

DOI: 10.11992/tis.201801009

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.TP.20180412.1148.008.html>

基于 Object Proposals 并集的显著性检测模型

赵闰霞¹, 蹇木伟^{1,2}, 齐强¹, 王静¹, 王瑞红¹, 董军宇¹

(1. 中国海洋大学 信息科学与工程学院, 山东 青岛 266000; 2. 山东财经大学 计算机科学与技术学院, 山东 济南 250014)

摘 要: 针对当前常见的显著性检测模型得到的结果会包含大量的背景区域的缺点, 本文提出了基于 Object Proposals 并集的显著性检测模型。该模型首先对于输入图片生成一系列 Object Proposals, 并通过其并集计算得到背景图; 然后结合纹理特征和全局对比度得到初始显著图; 最后, 用得到的背景图对初始显著图进行背景抑制得到最终显著图。实验结果表明, 在通用 MSRA1000 数据集上, 本文提出的显著性模型与其他 5 种方法相比取得了很好的效果。

关键词: 显著性检测; object proposal; 超像素; 纹理; 背景图; 全局对比度; 边界连通性; 自底向上

中图分类号: TP391 **文献标志码:** A **文章编号:** 1673-4785(2018)06-0946-06

中文引用格式: 赵闰霞, 蹇木伟, 齐强, 等. 基于 Object Proposals 并集的显著性检测模型[J]. 智能系统学报, 2018, 13(6): 946-951.

英文引用格式: ZHAO Runxia, JIAN Muwei, QI Qiang, et al. Saliency detection model based on the union of Object Proposals[J]. CAAI transactions on intelligent systems, 2018, 13(6): 946-951.

Saliency detection model based on the union of Object Proposals

ZHAO Runxia¹, JIAN Muwei^{1,2}, QI Qiang¹, WANG Jing¹, WANG Ruihong¹, DONG Junyu¹

(1. College of Information Science and Engineering, Ocean University of China, Qingdao 266000, China; 2. School of Computer Science & Technology, Shandong University of Finance and Economics, Ji'nan 250014, China)

Abstract: In saliency detection, current existing models usually produce results containing many background regions. To improve the performance, a novel saliency detection model is proposed based on the union of object proposals. The model first generates a series of object proposals from the input pictures, and then gets the background map by computing the union, and then obtains the initial saliency map by combining the texture and global contrast. Finally, the final saliency map is derived by restraining the initial saliency map with the obtained background map. Experimental results on the general MSRA1000 dataset demonstrate that the proposed saliency model performs well compared to the other five existing methods.

Keywords: saliency detection; Object Proposal; superpixels; texture; background map; global contrast; boundary connectivity; bottom-up

近年来, 显著性检测在图像处理和计算机视觉领域中愈发重要, 并且得到了广泛研究和应用。显著性检测的本质是模拟人类的视觉注意机制提取图像中最引人注意的区域, 目前已被广泛应用于图像分割^[1]、目标识别^[2]、像压缩^[3]以及图像检索^[4]等领域。

显著性检测方法通常可分为两种模型: 自底向上 (bottom-up)^[5-10]和自顶向下 (top-down)^[11-12]。自底向上一般是快速的、数据驱动的显著性提取模型。Cheng 等^[5]提出了基于区域对比度的显著性检测算法, 用于测量全局对比度差异; Harel 等^[6]提出了一种基于图形的显著性模型检测图像中的显著物体; Erdem^[7]使用从局部图像块提取的图像特征的协方差矩阵来计算最终的显著性图; Murray 等^[8]设计了一种特殊的颜色外观显著性检测模

收稿日期: 2018-01-08. 网络出版日期: 2018-04-12.

基金项目: 国家自然科学基金项目 (61601427, 61602229).

通信作者: 蹇木伟. E-mail: 20173016@sdufe.edu.cn.

型; Shi^[9]通过图形模型提出了一种分层图像显著性目标检测方法, 以产生最终显著性等级; Yang等^[10]通过基于图形的流形排序实现图像元素(像素或区域)与前景或背景线索的相似性的排序。

与自底向上相反, 自顶向下通常是慢速的、任务依赖的、目标驱动的显著性提取。自顶向下的方法通常依赖于学习框架, 采用高层线索和先验知识进行显著性检测。文献[11]中提出了通过学习条件随机场进行特征组合, 并引入动态显著特征以实现显著目标的检测。最近, 文献[12]提出了一种结合字典学习和条件随机场的自上而下的视觉显著性模型。

一般情况下, 使用基于学习框架的显著性检测模型的检测效果会比较好, 但是其检测效果会很大程度上依赖于模型的训练样本, 训练的过程也会使计算复杂度增加。非学习类的显著性检测模型通常是以底层特征为线索进行显著性检测, 如对比度先验、中心先验、背景先验等。这些方法一般只对简单的场景有效, 当背景与检测对象相似时, 这类方法将无法准确地检测出目标物体。

本文提出基于 Object Proposals 并集的显著性检测模型。首先我们利用 Object Proposals 的有效信息预测背景区域的位置并得到背景图, 然后基于纹理特征以及全局对比度进行显著性检测, 最后将三者进行融合得到最终的显著图。

1 Object Proposals

Object Proposals 是一系列包含目标对象的候选区域, 它通常以预处理的形式广泛应用于各种视觉任务。Alexe 等^[13]提出了对象度量(objectness measure)的概念, 量化一个图像窗口覆盖对象的可能性; Cheng 等^[14]提出二值化规范梯度(BING)特征, 有助于用对象评分(objectness score)搜索对象。本文使用 Edge Boxes^[15]算法生成 Proposals, 得到的 Proposals 会包含丰富的对象级特征, 我们利用这些特征判断显著目标的位置并预测背景区域。

对于图像 I , 首先使用 SLIC 算法^[16]得到超像素图, $I = \{sp_1, sp_2, \dots, sp_n\}$ 。然后使用 Edge Boxes^[15]算法生成图像 I 的 Proposals, 这里面部分候选框能准确地标注到目标对象, 但也存在大量的候选框包含太多的背景区域或者不包含目标对象。以图 1 为例, (a) 中的 3 个候选框都标注到了目标对象, 且包含的背景区域比较少; (b) 中的候选框则没有包含目标对象; (c) 中的候选框虽然包含了目标对象, 但也包含了很多的背景区域。为了得到更准确的背景区域, 根据目标分数对得到的候选

框进行选择。一般分数靠前的候选框能够更准确地包含到目标物体, 所以选择前 N 个候选框作为目标集合 A , 在实验中, N 的取值为 100。



图 1 目标候选框

Fig. 1 Object Proposals

在目标集合 A 中, 针对图 1(b)、(c) 的现象, 对包含大量背景区域的候选框进行删除。受文献[17-19]的启发, 本文中使用基于超像素的边界连通性对候选框进行筛选。用 $\text{BndCon}(\cdot)$ 表示边界连通性, 其本质是描述一个区域与图像边界的连通程度^[20]。 $\text{BndCon}(\cdot)$ 的定义为

$$\text{BndCon}(sp_i) = \frac{\text{Len}_{\text{bnd}}(sp_i)}{\sqrt{\text{Area}(sp_i)}} \quad (1)$$

式中: $\text{Area}(sp_i)$ 的含义是区域中存在的超像素数目, 而 $\text{Len}_{\text{bnd}}(sp_i)$ 则是与边界相邻的超像素的数目。 $\text{BndCon}(sp_i)$ 的值越大, 则表明边界上的超像素数目越多, 也就是该区域是背景区域的可能性越大。若一个候选框计算得到的 $\text{BndCon}(sp_i)$ 很大, 我们则认为这个候选框包含了很多背景区域, 并从 A 中将其删除。最终得到的目标集合能尽量去除背景区域的影响, 更新后的目标集合记作 A^- 。

2 基于 Object Proposals 并集的显著性检测

2.1 构建基于 OP 并集的背景图

定义得到的目标集合 $A^- = \{b_1, b_2, \dots, b_n\}$, 其中 b_i 表示第 i 个候选框, n 表示候选框的个数。为了得到最终的背景图, 以候选框中包含的超像素为单位进行计算^[17], 超像素 sp_i 的目标分数 $O(p_i)$ 定义为

$$O(sp_i) = \sum_{j=0}^n o(b_j) \cdot \delta, \quad sp_i \in b_j \quad (2)$$

$o(b_j)$ 表示候选框 b_j 的目标分数, 若超像素 sp_i 属于候选框 b_j , $\delta(\cdot)$ 则取值为 1, 否则为 0。

通过计算 A^- 中所有超像素的 (p_i) 值, 可以得到最终的 OP 并集, 然后用 φ 划分目标对象和背景区域, 如下:

$$\varphi(sp_i) = \begin{cases} 0, & O(sp_i) \leq \varphi \\ O(sp_i), & \text{其他} \end{cases} \quad (3)$$

式中 $\varphi = \frac{\mu}{N} \sum_{i=1}^N O(sp_i)$ 。当目标分数小于等于阈

值 φ 时, 将其定义为背景区域, 大于 φ 的区域不做处理, 如图2所示。图2中第1列为原图, 第4列为得到的背景图, 背景图中黑色区域表示背景, 与原图相比可知, 我们得到的背景图能够检测到大部分背景区域。

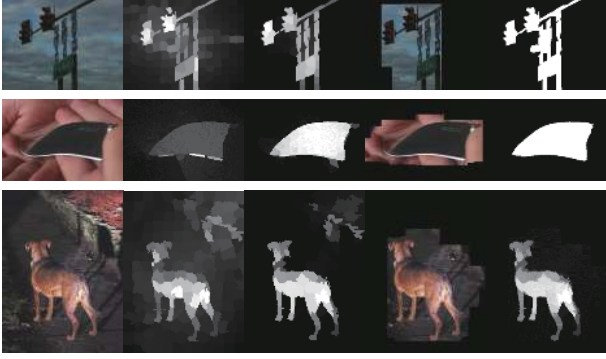


图2 本文的显著性检测模型

Fig. 2 Our proposed saliency model

2.2 纹理显著性

纹理^[21-22]通常反应的是图像的某种局部化性质, 一般表现为图像像素点某邻域内灰度级或者颜色的变化。在图像中彩色像素并不是一个随机的组合, 而是有一定含义的排列, 不同排列会形成不同的纹理, 这些纹理特征则可用于对图像中的空间信息进行一定程度的定量描述。

在本文中, 使用选择性纹理^[23]获得纹理显著性。首先, 将纹理表达式通过 k-means^[24]算法进行聚类, 并将聚类中心作为代表纹理^[23], 最后每个纹理特征都可以用其最近的代表纹理进行表示。使用的滤波器组是高斯导数以及其希尔伯特变换沿着 y 轴旋转得到的, 具体为

$$f_g(x, y) = \frac{d^2}{dy^2} \left(\frac{1}{\mu} \exp\left(\frac{y}{\sigma}\right)^2 \exp\left(\frac{x}{\sigma l}\right)^2 \right) \quad (4)$$

$$f_h(x, y) = \text{Hilbert}(f_g(x, y)) \quad (5)$$

式中: μ 是标准化常数, σ 和 l 分别是尺度参数和方向参数。对于图像 I 的第 i 个像素 I_i 的纹理显著性定义如下:

$$S_1(I_i, R_i) = \sum_{I_j \neq I_i, I_j \in R_i} D(I_i, I_j) \quad (6)$$

式中: R_i 表示 I_i 所在的区域, $D(I_i, I_j)$ 是 I_i 和 I_j 的纹理特征的距离, 其定义如下:

$$D(I_i, I_j) = \|t_i - t_j\|_2 \quad (7)$$

式中 t_i 和 t_j 分别是像素 I_i 和 I_j 的纹理特征。图2中的第3列是选择性纹理检测^[23]的结果, 可以看出纹理显著性检测能够区分出不同纹理的对象, 但结果依然会包含背景中的部分纹理内容, 比如第3行的纹理检测结果。

2.3 全局对比度显著性

空间加权对比度在文献[5, 18]中已被证明是

有效的显著性检测方法。在文献[5]中是将图片分割成若干区域, 计算区域及颜色对比度, 再用每个区域与其他区域对比度加权和定义此区域的显著性。本文中直接在第1节中获得的超像素基础上进行计算。对于任意超像素 sp_i , 在 CIE LAB 色彩空间中计算超像素 sp_i 的颜色平均值 c_i 以及平均像素坐标 d_i , 并将 d_i 归一化到 $[0, 1]$ 。对于超像素 sp_i , 其空间加权区域对比度为

$$S_2(sp_i) = \sum_{j \neq i} \|c_i - c_j\| \cdot \exp\left(-\frac{\|d_i - d_j\|}{2\sigma_p^2}\right) \quad (8)$$

式中: σ_p 为空间加权参数, σ_p 越小, 空间权值的影响就会越大, 这样就可以使较近的区域对当前区域的显著性值产生较大的影响。本文中, σ_p^2 取值 0.2。图2中的第2列是全局对比度的检测结果, 全局对比度图能进一步凸显前景区域, 加深纹理检测图中的显著目标的显著程度。

2.4 显著性融合

将得到的背景图(2.1节)、纹理显著图(2.2节)以及全局对比显著图(2.3节)进行融合。首先将纹理显著图和全局对比度显著图进行融合, 旨在结合二者的优势, 相互补充, 优化显著图。

$$\text{Sal} = \text{norm}(S_1) + \text{norm}(S_2) \quad (9)$$

式中 $\text{norm}(\cdot)$ 是对得到的显著图进行归一化。在这里选择“+”是因为纹理检测和全局对比度检测可以相互补充, 此外, 通过加运算可以使前景区域分布更均匀^[19]。

在得到融合显著图 Sal 后, 使用 2.1 节中得到的背景图对显著图 Sal 的背景区域进行抑制, 从而提高显著图的准确率。

$$S = \text{Sal} \times \Phi(O(p_i) \leq \varphi) \quad (10)$$

式中 $\Phi(O(p_i) \leq \varphi)$ 是指背景图中的背景区域。图2中的第5列是融合后的最终显著图 S , 得到的最终显著图在目标物体上显著性分布更均匀, 与背景的分割边界也更加清晰。

3 实验与比较

本文在 MSRA1000 公开数据集对本文的显著性模型进行了测试, 并与一些经典的算法进行比较: SR^[25]、HC^[5]、FT^[26]、LC^[27]、GS^[28]。实验过程中, 使用的代码均采用作者公开的代码。

图3是将本文的显著性检测模型与其他5种算法进行比较的效果图。从图中可以看出在 MSRA1000 数据集上本文提出的显著性检测模型比 SR、HC、FT、LC 相比效果有很大提升; 与 GS 相比, 本文的显著性检测图前景分布更均匀, 且背景区域得到了明显抑制。

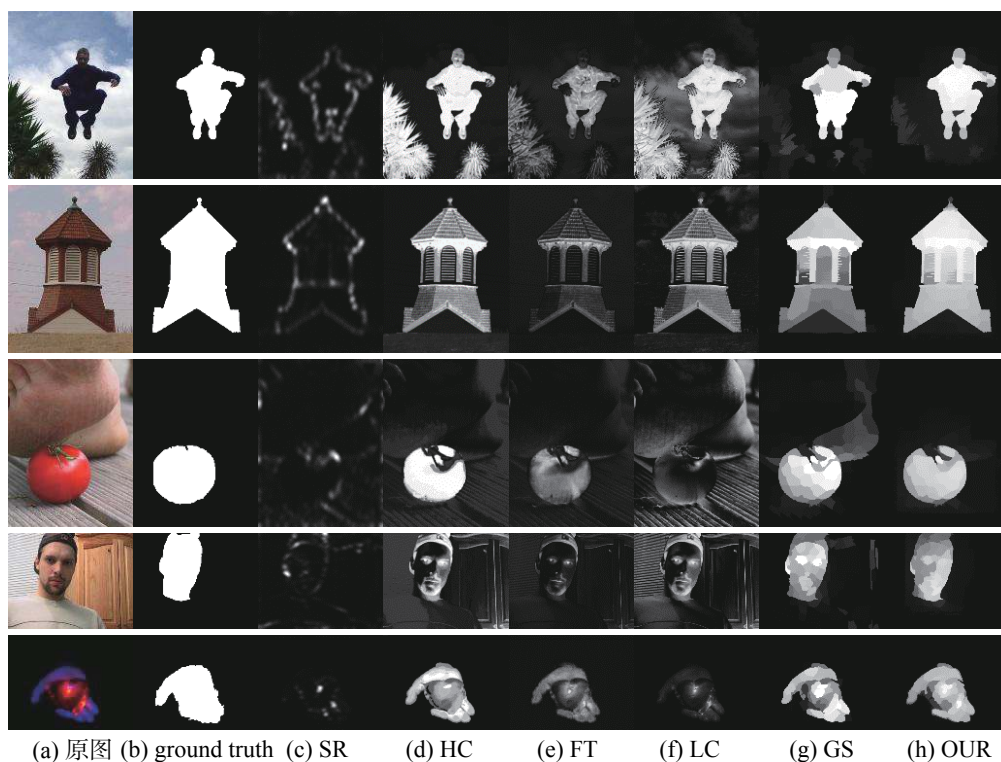


图 3 显著图的视觉效果对比图

Fig. 3 Visual comparison of saliency maps

为了更全面地测试方法的性能,使用 PR 曲线、 F -measure 进一步与其他方法进行比较。

对于得到的显著图,设定阈值 $T_f \in [0, 255]$ 来得到二值分割图,并与 ground truth 进行比较得到准确率和召回率。除此之外,我们还记算了 F -measure 衡量总体的检测效果,

$$F_\beta = \frac{(1 + \beta^2) \text{Precision} \times \text{Recall}}{\beta^2 \text{Precision} + \text{Recall}} \quad (11)$$

β^2 取值为 0.3。由图 4 ~ 5 可知,本文方法与 SR、HC、FT、LC 相比,准确率和召回率、 F -measure 都有很大提升。与 GS 相比,由图 5 可以看出,在召回率相同时,大部分情况下我们的方法准确率更高,也存在小部分结果其准确率微低于 GS,但从图 4 也能够看出在平均情况下,无论是准确度、召回率还是 F -measure,本文方法均略高于 GS。

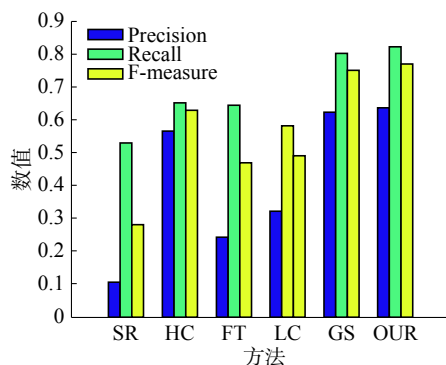


图 4 统计指标的对比结果

Fig. 4 Quantitative comparison results

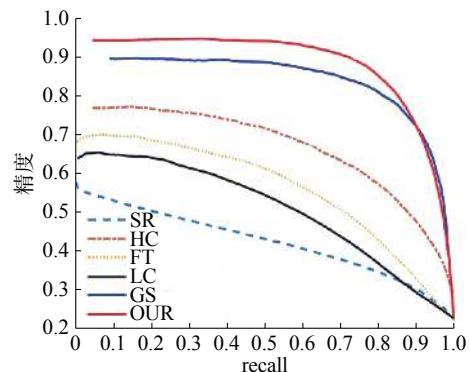


图 5 正确率-召回率曲线

Fig. 5 Precision-recall curves

4 结束语

本文提出的基于 Object Proposals 并集的显著性检测模型,首先利用 Object Proposals 计算背景图,然后将纹理和全局对比度结合进行显著性检测,最后利用背景图抑制背景区域得到最终显著图。实验结果表明,本文提出的显著性检测模型取得了较好的效果。

参考文献:

- [1] 王超, 王浩, 王伟, 等. 基于优化 ROI 的医学图像分割与压缩方法研究[J]. 重庆邮电大学学报: 自然科学版, 2015, 27(2): 279-284.

WANG Chao, WANG Hao, WANG Wei, et al. Study of optimized ROI based medical image segmentation and

- compression method[J]. Journal of Chongqing university of posts and telecommunications: natural science edition, 2015, 27(2): 279–284.
- [2] 杨林娜, 安玮, 林再平, 等. 基于空间距离改进的视觉显著性弱小目标检测[J]. 光学学报, 2015, 35(7): 0715004.
YANG Linna, AN Wei, LIN Zaiping, et al. Small target detection based on visual saliency improved by spatial distance[J]. Acta optica sinica, 2015, 35(7): 0715004.
- [3] ITTI L. Automatic foveation for video compression using a neurobiological model of visual attention[J]. IEEE transactions on image processing, 2004, 13(10): 1304–1318.
- [4] CHEN Tao, CHENG Mingming, TAN Ping, et al. Sketch2photo: internet image montage[J]. ACM transactions on graphics, 2009, 28(5): 1–10.
- [5] CHENG Mingming, MITRA N J, HUANG Xiaolei, et al. Global contrast based salient region detection[J]. IEEE transactions on pattern analysis and machine intelligence, 2015, 37(3): 569–582.
- [6] HAREL J, KOCH C, PERONA P. Graph-based visual saliency[C]//Proceedings of the 19th International Conference on Neural Information Processing Systems. Cambridge, MA, USA, 2007: 545–552.
- [7] ERDEM E, ERDEM A. Visual saliency estimation by nonlinearly integrating features using region covariances[J]. Journal of vision, 2013, 13(4): 11.
- [8] MURRAY N, VANRELL M, OTAZU X, et al. Saliency estimation using a non-parametric low-level vision model [C]//Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition. Washington, DC, USA, 2011: 433–440.
- [9] SHI Jianping, YAN Qiong, XU Li, et al. Hierarchical image saliency detection on extended CSSD[J]. IEEE transactions on pattern analysis and machine intelligence, 2016, 38(4): 717–729.
- [10] YANG Chuan, ZHANG Lihe, LU Huchuan, et al. Saliency detection via graph-based manifold ranking[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Washington, DC, USA, 2013: 3166–3173.
- [11] LIU Tie, YUAN Zejian, SUN Jian, et al. Learning to detect a salient object[J]. IEEE transactions on pattern analysis and machine intelligence, 2011, 33(2): 353–367.
- [12] YANG Jimei, YANG M H. Top-down visual saliency via joint CRF and dictionary learning[C]//Proceedings of Conference on Computer Vision and Pattern Recognition (CVPR). Washington, DC, USA, 2012: 2296–2303.
- [13] ALEXE B, DESELAERS T, FERRARI V. Measuring the objectness of image windows[J]. IEEE transactions on pattern analysis and machine intelligence, 2012, 34(11): 2189–2202.
- [14] CHENG Mingming, ZHANG Ziming, LIN Wenyan, et al. BING: Binarized normed gradients for objectness estimation at 300fps[C]//Proceedings of IEEE conference on computer vision and pattern recognition. Columbus, OH, USA, 2014: 3286–3293.
- [15] ZITNICK C L, DOLLÁR P. Edge boxes: Locating object proposals from edges[C]//Proceedings of European Conference on Computer Vision. Cham: Springer, 2014: 391–405.
- [16] ACHANTA R, SHAJI A, SMITH K, et al. SLIC superpixels compared to state-of-the-art superpixel methods[J]. IEEE transactions on pattern analysis and machine intelligence, 2012, 34(11): 2274–2282.
- [17] HUANG Fang, QI Jinqing, LU Huchuan, et al. Salient object detection via multiple instance learning[J]. IEEE transactions on image processing, 2017, 26(4): 1911–1922.
- [18] PERAZZI F, KRÄHENBÜHL P, PRITCH Y, et al. Saliency filters: Contrast based filtering for salient region detection[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Providence, RI, USA, 2012: 733–740.
- [19] XI Tao, ZHAO Wei, WANG Han, et al. Salient object detection with spatiotemporal background priors for video[J]. IEEE transactions on image processing, 2017, 26(7): 3425–3436.
- [20] ZHU Wangjiang, LIANG Shuang, WEI Yichen, et al. Saliency optimization from robust background detection [C]//Proceedings of IEEE conference on computer vision and pattern recognition. Columbus, OH, USA, 2014: 2814–2821.
- [21] 刘丽, 匡纲要. 图像纹理特征提取方法综述[J]. 中国图象图形学报, 2009, 14(4): 622–635.
LIU li, KUANG Gangyao. Overview of image textural feature extraction methods[J]. Journal of image and graphics, 2009, 14(4): 622–635.
- [22] 王佐成, 薛丽霞. 一种新的纹理基元发现及表达方法[J]. 重庆邮电大学学报: 自然科学版, 2011, 23(1): 115–120.
WANG Zuocheng, XUE Lixia. A new representation method of image texon[J]. Journal of Chongqing university of posts and telecommunications: natural science edition, 2011, 23(1): 115–120.
- [23] WANG Qi, YUAN Yuan, YAN Pingkun. Visual saliency by selective contrast[J]. IEEE transactions on circuits and systems for video technology, 2013, 23(7): 1150–1155.
- [24] HARTIGAN J A, WONG M A. Algorithm as 136: a k-means clustering algorithm[J]. Journal of the royal statistical society. series c (applied statistics), 1979, 28(1): 100–108.
- [25] HOU Xiaodi, ZHANG Liqing. Saliency detection: a spectral residual approach[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis, MN, USA, 2007: 1–8.
- [26] ACHANTA R, HEMAMI S, Estrada F, et al. Frequency-

tuned salient region detection[C]//Proceedings of Conference on Computer vision and pattern recognition. Miami, FL, USA, 2009: 1597–1604.

- [27] ZHAI Yun, SHAH M. Visual attention detection in video sequences using spatiotemporal cues[C]//Proceedings of the 14th ACM international conference on Multimedia. New York, NY, USA, 2006: 815–824.
- [28] WEI Y, WEN F, ZHU W, et al. Geodesic saliency using background priors[J]. Computer vision-ECCV 2012, 2012: 29–42.

作者简介:



赵闰霞, 女, 1993 年生, 硕士研究生, 主要研究方向为图像处理。



蹇木伟, 男, 1982 年生, 教授, 博士生导师, CCF 计算机视觉专委会委员, CCF 多媒体专委会委员, CCF 机器学习与模式识别通讯委员, 山东数媒专委会委员等。主要研究方向为图像处理、模式识别、多媒体计算、机器学习、认知科学。主持国家自然科学基金等研究课题 10 余项。以第一发明人或第一申请人被授予 3 项国家专利, 其中 1 项国家发明专利和 2 项国家实用新型专利。发表学术论文 50 余篇。被 SCI 检索的国际期刊论文 14 篇、被 EI 检索论文 40 余篇。



齐强, 男, 1990 年生, 硕士研究生, 主要研究方向为图像处理、模式识别、水下视觉。