

文章编号:1004-9037(2014)01-0060-06

# 基于视觉单词选择的高分辨率遥感图像飞机目标检测

李士进 仇建斌 於 慧

(河海大学计算机与信息学院,南京,210098)

**摘要:** 高分辨率遥感图像中飞机目标的检测和识别具有重要的军事和民用价值,针对以往方法易受灰度分布和形态变化及伪装干扰等缺点,提出一种基于视觉词袋模型的高分辨率遥感图像飞机目标检测的新方法。为了精简飞机视觉码本得到最具鉴别力的视觉单词,结合相关性及冗余度分析去除视觉码本中不相关、弱相关以及冗余的视觉单词,选择对飞机目标检测最为重要的视觉单词,减少了计算复杂度,提高了算法的检测性能。

**关键词:** 遥感图像;飞机检测;特征选择;视觉词袋

**中图分类号:** TP391.4

**文献标志码:** A

## Aircraft Detection in High-resolution Remote Sensing Imagery Based on Visual Words Selection

Li Shijin, Qiu Jianbin, Yu Hui

(College of Computer and Information, Hohai University, Nanjing, 210098, China)

**Abstract:** It is difficult to detect the target of aircrafts in high resolution remote sensing images. So a bag-of-visual-words model is proposed. In order to obtain the most discriminative visual words and make visual codebook effective and compact, the visual words that are irrelevant, weakly irrelevant and redundant must be removed from the visual codebook. The method integrating relevance analysis with redundancy analysis is used to prune out those useless visual words. Finally, the most important visual words are chosen to describe aircrafts in high resolution remote sensing images, which helps reducing the computation in the following test stage and also improving the efficiency.

**Key words:** remote sensing imagery; aircraft detection; feature selection; bag-of-visual-words

## 引 言

随着遥感技术的迅速发展,遥感图像的空间分辨率越来越高,对遥感图像中感兴趣的目标进行定位和辨识已经成为一个重要的研究方向<sup>[1-5]</sup>。作为一种重要的战略目标,基于高分辨率遥感图像的飞机目标检测,无论在军事还是民用方面都有重要的意义。

目前,基于遥感图像的飞机目标检测技术,大多采用由下而上的数据驱动型策略,通常采用先分割或边缘提取<sup>[1-3]</sup>,然后采用模板匹配<sup>[6]</sup>的方法,但是由于遥感图像质量和飞机阴影的影响,通常分割后飞机目标会断开成几个区域,很难完整提取出飞

机目标边缘轮廓,所以通常做法是再进行区域合并或者边缘的连接。这种方法不仅实现起来很复杂而且抗干扰能力很弱,同时由于飞机的类型很多,很难用一个统一的模板来检测所有的飞机目标,因此往往检测结果不令人满意。文献[7]提出了一种采用自上而下知识驱动型策略的飞机检测方法,但是该方法对图像灰度值分布有很强的依赖性,对于有伪装的飞机目标或者飞机与背景区分程度较弱的图像,存在很多漏检的情况。

近年来,基于视觉词袋的图像表示已被广泛地应用到计算机视觉和多媒体领域<sup>[8]</sup>,本文将视觉词袋模型引入到高分辨遥感图像中用于飞机目标检测,对飞机图像训练集构建描述飞机特性的视觉码本,并将此作为检测依据,通过特征选择有效压缩

了视觉码本的规模,选出了对飞机检测最为重要的特征,从而减少后续检测的计算量,提高了算法的检测性能。

## 1 视觉词袋模型

视觉词袋是目前比较流行的基于局部属性的图像表示方法,源于文档分析领域中的词袋表示(bag-of-words)。López-Sastre 等将监督学习的聚类方法引入到视觉词汇表的构建过程,增加了视觉单词的语义表达力和区别力<sup>[9]</sup>;Elsayad 等针对视觉单词空间排列信息的缺失,采用高斯混合模型来描述视觉单词的空间分布<sup>[10]</sup>,王宇新等提出了一种用于图像场景分类的空间视觉词袋模型<sup>[11]</sup>。霍华等对图像进行多尺度变换并构建多尺度视觉词汇表,将图像表示为不同尺度特征<sup>[12]</sup>。

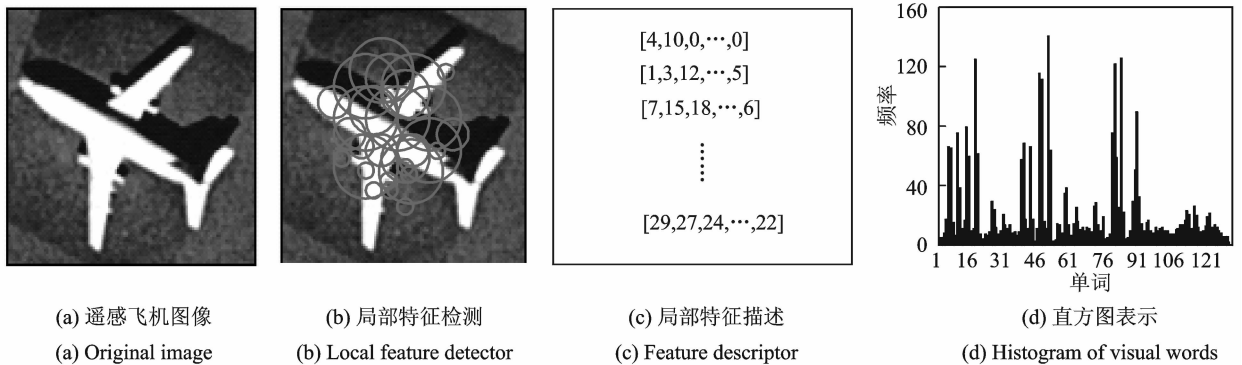


图 1 遥感飞机图像视觉词袋方法的表示过程

Fig. 1 The procedure of bag-of-visual-words extraction from remote sensing images

## 2 基于视觉单词选择的高分辨率遥感飞机图像目标检测

### 2.1 视觉单词选择算法

在飞机目标的检测过程中,视觉码本太小,一些不相似的描述符可能会与相同的视觉单词相对应,从而缺少鉴别能力;视觉码本太大时,某些相似的描述符可能会被匹配到不同的视觉单词上去,造成冗余,且需要更多的存储和计算资源。因此有必要对视觉码本进行特征选择以去除视觉码本中相关性不高、冗余度大的视觉单词。

Dash 和 Liu 提出一种特征选择框架<sup>[15]</sup>,文献<sup>[16]</sup>提出将互信息、比值比和线性支持向量机结合起来选取最具信息量的视觉单词。文献<sup>[17]</sup>使用 Boosting 特征选择方法从多分辨率视觉码本中

视觉词袋模型的基本思想是将图像通过关键点检测或按固定尺寸分割成同等大小的图像片段并通过某种算法表示为局部视觉特征。通过对局部视觉特征集的聚类构建描述目标特性的视觉码本,由每个聚类中心代表的视觉单词组成,并用一个可供分类的特征向量直方图表示<sup>[13]</sup>。

在高分辨率遥感飞机图像目标识别中,首先提取所有飞机图像的 SIFT 局部特征,用 K-means<sup>[14]</sup>聚类算法对这些特征进行聚类形成视觉单词,然后将所有视觉单词组成一个可以描述高分辨率遥感飞机图像形态特征的视觉码本。任一高分辨率遥感飞机图像均可表达为该幅图像提取的 SIFT 局部特征关于该视觉码本中各视觉单词出现频次的直方图。

基于视觉词袋表示方法的遥感图像中飞机检测的过程如图 1 所示。

选取最具鉴别力的视觉单词。文献<sup>[18]</sup>提出了一种基于相关性及冗余度分析的新的特征选择结构,具体如图 2 所示。从视觉单词间的相关性及冗余度角度考虑,本文采用文献<sup>[18]</sup>提出的特征选择框架进行视觉单词选择。衡量两个视觉单词之间的相关性时,选用基于熵概念的对称不确定性作为相关性度量。随机变量  $X$  的信息熵如下式所示

$$H(X) = - \sum_i P(x_i) \log_2(P(x_i)) \quad (1)$$

其中  $P(x_i)$  为  $X$  每个取值的概率密度函数,在  $Y$  给定时  $X$  的条件信息熵定义为

$$H(X | Y) = - \sum_j P(y_j) \sum_i P(x_i | y_j) \log_2(P(x_i | y_j)) \quad (2)$$

其中  $P(x_i | y_j)$  为两者的条件概率。

在  $Y$  给定的条件下  $X$  的信息增益为

$$IG(X | Y) = H(X) - H(X | Y) \quad (3)$$

最终的非线性相关性度量为对称不确定性,计算公

式为

$$SU(X, Y) = 2 \left[ \frac{IG(X | Y)}{H(X) + H(Y)} \right] \quad (4)$$

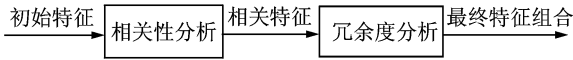


图 2 基于相关性及冗余度分析的特征选择结构

Fig. 2 The framework of feature selection based on relevance and redundancy analysis

## 2.2 算法描述

### 2.2.1 训练阶段

首先提取所有飞机训练图像的 SIFT 局部特征并进行汇总,用 K-means 方法进行聚类生成  $K$  个视觉单词,构成高分辨率遥感飞机图像初始视觉码本。然后对初始视觉码本进行视觉单词选择,得到最终精简的高分辨率遥感飞机图像视觉码本。

视觉单词选择具体算法步骤如下:

(1)对视觉码本中的视觉单词  $F_i$ ,利用式(4)计算其与类别  $C$  之间的类别相关性,记为  $SU_{F_i, c}$ ,  $1 \leq i \leq K$  ( $K$  为 K-means 的聚类数目,即初始飞机视觉码本中的视觉单词数目)。

(2)根据预先设定的阈值  $\delta$  对视觉码本中的视觉单词进行相关性分析。对任意  $SU_{F_i, c}$ ,  $1 \leq i \leq K$ , 当  $SU_{F_i, c} \geq \delta$ ,  $1 \leq i \leq K$  时,该视觉单词  $F_i$  被保留并加入到集合  $W'_{list}$  中,否则  $F_i$  被去除。

(3)对初始视觉码本中的所有视觉单词都执行步骤(2)后相关性分析结束。按  $SU_{F_i, c}$  值对  $W'_{list}$  中的所有视觉单词进行降序排列。

(4)对  $W'_{list}$  中每一当前排序第一的视觉单词  $F_j$ ,利用式(4)计算该视觉单词与  $W'_{list}$  其他任意视觉单词  $F_i$ ,  $i \neq j$  之间的对称不确定性  $SU_{F_i, F_j}$ ,若  $SU_{F_i, F_j} > SU_{F_i, c}$ ,则说明  $F_i$  与视觉单词  $F_j$  冗余,将  $F_i$  从  $W'_{list}$  中去除。当  $F_j$  与  $W'_{list}$  中所有其他视觉单词都计算一轮后,将  $F_j$  加入到集合  $W_{list}$  中,同时在  $W'_{list}$  中删除  $F_j$ 。

(5)重复(4)中的操作直到  $W'_{list}$  为空,此时所有的视觉单词都已被处理,冗余度分析完毕。 $W_{list}$  即是最终精简且具有鉴别力的高分辨率遥感飞机图像视觉码本。

得到新的飞机目标视觉码本后,必须对训练集中的所有图像根据新的视觉码本统计出各自的直方图信息,完成训练过程。针对每幅图像,根据它的 SIFT 局部特征形成视觉单词直方图的过程如下所示:

(1)对每幅图像的每个 SIFT 特征,计算它与视觉码本中各个视觉单词之间的欧氏距离。

(2)找出与对应视觉单词距离最近的 SIFT 局部特征,将隶属于该视觉单词的 SIFT 特征个数加 1。

(3)重复步骤(1)~(2),直到该图像的所有 SIFT 特征都处理完毕,此时得到该幅图像关于新视觉码本中所有视觉单词的分布直方图。

### 2.2.2 测试阶段

对测试集中的每幅测试图像提取它的 SIFT 局部特征,按照构造训练直方图相同的方法构造测试直方图,计算与训练集中所有训练直方图之间的相似程度,具体的测试过程为:

(1)对测试图像,计算该图像测试直方图与训练集中每一训练图像的直方图的相似度——直方图相交值。

(2)如果直方图相交值大于预先设定的阈值,则认为该测试图像为飞机目标,测试过程结束;否则转(3)。

(3)与训练集中其他未比较的训练直方图继续步骤(1)~(2),如果遍历整个训练集都没有找到满足步骤(2)中条件的训练图像,则认为该测试图像不是飞机图像,测试过程结束。

## 3 实验结果及分析

为了综合评判视觉码本的优劣,本文采用  $F_1$ —Measure 作为最终的度量指标,该指标是根据检测查全率  $R$  和查准率  $P$  定义的一个综合评价指标,具体的计算公式为

$$F_1 = \frac{2P \times R}{P + R} \quad (5)$$

本文在 60 幅快鸟高分辨率遥感飞机大图像中选取测试图像,这 60 幅大图像中共包含 238 架飞机目标,共切分了 325 幅小测试图像,其中 219 幅含飞机目标,106 幅不含飞机目标。图 3 给出了其中的一些代表性样本。从图 3 中可以看出本文训练集不仅包含了不同类型、不同大小的飞机目标,还包含一些伪装的飞机目标,能够体现训练样本的多样性。

提取这些飞机训练图像的 SIFT 局部特征汇总后共有 6 100 个,每个 SIFT 特征有 128 维。接着对这 6 100 个局部特征用 K-means 方法进行聚类生成视觉单词,构成高分辨率遥感飞机图像视觉码本。

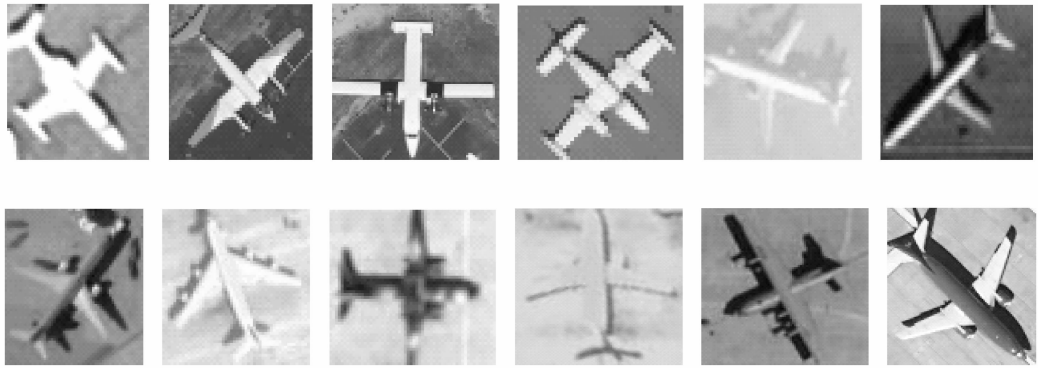


图 3 遥感飞机图像训练集包含的各式飞机类型  
Fig. 3 Typical airplane examples in the training set

本文在聚类阶段尝试了 3 个聚类数  $K$ , 分别形成了包含 200, 500 和 1 000 个视觉单词的初始飞机图像视觉码本。分别对这 3 个视觉码本进行视觉单词选择算法, 结果如表 1 所示。虽然后两种情况视觉单词数较多, 但冗余度较高, 检测结果反而较差(检测率最高只有 75% 左右), 因此下文只报告 200 个视觉单词的检测结果。

表 1 不同聚类数进行单词选择结果

Table 1 The feature selection results with different cluster numbers

聚类数 $K$	相关性分析去除视觉单词数	冗余度分析去除视觉单词数	最终视觉码本中包含单词数
200	163	25	12
500	223	274	3
1 000	186	812	2

图 4 和图 5 给出了几个检测的例子。对图 4~

5 的原始图像(a)分别提取 SIFT 特征, 第(d)列框内的图像为最终得到的检测结果, 可以发现, 大部分的飞机目标被正确检测, 同时一些 SIFT 特征点个数满足要求的背景区域也被选中, 如图 5。另外还存在一些漏检情况, 其原因主要是由于图像 SIFT 特征点密集, 会有很多区域重叠在一起而且这些区域的中心点都相距很近, 经过区域合并后这些相邻区域的中心点都会被归为一类, 最终只算作一个目标区域, 造成多个靠近的飞机漏检。

用选择视觉单词后精简的飞机视觉码本对 325 幅测试图像进行测试, 与不进行单词选择时, 直接采用原始所有 200 个视觉单词对训练集构建训练直方图进行飞机识别时的情况对比, 得到的测试结果如表 2 所示。可以看出, 将视觉词袋引入到高分辨率遥感图像中用于飞机目标检测识别是可行的。

表 2 不同方法对相同测试集测试的情况

Table 2 Comparison of detection results with/without feature selection

方法	阈值	正确检测	漏检	虚警数	查全率 $R/\%$	查准率 $P/\%$	$F_1$ -Measure/ $\%$
不进行单词选择	0.2	200	38	45	84.03	81.63	82.81
单词选择	0.75	213	25	43	89.50	83.20	86.24

对于同样的测试集, 若不进行单词选择, 当测试直方图与训练直方图的相交值同样大于等于 0.75 时, 不能正确识别出任何一幅测试图像, 当相似度阈值取 0.2 时识别结果最好, 但此时 3 个评判指标都不及本文选取的包含 12 个视觉单词的视觉码本对应的评判指标, 且此时测试图像与训练图像的相似度很低(只有 0.2), 说明原有 200 个单词中不相关、弱相关以及冗余单词的存在严重影响了飞机目标的识别, 在视觉单词数较多时不能准确地描

述对飞机目标识别最为重要的特征, 说明了特征选择的重要性以及必要性。

本文还分别用初始聚类成 500 和 1 000 类时选择出的视觉码本对测试图像进行测试, 两者的查全率虽然都比较高, 但是查准率却远远低于初始聚类成 200 类时选择出的飞机码本(只有 65% 左右), 说明此时它们几乎将第二类不含飞机目标的测试图像都错识别为飞机, 说明此时的视觉码本包含的视觉单词鉴别能力较弱, 不能准确区分飞机目标和

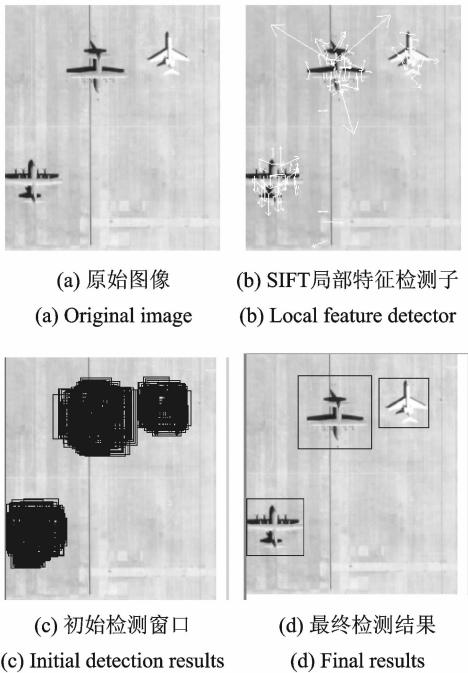


图 4 包含多个飞机目标的高分辨率遥感图像检测过程

Fig. 4 Illustration of the aircraft detection procedure

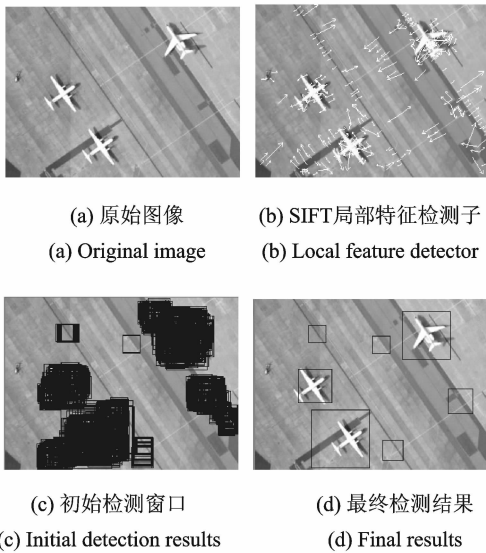


图 5 误检测情况

Fig. 5 Illustration of the aircraft detection procedure with false alarms

非飞机目标,因而这两组视觉码本对飞机图像来说不具有较强鉴别能力。

文献[7]中的基于圆周频率滤波法的飞机检测方法,是近年来提出的一种比较有效的方法,在本文数据集上的查准率为 93.23%,查全率为 75.21%, $F_1$  指标为 83.26%,低于本文提出的基于视觉单词

选择的检测方法的结果(86.24%)。基于圆周频率滤波法的飞机检测,在飞机与背景反差比较大的遥感图像中,检测结果比较理想。但是在复杂背景下,特别是飞机有伪装时,该方法漏检情况严重,造成其查全率较低。而本文提出的飞机检测方法大大提高了查全率,降低了飞机漏检率。飞机检测一般应用于军事领域,能否及时检测出全部飞机目标对作战指挥员把握敌我双方动态,快速作出反应,打赢未来高科技条件下信息战争显得尤为重要。因此,以稍低的查准率代价换取较高查全率是值得的。本文方法比文献[7]中圆周频率滤波法更加高效,也更具应用价值。

## 4 结束语

本文从视觉词袋单词选择的角度提出了一种高分辨率遥感飞机图像目标检测的方法。首先对训练集中的图像提取 SIFT 局部特征进而通过 K-means 聚类方法聚类形成视觉单词,所有视觉单词组成描述高分辨率遥感飞机图像的视觉码本。接着针对初始视觉码本视觉单词数较多、且存在冗余和噪声的特点,结合相关性及冗余度分析精简了视觉单词码本的规模,选出了最能描述飞机目标特点的较少的视觉单词,提高了视觉码本的鉴别能力,减少了后续飞机检测识别的计算量,提高了检测效率,并且通过实验验证了提出方法的检测率比较理想。本文从特征选择角度研究了视觉词袋模型在飞机检测中的应用,其前提是视觉单词已经生成。近年来,稀疏编码特别是字典学习在图像识别中得到了较广泛的关注<sup>[19]</sup>。如何利用稀疏学习获得飞机图像更加有效的特征表示是今后值得研究的方向之一。

### 参考文献:

- [1] 徐大琦,倪国强,许廷发. 中高分辨力遥感图像中飞机目标自动识别算法研究[J]. 光学技术, 2006, 32(6): 855-862.  
Xu Daqi, Ni Guoqiang, Xu Tingfa. Study on the algorithm for automatic plane classification from remote sensing images with mid-high resolution[J]. Optical Technique, 2006, 32(6): 855-862.
- [2] 王树国,黄勇杰,张生. 可见光图像中飞机目标的特征选择及提取[J]. 哈尔滨工业大学学报, 2010, 42(7): 1056-1059.  
Wang Shuguo, Huang Yongjie, Zhang Sheng. Selection and extraction of features of aircraft in optical image[J]. Journal of Harbin Institute of Technolo-

- gy, 2010, 42(7): 1056-1059.
- [3] 杨桃, 张柏, 王宗明, 等. 基于阴影搜索法的飞机目标遥感图像分割研究[J]. 地理与地理信息科学, 2006, 22(1): 48-50.  
Yang Guang, Zhang Bai, Wang Zongming, et al. Study on segmentation of airplane object in remote sensing based on shadow-searching method[J]. Geography and Geo-Information Science, 2006, 22(1): 48-50.
- [4] 潘诚. 极化 SAR 图像中的机场和飞机目标检测[D]. 上海:复旦大学,2012.  
Pan Cheng. Detection of airport and airplane using PolSAR images[D]. Shanghai: Fudan University, 2012.
- [5] 张亚飞, 朱敏慧. 基于混沌理论和形态学变换的海面分布目标检测[J]. 数据采集与处理. 2008, 23(2): 123-128.  
Zhang Yafei, Zhu Minhui. Detection of sea-surface distributed targets based on chaos theory and morphological transform[J]. Journal of Data Acquisition and Processing, 2008, 23(2): 123-128.
- [6] 马时平, 毕笃彦, 陈岚岚. 基于图像匹配技术的飞机识别[J]. 计算机工程, 2004, 30(5): 159-160.  
Ma Shiping, Bi Duyan, Chen Lanlan. Airplane recognition based on image matching technology[J]. Computer Engineering, 2004, 30(5): 159-160.
- [7] 蔡红苹, 耿振伟, 粟毅. 遥感图像飞机检测新方法——圆周频率滤波法[J]. 信号处理, 2007, 23(4): 539-543.  
Cai Hongping, Geng Zhenwei, Su Yi. A new method to detect air planes in remote sensing image—Circle-frequency filter[J]. Signal Processing, 2007, 23(4): 539-543.
- [8] 周文罡. 基于局部特征的视觉上下文分析及其应用[D]. 合肥:中国科学技术大学, 2011.  
Zhou Wengong. Visual context analysis based on local features and its applications[D]. Hefei: University of Science and Technology of China, 2011.
- [9] López-Sastre R J, Tuytelaars T, Acevedo-Rodríguez F J, et al. Towards a more discriminative and semantic visual vocabulary[J]. Computer Vision and Image Understanding, 2011, 115(3): 415-425.
- [10] Elsayad I, Martinet J, Urruty T, et al. A new spatial weighting scheme for bag-of-visual-words[C]//2010 International Workshop on Content-Based Multimedia Indexing (CBMI). [S.l.]:IEEE, 2010; 1-6.
- [11] 王宇新, 郭禾, 何昌钦, 等. 用于图像场景分类的空间视觉词袋模型[J]. 计算机科学, 2011, 38(8): 265-268.  
Wang Yu Xin, Guo He, He Cangqin, et al. Bag of spatial visual words model for scene classification[J]. Computer Science, 2011, 38(8): 265-268.
- [12] 霍华, 赵刚. 基于改进视觉词袋模型的图像标注方法[J]. 计算机工程, 2012, 38(22): 276-278, 282.  
Huo Hua, Zhao Guang. Image annotation method based on improved BOVW model[J]. Computer Engineering, 2012, 38(22): 276-278, 282.
- [13] Lowe D G. Distinctive image features from scale-invariant key points[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [14] MacQueen J. Some methods for classification and analysis of multivariate observations[C]//Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability. [S.l.]: University of Calif. Press, 1967, (1): 281-297.
- [15] Dash M, Liu H. Feature selection for classification[J]. Intelligent Data Analysis, 1997, 1(3): 131-156.
- [16] Jurie F, Triggs B. Creating efficient codebooks for visual recognition [C]//Tenth IEEE International Conference on Computer Vision, ICCV 2005. [S.l.]:IEEE, 2005, 1: 604-610.
- [17] Wang L. Toward a discriminative codebook; code-word selection across multi-resolution [C]//IEEE Conference on Computer Vision and Pattern Recognition, CVPR'07. [S.l.]:IEEE, 2007;1-8.
- [18] Yu L, Liu H. Efficient feature selection via analysis of relevance and redundancy[J]. Journal of Machine Learning Research, 2004, 5: 1205-1224.
- [19] Yang Jianchao, Yu Kai, Gong Yihong, et al. Linear spatial pyramid matching using sparse coding for image classification. [C]//2009 IEEE Conference on Computer Vision and Pattern Recognition, CVPR2009. [S.l.]: IEEE Press, 2009;1794-1801.

作者简介:李士进(1973-),男,博士,教授,研究方向:模式识别与计算机视觉,E-mail:lishijin@hhu.edu.cn;仇建斌(1988-),女,硕士研究生,研究方向:机器学习与特征选择;於慧(1987-),女,硕士研究生,研究方向:遥感图像处理与图像检索。

