

# 基于多视觉词典的显著性加权图像检索方法

孔超<sup>1,2,3</sup> 张化祥<sup>1,3</sup> 生海迪<sup>1,3</sup>

(1. 山东师范大学信息科学与工程学院, 济南, 250014; 2. 国网技术学院, 济南, 250002; 3. 山东省分布式计算机软件新技术重点实验室, 济南, 250014)

**摘要:** 针对视觉词典在图像表示与检索方面的应用需求, 本文提出了一种基于多视觉词典与显著性加权相结合的图像检索方法, 实现了图像多特征的显著性稀疏表示。该方法首先划分图像为小块, 提取图像块的多种底层特征, 然后将其作为输入向量, 通过非负稀疏编码分别学习图像块多种特征对应的视觉词典, 将得到的图像块稀疏向量经过显著性汇总方法引入空间信息并作显著性加权处理, 形成整幅图像的稀疏表示, 最后采用提出的 SDD 距离计算方式进行图像检索。在 Corel 和 Caltech 通用图像集上进行仿真实验, 与单一视觉词典的方法对比, 结果表明本文方法能够有效提高图像检索的准确率。

**关键词:** 多视觉词典; 非负稀疏编码; 显著性加权; 相似性度量

**中图分类号:** TP391 **文献标志码:** A

## Image Retrieval Based on Saliency Weighted for Multiple Visual Dictionaries

Kong Chao<sup>1,2,3</sup>, Zhang Huaxiang<sup>1,3</sup>, Sheng Haidi<sup>1,3</sup>

(1. School of Information Science & Engineering, Shandong Normal University, Jinan, 250014, China; 2. State Grid of China Technology, Jinan, 250002, China; 3. Shandong Provincial Key Laboratory for Novel Distributed Computer Software Technology, Jinan, 250014, China)

**Abstract:** In view of application requirements of visual dictionary in image representation and retrieval, this paper proposes an image retrieval method based on the combination of multiple visual dictionaries and saliency weight, which can represent image features with saliency and sparsity. Firstly, the image is divided into blocks, and different kinds of underlying features of image blocks are extracted. Secondly, the image block features are used to learn the multiple visual dictionaries through non-negative sparse coding. The spatial information and saliency are introduced into the sparse vectors for the image blocks by the saliency pooling method, and saliency weight is introduced to form the sparse representation of the entire image. Finally, a proposed SDD distance is used for image retrieval. Compared with the method of single visual dictionary on common image dataset Corel and Caltech, Experimental results demonstrate that the proposed method can effectively improve the image retrieval accuracy.

**Key words:** multiple visual dictionaries; non-negative sparse coding; saliency weighted; similarity measure

## 引言

随着互联网技术的迅速发展和多媒体需求的不断提高,如何让用户高效地分析、管理和检索图像,已成为多媒体处理的研究热点之一。现阶段,基于文本与基于内容的图像检索方法得到了广泛应用,但仍存在一些研究难点,如底层特征与高层语义间存在“语义鸿沟”,只通过底层特征对图像进行描述不能很好地理解图像语义,导致图像检索准确率不高<sup>[1]</sup>。为此,研究者从多个方面提出了一系列的图像表示和检索方法,大致分为以下两方面:(1)学习图像的稀疏表示。文献[2]采用核稀疏表示与空间金字塔匹配相结合,减少了特征量化过程中的损失,生成判别性的稀疏编码,具有很好的性能。文献[3]提出了一种新的词典学习方法,将 Fisher 判别准则强加给稀疏系数,使它们有较小的类内散布和较大的类间散布,使得该方法更具有判别性。根据人脑的视觉分级机理,图像高层特征是其底层特征的组合,从底层到高层的特征表示是一个逐步抽象的过程<sup>[4]</sup>。稀疏编码方法可以在图像底层特征基础上抽象出图像高一层特征,实现从图像高层语义到底层特征的映射。传统稀疏编码从自然图像中随机选取若干图像子块组成一个训练集,通过学习得到一系列的基向量(即视觉词典),再采用稀疏分解算法实现图像稀疏表示<sup>[5,6]</sup>。文献[7]先提取图像的 Sift 特征,再对提取的特征进行稀疏编码,实现对特征的进一步抽象。但标准稀疏编码的系数具有可正可负性,可能会产生相互抵消的消极作用,非负的稀疏编码更适合构造视觉词典<sup>[8]</sup>,因此本文选用非负稀疏编码构造视觉词典。多数图像检索方法在特征提取阶段只提取图像的单一底层特征,往往不能很好满足用户的检索意图,因此本文融合图像的多种底层特征对图像进行表示,在一定程度上可以减小图像底层特征和高层语义间存在的“语义鸿沟”。(2)图像特征的空间量化。文献[9]提出了一种词袋模型的空间排列方法,以检测点为原点分割图像为 4 个象限,为词袋模型中的视觉单词进行空间位置编码,有效地加入了空间信息。文献[10]提出了一种新的多层次结构化图像压缩编码方法,编码中引入了语义和空间层次。显著性检测是根据图像亮度、颜色和对对比度等全局和局部特征分析图像的一种方法,在图像处理领域应用广泛,是计算机视觉方面的研究热点。显著性检测可以有效识别图像的兴趣区域,提高图像检索的准确率。

综上所述,本文结合非负稀疏编码和图像空间的显著性分析,提出了基于多视觉词典的显著性加权图像检索方法,并采用一种新的稀疏词典相似性度量方式进行图像检索。首先对图像进行重叠分块提取特征,通过非负稀疏编码学习特征词典。然后计算图像块的稀疏向量并进行显著性汇总,加入显著性和空间信息,形成整幅图像多尺度上的稀疏向量表示。最后使用提出的相似性度量方式进行图像检索。

## 1 非负稀疏词典学习

首先对训练图像进行重叠分块得到  $n$  个图像块,然后每个图像块可以提取  $p$  维特征。假设  $\mathbf{X} \in \mathbf{R}^{n \times p}$  是  $n \times p$  维的图像块特征矩阵,  $\mathbf{x}_i$  表示特征矩阵  $\mathbf{X}$  中的第  $i$  列(即第  $i$  个分块的特征向量)。通过最小化式(1)中的目标函数,可以得到图像特征的稀疏向量集合  $\mathbf{Z} = \{\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n\}$  以及一组  $d$  维的视觉词典  $\mathbf{D}$ 。即有

$$\min_{\mathbf{D}, \mathbf{Z}} \sum_{i=1}^n \|\mathbf{D}\mathbf{z}_i - \mathbf{x}_i\|_2 + \lambda \|\mathbf{z}_i\|_1 \quad \text{s. t.} \quad \mathbf{z}_i \geq 0 \quad (1)$$

式中:  $\|\mathbf{D}\mathbf{z}_i - \mathbf{x}_i\|_2$  是视觉词典与稀疏向量重建图像块的特征所产生的误差项;  $\lambda \|\mathbf{z}_i\|_1$  是稀疏性惩罚项,确保编码系数  $\mathbf{z}_i$  的稀疏性;正则化系数  $\lambda$  用于平衡误差项和稀疏惩罚项;  $\mathbf{z}_i \geq 0$  用于保证稀疏向量非负。

当固定系数  $\mathbf{Z}$  时,式(1)相对于变量  $\mathbf{D}$  是凸的。当固定词典  $\mathbf{D}$  时,式(1)相对于变量  $\mathbf{Z}$  也是凸的,但是二者不能同时进行优化,可通过重复的迭代选择固定  $\mathbf{Z}$  或  $\mathbf{D}$ ,对另外一个变量进行优化,每次迭代分为如下两步<sup>[11]</sup>:

(1) 固定词典  $D$ , 通过调整系数  $z_i$ , 使得目标函数最小 (即求解最小绝对压缩和选择算子 (Least absolute shrinkage and selection operator, LASSO) 问题)。

(2) 固定系数  $z_i$ , 通过调整词典  $D$ , 使得目标函数最小 (即求解凸二次规划 (Quadratic programming, QP) 问题)。

给定一个新图像块时, 由以上得到的视觉词典, 通过求解一个 LASSO 问题即可得到其稀疏向量  $z$ , 这个稀疏向量即为图像块的特征所对应的稀疏向量。

## 2 基于多视觉词典的显著性加权图像检索

本文提出的图像检索方法主要包括特征视觉词典的学习和查询图像的检索两部分, 其具体流程如图 1 所示。

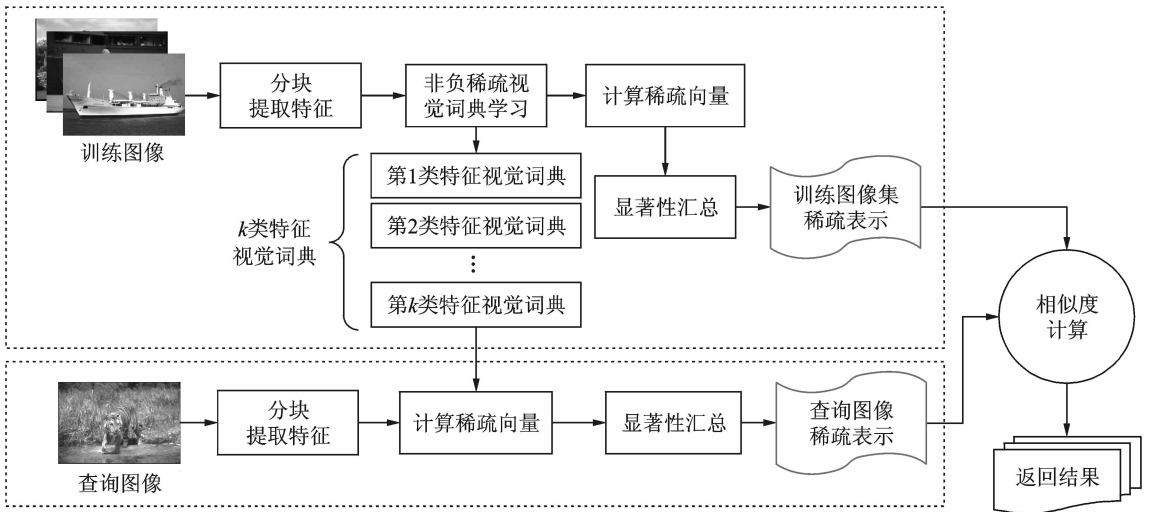


图 1 图像检索方法流程图

Fig. 1 Procedure of the image retrieval method

在特征视觉词典学习阶段, 先对图像训练集的大量图像块分别提取多类特征, 将提取出的每一类特征依次作为非负稀疏编码的输入, 分别学习多个视觉词典, 再采用显著性汇总的方法量化每类特征的稀疏向量。最后将多类特征稀疏向量结合在一起得到训练图像的稀疏表示。

在查询图像检索阶段, 先对查询图像提取分块多类特征, 通过学习到的多个视觉词典, 对于每个特征块的不同特征, 分别得到基于对应视觉词典的稀疏向量。再经过显著性汇总得到整个查询图像基于不同视觉词典的稀疏向量, 多向量结合之后生成查询图像的稀疏表示。最后计算与训练图像间的相似度并按照相似度由大到小排序返回查询结果。

### 2.1 显著性汇总

构建整幅图像的稀疏表示中, 如果将图像所有子块的稀疏向量简单连接起来, 则该图像的稀疏表示维数是其图像块稀疏向量维数的几百甚至上千倍, 这就为之后的相似性度量带来很大的计算开销并且得到的稀疏表示对图像尺度、方向等变化较敏感。通过汇总方法, 可以在多个尺度下对图像所有图像块的稀疏向量进行统计, 加入其空间分布信息。传统的汇总方法有 Average-pooling 和 Max-pooling 等<sup>[12]</sup>。Average-pooling 法只选取区域特征的平均值作为该区域的特征。Max-pooling 法选择区域特征的最大值作为该区域特征, 比 Average-pooling 法更具有鲁棒性, 但是没有考虑图像的显著性问题, 不能过滤掉

背景或非目标区域的噪声。

因此本文提出在图像多个尺度上进行显著性加权汇总的图像稀疏表示方法。首先,得到每个图像子块基于视觉词典的稀疏向量。然后在图像上以  $l$  尺度重新划分为  $l^2$  个不重叠的均匀小块,对每个小块内包含的多个子块系数进行显著性加权。最后将所有尺度上得到的新向量组合成整幅图像的稀疏表示,即有

$$z_i' = \sum_{j=1}^l \sum_{k=1}^d \left( \frac{\text{sal}}{\text{SAL}_i + \epsilon} \right) \cdot z_{ij} \quad (2)$$

式中:  $\text{sal}$  表示提取特征时所划分图像块的显著性值,  $\text{SAL}_i$  则表示汇总时所划分子块的显著性值,显著性检测算法由文献[13]给出。  $\frac{\text{sal}}{(\text{SAL}_i + \epsilon)}$  即为特征的显著性值在汇总子块的显著性值中的贡献度,  $\epsilon$  为平滑因子,用于保证显著性权值连续。  $z_i$  表示图像分块第  $i$  块的稀疏向量,  $z_{ij}$  即为其  $d$  维稀疏向量中的第  $j$  维。

图 2(a) 表示传统 Max-pooling 方法,图 2(b) 表示本文方法,其中蓝色柱是无显著性加权的特征,红色柱是经过显著性加权之后特征,如果经过显著性加权之后超过原最大值,则选择其作为显著性汇总的选择,否则不改变。

将不同尺度统计得到的稀疏向量作为整个图像的稀疏表示,引入多层次的空间信息,即有

$$Z' = \sum_i z_i' \quad (3)$$

## 2.2 多视觉词典图像表示

研究表明,由于通过单一特征表示图像内容存在判别性和描述性较差的问题,在图像检索中采用多种特征融合的特征表示方法所得到的准确度远比采用单一特征的要高<sup>[14-15]</sup>。因此本文提出了一种融合多特征生成多稀疏视觉词典的图像表示方法。

对图像块提取  $m$  种不同类型的底层特征,再利用得到的每一种底层特征来学习其对应的视觉词典。这样,一幅图像就可以由若干图像块的多类特征稀疏向量来表示。假设  $Z'_k$  是图像  $x$  第  $k$  类特征视觉词典稀疏向量进行显著性汇总之后稀疏表示,  $w_k$  是第  $k$  类特征稀疏表示的权值,则图像  $x$  的多类特征稀疏表示为

$$I(x) = \sum_{k=1}^m w_k Z'_k \quad (4)$$

## 2.3 相似性度量

在图像检索的相似性度量方面,常见的图像词典表示相似性度量方法有归一化词典距离<sup>[16]</sup> (Normalized dictionary distance, NDD) 和快速压缩距离<sup>[17]</sup> (Fast compression distance, FCD) 等方法。本文基于以上方法提出了一种综合考虑稀疏表示非零元素位置与差值的相似性度量方法。

假设  $x$  表示查询图像,则其对应的多特征稀疏表示为  $I(x)$ , 图像训练集的稀疏表示集合为  $Y = \{I(1), I(2), \dots, I(y)\}$ 。定义查询图像稀疏向量和训练集图像稀疏向量间的相似度计算公式为

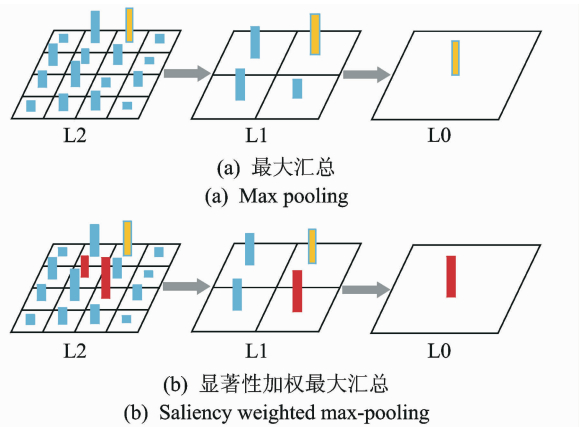


图 2 显著性加权最大汇总

Fig. 2 The method of the saliency weighted max-pooling

$$\text{SDD} = \left(1 - \frac{C(x \cap y)}{\max\{C(x), C(y)\}}\right) \times \| \mathbf{I}(x) - \mathbf{I}(y) \|_2 \quad (5)$$

式中: $C(x)$ 和 $C(y)$ 分别代表图像 $x$ 和 $y$ 的稀疏向量按阈值 $\tau$ 二值化后非零项的个数; $C(x \cap y)$ 表示两个二值稀疏表示同时不为零的元素个数; $\| \mathbf{I}(x) - \mathbf{I}(y) \|_2$ 为 $L_2$ 范数(即两稀疏表示的欧式距离)。

### 3 实验结果及分析

实验选用 Core10K 和 Caltech256 图像数据集,从中随机 10 类图像,每类 70 幅,共 700 幅图像,其中训练图像每类 60 幅,共 600 幅,测试图像每类 10 幅,共 100 幅。实验平台配置为 Intel(R)Core™ CPU3.3GHZ, RAM2.0GB 和 MATLAB 2012a。

#### 3.1 子块划分和特征提取

通过一个 16 像素 $\times$ 16 像素的滑动窗口将训练图像划分成图像块,每次的移动步长为 8 个像素,提取每个图像块的底层特征。本文分别提取图像块的 HSV 颜色特征、Gabor 纹理特征以及 Sift 特征。HSV 颜色特征中将  $\mathbf{H}$  分量分为 8 份, $\mathbf{S}$  分量分为 3 份, $\mathbf{V}$  分量分为 3 份,得到 72 维 HSV 特征向量( $\mathbf{L}_{\text{hsv}} = 8\mathbf{H} + 3\mathbf{S} + 3\mathbf{V}$ )。Gabor 纹理特征是对分块采用 4 尺度、6 方向的 Gabor 滤波器滤波时每个方向梯度上的均值和方差,从而得到 48 维纹理特征。SIFT 特征对尺度、旋转以及一定视角和光照变化等图像变化都具有不变性,并且具有很强的可区分性<sup>[18]</sup>,提取 Sift 特征时将图像块均匀分成 16 个小区域,在每个小区域上获取 8 个方向梯度,形成 128 维稠密 Sift 特征。为了解决特征向量间数量级不一致问题,对各特征向量进行内部归一化处理。

#### 3.2 实验设置

实验具体设置如下。

(1)在稀疏编码中,随机选取每幅训练图像的部分图像块,将这些图像块所提取的底层特征作为视觉词典训练的输入,以降低计算量。首先随机选择一组  $d$  维特征作为初始视觉词典。然后迭代训练稀疏向量  $\mathbf{Z}$  和视觉词典  $\mathbf{D}$ ,使得式(1)中目标函数最小化,其中正则化参数  $\lambda$  的值经实验对比后设定为 0.0001 时检索准确率最高,用于平衡误差项和稀疏项。设置迭代 50 次,单个词典维数  $d$  设置为 400。实验表明,词典维数的增加会在一定程度上增加检索的准确率,但这种增长较小。

(2)在显著性汇总中,实验直接采用文献[13]中的显著性检测算法,取其经过 Grabcut 之后的显著性图。对小于图像像素总和 5% 的显著性区域不予考虑(去除噪声),图像块的显著性值定义为图像块内显著性区域所占比例,即在 $[0, 1]$ 区间内。汇总的尺度分别为 1, 2 和 4,则汇总后的图像稀疏向量则为 8400 维。实验表明,多词典图像表示权值设为 $[0.4, 0.3, 0.3]$ 时,具有最优的检索性能。

(3)在相似性度量中,设置对图像的稀疏向量进行二值化的阈值  $\tau = 0.0005$ ,实验对比方法默认采用欧式距离。

#### 3.3 Core10K 数据集

本文采用图像检索常用的性能指标查全率和查准率来检测图像检索方法的有效性,其定义如下

$$\text{Precision} = \frac{T_M}{M}, \quad \text{Recall} = \frac{T_M}{N} \quad (6)$$

式中: $T_M$ 是图像检索返回结果中正确的图像数; $M$ 是图像检索返回的结果总数; $N$ 是图像训练集中与查询图像相同语义类的图像数,在这里  $N = 60$ 。

表 1 反映出本文方法( $k = 3, k$  为词典个数)在该数据集上相对于单一特征视觉词典方法( $k = 1$ )在查准率和查全率方面均有所提升,并且随着词典维数的增多有略微的增加。

表 1  $M=30$  时不同方法在多个词典维数下的性能比较Tab. 1 Performance comparison of different methods in several dictionary dimensions while  $M=30$ 

方法	$d=400$		$d=600$		$d=800$	
	Precision	Recall	Precision	Recall	Precision	Recall
H. G. S+SC( $k=1$ )	0.713 0	0.368 4	0.716 3	0.370 8	0.726 2	0.375 2
HSV+SC( $k=1$ )	0.711 6	0.367 6	0.720 9	0.372 5	0.732 0	0.378 2
Gabor+SC( $k=1$ )	0.531 1	0.274 4	0.545 3	0.281 7	0.548 3	0.283 3
Sift+SC <sup>[7]</sup> ( $k=1$ )	0.354 2	0.183 0	0.468 3	0.242 0	0.488 7	0.252 5
SC+欧式距离( $k=3$ )	0.781 1	0.403 6	0.787 9	0.407 1	0.795 6	0.411 1
本文方法( $k=3$ )	0.816 3	0.421 8	0.838 5	0.431 6	0.843 6	0.435 9

图 3 给出了不同方法在词典维数  $d=400$  时图像检索的查准率-查全率曲线。其中单个特征的 HSV, Gabor 和 Sift 视觉词典, 以及多种特征训练一组词典的方法 (H. G. S+SC), 在图 3 的 PR 曲线中均低于本文方法。由图 3 可看出本文提出的 SDD 距离对图像检索性能有所提高。图 4 反映出显著性汇总与 Max-pooling 相比具有明显的作用。

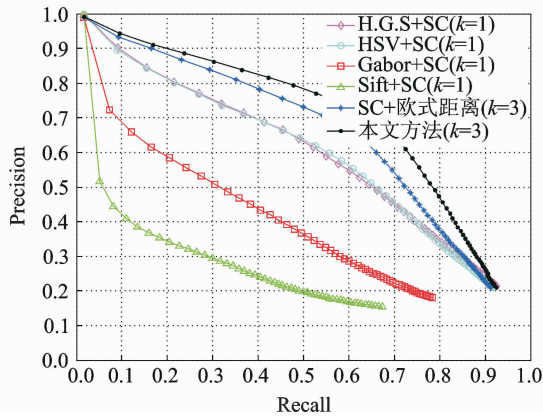
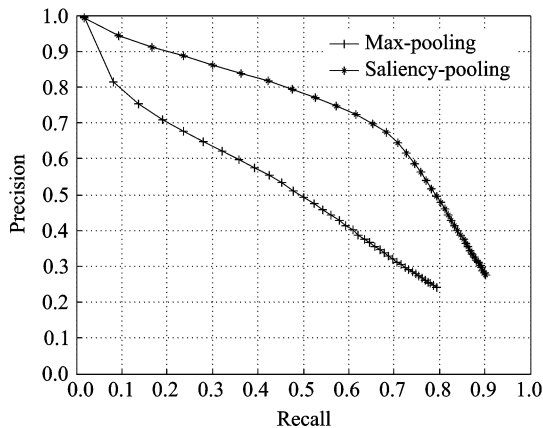
图 3  $d=400$  时图像检索查准率-查全率曲线Fig. 3 Precision-Recall curves of image retrieval while  $d=400$ 

图 4 显著性汇总和 Max-pooling 的检索结果对比

Fig. 4 Retrieval result comparison of saliency weight max-pooling and max-pooling

图5列举了图像检索部分语义类图像的平均查准率和查全率。其中,花类图像和食物类图像的检索结果比较好,老虎类图像的结果较差。图像检索方法的部分实验结果如图6所示,最左边的为查询图像,后面的即为图像检索结果。由图6可看出,在返回图像数量较少的情况下,其查准率比较高。

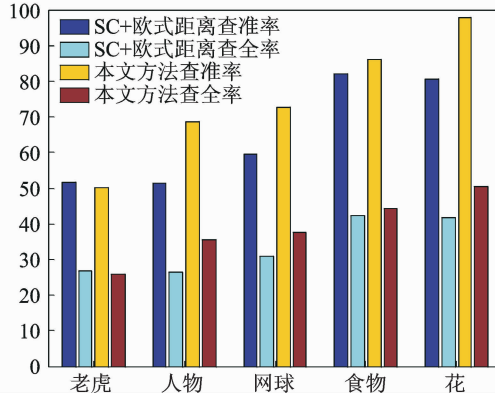


图5  $M=30$  时部分语义类图像的检索结果对比

Fig.5 Retrieval result comparison of several image semantic classes while  $M=30$

### 3.4 Caltech256 数据集

在该数据集上,采用与 Corel10K 数据集相同的实验设置,实验结果表明本文方法在 Caltech 数据集上也表现出了较好的性能,具体的图像检索查准率与查全率如表2所示。

表2 Caltech 上不同方法在多个词典维数下性能比较

Tab.2 Performance comparison of different methods in several dictionary dimensions on Caltech data set

方法	$d=400$		$d=500$	
	Precision	Recall	Precision	Recall
H. G. S+SC( $k=1$ )	0.627 5	0.305 8	0.630 3	0.312 8
HSV+SC( $k=1$ )	0.638 0	0.326 5	0.635 0	0.318 6
Gabor+SC( $k=1$ )	0.548 4	0.272 4	0.534 9	0.269 1
Sift+SC <sup>[7]</sup> ( $k=1$ )	0.656 0	0.332 1	0.676 2	0.347 2
SC+欧式距离( $k=3$ )	0.680 1	0.344 5	0.692 1	0.359 4
本文方法( $k=3$ )	0.706 2	0.350 8	0.710 4	0.360 3

## 4 结束语

本文通过结合非负稀疏编码和图像显著性分析,提出了基于多视觉词典的显著性加权图像表示方法,通过非负稀疏编码分别学习图像块多种特征对应的视觉词典,将得到的图像块稀疏向量经过显著性汇总方法引入空间信息并作显著性加权处理,并采用一种新的相似性度量方式进行图像检索。在 Corel 和 Caltech 图像集上的仿真实验结果表明,本文提出的方法具有较好的图像检索性能。目前,本文方法仅尝试了3种特征的词典组合及其在小规模图像数据集进行实验,尚未考虑词典的自动更新问题,今后的工作将围绕这些问题以及对多视觉词典进行语义关联等方面展开,构建符合人视觉特性的图像表示,以更准确地表示图像语义。



图 6 部分图像检索结果

Fig. 6 Partial image retrieval results

## 参考文献:

- [1] Liu Chengjun, Song Guangwei. A method of measuring the semantic gap in image retrieval: Using the information theory [C]// International Conference on Image Analysis and Signal Processing. Hubei: IEEE, 2011: 287-291.
- [2] Gao Shenghua, Tsang I W H, Chia L T. Kernel sparse representation for image classification and face recognition [C]// The 11th European Conference on Computer Vision. Greece: Springer, 2010: 1-14.
- [3] Yang Meng, Zhang Lei, Feng Xiangchu, et al. Fisher discrimination dictionary learning for sparse representation [C]// International Conference on Computer Vision. Barcelona: IEEE, 2011: 543-550.
- [4] Hubel D H, Wiesel T N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex [J]. The Journal of Physiology, 1962, 160(1): 106-154.
- [5] Yang Meng, Zhang Lei, Yang Jian. Robust sparse coding for face recognition [C]// Conference on Computer Vision and Pattern Recognition. Providence: IEEE, 2011: 625-632.
- [6] Yu Kai, Lin Yuanqing, Lafferty J. Learning image representations from the pixel level via hierarchical sparse coding [C]// Conference on Computer Vision and Pattern Recognition. Providence: IEEE, 2011: 1713-1720.
- [7] Shi Jun, Jiang Zhiguo, Feng Hao, et al. SIFT-based elastic sparse coding for image retrieval [C]// The 19th International Con-



ference on Image Processing. Orlando:IEEE,2012:2437-2440.

- [8] 江爱文,王春恒,肖柏华. 基于 GaborSIFT+NNScSPM 图像特征抽取算法研究[J]. 自动化学报,2011,37(10):1183-1189.  
Jiang Aiwen,Wang Chunheng,Xiao Bohua. An image feature extraction method based on GaborSIFT+NNScSPM[J]. Acta Automatica Sinica,2011,37(10):1183-1189.
- [9] Penatti O A B, Silva F B, Valle E, et al. Visual word spatial arrangement for image retrieval and classification[J]. Pattern Recognition,2014,47(2):705-720.
- [10] Li Lijia,Zhu Jun,Su Hao, et al. Multi-level structured image coding on high-dimensional image representation[C]//The 11th Asian Conference on Computer Vision. Korea:Springer,2013:147-161.
- [11] Yang Jianchao, Yu Kai, Gong Yihong, et al. Linear spatial pyramid matching using sparse coding for image classification [C]// Conference on Computer Vision and Pattern Recognition. Miami:IEEE,2009:1794-1801.
- [12] Yang Jimei, Yang M H. Learning hierarchical image representation with sparsity, saliency and locality[C]//Proceedings of the British Machine Vision Conference. Dundee:BMVA,2011:19. 1-19. 11.
- [13] Cheng Mingming, Zhang Guoxin, Mitra N J, et al. Global contrast based salient region detection[C]// Conference on Computer Vision and Pattern Recognition. Providence:IEEE,2011:409-416.
- [14] Sun Lang, Tang Yan, Zhang Hong. An effective image retrieval method based on multi-features[J]. Journal of Software, 2014,9(4):931-937.
- [15] 刘帅,李士进,冯钧. 多特征融合的遥感图像分类[J]. 数据采集与处理,2014,29(1):108-115.  
Liu Shuai, Li Shijin, Feng Jun. Remote sensing image classification based on adaptive fusion of multiple features[J]. Journal of Data Acquisition and Processing,2014,29(1):108-115.
- [16] Macedonas A, Besiris D, Economou G, et al. Dictionary based color image retrieval[J]. Journal of Visual Communication and Image Representation,2008,19:464-470.
- [17] Cerra D, Datcu M. A fast compression-based similarity measure with applications to content-based image retrieval[J]. Journal of Visual Communication and Image Representation,2012,23:293-302.
- [18] Zhang Dengsheng, Islam M M, Guojun Lu. A review on automatic image annotation techniques [J]. Pattern Recognition, 2012,45:345-362.

#### 作者简介:



**孔超**(1989-),男,硕士研究生,研究方向:机器学习、图像检索, E-mail: kongchao-0626@163.com。



**张化祥**(1966-),男,教授,博士生导师,研究方向:机器学习,模式识别及 Web 挖掘等。



**生海迪**(1989-),女,硕士研究生,研究方向:数字图像处理。

