

局部位置特征与全局轮廓特征 相结合的图像分类方法

李雅倩¹, 吴超², 李海滨¹, 刘彬²

(1. 燕山大学工业计算机控制工程河北省重点实验室, 河北秦皇岛 066004;
2. 燕山大学信息科学与工程学院, 河北秦皇岛 066004)

摘要: 在空间金字塔词袋模型的基础上, 针对其空间信息利用不足的问题, 本文先计算图像中每一个字典向量的相对位置分布来提取出局部位置特征. 然后, 用非下采样轮廓波变换和线性判别分析来生成图像的全局轮廓特征. 最后, 通过局部位置特征与全局轮廓特征相结合的方式提高空间信息利用率, 从而提高场景和物体图像分类正确率. 为了检验方法的可行性, 本文分别在数据库 Caltech 101、MSRC 和 15 Scene 上进行实验. 实验结果证明, 本文提出的方法进一步利用了空间信息, 从而提高了分类正确率.

关键词: 图像分类; 词袋模型; 局部位置特征; 非下采样轮廓波变换; 全局轮廓特征

中图分类号: TP391.41 **文献标识码:** A **文章编号:** 0372-2112 (2018)07-1726-06

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2018.07.026

Image Classification Method Combining Local Position Feature with Global Contour Feature

LI Ya-qian¹, WU Chao², LI Hai-bin¹, LIU Bin²

(1. Key Lab of Industrial Computer Control Engineering of Hebei Province, Yanshan University, Qinhuangdao, Hebei 066004, China;
2. Institute of Information Science and Engineering, Yanshan University, Qinhuangdao, Hebei 066004, China)

Abstract: Based on Spatial Pyramid Matching method, aiming at the insufficient utilization of spatial information, firstly, local position feature is extracted by computing relative position distribution of each dictionary vector in image. Then, global contour feature is generated through Nonsubsampled Contourlet Transform and Linear Discriminant Analysis. Finally, Spatial information is enhanced by combining local position feature with global contour feature, which consequently improves the accuracy of scene and object classification. Extensive experiments are performed on Caltech 101, MSRC and 15 Scene datasets respectively. The experimental results show that the proposed method further utilizes the spatial information, and thus improves the accuracy of image classification.

Key words: image classification; bag of words; local position feature; nonsubsampled contourlet transform; global contour feature

1 引言

随着图像分类的飞速发展, 该领域的学者们研发与改进了许多非常有效的分类方法. 文献[1]首次引入术语词袋模型(Bag of Words, BoW), 利用频率统计方法来解决图像分类问题. BoW 在统计全局直方图时, 没有考虑特征的分布情况, 造成了特征空间分布信息的丢失. 因此, 很多基于 BoW 的方法都是在特征空间分布信息的利用上进行改进, 以得到更好的图像分类性能.

为了解决 BoW 没有考虑空间信息的问题, 文献

[2]提出了空间金字塔词袋模型(Spatial Pyramid Matching, SPM), 通过对图像进行不同等级的网格划分来利用空间信息. Vector of Locally Aggregated Descriptors^[3] (VLAD) 分别累加一幅图像中各个字典向量与其对应的 SIFT 特征向量的差值, 从而求出这一幅图像中各个字典向量周围特征向量的分布特征. VLAD 采用的是硬分配, 在分配 SIFT 特征时有很大几率造成误差, 同时这些分布特征按照字典顺序串联起来, 会产生高维 VLAD 特征, 增加分类的计算难度. 文献[4]提出了一种池化方法 BossaNova, 该方法通过设置区间阈值来确定 SIFT

特征归属于哪个字典向量,根据区间阈值内的 SIFT 和相应的字典向量的欧氏距离对这些向量分级,然后统计每个字典周围各级 SIFT 的个数直方图,最后将这些直方图进行串联. BossaNova 通过距离分级来加入空间信息,但是没有考虑图像中的实际空间信息. 文献[5]考虑了特征向量之间相对距离和角度的共同影响,设计出 Soft Pairwise Spatial Angle-Distance Histograms (SPS_{sd})来提取相似特征的分布信息,虽然有效利用了局部空间信息但并未考虑图像的整体特征信息. 文献[6]先用 Shi-Tomasi corner 和 Itti saliency map 提取出物体所在的区域部分,然后用优化过的 SPM 对这部分进行处理. 由于忽略了背景信息,该方法无法有效处理多物体和场景图像分类问题. 文献[7]对目前改进的词袋模型方法进行了总结与分类,并提出了 Higher-order Occurrence Pooling 方法. 该方法除了通过聚合每幅图像中字典特征的共现关系系数来使得特征更具有辨识度,

还可以解决多个局部特征的融合问题,但该方法依然没有有效利用图像中蕴含的空间信息. 通过不断地改进与优化,这些方法的分类性能也在不断提高,但是空间信息利用不足的情况依然没有得到有效的改善.

SPM 结构合理、具有一定的分类能力. 但是,在池化各个图像块的特征直方图时,SPM 忽略了图像块内的特征向量之间的空间信息. 串联特征直方图时,直方图之间固定的排列顺序也导致块间空间信息的利用过于单一. 因此,本文针对 SPM 空间信息利用不足的问题,分别从图像的局部与全局两个方面提取空间信息,编码成局部位置特征与全局轮廓特征. 然后,在 SPM 特征的基础上,串联入局部位置特征与全局轮廓特征,以此提高对空间信息的描述能力,从而提高分类正确率.

2 本文方法

本文方法整体流程如图 1 所示.

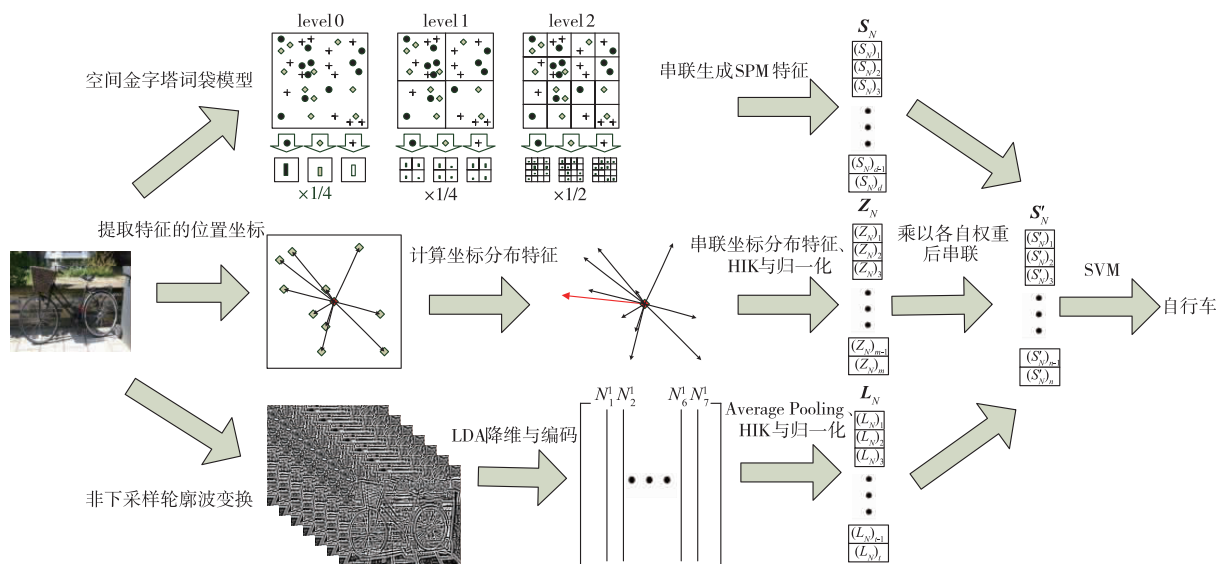


图1 本文方法流程示意图

2.1 局部位置特征

受到 VLAD 和文献[8]的启发,本文方法通过计算图像中每一个字典向量周围 Dense-SIFT 特征坐标的相对位置分布,来提取局部位置特征. 为了配合局部位置特征提取,在词袋模型特征编码过程中,本文使用 Localized Soft-assignment Coding (LSAC)^[4,9]对 Dense-SIFT 进行编码.

设特征 T 的坐标 $W = [w_1, w_2, \dots, w_m] \in \mathbb{R}^{2 \times m}$, $B_j = [b_1, b_2, b_3, b_4, b_5] \in \mathbb{R}^{d \times 5}$ 是与特征向量 t_j 欧氏距离最近的 5 个字典向量. 则经过 LSAC 编码后的特征 T' 如式(1)所示:

$$T' = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,m} \\ \vdots & \ddots & \vdots \\ \alpha_{n,1} & \cdots & \alpha_{n,m} \end{pmatrix} \quad (1)$$

每一列代表相应特征向量的视觉单词,其中系数 $\alpha_{i,j}$ 的计算公式如式(2)所示;其中 $\beta_i = \sigma_i^{-2}$, σ_i 是字典聚

$$\alpha_{i,j} = \begin{cases} \frac{\exp - \beta_i d_2(t_j, c_i)}{\sum_{r=1}^5 \exp - \beta_r d_2(t_j, b_r)}, & c_i \in B_j \\ 0, & c_i \notin B_j \end{cases} \quad (2)$$

类中心 c_i 的标准偏差, $d_2(t_j, c_i)$ 是特征向量 t_j 与字典向量 c_i 的欧氏距离. 在第 i 行中找出所有 $\alpha_{i,j} (i, j = 1, 2, \dots, m) \neq 0$ 的系数,其下标 j 与特征向量 t_j 的下标相对应,表示这些特征向量是由字典向量 c_i 所编码,且系数

为 $\alpha_{i,j}$. 图像 N 中, 由 c_i 编码的特征向量的位置坐标分布记为 $z_i (i = 1, 2, \dots, n)$. 因此, 本文利用字典向量 c_i 在视觉单词 j 中的系数和特征 t_j 的位置坐标来提取局部位置特征, 具体步骤如下:

(1) 找出所有由字典向量 c_i 编码的特征向量的位置坐标, 求出坐标平均值 p_i .

(2) 由式(3)求出每一个字典特征 c_i 下的坐标分布 z_i , 其中 w_j 表示由 c_i 编码的特征向量的位置坐标.

$$z_i = \sum_{j|\alpha_{ij} \neq 0} \alpha_{ij} \times (w_j - p_i) \quad (3)$$

(3) 将图像 N 中所有坐标分布按照字典向量的排列顺序串联得到局部位置特征 $Z_N = [z_1^T, z_2^T, \dots, z_n^T]^T$.

(4) 用 Histogram Intersection Kernel (HIK) [10] 对全部图像的局部位置特征 Z_N 进行编码, 使编码后的 Z_N 更具有可区分性, 并对其进行二范数归一化处理.

2.2 全局轮廓特征

首先, 本文方法使用非下采样轮廓波变换 (Nonsample Contourlet Transform, NSCT) [11] 来提取轮廓图像. NSCT 先把原图像分解为一个低频子带和一个高频子带, 然后把高频子带分解为多个方向的子带. 将这两个步骤在低频部分反复迭代, 直到达到要求的分解级数为止. 设 NSCT 的分解层数为 L , 第 $l = (1, 2, \dots, L)$ 层的分解级数的系数为 k_l , 则分解级数为 2^{k_l} , 可得到 2^{k_l} 个方向子带. 因此, 经过 NSCT 分解后可得到 $1 + \sum_{l=1}^L 2^{k_l}$ 个子带图像, 且分解后各级图像大小与原图像相同 [12], 图 2 为 NSCT 的结构和频域分解示意图.

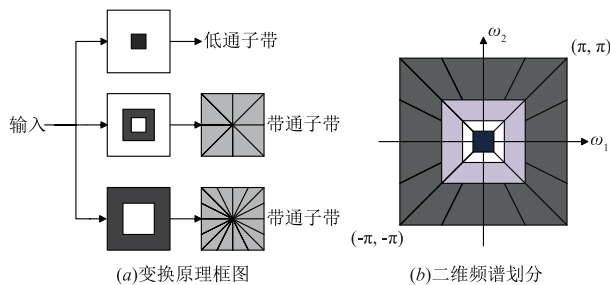


图2 非下采样轮廓波变换示意图

例如: 将 NSCT 的分解参数设置为 $[0, 1, 2]$, 最终一共生成 7 幅轮廓图像和 1 幅低通子带图像. 分解出的轮廓图像如图 3 所示, 这 7 幅图像分别由带通子带图像按不同方向分解而成, 足以表示原图像的轮廓信息. 实验中舍弃低通子带图像, 将第 $N (N = 1, 2, \dots, k)$ 幅图像分解、处理成 7 个轮廓特征向量. 把全部图像的轮廓特征按照不同的分辨率和方向组合成 7 大类, 每一类包含有同一分辨率和方向的全部轮廓特征, 表示为 $N_i (i = 1, 2, \dots, 7)$. 实验证明, 分解图像不超过 7 幅, 能够在时间较短的情况下得到较高的分类正确率, 具体分析将在实验部分给出.

在实验部分给出.

然后, 为了在降维的同时提高特征的可区分性, 本文方法用线性判别分析 (Linear Discriminant Analysis, LDA) [13, 14] 对 7 大类 $N_i (i = 1, 2, \dots, 7)$ 分别进行降维, 将结果表示为 $N_i^l (i = 1, 2, \dots, 7)$.

全局轮廓特征具体生成步骤如下所示:

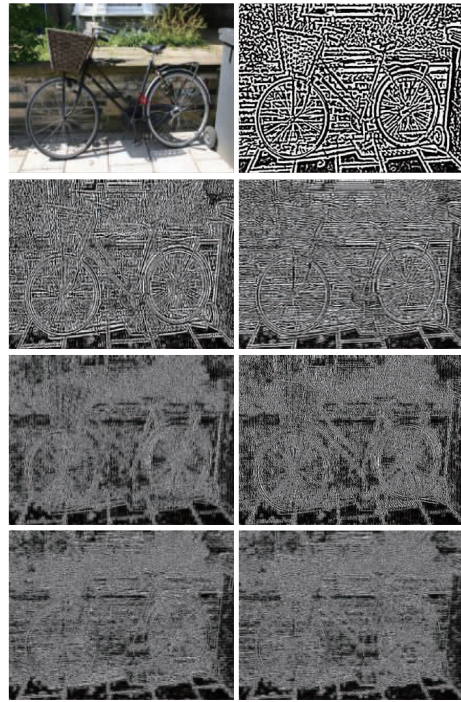


图3 7幅轮廓图像示意图

(1) 将第 N 幅图像缩放至统一尺寸, 然后对其进行轮廓特征变换, 使其分解为 7 幅轮廓图像.

(2) 将每幅轮廓图像按列处理为轮廓特征向量 $N_i (i = 1, 2, \dots, 7)$.

(3) 使用 LDA 将全部图像的 7 类 $N_i (i = 1, 2, \dots, 7)$ 分别降维成 $N_i^l (i = 1, 2, \dots, 7)$.

(4) 用 K-Means 对降维后的所有轮廓特征 $N_i^l (i = 1, 2, \dots, 7)$ 进行聚类, 生成轮廓特征字典, 然后用字典和 1-NN 方法将各个轮廓特征编码为轮廓视觉单词. 用 Average Pooling 将一幅图像的 7 个轮廓视觉单词池化为一个完整的轮廓视觉特征 L_N , 其维数与轮廓字典维数相同.

(5) 为了再次加强轮廓视觉特征的可区分性, 用 HIK 将其统一编码为最终的轮廓视觉特征 L_N , 并对其进行二范数归一化处理.

2.3 优化空间金字塔模型

设第 $N (N = 1, 2, \dots, k)$ 幅图像的 SPM 特征为 S_N , 局部位置特征 Z_N 与全局轮廓特征 L_N 的串联权重分别为 ε_z 和 ε_l , 图像的最终特征向量为 S'_N . 具体有如下步骤.

(1) 提取出全部图像的 Dense-SIFT 特征, 通过 K-means 生成字典. 使用字典和 LSAC 将全部图像的 Dense-SIFT 特征全部编码为视觉单词.

(2) 用 Max Pooling 把各个图像块的视觉单词池化成一个视觉特征向量, 其维数与字典维数相同. 然后把这些向量串联起来生成 SPM 特征 S_N , 并对其进行二范数归一化处理.

(3) 分别计算并生成各个图像的局部位置特征 Z_N 与全局轮廓特征 L_N , 详见 2.1 和 2.2.

(4) 分别将每一幅图像的 $S_N, \varepsilon_Z \cdot Z_N, \varepsilon_L \cdot L_N$ 这三个向量串联生成 S'_N , 如式(4)所示.

$$S'_N = [S_N^T, \varepsilon_Z Z_N^T, \varepsilon_L L_N^T]^T \quad (4)$$

(5) 用 LIBSVM^[15] 进行分类, 核函数使用的是自定义核函数 HIK. 除此之外, 其他步骤均与 SPM^[2] 一致.

3 实验结果与分析

本文分别在物体图像数据库 Caltech 101^[2]、MSRC^[16] 和场景图像数据库 15 Scene^[2] 上进行实验. 实验环境为 Intel (R) Core (TM) i5-4200M @ 2.50GHz 双核 CPU, 4.00GB RAM, MATLAB 2012a. 所有实验均参照文献[2]的实验步骤与参数进行, 其中 Dense-SIFT 的提取参数为单层 16×16 像素块、间隔 8 像素.

3.1 物体类别分类

3.1.1 Caltech 101 数据库

本实验使用 101 类物体图像, 参照文献[2]实验选取图像, 总共选中 5978 幅图像进行实验. 表 1 中 4 幅轮廓图像的 NSCT 分解参数为 $[0, 0, 1]$, 7 幅为 $[1, 0, 2]$. 由大量实验数据得出, 设置 NSCT 的分解参数使得最高频段包含的轮廓图像最多, 而其他各个频段包含的轮廓图像数目均衡时, 物体类别分类正确率最高. 此外, 使用 4 幅轮廓图像能够让本文方法在较短时间内获得较高的物体类别分类正确率, 因此, 本文使用 4 幅轮廓图像的实验结果与其他已发表的实验结果进行比较.

表 1 Caltech 101 结果比较

方法	ε_Z	ε_L	字典 维数	轮廓字 典维数	图像 轮廓数	时间 (分钟)	正确率 (%)
SPM	0	0	400	-	-	226	64.13
SPM + 局部位置	1	0	400	-	-	231	72.15
SPM + 全局轮廓	0	1	400	30	4	244	75.91
SPM + 全局轮廓	0	1	400	30	7	274	76.29
SPM + 局部 + 全局	1	1	400	30	4	249	76.9
SPM + 局部 + 全局	1	1	400	30	7	279	77.93

如表 2 所示, 在仅需较低字典维数的同时本文方法对物体图像的分类正确率确实得到了显著的提高.

图 4 中, 第一行和第二行分别列出了本文方法在实验中正确率较高的和较低的部分类别图像. 仔细分析实验结果并观察相应图像, 可以发现正确率高的类别图像具有背景不嘈杂, 并且物体本身轮廓清晰的特点. 如果换成背景复杂的物体图像本方法是否依然有效呢? 为了回答这个疑问, 本文在 MSRC 数据库上进行实验.

表 2 几种方法在 Caltech101 数据库上的实验结果

方法	ε_Z	ε_L	字典维数	轮廓字典维数	正确率 (%)
SPM ^[2]	-	-	200	-	64.60
SPS _{ad} ⁺ ^[5]	-	-	200	-	68.40
LLC ^[17]	-	-	2048	-	73.44
SLC ^[18]	-	-	2048	-	74.90
文献[7]	-	-	600	-	83.9
本文方法	1	1	400	30	76.9

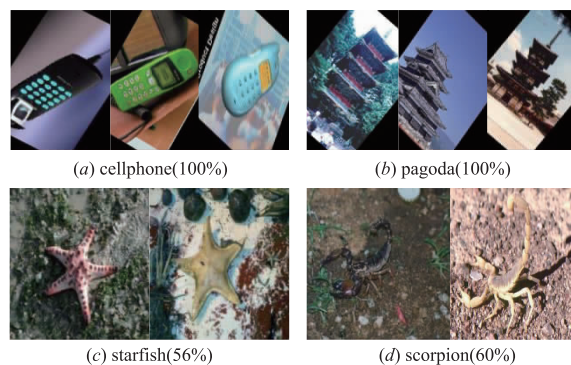


图 4 正确率较高与较低类别的部分图像

3.1.2 MSRC 数据库

MSRC 与其他数据库相比, 该数据库图像背景更复杂 (如图 5 所示), 因此, 该数据库的实验结果正好可以用来回答上面的疑问.



图 5 MSRC 数据库图像

本文选出 MSRC 的 18 个类别, 又从 MSRC-21^[6] 数据库选出了 3 个类别, 一共选择出 21 类物体图像进行实验, 其中每个类别随机选择 30 幅图像进行训练, 在剩下的图像中再选择 30 幅图像进行测试, 共使用 1260 幅图像进行实验. 如表 3 所示, 在对物体背景嘈杂的图像分类时, 嘈杂的背景会对整体轮廓的可区分度造成一定的影响. 因此, 单独加入全局轮廓特征的提升效果没有局部位置特征的提升效果显著.

表 3 MSRC 结果比较

方法	ε_Z	ε_L	字典 维数	轮廓字 典维数	图像 轮廓数	时间 (分钟)	正确率 (%)
SPM	0	0	400	-	-	98	87.61
SPM + 局部坐标	12	0	400	-	-	100	93.65
SPM + 全局轮廓	0	0.2	400	30	4	100	90.1
SPM + 全局轮廓	0	0.2	400	30	7	106	91.27
SPM + 局部 + 全局	12	0.2	400	30	4	102	93.97
SPM + 局部 + 全局	12	0.2	400	30	7	108	94.13

MSRC 大多用于图像分割与识别实验,与 MSRC-21 的相似性较高,因此表 4 中仅列出了一些方法在 MSRC-21 数据库上的实验结果.可以看出,即使在复杂的背景下,本文方法仍然具有较高的分类正确率.

表 4 几种方法的分类结果

方法	ε_Z	ε_L	字典维数	轮廓字典维数	正确率(%)
文献[6]	-	-	265	-	71
SPS _{ad} + [5]	-	-	400	-	83.50
本文方法	12	0.2	400	30	93.97

通过大量实验,我们总结出如下规律:对于背景简单的物体图像进行分类时,使得 ε_Z 与 ε_L 大致相同且不超过 3,可以得到较高的正确率.对于背景复杂的物体图像进行分类时,复杂的背景会对全局特征的利用造成一定的影响.因此,应加大局部特征的权重,使得 ε_Z 是 ε_L 的 15 倍及以上,且 ε_L 小于 0.5,可以得到较高的正确率.

3.2 场景类别分类

本文选用 15 Scene 数据库全部 15 种场景图像进行实验,从每一类图像中随机抽取 100 幅用于训练,剩余全部图像用于测试,总共使用 4485 幅图像进行实验.表 5 中,4 幅轮廓图像的 NSCT 分解参数为[0,0,1].

表 5 15 Scene 结果比较

方法	ε_Z	ε_L	字典 维数	轮廓字 典维数	图像 轮廓数	时间 (分钟)	正确率 (%)
SPM	0	0	400	-	-	81	80.1
SPM + 局部坐标	0.15	0	400	-	-	84	81.64
SPM + 全局轮廓	0	0.35	400	30	4	91	85.49
SPM + 全局轮廓	0	0.35	400	30	7	114	86.2
SPM + 局部 + 全局	0.15	0.35	400	30	4	94	85.79
SPM + 局部 + 全局	0.15	0.35	400	30	7	117	86.46

数据库各类场景之间可能包含有相似的局部物体,但是整体场景却不相似,如图 6 所示,上面 3 张是场景 kitchen,下面 3 张是场景 living room,可以看出这两类场景包含的物体非常相似.因此在场景分类中,局部位置特

征对 SPM 的提升效果没有轮廓特征的提升效果显著.



图 6 kitchen与living room场景图像

设置 NSCT 的分解参数使得各频段的轮廓图像数随着频率段的升高而增加时,场景类别分类正确率最高.同样,使用 4 幅轮廓图像能够让本文方法在较短时间内获得较好的场景图像分类结果,因此,使用 4 幅轮廓图像的实验结果与其他已发表的实验结果进行比较.如表 6 所示,在仅需较低字典维数的同时本文方法确实有效地提高了场景图像的分类正确率.通过大量实验总结出:对于场景图像进行分类时,局部特征比全局特征对正确率的影响更大.因此,应减小局部特征的权重,使得 ε_L 是 ε_Z 的 2 倍,且 ε_Z 与 ε_L 均小于 0.5,可以得到较高的正确率.

表 6 几种方法在 15 Scene 数据库上的实验结果

方法	ε_Z	ε_L	字典维数	轮廓字典维数	正确率(%)
SPM ^[2]	0	0	200	-	81.1
LLC ^[18]	-	-	1000	-	81.73
SLC ^[18]	-	-	1024	-	81.89
SPS _{ad} + [5]	-	-	200	-	83.70
文献[7]	-	-	600	-	90.1
本文方法	0.15	0.35	400	30	85.79

从两个数据库的实验结果表 2 和表 6 可以看出,文献[7]的正确率高于本文方法,是目前非 CNN 分类方法中正确率最高的方法,但该方法没有有效利用图像的空间信息.本文方法主要侧重于对图像空间信息的利用,从实验来看,本文方法有效利用了图像局部和全局的空间信息,在仅需较低字典维数的同时提高了场景和物体图像的分类正确率.

4 结论

本文分别提出了图像的局部位置特征与全局轮廓特征的概念,通过利用相似特征的局部位置信息和图像的全局轮廓信息,有效提高了空间信息的利用率,从而提高了场景和物体图像分类正确率.但是,随着数据

库的变换,局部位置特征与全局轮廓特征各自的最优串联权重值也在变换.因此,在今后的学习与研究中,本文方法还可以在最优串联权重取值不统一的问题上继续进行改进.同时,我们也会将这两个空间特征与更具辨识度的视觉特征相结合,以进一步提高算法性能.

参考文献

- [1] Csurka G, Dance C R, Fan L, et al. Visual categorization with bags of keypoints[J]. Workshop on Statistical Learning in Computer Vision Eccv, 2004, 44(247): 1–22.
- [2] Lazebnik S, Schmid C, Ponce J. Beyond bags of features: spatial pyramid matching for recognizing natural scene categories[A]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition [C]. New York, USA: IEEE, 2006. 2169–2178.
- [3] Jégou H, Douze M, Schmid C, et al. Aggregating local descriptors into a compact image representation[A]. IEEE Conference on Computer Vision and Pattern Recognition [C]. San Francisco, USA: IEEE, 2010. 3304–3311.
- [4] Avila S, Thome N, et al. Pooling in image representation: the visual codeword point of view[J]. Computer Vision and Image Understanding, 2013, 117(5): 453–465.
- [5] Khan R, Barat C, Muselet D, et al. Spatial histograms of soft pairwise similar patches to improve the bag-of-visual-words model[J]. Computer Vision & Image Understanding, 2014, 132: 102–112.
- [6] Li W S, Dong P, Xiao B, et al. Object recognition based on the region of interest and optimal bag of words model[J]. Neurocomputing, 2015, 172(8): 271–280.
- [7] Koniusz P, Yan F, Gosselin P, et al. Higher-order occurrence pooling for bags-of-words: visual concept detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(2): 313–326.
- [8] 艾列富. 基于内容的大规模图像索引与检索方法研究[D]. 湖北武汉: 华中科技大学, 2014.
AI Lie-fu. Research on Content-based Indexing and Retrieval Methods for Large Scale Images[D]. Hubei, Wuhan: Huazhong University of Science & Technology, 2014. (in Chinese)
- [9] Liu L Q, et al. In defense of soft-assignment coding[A]. Proceedings of the International Conference on Computer Vision [C]. Barcelona, Spain, 2011. 2486–2493.
- [10] Grauman K, Darrell T. Pyramid match kernels: discriminative classification with sets of image features[A]. Tenth IEEE International Conference on Computer Vision [C]. Beijing, China: IEEE, 2005. 1458–1465.
- [11] Cunha A L D, et al. The nonsubsampling contourlet transform: theory, design, and applications[J]. IEEE Transactions on Image Processing, 2006, 15(10): 3089–3101.
- [12] 倪伟. 基于多尺度几何分析的图像处理技术研究[D]. 陕西西安: 西安电子科技大学, 2008.
NI Wei. Research on Image Processing Algorithms via a Multiscale Geometric Analysis [D]. Xi'an, Shaanxi: Xidian University, 2008. (in Chinese)
- [13] 何国辉, 甘俊英. PCA-LDA 算法在性别鉴别中的应用[J]. 计算机工程, 2006, 32(19): 208–210.
HE Guo-hui, GAN Jun-ying. Application of PCA and LDA on gender classification[J]. Computer Engineering, 2006, 32(19): 208–210. (in Chinese)
- [14] 汤鹏. 基于 LDA 的特征提取及其在人脸识别中的应用[D]. 河北保定: 河北大学, 2015.
TANG Peng. LDA Based Feature Extraction and Its Application To Face Recognition [D]. Baoding, Hebei: Hebei University, 2015. (in Chinese)
- [15] Chang C C, Lin C J. LIBSVM: A library for support vector machines[J]. Acm Transactions on Intelligent Systems & Technology, 2011, 2(3): 389–396.
- [16] Microsoft Research Cambridge Object Recognition Image Database [DB]. <https://www.microsoft.com/en-us/download/details.aspx?id=52644>, 2018.
- [17] Wang J Y, Yang J C, Yu K, et al. Locality-constrained linear coding for image classification[A]. IEEE Conference on Computer Vision and Pattern Recognition [C]. San Francisco, USA, 2010. 3360–3367.
- [18] 肖文华, 包卫东, 等. 一种用于图像分类的语义增强线性编码方法[J]. 电子与信息学报, 2015, 37(4): 791–797.
XIAO Wen-hua, BAO Wei-dong, et al. A semantic enhanced linear coding for image classification[J]. Journal of Electronics & Information Technology, 2015, 37(4): 791–797. (in Chinese)

作者简介



李雅倩 女. 1982年4月出生, 湖南岳阳人. 2010年毕业于法国鲁昂应用科学学院信息学专业, 现为燕山大学电气工程学院副教授, 研究方向为模式识别.
Email: yaqianli@126.com



吴超 (通信作者) 男. 1990年10月出生, 河南安阳人. 燕山大学在读博士研究生, 主要研究方向为机器学习与模式识别.
Email: yddqb410@163.com