\|_F^2 + \beta \|Y-UB\|_F^2 +$$

$$\lambda \text{Tr}(U^T L_u U)$$

由 $\frac{\partial J}{\partial A}=0$ 与 $\frac{\partial J}{\partial B}=0$ 可得 $A=U^T X$ 及 $B=U^T Y$ 。进一步可

有:

$$\begin{aligned} J(A, B, U, L_u) &= (1-\beta) \|X-UA\|_F^2 + \beta \|Y-UB\|_F^2 + \\ &\quad \lambda \text{Tr}(U^T L_u U) \\ &= (1-\beta) \text{Tr}((X^T - A^T U^T)(X-UA)) + \\ &\quad \beta \text{Tr}((Y^T - B^T U^T)(Y-UB)) + \lambda \text{Tr} \\ &\quad (U^T L_u U) \\ &= \text{Tr}[(1-\beta)XX^T + \beta YY^T] - \text{Tr}[U^T \\ &\quad ((1-\beta)XX^T + \beta YY^T)U] + \lambda \text{Tr} \\ &\quad (U^T L_u U) \\ &= \text{Tr}(G) - \text{Tr}[U^T(G-\lambda L_u)U] \end{aligned}$$

为此,最优化问题(4)等价于:

$$\begin{aligned} \max_{U \in \mathbb{R}^{n \times p}} & U^T(G-\lambda L_u)U \\ \text{s. t. } & U^T U = I \end{aligned} \quad (6)$$

不难看出,式(6)的最优解 $U=[u^1, \dots, u^p] \in \mathbb{R}^{n \times p}$ 由矩阵 $G-\lambda L_u$ 的 p 个最大特征值对应的特征向量组成。定理 2 得证。

在计算复杂度方面,由于 LSPSA 模型涉及 $n \times n$ 矩阵 $G-\lambda L_u$ 的特征值求解(由式(6)给出),因此其同 MORP、CCA、PLS 具有同样的计算复杂度,即 $O(n^3)$ 。

4 局部结构保持的共享子空间分析模型的在线扩展

如前所述,对于新测试样本(Out-of-Sample) $\{x_{n+1},$

$$L_u^{[n+1]} = \begin{bmatrix} L_u^{[n]} & (L_{n+1, \phi^n})^T \\ L_{n+1, \phi^n} & L_{n+1, n+1} \end{bmatrix} = I^{[n+1]} - \begin{bmatrix} (D^{[n]})^{-0.5} W^{[n]} (D^{[n]})^{-0.5} & (D^{[n]})^{-0.5} (W_{n+1, \phi^n})^T (D_{n+1, n+1})^{-0.5} \\ (D_{n+1, n+1})^{-0.5} W_{n+1, \phi^n} (D^{[n]})^{-0.5} & (D_{n+1, n+1})^{-0.5} W_{n+1, n+1} (D_{n+1, n+1})^{-0.5} \end{bmatrix} \quad (7)$$

其中, $D^{[n+1]} = \begin{bmatrix} D^{[n]} & 0 \\ 0 & D_{n+1, n+1} \end{bmatrix}$ 。令 $\Phi^{n+1} = (A^{n+1}, B^{n+1},$

U^{n+1}), 则式(4)转化为:

$$\begin{aligned} \min_{\Phi^{n+1}} & (1-\beta) \|X^{n+1} - U^{n+1} A^{n+1}\|_F^2 + \beta \|Y^{n+1} - \\ & U^{n+1} B^{n+1}\|_F^2 + \lambda \text{Tr}((U^{n+1})^T L_u^{n+1} U^{n+1}) \end{aligned} \quad (8)$$

s. t. $(U^{n+1})^T U^{n+1} = I^{[n+1]}$

其中, $A^{n+1} \in \mathbb{R}^{p \times s}$ 和 $B^{n+1} \in \mathbb{R}^{p \times t}$ 为原有训练样本加入新测试

样本后获得的载入矩阵, $U^{n+1} = \begin{bmatrix} U^n \\ u_{n+1}^T \end{bmatrix} \in \mathbb{R}^{(n+1) \times p}$ 为所有样本

$$\begin{aligned} L_u^{[n+1]} &= \begin{bmatrix} L_u^{[n]} & (L_{n+1, \phi^n})^T \\ L_{n+1, \phi^n} & L_{n+1, n+1} \end{bmatrix} \approx I^{[n+1]} - \begin{bmatrix} (\tilde{D}^{[n]})^{-0.5} \tilde{W}^{[n]} (\tilde{D}^{[n]})^{-0.5} & (\tilde{D}^{[n]})^{-0.5} (W_{n+1, \phi^n})^T (D_{n+1, n+1})^{-0.5} \\ (D_{n+1, n+1})^{-0.5} W_{n+1, \phi^n} (\tilde{D}^{[n]})^{-0.5} & (D_{n+1, n+1})^{-0.5} W_{n+1, n+1} (D_{n+1, n+1})^{-0.5} \end{bmatrix} \\ &= \begin{bmatrix} L_u^{[n]} & -(\tilde{D}^{[n]})^{-0.5} (W_{n+1, \phi^n})^T (D_{n+1, n+1})^{-0.5} \\ - (D_{n+1, n+1})^{-0.5} W_{n+1, \phi^n} (\tilde{D}^{[n]})^{-0.5} & 1 - (D_{n+1, n+1})^{-0.5} W_{n+1, n+1} (D_{n+1, n+1})^{-0.5} \end{bmatrix} \end{aligned} \quad (9)$$

与此同时,当训练样本足够大时,新加入的测试样本也不会对样本在共享子空间中的投影结构造成很大影响,或者说 $U^{n+1}, A^{n+1}, B^{n+1}$ 近似不变,即有 $A^{n+1} \approx \tilde{A}^n, B^{n+1} \approx \tilde{B}^n, U^{n+1} =$

$$\begin{bmatrix} U^n \\ u_{n+1}^T \end{bmatrix}。$$

通过上述逼近,由式(8)给出的优化问题可退化为:

$$\begin{aligned} \min_{u_{n+1}} & (1-\beta) \left\| \begin{bmatrix} X^n \\ x_{n+1}^T \end{bmatrix} - \begin{bmatrix} \tilde{U}^n \\ u_{n+1}^T \end{bmatrix} \tilde{A}^n \right\|_F^2 + \\ & \beta \left\| \begin{bmatrix} Y^n \\ y_{n+1}^T \end{bmatrix} - \begin{bmatrix} \tilde{U}^n \\ u_{n+1}^T \end{bmatrix} \tilde{B}^n \right\|_F^2 + \end{aligned}$$

y_{n+1} }, 如果要在共享子空间中获取与之对应的 u_{n+1} , MORP 则需根据式(4)重新求解一个复杂度为 $O((n+1)^3)$ 的特征值问题。当样本规模很大时,此时的高计算复杂度将很难满足一些实时性要求较高的在线处理需求,如检索等。为此,本文提出一种模型逼近技术,用以实现局部结构保持的共享子空间分析(LSPSA)的在线扩展,解决在线获取新测试样本 $\{x_{n+1}, y_{n+1}\}$ 在共享子空间中表征 u_{n+1} 的高复杂度问题。

为便于问题描述,这里重新引入一些符号说明。用 X^n 和 Y^n 分别表示训练样本集合,上标 n 表示对应矩阵由原始 n 个训练样本构造, $X^{n+1} = \begin{bmatrix} X^n \\ x_{n+1}^T \end{bmatrix} \in \mathbb{R}^{(n+1) \times s}$ 和 $Y^{n+1} =$

$\begin{bmatrix} Y^n \\ y_{n+1}^T \end{bmatrix} \in \mathbb{R}^{(n+1) \times t}$ 表示由一个新加入的测试样本与原始训练样本组成的矩阵,上标 $n+1$ 表示对应矩阵由原始训练样本与新加入样本共同构造。定义由训练样本 X^n 和 Y^n 构造的关系矩阵为 $\tilde{W}^{[n]}$, 矩阵上标 $[n]$ 表示该矩阵为 $n \times n$ 的方阵,相应的 Laplacian 矩阵 $\tilde{L}_u^{[n]} = I^{[n]} - (\tilde{D}^{[n]})^{-0.5} \tilde{W}^{[n]} (\tilde{D}^{[n]})^{-0.5}$,

$I^{[n]}$ 为单位阵,且有 $\tilde{D}^{[n]} = \text{diag}(\sum_{j=1}^n W_y^{[n]})$ 。根据式(4),经离线学习可获得载入矩阵 \tilde{A}^n 和 \tilde{B}^n 及相应的共享子空间表征 $\tilde{U}^n = [\tilde{u}^1, \dots, \tilde{u}^p] \in \mathbb{R}^{n \times p}$ 。当加入新测试样本 $\{x_{n+1}, y_{n+1}\}$ 之后,具有 $n+1$ 个节点的图关系矩阵则为: $W^{[n+1]} = \begin{bmatrix} W^{[n]} & (W_{n+1, \phi^n})^T \\ W_{n+1, \phi^n} & W_{n+1, n+1} \end{bmatrix}$, 其中 $\phi = [1, \dots, n]$ 为训练样本索引,

相应的 Laplacian 矩阵为:

$$\tilde{L}_u^{[n+1]} = I^{[n+1]} - \begin{bmatrix} (D^{[n]})^{-0.5} W^{[n]} (D^{[n]})^{-0.5} & (D^{[n]})^{-0.5} (W_{n+1, \phi^n})^T (D_{n+1, n+1})^{-0.5} \\ (D_{n+1, n+1})^{-0.5} W_{n+1, \phi^n} (D^{[n]})^{-0.5} & (D_{n+1, n+1})^{-0.5} W_{n+1, n+1} (D_{n+1, n+1})^{-0.5} \end{bmatrix} \quad (7)$$

在共享子空间中的表征,其中 $U^n = [u^1, \dots, u^p] \in \mathbb{R}^{n \times p}$, $u_{n+1} \in \mathbb{R}^p$ 为测试样本 $\{x_{n+1}, y_{n+1}\}$ 在共享子空间的表征。

当训练样本足够大,可以认为新的测试样本的加入不会改变原有训练样本的整体几何结构关系,则有如下图逼近:

$$W^{[n+1]} \approx \begin{bmatrix} \tilde{W}^{[n]} & (W_{n+1, \phi^n})^T \\ W_{n+1, \phi^n} & W_{n+1, n+1} \end{bmatrix} \text{ 及 } D^{[n+1]} = \begin{bmatrix} \tilde{D}^{[n]} & 0 \\ 0 & D_{n+1, n+1} \end{bmatrix}。$$

进一步把近似逼近后的 $W^{[n+1]}$ 与 $D^{[n+1]}$ 代入式(7),可得:

$$\lambda \text{Tr} \left(\begin{bmatrix} \tilde{U}^n \\ u_{n+1}^T \end{bmatrix}^T \begin{bmatrix} L_u^{[n]} & (L_u^{[n]})^T \\ L_u^{[n]} & L_u^{[n+1]} \end{bmatrix} \begin{bmatrix} \tilde{U}^n \\ u_{n+1}^T \end{bmatrix} \right) \quad (10)$$

其中, $L_u^{[n+1]} = - (D_{n+1, n+1})^{-0.5} W_{n+1, \phi^n} (\tilde{D}^{[n]})^{-0.5}$ 。由于 \tilde{U}^n 具有正交性,即 $(\tilde{U}^n)^T \tilde{U}^n = I^{[n]}$, 式(10)中忽略了正交约束 $(\tilde{U}^{n+1})^T \tilde{U}^{n+1} = I^{[n+1]}$ 。定理 3 给出了该最优化问题的解。

定理 3 令式(10)中 $\tilde{U}^n = [\tilde{u}^1, \dots, \tilde{u}^p]$, $\tilde{A}^n = (\tilde{U}^n)^T X^n$ 及 $\tilde{B}^n = (\tilde{U}^n)^T Y^n$ 为通过离线方式根据式(4)得到,对于给定的测试样本 $x_{n+1} \in \mathbb{R}^s$ 与 $y_{n+1} \in \mathbb{R}^t$, 其对应的在共享子空间的表征为 $u_{n+1} \in \mathbb{R}^p$, 则 u_{n+1} 可由如下式(11)给出:

$$u_{n+1} = [(\tilde{U}^n)^T G \tilde{U}^n + \lambda L_u^{[n+1]} I^{[p]}]^{-1} \cdot [(1-\beta)(\tilde{U}^n)^T X^n x_{n+1} + \beta (\tilde{U}^n)^T Y^n y_{n+1} - \lambda (\tilde{U}^n)^T (L_u^{[n]})^T] \quad (11)$$

其中, $G = (1 - \beta)X^n(X^n)^T + \beta Y^n(Y^n)^T$ 。证明略。

为获取给定测试样本在共享子空间的表征 u_{n+1} , 由式(11)可以看出在线扩展算法的计算复杂度为 $O(n^2 \times p)$, 而通过式(8)的在线非扩展算法的计算复杂度为 $O((n+1)^3)$, 一般情况下, 共享子空间的维数 p 远小于样本的数量。当样本量比较大时, 在线扩展算法极大地降低了计算的复杂度。需注意的是, 对于测试样本, 上述复杂度分析没有考虑获得式(11)中的 $W_{n+1, p}^n$ (由 L_n^{21} 给出) 的计算代价。

5 实验分析

5.1 数据库说明

本文采用 UCI 多特征手写体数据库^[13] 分别进行分类和检索实验, UCI 多特征手写数据库包含 0-9 共 10 个数字的多特征数据集, 每类有 200 个样本, 从二值化手写体数字图像中抽取 6 个视角特征: mfeat_fou(76 维, 简称为 fou)、mfeat_fac(216 维, fac)、mfeat_kar(64 维, kar)、mfeat_pix(240 维, pix)、mfeat_zer(47 维, zer)、mfeat_mor(6 维, mor)。

另外, 我们采用视频广告数据库^[14] 中的一个子集进行分类实验来测试 LSPSA 算法的性能。视频广告数据库是一个用于广告检测的数据库, 该数据库来自电视频道搜集的视频, 对其进行镜头分割, 每个镜头被标注为广告和非广告镜头两类, 对每个镜头分别提取 432 维的音频(Audio)和视觉(Visual)特征形成两个视角的描述。在实验中, 我们随机选取了一个包含 3000 个样本的子集, 其中非广告镜头 2100 个, 广告镜头 900 个。

5.2 实验结果分析

表 1 给出了不同算法在 UCI 多特征手写体数据库上的分类性能比较。在该实验中, 任选两个视角数据集分别作为 X 和 Y, 共有 $C_2^6 = 15$ 种组合方式。对于每种组合, 每类随机选择 50% 的样本用于训练, 剩余样本用于测试。实验重复执行 10 次, 取平均结果作为最终衡量指标。此外, 每种算法的参数均由算法性能取得最优时的值确定。从表 1 可以看出, 在 15 种组合方式中, LSPSA 算法与其他算法相比, 除了在组合 8 上比 PCA 性能略差外, 在其他 14 种组合情况下均取得最优性能。与 CCA、MORP 相比, LSPSA 表现出明显的优势。这表明, LSPSA 中引入的局部结构保持约束项使共享子空间在保留原始数据结构的同时, 也保留了更多的判别信息。需说明这里提到的 PCA 指的是把由 X 视角数据与 Y 视角数据串联后再进行降维, 因而本质上并不是共享子空间方法。

表 1 不同算法在 UCI 手写体数据库上的分类性能比较

视角 X	视角 Y	CCA	PCA	PLS	MORP	LSPSA
fac	fou	0.8696	0.9625	0.9633	0.9601	0.9748
fac	kar	0.9602	0.9647	0.9677	0.9548	0.9704
fac	mor	0.7478	0.9029	0.8892	0.944	0.9626
fac	pix	0.9397	0.9707	0.9718	0.9602	0.9722
fac	zer	0.8465	0.9642	0.9474	0.9606	0.9698
fou	kar	0.8948	0.9560	0.9536	0.9448	0.9676
fou	mor	0.7536	0.7211	0.7572	0.7641	0.8112
fou	pix	0.8252	0.9753	0.961	0.9628	0.9744
fou	zer	0.8248	0.8250	0.8048	0.8138	0.8542
kar	mor	0.787	0.8735	0.8704	0.9148	0.9471
kar	pix	0.9580	0.958	0.9650	0.9584	0.9665
kar	zer	0.8998	0.9509	0.9278	0.9426	0.9596
mor	pix	0.7288	0.9031	0.8838	0.9531	0.9651
mor	zer	0.7202	0.6896	0.6736	0.7301	0.8229
pix	zer	0.8198	0.9682	0.9549	0.9614	0.9688

表 2 和表 3 反映了在多特征数据集上, 对于组合 (fac, fou) 和 (fac, kar), 共享子空间的维度对各种算法分类准确率的影响。从表中可以看出, 4 种共享子空间算法的最优子空间维度比较接近。对两个组合来说, 不同共享子空间算法均在 20~30 维左右取得最优性能。

表 2 5 种算法对于组合 (fac, fou) 在不同维度的分类性能比较

维数	CCA	PCA	PLS	MORP	LSPSA
10	0.8511	0.9527	0.9385	0.9471	0.9661
20	0.8696	0.9625	0.9593	0.9601	0.9748
30	0.8605	0.9661	0.9633	0.9472	0.9719
40	0.8517	0.9683	0.963	0.9416	0.9679
50	0.8441	0.9682	0.9615	0.9232	0.9617
60	0.831	0.9695	0.9628	0.9085	0.9477

表 3 5 种算法对于组合 (fac, kar) 在不同维度的分类性能比较

维数	CCA	PCA	PLS	MORP	LSPSA
10	0.9479	0.9487	0.9492	0.9485	0.9551
20	0.9602	0.9579	0.9585	0.9519	0.9681
30	0.9547	0.9643	0.9664	0.9548	0.9704
40	0.9495	0.9647	0.9677	0.9525	0.9702
50	0.945	0.9641	0.9676	0.9475	0.9694
60	0.9388	0.9644	0.9674	0.9389	0.9686

表 4 给出了 5 种算法在视频广告数据库上的分类性能比较。在该实验中, 每类随机选择 30% 的样本用于训练, 剩余样本用于测试。实验重复执行 5 次, 取平均结果作为最终衡量指标。从表 4 可以看出, LSPSA 算法的分类性能明显优于 CCA、PLS、MORP 及 PCA。这也进一步说明 LSPSA 模型中的局部结构保持项使数据的分类性能有了明显提高。

表 4 不同算法在 UCI 手写体数据库上的分类性能比较

视角 X	视角 Y	CCA	PCA	PLS	MORP	LSPSA
Audio	Visual	0.741811	0.9739	0.7784	0.9660	0.9840

此外, 本文还对 LSPSA 与其他共享子空间算法进行了检索性能的测试。图 1 和图 2 分别给出了在特征组合 (fou, kar) 与 (kar, zer) 上检索性能的比较。在该实验中, 每类随机选择 90% 的样本用于构造检索数据库, 剩余样本用作查询样例, 其它设置与分类实验相同。从图 1 和图 2 中可以看出, LSPSA 的检索性能明显优于其他算法, 而且随着返回列表范围的增大, 本文提出的 LSPSA 的优势体现得更加明显。与 MORP 算法的性能比较表明, 本文提出的引入保持局部结构的图约束项是必要的。

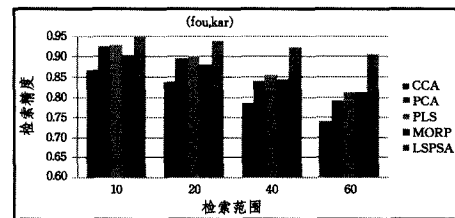


图 1 组合 (fou, kar) 的检索性能

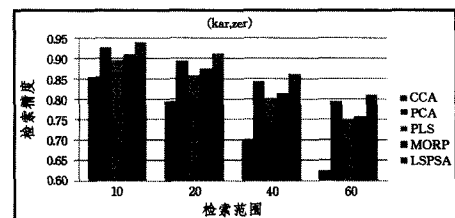


图 2 组合 (kar, zer) 的检索性能

如前所述, 针对新测试样本 $\{x_{n+1}, y_{n+1}\}$, 在获取其相应

的共享子空间表征 u_{n+1} 的计算复杂度方面,本文提出的 LSPSA 的在线扩展的复杂度由 $O((n+1)^3)$ 降为 $O(n^2 \times p)$ 。这里,我们进一步通过检索实验验证 LSPSA 的在线扩展的有效性。图 3 给出了多特征组合为 (fac, fou) 且检索数据规模为 2000 时的检索性能。从中可以看出,随着返回列表范围的变化,在线扩展与在线非扩展在检索性能方面取得了非常相似的检索结果。但在检索效率方面,图 4 表明在线扩展的效率是在线非扩展效率的近 2100 倍。从图 4 中也可看出,随着样本量增大,在线扩展算法的计算效率优势将会更加明显。

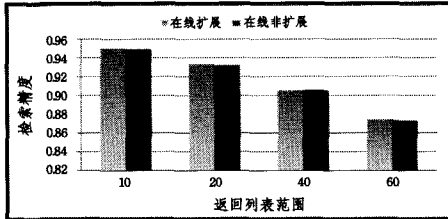


图 3 在线扩展和在线非扩展情况下检索性能比较

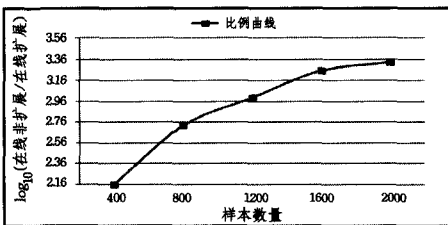


图 4 在线扩展和在线非扩展情况下检索效率比较

结束语 针对多视角分析问题,通过在 MORP 模型的基础上引入图约束,本文提出了一种局部结构保持的共享子空间分析(LSPSA)算法,它在把多视角数据映射到一个共享子空间的同时,保持原始多视角特征空间与共享子空间中的数据具有相近的局部几何结构关系,从而避免多视角数据在共享子空间的过拟合问题。此外,针对新测试样本的在线非扩展问题,本文还提出了一个图模型逼近的在线扩展方法,使得获取新测试样本在共享子空间表征的复杂度大大降低。

参 考 文 献

[1] Estellers V, Gurban M, Thiran J. On Dynamic Stream Weighting for Audio-Visual Speech Recognition [J]. IEEE Transactions on

(上接第 56 页)

[10] Bay H, Tuytelaars T, Van Gool L. Surf: Speeded up robust features[M]. Computer Vision-ECCV 2006. Springer Berlin Heidelberg, 2006: 404-417

[11] Rosten E, Drummond T. Machine learning for high-speed corner detection[M] // Computer Vision-ECCV 2006. Springer Berlin Heidelberg, 2006: 430-443

[12] Calonder M, Lepetit V, Strecha C, et al. BRIEF: binary robust independent elementary features[M] // Computer Vision-ECCV 2010. Springer Berlin Heidelberg, 2010: 778-792

[13] Leutenegger S, Chli M, Siegwart R Y. BRISK: Binary robust invariant scalable keypoints[C] // 2011 IEEE International Conference on Computer Vision (ICCV). IEEE, 2011: 2548-2555

[14] Rublee E, Rabaud V, Konolige K, et al. ORB: an efficient alter-

Audio, Speech, and Language Processing, 2012, 20 (4): 1145-1157

[2] Li Hao-jie, Wang Xiao-hui, Tang Jin-hui. Combining global and local matching of multiple features for precise item image retrieval[J]. ACM/Springer Multimedia System Journal, 2013, 19 (1): 37-49

[3] Kalamaras I, Mademlis A, Malassiotis S, et al. novel framework for retrieval and interactive visualization of multimodal data[J]. Electronic Letters on Computer Vision and Image Analysis, 2013, 12(2): 28-39

[4] Hotelling H. Relations between two sets of variates [J]. Biometrika, 1936, 28(3/4): 321-377

[5] 王惠文. 偏最小二乘回归的线性与非线性方法[M]. 北京: 国防工业出版社, 1999

[6] Akaho S. A kernel method for canonical correlation analysis [C] // IMPS 2001 International Meeting of Psychometric Society, Feb. 2007

[7] 彭岩, 张道强. 半监督典型相关分析算法[J]. 软件学报, 2008, 19 (11): 2822-2832

[8] Rosipal R, Trejo L J. Kernel partial least squares regression in reproducing kernel Hilbert space[J]. Journal of Machine Learning Research, 2001(2): 97-123

[9] 王珏, 周志华, 周傲英. 机器学习及其应用[M]. 北京: 清华大学出版社, 2006

[10] Sharma A, Kumar A, Daume H III, et al. Generalized MultiviewAnalysis: A Discriminative Latent Space[C] // IEEE Conference on Computer Vision and Pattern Recognition, 2012: 2160-2167

[11] Chen Yao-nan, Lin H T. Feature-aware label space dimension reduction for multi-label classification[C] // Advances in Neural Information Processing Systems, 2012, 25: 1538-1546

[12] Yu Shi-peng, Yu Kai, Tresp V, et al. Multi-Output regularized feature projection[J]. IEEE Conference on knowledge and data engineering, 2006, 18(12): 1600-1613

[13] <http://www.ics.uci.edu/~mllearn/MLSummary.html>

[14] Liu Nan, Zhao Yao, Zhu Zhen-feng, et al. Exploiting Visual-Audio-Textual Characteristics for Automatic TV Commercial Block Detection and Segmentation[J]. IEEE Transaction on Multimedia, 2011, 13(5): 961-973

native to SIFT or SURF[C] // 2011 IEEE International Conference on Computer Vision (ICCV). IEEE, 2011: 2564-2571

[15] Quinlan J R. Induction of decision trees[J]. Machine learning, 1986, 1(1): 81-106

[16] 邓黎, 王春鸿, 饶长辉. 基于二进制稳健独立基元特征的扩展目标快速检测[J]. 中国激光, 2012, 39(B06): 327-331

[17] 郭莉莎, 等. 基于多尺度 FAST-9 的图像快速匹配算法[J]. 计算机工程, 2012, 38(12)

[18] <http://www.ee.columbia.edu/ln/dvmm/researchProjects/FeatureExtraction/NearDuplicateByParts/INDDetection.html>, 2004

[19] Tan H K, Ngo C W, Chua T S. Efficient mining of multiple partial near-duplicate alignments by temporal network[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2010, 20(11): 1486-1498