

# 基于图正则化局部特征编码算法的 图像分类方法

杨 赛<sup>1</sup>, 赵春霞<sup>2</sup>, 胡 彬<sup>3</sup>, 陈 峰<sup>1</sup>

(1. 南通大学电气工程学院, 江苏南通 226019; 2. 南京理工大学计算机科学与工程学院, 江苏南京 210094;  
3. 南通大学计算机科学与技术学院, 江苏南通 226019)

**摘 要:** 为了解决经典局部特征编码方法会产生相似局部特征之间编码系数不一致的问题, 本文提出一种图正则化局部特征编码算法. 该算法在对初始编码矢量所定义的能量化函数中引入正则化项, 保证空间上相邻外观上相似的局部特征之间的编码矢量尽可能一致. MSRcv2、Caltech101、Scene 15 以及 Indoor 67 四个公开数据集上的实验结果表明本文方法能够提高硬分配、软分配、稀疏编码、局部约束线性编码以及局部软分配五种经典编码方法的性能, 并且基于本文编码算法的图像分类方法在上述四个公开数据集上的平均分类正确率分别达到了 91.13%、76.02%、83.76%、44.78%.

**关键词:** 词袋模型; 编码算法; 图模型; 图像分类

**中图分类号:** TP391.4

**文献标识码:** A

**文章编号:** 0372-2112 (2017)08-1882-06

**电子学报 URL:** <http://www.ejournal.org.cn>

**DOI:** 10.3969/j.issn.0372-2112.2017.08.011

## An Image Classification Method Using Graphically Regularized Coding Algorithm

YANG Sai<sup>1</sup>, ZHAO Chun-xia<sup>2</sup>, HU Bin<sup>3</sup>, CHEN Feng<sup>1</sup>

(1. School of Electrical Engineering, Nantong University, Nantong, Jiangsu 226019, China;

2. School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, Jiangsu 210094, China;

3. School of Computer Science and Technology, Nantong University, Nantong, Jiangsu 226019, China)

**Abstract:** In order to solve the problem that current coding schemes lost consistence between similar local features, this paper proposes a new graphically regularized coding algorithm. This algorithm used any current coding scheme to get the initial coding coefficients, and utilized a regularized term to preserve locality constraints both in the feature space and the spatial domain of the image. Experimental results on popular benchmark datasets show that our method improves the performances of the current coding algorithms, and the average classification accuracies of our proposed method in MSRcv2, Caltech101, Scene15, Indoor 67 and UIUC-sport has reached 91.13%, 76.02%, 83.76%, 44.78% and 89.05% respectively.

**Key words:** bag-of-feature; coding algorithm; graphical model; image classification

## 1 引言

如今面对海量的图像信息, 图像的自动分类已成为当前研究的新热点, 而词袋模型因其简单高效已经成为当前流行的图像特征表示方法. 该模型按照某种编码规则得到局部特征在视觉词汇向量张成的语义空间中的投影系数, 将投影向量整合后表征图像. 其中的

编码算法是局部特征与视觉词汇向量之间的联系纽带, 是词袋模型的核心步骤.

标准词袋模型中的编码操作作为硬分配 (Hard-assignment Coding, HAC)<sup>[1]</sup>, 该方法选择距离最近的视觉词汇表征局部描述矢量, 会产生很大的量化误差; 最直观的改进方法是软分配 (Soft-assignment Coding, SAC)<sup>[2]</sup>, 该方法使用所有视觉词汇向量的加权和表示局部描述矢量;

收稿日期: 2016-02-22; 修回日期: 2016-10-13; 责任编辑: 马兰英

基金项目: 江苏省普通高校自然科学研究面上项目 (No. 16KJB520037); 国家自然科学基金 (No. 61602150); 江苏省自然科学基金 (No. BK20151273); 南通市科技项目前沿与关键技术 (No. MS22015100); 江苏省博士后科研资助计划项目 (No. 1601013B)

Yang 等<sup>[3]</sup>学者提出的稀疏编码(Sparse Coding, SC)使用有限个视觉词汇向量对局部特征进行编码,同样可以尽量减少量化误差;针对 SC 求解  $l_1$  范数非常耗时的问题, Wang 等<sup>[4]</sup>提出一种局部约束的线性编码(Local-constrained Linear Coding, LLC)算法,该方法对局部特征与视觉词汇向量之间局部关系进行约束;针对软分配编码方法忽略了局部特征与视觉词汇之间的势场关系, Liu 等<sup>[5]</sup>提出的局部软分配编码(Localized Soft-assignment Coding, LSC)使用核函数将局部特征分配到  $k$  个最近邻的视觉单词. 然而上述经典编码算法都只是对图像中的每个局部特征单独进行编码,忽略了它们之间的空间上下文信息;并且由于视觉码书通常是过完备的,外观相似的局部特征会选择不同的视觉单词进行响应,会产生相似的局部特征之间的编码系数不一致的现象. 为了解决上述问题,本文提出一种图正则化(Graphically Regularized, GR)局部特征编码算法.

## 2 典型局部特征编码算法

假设给定图像数据集为  $I = \{I_1, I_2, \dots, I_N\}$ , 对其中每幅图像提取 SIFT 特征, 随机抽取所有训练图像局部特征集合中的一个子集  $Y$ , 使用  $K$  均值算法对  $Y$  聚类得到  $K$  个视觉词汇向量, 其中第  $k$  个记为  $w_k \in R^{D \times 1}$ ,  $k = 1, 2, \dots, K$ ,  $D$  为 SIFT 特征的维数, 则  $K$  个词汇向量组成视觉码书  $W = [w_1, w_2, \dots, w_K]^T \in R^{D \times K}$ . 将图像  $I_j \in I$  中提取的  $M$  个 SIFT 特征记为  $X = \{x_1, x_2, \dots, x_M\}$ , 相应的坐标矢量记为  $L = \{l_1, l_2, \dots, l_M\}$ , 则其中的第  $m$  个 SIFT 特征为  $x_m$ , 坐标特征为  $l_m$ , 设  $c_m \in R^{K \times 1}$  表示相应的编码特征矢量, 则典型编码算法的表示如下:

### (1) 硬分配编码

$$c_{mk} = \begin{cases} 1, & \text{如果 } k = \arg \min_{k=1, \dots, K} \|x_m - w_k\|_2^2 \\ 0, & \text{否则} \end{cases} \quad (1)$$

式(1)中,  $\|\cdot\|_2$  为欧式距离,  $c_{mk}$  为第  $m$  个编码特征第  $k$  维上的数值.

### (2) 软分配编码

$$c_{mk} = \frac{\exp(-\beta \|x_m - w_k\|_2^2)}{\sum_{k=1}^K \exp(-\beta \|x_m - w_k\|_2^2)} \quad (2)$$

式(3)中,  $\beta$  为光滑系数.

### (3) 稀疏编码

$$c_m = \arg \min_{c \in R^K} \|x_m - Wc_m\|_2^2 + \alpha \|c_m\|_1 \quad (3)$$

式(3)中,  $\|\cdot\|_1$  为  $l_1$  范数,  $\alpha$  为正则系数

### (4) 局部约束线性编码

$$\begin{aligned} \min_c \sum_{m=1}^M \|x_m - Wc_m\|^2 + \lambda \|w_k \Theta d_m\|^2 \\ \text{s. t. } |c_m|_1 = 1, \forall m \end{aligned} \quad (4)$$

式(4)中,  $d_m \in R^K$  表示  $x_m$  与  $w_k$  之间的距离矢量,  $\Theta$  表示  $d_m$  与  $c_m$  中对应元素的乘积.

### (5) 局部软分配编码

$$\begin{aligned} c_{mk} &= \frac{\exp(-\beta \hat{D}(x_m, w_k))}{\sum_{s=1}^s \exp(-\beta \hat{D}(x_m, w_s))} \\ \hat{D}(x_m, w_s) &= \begin{cases} D(x_m, w_s), & \text{如果 } w_s \in N_k(x_m) \\ \infty, & \text{否则} \end{cases} \end{aligned} \quad (5)$$

式(5)中,  $\hat{D}(x_m, w_s)$  表示  $x_m$  与  $w_s$  之间的距离,  $N_k(x_m)$  表示  $x_m$  的  $k$  个近邻视觉词汇向量.

## 3 图正则化局部特征编码算法

### 3.1 算法描述

第2节中典型局部特征编码算法表达式可以归结为:

$$c_{ml} = C(x_m, W) \quad (6)$$

式(6)中,  $C(\cdot)$  为局部特征编码函数. 将编码矢量  $c_{ml}$  作为初始编码矢量, 则所有编码矢量组成初始编码矢量矩阵  $C_l = [c_{1l}, c_{2l}, \dots, c_{Ml}] \in R^{M \times K}$ . 对于图像  $I_j \in I$  中的  $M$  个 SIFT 特征  $X = \{x_1, x_2, \dots, x_M\}$  以及坐标矢量  $L = \{l_1, l_2, \dots, l_M\}$ , 将其表示为一个带权无向图  $G(O, E, V)$ . 图  $G$  的节点为每一个 SIFT 特征, 所有节点集合为  $O$ , 任意两个节点由一条边连接起来, 并且被赋一权值表征它们之间的相似程度, 所有边的集合记为  $E$ , 所有权值的集合为  $V$ , 其中第  $ij$  个元素  $v_{ij}$  的计算公式为:

$$v_{ij} = \begin{cases} e^{-\left\{\frac{\|x_i - x_j\|}{\sigma_i^2}\right\}} * e^{-\left\{\frac{\|l_i - l_j\|}{\sigma_l^2}\right\}}, & \text{如果 } l_j \in N(l_i) \\ 0, & \text{否则} \end{cases} \quad (7)$$

式(7)中,  $x_i, x_j$  分别表示图像中第  $i$  个和第  $j$  个 SIFT 特征,  $l_i, l_j$  分别表示其相应的坐标特征,  $N(l_i)$  表示第  $i$  个 SIFT 特征的  $K$  个空间近邻.

假设对集合  $X$  进行编码后得到编码矢量矩阵为  $C = [c_1, c_2, \dots, c_M] \in R^{K \times M}$ , 对其定义下面的能量函数:

$$E(c_i) = (c_i - c_{il})^2 + \gamma \sum_{i,j=1}^M v_{ij} (c_i - c_j)^2 \quad (8)$$

式(8)中,  $c_i$  和  $c_j$  表示第  $i$  和第  $j$  个 SIFT 特征的编码矢量,  $c_{il}$  表示第  $i$  个 SIFT 特征的初始编码矢量. 等式(8)中右边的第一项为拟合项, 表示 SIFT 特征的编码矢量与初始编码矢量之间的方差, 确保两者之间的误差尽可能的小; 第二项为正则化项, 表示空间上相邻 SIFT 特征编码矢量之间的方差, 保证空间上相邻外观上相似的局部特征之间的编码矢量尽可能一致, 图正则化系数  $\gamma$  控制两项的比重. 对式(8)求微分并令其为零最终得到:

$$C = \mu(D - V + \mu I)^{-1} C_l \quad (9)$$

式(9)中,  $V$  为  $M \times M$  的对称矩阵的权值集合, 即使用公式(7)计算其中的每一个元素,  $D$  为  $M \times M$  的对角矩

阵, 对角线上的元素为  $d_i = \sum_j v_{ij}, \mu = 1/(2\gamma)$ .

### 3.2 局部特征与编码矢量之间的相似性匹配

与文献[6]一致, 对图像数据集中某幅图像提取 SIFT 特征, 分别使用 HAC 和 GR-HAC 对局部特征进行编码, 并使用直方图交叉计算局部特征以及编码矢量两两之间的相似系数, 局部特征与编码矢量的相似性匹配关系如图 1 所示. 由图 1(a) 可知, 局部特征与编码矢量构成的坐标点处于坐标轴的底部和顶部, 几乎没有点出现在  $y=x$  这条线的附近, 表明 HAC 中相似局部特征之间的编码系数不一致的现象非常严重; 由图 1(b) 可知, 局部特征与编码矢量之间的相似性匹配关系满足线性函数, 验证了本文方法的有效性.

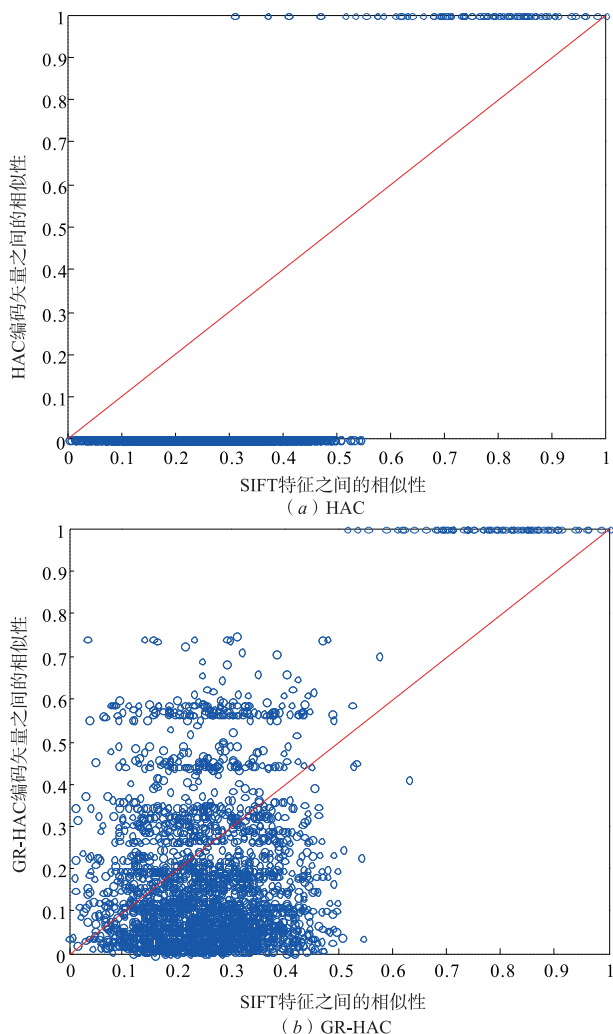


图1 局部特征与编码矢量之间相似性匹配关系图

## 4 实验与分析

实验选取 MSRcv2、Caltech101、Scene 15 以及 Indoor67 四个公开数据库验证本文方法的性能. 对所有的图像数据集, 提取每幅图像的密集 SIFT 特征, 采样步长

为 8, 图像小块的大小分别为  $16 \times 16, 32 \times 32, 64 \times 64$ ,  $K$  的值为 1500, 空间金字塔模型为 3 层.

### 4.1 MSRcv2 数据集

MSRcv2 数据集<sup>[7]</sup>包含 15 种类别, 每个类别的图像数目为 30, 示例图片如图 2 所示. 使用其中的 9 种类别作为实验对象, 随机抽取每类中的 15 幅图像为训练数据, 剩余图像为测试数据, 重复抽取 100 次, 实验结果取各次平均值. 本文方法与 HAC、SAC、SC、LLC、LSC 的对比结果见表 2, 由此可知本文方法都不同程度地提高了五种经典编码方法的性能. 基于 GR-LSC 编码算法的图像分类方法与目前主流方法的对比结果如表 1 所示, 由此可知与对比方法中性能最优的文献[9]相比较, 本文方法仍然高于其 0.43%.

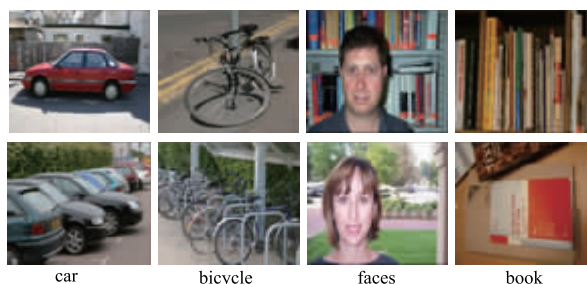


图2 MSRcv2数据集的示例图片

表1 MSRcv2 数据集上本文方法与主流图像分类方法的对比结果(%)

分类方法	正确率
Zhang <sup>[7]</sup>	$80.4 \pm 2.5$
Sawaree <sup>[8]</sup>	80.00
Su <sup>[9]</sup>	$90.7 \pm 1.8$
本文方法	<b><math>91.13 \pm 2.10</math></b>

表2 MSRcv2 数据集上本文方法与典型编码算法的对比结果(%)

编码算法	分类正确率	本文方法	分类正确率
HAC	$88.27 \pm 2.60$	GR-HAC	$89.65 \pm 2.39$
SAC	$90.15 \pm 2.46$	GR-SAC	$90.53 \pm 2.62$
SC	$90.25 \pm 2.34$	GR-SC	$90.80 \pm 2.24$
LLC	$90.16 \pm 2.36$	GR-LLC	$90.91 \pm 2.13$
LSC	$90.33 \pm 2.18$	GR-LSC	$91.13 \pm 2.10$

### 4.2 Caltech101 数据集

Caltech101 数据集<sup>[10]</sup>包含 102 种类别, 每个类别的图像数目在 40 到 800 幅不等, 示例图片如图 3 所示. 随机抽取每类中的 30 幅图像为训练数据, 剩余图像为测试数据, 重复抽取 10 次, 实验结果取各次平均值. 本文方法与 HAC、SAC、SC、LLC、LSC 的对比结果见表 4, 由此可知本文方法都不同程度地提高了五种经典编码方法的性能. 基于 GR-LLC 的图像分类方法与目前主流方法的对比结果如表 3 所示, 由此可知与对比方法中性能最优的文献[13]相比较, 本文方法仍然高于其 0.34%.

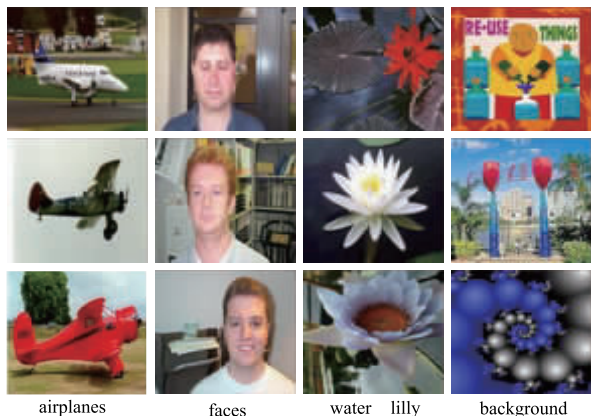


图3 Caltech101 数据集的示例图片

表 3 Caltech101 数据集上本文方法与主流图像分类方法的对比结果(%)

分类方法	正确率	分类方法	正确率
Griffin <sup>[11]</sup>	67.6	Baureau <sup>[12]</sup>	71.50 ± 1.1
Zhang&liu <sup>[13]</sup>	75.68 ± 0.89	Wang <sup>[14]</sup>	74.47 ± 0.46
Shabou <sup>[15]</sup>	73.23 ± 0.81	Zhang <sup>[16]</sup>	75.02 ± 0.74
Jia <sup>[17]</sup>	75.3 ± 0.7	Wu <sup>[18]</sup>	73.4 ± 1.2
本文方法	<b>76.02 ± 0.88</b>		

表 4 Caltech101 数据集上本文方法与典型编码算法的对比结果(%)

编码算法	分类正确率	本文方法	分类正确率
HAC	72.76 ± 0.52	GR-HAC	73.82 ± 0.83
SAC	74.24 ± 1.03	GR-SAC	74.96 ± 1.06
SC	74.81 ± 0.76	GR-SC	75.88 ± 0.72
LLC	75.16 ± 0.93	GR-LLC	<b>76.02 ± 0.88</b>
LSC	74.97 ± 0.97	GR-LSC	75.91 ± 1.22

### 4.3 Scene 15 数据集

Scene 15 数据集<sup>[19]</sup>包含 15 种类别的自然与人工场景图像,每个类别的图像数目在 200 到 400 幅不等,示例图片见图 4。随机抽取每种类别中的 100 幅图像为训练数据,剩余图像为测试数据,重复抽取 10 次,实验结果取各次平均值。本文方法与 HAC、SAC、SC、LLC、LSC 的对比结果见表 6,由此可知本文方法都不同程度地提高了五种经典编码方法的性能。基于 GR-LLC 的图像分

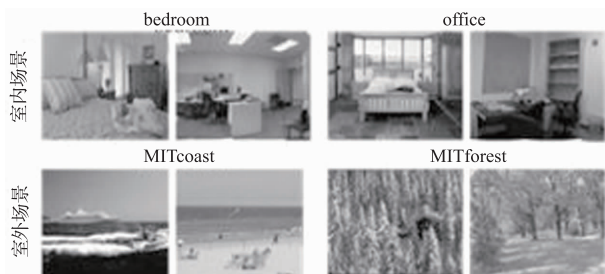


图4 Scene 15数据集的示例图片

类方法与目前主流方法的对比结果如表 5 所示,由此可知与对比方法中性能最优的文献[21]相比较,本文方法仍然高于其 0.82%。

表 5 Scene 15 数据集上本文方法与主流图像分类方法的对比结果(%)

分类算法	分类正确率	分类算法	分类正确率
Shabou <sup>[15]</sup>	82.67 ± 0.51	Lazebik <sup>[19]</sup>	81.40 ± 0.50
Yang <sup>[20]</sup>	82.51 ± 0.43	Theriat <sup>[21]</sup>	82.94 ± 0.57
本文方法	<b>83.76 ± 1.07</b>		

表 6 Scene 15 数据集上本文方法与典型编码算法的对比结果(%)

编码算法	分类正确率	本文方法	分类正确率
HAC	81.79 ± 0.53	GR-HAC	82.76 ± 0.73
SAC	82.77 ± 0.46	GR-SAC	83.55 ± 0.72
SC	83.25 ± 0.52	GR-SC	83.47 ± 0.67
LLC	83.23 ± 0.57	GR-LLC	<b>83.76 ± 1.07</b>
LSC	83.38 ± 0.74	GR-LSC	83.63 ± 0.67

### 4.4 Indoor 67 数据集

Indoor 67 数据集<sup>[22]</sup>包含 67 种类别的室内场景图像,每种类别中的图像数目在 100 幅以上,示例图片如图 5 所示。使用数据集相关网站上提供的固定划分方式,即每类中的 80 幅图像被用作为训练,每类中的 20 幅图像被用作为测试。本文方法与 HAC、SAC、SC、LLC、LSC 的对比结果见表 8,由此可知本文方法都不同程度地提高了五种经典编码方法的性能。基于 GR-LSC 的图像分类方法与目前主流方法的对比结果如表 7 所示,由此可知与对比方法中性能最优的文献[24]相比较,本文方法仍然高于其 2.98%。

表 7 Indoor 67 数据集上本文方法与主流图像分类方法的对比结果(%)

分类算法	分类正确率	分类算法	分类正确率
Quattoni <sup>[22]</sup>	26.5	Li <sup>[23]</sup>	37.6
Bo <sup>[24]</sup>	41.8		
本文方法	<b>44.78</b>		



图5 Indoor 67数据集的示例图片

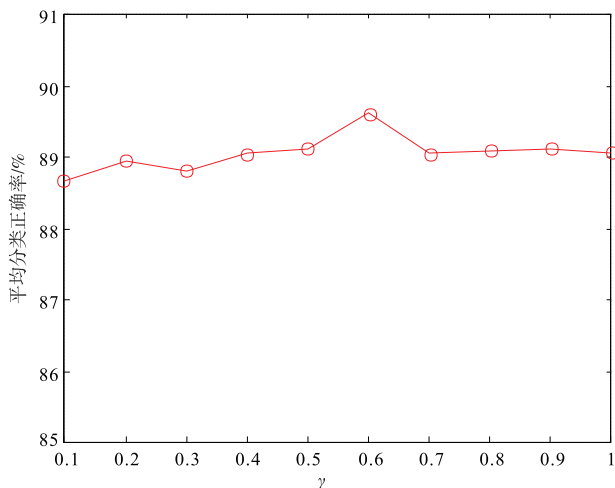


表 8 Indoor 67 数据集上本文方法与典型编码算法的对比结果 (%)

编码算法	分类正确率	本文方法	分类正确率
HAC	42.76	GR-HAC	44.17
SAC	42.98	GR-SAC	43.85
SC	43.36	GR-SC	44.42
LLC	43.58	GR-LLC	44.28
LSC	43.66	GR-LSC	<b>44.78</b>

#### 4.5 图正则化系数 $\gamma$ 的选择

图正则化系数  $\gamma$  为本文方法的重要参数,为此本节在 MSRcv2 数据集上给出参数对分类性能的影响曲线.在区间  $[0.1, 1]$  中以 0.1 为采样步长均匀选取一系列  $\gamma$  的值,各种参数下 GR-HAC 的平均分类正确率如图 6 所示,由此可知当  $\gamma$  的值为 0.5–0.7 之间时,它们之间的分类性能相差不大,因此实验过程中图正则化系数被设为 0.5–0.7 之间.

图 6 图正则化系数  $\gamma$  性能影响曲线

## 5 总结

针对目前经典编码算法忽略了局部特征之间的空间上下文信息的问题,本文提出一种图正则化局部特征编码方法.该算法在对初始编码矢量所定义的能量化函数中引入正则化项,从而使得相似的局部特征之间的编码系数保持一致性. MSRcv2、Caltech101、Scene 15 以及 Indoor 67 四个公开数据集上的实验结果表明本文方法经典编码方法的性能,并且基于本文编码算法的图像分类方法在上述五个公开数据集上取得了最高的平均分类正确率.

#### 参考文献

[1] Csúrká G, Dance C R, Fan L X, et al. Visual categorization with bags of keypoints [A]. Proceedings of the 8th European Conference on Computer Vision [C]. Prague, CZE:

Springer-Verlag, 2004. 1–22.

- [2] Van Gemert J C, Veenman C J, Smeulders A W M, et al. Visual word ambiguity [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32 (7): 1271–1284.
- [3] Yang J C, Yu K, Gong Y H, Huang T. Linear spatial pyramid matching using sparse coding for image classification [A]. Proceedings of the 22nd International Conference on Computer Vision and Pattern Recognition [C]. Miami, USA: IEEE Computer Society, 2009. 1794–1801.
- [4] Wang J J, Yang J C, Yu K, et al. Locality-constrained linear coding for image classification [A]. Proceedings of the 23rd International Conference on Computer Vision and Pattern Recognition [C]. San Francisco, USA: IEEE Computer Society, 2010. 3360–3367.
- [5] Liu L Q, Wang L, Liu X W. In defense of soft-assignment coding [A]. Proceedings of the 13th International Conference on Computer Vision [C]. Barcelona, Spain: IEEE Computer Society, 2011. 2486–2493.
- [6] Gao S, Tsang I, Huang T, Chia L. Laplacian sparse coding, hypergraph laplacian sparse coding, and application [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35 (1): 92–104.
- [7] Zhang Y, Chen T. Efficient kernels for identifying unbounded-order spatial features [A]. Proceedings of the 22nd International Conference on Computer Vision and Pattern Recognition [C]. Miami, USA: IEEE Computer Society, 2009. 1762–1769.
- [8] Sawaree S, Winn J, Criminisi A. Discriminative object class models of appearance and shape by correlations [A]. In: Proceedings of the 19th International Conference on Computer Vision and Pattern Recognition [C]. New York, USA: IEEE Computer Society, 2006. 2033–2040.
- [9] Su Y, Jurie F. Visual word disambiguation by semantic contexts [A]. Proceedings of the 13th IEEE International Conference on Computer Vision [C]. Barcelona, Spain: IEEE Computer Society, 2011. 311–318.
- [10] Li F F, Fergus R, Perona P. Learning generative visual models from few training examples: an incremental Bayesian approach tested on 101 object categories [J]. Computer Vision and Image Understanding, 2007, 106 (1): 59–70.
- [11] Griffin G, Holub A, Perona. Caltech256 object category dataset [R]. Pasadena: CIT, 2007. 1–20.
- [12] Boureau Y L, Boch F, LeCun Y, Ponce J. Learning mid-level features for recognition [A]. In: Proceedings of the 23rd International Conference on Computer Vision and Pattern Recognition [C]. San Francisco, USA: IEEE Computer Society, 2010. 1–8.
- [13] Zhang C J, Liu J, Tian Q, et al. Image classification by

- non-negative sparse coding, Low-Rank and sparse decomposition [A]. Proceedings of the 24th International Conference on Computer Vision and Pattern Recognition [C]. Colorado Springs, USA: IEEE Computer Society, 2011. 1673 – 1680.
- [14] Wang Z, Feng J, Yan S, et al. Linear distance coding for image classification [J]. IEEE Transaction on Image Processing, 2013, 22(2): 537 – 548.
- [15] Shabou A, Le Borgne H. Locality-constrained and spatially regularized coding for scene categorization [A]. Proceedings of the 25th International Conference on Computer Vision and Pattern Recognition [C]. Providence, USA: IEEE Computer Society, 2012. 3618 – 3625.
- [16] Zhang T Z, Ghanem B, Liu S. Xu C S, et al. Low-rank sparse coding for image classification [A]. Proceedings of the 14th International Conference on Computer Vision [C]. Sydney, Australia: IEEE Computer Society, 2013. 281 – 288.
- [17] Jia Y Q, Huang C, Darrell T. Beyond spatial pyramids: receptive field learning for pooled image features [A]. Proceedings of the 25th International Conference on Computer Vision and Pattern Recognition [C]. Providence, USA: IEEE Computer Society, 2012. 3370 – 3377.
- [18] Wu Z F, Huang Y Z, Wang L, Tan T N. Group encoding of local features in image classification [A]. Proceedings of the 19th International Conference on Image Processing [C]. Orlando, USA: IEEE Computer Society, 2012. 1505 – 1508.
- [19] Lazebnik S, Schmid C, Ponce J. Beyond bags of features: spatial pyramid matching for recognizing natural scene categories [A]. Proceedings of the 19th International Conference on Computer Vision and Pattern Recognition [C]. New York, USA: IEEE Computer Society, 2006. 2169 – 2178.
- [20] Yang Y, Newsam S. Spatial pyramid co-occurrence for image classification [A]. Proceedings of the 13th IEEE International Conference on Computer Vision [C]. Barcelona, Spain: IEEE Computer Society, 2011. 1465 – 1472.
- [21] Thériault C, Thome N, Cord M. Extended coding and pooling in the HMAX model [J]. IEEE Transaction on Image Processing, 2013, 22(2): 764 – 777.
- [22] Quattoni A, Torralba A. Recognizing indoor scenes [A]. Proceedings of the 22nd International Conference on Computer Vision and Pattern Recognition [C]. Miami, USA: IEEE Computer Society, 2009. 413 – 420.
- [23] Li L, Su H, Xing E, Li F F. Object bank: a high-level image representation for scene classification and semantic feature sparsification [A]. Proceedings of the 24th Annual Conference on Neural Information Processing Systems [C]. Vancouver B. C., Canada: MIT Press, 2010. 1378 – 1386.
- [24] Bo L, Ren X, Fox D. Hierarchical matching pursuit for image classification: architecture and fast algorithm [A]. Proceedings of the 25th Annual Conference on Neural Information Processing Systems [C]. Vancouver B. C., Canada: MIT Press, 2011. 2115 – 2123.

#### 作者简介



杨 赛 (通信作者) 女, 1981 年出生, 山东滨州人. 2015 年获南京理工大学博士学位, 现为南通大学讲师, 研究方向为计算机视觉与机器学习.

E-mail: yangsai166@126.com



赵春霞 女, 1964 年出生, 北京人. 1985 年、1988 年和 1998 年在哈尔滨工业大学分别获得工学学士、工学硕士和工学博士学位, 现为南京理工大学教授, 博士生导师. 主要研究方向为地面智能机器人与复杂环境理解.