

文章编号: 2095-2163(2021)01-0051-04

中图分类号: TP391

文献标志码: A

一种快速低秩的判别子字典学习算法及图像分类

赵雅, 王顺政, 吕文涛, 王成群
(浙江理工大学信息学院, 杭州 310018)

摘要: 本文提出了一种快速低秩的判别子字典学习算法。在训练阶段, 构造一个子字典的低秩约束项和拉普拉斯矩阵正则化项, 加入判别字典学习的目标函数中。将原始样本映射到一个新的空间中, 使同一类别的相邻点彼此靠近, 同时增强子字典对同类样本的重构能力, 针对每类样本的判别性特征, 学习出相应的学习字典。在测试阶段, 利用 kNN 分类器估计测试样本的类别标签。同时, 将算法应用在 3 种数据集上, 与其他的字典学习算法进行比较, 取得了较好的分类结果。
关键词: 子字典; 判别字典; 拉普拉斯矩阵; 图像分类

A fast and low-rank discriminant sub-dictionary learning method for image classification

ZHAO Ya, WANG Shunzheng, LV Wentao, WANG Chengqun

(School of Informatics Science and Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China)

[Abstract] This paper proposes a fast, low-rank discriminative sub-dictionary learning algorithm. In the training phase, the low-rank constraint terms of the sub-dictionary and the Laplacian matrix regularization terms are constructed, and the objective function of the discriminative dictionary learning is added. The original sample is mapped to the new space so that adjacent points of the same category are closed to each other. At the same time, the sub-dictionary's ability is enhanced to reconstruct similar samples, and the corresponding learning dictionary is learnt according to the discriminative characteristics of each sample. In the testing phase, the kNN classifier is used to estimate the class label of the test sample. Finally, the algorithm are applied to three public data sets compare with other dictionary learning algorithms. The proposed algorithm has achieved better classification results.

[Key words] sub-dictionary; discriminant dictionary; Laplacian matrix; image classification

0 引言

判别字典学习是稀疏表示问题的一个研究分支, 主要是通过重构训练样本得到样本的学习字典, 并通过构造不同的约束项模型来增加字典的判别性能。字典学习包括无监督字典学习和有监督字典学习。其中, 无监督字典学习主要是通过所有训练信号重建并优化字典, 而不给出任何标签信息。典型的无监督字典学习有 KSVD 算法^[1], MOD 算法^[2]等。Zheng 等人^[3]给出了使用拉普拉斯算子的无监督字典的图形正则化稀疏编码, 并验证了其在分类和聚类上的有效性。但该类算法无法有效利用样本的标签信息, 分类性能不一定是最佳的。相应地, 有监督字典学习根据训练样本的标签信息学习出判别性字典。例如文献[4-7], 就是典型的有监督学习字典。文献[5]提出了一个 LSDDL 算法, 针对样本

的局部特征和几何结构, 结合样本的标签信息进行字典学习。但该算法容易受到训练样本个数的影响, 且该方法仅有一个被所有类别共享的字典, 会使字典元素和类别间产生误差。

针对这种情况, 研究提出了一种快速低秩的判别子字典学习算法。在判别字典学习的目标函数的基础上, 构造一个子字典的低秩约束项和拉普拉斯矩阵正则化项。将原始样本映射到一个新的空间中, 使同一类别的相邻点彼此靠近, 同时增强子字典对同类样本的重构能力, 针对每类样本的判别性特征, 学习出相应的学习字典。然后, 利用现有的分类器估计测试样本的类别标签, 得到算法的分类准确率。本文在 3 种数据集上的实验结果证明了该方法的有效性。

基金项目: 国家自然科学基金(61601410, U1709219); 浙江省重点研发项目(2018C01093); 浙江理工大学基础研究基金(2019Q040)。

作者简介: 赵雅(1996-), 女, 硕士研究生, 主要研究方向: 信号与信息处理; 王顺政(1996-), 男, 硕士研究生, 主要研究方向: 电子与通信工程; 吕文涛(1982-), 男, 博士, 副教授, 主要研究方向: 稀疏表示、机器学习; 王成群(1981-), 男, 博士, 讲师, 主要研究方向: 稀疏表示、图像处理。

通讯作者: 吕文涛 Email: 21447476@qq.com

收稿日期: 2020-11-06

1 基本理论

1.1 字典学习

设 $\mathbf{D}_C = [d_1, \dots, d_K] \in R^{n \times K}$ 表示第 C 类样本的学习字典。针对样本集, 令 $\mathbf{Z}_C = [Z_1^C, \dots, Z_{N^C}^C] \in R^{n \times N^C}$ 为第 C 类训练样本, 其中 $Z_i^C (i=1, \dots, N^C)$ 表示 C 类样本中具有 n 维维数的第 i 个图像样本。同时, 令 $\mathbf{V}_C = [V_1^C, \dots, V_{N^C}^C] \in R^{K \times N^C}$ 表示为 \mathbf{Z}_C 在 \mathbf{D}_C 上的稀疏表示矩阵。可以得到传统的字典学习模型:

$$\min_{\mathbf{D}_C, \mathbf{V}_C} \|\mathbf{Z}_C - \mathbf{D}_C \mathbf{V}_C\|_F^2 + \lambda \|\mathbf{V}_C\|_p \quad (1)$$

$$\text{s. t. } \|d_i\| = 1, i = 1, \dots, K.$$

模型第一项为训练样本的误差重构项, 第二项是稀疏正则项, $\lambda > 0$ 为约束系数。 $\|\cdot\|_p$ 为 l_p 范数, 当 $p=1$, 表示 lasso 稀疏模型; 当 $p=2$, 表示岭回归。由于公式(1)中的 \mathbf{D}_C 或 \mathbf{V}_C 都可以单独得到相关的凸函数, 但不能同时得到两者的凸函数, 因此通常保持一个变量不变, 交替求解另一个变量, 通过迭代求得最优解。

1.2 拉普拉斯矩阵

定义一个顶点数量为 N 的无向图 G 的拉普拉斯矩阵 \mathbf{L} 为:

$$\mathbf{L} = \mathbf{M} - \mathbf{W}. \quad (2)$$

其中, \mathbf{W} 为图 G 的邻接矩阵, 维度是 $N \times N$, 用于表示节点之间的邻接关系。将邻接矩阵 \mathbf{W} 的每一列元素加起来得到 N 个数, 再以其为对角线构造一个维度是 $N \times N$ 的对角矩阵 \mathbf{M} , 该对角矩阵 \mathbf{M} 则被称为图的度矩阵。由公式(2)可以看出, 拉普拉斯矩阵 \mathbf{L} 是一个对称的半正定矩阵, 即对于任何向量 \mathbf{V} , \mathbf{L} 都能满足 $\mathbf{V} \mathbf{L} \mathbf{V}' \geq 0$ 。

2 本文算法

2.1 算法流程

为了设计一个具有判别性的字典, 令来自同一类别的样本可以具备更为相近的稀疏表示, 提出了一种基于子字典的判别字典学习算法。算法的整体流程如图1所示。

2.2 训练过程

基于上述字典学习算法, 本文提出以下判别字典学习算法模型:

$$\begin{aligned} (\mathbf{D}_C, \mathbf{V}_C) = \operatorname{argmin}_{\mathbf{D}_C, \mathbf{V}_C} & \|\mathbf{Z}_C - \mathbf{D}_C \mathbf{V}_C\|_F^2 + \\ & \lambda \sum_{i=1}^{N^C} \|\mathbf{V}_i^C\|_2^2 + \beta \|\mathbf{D}_C\| + \\ & \gamma \operatorname{tr}(\mathbf{V}_C \mathbf{L} (\mathbf{V}_C)^T), \end{aligned} \quad (3)$$

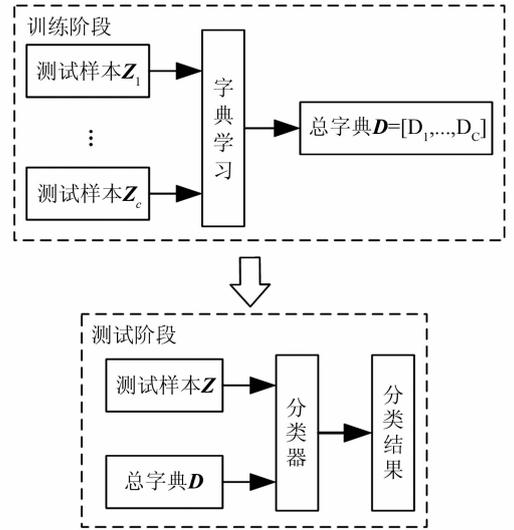


图1 算法流程图

Fig. 1 Algorithm flowchart

其中, 前两项为字典学习的基础模型, 后两项分别为子字典的低秩约束项和拉普拉斯矩阵正则化项。 $\beta > 0$ 和 $\gamma > 0$ 为相应的稀疏系数。

由于目标函数现在是非凸的, 因此通过迭代更新字典和稀疏表示得到学习字典。首先固定字典, 更新稀疏表示矩阵。即第 C 类样本的稀疏表示 \mathbf{V}_C 有:

$$\begin{aligned} \mathbf{V}_C^* = \operatorname{argmin}_{\mathbf{V}_C} & \|\mathbf{Z}_C - \mathbf{D}_C \mathbf{V}_C\|_F^2 + \\ & \lambda \sum_{i=1}^{N^C} \|\mathbf{V}_i^C\|_2^2 + \beta \|\mathbf{D}_C\| + \gamma \operatorname{tr}(\mathbf{V}_C \mathbf{L} (\mathbf{V}_C)^T), \end{aligned} \quad (4)$$

其中, 拉普拉斯矩阵正则化项可以展开成:

$$\operatorname{tr}(\mathbf{V}_C \mathbf{L} (\mathbf{V}_C)^T) = \sum_{i,j} L_C(i,j) (\mathbf{V}_j^C)^T \mathbf{V}_i^C, \quad (5)$$

对第 C 类样本逐个更新稀疏表示矩阵, 即:

$$\begin{aligned} \mathbf{V}_i^{C*} = \operatorname{argmin}_{\mathbf{V}_i^C} & \|\mathbf{Z}_i^C - \mathbf{D}_C \mathbf{V}_i^C\|_F^2 + \lambda \|\mathbf{V}_i^C\|_2^2 + \\ & \beta \|\mathbf{D}_C\| + \gamma \sum_{k,j} L_C(k,j) (\mathbf{V}_j^C)^T \mathbf{V}_k^C, \end{aligned} \quad (6)$$

令 $G(\mathbf{V}_i^C)$ 代表式(6), 对其求偏导可得:

$$\frac{\partial [G(\mathbf{V}_i^C)]}{\partial \mathbf{V}_i^C} = 2\mathbf{D}_C^T (\mathbf{D}_C \mathbf{V}_i^C - \mathbf{Z}_i^C) + 2\lambda \mathbf{V}_i^C +$$

$$2\gamma \mathbf{V}_C L_C, \quad (7)$$

$$\frac{\partial^2 [G(\mathbf{V}_i^C)]}{\partial \mathbf{V}_i^C \partial (\mathbf{V}_i^C)^T} = 2[\mathbf{D}_C^T \mathbf{D}_C + (\lambda + \gamma L_C(i,i)) \mathbf{I}], \quad (8)$$

其中, \mathbf{I} 为单位矩阵。令一阶导数(7)等于0, 可以得到 \mathbf{V}_i^C 的解析解:

$$\begin{aligned} \mathbf{V}_i^{C*} = & [\mathbf{D}_C^T \mathbf{D}_C + (\lambda + \gamma L_C(i,i)) \mathbf{I}]^{-1} \\ & (\mathbf{D}_C^T \mathbf{Z}_i^C - \gamma \sum_{i \neq j} L_C(i,j) \mathbf{V}_j^C), \end{aligned} \quad (9)$$

然后固定稀疏表示矩阵,更新子字典:

$$\mathbf{D}_c^* = \underset{\mathbf{D}_c}{\operatorname{argmin}} \|\mathbf{Z}_c - \mathbf{D}_c \mathbf{V}_c\|_F^2 + \beta \|\mathbf{D}_c\|, \quad (10)$$

令 $P(\mathbf{D}_c)$ 代表式(10),同样可以得到:

$$\frac{\partial [P(\mathbf{D}_c)]}{\partial \mathbf{D}_c} = 2(\mathbf{D}_c \mathbf{V}_c - \mathbf{Z}_c) \mathbf{V}_c^T + \beta \mathbf{I}, \quad (11)$$

$$\frac{\partial^2 [P(\mathbf{D}_c)]}{\partial \mathbf{D}_c \partial \mathbf{D}_c^T} = 2\mathbf{V}_c \mathbf{V}_c^T, \quad (12)$$

式(11)可证是半正定的,即对 \mathbf{D} 是凸函数。令一阶导数等于零,可以得到 \mathbf{D}_c 的解析解:

$$\mathbf{D}_c^* = (2\mathbf{Z}_c \mathbf{V}_c^T - \beta \mathbf{I})(2\mathbf{V}_c \mathbf{V}_c^T + \tau \mathbf{I})^{-1}, \quad (13)$$

其中, $\tau > 0$ 且取极小值,以确保 $(2\mathbf{V}_c \mathbf{V}_c^T + \tau \mathbf{I})^{-1}$ 为非奇异函数。基于上述分析,将算法总结在算法1中。

算法1 快速低秩的判别子字典学习算法

输入:训练样本 $\mathbf{Z} = [\mathbf{Z}_1, \dots, \mathbf{Z}_C]$; 参数 λ, β, γ

输出:学习字典 $\mathbf{D} = [\mathbf{D}_1, \dots, \mathbf{D}_C]$

1. 计算 L_c 。参考下式初始化 \mathbf{D}_{c0} 和 \mathbf{V}_{c0} , 将 \mathbf{D}_{c0} 的每个原子归一化,即:

$$(\mathbf{D}_c, \mathbf{V}_c) = \underset{\mathbf{D}_c, \mathbf{V}_c}{\operatorname{argmin}} \|\mathbf{Z}_c - \mathbf{D}_c \mathbf{V}_c\|_F^2 + \lambda \sum_{i=1}^{N_c} \|\mathbf{V}_i^c\|_2^2 + \beta \|\mathbf{D}_c\| + \gamma \operatorname{tr}(\mathbf{V}_c \mathbf{L}(\mathbf{V}_c)^T), \quad (14)$$

执行下列步骤直到收敛;

2. 固定 \mathbf{D}_c , 更新 \mathbf{V}_c , 即:

$$\mathbf{V}_i^{c*} = [\mathbf{D}_c^T + (\lambda + \gamma L_c(i, i)) \mathbf{I}]^{-1} (\mathbf{D}_c^T \mathbf{Z}_i^c - \gamma \sum_{i \neq j} L_c(i, j) \mathbf{V}_j^c), \quad (15)$$

3. 固定 \mathbf{V}_c , 更新 \mathbf{D}_c :

$$\mathbf{D}_c^* = (2\mathbf{Z}_c \mathbf{V}_c^T - \beta \mathbf{I})(2\mathbf{V}_c \mathbf{V}_c^T + \tau \mathbf{I})^{-1}. \quad (16)$$

2.3 测试过程

在训练过程得到总字典 $\mathbf{D} = [\mathbf{D}_1, \dots, \mathbf{D}_C]$ 后,通过求解岭回归得到测试样本 z 的稀疏表示 \mathbf{V} , 即:

$$\mathbf{V} = \underset{\mathbf{V}}{\operatorname{argmin}} \|\mathbf{z} - \mathbf{D}\mathbf{V}\|_2^2 + \lambda \|\mathbf{V}\|_2^2 =$$

$$(\mathbf{D}^T \mathbf{D} + \lambda \mathbf{I})^{-1} \mathbf{D}^T \mathbf{z}. \quad (17)$$

将训练样本和测试样本的稀疏表示输入 kNN 分类器中,估计测试样本的类别标签并得到算法的分类准确度。

3 实验结果与分析

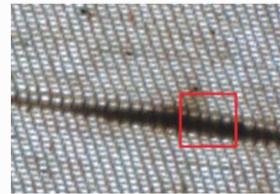
本文在3种数据集 (<http://users.umiacs.umd.edu/~zhuolin/projectlksvd.html>) 上进行实验,具体是:Extended YaleB、AR 和织物疵点数据集。根据每个数据集的特征,通过交叉验证对参数进行调整,本

文使用的最佳参数设置见表1。同时,将实验结果与 FDDL^[8]、LC-KSVD^[9]、LRSDDL^[7]、LSDDL^[5] 算法进行对比。所有实验均在 Windows 7 的电脑上进行,电脑处理器 Intel(R) Core(TM) i5-6500 CPU@ 3.20GHz,内存 4 G,使用的软件为 Matlab。下文将对数据集和实验结果进行详细的阐释与分析。

(1) Extended YaleB 数据集^[10]:由来自 38 个人在 64 种照明条件下的 2 414 张人脸图像组成。所有图像尺寸均调整为 192×168。每类图像均选取 32 张图片进行训练,其他图片作为测试集。随机人脸特征的维数为 504,实验重复 10 次,并将结果的平均值作为分类准确度。实验结果见表2,可以看出,本文提出的方法比 FDDL、LC-KSVD、LRSDDL 和 LSDDL 算法具有更高的精度。

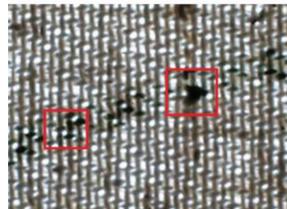
(2) AR 数据集^[10]:由 126 个人的 4 000 多张彩色人脸图像组成。每人都有 26 张正面人脸图。本文选用 50 名女性,50 名男性,共 2 600 张图像,随机选取每类 20 张作为训练集,其余的用于测试。AR 数据集的随机人脸特征维数为 540,结果同样由 10 次重复试验的平均值得到。由表2可以看出,本文所提出的方法获得了更好的性能。

(3) 织物疵点数据集:是由 20 类不同的疵点、共 3 300 张图像组成。几种典型的织物疵点图像如图2所示。本文选择每类 30 个样本用于训练,剩余样本用于测试。研究中使用灰度共生矩阵、方向梯度直方图及局部二值模式对该数据集进行特征提取。织物疵点图像的特征维数为 108。实验结果见表2,本文的方法再次取得最佳结果。

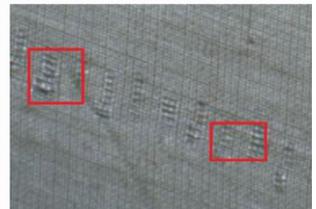


(a) 双纬

(a) Double weft



(b) 油污
(b) Oil stain



(c) 疵点
(c) Defect

图2 一些典型的织物疵点图像

Fig. 2 Some typical fabric defect images

表1 3个数据集的参数设置

Tab. 1 Parameter settings on three databases

	Extended YaleB	AR	织物疵点
λ	1×10^{-4}	5×10^{-6}	1×10^{-6}
β	1×10^{-3}	5×10^{-6}	2×10^{-8}
γ	0.02	0.01	0.16

表2 3个数据集的分类准确度

Tab. 2 Classification accuracy on three databases

字典学习 算法	Extended YaleB	AR	织物疵点	%
FDDL	96.7	97.5	61.9	
LC-KSVD	96.5	97.7	62.2	
LRSDDL	97.9	98.3	63.8	
LSDDL	97.0	97.9	63.7	
The proposed	98.3	98.6	66.4	

4 结束语

本文提出了一种基于子字典的判别字典学习算法,并在3个公开数据集中验证了算法的有效性。该算法针对每类样本得到对应的学习字典,加入子字典的重构误差项,同时利用了原始特征的判别结构,通过拉普拉斯矩阵正则化项提高相同类别稀疏表示的相似性。通过仿真实验结果可以看出,本文提出的算法与一些现有算法相比,能够获得更高的分类性能。后续工作可以在对象分类、纹理识别等更广泛的分类任务中,结合多种映射函数进行研究。

(上接第50页)

- [9] ELDAWY A, MOKBEL M F. SpatialHadoop: A MapReduce framework for spatial data [C]// IEEE 32nd International Conference on Data Engineering. Helsinki, Finland: IEEE, 2016:1352-1363.
- [10] AKDOGAN A, DEMIRYUREK U, BANAEI-KASHANI F, et al. Voronoi-based geospatial query processing with MapReduce [C]// 2010 IEEE Second International Conference on Cloud Computing Technology and Science. NW Washington, DC: IEEE Computer Society, 2010:9-16.
- [11] NISHIMURA S, DAS S, AGRAWAL D, et al. Md-HBase: A scalable multi-dimensional data infrastructure for location aware

参考文献

- [1] AHARON M, ELAD M, BRUCKSTEIN A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation [J]. IEEE Transactions on Signal Processing, 2006, 54(11):4311-4322.
- [2] OLSHAUSEN B A, FIELD D J. Sparse coding with an overcomplete basis set: A strategy employed by V1? [J]. Vision Research, 1997, 37(23):3311-3325.
- [3] ZHENG M, BU J, CHEN C A, et al. Graph regularized sparse coding for image representation [J]. IEEE Transactions on Image Processing, 2011, 20(5):1327-1336.
- [4] 毛丽珍, 汤红忠, 范朝冬, 等. 低秩判别性字典学习及组织病理图像分类算法 [J]. 小型微型计算机系统, 2019, 40(9):1881-1885.
- [5] GUO Jun, GUO Yanqing, LI Yi, et al. Locality sensitive discriminative dictionary learning [C]// 2015 IEEE International Conference on Image Processing. Quebec, Canada: IEEE, 2015:1558-1562.
- [6] LV Wentao, DAI Kaiyan, WU Long, et al. Runway detection in SAR images based on fusion sparse representation and semantic spatial matching [J]. IEEE Access, 2018, 6:27984-27992.
- [7] VU T H, MONGA V. Fast low-rank shared dictionary learning for image classification [J]. IEEE Transactions on Image Processing, 2017, 26(11):5160-5175.
- [8] YANG Meng, ZHANG Lei, FENG Xiangchu, et al. Fisher discrimination dictionary learning for sparse representation [C]// ICCV'11: Proceedings of the 2011 IEEE International Conference on Computer Vision. Barcelona, Spain: IEEE, 2011:543-550.
- [9] JIANG Zhuolin, LIN Zhe, DAVIS L S. Label consistent K-SVD: Learning a discriminative dictionary for recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(11):2651-2664.
- [10] CHEN Yefei, SU Jianbo. Sparse embedded dictionary learning on face recognition [J]. Pattern Recognition, 2017, 64:51-59.

services [C]// 2011 12th IEEE International Conference on MDM. Lulea, Sweden: IEEE, 2011, 1: 7-16.

- [12] ALY A M, MAHMOOD A R, HASSAN M S, et al. AQWA: Adaptive query-workload-aware partitioning of big spatial data [J]// Proceedings of the VLDB Endowment, 2015, 8(13):2062-2073.
- [13] CHEN Zhida, CONG Gao, ZHANG Zhenjie, et al. Distributed publish/subscribe query processing on the spatio-textual data stream [C]// IEEE International Conference on Data Engineering. San Diego, CA, USA: IEEE, 2017: 1095-1106.