

文章编号 1004-924X(2017)03-0772-07

结合局部特征及全局特征的显著性检测

蔡 强, 郝佳云*, 曹 健, 李海生

(北京工商大学 计算机与信息工程学院, 北京 100048)

摘要:针对目前大多数显著性检测方法中采用背景种子以及局部区域对比度显著性检测模型的缺点,本文提出了一种综合考虑局部特征以及全局特征的显著性检测算法。在对图像进行分割之后,算法首先融合了采用多特征方式生成的背景显著图与采用前景区域对比度方式生成的前景显著图,之后使用高斯滤波器对融合后的结果进行优化形成局部特征显著图。其次,在局部特征显著图的基础上提取多种特征的样本集合进行训练,从而得到全局特征显著图。算法最后将第一步生成的局部特征显著图与全局特征显著图进行结合生成最终的显著图。实验部分验证了算法各部分的有效性,并且在 3 个公开数据集上对文章方法与近年来优秀的显著性检测算法进行了对比,实验结果显示,本文算法在 CSSD 数据集上的准确率、召回率以及 F-measure 分别达到了 0.837 5、0.743 4 和 0.813 7,在其它数据集上也有良好表现。实验表明,本文算法能够有效抑制背景区域,并且高亮前景区域,更好地检测出显著目标。

关键词:多特征;显著性检测;高斯滤波器;局部特征;全局特征

中图分类号:TP391 **文献标识码:**A **doi:**10.3788/OPE.20172503.0772

Salient detection via local and global feature

CAI Qiang, HAO Jia-yun*, CAO Jian, LI Hai-sheng

(School of Computer and Information Engineering,
Beijing Technology and Business University, Beijing 100048, China)

* Corresponding author, E-mail: haojiayun_happy@163.com

Abstract: Due to the most of existing salient detection methods have some disadvantages on using background seeds and local area contrast salient detection model, a visual saliency detection algorithm named salient detection, which combines local feature and global feature, was proposed. After image segmentation, the algorithm firstly applied a background image created by multi-feature methods and a foreground saliency image created by foreground area contrast method, then, the fusion results was optimized by using Gaussian filter and the local feature saliency image was formed. Secondly, the sample set of various features was collected based on the local feature saliency image for practice and finally the global feature saliency image was obtained. At last, it combined the local feature saliency image produced in the first step with the global feature saliency image and created the final saliency image. In part of experiment, the proposed algorithm showed great results of precision, recall rate and F-measure on CSSD data set, with values of 0.837 5, 0.743 4 and 0.813 7 respectively, the performance on other data set was also perfect. The results show that the proposed algorithm can effectively suppress the background area, highlight foreground area and detect the salient target better.

Key words: multiple feature; salient detection; gaussian filter; local feature; global feature

收稿日期:2016-12-21; **修订日期:**2017-01-15.

基金项目:北京市教委科研计划一般项目(No. SQKM201610011010);北京市自然科学基金资助项目(No. 4162019);北京市科技计划课题(No. Z161100001616004)

1 引言

近年来,显著性检测在计算机视觉领域愈发重要。显著性检测的本质是一种视觉注意模型,它利用视觉注意机制得到图像中最容易引起注意的显著部分,并用一幅灰度图像表示其显著度。显著性检测通常作为预处理步骤用于降低计算复杂度,因此在图像分割^[1],视觉跟踪^[2]及图像分类^[3]等视觉工作中都有重要应用。

目前,显著性检测的方法可以分为两类:自下而上目标驱动型的方法以及自上而下任务驱动型的方法。在一些采用自下而上的算法中,Cheng^[4]采用基于直方图对比度以及局部对比度的方法,用空间距离和区域像素加权颜色差异来确定区域的显著值。在文献[5]中,Sun等人从视觉视点出发,致力于寻找具有超高斯分布的位置以此得到显著图。Perrazzi^[6]等人将颜色自身的独立性与颜色分布对显著图的贡献结合起来,采用滤波器进行加速,得到的显著图结果比较均匀。

考虑到采用自下而上的显著性检测算法可能会造成全局信息缺失,因此后来出现了一些采用自上而下的算法,在文献[7]中,采用贝叶斯模型,融合了自下而上和自上而下的方法进行显著性计算。Borji^[8]等人将方向,颜色通道直方图等底层特征及水平线检测,人脸检测等高层特征结合起来,用于表示每个像素。Tong^[9]使用多特征对比度计算得出自下而上的显著图,然后融合经过自上而下的显著图得到了最终显著图。另外,还有很多算法使用了机器学习^[10-11]中的方法,在文献[12]中,将显著性滤波算法与改进脉冲耦合神经网络算法相结合,提高了显著物体检测的准确性。

自下而上的目标驱动型方法主要针对底层信息以及细节信息,通过利用底层线索来构建显著图。自上而下的任务驱动型方法则主要针对全局信息以及样本中具有代表性的特征进行训练。然而,自下而上的方法难以关注全局形状信息,自上而下的方法又可能会比较粗糙。

为了更好地利用两类方法的优点,抑制两者的缺点,本文提出一种结合局部特征以及全局特征的显著性检测模型。通过结合背景种子和前景种子并且采用多特征的方式生成了自下而上的显

著图,另外,采用自上而下方法得到了更加关注全局信息的显著图,将二者融合后形成最终显著图,有效地提高了检测显著物体的准确率。

2 结合多尺度多特征的显著性检测

本文算法的实现框图如图1所示,先将输入图像进行分割得到超像素,通过多特征^[13]计算每个超像素与边缘超像素之间的对比度,进而得到基于背景的显著图,之后通过自适应阈值对图像进行分割来选择前景种子,通过计算与前景种子之间的相似度得到基于前景的显著图,将得到的基于前景及背景的显著图进行融合后,使用高斯滤波器进行优化,得到局部特征显著图。之后,通过对输入图像进行训练学习得到一个强分类器,利用其得到全局特征显著图。最后将第一步得到的局部特征显著图与第二步进行训练学习后得到的全局特征显著图进行融合,从而得到最终的显著图。

2.1 多特征的背景显著图

为了更好地抓住输入图像的结构信息,本文采用超像素而不是像素作为最小处理单元,采用SLIC算法^[14]来进行超像素分割。基于边缘优先性^[15],假设图像的边缘区域更容易成为背景。通过RGB、CIELab以及LBP 3种特征计算不同区域对应的欧氏距离,得到每个区域的显著值。将输入的图片分割成 K 个超像素,则对于区域 r_i (其中, $i \in \{1, 2, \dots, K\}$)与背景 b_j (其中, $j \in \{1, 2, \dots, K_b\}$, K_b 是边缘超像素的个数),可得到区域的显著值为:

$$M_0 = W(r_i) \times \sum_F \left[\frac{1}{K_b} \sum_{j=1}^{K_b} d_F(r_i, b_j) \right]. \quad (1)$$

其中: $W(r_i)$ 是通过 r_i 中心与图像中心归一化空间距离计算得到的中心先验权重。 F 是RGB、CIELab、LBP三种特征中的一种, $d_F(r_i, b_j)$ 是图像超像素区域 r_i 与边缘超像素区域 b_j 之间在特征 F 下的欧式距离。通过公式(1)可以得到每个超像素的显著值,将超像素区域内的像素设定为该值,从而得到像素水平的显著值 $\hat{M}_0(x, y)$,其中 (x, y) 代表像素的坐标,即得到显著图 S_0 。

2.2 基于前景种子的显著图

显著图 S_0 中一些属于背景的超像素可能并没有作为背景种子包含在边缘区域中,这就会导

致存在一定的背景噪音,因此本文在显著图 S_0 的基础上采用基于前景种子的方式抑制背景噪音。

首先使用文献[16]中的方法将基于背景的显著图进行二值化,之后选择在基于背景的显著图下显著值大于阈值的超像素作为前景种子。因为与前景种子更加相似的区域更容易成为显著区域,所以,通过计算每一个超像素与前景种子之间的空间、颜色相似度来计算其显著值:

$$M_1 = \sum \frac{\lambda}{d_c(r_i, f_j) + \vartheta d_l(r_i, f_j)}. \quad (2)$$

其中: $d_c(r_i, f_j)$ 和 $d_l(r_i, f_j)$ 是超像素 r_i (其中, $i \in \{1, 2, \dots, K\}$) 与前景 f_j (其中, $j \in \{1, 2, \dots, K_f\}$, K_f 是前景种子的个数) 的颜色和欧式空间距离。在实验中,权重 λ 和 ϑ 均被置为 1, 用于平衡颜色和空间距离在计算显著值时的重要性。

与计算基于背景的显著图类似,得到像素级别的显著值 $\tilde{M}_1(x, y)$ 以及基于前景的显著图 S_1 。

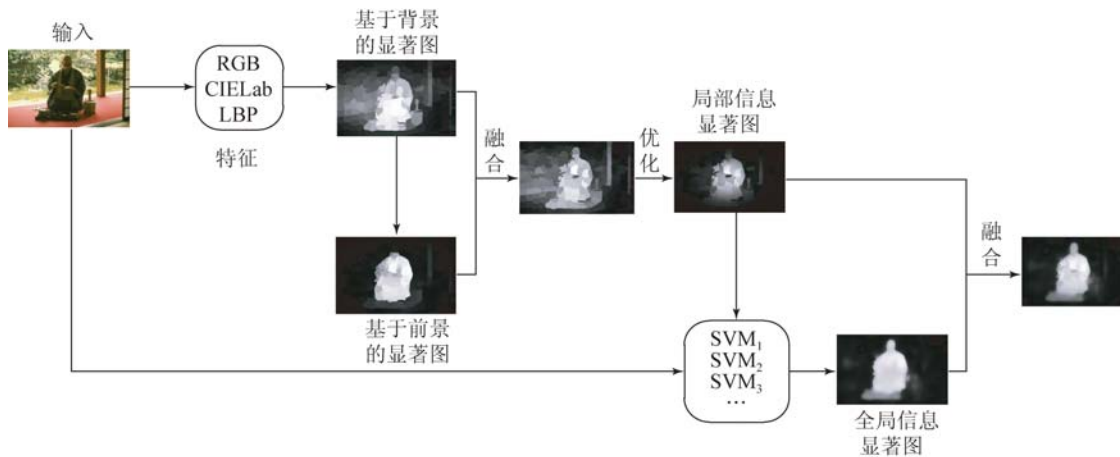


图 1 本文算法主要步骤

Fig. 1 Main steps of our algorithm

2.3 高斯滤波器优化

在 2.1 小节中得到的显著图可以更好地高亮物体区域,而 2.2 小节中得到的显著图可以更好地抑制背景噪音,将两者通过联合函数进行融合可以得到融合后的显著图:

$$\tilde{M}_2(x, y) = \tilde{M}_0(x, y) * (1 - \exp(-\beta * \tilde{M}_1(x, y))), \quad (3)$$

其中, β 用于平衡基于前景和基于背景的显著图在最终显著图中的比重,实验中将其置为 6。

为了进一步抑制背景噪音,本文采用高斯滤波器对显著图进行优化。则经过高斯加权图为 $G(x, y)$ 的滤波器进行优化后的显著图为:

$$\tilde{M}_3(x, y) = \tilde{M}_2(x, y) \times G(x, y), \quad (4)$$

$$G(x, y) = \exp(-9((x - d_x)/\omega)^2 - 9((y - d_y)/h)^2). \quad (5)$$

其中: ω 和 h 是输入图像的宽和高, (d_x, d_y) 是图像中心的位置坐标。通过公式(4)和(5)得到了经过高斯滤波器优化后的显著图 S_3 , 经过优化之后,背景噪声得到了抑制。

2.4 多核分类器学习模型

使用多个弱分类器进行迭代学习得到一个强分类器。因为多核模型相对于单核模型更加灵活,与支持向量机结合能获得比单核更优的性能^[17]。所以本文使用 Sigmoid, 径向基核, 线性核以及多项式核 4 个核函数进行加权求和构造多核函数:

$$K(x, z) = \sum_{j=1}^M \beta_j f_j(x, z), \beta_j \geq 0, \sum_{j=1}^M \beta_j = 1, \quad (6)$$

其中: β_j 是核函数 f_j 对应的权重, M 是具备不同核函数或采用不同特征的弱分类器的数量。

为了得到最后的强分类器,在训练和测试样本时,本文使用输入图片的前景部分作为正样本,背景部分作为负样本进行训练。对于每一个弱分类器,假设样本总数为 N ,那么初始化样本的权重相同,均为 $\omega_i(k) = 1/N, k = 1, 2, \dots, N$ 。把每一次的迭代结果都当做一个弱分类器,那么由公式(6)可知,经过 I 次迭代后,可以得到相应的 I

个 β_j 用于计算最后的强分类器。弱分类器在最终的强分类器计算过程中的权重表达式为:

$$\beta_j = \alpha_j \cdot \frac{1}{2} \left[1 + \operatorname{sgn} \left(\ln \frac{1 - e_j}{e_j} \right) \right], \quad (7)$$

$$\alpha_j = \frac{1}{2} \ln \frac{1 - e_j}{e_j}. \quad (8)$$

其中: e_j 是弱分类器的误差率,分类的误差率越小,权重值越大,即该弱分类器在最终的分类器中作用越大。在下一轮迭代中,样本具有新的权重分布,样本权重更新表达式为:

$$\omega_{i+1}(k) = \frac{\omega_i(k)}{Z_i} \cdot \exp(-\beta_i l_k y_i(r_k)), \quad (9)$$

其中: l_k 是样本的二值标签, $y_i(r_k)$ 是弱分类器中每一个单核单特征弱分类器所对应的目标函数,可以通过训练学习的过程得出。 Z_i 是规范化因子,其表达式为:

$$Z_i = 2 \sqrt{e_i(e_i - 1)}. \quad (10)$$

最后的强分类器是经过 I 次迭代后的 I 个弱分类器加权得到的。通过分类器,最终可以得到像素级别的显著值 $\widetilde{M}_4(x, y)$ 。

本文最终的显著图由底层细节信息相对丰富的显著图 S_3 和多核分类器模型得到的显著图组成。为了平衡两种显著图的优缺点,本文将两种显著图进行融合:

$$\widetilde{M}(x, y) = \eta \widetilde{M}_3(x, y) + (1 - \eta) \widetilde{M}_4(x, y), \quad (11)$$

其中: η 为平衡因子,实验中将其设置为 0.7。

3 实验结果及分析

在 3 个公开的数据集上对本文所提出的算法进行评价,并将其与 11 种显著性算法 IT^[18], GB^[19], SR^[20], FT^[1], LR^[21], SF^[22], GS^[3], GC^[23], BL^[9], AMC^[24], GBMR^[25] 进行了对比。

本文采用准确率-召回率 (PR) 曲线、F-measure 作为评价指标对算法进行定量评价。计算 PR 曲线的过程又称为固定阈值分割,通过在 0~255 之间设定阈值,可以对显著图进行二值化,将二值化之后的结果与真值作比较,可以计算测试数据集中所有图片的平均准确率以及召回率,通过对比 ROC 曲线下的面积 (AUC) 来评估

算法的好坏。

计算 F-measure 的过程又称为自适应阈值分割,即所选取的阈值由待测评的显著图本身决定。F-measure 作为整体性能的评价机制,用于综合评定准确率以及召回率。其计算公式如下:

$$F_\beta = \frac{(1 + \beta^2) \times \text{precision} \times \text{recall}}{\beta^2 \times \text{precision} + \text{recall}}, \quad (12)$$

其中: precision 为准确率, recall 为召回率,根据最早应用 F-measure 的文献[26]建议,设置 $\beta = 0.3$ 。

本文提出的算法中每一步的生成结果均对下一步有不同程度的影响。因此,对算法中每一个部分生成的显著图进行定量测量,判断每一步对于最终生成显著图的影响程度,如图 2 所示。其中, background 表示基于背景生成的结果, foreground 表示基于前景生成的结果, union 表示融合前景及背景区域结果后生成的显著图, our 表示结合局部及全局特征结果后生成的显著图。由图 2 可以看出,本文方法最终生成的显著图结果具有更高的准确率和召回率。

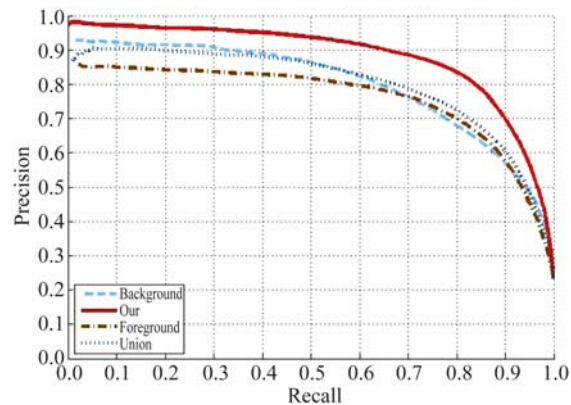


图 2 在 CSSD 图像集上对算法各个步骤的 PR 曲线

Fig. 2 PR curves of each step of proposed algorithm on CSSD data set

本文在 CSSD, ECSSD, DUT-OMRON 3 个数据集上对本文提出算法与部分现有优秀算法进行了定性以及定量比较。图 3 是本文算法与现有的比较流行的 8 种算法的对比图。对比图表明,本文方法能够有效抑制背景区域,并且高亮前景区域,能够更好地检测出显著目标。

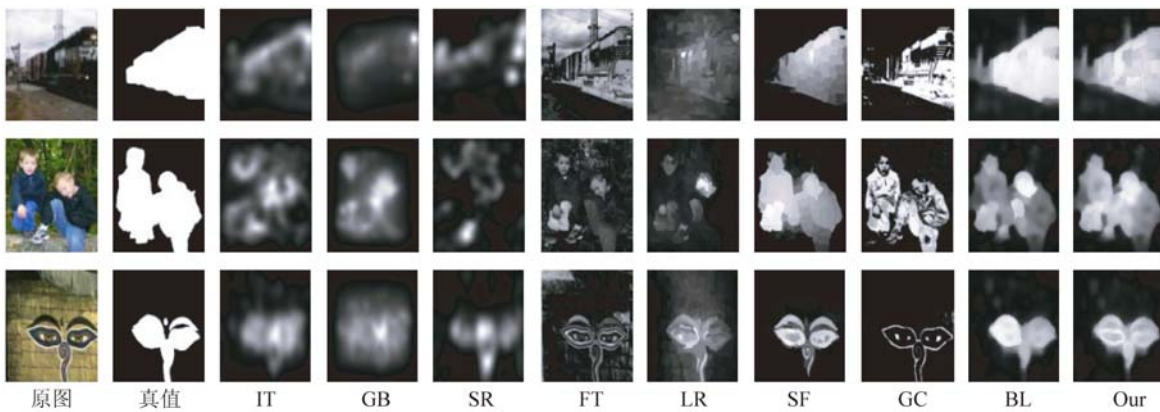


图 3 本文算法与现有 8 种算法对比结果

Fig. 3 Comparison results of proposed algorithm with other eight methods

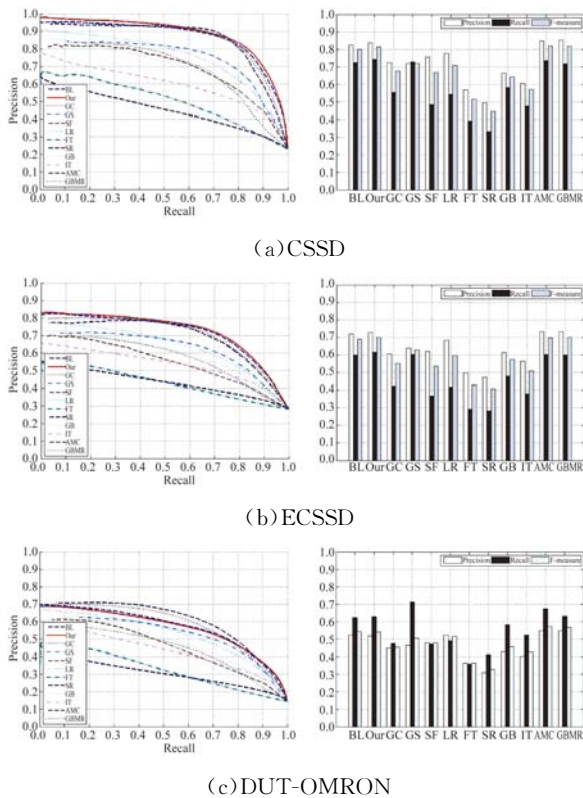


图 4 本文算法与现有 11 种算法在 CSSD、ECSSD、DUT-OMRON 3 个数据集上的对比结果

Fig. 4 Comparison results of proposed algorithm with the other 11 methods on the CSSD, ECSSD, DUT-OMRON data sets

图 4 为本文算法与 11 种算法分别在 CSSD, ECSSD, DUT-OMRON 3 个数据集上的比较结果。在 CSSD 数据集上,本文算法的准确率、召回

率以及 F-measure 分别达到了 0.837 5, 0.743 4 和 0.813 7; 在 ECSSD 数据集上分别为 0.728 3, 0.613 7, 0.698 2; 在 DUT-OMRON 数据集上分别为 0.520 2, 0.625 6, 0.541 2。从图 4 中的 ROC 曲线下面积大小以及柱状图中柱体高度不难看出本文算法相较于其它算法在准确度、召回率和 F-measure 上均得到了一定程度上的提高。

4 结 论

本文提出一种结合局部特征以及全局特征的显著性检测模型,其充分利用了前景信息并且抑制了背景信息,另外还充分考虑了局部信息和全局信息。本文首先采用多特征的方式得到基于背景的显著图,在此基础上,通过计算前景种子和超像素之间的距离生成了基于前景的显著图,并且使用高斯滤波器对前景和背景相结合的显著图结果进行优化。然后训练分类器对之前生成的显著图进行进一步优化,最后将高斯滤波器优化后的显著图和分类器输出得到的显著图进行融合,得到最终显著图。实验结果表明本文方法相较于其它的显著性检测模型在准确率以及召回率上均有一定程度的提高,算法在 CSSD 数据集上的平均准确率为 0.837 5, 平均召回率为 0.743 4, 相较于主流算法有一定优势,说明了结合局部特征以及全局特征进行显著性检测的可行性。

参考文献:

- [1] ITTI L, KOCH C, NIEBUR E. A model of saliency-based visual attention for rapid scene analysis [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 1998, 20(11):1254-1259.
- [2] ZHANG C J, XUE Z, ZHU X B, *et al.*. Boosted random contextual semantic space based representation for visual recognition [J]. *Information Sciences*, 2016, 369:160-170.
- [3] JIANG H, WANG J, YUAN Z, *et al.*. Salient object detection: a discriminative regional feature integration approach [C]. *IEEE Conference on Computer Vision and Pattern Recognition*. *IEEE Computer Society*, 2013:2083-2090.
- [4] CHENG M M, MITRA N J, HUANG X, *et al.*. Global contrast based salient region detection [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2011, 37(3):409-416.
- [5] SUN X S, YAO H X, JI R R, *et al.*. Toward statistical modeling of saccadic eye-movement and visual saliency [J]. *IEEE Transactions on Image Processing*, 2014, 23(11):4649-4662.
- [6] PERAZZI F, KRAHENBUHL P, PRITCH Y, *et al.*. Saliency filters: contrast based filtering for salient region detection [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2012: 733-740.
- [7] ZHANG L, TONG M H, MARKS T K, *et al.*. SUN: a bayesian framework for saliency using natural statistics [J]. *Journal of Vision*, 2008, 8(7): 1-20.
- [8] BORJI A. Boosting bottom-up and top-down visual features for saliency estimation [C]. *IEEE Conference on Computer Vision and Pattern Recognition*. *IEEE Computer Society*, 2012:438-445.
- [9] TONG N, LU H, XIANG R, *et al.*. Salient object detection via bootstrap learning [C]. *The IEEE Conference on Computer Vision and Pattern Recognition*, 2015:1884-1892.
- [10] ZHANG X Y, WANG S, YUN X. Bidirectional active learning: a two-way exploration into unlabeled and labeled data set [J]. *IEEE Transactions on Neural Networks & Learning Systems*, 2015, 26(12):3034-3044.
- [11] ZHANG X Y, WANG S, ZHU X, *et al.*. Update vs. upgrade: modeling with indeterminate multi-class active learning [J]. *Neurocomputing*, 2015, 162:163-170.
- [12] 贾松敏,徐涛,董政胤,等.采用脉冲耦合神经网络的改进显著性区域提取方法[J].*光学精密工程*, 2015, 23(3): 819-826.
- JIA S M, XU T, DONG ZH Y, *et al.*. Improved saliency region extraction algorithm with PCNN [J]. *Opt. Precision Eng.*, 2015, 23(3): 819-826. (in Chinese)
- [13] 张颖颖,张帅,张萍,等.融合对比度和分布性的图像显著性区域检测[J].*光学精密工程*, 2014, 22(4): 1012-1019.
- ZHANG Y Y, ZHANG SH, ZHANG P, *et al.*. Detection of salient maps by fusion of contrast and distribution [J]. *Opt. Precision Eng.*, 2014, 22(4): 1012-1019. (in Chinese)
- [14] ACHANTA R, SHAJI A, SMITH K, *et al.*. Slic superpixels [R]. *School of Computer and Communications Sciences, EPFL Technical Report 149300*, 2010.
- [15] BORJI A, SIHITE D N, ITTI L. Salient object detection: a benchmark [J]. *IEEE Transactions on Image Processing*, 2015, 24(12):414-429.
- [16] OHTSU N. A threshold selection method from gray-level histograms [J]. *IEEE Transactions on Systems Man & Cybernetics*, 1979, 9(1):62-66.
- [17] 高恒振,万建伟,粘永健,等.组合核函数支持向量机高光谱图像融合分类[J].*光学精密工程*, 2011, 19(4):878-883.
- GAO H ZH, WAN J W, NIAN Y J, *et al.*. Fusion classification of hyperspectral image by composite kernels support vector machine [J]. *Opt. Precision Eng.*, 2011, 19(4): 878-883. (in Chinese)
- [18] SIAGIAN C, ITTI L. Rapid biologically-inspired scene classification using features shared with visual attention [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2007, 29(2):300-12.
- [19] HAREL J, KOCH C, PERONA P. Graph-based visual saliency [C]. *Neural Information Processing Systems*, 2006:545-552.
- [20] HOU X, ZHANG L. Saliency detection: a spectral residual approach [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2007: 1-8.
- [21] MANNAN S K, KENNARD C, HUSAIN M. The role of visual saliency in directing eye movements in visual object agnosia[J]. *Current Biolo-*

- gy, 2009, 19(6):247-248.
- [22] REYNOLDS J H, DESIMONE R. Interacting roles of attention and visual salience in V4 [J]. *Neuron*, 2003, 37(5):853-863.
- [23] RUTISHAUSER U, WALTHER D, KOCH C, *et al.*. Is bottom-up attention useful for object recognition [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2004:37-44.
- [24] JIANG B, ZHANG L, LU H, *et al.*. Saliency detection via absorbing markov chain [C]. *IEEE International Conference on Computer Vision*, 2013: 1665-1672.
- [25] YANG C, ZHANG L, LU H, *et al.*. Saliency detection via graph-based manifold ranking [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2013:3166-3173.
- [26] ACHANTA R, HEMAMI S, ESTRADA F, *et al.*. Frequency-tuned salient region detection [C]. *IEEE International Conference on Computer Vision and Pattern Recognition*, 2009:1597-1604.

作者简介:



蔡 强(1969—),男,重庆永川人,博士,教授,1991年、2003年于北京航空航天大学分别获得学士、博士学位,1994年于北京轻工业学院获得硕士学位,现为北京工商大学计算机与信息工程学院院长,主要从事计算机图形学、计算几何、科学可视化、智能信息处理等方面的研究。E-mail: caiq@th. btbu. edu. cn



郝佳云(1993—),女,山西长治人,学士,2015年于北京工商大学获得学士学位,主要从事显著性检测、知识图谱等方面的研究。E-mail: haojiayun_happy@163. com