

文章编号: 1006-4710(2015)03-0360-06

# 显著目标存在性检测算法研究

李牧<sup>1</sup>, 张晓晖<sup>1</sup>, 杨玉峰<sup>1</sup>, 焦玲玲<sup>2</sup>, 曹向海<sup>2</sup>

(1. 西安理工大学 工程训练中心, 陕西 西安 710048; 2. 西安电子科技大学 电子工程学院, 陕西 西安 710071)

**摘要:** 显著目标检测得到越来越多的关注与研究,但用于显著目标检测的场景图像中有时并不包含任何显著目标,故由这些图像提取的显著图中只会产生虚假的显著目标,从而影响后续的处理。针对这一问题提出一种图像显著目标存在性检测方法,该方法基于中心周边直方图计算出的显著图,提取目标区域与图像中心点距离、目标区域位置分布方差、目标区域在图像边缘的分布、目标区域分布熵、图像显著图的直方图等5种特征进行分类,并利用投票的方式最终确定输入图片是否包含显著目标。基于微软亚洲多媒体图像数据库和网络图像数据库验证了本文算法的有效性,实验表明该方法的检测效果优于现有方法。

**关键词:** 显著目标; 存在性检测; 特征提取; 分类

**中图分类号:** TP391.4      **文献标志码:** A

## Research on the existence detection algorithm of salient objects

LI Mu<sup>1</sup>, ZHANG Xiaohui<sup>1</sup>, YANG Yufeng<sup>1</sup>, JIAO Lingling<sup>2</sup>, CAO Xianghai<sup>2</sup>(1. Engineering Center of Xi'an University of Technology, Xi'an University of Technology, Xi'an 710048, China;  
2. Faculty of Electronic Engineering, Xidian University, Xi'an 710071, China)

**Abstract:** More and more researches are conducted on salient object detection, but some of pictures don't contain any salient object at all, saliency maps extracted from these images tend to contain false salient objects whereby affecting the following-up processing. An approach of existence detection of salient objects is proposed to address this problem. This algorithm extracts five features based on saliency map which is extracted with Center-Surrounding Histogram(CSH) algorithm. These features are distance between salient regions and the center of the image, distribution of salient regions, distribution of salient regions near the edges of the image, entropy of salient regions and histogram of the saliency map of the image. These five features are fed into classifiers and the final result is obtained based on major voting. Experiments based on Microsoft Research Asia Multimedia database and Web Image Database show that the detection outcomes by the proposed method is superior to the existing approaches.

**Key words:** salient object; existence detection; feature extraction; classification

人类大脑和视觉系统会很容易发现图像中某些显著的目标或区域。通过对视觉注意机制的研究,可以使人们合理分配计算资源到更有价值的信息处理任务上(比如在机器人视觉应用方面)。近年来,众多研究者在生物、心理、神经系统、计算机视觉等领域对人类视觉注意系统进行了研究<sup>[1]</sup>。显著目标检测是视觉注意机制研究的一个重要分支,而其实际应用也多种多样,例如图像修剪,图像在移动设备

上的自适应显示,机器人视觉领域的目标识别、障碍物避让等<sup>[2]</sup>。显著性源自于独一无二性、不可预测性、稀有性等,一般归于图像属性,如色彩、梯度、边缘和边界的变化。显著性检测会高亮那些相对于周边区域突出的,在色彩、纹理、梯度等属性方面独特的区域,这是目标检测、目标识别等任务的第一步。但是在实际应用环境中,有些图像有可能不包含任何显著目标,而只包含背景信息,如图1所示。而基

**收稿日期:** 2015-04-08

**基金项目:** 国家自然科学基金资助项目(61405157)。

**作者简介:** 李牧,男,工程师,研究方向为高分辨率雷达信号处理、数字图像处理及多传感器融合技术。

E-mail: xalimu@xaut.edu.cn。

于现有的方法进行显著目标检测,往往会高亮一些毫无意义的区域,因此有必要展开显著目标存在性检测方面的研究。

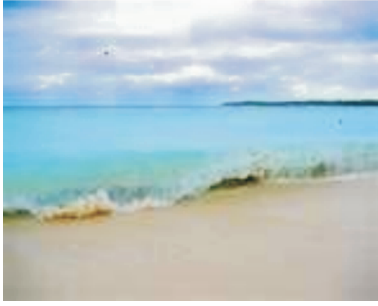


图1 不包含任何显著目标的图像  
Fig.1 Image without salient objects

## 1 相关研究

Itti等<sup>[3]</sup>最早提出了视觉注意模型的实现方法,该模型从输入图像中提取亮度、方向和颜色三个特征图,然后计算像素与其周围像素的差异作为该像素的显著性。Liu等<sup>[4]</sup>利用多尺度对比、中心周边直方图和色彩空间分布三种特征,通过条件随机场学习兴趣区域。Hou等<sup>[5]</sup>考察图像背景的特性,从图像中提取跟这些特性不符的区域就是显著区域,这种方法对光照和噪声不敏感。Achanta等<sup>[6]</sup>提出基于频率调谐的显著目标检测算法,对输入图像进行高斯平滑,每个像素的显著度则为该像素与整幅图像均值的差异。这些方法都采取了自下而上的策略,且不考虑任何特定目标属性。一些自上而下的方法如文献<sup>[7]</sup>,使用目标对象的检测结果作为权重,更好的高亮目标区域。Judd等<sup>[8]</sup>使用诸如人脸、人物、动物等目标检测器来检测显著区域。Shen等<sup>[9]</sup>在显著性计算中综合了高级语义、色彩、位置等先验目标信息,Luo等<sup>[10]</sup>基于显著密度最大化检测显著目标。

然而上述方法都没有考虑图像中是否存在显著目标,而是假定图像中一定存在显著目标,这必然导致在处理一些纯背景图像时,会高亮一些毫无意义的区域作为显著目标。因此有必要开展显著目标存在性方面的研究。早期的研究尝试通过设定阈值来解决这一问题(如文献<sup>[11]~[12]</sup>)。这类方法是基于感兴趣目标的显著值高而背景像素的显著值低的假设,然而这种方法对于那些背景杂乱、变化较大的图像效果会很差,因为这种图像的背景像素也具有较高的显著值,无法找到一个合适的阈值来抑制背景像素。

王鹏等<sup>[13]</sup>提出利用全局特征和多种显著图的信

息来预测显著目标的存在性。Christian Scharfenberger等<sup>[14]</sup>提出了一种利用显著图的直方图特征来区分不含目标图和包含目标图的方法,但上述两种方法对于具有复杂背景的图片来说,效果不是很好。

## 2 算法框架

本文提出了一种基于中心周边直方图(Center-Surrounding Histogram,CSH)的显著图提取方法,并利用目标的位置分布特点以及目标自身的分布特性,提取多种特征来检测显著目标的存在性。提取的特征包括:目标区域到图像中心的距离、目标区域位置分布方差、目标区域在图像边缘的分布、目标区域分布熵,显著图的直方图等5种特征。然后将这些特征分别输入分类器进行分类,最终基于5个分类器的投票结果确定图像是否包含显著目标。图2为本文算法的框图。

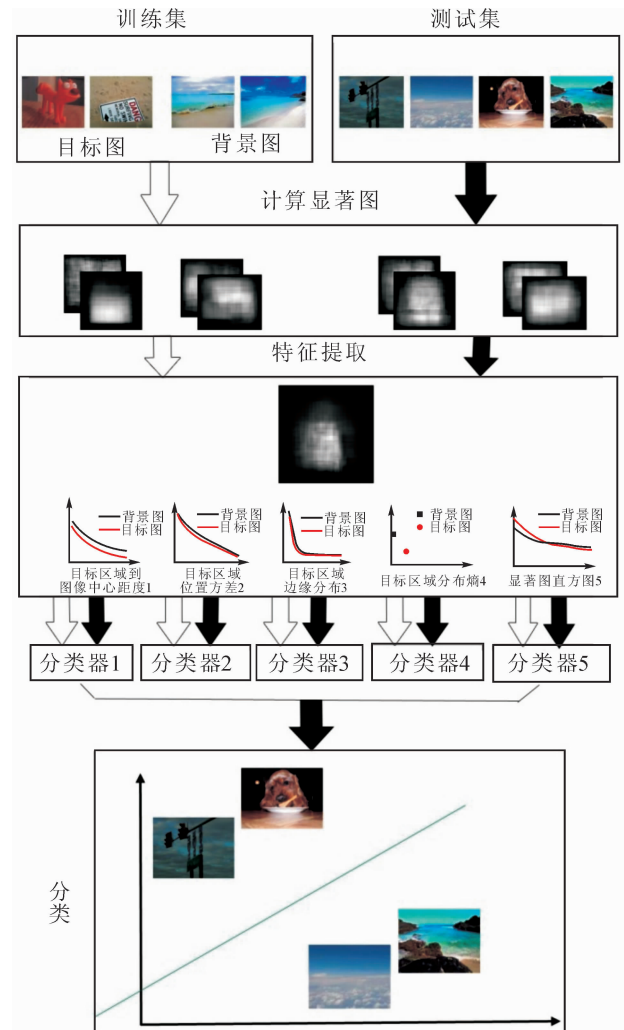


图2 算法框图

Fig.2 Flowchart of our algorithm

其中空心箭头代表训练阶段流程,实心箭头代

表测试阶段流程。

### 3 目标存在性检测算法

#### 3.1 CSH 显著性检测算法

不同的显著性检测方法得到不同质量和性质的显著图,本文基于中心周边直方图提取特征。假设显著目标由一个方形区域  $R$  所包含,围绕该区域构建一个同样面积的方框形区域  $R_s$ ,为了测量该目标与周围区域的差异,这里采用直方图的卡方距离  $\chi^2$  进行计算。方形区域  $R$  的长度( $W$ )、宽度( $H$ )的取值范围为 $[0.1, 0.7]$ ,因为目标往往具有不同的形状及大小,这里  $R$  采用 5 种不同的长宽比 0.5、0.75、1.0、1.5、2.0,根据  $W$ (或  $H$ )以及长宽比求得  $H$ (或  $W$ ),取  $W$  和  $H$  的最小值。对于每个像素  $x$ ,通过改变以其为中心的方形区域  $R(x)$  的长宽比以及大小,寻找出以每个像素  $x$  为中心的最优目标区域  $R^*(x)$ :

$$R^*(x) = \arg \max_{R(x)} \chi^2(R(x), R_s(x)) \quad (1)$$

式中,  $R_s(x)$  是指以像素  $x$  为中心,包含  $R(x)$  的方框型区域。

则可得到像素  $x$  的显著性为:

$$f(x) = \sum_{\{x' | x' \in R^*(x)\}} W_{xx'} \chi^2(R^*(x'), R_s^*(x')) \quad (2)$$

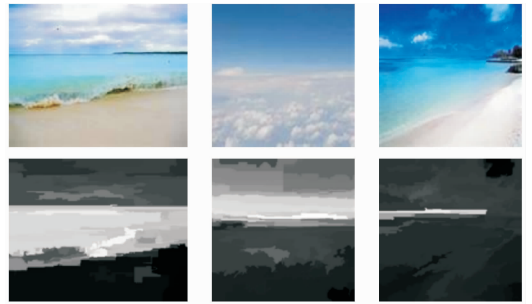
式中,  $x'$  是指以其为中心的方形区域内包含了像素  $x$  的那些像素点,  $R^*(x)$  是以像素  $x$  为中心,并包含像素  $x$  的方形区域,  $R_s^*(x')$  则是围绕在  $R^*(x')$  周围的方框型区域,  $w_{xx'} = \exp(-0.5\sigma_x^{-2} \|x - x'\|^2)$  为高斯权值,方差  $\sigma_x^2$  设为  $R^*(x)$  面积的  $1/3$ ,得到的显著性值最终归一化到 $[0, 1]$ 范围内。

#### 3.2 特征提取

图 3 给出了一组包含以及不包含显著目标的图像及其对应的显著图,其中图 3(a)是不含显著目标的图像及其显著图,图 3(b)是包含显著目标的图像及其显著图。从图中可以明显看出,不含显著目标图像的显著区域比较分散或者根本没有显著区域,而包含显著目标的图像会产生紧凑的、封闭的显著区域,且显著区域通常不会同时与多个图像边界有连接。基于此观察,利用 CSH 显著图构建以下特征。

##### 3.2.1 目标区域和图像中心的距离

对显著图首先进行归一化,然后对显著图设置等间隔阈值进行二值化处理,在每次二值化后,找出目标区域位置,根据(3)式计算目标区域位置和图像中心的距离。图像包含显著目标时该参数较小,不包含显著目标时该参数较大。采用该特征的依据是显著目标通常位于图像中心位置。



(a) 不包含显著目标的图像及其显著图



(b) 包含显著目标的图像及其显著图

图 3 包含目标图像与不包含目标图像以及其对应显著图  
Fig. 3 Images without and with salient objects and their corresponding saliency maps

$$f_{\text{OCD}}(k) = \sum_{n=1}^{N_k} (\text{indr}_k(n) - W/2)/N_k + \sum_{n=1}^{N_k} (\text{indc}_k(n) - H/2)/N_k \quad (3)$$

式中,  $f_{\text{OCD}}(k)$  表示目标区域和图像中心的距离,  $N_k$  表示在第  $k$  个阈值下目标区域的像素数目,  $\text{indr}_k(n)$ 、 $\text{indc}_k(n)$  分别表示在第  $k$  个阈值下,第  $n$  个目标像素的行索引以及列索引。  $k$  取  $1, 2, \dots, 10$ , 则  $f_{\text{OCD}}$  为包含 10 个元素的列向量。

##### 3.2.2 目标区域位置分布方差

对归一化的显著图设置等间隔阈值进行二值化处理,在每次二值化后,计算目标区域位置的分布方差。图像包含显著目标时该参数较小,不包含显著目标时该参数较大。采用该特征的依据是显著目标分布较为紧凑,因此其显著图中对应的目标区域也比较紧凑,而由未包含显著目标的图像得出的显著图中,目标区域的分布则通常较为松散。

$$f_{\text{OPD}}(k) = \sum_{n=1}^{N_k} (\text{indr}_k(n) - E\text{indr}_k)/N_k + \sum_{n=1}^{N_k} (\text{indc}_k(n) - E\text{indc}_k)/N_k \quad (4)$$

式中,  $f_{\text{OPD}}(k)$  表示目标区域位置分布方差,为包含 10 个元素的列向量,  $E\text{indr}_k$  为  $\text{indr}_k(n)$  的均值,  $E\text{indc}_k$  为  $\text{indc}_k(n)$  的均值。

### 3.2.3 目标区域在图像边缘的分布

同样利用阈值对显著图进行等间隔的二值化,在每次二值化后,计算显著图四个边缘 10 个像素宽度内所有像素值的和与显著图周长的比值,用以表征图像边缘部分显著性分布特性。图像包含目标时该参数较小,不包含目标时该参数较大。采用该特征的依据是显著目标通常不与多个边界同时接触,因此边界区域的显著值通常较小。

$$f_{\text{OBD}}(k) = \frac{E_{tk} + E_{bk} + E_{lk} + E_{rk}}{2(W + H)} \quad (5)$$

式中, $f_{\text{OBD}}(k)$ 表示目标区域在图像边缘的分布,同样为具有 10 个元素的列向量, $E_{tk}$ 、 $E_{bk}$ 、 $E_{lk}$ 、 $E_{rk}$ 分别为在阈值  $k$  下紧邻显著图上边缘、下边缘、左边缘、右边缘的 10 个像素宽度范围内目标像素的显著性之和。

### 3.2.4 目标区域分布熵

计算显著图目标区域的分布熵。熵可以表明显著图包含信息量的大小,从侧面反映了显著值的分布情况,在显著图比较均匀时该值较小,反之则较大。该特征的提取依据如下:包含显著目标的图像对应的显著图中,目标区域和背景区域都较为平滑,而未包含显著目标的图像对应的显著图则通常不具有这个特点。

$$f_{\text{OE}} = - \sum (p \log_2 p) \quad (6)$$

式中, $f_{\text{OE}}$ 表示目标区域分布熵, $p$ 为显著值的分布概率。

### 3.2.5 显著图的直方图

该特征  $f_{\text{Hist}}$ 由显著图的 20 维直方图构成,由图 3 中显著图的分布特性可知,包含目标显著图的直方图相对于不含目标显著图的直方图,通常在低显著值处会有一个更为明显的峰值。

因为图像中背景和目标分布的复杂性,单独靠一种特征很难对所有图像进行有效的分类,因此本文综合利用上述 5 种特征进行分类。

## 3.3 利用随机森林分类器检测目标存在性

本文利用随机森林分类器(Random Forest, RF)<sup>[15]</sup>进行图像的分类。将 3.2 节中提取的 5 种特征分别输入到 5 个分类器中进行分类,输出 0 表示不包含目标,1 表示包含目标,采用投票的方式确定最终的分类结果:

$$\text{ClassType} = \text{round}\left(\frac{C_1 + C_2 + C_3 + C_4 + C_5}{5}\right) \quad (7)$$

式中, $\text{ClassType}$ 表示待检测图像的类属,round 表示四舍五入的取整函数。 $C_1$ 、 $C_2$ 、 $C_3$ 、 $C_4$ 、 $C_5$ 为分别

输入 5 个特征值时的分类结果。计算  $C_1$ 、 $C_2$ 、 $C_3$ 、 $C_4$ 、 $C_5$  的分类结果均值,均值大于 0.5 表示该图像可能包含目标,则给出分类值 1;均值小于 0.5 表示该图像可能不包含目标,则给出分类值 0。

由于一般图像数据库中含目标图像和不含目标图像的数量高度不平衡,而 RF 分类器往往更偏向于数量较大的类,因此训练样本不平衡会影响分类器的准确性。一般有两种方法解决此类问题:一种是调整不同类别的权值,另一种是对数量大的类别进行抽取<sup>[16]</sup>。在本实验中采用抽取的方式得到一个平衡的 RF 分类器。

## 4 实验结果

本文基于两个数据库对本文算法进行评估。其中一个包含 10000 幅有目标图像的微软亚洲多媒体图像数据库(Microsoft Research Asia Multimedia, MSRA)。另一个是由网络图像数据库(Web Image Database)中仔细挑选出的 4115 幅不包含显著目标的背景图像组成的数据库,这些背景图像是在网络上通过搜索“沙漠”、“海洋”、“森林”等关键词,并经过进一步筛选得到的。因此,所用数据库由 10000 幅包含目标的图像和 4115 幅不含目标的图像构成,所有图像尺寸均调整为  $130 \times 130$  像素。

RF 分类器参数设置:树的数目为 200,最小叶节点设为 4。仿真实验采用的平台是:处理器为 Intel Core i7 4700MQ,主频为 2.4 GHz,内存 4 GB 的 PC 机,软件平台为:Windows 7 旗舰版 32 位操作系统和 Matlab R2014a。

同时,为了更全面的进行性能评估,又选用了另外三种常用的显著图提取同样特征作为对比,分别为区域对比(Region Contrast, RC)显著图<sup>[17]</sup>、多尺度对比(Multi-scale Contrast, MC)显著图以及色彩空间分布(Color Spatial-Distribution, CSD)显著图。

在构建的图像库中,每次分别从含目标图像集和背景图像集中随机选择 1000 幅图像用于训练,剩余图像用于测试。

### 4.1 不同显著图分类对比

本文用 10 折交叉验证(10-fold cross validation)的方法计算准确度(Accuracy)、精确度(Precision)、召回率(Recall)和  $F$ -measure<sup>[18]</sup>,来评价该算法的性能。以上四种指标定义如下:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (9)$$

$$Recall = \frac{TP}{TP + FN} \quad (10)$$

$$F\text{-measure} = \frac{(1 + \beta)Precision \times Recall}{\beta \times Precision + Recall} \quad (11)$$

式中,  $TP$  为分类正确的含目标图像数量,  $TN$  为分类正确的不含目标图像数量,  $FP$  为分类错误的含目标图像数量,  $FN$  为分类错误的含目标图像数量,  $\beta$  设为 2。

表 1 给出了不同显著图结合 RF 分类器的结果。从表 1 中可以看出, 用 CSH 显著图提取的特征向量能够产生更高的分类准确度, 达到 92.2%, 同时精度、召回率以及  $F\text{-measure}$  分别达到 93.5%、96.1% 和 93.96%, 而 MC 显著图分类准确度仅仅为 69.4%。

表 1 不同显著图分类性能对比

Tab. 1 Performance comparison of different salient detection algorithms

显著图	准确度/%	精确度/%	召回率/%	$F\text{-measure}/\%$
CSH	92.2	93.5	96.1	93.9
CSD	73.6	89.6	72.9	86.3
RC	71.1	86.2	74.2	83.9
MC	69.4	87.2	68.8	83.5

#### 4.2 不同分类器对准确度的影响

实验中也研究了不同分类器对分类结果的影响, 分别采用支持矢量机分类器(SVM)、贝叶斯分类器(Bayesian)以及 RF 分类器进行训练和分类, 如表 2。从表 2 中可以看出, 对于贝叶斯分类器而言, 不同的显著图对它的分类结果影响不大, 准确度基本保持在 74% 左右。而对于 RF 分类器和 SVM 分类器而言, 不同的显著图对其分类结果影响较大, CSH 显著图能够使两种分类器的准确度分别达到 92.18% 和 91.72%, 而基于 MC 显著图的准确度只有 69.37% 和 66.81%。RF 分类器基于四种显著图的分类性能都比 SVM 要好一些。

表 2 不同分类器对分类准确度影响

Tab. 2 The Influences of different classifiers

显著图	RF/%	SVM/%	Bayesian/%
CSD	73.6	72.5	71.8
CSH	92.1	91.7	75.8
RC	71.8	70.4	75.8
MC	69.3	66.8	74.2

#### 4.3 与现有方法的性能对比

这里将本文的算法与现有两种显著目标存在性检测算法进行对比。一种是文献[14]中提出的算法(简称 HFED), 该方法利用显著图的直方图特征进行显著目标存在性的判别。另一种是文献[13]中提

出的算法(简称 GFED), 利用全局特征和多种显著图的信息来判断显著目标的存在性。表 3 为本文方法与现有两种方法的性能对比。结果表明, 本文的算法要优于这两种算法。

表 3 本文方法与现有两种方法性能对比  
Tab. 3 Performance comparison between existing algorithms and our proposed algorithm

方法	GFED/%	HFED/%	本文算法/%
准确度	81.20	88.10	92.17
精度	86.20	89.60	93.50
召回率	87.20	90.01	96.10
$F\text{-measure}$	86.51	88.99	93.90

## 5 结 论

本文围绕图像显著目标存在与否展开研究, 通过分析存在显著目标与不存在显著目标的显著图之间的差异, 提取了 5 种不同特征, 并利用不同分类器对显著目标存在性进行了全面评估。实验表明, 基于中心周边直方图显著图提取特征, 并结合随机森林分类器来检测显著目标的存在性是一种可靠的方法。

#### 参考文献:

- [1] Borji A, Itti L. State-of-the-art in visual attention modeling [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(1): 185-207.
- [2] Borji A. What is a salient object? a dataset and a baseline model for salient object detection [J]. IEEE Transactions on Image Processing, 2015, 24(2): 742-756.
- [3] 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.
- [4] 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.
- [5] Hou X, Zhang L. Saliency detection: a spectral residual approach [C]// IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, 2007: 1-8.
- [6] Achanta R, Hemami S, Estrada F, et al. Frequency-tuned salient region detection [C]// IEEE Conference on Computer Vision and Pattern Recognition, Miami, 2009: 1597-1604.
- [7] Frintrop S, Backer G, Rome E. Goal-directed search with a top-down modulated computational attention system [C]// Lecture Notes in Computer Science, Vienna, 2005: 117-124.
- [8] Judd T, Ehinger K, Durand F, et al. Learning to pre-

- dict where humans look [C]// IEEE 12th International Conference on Computer Vision, Kyoto, 2009: 2106-2113.
- [9] Shen X, Wu Y. A unified approach to salient object detection via low rank matrix recovery [C]// IEEE Conference on Computer Vision and Pattern Recognition, Providence, 2012: 853-860.
- [10] Luo Y, Yuan J, Xue P, et al. Saliency density maximization for object detection and localization [C]// Lecture Notes in Computer Science, Queestown, 2011: 396-408.
- [11] Xu T, Chenkov N, Kühnlenz K, et al. Autonomous switching of top-down and bottom-up attention selection for vision guided mobile robots [C]// International Conference on Intelligent Robots and Systems, ST. Louis, 2009: 4009-4014.
- [12] Xu T, Wu H, Zhang T, et al. Environment adapted active multi-focal vision system for object detection [C]// IEEE International Conference on Robotics and Automation, Kobe, 2009: 2418-2423.
- [13] Wang P, Wang J, Zeng G, et al. Salient object detection for searched web images via global saliency [C]// IEEE Conference on Computer Vision and Pattern Recognition, Providence, 2012: 3194-3201.
- [14] Scharfenberger C, Waslander S L, Zelek J S, et al. Existence detection of objects in images for robot vision using saliency histogram features [C]// International Conference on Computer and Robot Vision, Regina, 2013: 75-82.
- [15] Breiman L. Random forests [J]. Machine Learning, 2001, 45(1): 5-32.
- [16] Chen C, Liaw A, Breiman L. Using random forest to learn imbalanced data [R]. University of California, Berkeley, San Francisco, 2004.
- [17] Cheng M, Mitra N J, Huang X, et al. Global contrast based salient region detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(3): 569-582.
- [18] Powers D M. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation [J]. Journal of Machine Learning Technologies, 2011, 2(1): 37-63.

(责任编辑 周蓓)