

文章编号: 2095-2163(2021)08-0011-04

中图分类号: TP391

文献标志码: A

基于判别子字典学习的图像分类优化方法

赵雅, 钟佳莹, 吕文涛

(浙江理工大学 信息学院, 杭州 310018)

摘要: 本文提出了一种基于判别子字典学习算法的图像分类优化方法。在判别字典学习算法的基础上,引入字典矩阵的正则化约束项,针对每一类图像学习其对应的特定字典,使字典中包含该类别的特定原子,规避不同子字典之间原子的相关性。同时,引入标签信息矩阵和拉普拉斯正则化矩阵,使大系数集中在某一类别的特定原子上,属于同一类别的样本彼此靠近,从而提高字典的判别能力。将该算法应用在3种不同的数据集上,实验结果证明了所提方法的有效性。

关键词: 字典学习; 拉普拉斯矩阵; 子字典; 图像分类

Image classification optimization method based on discriminant sub-dictionary learning

ZHAO Ya, ZHONG Jiaying, LV Wentao

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

[Abstract] This paper proposes an image classification optimization method based on discriminant sub-dictionary learning algorithm. On the basis of the discriminative dictionary learning algorithm, the regularization constraint item of the dictionary matrix is introduced, and the corresponding specific dictionary is learned for each type of image, so that the dictionary contains the specific atom of the category, and avoids the correlation of atoms between different sub-dictionaries. At the same time, the label information matrix and the Laplacian regularization matrix are introduced, so that large coefficients are concentrated on specific atoms of a certain category, and samples belonging to the same category are close to each other, thereby improving the discriminative ability of the dictionary. The algorithm is applied to three different data sets, and the effectiveness of the proposed method is proved through experimental results.

[Key words] dictionary learning; laplacian matrix; sub-dictionary; image classification

0 引言

近年来,判别字典学习算法在处理视觉跟踪、计算机视觉和图像处理等各个领域都有着广泛的应用。文献[1]提出了一种通过共享特征的、具有特征结构的字典学习算法,针对每一类样本进行字典学习,并对字典进行聚类,从而将同一学习字典下相关联的数据点聚集在一起。文献[2]将原始数据的局部关系,整合到基本字典学习框架中。DDL (Discriminant Dictionary Learning, 判别字典学习)算法模型,考虑了类内局部信息和类间特征模糊性。由于判别字典学习依赖于预先设置的系数矩阵的初始值和样本数量,因此该方法不建议直接应用于类间特征模糊度较大的图像目标。

基于以上描述,本文在文献[3]的基础上,改进并提出了基于判别子字典学习的图像分类优化方

法。在目标函数的基础上,构造一个子字典的低秩约束项、标签信息约束项和拉普拉斯矩阵正则化项。针对每一类图像,学习其对应的特定字典,使字典中包含该类别的特定原子,规避不同子字典之间原子的相关性。通过标签信息约束项,将大系数集中在某一类别的特定原子上,加强判别能力。同时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,增强子字典对同类样本的重构能力。为了提高分类的准确率和规避特征图像的字典中的重复原子,在分类过程中引入重建残差,通过重建残差估计测试样本的类别标签,提高字典的判别能力,得到算法的分类准确率。

1 本文算法

1.1 字典学习模型构建

设 $D_C = [d_1, \dots, d_k] \in R^{n \times k}$ 表示第 C 类样本的

基金项目: 国家自然科学基金(61601410);浙江省重点研发项目(2018C01093);浙江理工大学基础研究基金(2019Q0040)。

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;钟佳莹(1996-),女,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。

通讯作者: 吕文涛 Email:21447476@qq.com

收稿日期: 2021-05-29

学习字典, $Z_c = [Z_1^c, \dots, Z_{N^c}^c] \in R^{n \times N^c}$ 表示第 C 类训练样本。在提出的快速低秩的判别子字典学习算法[3]的基础上, 加强子字典的约束重构能力, 结合样本的标签信息提出以下目标函数:

$$\min_{D, V} \sum_{c=1}^C \left\{ \|Z_c - D_c V_c\|_F^2 + \lambda \sum_{i=1}^{N^c} \|V_i^c\|_2^2 + \beta \|D_c\|_F^2 \right\} + \eta \sum_{c=1}^C \|H_c V_c\|_F^2 + \gamma \text{tr}(V L V^T). \quad (1)$$

其中, $V_c = [V_1^c, \dots, V_{N^c}^c] \in R^{K \times N^c}$, 表示为 Z^c 在 D_c 上的稀疏表示矩阵; H_c 为标签矩阵; \tilde{L} 为拉普拉斯正则化矩阵。前两项分别为重建误差项和正则化项, 第三项为子字典的约束项, 后两项分别为标签约束项和拉普拉斯正则化项, λ , β , η 和 γ 是其约束系数。

1.2 标签矩阵

针对不同类别的样本, 为了提高字典的判别能力, 通过标签矩阵使大系数主要集中在同一类别的原子, 尽可能地保留同一类样本的标签属性并减少其它类别原子产生的误差影响, 如图1所示。定义一个标签矩阵 $H_c \in R^{K \times K}$, 标签矩阵的系数可定义为:

$$H_c(m, n) = \begin{cases} 1, & m = n \text{ 且 } m \in \Gamma_c, \\ 0, & \text{其它}. \end{cases} \quad (2)$$

其中, $H_c(m, n)$ 表示相应向量的第 (m, n) 个分量。

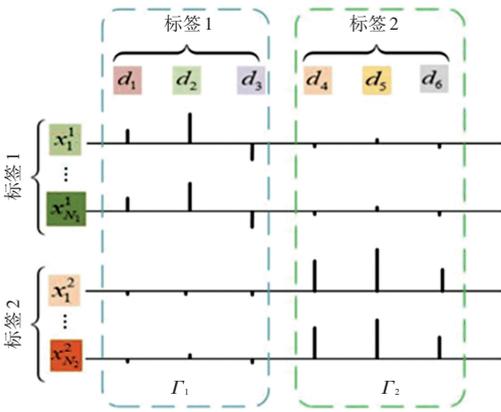


图1 标签矩阵示意图

Fig. 1 Schematic diagram of label matrix

1.3 拉普拉斯正则化矩阵

对于目标函数中的拉普拉斯组正则化项, 由于有 C 个不同类别的样本, 因此构建图时, 将属于同一类别的样本对应的顶点相互靠近并两两相连, 构成一个紧密相连子图, 从而形成 C 个互不相连的子图。已知训练样本和稀疏编码矩阵可以由一个 N

个点组成的直线来表示训练样本中第 K 个元素的图映射^[1]:

$$f_k = \begin{bmatrix} f_k^1 \\ \vdots \\ f_k^K \end{bmatrix} \in R^{N \times 1}, \quad \forall k = 1, 2, \dots, K, \quad (3)$$

可以得到拉普拉斯映射中图的总间距:

$$\text{Var} = \sum_1^K f_k^T \tilde{L} f_k = \text{tr} \left(\begin{bmatrix} f_1^T \\ \vdots \\ f_K^T \end{bmatrix} \tilde{L} [f_1, \dots, f_K] \right) = \text{tr}(V \tilde{L} V^T), \quad (4)$$

则式(4)可以表示为:

$$\text{tr}(V \tilde{L} V^T) = \sum_{c=1}^C \text{tr}(V_c \tilde{L}_c V_c^T), \quad (5)$$

综上, 对于第 C 类样本, 本节所提出的目标函数可以表示为:

$$\arg \min_{D_c, V_c} \|Z_c - D_c V_c\|_F^2 + \lambda \sum_{i=1}^{N^c} \|V_i^c\|_2^2 + \beta \|D_c\|_F^2 + \eta \|H_c V_c\|_F^2 + \gamma \text{tr}(V_c \tilde{L}_c V_c^T). \quad (6)$$

2 算法求解

由于目标函数现在是非凸函数, 因此通过迭代逐步更新字典和稀疏表示矩阵来求解。本文分别从训练过程与测试过程上对算法进行分析。

2.1 训练过程

针对第 C 类样本, 首先初始化学学习字典 D_0^c 和稀疏表示矩阵 V_0^c 。随机初始化学学习字典, 在训练样本 Z_c 中随机选取 K 个样本作为初始字典 D_0^c 的原子。由基础的字典学习目标函数得到初始化的稀疏表示矩阵 V_0^c :

$$V_0^c = \arg \min_{V_c} \|Z_c - D_0^c V_c\|_F^2 + \lambda \|V_c\|_F^2 = ((D_0^c)^T D_0^c + \lambda I)^{-1} (D_0^c)^T Z_c, \quad (7)$$

然后, 固定字典, 更新稀疏表示矩阵。由式(6)可以得到第 C 类样本的稀疏表示矩阵 V_c :

$$V_c^* = \arg \min_{V_c} \|Z_c - D_c V_c\|_F^2 + \lambda \sum_{i=1}^{N^c} \|V_i^c\|_2^2 + \eta \|H_c V_c\|_F^2 + \beta \text{tr}(V_c \tilde{L}_c (V_c)^T), \quad (8)$$

对第 C 类样本逐个更新稀疏表示矩阵, 即

$$v_i^{c*} = \arg \min_{V_i^c} \|z_i^c - D_c v_i^c\|_F^2 + \lambda \|v_i^c\|_2^2 + \eta \|H_c v_i^c\|_F^2 + \beta \sum_{k,j} \tilde{L}_c(k, j) (v_j^c)^T v_k^c, \quad (9)$$

可以得到 v_i^c 的解析解:

$$\hat{v}_i^c = [D_c^T D_c + \eta H_c^T H_c + (\lambda + \beta \tilde{L}_c(i, i)) I]^{-1}$$

$$(D_c^T Z_i^c - \beta \sum_{i \neq j} \tilde{L}_c(i, j) v_j^c), \quad (10)$$

固定稀疏表示矩阵, 逐列更新子字典。即第 C 类样本对应的字典求解公式为:

$$\min_{D_c} \|Z_c - D_c V_c\|_F^2 + \beta \|D_c\|_F^2, \quad (11)$$

为了充分利用已经更新的字典, 选择按原子依次更新字典, 则式(11)可以改为:

$$\min_{d_i} \|P_c - d_i v_i^c\|_F^2 + \gamma \|d_i\|_F^2, \quad (12)$$

其中, $P_c = Z_c - \sum_{j \neq k} d_j v_j^c$ 。

对上式求解, 可以得到

$$\hat{d}_i^c = \frac{P_c (v_i^c)^T}{\|v_i^c\|_2 + \gamma I}. \quad (13)$$

2.2 测试过程

利用给定的训练数据 Z 可以学习得到一组优化的字典 $D_{len} = [D_{len}^1, D_{len}^2, \dots, D_{len}^C]$ 。然后, 给定一个测试样本 z_i , 可以通过求解式(14), 在 C 个学习子字典 D_{len}^c 上, 为测试样本 z_i 获得对应的编码向量 \hat{v}_i 。

$$\hat{v}_i = \arg \min_v \|z_i - D_{len}^c v\|_2^2 + \lambda \|v\|_2^2 = ((D_{len}^c)^T D_{len}^c + \lambda I)^{-1} (D_{len}^c)^T z_i, \quad (14)$$

理想情况下, 所获得的稀疏编码向量 \hat{v}_i 中的原子, 除了与 z_i 所属类别相关的原子外, 其它原子应为 0。因此, 针对这种情况, 定义一个选择算子 $\delta_i(\hat{v}_i)$, 使获得的算子元素除了所属类标签的原子外, 其余全部为 0。对测试样本在这一子类上计算重建残差:

$$\varepsilon_i(z_i) = \|z_i - D_{len}^c \delta_i(\hat{v}_i)\|_2^2, c = 1, 2, \dots, C, \quad (15)$$

由于有 C 个子字典, 因此测试样本 z_c 可以得到 C 个重建残差, 对每个重建残差进行判断, 该测试样本的最佳类别就是其最小残差所对应的类别的归属类:

$$\text{label}(z_i) = \arg \min_{c=1, \dots, C} \varepsilon_i(z_i), \quad (16)$$

综上, 基于子分类的判别字典学习图像分类优化算法步骤如下:

(1) 输入训练样本 $Z = [Z_1, \dots, Z_c]$ 以及参数 $\lambda, \beta, \gamma, \eta$;

(2) 计算 \tilde{L}_c : 通过式(17) 初始化 D_0^c 和 V_0^c 。

$$V_0^c = \arg \min_{S_c} \|Z_c - D_0^c V_c\|_F^2 + \lambda \|V_c\|_F^2 =$$

$$((D_0^c)^T D_0^c + \lambda I)^{-1} (D_0^c)^T Z_c. \quad (17)$$

(3) 固定字典: 通过式(10) 逐列更新稀疏表示系数;

(4) 固定稀疏表示系数: 通过式(13) 逐列更新字典;

(5) 重复步骤 3 和步骤 4, 直到算法收敛或达到所设定的迭代次数;

(6) 输出学习字典 D_{len} 。

3 实验结果与分析

本节将提出的新算法与多种方法进行比较。所有算法均基于 Matlab R2016b 进行编写, 并在联想天翼的 windows 7 系统中运行, 具有 3.20-GHz Intel Core i5-6500 CPU 和 4.0GB 内存。

(1) ExtendedYaleB 数据集^[4]: 由来自 38 个人在 64 种照明条件下的 2 440 张人脸图像组成。每类图像约有 64 张, 图像尺寸设置为 192×168。选取每类的 32 张图像作为训练集, 剩余图像作为测试集。通过交叉验证获得参数, 见表 2。为了确保算法的准确性, 选取 10 次实验结果的平均值, 作为最终准确率, 其结果见表 3。由此可见, 本文所提算法具有更好的精度, 且高于文献[3]的算法准确率。

(2) Caltech 101 数据集^[5]: 由来自 101 个对象类别和一个背景类别的 9 144 幅图像组成。每类图像的张数不等, 类别包含蝴蝶、美洲狮、美洲驼、人脸、手风琴、海豚等目标。该数据集中各类样本之间不仅千差万别, 同类样本内也在形状上各有差异, 该数据集具有较高的复杂性。针对该情况, 本文实验选取了 6 类图片, 每类图像中选用 32 张、64 张和 80 张进行训练, 类别的剩余图片进行测试。

实验首先对每类图片进行特征提取, 从 16×16 色块中提取筛选描述符, 这些色块使用步长为 6 的网格进行密集采样; 然后基于提取的具有 3×1、1×2 和 4×4 网格的筛分特征提取空间金字塔特征^[6]。采用标准的 k 均值聚类训练, 用于空间金字塔的码本, 其中 $k = 1 024$ 。最后, PCA 将空间金字塔特征从 21 504 维缩减为 3 000 个维度。由表 3 可见本文所提算法较其他算法获得了更好的性能。

(3) 天池布匹疵点数据集: 由 32 类不同疵点共 1 996 张图像组成。每类图像张数不等, 分别有擦洞、吊经、污渍、粗纱、破边等多种类别。将天池布匹疵点图像分为正常图像与疵点图像两种类别, 选取每类 200 张图像作为训练样本, 剩余图像作为测试样本。图像尺寸均调整为 256×256。在特征提取阶段, 为了获得更准确的特征信息, 将疵点图像等分为

16×16 的图像块。对每个图像块进行 LBP 算子子分类,然后对每个窗口提取 GLCM 和 HOG 特征,将所有图像块的特征组成该图像的特征向量,用来表示该图像的特征信息。以这种方法,使每个图像的特征信息以不同的结果多次出现在特征向量中。实验结果见表 3,本文的算法精度同样取得了明显的提升。

表 2 不同数据集的参数设置

Tab. 2 Parameter settings on different databases

	λ	β	γ	η
ExYaleB	7×10^{-3}	5×10^{-2}	5×10^{-3}	2×10^2
Caltech101	6×10^{-2}	1×10^{-4}	1×10^{-4}	0.1
天池布匹疵点	2×10^{-4}	5×10^{-5}	1×10^{-4}	0.5

表 3 不同数据集的分类准确率

Tab. 3 Classification accuracy on different databases

算法	ExYaleB	Caltech101	天池布匹疵点
FDDL	96.05±0.75	93.22±0.56	67.95±0.48
LC-KSVD	94.11±1.82	93.33±0.81	58.46±0.42
LRSDDL	95.73±0.97	94.89±0.77	57.56±1.84
LSDDL	96.20±0.81	84.46±0.62	68.75±0.29
文献[2]算法	98.06±0.41	95.78±0.51	71.41±0.93
本文算法	98.16±0.43	97.82±0.47	76.35±0.56

4 结束语

在本文中,研究了一个基于判别子字典学习的图像分类优化方法,主要针对如何减少类间模糊和

类内差异的问题。通过拉普拉斯正则化矩阵、标签信息矩阵和字典约束项,设计特定于每个子类的字典,使子字典中包含该类别的特定原子,减弱不同子字典之间原子的相关性,放大该类别特定原子的系数,将样本的类别标签信息和空间几何结构整合到字典学习的目标函数中,从而增加子字典的判别重构能力。同时,本文在 3 个不同的图像数据集上进行了实验,证明了该方法的可靠性。后续工作可以结合特定图像的特征提取方法,在图像分类、纹理识别等方面做进一步研究。

参考文献

- [1] WANG X, GU Y. Cross-label suppression: A discriminative and fast dictionary learning with group regularization [J]. IEEE Transactions on Image Processing, 2017, 26(8): 3859-3873.
- [2] GUO J, GUO Y, LI Y, et al. Locality sensitive discriminative dictionary learning [C]//IEEE International Conference on Image Processing. IEEE, 2015.
- [3] 赵雅,王顺政,吕文涛,等. 一种快速低秩的判别子字典学习算法及图像分类[J]. 智能计算机与应用,2021,11(1):51-54.
- [4] SCHÖLKOPF B, PLATT J, HOFMANN T. Efficient sparse coding algorithms[M]. MIT Press, 2007.
- [5] MENG Y, LEI Z, FENG X, et al. Fisher Discrimination Dictionary Learning for sparse representation [C]// IEEE International Conference on Computer Vision. IEEE, 2012:543-550.
- [6] JIANG Z, ZHE L, 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.
- [7] LI Y, ZHAO W, PAN J. Deformable Patterned Fabric Defect Detection With Fisher Criterion-Based Deep Learning [J], IEEE Transactions on Automation Science and Engineering, 2017, 14(2): 1256-1264.
- [8] 石美红,张正,郭仙草,等. 基于显著纹理特征的织物疵点检测方法[J]. 纺织学报,2016, 37(10): 42-52.
- [9] LI Y, ZHAO X, CHENG T. Heterogeneous Computing Platform Based on CPU + FPGA and Working Modes [C]//2016 12th International Conference on Computational Intelligence and Security (CIS), 2016: 669-672.
- [10] 杨柳. 基于机器视觉的机织物疵点实时检测平台研发[D]. 上海: 东华大学,2018.
- [11] SILVA J, SKLYAROV V, SKLIAROVA I. Comparison of On-chip Communications in Zynq-7000 All Programmable Systems-on-Chip [J], IEEE Embedded Systems Letters, 2015, 7(1): 31-34.
- [12] 田宸玮,王雪纯,杨嘉能,等. 织物瑕疵检测方法研究进展[J]. 计算机工程与应用, 2020, 56(12): 8-18.
- [13] LI F, YUAN L, ZHANG K, et al. A defect detection method for unpatterned fabric based on multidirectional binary patterns and the gray-level co-occurrence matrix [J]. Textile Research Journal. 2020;90(7-8):776-796.
- [14] 李文青. 基于织物纹理分类的分层瑕疵检测系统研究[D]. 上海: 东华大学, 2020.
- [15] BRUNELLI, ROBERTO. Template matching techniques in computer vision: theory and practice [M]. John Wiley & Sons, 2009:7-38.
- [16] 胡景华. 基于 AXI 总线的 SoC 架构设计与分析[D]. 上海: 上海交通大学,2013.
- [17] ZHANG X, WANG J, WANG Y, et al. BRAM-based asynchronous FIFO in FPGA with optimized cycle latency [C]//2012 IEEE 11th International Conference on Solid-State and Integrated Circuit Technology, Xi'an, 2012: 1-3.
- [18] CONG J, WEI P, YU C H, et al. Automated Accelerator Generation and Optimization with Composable, Parallel and Pipeline Architecture [C]//2018 55th ACM/ESDA/ IEEE Design Automation Conference (DAC), San Francisco, CA, 2018: 1-6.

(上接第 10 页)