

文章编号 1004-924X(2013)02-0531-08

视觉显著目标的自适应分割

赵宏伟, 陈 霄, 刘萍萍*, 耿庆田

(吉林大学 计算机科学与技术学院, 吉林 长春 130022)

摘要: 基于视觉注意模型和最大熵分割算法, 提出了一种自适应显著目标分割方法来分离目标和复杂背景, 以便快速准确地从场景图像中检测出显著目标。首先, 通过颜色、强度、方向和局部能量 4 个特征通道获取图像的显著图; 通过引入局部能量通道来更好地描述了显著目标的轮廓。然后, 根据显著图中像素灰度的强弱构建不同的目标检测蒙板, 将每个蒙板作用于原图像作为预分割的结果, 再计算每个预分割图像的熵。最后, 利用最大熵准则估计图像目标熵, 根据预分割图像的熵和目标熵判断选取最优显著目标分割图像。实验结果表明: 本文算法检测的显著目标更为完整, 分割性能 F-measure 达到 0.56, 查全率和查准率分别为 0.69 和 0.41, 相对于传统方法更为有效准确, 实现了在复杂背景下对显著目标的有效准确检测。

关键词: 视觉显著目标; 局部能量; 目标检测; 最大熵准则; 自适应分割

中图分类号: TP391 **文献标识码:** A **doi:** 10.3788/OPE.20132102.0531

Adaptive segmentation for visual salient object

ZHAO Hong-wei, CHEN Xiao, LIU Ping-ping*, GENG Qing-tian

(College of Computer Science and Technology, Jilin University, Changchun 130012, China)

* Corresponding author, E-mail: liupp@jlu.edu.cn

Abstract: On the basis of a visual attention model and a maximum entropy segmentation method, an adaptive segmentation method was proposed to segment the object from a complex background in the scene image and to detect a salient object effectively and accurately. First, the feature of original image was extracted via four channels on color, intensity, orientation and local energy. The profile of object feature was described more accurately by combining the channel of local energy with a simple biologically-inspired model. Then, object detection masks were constructed to remove background gradually according to the gray intensity of the pixels in the saliency map. By taking blend masks with the original image as a pre-segmentation result, the entropy of pre-segmentation images was computed. Finally, the entropy of salient object was estimated via maximization information entropy principle and the optimized image extraction for the salient object was obtained by estimating the relationship of entropy between salient object and masks in the saliency map. Experimental results indicate that the salient object detected by proposed method is more integrity, the F-measure of segmentation performance is 0.56, and the precision ratio and the recall ratio of detection are 0.69 and 0.41, respectively. The proposed method is more reasonable and effective than the traditional method, and it can satisfy

收稿日期: 2013-01-11; 修订日期: 2013-01-19.

基金项目: 国家自然科学基金资助项目 (No. 61101155); 吉林省科技发展计划资助项目 (No. 20101504)

the requirements of detecting the salient objects from complex backgrounds.

Key words: visual salient object; local energy; object detection; maximization entropy; adaptive segmentation

1 引言

人的视觉系统可以在未知环境中不受复杂背景影响即快速准确地搜索到场景中的显著目标,这种视觉特性是一个自底向上的过程,通常称作视觉选择注意。研究者们一直追求实现一种有效的目标检测跟踪方法^[1-2]。为了将显著目标从复杂背景中分离,研究者们使用显著性检测方法定位目标,因此,准确地分割目标和背景已成为该领域的研究热点。根据目标显著性强弱,融合图像分割思想,自适应分析目标并逐渐消除背景,整个过程为自底向上,无需训练,可以有效地检测出显著目标。

目前,基于图像显著性的目标检测方法大体分为二类:第一类为基于视觉生理特性的显著性分析方法^[3]。该方法的代表性模型为 Itti 视觉选择注意模型,由 Itti 等人在 1998 年提出,其基本思想是通过模拟生物视觉特性,建立颜色、方向以及强度通道并强化图像中的相应特征,然后融合各个通道特征图构成显著图;该方法建立在局部特征的差异上,忽略了目标信息在图像中的完整性。Z. Li^[4]将 Itti 模型应用于卫星图像中的目标检测,并在原通道基础上将加入 L、X、T 连接型通道,熵通道以及局部方差等通道弥补在显著性分析时缺乏对目标完整性的考虑,但该方法是针对卫星图像中地形地貌等特征建立的,应用较受局限。第二类为基于数学统计模型的显著性分析方法。通常图像的局部显著性与全局特征分布呈反比,Hou 等^[5]利用这一思想,假设图像整体由目标和背景构成,其中背景信息冗余相似,而目标信息孤立显著,由场景图像的傅里叶统计特性得到图像的残余频谱作为显著特征,但该方法适合对较小目标进行检测。Chen 等人^[6]提出一种基于区域显著对比的目标检测方法,同时考虑全局的对比度以及空间的相关性,该算法速度和准确率都很优秀,但算法在目标检测时的分割方法采用经验固定阈值,配合膨胀腐蚀等算法优化,未达到普遍的自适应分割效果。另外,基于图论的

显著目标检测方法也备受关注,胡正平^[7]等利用马尔科夫随机游走的方法,在图像分块中找出全局孤立性特征,再通过 k-regular 提取局部显著特征,将两者结合获得显著图,该方法只考虑到区域特征而没有考虑目标的边缘信息,使得算法还不够完善。赵宏伟等人^[8]通过边缘检测和颜色聚类对视网膜图像中的目标进行编码,但很难应用于复杂背景场景图像。

综上,无论基于生理学还是基于数学方法建立的显著性目标检测方法都非常注重从图像中的局部差异和目标孤立完整性考虑显著性问题。本文从生理学角度出发,在传统三通道模型中加入局部能量通道,通过颜色、方向以及强度通道获取图像中局特征差异,通过局部能量获取轮廓信息以强化显著目标的全局孤立性,并且通过最大熵准则自动判别目标显著性,由此提出一种既符合视觉生理特性,且可以自适应分离背景的目标检测算法。

2 改进的视觉显著图

传统视觉选择注意模型通常会采用 Gabor 函数作为方向通道,强化场景图像中的方向信息,Gabor 函数符合视觉系统感受野细胞结构,对图像中心区域的信息有强化作用,而周边特征提取相对较弱,这种特性使得在对复杂场景下目标进行检测时,容易忽略偏离中心的目标特征信息,局部能量的引入可以很好地弥补这一弱点。并且,局部能量与局部相位正相关,可以较好地描述目标的轮廓信息^[11]。

2.1 局部能量特征的获取

神经生理学研究发现,视觉处理的前期阶段主要是物体的边缘特征或线特征,Andrea Perna^[9]等人用单纯的边缘,相位变换的边缘以及带有幅度噪声的轮廓线刺激初级视觉皮层,发现相位变换的边缘对初级视觉皮层刺激最强,说明相位一致性是在初级视觉皮层编码的。虽然相位一致性可以很好地描述边缘,但在实际实验过程中发现,由于边缘信息较为稀疏,因此经过多通道融

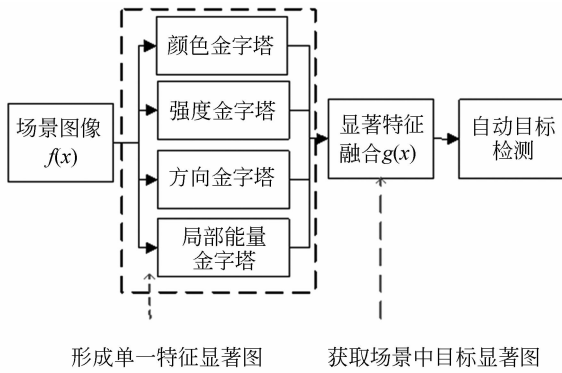


图 2 显著图的生成

Fig. 2 Getting saliency map

道,该函数符合感受野细胞结构,并且可以很好地响应图像中的方向信息。由 Gabor 函数图像可以看出,此函数对图像中心区域的特征信息有较强的响应,而分布在图像非中心区域的特征信息响应较弱,这一特性使得在对复杂场景下目标进行检测时,偏离中心的目标特征信息容易丢失,而局部能量可以很好地弥补这一弱点。如图 3 所示。

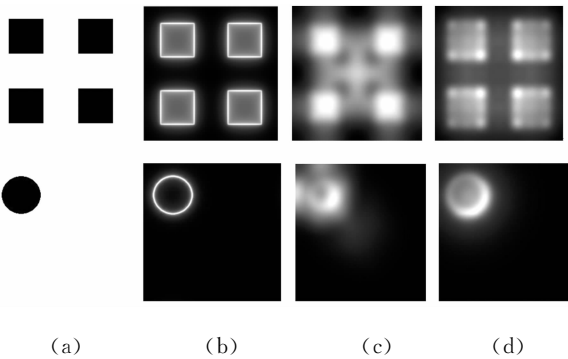


图 3 不同方法结果的对比((a)原图像;(b)局部能量特征图;(c)三通道显著图;(d)本文方法显著图)

Fig. 3 Comparison of results with different methods ((a) Original images; (b) Feature maps of local energy; (c) Saliency maps of Itti's method; (d) Saliency maps of proposed method)

局部能量与相位一致性成正相关性(由公式 2 可知),所以利用局部能量作为显著性通道既弥补了相位一致性特征信息稀疏离散的弱点,同时也保留了相位一致性对图像整体都具有较好的特征响应的特点。如图 3 所示,第一列为原图像,第二列为局部能量特征图,第三列为颜色,方向以及强度三通道显著图,第四列为添加局部能量的四

通道显著图。可以看出,局部能量特征对图像中心区域和非中心区域特征都有较好响应,并且,融合局部能量特征通道以后,弱化了 Gabor 函数中心强化的特性,使非中心区域的目标特征得到加强,保证了目标检测的完整性。

3 显著目标自动检测

3.1 基于最大熵准则的目标熵获取

信息论中,熵是用来衡量信息不确定性的量度,是平均信息量的表征。Shannon^[13]提出熵定义:

$$H(A) = - \sum_{i=1}^n p(i) \log p(i). \quad (11)$$

熵理论同样可以应用于图像中^[14],式(11)中 $p(i)$ 表示图像中每级灰度出现的概率, $H(A)$ 为图像的熵。

本文通过最大熵准则获得目标熵。最大熵的理论通常用于图像分割^[14],其原理认为图像中的目标区域和背景区域内的灰度变化具有相近性,灰度起伏平缓,当目标区域和背景区域的信息量都很大时,图像的熵达到最大,最大熵对应的灰度为图像分割的阈值。如果此时将阈值增大或者减少,分割就会出现目标区域和背景区域间的混淆。当目标区域和背景区域的熵相加达到最大时,背景和背景分割达到最优,此时目标区域的熵表示图像中目标的信息量。

已知大小为 $M \times N$ 的图像 $f(x, y)$, 图像的灰度集合 $G = \{0, 1, \dots, L-1\}$, n_i 表示图像中灰度值为 i 的像素出现的总次数, p_i 是灰度 i 在图像中出现的概率,表示为:

$$p_i = \frac{n_i}{M \times N}, \text{ 其中 } i \in G. \quad (12)$$

按照最大熵分割原理,将图像分成目标区域和背景区域两部分,两个区域的概率分别表示为:

$$p_O(t) = \sum_{i=0}^t p_i, \quad (13)$$

$$p_B(t) = \sum_{i=0}^t p_i, \quad (14)$$

并且有 $p_O(t) + p_B(t) = 1$, 因此图像中的目标区域和背景区域的熵分别表示为:

$$H_O(t) = \ln \sum_{i=0}^t \frac{p_i}{p_O(t)}, \quad (15)$$

$$H_B(t) = \ln \sum_{i=t+1}^{L-1} \frac{p_i}{p_B(t)} = \ln \sum_{i=t+1}^{L-1} \frac{p_i}{1 - p_0(t)}. \quad (16)$$

图像总体的熵为:

$$H(t) = H_0(t) + H_B(t). \quad (17)$$

当 $H(t)$ 达到最大时,此时 $H_0(t)$ 值为目标区域的熵。

3.2 自适应目标检测实现

在第二节中获得了四通道的显著图 $g(x)$,其中在 $g(x)$ 中灰度越强的区域表示原图像中 $f(x)$ 的特征显著性越强。将 $g(x)$ 图像按照灰度的强弱划分 99 个等级, x_i 中存在一点 x_i ,且 $g(x_i) \in g(x_i)$,将显著图像分为两部分, $g(M)$ 与 $g(m)$,其中 $g(M) > g(x_i)$ 为图像中的显著区域点集合, $g(m) < g(x_i)$ 为图像中的非显著区域点集合。令 $g(M) = 1, g(m) = 0$ 得到显著区域的二值蒙板 $g(t)$,图像中显著区域图像 $f(M)$ 表示为:

$$f(M) = (f * g)(x). \quad (18)$$

计算 $f(M)$ 的熵 $H(M)$,当 $H(M) \approx H_0(t)$ 时, $f(M)$ 最终为目标检测图像。

算法的实现框架如图 4 所示。

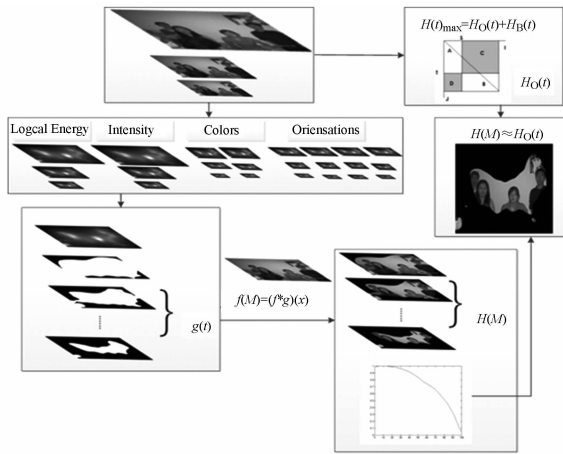


图 4 本文算法流程图

Fig. 4 Flow chart of proposed method

首先构建图像金字塔,同时根据图像最大熵准则求出目标熵,然后通过颜色、强度、方向以及局部能量通道获取显著图,接着根据显著性的强弱将显著图排序,得到目标检测二值蒙板并作用于原图像,计算每个加载蒙板后图像的熵,若与目标熵最为接近,此图就作为显著目标检测的结果。

4 仿真实验

本文选取 MSRA 数据集作为实验图片^[6-7],

并利用其配备的标注作为目标检测验证依据。在实验一中 ROC 曲线调节显著模型的参数。实验二通过目标分割结果图像主观评价算法,再通过查全率,查准率和 F-measure 方法检验算法客观评价目标分割的性能。

实验一:选取 5 幅图像样本,通过 ROC 曲线设定通道权重,主要设置方向通道和局部能量通道的权重,其他通道权重均为 1,在局部能量通道内,将维度设置为 3,最小波长设置为 3,倍频设置为 2,这样可以较好地描述显著目标的轮廓。如图 5 所示,当将局部能量通道的权值 L 设置为 1.3,方向通道的权值 G 设置为 0.7 时,视觉显著模型达到最优,从图 5 中可以看出 $L=1.3, G=0.7$ 时,ROC 曲线面积大于 $L=1.4, R=0.6$ 和 $L=1.1, R=0.9$,并且可以看出增加局部能量通道后的改进模型较传统 Itti 模型性能有所提高。

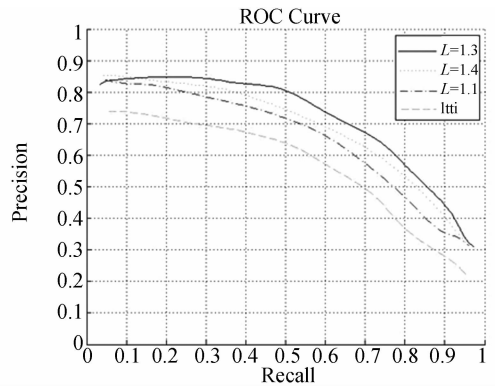


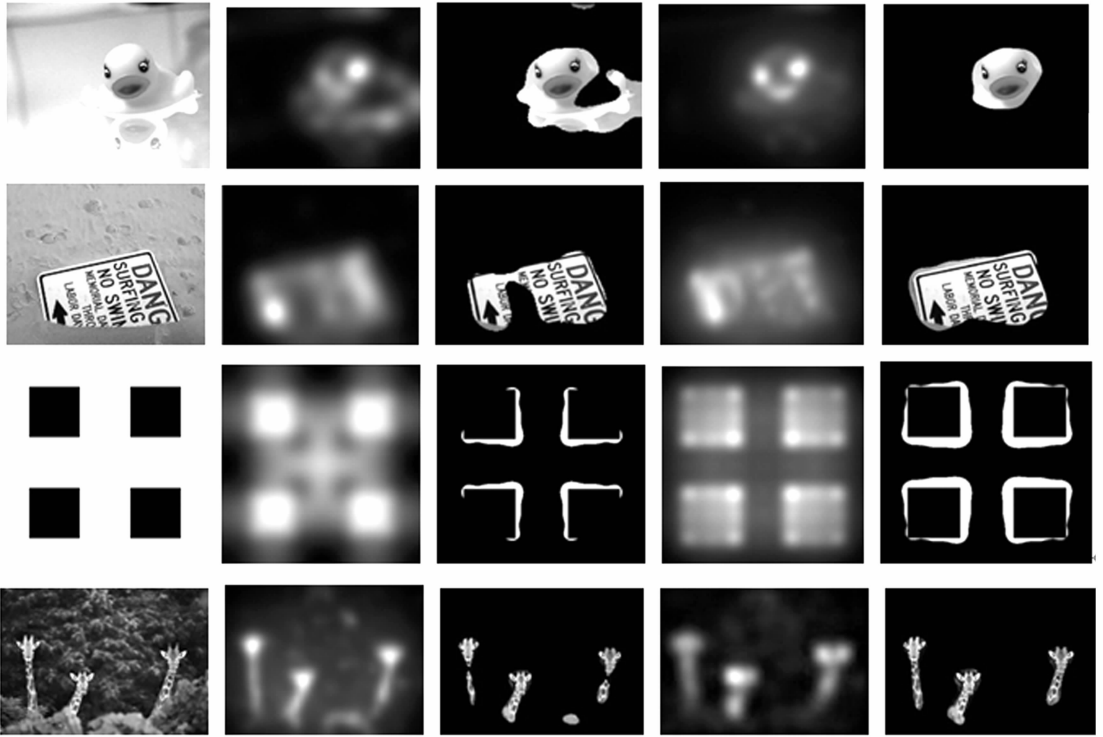
图 5 ROC 曲线调节通道权重

Fig. 5 ROC curves for selecting the weight of channel

实验二:将本文提出的算法与 Itti 模型算法进行对比,如图 6 所示,图像第一列为原图像,第二列为 Itti 显著图像,第三列为基于 Itti 模型的显著目标检测算法,第四列为本文算法构成的显著图,最后一列为本文自动目标检测算法。可以看出,本文算法很好地保证了目标的完整性,第一行和第二行实验图像为目标分布在场景的中心部位,从目标检测图像中可以看出在 Itti 模型中,目标由于缺乏轮廓的信息,使得检测时小鸭子和标牌部分显著性缺失,而本文算法在此两副图像中保证了目标的完整性。第三行图像为检测算法设计的多目标测试图像,四个方框目标分布在图像的四个角落,由于 Itti 模型目标检测结果过于强

调图像中心,所以检测图像外围目标时检测到的特征不够完整,而本文算法通过局部能量强化目标轮廓,弱化 Gabor 函数突出中心的特点,使目标检测更为完整,同样在第四列图像也体现了本

文算法的优越性。第四行图像为多目标场景图像,从检测结果看,Itti 算法在下部出现少量误报,并且和本文方法比,分布在场景两侧的长颈鹿的颈部与背景混淆。



(a) 原图像 (b) Itti 方法显著图 (c) Itti 方法显著目标检测结果 (d) 本文方法显著图 (e) 本文目标检测方法结果
 (a) Original images (b) Saliency maps of Itti's method (c) Salient objects of Itti's method (d) Saliency maps of proposed method (e) Salient objects of proposed method

图 6 不同方法实验结果比较

Fig. 6 Comparison of experimental results with different methods

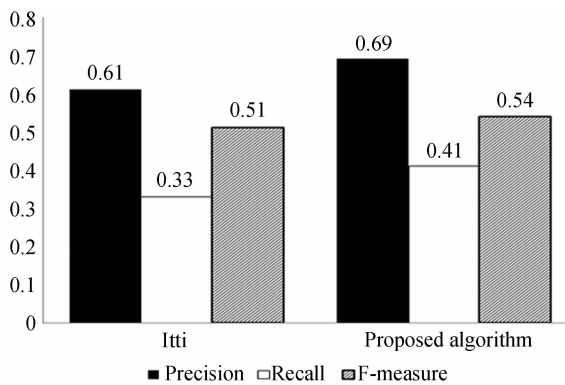


图 7 Itti 模型和本文算法的查准率、查全率和 F-measure
 Fig. 7 Comparison of average precision, recall, and F-measure values of Itti's method and the proposed framework

本文利用客观的评价方法测试了本文提出的显著性目标检测方法。如图 7 所示,本文算法在查全率和查准率上均高于 Itti 模型,为了更好地衡量本算法性能,利用 F-measure 方法^[15]算法。F-measure 方法公式:

$$F\text{-measure} = \frac{(1 + \alpha) * \text{Precision} * \text{Recall}}{(\alpha * \text{Precision} + \text{Recall})}, \quad (19)$$

其中: Precision 为查准率, Recall 为查全率, $\alpha = 0.3$ 。从图 7 中可以看出,本文算法的 F-measure 高于 Itti 模型。

5 结 论

本文将视觉显著模型与图像分割方法相结

合,提出自适应显著目标分割方法,实现了显著目标和复杂背景的分隔。在视觉显著模型中,通过新增局部能量描述目标轮廓特征,使显著目标的显著性在整幅图像中更加完整。在自适应分割过程中,利用最大熵准则估计图像中目标熵,再计算由显著图像构成预分割图像的熵,以熵作为标准,选取最优目标检测结果。本方法依据视觉注意模

型,在符合视觉生理学基础上,将反映目标轮廓信息的局部能量特征引入到显著模型中,按照显著性强弱划分目标和背景,通过最大熵准则估计显著区域目标。在对比实验中,本文算法检测显著目标更为完整,查全率、查准率和 F-measure 分别为 0.69、0.41 和 0.54,均优于传统方法。该方法基本适用于在复杂场景下检测视觉显著目标。

参考文献:

- [1] 丛明煜,何文家,逮力红,等. 复杂背景成像条件下运动点目标的轨迹提取[J]. 光学精密工程,2012,20(7): 211-217.
CONG M Y, HE W J, LU L H, *et al.*. Trace extraction of moving point targets in complex background images [J]. *Opt. Precision Eng.*, 2012, 20(7): 211-217. (in Chinese)
- [2] 曾文静,万磊,张铁栋,等. 复杂海空背景下弱小目标的快速自动检测[J]. 光学精密工程,2012,20(2): 196-205.
ZHEN W J, WAN L, ZHANG T D, *et al.*. Fast detection of weak targets in complex sea-sky background [J]. *Opt. Precision Eng.*, 2012, 20(2): 196-205. (in Chinese)
- [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] LI Z, LTTI L. Saliency and gist features for target detection in satellite images [J]. *IEEE Transactions on Image Processing*, 2011, 20(7): 2017-2029.
- [5] HOU X D, HANG L Q. Saliency detection: a spectral residual approach [C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis, USA: IEEE*, 2007: 1-8.
- [6] CHENG M M, ZHANG G X, MITRA N J, *et al.*. Global contrast based salient region detection [C]. *IEEE CVPR, Colorado Springs, USA*, 2011: 409-416.
- [7] 胡正平,孟鹏权. 全局孤立性和局部同质性图表示的随机游走显著目标检测算法 [J]. 自动化学报, 2011, 37(10): 1279-1284.
HU ZH P, MENG P Q. Graph presentation random walk salient object detection algorithm based on global Isolation and local homogeneity [J]. *Acta Automatica Sinica*, 2011, 37(10): 1279-1284. (in Chinese)
- [8] 赵宏伟,王慧,刘萍萍,等. 有指向性的视觉注意计算机模型 [J]. 计算机研究与发展, 2009, 46(7): 1192-1197.
ZHAO H W, WANG H, LIU P P, *et al.*. A computer model of directional visual attention [J]. *Journal of Computer Research and Development*, 2009, 46(7): 1192-1197. (in Chinese)
- [9] ANDREA P, MICHELA T, DOMENICO M, *et al.*. BOLD response to spatial phase congruency in human brain [J]. *Vision of Journal*, 2008, 8(10): 1-15.
- [10] LINDA H, AAPO H, SIMO V. Representation of cross-frequency spatial phase relationships in human visual cortex [J]. *The Journal of Neuroscience*, 2009, 29(45): 14342-14351.
- [11] VENKATESH S, OWENS R A. An energy feature detection scheme [C]. *The International Conference on Image Processing, Singapore*, 1989: 553-557.
- [12] MICHAEL F, GERALD S. The monogenic signal [J]. *IEEE Transactions on Signal Processing*, 2001, 49: 3136-3144.
- [13] 许元男,赵远,刘丽萍,等. 基于 Rényi 熵的显著图生成与目标探测[J]. 光学精密工程, 2010, 18(3): 723-731.
XUE Y N, ZHAO Y, LIU L P, *et al.*. Rényi entropy-based saliency map generation and target detection [J]. *Opt. Precision Eng.*, 2010, 18(3):

723-731. (in Chinese)

- [14] TAMAYO N, TRAVER V J. Entropy-based saliency computation in log-polar images [C]. *3rd International Conference on Computer Vision Theory and Applications, Funchal, Madeira, Portugal*, 2008:501-506.

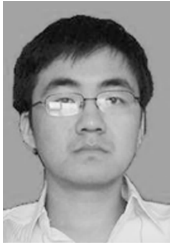
gal, 2008:501-506.

- [15] ACHANTA R, HEMAMI S, ESTRADA F, *et al.*. Frequency-tuned salient region detection [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2009:409-414.

作者简介:



赵宏伟 (1962—)男, 辽宁沈阳人, 教授, 博士生导师, 主要从事智能信息系统与嵌入式技术方面的研究. E-mail: zhaohw@jlu.edu.cn



陈霄 (1983—)男, 吉林长春人, 博士研究生, 2006年、2010年于吉林大学分别获得学士、硕士学位, 主要从事机器视觉、认知计算方面的研究. E-mail: xiaoxiaocctv5@126.com



刘萍萍 (1979—), 女, 吉林长春人, 博士, 讲师, 2001年、2004年和2009年于吉林大学分别获得学士学位、硕士学位和博士学位, 主要从事机器视觉、模式识别、嵌入式系统等方面的研究. 通信作者, E-mail: liupp@jlu.edu.cn



耿庆田 (1972—), 男, 江苏邳县人, 博士研究生, 讲师, 2005年于吉林大学获得硕士学位, 主要从事计算机网络与智能信息系统等方面的研究. E-mail: qt-geng@163.com

(本栏目编辑: 李自乐)

(版权所有 未经许可 不得转载)