

# 基于 BoC-BoF 特征的图像检索方法研究

冯进丽<sup>1</sup> 杨红菊<sup>1,2</sup>

(山西大学计算机与信息技术学院 太原 030006)<sup>1</sup>

(山西大学计算智能与中文信息处理教育部重点实验室 太原 030006)<sup>2</sup>

**摘 要** 为了优化基于内容的图像检索方法,提出了一种融合特征来表征图像内容。首先,提取基于 RootSift 描述子的特征词袋(Bag-of-Features, BoF)表示向量,获得图像的边缘和形状信息;其次,采用基于 HSV 的颜色词袋(Bag-of-Colors, BoC)表示向量来代替传统颜色直方图方法,获取图像的颜色信息;最后,将 BoF 表示向量和 BoC 表示向量相融合,形成 BoC-BoF 特征向量。BoC-BoF 特征有效地实现了全局特征和局部特征的融合。两个数据集检索的实验结果表明,该方法比其它方法更加有效。

**关键词** 颜色词袋,特征词袋,图像检索,RootSift

**中图分类号** TP391.41 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2015.4.061

## Image Retrieval Method Research Based on BoC-BoF Feature

FENG Jin-li<sup>1</sup> YANG Hong-ju<sup>1,2</sup>

(School of Computer and Information Technology, Shanxi University, Taiyuan 030006, China)<sup>1</sup>

(Key Laboratory of Computational Intelligence & Chinese Information Processing of  
Ministry of Education, Shanxi University, Taiyuan 030006, China)<sup>2</sup>

**Abstract** The fusion feature for representing content of images was investigated in order to optimize content-based image retrieval method. Firstly, RootSift-based Bag-of-Features (BoF) were extracted, which capture shape and edge information. Then, Bag-of-Colors (BoC) based on HSV were adopted to replace the traditional color histogram quantization method was adopted, which capture color information. Lastly, BoC-BoF algorithm which integrates BoC vectors and BoF vectors was proposed. BoC-BoF algorithm effectively realizes the integration of global features and local features. The obtained impressive results show that this algorithm is more effective than other methods in two datasets of this paper.

**Keywords** Bag-of-colors, Bag-of-features, Image retrieval, RootSift

## 1 引言

随着互联网技术的发展,网络上存放的图像文档以及相关资料逐年增加,目前已有数以亿计的网络图片。如何在如此大规模的图像数据库中快速准确地检索和定位到用户所要的图片,已经成为计算机视觉领域的一个重要问题。鉴于这一现象,基于内容的图像检索方法受到了广泛的关注<sup>[1]</sup>。图像的尺度变化、光照变化、视角差异、背景杂乱以及目标遮挡等现象给图像检索带来巨大困难。寻找一种鲁棒性好的图像特征成为图像检索的首要任务。SIFT<sup>[2]</sup>是 David G. Lowe 提出的一种对图像的尺度、光照、旋转以及噪声干扰等表现出较好鲁棒性的特征。直接通过 SIFT 匹配算法<sup>[2]</sup>实现图像检索,当数据库图片数量增多时,特征匹配所花时间急剧增加,无法满足用户的时间要求。2003 年 Josef Sivic 等<sup>[3]</sup>借鉴文本信息检索的 Bag-of-Words 模型提出了 BoF 表示向量,在一定程度上消除了直接应用局部特征存储代价大、匹配算法复杂

度较高的缺陷。首先把训练集中的 SIFT 描述子进行 K-means 聚类,簇集中心定义为视觉单词;然后把每幅图像的 SIFT 描述子量化到视觉单词上,通过 TF-IDF 表示对查询图片和数据库中的图片进行相似度评测。

近几年,基于 SIFT 描述子的 BoF 模型在图像检索中被广泛应用<sup>[4,5]</sup>。SIFT 通过统计图像梯度分布可以捕捉局部物体的形状和边缘信息,但对于背景复杂、噪声干扰明显的图像,提取到的特征包含太多干扰信息,因此该方法比较适用于噪声干扰少、背景简单的图像库。近年来大量学者提出将传统的底层特征(如颜色、纹理和形状等)与中层特征 SIFT 融合的方法。文献<sup>[6]</sup>提出 SIFT-LBP 实现了底层纹理特征与 SIFT 特征的融合,使得检索准确率有了显著提高。文献<sup>[7]</sup>提出 RGB-SIFT,首先提取 R、G、B 3 个通道,然后分别计算每个通道的 128 维 SIFT 描述子,再通过级联得到 384 维 RGB-SIFT 描述子,该方法综合考虑了颜色特征与 SIFT 特征,提高了特征描述力。

到稿日期:2014-05-27 返修日期:2014-07-25 本文受国家自然科学基金(61201453),山西省自然科学基金(2012011014-4),山西省教育厅专项(20110002)资助。

冯进丽(1986—),女,硕士生,主要研究方向为计算机视觉与模式识别,E-mail:fengjinlifengdou@163.com;杨红菊(1975—),女,博士,副教授,主要研究方向为模式识别、图像检索和图像处理。

文献[8]提出基于 Lab 空间的颜色词袋(BoC)方法来表征图像的全局颜色特征。BoC 表示方法消除了传统颜色直方图量化范围大、对于颜色较一致的图像库检索效果差的缺陷。文献[9]提出了一种 SIFT 的改进算法 RootSift,该算法在用欧氏距离进行相似度比较中呈现良好性能。由于 HSV 空间能更加准确地反映人类视觉对色彩的理解,而且不同颜色在 HSV 颜色空间中的欧氏距离和视觉距离大体一致,因此本文给出了一种新的基于 HSV 空间的 BoC 构造方法,并将其与基于 RootSift<sup>[9]</sup>的 BoF 表示向量相融合,得到 BoC-BoF 融合特征。该特征从颜色、纹理、形状 3 方面对图像进行描述,实现了全局特征与局部特征的融合。实验结果表明,基于 HSV 空间的 BoC 特征向量比传统颜色直方图检索性能更好,BoC-BoF 算法与其它算法相比在图像检索中拥有较高检索准确率。

## 2 基于 BoC-BoF 的图像检索

### 2.1 BoF 向量表示

特征词袋(BoF)模型借鉴文本信息处理的词包(Bag-of-Words)算法,将图像表示成视觉关键词统计直方图。所谓视觉关键词,就是图像局部区域特征(如纹理、兴趣点描述子)经过聚类形成的聚类中心。一幅图像的视觉关键词的集合就是特征词袋(BoF)。构建特征词袋过程如图 1 所示。

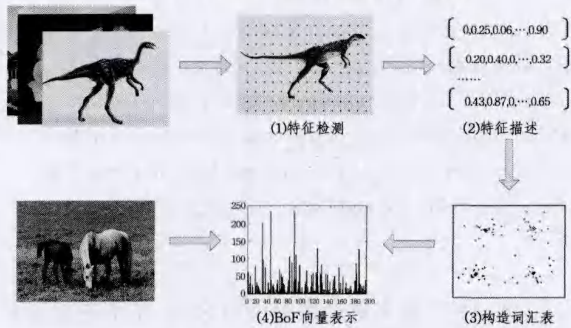


图 1 特征词袋构建流程

(1)特征检测。特征检测即在图像中检测稳定特征点或区域,并确定其位置、方向、尺度及其它仿射变换参数。特征检测分为稀疏检测<sup>[3]</sup>和稠密检测<sup>[10,11]</sup>。稀疏检测是指选取易于定位的角点或视觉上突出的区域作为特征点,如 DoG 和 MSER。稠密采样是指随机<sup>[11]</sup>或按照某种规则<sup>[5]</sup>从图像中抽取大量的点作为特征点。本文选取稠密检测和稀疏检测两种方法进行实验。稠密检测实现方法:像素间隔为 8 的像素点作为特征点,如图 1 所示。稀疏检测实现方法:建立高斯差分金字塔(DoG),求取每一层的极值点作为特征点<sup>[2]</sup>。

(2)特征描述。在特征检测的特征点周围提取一块图像区域,随后从该图像块抽取特征得到特征点的描述向量。本文采用 RootSift<sup>[9]</sup>描述子作为描述向量。首先在特征点周围取一个  $4 \times 4$  的棋盘格状网格,每个网格里有  $4 \times 4$  个像素;然后计算每个网格中的像素点的梯度,统计得到 8 柄梯度方向直方图;最后将 16 个直方图级联起来构建 128 维 SIFT 向量,根据文献[9]对 SIFT 向量 L1-规范化并求平方得到 RootSift 描述子。

(3)构建视觉词汇表。完成步骤(2)以后,若不考虑空间

信息,一幅图像可视为一个 RootSift 描述子的集合。通过 K-均值聚类将图像库所有的 RootSift 描述子进行分组,簇集中心便是构建的视觉词汇。

(4)图片的特征向量表示。假设由步骤(3)得到  $k$  个视觉词汇,词汇表表示为  $F = \{F_1, F_2, \dots, F_k\}$ 。取图片  $D_j$  的 RootSift 集合,通过近邻匹配将每个 RootSift 用离它最近的视觉词汇表示,统计视觉词率表示该幅图像。为了避免相似度计算时太大词频值权重大的问题,采用取平方的方法将词频值进行拉伸。实验表明,这一简单的处理使得检索准确率得到了很大的改善。通过上述步骤,图片  $D_j$  便可表示为  $D_j = [\omega_{1,j}, \omega_{2,j}, \dots, \omega_{k,j}]$ ,  $\omega_{i,j}$  表示数据库中第  $j$  张图片的词汇  $F_i$  的权重。

假设检索图像库有  $N$  幅图像,只需一个矩阵便可表示该图像库的所有图片:

$$\begin{bmatrix} \omega_{1,1} & \omega_{2,1} & \dots & \omega_{k,1} \\ \omega_{1,2} & \omega_{2,2} & \dots & \omega_{k,2} \\ \vdots & \vdots & \vdots & \vdots \\ \omega_{1,N} & \omega_{2,N} & \dots & \omega_{k,N} \end{bmatrix}$$

### 2.2 BoC 向量表示

颜色直方图是使用最为广泛的颜色特征,其实现方法为:首先将颜色空间量化成若干个颜色区间,每个区间成为颜色直方图的一个柄,然后统计落在每个柄的像素数量,最后进行归一化处理得到颜色直方图。这样的方法简单方便,但存在两个弊端:①由于离散的量化级方式,位于量化边界上或边界附近的灰度值只能归入到一个量化级,这就割裂了相近颜色的相关性和连续性,引起量化级误差。②量化范围比较大,对于颜色较一致的图像库的检索效果必然很差。为了解决这些问题,本文提出了基于 HSV 空间的颜色词袋(BoC)算法。该算法由无监督的机器学习方法 K-means 对检索图像库所有 HSV 值进行聚类训练,从而得到颜色调色板,然后通过统计颜色调色板每一种颜色数量来构建 BoC 特征向量。BoC 特征向量实现步骤如下:

(1)被检索图像库统一化预处理。将图像库所有图像转化为 HSV 空间,同时把图像大小统一为  $256 \times 256$  像素。

(2)提取图像块主颜色。首先将图像分割成  $16 \times 16$  像素的图像块,即每一幅图像由 256 个图像块组成;然后计算每一个图像块,HSV 众数值作为对应图像块主颜色;最后收集图像的主颜色,一幅图像由 256 个主颜色组成。

(3)K-means 聚类构建颜色调色板。假设图像库有  $N$  幅图像,由步骤(2)提取得到  $m = 256 \times N$  个主颜色值,通过 K-means 聚类方法训练这些颜色值得到  $t$  个聚类中心。

Step 1(初始化) 随机选取  $t$  个主颜色值作为聚类质心点  $(\mu_1, \mu_2, \dots, \mu_t)$ ;

Step 2(分配归类) 通过式(1)训练每一个主颜色  $x_i$  ( $i = 1, 2, \dots, m$ ) 所属类,其中  $c_i$  表示主颜色  $i$  与  $t$  个聚类质心距离最近的那一类,  $c_i$  的取值为 1 到  $t$  中的一个;

$$c_i = \arg \min_j \|x_i - \mu_j\|^2 \quad (1)$$

Step 3(修正聚类中心) 通过式(2)修正簇类中心;

$$\mu_j = \frac{\sum_{i=1}^m 1\{c_i=j\} x_i}{\sum_{i=1}^m 1\{c_i=j\}} \quad (2)$$

Step 4(迭代终止判断) 当前后两次聚类中心不再变化时,得到的  $t$  个聚类中心便是颜色词汇,即颜色调色板。图 2 所示为训练 Corel 数据库<sup>[12]</sup> 1000 幅图像  $t$  取 16、32、64 得到的 BoC 颜色调色板。

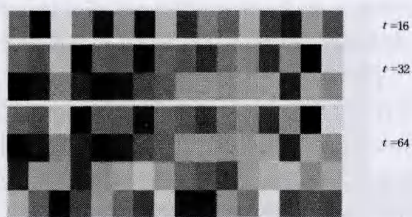


图 2 Corel 数据库训练得到的 BoC 调色板

(4) BoC 特征向量表示。首先提取所要表示图片的全部像素 HSV 值;然后将每个像素的 HSV 值映射为距离最近的颜色词汇;其次统计该图颜色词汇词频,得到颜色词汇直方图;最后为了降低频率较大颜色词汇在一幅图的 BoC 表示中的权重,同时突出频率低且有代表性的颜色词汇,进一步对得到的颜色词汇统计直方图进行开方归一化处理,最终得到的向量即为 BoC 特征向量。

### 2.3 BoC-BoF 特征融合

BoC 特征向量虽然很好地统计了图像颜色全局分布,但缺少空间分布信息。如图 3 所示,颜色空间分布明显不相同的图像具有相同的 BoC 特征向量,因此很明显单一利用 BoC 特征向量进行检索,检索效果会与人的视觉感受不相吻合。RootSift 通过统计图像梯度分布捕捉局部物体的形状和边缘信息,但容易受背景和噪声的干扰。针对这两个缺陷,本文将一定程度上互补的 BoC 特征向量和 BoF 特征向量相融合来提高图像特征描述力。由于 BoF 特征向量是基于局部特征的统计量,BoC 特征向量是基于全局特征的统计量,因此该融合巧妙地实现了全局特征和局部特征融合,使得融合后的特征具有更佳描述力和区分力。

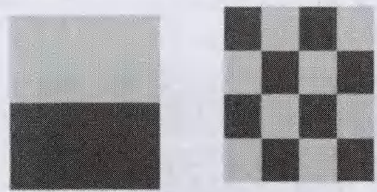


图 3 颜色特征相同空间分布不同的两幅图

由于两种算法得到的视觉特征的物理意义不同,于是取值范围差别比较大,相似距离可比性差,因此把特征直接相加的处理方法是不恰当的,也不能达到特征融合的最终目标,从而失去了多特征检索意义。为了解决不同特征之间的差异性,在利用综合特征进行图像检索时,需要对不同特征向量进行外部归一化处理<sup>[13]</sup>,目的是确保不同特征向量在融合加权时处于平等的地位。

外部特征归一化的本质是对图像库中所有图像之间的相似距离进行归一化,其实现过程如下:

(1) 计算图像库(图像数量为  $M$ )中每对图像  $I, J$  的特征向量  $F_I, F_J$  的相似距离:

$$D(I, J) = \text{dis}(F_I, F_J), I, J = 1, 2, \dots, M, \text{且 } I \neq J \quad (3)$$

(2) 计算由步骤(1)得到的  $M(M-1)/2$  个距离值的均值  $M_D$  和标准差  $\delta_D$ 。

(3) 对于样本图像  $Q$ , 计算其与图像库中所有图像的相似距离并记为  $D_{1Q}, D_{2Q}, \dots, D_{MQ}$ 。

(4)  $D_{1Q}, D_{2Q}, \dots, D_{MQ}$  先进行高斯归一化,再作线性变换:

$$D_{iQ} = \left[ \frac{D_{iQ} - M_D}{3\delta_D} + 1 \right] / 2 \quad (4)$$

归一化的值  $D_{iQ}$  几乎全部都落到了  $[0, 1]$  区间上。

对 BoC 表示向量和 BoF 表示向量进行外部归一化处理,确保了融合检索时两种特征向量在相似度计算时的地位相同。通过结合 BoC 特征向量相似距离和 BoF 特征向量相似距离,将基于 BoC 特征向量的图像检索与基于 BoF 特征向量的图像检索融合为基于 BoC-BoF 特征的图像检索。设  $q$  为查询图像,  $p$  为数据库中一幅图像,  $D_c$  代表基于 BoC 特征外部归一化的  $q$  与  $p$  的相似距离,  $D_f$  代表基于 BoF 特征外部归一化的  $q$  与  $p$  相似距离,则两幅图像间的加权融合相似距离计算如下:

$$D(q, p) = \omega_c D_c + \omega_f D_f, \omega_c + \omega_f = 1 \quad (5)$$

其中,  $\omega_c$  和  $\omega_f$  分别表示 BoC 特征向量和 BoF 特征向量的权重。通过计算数据库中每幅图像与查询图像的加权融合相似距离,按距离由小到大排序,将前若干幅图像作为检索结果返回。

## 3 实验与结果分析

基于内容的图像检索大致可以分为两大类:全局相似性图像检索和局部相似性图像检索。全局相似性检索的目标是检索与查询图像在整体属性上相似的图像,如同样的蓝色背景、同样的大片草地、同样有大部分建筑物或者具有相似布局的图像。局部相似图像检索又称为“物体检索”,目的是找出与查询图像具有相似局部内容的图像。例如,找出与查询图像包含同一栋大楼、同一本书、同一个人脸的图像。本文选取两个基准数据库来验证本文所提方法对两种类型的基于内容的图像检索的适用性。

### 3.1 全局相似性图像检索

本文在 Windows 2007 操作系统上使用 MATLAB 2013a 编程平台进行实验。使用全局相似性图像检索基准数据库 Corel<sup>[12]</sup> 对提出的算法进行验证。Corel 数据库包含 10 类图像(非洲、沙滩、建筑、车、恐龙、大象、花儿、马、雪山、食物),每一类包含 100 张图像,共 1000 张图像,图像的尺寸大小为  $256 \times 384$  或  $384 \times 256$  个像素。分别用如下 3 种评价方法对文中所提算法进行验证。

(1) ARP 评价

平均检索查准率(Average Retrieval Precision, ARP)<sup>[6]</sup> 是一种有效的验证检索效果优劣的性能评价标准。假设查询图像  $i$  的检索返回图像数量是  $B(i)$ , 返回图像与查询图像  $i$  属于同一类的数量为  $A(i)$ , 则查询图像  $i$  的查准率  $P(i)$  大小为:

$$P(i) = \frac{|A(i) \cap B(i)|}{|B(i)|} \quad (6)$$

平均检索查准率(ARP)为同一类别所有图像作为查询图像所得查准率的平均值,其定义如下:

$$ARP(ID_m) = \frac{1}{N_{id(i)=ID_m}} \sum P(i) \quad (7)$$

其中,  $ID_m$  ( $m$  为类别索引) 是图像类别,  $id(i)$  是查询图像  $i$  所属类别索引,  $N$  是属于类别  $ID_m$  的图像数量。在本文实验中, 检索返回数量  $B(i)$  取 100, 恰好为每一类别所含有的图像数  $N$ 。假设某一类别 ARP 为 0.5 表示将属于该类的所有图像分别作为查询图像返回 100 幅图片, 平均检索正确的图片数量为 50 幅。

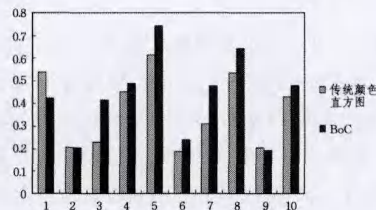


图 4 BoC 与传统颜色直方图的 ARP 对比图

为了说明 BoC 颜色特征优于传统颜色量化方法, 本文将 256 维 BoC 特征与 (16; 4; 4) 非均匀量化 HSV 颜色空间的

256 维传统颜色直方图<sup>[14]</sup> 进行检索比较, 结果如图 4 所示。由图可见, 除了 1、2、9 类即非洲、沙滩、雪山的 ARP 值稍微低于传统颜色直方图外, 其它类别的 ARP 值远远高于传统颜色直方图。

为验证 BoC-BoF 特征的有效性, 本文将其与近几年在图像检索、分类实验应用广泛的两种特征进行比较说明, 结果如表 1 所列。同时本文还探索了密集采样和稀疏采样对全局相似性图像检索的适用性。其中 BoF 与 BoC 的词汇数量分别为 200、256,  $w_c$  取 0.8。分析表 1: ①由第 3、4 列可知, 在全局相似性图像库检索中密集采样明显优于稀疏采样; ②由第 1、2、6 列可知, BoC-BoF 的平均 ARP 值高于文献[5, 16]的特征, 这表明 BoC-BoF 是一种比较有效的图像特征, 在全局相似性图像检索中呈现良好的性能; ③文献[5, 16]的特征维数都是上千维, 而 BoC-BoF 的特征维数只为  $200 + 256 = 456$  维, 所以在相似度计算时 BoC-BoF 的速度远远快于文献[5, 16]的特征的速度。

表 1 本文方法与传统方法的 ARP 值

类别	SPM (文献[5])	Phog-Phow (文献[16])	BoF (密集采样)	BoF (DoG 采样)	BoC-BoF (DoG 采样)	BoC-BoF (密集采样)
非洲	0.3524	0.3562	0.4048	0.3633	0.4410	0.5164
沙滩	0.3208	0.3182	0.3081	0.1020	0.2061	0.3016
建筑	0.2657	0.2667	0.2631	0.2344	0.4142	0.3441
车	0.7498	0.7519	0.7883	0.4105	0.5136	0.8117
恐龙	0.9428	0.9421	0.8691	0.4227	0.7441	0.9077
大象	0.4299	0.4330	0.4280	0.3552	0.2643	0.3908
花	0.7906	0.7882	0.7156	0.1462	0.4299	0.7132
马	0.6936	0.6914	0.5748	0.4550	0.6626	0.6720
雪山	0.3176	0.3145	0.2855	0.1613	0.2125	0.2831
食物	0.3411	0.3459	0.3645	0.2389	0.4882	0.5178
平均 ARP	0.5204	0.5208	0.5002	0.2889	0.4376	0.5458

### (2) 查全率-查准率评价

ARP 从查询返回前 100 幅图像准确率的角度体现图像检索效果, 缺乏全面性。为了突出全面性, 本文同时采用查全率-查准率进行实验验证。实验方法: 取每类图像的前 10 幅图像作为查询图像, 共 100 幅, 将查准率、查全率的平均值作为结果。图 5 给出了传统颜色直方图、BoC、BoF、BoC-BoF 4 种算法的查全率-查准率。由图 5 可知: ①在查全率相同时, BoC 比传统颜色直方图拥有更高的查准率, 这表明 BoC 特征描述力优于传统颜色直方图; ②本文提出的 BoC-BoF 算法较单一 BoC 或 BoF 特征拥有更高的图像检索准确率。

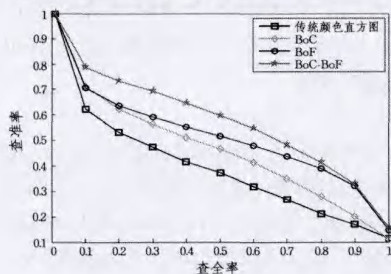


图 5 4 种算法的查全率-查准率的对比图

### (3) 可视化评价

仔细比对 BoC、BoF、BoC-BoF 3 种算法的前 6 幅检索返回结果, 从中选取一组较直观的结果进行说明, 如图 6 所示。通过比较发现, BoC 算法、BoC-BoF 算法都得到了 100% 的准

准确率, 但 BoC-BoF 算法更符合查询图像的空间布局, 即返回的图像前景都是以非洲人头像为主的空间格局。另外, 显而易见 BoF 检索结果有 4 幅不正确, 非常不理想, 但是经过融合后 BoC-BoF 算法实现了颜色-纹理-形状三位一体的特征描述, 返回的结果特别理想。



图 6 3 种算法的可视化检索结果对比

## 3.2 局部相似性图像检索

本小节选用 Kentucky 大学的图像检索数据库<sup>[15]</sup> 作为基准数据库, 该数据库大约一万张图像, 其中每 4 张图像构成一类, 分别是同一个物体在不同光照和视角变化下的照片, 如图 7 所示。由于实验平台限制, 从中抽取 200 类 800 张图像作为本文算法评估数据库。采用文献[15]中提出的评测准则“Top-4-Score”, 即让图像库中每一幅图像进行检索, 返回前 4 张图像中正确结果的平均数量, 当 Top-4-Score 等于 4 时结果最理想。

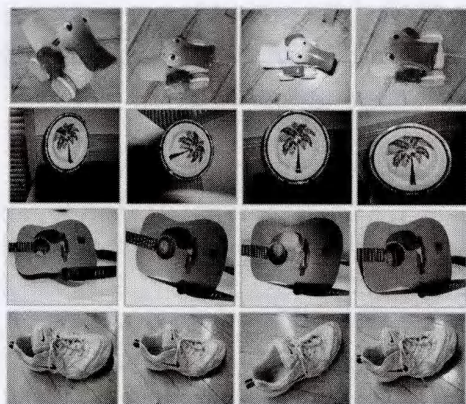


图7 Kentucky大学的图像检索数据库图像示例

表2 5种算法的 Top-4-Score 值

	BoC	BoF (DoG 采样)	BoF (密集采样)	SPM (文献[5])	Phog-Phow (文献[16])	BoC-BoF (DoG 采样)	BoC-BoF (密集采样)
Top-4-Score	3.1913	2.6325	2.4162	3.0488	3.0500	3.3875	3.3163

**结束语** 本文的贡献可归纳为3点:1)通过实验证明密集采样适用于全局相似性图像检索,稀疏采样适用于局部相似性图像检索。2)提出了基于HSV空间的BoC表示方法来代替传统的颜色量化方法,BoC模型消除了量化范围大、在颜色较一致的图像库检索效果差的缺陷。由实验结果可见,BoC向量的检索性能高于传统的颜色量化模型。3)提出了BoC-BoF算法,基于该算法的图像检索在一定程度上改善了BoF向量只适应于噪声干扰少、背景简单的图像库检索的弊端。下一步考虑提升实验平台,将该方法应用于更大规模的图像库,并学习更多的方法优化BoF算法和BoC算法。

### 参考文献

[1] Datta R, Joshi D, Li J, et al. Image retrieval: Ideas, influences, and trends of the new age [J]. ACM Computing Surveys (CSUR), 2008, 40(2): 1-60

[2] Lowe D G. Distinctive image features from scale-invariant key points [J]. International journal of computer vision, 2004, 60(2): 91-110

[3] Sivic J, Zisserman A. Video Google: A text retrieval approach to object matching in videos [C]//Ninth IEEE International Conference on Computer Vision, 2003. IEEE, 2003: 1470-1477

[4] O'Hara S, Draper B A. Introduction to the bag of features paradigm for image classification and retrieval [J]. arXiv preprint arXiv: 1101. 3354, 2011: 1-25

[5] Lazebnik S, Schmid C, Ponce J. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories [C]// 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, 2006, 2: 2169-2178

[6] Yu J, Qin Z, Wan T, et al. Feature integration analysis of bag-of-features model for image retrieval [J]. Neurocomputing, 2013, 120: 355-364

BoC、BoF、BoC-BoF以及文献[5, 16]5种算法的 Top-4-Score 实验结果如表2所列,其中BoF分别采用DoG稀疏采样和密集采样两种特征点提取方法,BoC、BoF的词汇数量为256、200,  $\alpha_i$ 取0.3。表2中:①由第2、3列可见,在局部相似性图像库检索中稀疏采样明显优于密集采样。②由第4-7列可知,BoC-BoF的平均ARP值高于文献[5, 16]提出的特征,表明BoC-BoF在局部相似性图像检索中也表现出良好性能。

通过全局相似性图像检索和局部相似性图像检索实验数据的综合比较,可以发现:密集采样包含了丰富的前景与背景信息,较适用于全局相似性图像检索;稀疏采样如DoG,突出了图像前景的边缘轮廓信息,在局部相似性检索即物体检索中表现出较高的检索准确率。

[7] Van De Sande K E A, Gevers T, Snoek C G M. Evaluating color descriptors for object and scene recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(9): 1582-1596

[8] Wengert C, Douze M, Jégou H. Bag-of-colors for improved image search [C]// Proceedings of the 19th ACM International Conference on Multimedia. ACM, 2011: 1437-1440

[9] Arandjelovic R, Zisserman A. Three things everyone should know to improve object retrieval [C]// 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2012: 2911-2918

[10] Jurie F, Triggs B. Creating efficient codebooks for visual recognition [C]// Tenth IEEE International Conference on Computer Vision, 2005 (ICCV 2005). IEEE, 2005, 1: 604-610

[11] Nowak E, Jurie F, Triggs B. Sampling strategies for bag-of-features image classification [C]// Computer Vision-ECCV 2006. Springer Berlin Heidelberg, 2006: 490-503

[12] Wang J Z, Li J, Wiederhold G. SIMPLiCity: Semantics-sensitive integrated matching for picture libraries [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23(9): 947-963

[13] 刘忠伟, 章毓晋. 综合利用颜色和纹理特征的图像检索 [J]. 通信学报, 1999, 20(5): 36-40

[14] 陈秀新, 贾克斌. 三维量化颜色直方图在彩色图像检索中的应用 [J]. 计算机应用与软件, 2012, 29(9): 31-32

[15] Nister D, Stewenius H. Scalable recognition with a vocabulary tree [C]// 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, 2006, 2: 2161-2168

[16] Bosch A, Zisserman A, Munoz X. Image classification using random forests and ferns [C]// 2007 IEEE 11th International Conference on Computer Vision. Riode Janeiro, Brazil: IEEE Press, 2007: 1-8