

基于加权高阶奇异值分解的支持张量机图像分类

刘亚楠^{1,2}, 涂铮铮¹, 罗 斌¹

(1 安徽大学 计算机科学与技术学院, 安徽 合肥 230039; 2 合肥师范学院 计算机科学与技术系, 安徽 合肥 230601)

摘 要: 为了有效提高图像分类的准确率, 充分利用图像本身的结构信息并压缩图像数据, 首先构造三阶图像特征张量, 利用非负矩阵分解(NMF)在张量子空间降维, 提出了一种基于二维主成分分析(2DPCA)来得到 NMF 初始点的方法, 保证了图像信息的有效利用. 然后, 为了保持降维后的张量子空间所在的流形空间的本征结构, 根据图像类标构造权值矩阵, 并把图像集合构造造成四阶张量实现图像的分类. 通过对两个图像数据库的实验, 表明该方法能有效提升图像分类的准确率.

关键词: 高阶奇异值分解; 非负矩阵分解; 支持张量机; 二维主成分分析

中图分类号: TP391

文献标识码: A

文章编号: 1000-7180(2014)05-0028-04

Support Tensor Machine Image Classification Based on Weighted High-order Singular Value Decomposition

LIU Ya-nan^{1,2}, TU Zheng-zheng¹, LUO Bin¹

(1 School of Computer Science and Technology, Anhui University, Hefei 230039, China;

2 Department of Computer Science and Technology, Hefei Normal University, Hefei 230601, China)

Abstract: To improve the accuracy of image classification, fully use the structural information of the data, and compress the image data, first, third order tensor image features are constructed. Then, non-negative matrix factorization (NMF) is used for dimension reduction. A method of choosing a good starting point is proposed using two-dimensional principal component analysis (2DPCA), which uses the information of the image effectively. Next, in order to maintain the intrinsic structure of the manifold for tensor subspace, weighted matrix is derived according to the labels of images. Meanwhile the set of images is used to construct a fourth order tensor. A method for classification is proposed by using weighted high-order singular value decomposition for support tensor machine. The experimental results on two image databases show that the proposed can effectively improve the accuracy of image classification.

Key words: high-order singular value decomposition; non-negative matrix factorization; support tensor machine; two-dimensional principal component analysis

1 引言

图像分类在图像检索、数据挖掘和模式识别等领域有着广泛的应用价值. 图像数据本身在计算机中往往以矩阵形式存在, 但是传统的用于机器学习算法的图像表达一般采用向量模型.

传统的向量形式或者矩阵形式都可以称作是一

阶张量或者二阶张量^[1], 而 2007 年 Tao^[2] 提出了一个统一的监督张量学习(STL)框架, 可以输入任意维数的图像数据特征, 将传统支持向量机(SVM)扩展到支持张量机(STM). 在图像表达上选用张量模型, 有效的利用了图像的内在几何结构信息, 大幅提升了遥感图像分类、视频语义概念检测等的性能^[3]. 同时, 借助张量的概念, He 和 Cai^[4] 提出了张量子

收稿日期: 2013-06-29; 修回日期: 2013-08-23

基金项目: 国家自然科学基金项目(61073116); 高校省级优秀青年人才基金重点项目(2011SQRL129ZD)

空间分析(TSA)算法,Luo^[5]等提出了判别高阶奇异值分解(DHOSVD),自适应的选择张量子空间.传统的方法不仅在寻找低维张量子空间的过程中没有考虑到数据的非负性,而且在对图像进行处理的过程中仍然选用向量模型,不能有效利用张量运算的优势.

本文在 STL 框架下,提出了一种基于加权 HOSVD 的支持张量机图像分类方法.首先搭建了一套支持张量机图像分类框架,然后针对该系统中的若干问题进行了研究,在利用图像结构信息的同时运用了张量运算.

2 张量几何相关运算

算法中用到的张量几何^[1]的相关定义和定理如下:

定义 1 (张量矩阵展开) 是将一个张量中的元素重新排列,得到一个矩阵的过程,张量 $\mathbf{X} \in R^{I_1 \times I_2 \times \dots \times I_N}$ 的 n 模展开矩阵表示为 $\mathbf{X}_{(n)} \in R^{I_n \times (I_1 \times \dots \times I_{n-1} \times I_{n+1} \times \dots \times I_N)}$.

定义 2 (张量乘法) 张量与矩阵的乘法由 n 模乘积所定义.具体地,张量 $\mathbf{X} \in R^{I_1 \times I_2 \times \dots \times I_N}$ 和矩阵 $\mathbf{U} \in R^{J_n \times I_n}$ 的 n 模乘积表示为 $\mathbf{X} \times_n \mathbf{U}$,这是一个 $I_1 \times I_2 \times \dots \times I_{n-1} \times J_n \times I_{n+1} \times \dots \times I_N$ 阶张量,定义如下:

$$(\mathbf{X} \times_n \mathbf{U})_{i_1 i_2 \dots i_n \dots i_N} = \sum_{i_n} a_{i_1 i_2 \dots i_n \dots i_N} \cdot u_{j_n i_n} \quad (1)$$

定理 1(HOSVD) 每个张量 $\mathbf{X} \in R^{I_1 \times I_2 \times \dots \times I_N}$ 均可唯一分解为 $\mathbf{X} = \mathbf{T} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \times \dots \times_N \mathbf{U}_N$,使得:
(1) $\mathbf{U}_{(n)}$ 均为 $I_n \times I_n$ 维正交矩阵;(2) $\mathbf{T}_{(n)}$ 的行都是正交的;(3) 对于任意 n , 都有 $\|\mathbf{T}_{i_n=1}\| \geq \|\mathbf{T}_{i_n=2}\| \geq \dots \geq \|\mathbf{T}_{i_n=N}\| \geq 0$.

张量 $\mathbf{T}$ 也称为核心张量,对于张量 $\mathbf{X}$,其 HOSVD 如下:

- (1) 对于任意 n , $\mathbf{U}_{(n)}$ 是张量 $\mathbf{X}$ 的 n 模展开矩阵 $\mathbf{X}_{(n)}$ 进行 SVD 分解得到的左矩阵;
- (2) 计算 $\mathbf{T} = \mathbf{X} \times_1 \mathbf{U}_1^T \times_2 \mathbf{U}_2^T \times \dots \times_N \mathbf{U}_N^T$.

3 降维方法

3.1 非负矩阵分解

定义 3 NMF 是发现非负的 $m \times r$ 的基矩阵 $\mathbf{W}$ 和 $r \times n$ 的系数矩阵 $\mathbf{H}$ 使^[6]

$$\mathbf{V} \approx \mathbf{WH} \quad (2)$$

通常设定 $r \ll \frac{m \times n}{m+n}$, 所以只有在 $\mathbf{W}$ 包含了随机变量 v 的本质特征时,才可能使 $\mathbf{V} \approx \mathbf{WH}$, 即可

以用很少量的数据描述很大的数据.

一般情况下,随机产生两个正的矩阵作为 NMF 迭代公式中 $\mathbf{W}$ 和 $\mathbf{H}$ 的初值,然后通过 KL 分散度的迭代公式计算出 $\mathbf{W}$ 和 $\mathbf{H}$. 但是这样得到的基矩阵 $\mathbf{W}$ 就是一个不确定的值,将会直接影响降维结果,进而影响分类结果.为了避免这种随机不确定性,并保持图像空间的结构信息和空间信息,本文采用 2DPCA 得到 $\mathbf{W}$ 和 $\mathbf{H}$ 的初值.

3.2 张量子空间降维

PCA 要求数据为向量形式.对于图像数据,首先需要把图像进行向量化才能进行 PCA 压缩.然而,图像被向量化后,维数通常会很高,而且方差矩阵是奇异的,计算上比较困难且特征向量计算不精确. Yang^[7] 提出利用图像矩阵直接投影,形成 2DPCA 算法.

具体地,设有一幅图像 I , 大小为 $m \times n$. 寻求一个 n 维单位向量,用线性变换

$$v = It \quad (3)$$

把 I 投影到 t 上,使得投影后得到的 x 的各分量方差和达到最大. 寻求 $n \times r$ 列单位正交矩阵 $\mathbf{T} = (t_1, t_2, \dots, t_r)$, 使得 $\mathbf{V} = \mathbf{IT}$ 的各列互不相关且各列的方差和达到最大.

r 个主轴 t_j , ($j = 1, 2, \dots, r$), 被用来对观测图像进行特征提取. 对一幅给定的观测图像 I , 由

$$v_j = It_j, (j = 1, 2, \dots, r) \quad (4)$$

得到一组投影向量 $v_1, v_2, \dots, v_r$ 称为图像 I 的主成分. r 个主成分构成了维数为 $m \times r$ 的矩阵 $\mathbf{V} = (v_1, v_2, \dots, v_r)$, 这样,把得到的 $\mathbf{V}$ 作为 NMF 中 $\mathbf{W}$ 的初值,把 $\mathbf{T}$ 的转置作为 $\mathbf{H}$ 的初值. 算法 1 给出了基于 2DPCA 的 NMF 张量子空间降维算法.

算法 1 张量子空间降维算法

- (1) 对图像提取特征,得到三阶特征张量集合 $\mathbf{X} = \{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N\}$, $\mathbf{X}_i \in R^{I_1 \times I_2 \times I_3}$, ($i = 1, 2, \dots, N$)
- (2) 得到 $\mathbf{X}_i$ 的 n 模展开矩阵 $\mathbf{X}_{(n)}$, ($n = 1, 2, 3$)
- (3) 利用 2DPCA 得到初始矩阵 $\mathbf{W}$ 和 $\mathbf{H}$, 对 $\mathbf{X}_{(n)}$, ($n = 1, 2, 3$), 进行 NMF 操作,得到基矩阵 $\mathbf{W}_{(n)}$, ($n = 1, 2, 3$)
- (4) 根据 $\mathbf{Y}_i = \mathbf{X}_i \times_1 \mathbf{W}_1 \times_2 \mathbf{W}_2 \times_3 \mathbf{W}_3$, ($i = 1, 2, \dots, N$) 可得到降维后的数据集 $\mathbf{Y} = \{\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_N\}$, ($i = 1, 2, \dots, N$), $\mathbf{Y}_i \in R^{J_1 \times J_2 \times J_3}$, ($J_1 < I_1, J_2 < I_2, J_3 < I_3$)

4 图像分类方法

给定张量训练样本集 $\mathbf{X} = \{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N\}$, (i

$= 1, 2, \dots, N$), 其中 $\mathbf{X}_i \in R^{I_1 \times I_2 \times \dots \times I_M}$ (M 为张量的阶数, 本文 $M = 3$) 以及所属类别标记 $y_i \in \{+1, -1\}$, 将正负两类数据分开所需的张量分类平面如下:

$$y(\mathbf{X}) = \mathbf{X} \prod_{k=1}^M \times_k \omega_k + b \quad (5)$$

为了确定该分类平面, 通过最优化式(6) 得到 ω_k 和 b 的值.

$$\begin{aligned} \min_{\omega_k |_{k=1}^M, b, \xi} \quad & \frac{1}{2} \left\| \bigotimes_{k=1}^M \omega_k \right\|_{Fro}^2 + c \sum_{i=1}^N \xi_i \\ \text{s. t.} \quad & y_i \left[\mathbf{X}_i \prod_{k=1}^M \times_k \omega_k + b \right] \geq 1 - \xi_i, 1 \leq i \leq N \\ & \xi_i \geq 0 \end{aligned} \quad (6)$$

其中, $\xi = [\xi_1, \xi_2, \dots, \xi_N]^T \in R^N$ 为松弛变量, $\omega_k \in R^{I_k} (1 \leq k \leq M)$, 判别函数为 $y(\mathbf{X}) = \text{sign}(\mathbf{X} \prod_{k=1}^M \times_k \omega_k + b)$.

$$\begin{aligned} L(\omega_k |_{k=1}^M, b, \xi, \alpha, \kappa) = & \frac{1}{2} \left\| \bigotimes_{k=1}^M \omega_k \right\|_{Fro}^2 + \\ & c \sum_{i=1}^N \xi_i - \sum_{i=1}^N \alpha_i (y_i [\mathbf{X}_i \prod_{k=1}^M \times_k \omega_k + b] - 1 + \xi_i) - \\ & \sum_{i=1}^N \kappa_i \xi_i = \frac{1}{2} \prod_{k=1}^M \omega_k^T \omega_k + c \sum_{i=1}^N \xi_i - \\ & \sum_{i=1}^N \alpha_i y_i [\mathbf{X}_i \prod_{k=1}^M \times_k \omega_k] - b \alpha^T \mathbf{y} + \sum_{i=1}^N \alpha_i - \alpha^T \xi - \kappa^T \xi \end{aligned} \quad (7)$$

为了解决该最优化问题式(6), 构造如式(7) 拉格朗日函数, α_i, κ_i 为拉格朗日因子. 对各变量求偏导可得

$$\begin{aligned} \partial_{\omega_j} L = 0 \Rightarrow \omega_j &= \frac{1}{\prod_{k=1, k \neq j}^M \omega_k^T \omega_k} \sum_{i=1}^N \alpha_i y_i (\mathbf{X}_i \bar{\times}_j \omega_j) \\ \partial_b L = 0 \Rightarrow \alpha^T \mathbf{y} &= 0 \\ \partial_{\xi} L = 0 \Rightarrow c - \alpha - \kappa &= 0 \end{aligned} \quad (8)$$

同时, 式(7) 的对偶问题如下:

$$\max_{\alpha, \kappa} \min_{\omega_k |_{k=1}^M, b, \xi} L(\omega_k |_{k=1}^M, b, \xi, \alpha, \kappa) \quad (9)$$

把式(8) 代入式(7) 可得如下最优化问题:

$$\begin{aligned} \max \quad & \frac{1}{2} \omega_j^T \sum_{i=1}^N \alpha_i y_i (\mathbf{X}_i \bar{\times}_j \omega_j) - \\ & \sum_{i=1}^N \alpha_i y_i [\mathbf{X}_i \prod_{k=1}^M \times_k \omega_k] + \sum_{i=1}^N \alpha_i \\ \text{s. t.} \quad & \alpha^T \mathbf{y} = 0 \\ & 0 \leq \alpha \leq c \end{aligned} \quad (10)$$

可以看出, 这是一个线性规划问题.

由式(8) 可知, 求解 ω_j 必须首先已知 α_i 和 $\omega_k |_{k=1}^M$, 然后使用交替投影方法^[2] 求解. 交替投影方法

是随机给定 α_i 和 $\omega_k |_{k=1}^M$ 的初值, 这样直接影响到支持张量机的分类性能. 为了避免这种随机性对分类性能的影响, 构建最近邻图 G 模拟张量数据集 $\mathbf{X}$ 所在的流形空间的局部几何结构, 定义 G 的权重矩阵 $\mathbf{W}$ 如下:

$$W_{ij} = \begin{cases} 1, & \text{当 } \mathbf{X}_i \text{ 和 } \mathbf{X}_j \text{ 为同一类别} \\ 0, & \text{其他} \end{cases} \quad (11)$$

利用该 $\mathbf{W}$ 与张量数据集 $\mathbf{X}$ 做加权 HOSVD, 以得到 α_i 和 $\omega_k |_{k=1}^M$ 的初值. 具体算法思路见算法 2. 随后, 利用加权 HOSVD 得到 α_i 和 $\omega_k |_{k=1}^M$ 的初值, 进而进行支持张量机分类, 具体算法思路见算法 3.

算法 2 加权 HOSVD

(1) 针对数据集 $\mathbf{X}, \mathbf{X} \in R^{I_1 \times I_2 \times I_3 \times N}$ 构造 4 阶张量;

(2) 求得 $\mathbf{X} \in R^{I_1 \times I_2 \times I_3 \times N}$ 与 $\mathbf{W} \in R^{N \times N}$ 的 4 模乘积 $\mathbf{X} \times_4 \mathbf{W}$, 记为 $\mathbf{A}, \mathbf{A} \in R^{I_1 \times I_2 \times I_3 \times N}$;

(3) 对 $\mathbf{A}$ 进行 n 模展开, 得到 $\mathbf{A}_{(n)}, (n = 1, 2, 3, 4)$;

(4) 对 $\mathbf{A}_{(n)}$ 分别进行 SVD 分解得到奇异值向量 $\mathbf{s}_{(n)}, (n = 1, 2, 3, 4)$.

算法 3 基于加权 HOSVD 的支持张量机分类算法

记数据集 $\mathbf{X}$, 采用张量子空间降维后的数据为 $\mathbf{Y}$, 通过算法 2 对数据集 $\mathbf{Y}$ 进行操作, 得到奇异值向量 $\mathbf{s}_{(n)}, (n = 1, 2, 3, 4)$, 依次设为 $\omega_k |_{k=1}^3$ 和 α 的初始值, 然后利用交替投影算法进行迭代优化直到 α_i 和 ω_k 都收敛, 完成分类器训练过程.

5 实验结果

为了验证所提算法在图像分类方面的有效性, 采用 Caltech 101 数据库和 PPMI^[8] 数据库进行测试, PPMI 数据库共有 7 类 (bassoon, erhu, flute, violin, frenchhorn, guitar, saxophone) 图像数据. Caltech 101 数据库共 101 类, 从中随机选取 10 种图像类别 (euphonium, faces_easy, leopards, motor-bikes, accordion, cellphone, car_side, dollar_bill, electric_guitar, soccer_ball) 来测试. 一般地为了在分类器性能与运行效率之间寻求平衡, 以正负标记 1:4 或 1:5 的比例来选取正负样本, 选用的正负比例是 1:5.

为了说明加权 HOSVD (WHOSVD) 和 STM 结合的有效性, 选用 5 种不同的分类方法对这两个数据库进行分类准确率的比较, 如图 1、2 所示, 图 1、2 中横坐标为图像类别, 纵坐标为分类准确率.

从图1、2中可以看出对于不同的3种降维方法,NMF+2DPCA+STM的分类准确率最高,其次为2DPCA+STM,NMF+STM的分类准确率最低.同时,本文方法比其他4种方法的分类结果都要高,这说明2DPCA+NMF降维方法与WHOSVD+STM分类方法的结合可以得到更好的分类结果,另外,对于大部分类别数据,如 faces_easy(图1)、saxophone(图2)等,所提方法与2DPCA+NMF+STM方法相比分类准确率有明显的提高,说明所提方法采用WHOSVD初始化STM是可行的,表明了所提方法的优势.

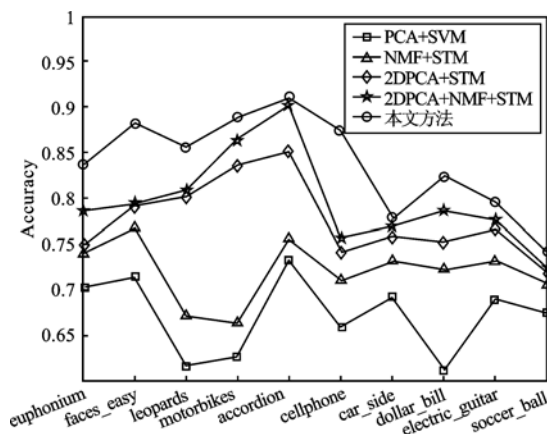


图1 Caltech 101 分类准确率结果比较

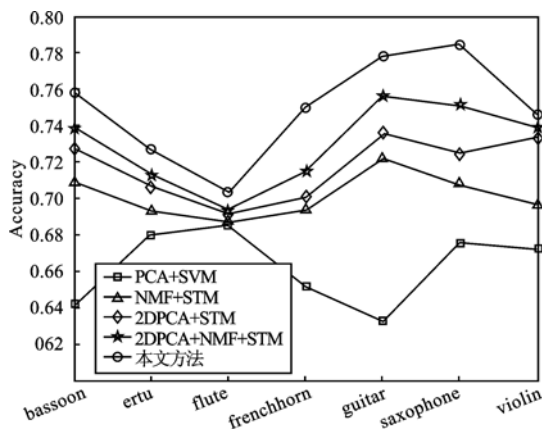


图2 PPMI 分类准确率结果比较

6 结束语

本文研究加权HOSVD支持张量机图像分类方法,在张量子空间中对图像进行分类.提出一种基于2DPCA的NMF方法,为NMF提供了一个好的

初始点,进而提出一种基于NMF的张量子空间降维方法,得到非负降维数据,以保证图像数据的整体直观性.实验结果证明本文提出的降维方法和分类方法相结合可以有效改善图像分类的准确性.为了进一步提高分类的精度,下一步将尝试结合图谱理论构造支持张量机图像分类方法.

参考文献:

- [1] Lathauwer L D. Signal processing based on multilinear algebra [D]. Belgium: Katholieke Universiteit Leuven, 1997.
- [2] Tao D C, Li X L, Wu X D, et al. Supervised tensor learning [J]. Knowledge and Information Systems, 2007, 13(1):1-42.
- [3] 张乐飞,黄昕,张良培. 高分辨率遥感影像的支持张量机分类方法[J]. 武汉大学学报. 信息科学版, 2012, 37(3):314-317.
- [4] He X F, Cai D, Niyogi P. Tensor subspace analysis [C]// Proceedings of the Advances in Neural Information Processing Systems. New York: NIPS, 2005: 499-506.
- [5] Dijun Luo, Heng Huang, Chris Ding. Discriminative high order SVD: adaptive tensor subspace selection for image classification, clustering, and retrieval [C]// 2011 International Conference on Computer Vision. Barcelona: ICCV, 2011:1443-1448.
- [6] Lee D D, Seung H S. learning the parts of objects by non-negative matrix factorization [J]. Nature, 1999, 401(6755):788-791.
- [7] Yang J, Zhang D, Frangi A F, et al. Two-dimensional PCA: a new approach to appearance-based face representation and recognition [J]. IEEE Trans Pattern Analysis and Machine Intelligence, 2004, 26(1): 131-137.
- [8] Yao B, Fei L. Grouplet: a structured image representation for recognizing human and object interactions [C]// 2010 IEEE Conference on Computer Vision and Pattern Recognition. San Francisco: CVPR, 2010:9-16.

作者简介:

刘亚楠 女,(1984—),博士,讲师.研究方向为模式识别与图像处理.