

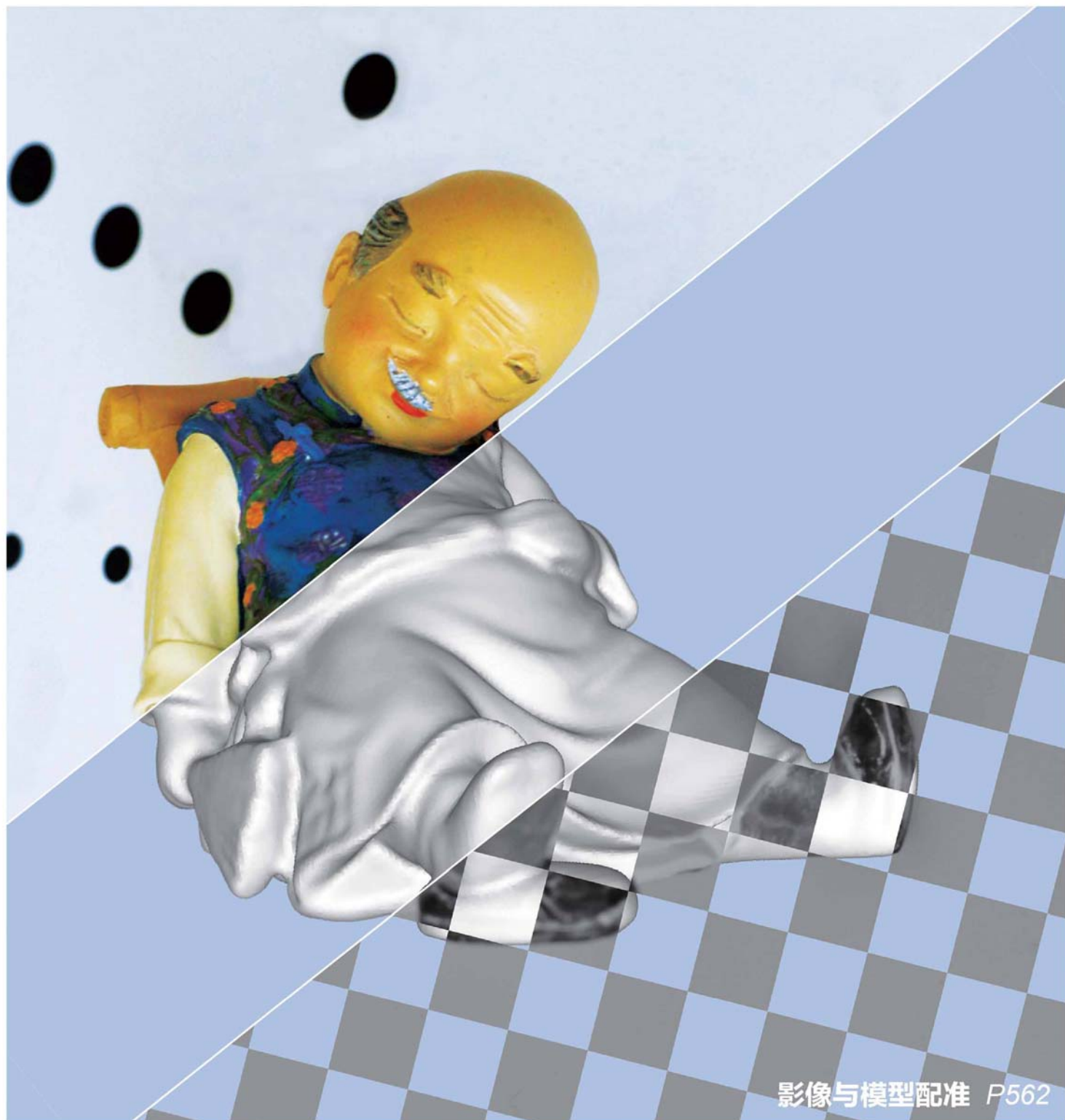


主办: 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象图形学报

2014
04
VOL.19

ISSN1006-8961
CN11-3758/TB



影像与模型配准 P562



第19卷第4期（总第216期）

2014年4月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊

版权声明

凡向《中国图象图形学报》投稿，均视为同意在本刊网站及CNKI等全文数据库出版，所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用，所有内容，未经许可，不得转载或以其他方式使用。

Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院

主办单位 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

主 编 李小文

编辑出版《中国图象图形学报》编辑出版委员会

邮政信箱 北京9718信箱

邮 编 100101

电子信箱 jig@irsa.ac.cn

电 话 010-64807995

网 址 www.cjig.cn

广告经营许可证 京朝工商广字第0361号

总 发 行 北京报刊发行局

订 购 全国各地邮局

海外发行 中国国际图书贸易集团有限公司
(邮政信箱：北京399信箱 邮编：100048)

印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

Title inscription: Song Jian

Monthly, Started in 1996

Superintended by Chinese Academy of Sciences

Sponsored by Institute of Remote Sensing and Digital Earth, CAS
China Society of Image and Graphics

Institute of Applied Physics and Computational Mathematics

Editor-in-Chief LI Xiaowen

Editor, Publisher Editorial and Publishing Board of Journal of
Image and Graphics

P.O.Box 9718, Beijing, P.R.China

Zip code 100101

E-mail jig@irsa.ac.cn

Telephone 010-64807995

Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals

Domestic All Local Post Offices in China

Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China))

Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB

ISSN 1006-8961

CODEN ZTTXFZ

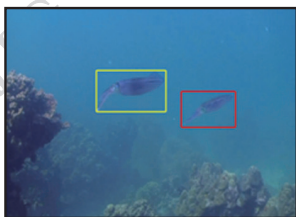
国外发行代号 M1406

国内邮发代号 82-831

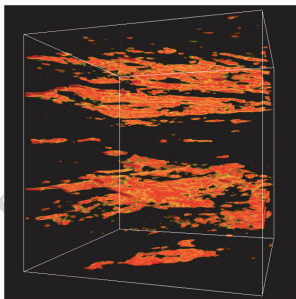
国内定价 50.00元



面泊松融合结合色彩变换的无缝纹理辐射处理(第512页)



融合视觉深度的特征计算与水下目标跟踪(第534页)



针对全空子数据体的GPU体绘制(第577页)

图像处理和编码

- 梯度矢量扩散控制实现边缘保持的彩色图像去噪
贾迪, 孟祥福, 张一飞, 董娜, 李思慧 0493
- 联合矩阵F范数的低秩图像去噪
刘新艳, 马杰, 张小美, 胡钊政 0502
- 面泊松融合结合色彩变换的无缝纹理辐射处理
周漾, 郑顺义, 黄荣永, 马电 0512
- 改进的局部方向模式纹理表示方法
刘海军, 常东超, 张凌宇 0520
- 图像插值空间大容量可逆数字水印算法
王继军 0527

图像分析和识别

- 融合视觉深度的特征计算与水下目标跟踪
王慧斌, 程勇, 陈哲 0534
- 融合颜色属性和空间信息的显著性物体检测
徐丹, 唐振民, 徐威 0541
- 采用人工鱼群的改进广义Hough变换目标定位
李志, 谢强 0549
- 非规则弯曲形变的不变量表示
李国祥, 夏国恩 0556

图像理解和计算机视觉

- 小物体3维重建中影像与几何模型的配准
郑顺义, 王晓南 0562
- 像素聚类改进二进制描述子鲁棒性
惠国保, 李东波 0569

计算机图形学

- 针对全空子数据体的GPU体绘制
李国和, 段忠祥, 吴卫江, 洪云峰, 刘智渊, 程远 0577

医学图像处理

- 基于无损水印的医学图像篡改检测和质量恢复
邓小鸿, 陈志刚, 毛伊敏 0583

遥感图像处理

- 特征点和不变矩结合的遥感图像飞机目标识别
曾接贤, 付俊, 符祥 0592
- 结合稀疏识别的自适应Wallis滤波在高分辨率影像控制点匹配中的应用
耿蕾蕾, 孙权森, 纪则轩, 刘佑鑫, 袁凯 0603

第9届和谐人机环境联合学术会议专栏

- 局部特征点的3维模型变换域水印算法
张建明, 周小梅, 王新宇, 詹永照 0613
- 足球视频搜索引擎中的用户偏好挖掘
胡雨成, 于俊清, 黄贤强, 何云峰, 管涛 0622
- 随机亮度差量化的二进制特征描述
庄东晔, 张冬明, 张勇东, 李锦涛 0630
- L_1 优化在网格去噪中的应用
王鹏, 王胜法, 曹俊杰, 李楠楠, 李波, 苏志勋 0637
- 多视点视频双向实时转码技术的设计与实现
孙立峰, 蔡飞飞, 杨士强 0645

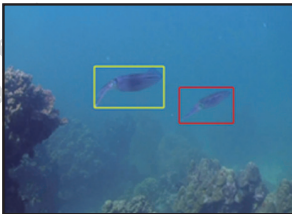
- 第五届数字地球高峰会议征文通知 封三

CONTENTS

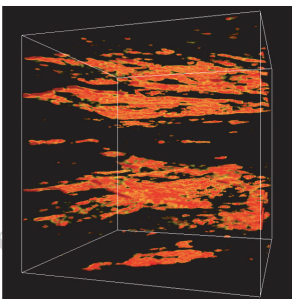
JOURNAL OF IMAGE AND GRAPHICS



Face-wise Poisson blending with color transfer in seamless texturing(P512)



Visual depth based feature calculation and underwater target tracking(P534)



GPU-based volume rendering for full-empty subdata blocks(P577)

Image Processing and Coding

Edge preserving denoising of color image through controlling the gradient vector Jia Di, Meng Xiangfu, Zhang Yifei, Dong Na, Li Sihui	0493
Image denoising of low-rank matrix recovery via joint Frobenius norm Liu Xinyan, Ma Jie, Zhang Xiaomei, Hu Zhaozheng	0502
Face-wise Poisson blending with color transfer in seamless texturing Zhou Yang, Zheng Shunyi, Huang Rongyong, Ma Dian	0512
Improved local directional pattern texture descriptor Liu Haijun, Chang Dongchao, Zhang Lingyu	0520
High capacity reversible watermarking for image interpolation space Wang Jijun	0527

Image Analysis and Recognition

Visual depth based feature calculation and underwater target tracking Wang Huibin, Cheng Yong, Chen Zhe	0534
Combining color names with spatial information for salient object detection Xu Dan, Tang Zhenmin, Xu Wei	0541
Improved generalized Hough transform using artificial fish swarm algorithm in target location Li Zhi, Xie Qiang	0549
The global descriptor for bending non-rigid shapes Li Guoxiang, Xia Guoen	0556

Image Understanding and Computer Vision

Images and geometric model registration in 3D reconstruction of small objects Zheng Shunyi, Wang Xiaohan	0562
Robust descriptor structured by pixel clustering Hui Guobao, Li Dongbo	0569

Computer Graphics

GPU-based volume rendering for full-empty subdata blocks Li Guohe, Duan Zhongxiang, Wu Weijiang, Hong Yunfeng, Liu Zhiyuan, Chen Yuan	0577
--	------

Medical Image Processing

Lossless watermarking algorithm for medical image's tamper detection and recovery with high quality Deng Xiaohong, Chen Zhigang, Mao Yimin	0583
---	------

Remote Sensing Image Processing

Aircraft target recognition in remote sensing images based on distribution of the feature points and invariant moments Zeng Jiexian, Fu Jun, Fu Xiang	0592
Adaptive enhancement via sparse recognition for automatic control point extraction in high-resolution images Geng Leilei, Sun Quansen, Ji Zexuan, Liu Jixin, Yuan Kai	0603

Special Issue on HHME 2013

Transform domain watermarking scheme of 3D models based on local feature points Zhang Jianming, Zhou Xiaomei, Wang Xinyu, Zhan Yongzhao	0613
User preference mining for soccer video search engine Hu Yucheng, Yu Junqing, Huang Xianqiang, He Yunfeng, Guan Tao	0622
Novel binary feature from intensity difference quantization Zhuang Dongye, Zhang Dongming, Zhang Yongdong, Li Jintao	0630
The application of l_1 -optimization in mesh denoising Wang Peng, Wang Shengfa, Cao Junjie, Li Nannan, Li Bo, Su Zhixun	0637
The design and implementation of real time MVC transcoding Sun Lifeng, Cai Feifei, Yang Shiqiang	0645

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2014)04-0541-08

论文引用格式: 徐丹,唐振民,徐威. 融合颜色属性和空间信息的显著性物体检测[J]. 中国图象图形学报, 2014, 19(4): 541-548. [DOI: 10.11834/jig.20140407]

融合颜色属性和空间信息的显著性物体检测

徐丹^{1,2}, 唐振民¹, 徐威¹

1. 南京理工大学计算机科学与工程学院, 南京 210094;
2. 江苏科技大学计算机科学与工程学院, 镇江 212003

摘要: 目的 提出了一种基于颜色属性和空间信息的显著性物体检测算法,并将其用于交通标志的检测。方法 首先,训练颜色属性得到颜色-像素值分布,据此将图像划分为不同的颜色聚簇,每个聚簇的显著性取决于其空间紧致性。其次,将每个聚簇分割为多个区域,用颜色属性描述子表示每个区域,计算区域的全局对比度。最后,同时考虑区域对比度和相应聚簇的空间紧致性,得到最终的显著图。在此基础上,将交通标志的先验知识转换为自上而下的显著性图,形成任务驱动的显著性模型,对交通标志进行检测。结果 在公开数据集上的测试结果表明,该算法取得最高92%的查准率,优于其他流行的显著性算法。在交通标志数据集上的检测取得了90.7%的正确率。结论 提出了一种新的显著性检测算法,同时考虑了区域的颜色和空间信息,在公开数据集上取得了较高的查准率和查全率,在对交通标志的检测中也取得了较好的结果。

关键词: 颜色属性;视觉显著性;区域对比度;空间紧致性;路标检测

Combining color names with spatial information for salient object detection

Xu Dan^{1,2}, Tang Zhenmin¹, Xu Wei¹

1. School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China;
2. School of Computer Science and Engineering, Jiangsu University of Science and Technology, Zhenjiang 212003, China

Abstract: **Objective** Visual saliency refers to the physical, bottom-up distinctness of image details. It is a relative property that depends on the degree to which a pixel or region is visually distinct from its background. The region-based contrast method (RC) achieves good results on public dataset. However, the over-segmentation of the method is considered as a preview contrast computation which contributes to the high precision. In this study, a novel bottom-up salient object detection method based on color names and spatial information is proposed, in which regional contrast and spatial compactness are considered as two factors for saliency evaluation. In addition, we express the prior knowledge of traffic signs with top-down saliency maps, combined with the bottom-up saliency maps, to detect traffic signs from real world images. **Method** First, we learn color names offline by a probabilistic latent semantic analysis model to find the mapping between color names and pixel values. The color names can be used for image segmentation and region description. Second, pixels are assigned into special color names according to their values to form different color clusters in the image. The saliency measure for every color cluster is evaluated by its intra spatial variance. The less the color cluster spreads the more salient it is. Third, every color cluster is divided into some local regions which are represented by color name descriptors. The regional contrast is evaluated by computing color distance between different regions in entire image. Last, the final saliency map is constructed by incorporating the color cluster's spatial compactness measure and the corresponding regional contrast. Note that in most cases, it is not the bottom-up saliency, but the most "interesting" object in an image that attracts attention. In

收稿日期:2013-05-25;修回日期:2013-10-18

基金项目:国家自然科学基金重大研究计划重点项目(90820306)

第一作者简介:徐丹(1981—),女,讲师,南京理工大学模式识别与智能系统专业在读博士研究生,主要研究方向为图象处理和模式识别。

E-mail: xudan_zj@gmail.com

this study, road signs are divided into three categories; every category has special color information. For each category, a class-specific distribution is constructed by the bag-of-words (BoW) model with training images to form the top-down saliency maps. Then the traffic signs are detected from the saliency maps, which are generated by combining the bottom-up saliency maps with top-down saliency maps. **Result** When evaluated using one of the largest publicly available datasets, our method outperforms several existing salient object detection methods with an achieved accuracy of 92%. The ROC curve generated by our method is better than the curves produced by other methods with the area under curve (AUC) of 0.9453. In addition, when tested on the traffic sign dataset constructed by ourselves, our method achieves a detection rate of 90.7%. **Conclusion** In the paper, we propose a novel saliency detection method, in which an image is treated as a composition of 11 basic color names. Every pixel belongs to different color names in a probabilistic manner. Based on this idea, the image is divided into some color clusters, followed by segmenting every cluster into local regions. For saliency measure evaluation, both spatial compactness of the color cluster and the region contrast are considered. Our approach achieves the best results compared with some state-of-the-art methods on the public dataset. Furthermore, it obtains good performance when considering task-dependent application for traffic sign detection.

Key words: color names; visual saliency; region contrast; spatial compactness; road sign detection

0 引言

人类能够很容易地注意到图像中的显著性区域,并分辨出图像中的重要部分。视觉显著性研究试图从神经学^[1-2]、认知心理学^[3]以及计算机视觉^[4-5]的角度探寻人类具有此能力的视觉机制。

视觉显著性是一个相对概念,可通过像素或区域的对比度来体现。Itti 等人^[4]提出的自下而上的显著性模型第一次实现了 Koch 和 Ullman 的基于特征整合理论的计算体系。此模型在不同的尺度空间将图像分解为低层次的属性,如颜色、强度、方向,在每个属性通道上形成显著图,并以中心一周围的差值作为局部对比度的滤波响应。Achanta 等人^[6]提出频率调谐方法,用单个像素和全体像素的平均色的色差定义该像素的显著性值。Cheng 等人提出一种基于区域的全局对比度方法 RC (region-based contrast method),以区域为单位计算颜色对比度^[7]。

以上算法均采用了颜色作为底层特征,本文将判别力更强的颜色属性^[8]概念引入显著性检测算法中。对 11 种基本颜色进行学习,得到其在不同颜色值上的分布概率,以此作为先验知识将图像中的像素分成不同的颜色聚簇。每一个颜色聚簇进而被划分成不同的区域并用颜色属性描述子表示。通过计算区域间颜色描述子的对比度得到区域显著值。

除了对比度,区域的空间紧致性也是衡量显著性的有效方法。Gopalakrishnan 等人^[9]提出颜色聚簇的相对扩展度概念,用某一聚簇和其他聚簇之间的相对距离之和衡量该聚簇的紧致性。Cheng 等

人^[7]在 RC 算法中引入空间信息,增加区域的空间影响效果,形成空间加权的对比度。本文中考虑每个聚簇的空间协方差,以此评价每个颜色聚簇的空间紧致性。最终的显著图依赖于区域对比度以及其所属聚簇的空间紧致性。

需要指出的是,很多情况下不是自底向上、低层次的显著性,而是更“有趣”、更“有意义”的物体引起人们的注意。Shen 等人^[10]提出一种低层次特征与高层次知识相结合的算法,将中心位置、暖色系以及人脸作为先验知识与自下而上的显著性集成在一起,取得了较好的检测结果。本文针对交通标志的特征形成自上而下的显著图,与基于颜色属性的低层次显著图相结合,完成对交通标志的检测。

1 自下而上的显著性

1.1 颜色属性训练

颜色属性是人类为自然界中存在的颜色赋予的语言标签^[8]。Berlin 和 Kay^[11]在语言学研究认为英语中包含 11 种基本颜色:黑色、蓝色、棕色、灰色、绿色、橙色、粉色、紫色、红色、白色以及黄色。对颜色属性的训练即要得到每种颜色属性与像素值之间的对应关系。训练数据来源于显著性数据集 MSRA (Microsoft Research Asia)^[12],每种颜色属性对应 100 幅训练图像。在训练过程中,颜色属性即为其相应训练图像的标签。

采用概率潜在语义分析(PLSA)^[13]算法训练颜色属性。PLSA 源于自然语言处理研究,将文档和词汇之间抽象的潜在主题以概率的形式表达出来。假

设 $D = \{d_1, \dots, d_n\}$ 和 $W = \{w_1, \dots, w_m\}$ 分别为文档和词汇集合, $Z = \{z_1, \dots, z_k\}$ 表示潜在主题集合。PLSA 只考虑单词在文档中的出现频率, 忽略其出现的先后次序。假设“文档—单词”之间以及它们所包含的潜在主题都是条件独立的, 那么, 一个在 $D \times W$ 上的联合概率分布 $p(d, w)$ 可定义为

$$p(d, w) = p(d)p(w|d) \quad (1)$$

$$p(w|d) = \sum_{z \in Z} p(w|z)p(z|d)$$

式中, $p(d)$ 表示文档 d 出现的概率, $p(w|d)$ 是单词 w 在文档 d 中的出现频率, $p(w|z)$ 是给定潜在主题 z 的条件下词汇 w 的出现概率, 而 $p(z|d)$ 则表示文档 d 中潜在主题 z 的分布概率。

PLSA 用于颜色属性训练的方式如下: 图像相当于文档 d , 在 LAB 空间创建其颜色直方图, 直方图中的每个 bin 相当于 PLSA 模型中的单词。直方图表

示了单词在文档中的出现概率, 即 $p(w|d)$ 。图像中存在的颜色属性对应文档中的隐藏主题 z , $p(z|d)$ 表示每幅图像中各种颜色属性的混合系数。最后, 采用最大期望 (EM) 算法估计 $p(z|d)$ 和 $p(w|z)$ 。 $p(w|z)$ 表示颜色属性在颜色值上的分布概率, 如红色对应的颜色值可由 $p(w|z = \text{red})$ 得到。这种颜色属性和像素值之间的关系为所有图像所共享。另外, 根据训练图像所属颜色属性 (标签) 得到 $p(z|d)$ 的先验分布, 图像中与标签相同的颜色被赋予较高的概率。为此, 引入长度为 11 (颜色属性数) 的向量 α_{l_d} (l_d 为图像标签) 和控制参数 c 。如果图像中的颜色 $z = l_d$, $\alpha_{l_d}(z) = c > 1$, 否则 $\alpha_{l_d}(z) = 1$ 。通过改变 c 的取值, 可控制图像标签对 $p(z|d)$ 的影响程度, c 的取值越大, 图像标签对 $p(z|d)$ 的影响越大。训练流程如图 1 所示。

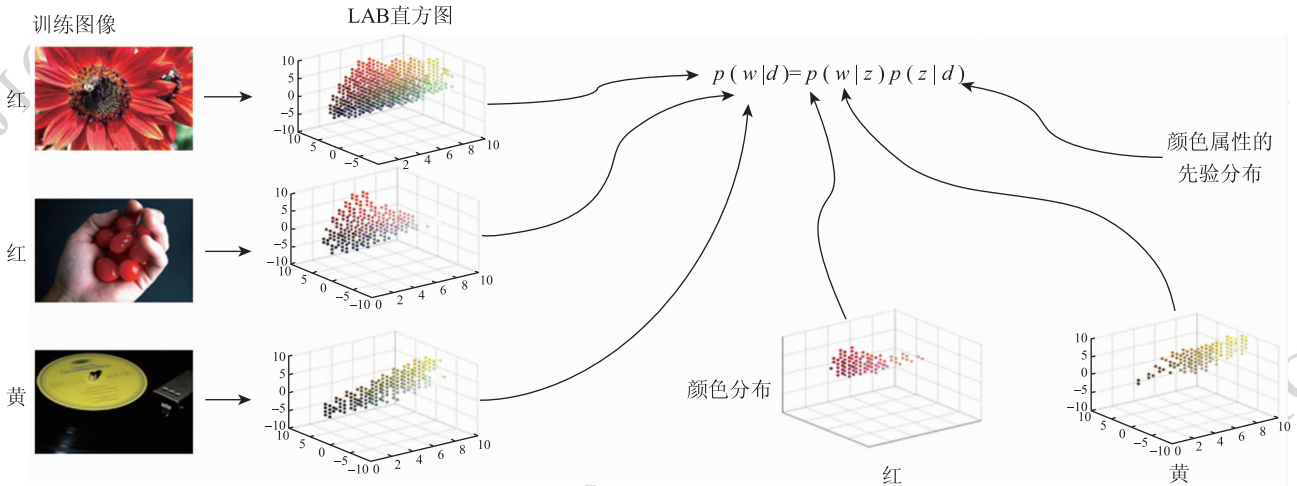


图 1 PLSA 用于颜色属性训练

Fig. 1 Overview of PLSA used in color name training

1.2 图像分割和区域表示

根据训练所得的颜色属性在像素值上的分布概率 $p(w|z)$, 利用贝叶斯定理计算输入图像中每个像素属于不同颜色属性的概率

$$p(z|w) \propto p(z)p(w|z) \quad (2)$$

由于输入图像不存在特定的颜色标签, 式中各颜色属性的先验分布 $p(z)$ 取值相同。根据 $p(z|w)$, 像素 X 属于 11 种颜色属性的概率可表示为 $\{p(cn_1|X), p(cn_2|X), \dots, p(cn_{11}|X)\}$ 。像素分配时, X 被分配至具有最大概率的颜色属性中。如图 2 所示, 花瓣对应的像素被分配成粉色, 因为其属于粉色的概率最大。像素分配后在图像中产生多个颜色聚簇, 每个颜色聚簇对应一个颜色属性。接下来将颜色聚簇分割成多

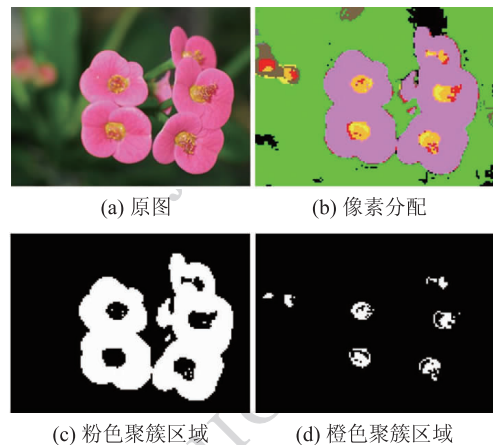


图 2 基于颜色属性的图像分割

Fig. 2 Image segmentation based on color names

个区域。以粉色聚簇为例,属于粉色的像素设置为1,否则为0,这样对聚簇的分割就转换成对二值图像的分割。

用颜色属性描述子(CN)表示区域特征。设 X 是区域 R 中的像素,其对应的各颜色属性概率为 $g(X) = \{p(cn_1|X), p(cn_2|X), \dots, p(cn_{11}|X)\}$,可由 $p(z|w)$ 得到。 $f(g(X))$ 是 $g(X)$ 在 R 中出现的频率,则区域 R 的颜色属性描述子定义为

$$CN_R = \frac{1}{N} \sum_{X \in R} f(g(X)) \cdot g(X) \quad (3)$$

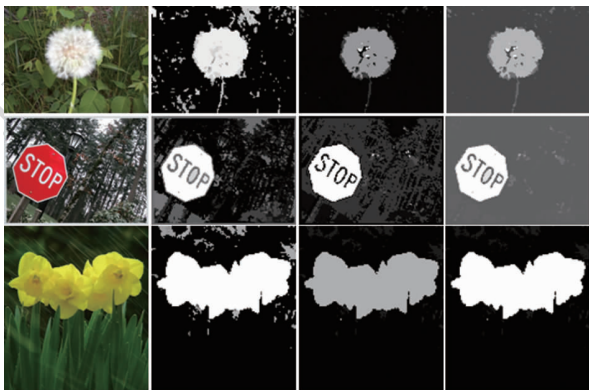
式中, N 为区域 R 中像素的总数。可见,区域的颜色属性描述子是该区域属于不同颜色属性的概率组成的向量,其中的每个元素是区域中所有像素属于该颜色属性的平均概率。Cheng等人^[7]采用RGB直方图作为区域描述子,其维数为1728。在实际操作中,为了提高计算速度只保留了大概85维。而CN描述子只有11维,提高了运算速度。

1.3 显著性检测

在计算显著性的时候考虑两个因素:一是区域与图像中其他区域的对比度,二是区域所属颜色聚簇的空间紧致性。首先计算区域对比度。在上文中将图像划分成多个颜色聚簇,进而将聚簇分割为多个区域,每个区域用颜色属性描述子表示。则聚簇 i 中第 k 个区域的对比度定义为此区域与图像中其他区域颜色描述子之间的距离,即

$$S_{ikcd} = \sum_{r_l \neq r_{ik}} D_c(CN_l, CN_{ik}) \quad (4)$$

式中, CN_l, CN_{ik} 分别是区域 r_l, r_{ik} 的颜色属性描述子, $D_c(CN_l, CN_{ik})$ 是 CN_l 和 CN_{ik} 之间的欧几里德距离。由此产生的显著图如图3所示。



(a) 原图 (b) 基于区域对比度的显著图 (c) 空间紧致性显著图 (d) 最终的显著图

图3 显著性物体检测

Fig. 3 Illustration of salient object detection

颜色聚簇的空间紧致性源于其在图像中的扩展程度,聚簇的扩展程度越大,其显著值越小。相对于显著性物体,背景具有更大的扩展度^[9]。这里用颜色聚簇的内部空间协方差(spatial variance)衡量聚簇扩展度。需要说明的是,每个像素以概率形式属于某一颜色属性。聚簇的水平协方差可表示为

$$V_h(c_i) = \frac{\sum_X p(c_i|X) \cdot |x - M_h(c_i)|^2}{\sum_X p(c_i|X)} \quad (5)$$

$$M_h(c_i) = \frac{\sum_X p(c_i|X) \cdot x}{\sum_X p(c_i|X)} \quad (6)$$

式中, X 是聚簇 c_i 中的像素, $p(c_i|X)$ 表示 X 属于聚簇 c_i 的概率, x 是 X 在图像中的横坐标值。垂直协方差 $V_v(c_i)$ 的定义与此类似。则聚簇 c_i 的协方差为 $V(c_i) = V_h(c_i) + V_v(c_i)$,进行归一化后,聚簇 c_i 与空间分布相关的显著性表示为

$$S_{isd} = p(c_i|X) \cdot (1 - V(c_i)) \quad (7)$$

由于图像是从整个场景中剪切出来的,靠近图像边缘的颜色也可能会有较小的空间协方差^[12]。引入中心权值以减小这种人为的影响,聚簇 c_i 的空间显著性定义为

$$S_{isd} = p(c_i|X) \cdot (1 - V(c_i)) \cdot (1 - D(c_i)) \quad (8)$$

式中, $D(c_i) = \sum_X p(c_i|X) \cdot d_X$, d_X 是像素 X 到图像中心的距离。

区域 k 的最终显著值取决于其所属聚簇 c_i 的空间紧致性和其自身的区域对比度。将空间紧致性显著图 S_{isd} 和区域对比度显著图 S_{ikcd} 归一化,生成区域 k 最终的显著值

$$S_{ik} = S_{isd} \cdot S_{ikcd} \quad (9)$$

2 自上而下的颜色显著性

我国的交通标志大致可分为禁止、警告、指示三大类,每一类均有其特定的颜色特征^[14]。利用这种颜色信息,生成与类别相关的显著图^[15]。每类选取50幅图像作为训练数据,部分训练图像如图4所示。对每幅训练图像 $I^i (i = 1, 2, \dots, 50)$,利用多尺度网格采样和Harris-Laplace兴趣点检测的方法提取其局部特征块 $x_{ij} (j = 1, 2, \dots, M^i)$ 。得到的特征块用CN描述子表示,并以此作为一个视觉单词。用K-Means算法合并词义相近的视觉词汇,构造一



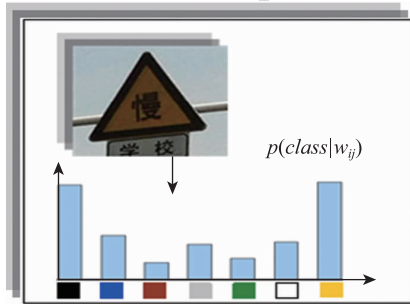
图4 部分训练图像

Fig. 4 Samples of training images

个包含 K 个词汇的单词表 $\mathbf{W}$ 。局部特征 x_{ij} 与类别 $class$ 相关的颜色显著性可定义为其向量化后的单词属于此类别的概率

$$a_t(x_{ij}, class) = p(class | w_{ij}) \quad (10)$$

特定类别分布

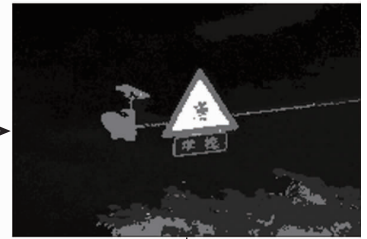


自上而下的显著图

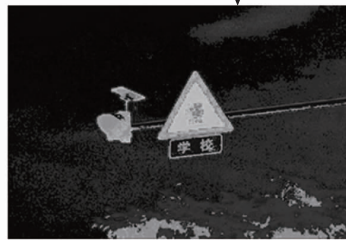
输入图像



自下而上的显著图



最终的显著图



显著物体分割

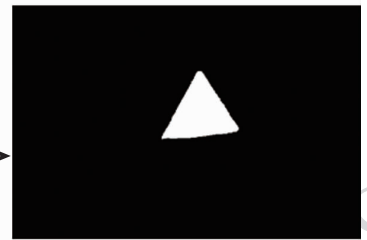


图5 交通标志显著图的生成和分割

Fig. 5 Generation and segmentation of saliency map of traffic sign

式中, w_{ij} 是图像 i 的第 j 个特征 x_{ij} 向量化后的视觉单词。 $p(class | w)$ 可由贝叶斯定理得到, 即

$$p(class | w) \propto p(w | class) p(class) \quad (11)$$

$p(w | class)$ 是训练数据中属于类 $class$ 的所有图像 I^{class} 中视觉单词的概率分布, 即

$$p(w_n | class) \propto \sum_{class} \sum_{j=1}^{M^i} \delta(w_{ij}, w_n) \quad (12)$$

式中, $w_n \in \mathbf{W}$, w_{ij} 是图像 i 的第 j 个特征 x_{ij} 向量化后的视觉单词。如果 $w_{ij} = w_n$, $\delta(w_{ij}, w_n) = 1$, 否则 $\delta(w_{ij}, w_n) = 0$ 。先验分布 $p(class)$ 由训练数据得到。对于输入图像, 计算图像中所有局部特征的 $p(class | w_{ij})$, 就可以构建类别相关的颜色显著图 $a_t(x_{ij}, class)$ 。将类别相关(自上而下)的显著性与自下而上的显著性相结合, 得到最终的显著图, 即

$$a(x_{ij}, class) = a_b(x_{ij}) a_t(x_{ij}, class) \quad (13)$$

式中, a_b 为第 1 节得到的自下而上的显著图 S_{ik} 。算法过程如图 5 所示。

3 实验结果分析

3.1 显著性物体检测实验

颜色属性的训练数据源于 MSRA 数据集, 每个颜色属性对应 100 幅图像。显著性物体检测采用

Achanta^[6] 提供的 1 000 幅图像数据集。实验包括算法本身各组成部分的比较以及与其他算法的比较。本文算法包括基于颜色属性的区域对比度 (CNRC) 计算和聚簇的空间紧致性度量。另外, 将本文算法与以下 5 种影响广泛的显著性检测的算法作比较: GB (graph-based)^[16]、FT (frequency-tuned)^[6]、CA

(context-aware)^[17]、MSSS (maximum symmetric surround saliency)^[18] 以及 RC(region contrast)^[7]。

检测结果评价采用固定阈值法获得二值图,与人工标注的显著性区域比较,得到准确率和召回率,改变阈值的取值($0 \leq T \leq 255$)形成准确率—召回率曲线,如图6所示。单独使用 CNRC 和空间紧致性时最大能够达到 87% 和 84% 的准确度,明显优于 GB、FT、CA 以及 MSSS 算法。结合了 CNRC 和空间紧致性的最终算法与 RC 曲线比较时,在召回率等于 0.8 时交叉,阈值变大时本文算

法的精度大于 RC。这是因为本文算法易于在显著性物体内部得到较高的显著值,但在阈值过小时不能覆盖整个显著性区域。本文算法和 RC 算法均需要将图像分割成区域,计算区域之间的对比度。前者采用判别力更强的颜色属性作为区域描述子。另外,本文算法考虑了区域所属颜色聚簇的空间紧致性。图7是本文算法与其他算法的受试者工作特征(ROC)曲线比较,AUC表示曲线下的面积。图8显示了不同方法产生的显著图的视觉效果。

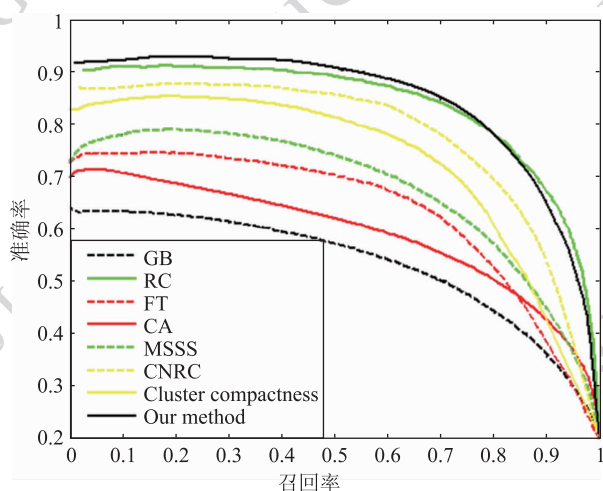


图6 在1000幅图像上与其他算法的准确率—召回率比较

Fig. 6 Precision-recall curves on 1000-image dataset compared with other methods

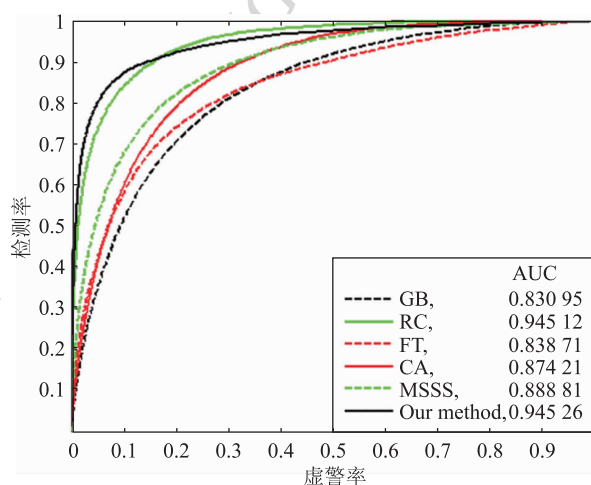


图7 在1000幅图像上与其他算法的ROC曲线比较

Fig. 7 ROC curves on 1000-image dataset compared with other methods

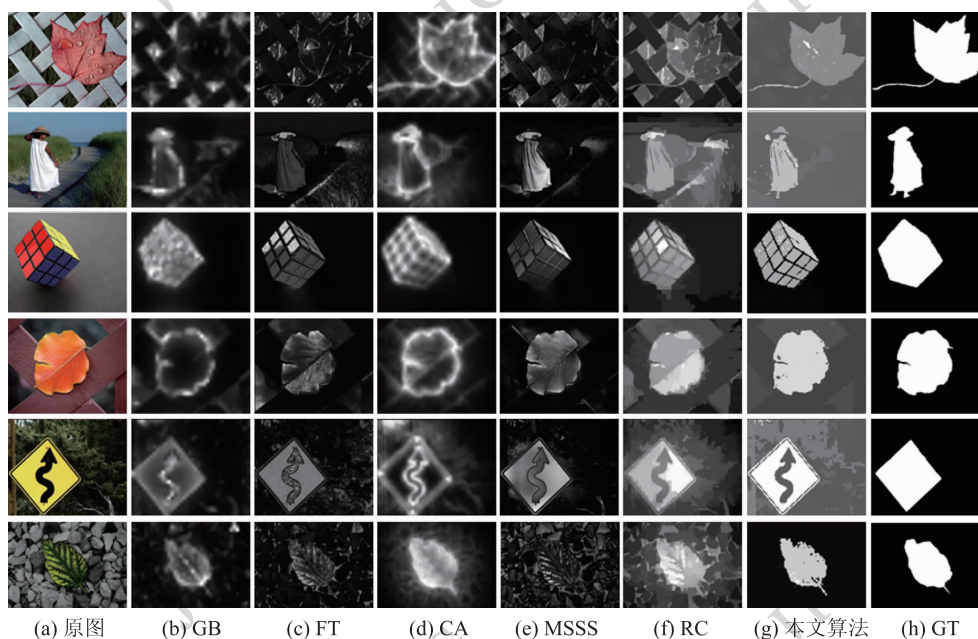


图8 不同算法产生的显著图比较

Fig. 8 Comparison of saliency maps produced by different methods

3.2 交通标志检测实验

1)数据集。为验证本文算法对交通标志检测的有效性,在实景拍摄的交通标志图像库上进行实验。该图像库主要由项目组采集,还有一部分路标由于其通用性,采用了西班牙研究者提供的图像^[19]。图像总数为 282,包含 300 个交通标志,所有图像的大小均为 640 × 480,包括多种天气和光照情况下的交通标志。

2)显著性分割。迭代使用 GrabCut 方法^[7,20]改善二值化显著图后的分割结果。首先进行初始化,选择固定阈值实验中召回率为 95% 时对应的阈值进行显著性分割,将分割结果作为 GrabCut 算法的初始值。然后迭代运行(最多 4 次)GrabCut 改善分割结果。每一次迭代后,用膨胀和腐蚀操作得到新

的 Trimap 以进行下一次迭代。交通标志的检测结果如表 1 所示。图 9 显示了部分交通标志显著性检测和分割的结果,可见,自上而下的显著性信息能够有效地突出图像中的交通标志,减小背景区域的显著值。另外,GrabCut 算法本身是高斯混合模型和图分割的迭代过程,靠近初始显著性物体的区域属于目标物体的概率更大,这有助于分割出完整的显著性物体。

表 1 检测结果统计

Table 1 Statistic of detection result

算法	正确检测路标数
自下而上的显著性算法	197
与自上而下算法结合的算法	272



(a) 原图 (b) 自下而上的显著图 (c) 自上而下的显著图 (d) 最终的显著图 (e) 显著物体分割

图 9 路标检测实例

Fig. 9 Samples of road sign detection

4 结 论

提出一种新的显著性物体检测方法,将图像看成 11 种基本颜色的组合,图像中的每一个像素以概率形式属于不同的颜色属性。由此,每幅图像可以被分成不同的颜色聚簇,每一个聚簇进而被分割成多个区域。然后,计算区域间颜色属性描述子的距离,得到区域的全局对比度。采用颜色属性空间分布的紧致性来衡量聚簇的空间显著性。区域的最终显著性由区域对比度和其所属聚簇的空间紧致性共同决定。在 1 000 幅图像库上的实验表明本文算法优于现有的一些显著性物体检测的算法。另外,还将本文算法与自上而下的显著性算法相结合,实现对交通标志的检测,在实景拍摄的图像库上取得了 90.7% 的检测率。

参考文献 (References)

[1] Desimone R, Duncan J. Neural mechanisms of selective visual attention[J]. Annual Review of Neuroscience, 1995, 18 (1) : 193-222. [DOI: 10. 1146/annurev. ne. 18. 030195. 001205]

[2] Mannan S K, Kennard C, Husain M. The role of visual salience in directing eye movements in visual object agnosia[J]. Current Biology, 2009, 19 (6) : 247-248.

[3] Wolfe J M, Horowitz T S. What attributes guide the deployment of visual attention and how do they do it? [J]. Nature Reviews Neuroscience, 2004, 5 (6) : 495-501. [DOI: 10. 1038/nrn1411]

[4] Itti L, Koch C, Niebur E. A model of saliency-based visual attention for rapid scene analysis[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20 (11) : 1254-1259. [DOI: 10. 1109/34. 730558]

[5] Bai X F, Wang W J, Liang J Y. An active contour model based on region saliency for image segmentation[J]. Journal of Computer

- Research and Development, 2012, 49 (12) : 2686-2695. [白雪飞, 王文剑, 梁吉业. 基于区域显著性的活动轮廓分割模型 [J]. 计算机研究与发展, 2012, 49 (12) : 2686-2695.]
- [6] Achanta R, Hemami S, Estrada F, et al. Frequency-tuned salient region detection [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2009: 1597-1604. [DOI: 10.1109/CVPR. 2009. 5206596]
- [7] Cheng M M, Zhang G X, Mitra N J, et al. Global contrast based salient region detection [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2011: 409-416. [DOI: 10.1109/CVPR. 2011. 5995344]
- [8] Bagdanov A D. Color attributes for object detection [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2012: 3306-3313. [DOI: 10.1109/CVPR. 2012. 6248068]
- [9] Gopalakrishnan V, Hu Y, Rajan D. Salient region detection by modeling distributions of color and orientation [J]. IEEE Transactions on Multimedia, 2009, 11 (5) : 892-905. [DOI: 10.1109/TMM. 2009. 2021726]
- [10] Shen X, Wu Y. A unified approach to salient object detection via low rank matrix recovery [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2012: 853-860. [DOI: 10.1109/CVPR. 2012. 6247758]
- [11] Berlin B, Kay P. Basic color terms: Their universality and evolution [M]. California: University of California Press, 1991.
- [12] Liu T, Yuan Z, Sun J, et al. Learning to detect a salient object [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33 (2) : 353-367. [DOI: 10.1109/TPAMI. 2010. 70]
- [13] Hofmann T. Probabilistic latent semantic indexing [C] // Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM, 1999: 50-57. [DOI: 10.1145/312624. 312649]
- [14] Xu S Q. Traffic sign detection and shape recognition in outdoor environments [J]. Journal of Image and Graphics, 2009, 14 (4) : 707-711. [许少秋. 户外交通标志检测和形状识别 [J]. 中国图象图形学报, 2009, 14 (4) : 707-711.] [DOI: 10.11834/jig. 20090422]
- [15] Khan F S, van de Weijer J, Vanrell M. Modulating shape features by color attention for object recognition [J]. International Journal of Computer Vision, 2012, 98 (1) : 49-64. [DOI: 10.1007/s11263-011-0495-2]
- [16] Harel J, Koch C, Perona P. Graph-based visual saliency [C] // Proceedings of Advances in Neural Information Processing Systems. British Columbia, Canada: Computer Science Bibliography, 2006: 545-552.
- [17] Goferman S, Zelnik-Manor L, Tal A. Context-aware saliency detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34 (10) : 1915-1926. [DOI: 10.1109/TPAMI. 2011. 272]
- [18] Achanta R, Susstrunk S. Saliency detection using maximum symmetric surround [C] // Proceedings of the IEEE International Conference on Imaging Processing. New York: IEEE Press, 2010: 2653-2656. [DOI: 10.1109/ICIP. 2010. 5652636]
- [19] Maldonado-Bascon S, Lafuente-Arroyo S, Gil-Jimenez P, et al. Road-sign detection and recognition based on support vector machines [J]. IEEE Transactions on Intelligent Transportation Systems, 2007, 8 (2) : 264-278. [DOI: 10.1109/TITS. 2007. 895311]
- [20] Rother C, Kolmogorov V, Blake A. Grabcut: interactive foreground extraction using iterated graph cuts [C] // ACM Transactions on Graphics. New York: ACM, 2004: 309-314. [DOI: 10.1145/1015706. 1015720]