

基于生物激励计算模型在图像显著性提取中的研究

纪超,刘慧英,邵刚,孙景峰

(西北工业大学 自动化学院,陕西 西安 710072)

摘要: 提出一种生物激励的显著性特征计算模型。首先通过注意块学习从眼动数据库中选择与视觉响应一致的稀疏基;然后基于稀疏基表达原理对图像建立计算模型并提取显著性特征:包括全局连续性、区域颜色对比以及局部复杂度对比特征;再仿照细胞调节原理,提出新的特征组合方法进行特征融合。最后将该算法在多个典型的场景中对感兴趣区进行提取实验,证明比其他算法具有优越性。并提出将此算法应用于虚拟与现实场景融合中,能良好地提取出真实场景中的有效区域和剔除虚景区域。

关键词: 机器视觉; 显著性提取; 生物激励; 稀疏表达

中图分类号: TP911 **文献标志码:** A **文章编号:** 1007-2276(2013)03-0823-06

Research on biologically-inspired computational model for image saliency detection

Ji Chao, Liu Huiying, Shao Gang, Sun Jingfeng

(School of Automation, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: A biologically-inspired model for the computer vision community was proposed. At first, a set of basis functions that accorded with visual responses to natural stimuli was learned by using eye-fixation patches from an eye-tracking dataset. Then image calculation model was established and features was derived based on the principle of sparse representation: including global continuity, regional color contrast, and local complexity contrast. And then refer to the principle that activity in cells responding to stimuli, a new feature combination theory was proposed to achieve features fusion. Afterwards, some experiments extracting regions of interest from typical scenes prove that this algorithm has superiority than other algorithms, and the algorithm was applied in virtual and reality interactivity. It can effectively extract effective regions and eliminate virtual scene area.

Key words: machine vision; saliency detection; biologically-inspired; sparse representation

收稿日期:2012-07-13; 修订日期:2012-08-15

基金项目:国家自然科学基金(40905011)

作者简介:纪超(1987-),男,博士生,主要从事机器视觉与图像处理方面的研究。Email:dachao9898@163.com

导师简介:刘慧英(1956-),女,教授,博士生导师,博士,主要从事机器视觉与图像处理、计算机控制及网络控制等方面的研究。

Email:lhy2005@nwpu.edu.cn

0 引言

近年来越来越多的生物学、神经学、心理学实验证明了视觉注意计算模型对于显著性提取发展的重要性。正确地提取图像的关键区域可大大提高图像分析处理效率,避免不必要的计算资源浪费。图像稀疏表达已成为目前研究的热点,尤其是探讨基于人眼视觉的基函数理论模型和构造方法,以及研究快速、有效的图像稀疏表达算法成为机器视觉和图像处理研究中的前沿课题。

神经生理学和成像理论研究表明,视觉关注的分析方法可分为自底向上和自顶向下两类模型:前一类是基于低级视觉,由数据驱动,自动捕获刺激;后一类是基于高级视觉,由知识和任务驱动,以自我意识决定视觉关注区域^[1]。

目前的视觉注意显著区域检测算法以局部对比度算法为主,其中以 Itti 提出的算法^[2]是最为代表性的模型;Achanta 等人^[3]通过多个尺度领域对比的差异叠加来计算像素的显著度,从而产生全分辨率的显著性图;Goferman 等人^[4]根据新颖的特征和颜色具有高显著度,并集中分布而不是分散在图像中的观点来计算显著区域;基于全局对比度的显著性区域计算方法有 Hou Xiaodi^[5]在频谱域上通过分析显著区域的特征构建相应的显著图,该算法与 Itti 的算法相比,在计算速度和检测效果上有一定的优势,但该算法的缺陷是倾向于小目标检测。

单纯基于全局的计算模型问题在于注意区域的尺寸不容易确定,造成注意区域不能包含完整的物体或是多个物体;单纯的视觉注意显著区域检测算法,主要依赖像素间的对比,忽略了感知物体的大小和位置等信息。文中从生物学规律出发,在稀疏基的基础上对显著图进行数学建模分析,并采用视觉注意显著计算模型提取图像的全局、区域以及局部特征,并采用特征独立联合作用的机制对特征进行线性与非线性组合得到显著图。

1 稀疏表达方式

稀疏编码原理认为视觉系统中简单细胞的感受野能够被某些基函数所刻画,这些基函数被认为是神经元的响应^[6]。因此一个图像块 A 能被表示成一

组基的线性组合,如公式(1)所示:

$$A = \sum \alpha_k A_k \quad (1)$$

式中: A_k 和 α_k 表示第 k 个基和对应的响应系数;

$F_k = A_k^{-1}$ 是第 k 个滤波函数, α_k 通过以下公式计算:

$$\alpha_k = \sum_{(x,y)} F_k(x,y) A(x,y) \quad (2)$$

参考文献[7]中表明,基函数是从大规模数据库上通过独立成分分析 ICA,随机采样学习所得。理论上,如果训练数据足够丰富,就可以得到一组完备基,但实际上要受计算机资源有限性的限制。因此实际应用中,基于具体任务训练的基会得到更好的性能。

文中实验中从眼动数据库中选择注意块,采用 FastICA 方法来学习基函数(EFLBF)。为了证明从注意点处选择的基对于提取显著物体的优势,文中从同一数据库中随机选择相同数量的模块,用同样的方法进行训练,形成一组对比基(RSLBF)。图 1、图 2 显示了从注意点上学习得到的基和随机选取学习得到的基。

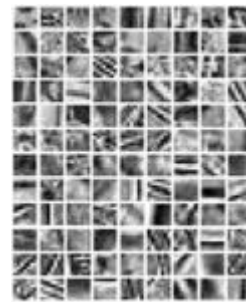


图 1 学习训练基

Fig.1 Eye-fixation patches learned basis function

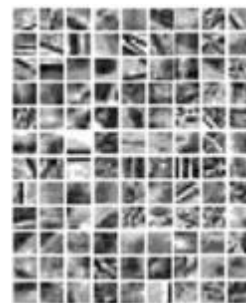


图 2 随机训练基

Fig.2 Random patches learned basis function

2 算法分析

文中根据以下三种稀疏结构机理和视觉注意计算模型对初级视觉神经元进行建模:(1) 稀疏基特征简单,与神经元活动相似,基于全局的方法倾向于将

大范围的目标和周围环境分开,并且连续物体更能引起注意;(2)一个神经元的激活性要受它的领域驱动,某区域的显著性主要是由它和周围区域的对比度决定,颜色特征是区域对比度中最为显著的特征;(3)复杂度对比影响着观察者的视觉搜索和对刺激的感知,可用稀疏基建模简化分析。

2.1 全局连续性特征

影响图像显著性特征检测的主要因素是噪声,首先需对图像进行预处理去噪,传统方法是采用高斯滤波器对图像进行平滑,但高斯滤波器在消除噪声的同时模糊了图像的边缘信息。文中使用自适应平滑滤波器,该滤波器是基于各向异性热扩散方程,既考虑到点自身的梯度特征又考虑到局部的结构特征^[8],由于显著边缘的梯度幅度较大,需要借用松弛迭代的思想,使得边缘点的幅度随迭代次数的增加而增强,而梯度幅度较小的噪声随着迭代而得到平滑。算法的具体实现步骤如公式(3)~(7)所示:

(1) 计算梯度

$$G_x(x, y) = 0.5 [P(x+1, y) - P(x-1, y)] \quad (3)$$

$$G_y(x, y) = 0.5 [P(x, y+1) - P(x, y-1)] \quad (4)$$

(2) 计算导热系数

$$c(x, y) = \exp \left[- \frac{G_x^2(x, y) + G_y^2(x, y)}{2\sigma^2} \right] \quad (5)$$

(3) 由 $P^l(x, y)$ 迭代计算 $P^{l+1}(x, y)$, 其中

$$P^0(x, y) = P(x, y) \quad (6)$$

$$P^{l+1}(x, y) = \frac{\sum_{i=-1}^1 \sum_{j=-1}^1 P^l(x+i, y+j) + c^l(x+i, y+j)}{\sum_{i=-1}^1 \sum_{j=-1}^1 c^l(x+i, y+j)} \quad (7)$$

导热系数 $c(x, y)$ 为梯度幅度的降函数,梯度幅度高处导热系数小,梯度幅度低处导热系数大,即孤立的梯度突变点将被平滑;而局部具有整体集团优势的突变点,即显著边缘将被增强。文中将滤波后的图像用稀疏基来对其建模,通过构建基函数的系数直方图得到概率分布,从而得到边缘连续信息。首先,通过公式(2)得到基函数的系数,然后构建一个系数直方图,直方图被转化成离散的概率分布 $P_n(x, y)$,用高斯核建模^[9]如公式(8)所示:

$$C_n(x, y) = \exp \left(- \frac{P_n(x, y)}{2\sigma^2} \right) \quad (8)$$

由于直方图可以描述系数的一致性,那么连续

性特征通过公式(9)对所有系数相加得到:

$$FC(x, y) = \sum_n C_n(x, y) \quad (9)$$

将原图经过自适应滤波,并在稀疏基表示的基础上提取连续性特征得到结果,如图3所示。

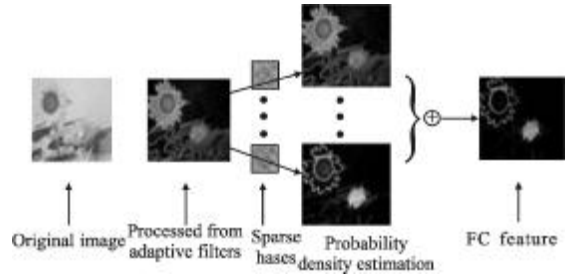


图3 连续性特征提取框架图

Fig.3 Framework of continuity feature extraction

2.2 区域颜色对比特征

一个像素的显著性和自身像素值无关,而是用它和图像中其他像素颜色的对比度定义。空间关系在人类注意力方面也起到非常大的作用,近邻的区域增大影响,较远的区域减少影响;相邻区域的高对比度比比较远区域的高对比度更容易引起视觉注意。

文中提出通过计算区域颜色对比度来加权区域稀疏基,将空间关系和区域级对比度计算相结合来定义图像区域的显著权值大小。首先用图像分割方法将输入图像分割成 n 个区域,然后通过稀疏基来表示每个区域,并建立颜色直方图,再通过公式(10)测量每个区域 r_i 与其他区域 r_k 的颜色对比度来计算它的显著权值大小。

$$w(r_k) = \sum_{r_i \neq r_k} r_i(n) d_k(r_k, r_i) \quad (10)$$

其中用区域 r_i 里的像素个数 $r_i(n)$ 来强调大区域的颜色对比度, $d_k(r_k, r_i)$ 为两个区域的颜色距离度量。两个区域 r_1 和 r_2 的颜色距离计算方法如公式(11)所示:

$$d_r(r_1, r_2) = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} w_c(c_{1,i}) w_c(c_{2,j}) d(c_{1,i}, c_{2,j}) \quad (11)$$

式中: $w_c(c_{k,i})$ 为第 i 个颜色 $c_{k,i}$ 在第 k 个区域 r_k 的所有 n_k 种颜色中出现的概率。使用区域颜色直方图构造概率密度函数计算权值,以强调主要的颜色之间的区别^[10]。

文中还在公式(10)的基础上引进空间权值来加强空间信息对于区域显著性的影响。对任意区域 r_k , 空间位置对区域显著性对比度的影响大小可通过公式(12)计算^[11]:

$$W(r_k) = \sum_{r_i \neq r_k} \exp(-d_s(r_k, r_i)/\sigma) w(r_k) \quad (12)$$

式中: $d_s(r_k, r_i)$ 为区域 r_k 和 r_i 的空间距离, 定义为两个区域重心的欧氏距离; σ 是调节空间作用的强度。 σ 越大, 空间作用的影响越小, 文中试验中选择 $\sigma=0.5$ 。最后按照公式(13)加权稀疏基。图 4 显示了整个处理过程:

$$FR(x, y) = \sum_{i=1}^n \alpha(x, y) W(r_k) \quad (13)$$

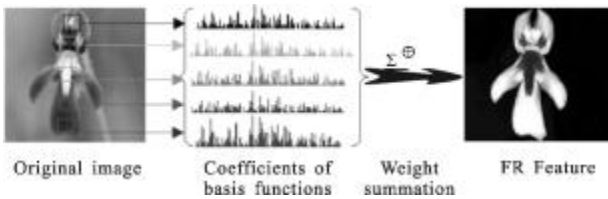


图 4 区域颜色对比特征计算框架图

Fig.4 Framework of regional color contrast

2.3 局部复杂度对比特征

参考文献[12]中提到每一个像素都包含一些信息, 图像的全部信息就是各部分对应的熵的集。根据一个区域所含的信息越多, 该区域就越显著, 提出特征集中的度量方法来估计局部的复杂度, 文中在稀疏基上进行建模, 首先通过公式(14)、(15)计算特征信息的平均值 \overline{SL} , 再将图像划分为指定的区域, 并计算各区域的 SL , 与 \overline{SL} 进行比较得出是否为特征集中区。

$$SL = \sum_{xy \in \Omega} (\beta E(x, y) + (1 - \beta) Q(x, y)) \quad (14)$$

$$\overline{SL} = SL / \Omega_{area} \quad (15)$$

式中: $E(x, y)$ 和 $Q(x, y)$ 表示熵和能量, 计算方法如公式(16)、(17)所示:

$$E(x, y) = \sum_i p_i \log p_i, p_i = \frac{|\alpha_i|}{\sum_j |\alpha_j|}, \alpha_i, \alpha_j \in \alpha(x, y) \quad (16)$$

$$Q(x, y) = \sum_i \alpha_i^2, \alpha_i \in \alpha(x, y) \quad (17)$$

复杂度特征计算方法如公式(18)所示:

$$FL(x, y) = \exp\left(-\frac{SL}{2\sigma^2}\right) \quad (18)$$

在公式(14)中, 参数 β 是平衡图像复杂图的熵和能量的因子。在文中试验中 $\beta=0.3$, 参考文献[13]指出显著图通常可划分为两种典型的对比模型: 光滑背景和复杂前景组合以及光滑前景和复杂背景组合模型。根据公式(18)可直接提取光滑背景复杂前

景模型中前景; 对于处理光滑前景复杂背景模型, 文中采用如图 5 所示的翻转算法。

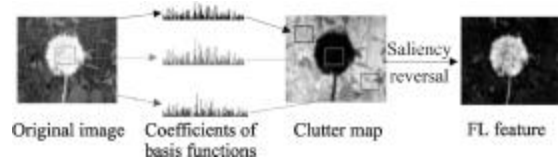


图 5 复杂度对比特征提取框架图

Fig.5 Framework of local complexity contrast

3 特征的独立与联合作用

参考文献[14-15]指出没有任何神经网络结构和相关领域的研究能证明特征之间完全是线性组合, 传统的特征组合方法一般不会考虑特征之间相互作用问题, 或者大多都是基于特定实验环境和经验做出的假设, 其权值指定都依赖经验值或机器学习。文中提出采用特征独立和联合作用的组合方式: 一部分来自特征的独立线性组合, 另一部分来自特征的联合作用部分。假设特征集合 $F = \{f_1, f_2, \dots, f_n\}$, 其中 N 为归一化因子。特征独立联合作用的组合方式如公式(19)所示:

$$S = \frac{1}{N} \left(\sum_{i=1}^n f_i^2 + \sum_{i=1, j=1, i \neq j}^n f_i * f_j \right) \quad (19)$$

文中在提取三种特征基础上, 通过特征独立和联合作用的方式将 $FC(x, y)$, $FR(x, y)$ 和 $FL(x, y)$ 组合起来, 形成最终的显著图 S 。具体计算公式如下:

$$S = \frac{1}{N} (FC^2(x, y) + FR^2(x, y) + FL^2(x, y) + FC(x, y) \cdot FL(x, y) + FC(x, y) \cdot FR(x, y) + FR(x, y) \cdot FL(x, y) + FC(x, y) \cdot FR(x, y) \cdot FL(x, y)) \quad (20)$$

4 测试实验

为了验证算法的有效性, 实验采用 Intel 酷睿 i5 处理器, 主频 3 GHz, 内存 2 GHz 的配置, 环境为 Matlab 2010 a。文中共做了两组评估实验。(1) 比较基于注意块学习基和随机选取基的性能; 图 6 显示了基于两种不同基下特征的 ROC 曲线。比较得出, 经过学习的基能够提高稀疏基的表达准确性。(2) 通过从原始库 MSRA 中公开数据集中挑选出来的 1000 幅图和它们的基准图作为测试数据, 将文中算法和现有的算法比较, 包括谱残余假设算法(SR)^[16]、频率协调算法(FT)^[3]、编码长度增量算法(ICL)^[17]、

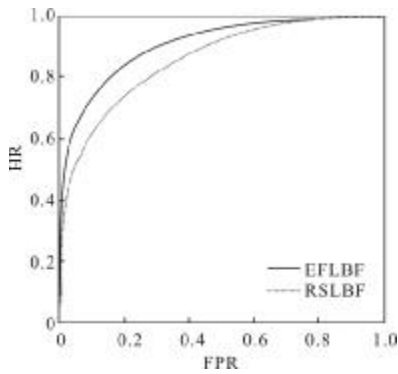


图 6 EFLBF 和 RSLBF 的 ROC 曲线
Fig.6 ROC curves of EFLBF and RSLBF

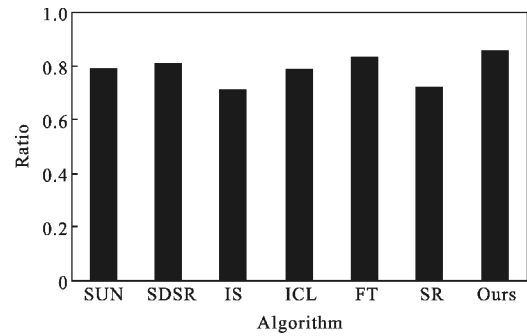


图 9 不同算法 ROC 曲线下的面积
Fig.9 The area under ROC curves of different algorithms

图像显著性(IS)^[18]、自相似算法(SDSR)^[19]以及统计学方法(SUN)^[20]。

从图 7 看出 SDSR、SUN 算法倾向于在边缘部分产生高显著性值,FT 虽然检索到显著区域,但区域与背景的灰度值相差不大,不利于后面的图像分割。文中算法可以很好提取边缘和内部信息,较好地抑制杂乱背景影响,能连续均匀地突出整个显著区域。算法性能比较也可以用 ROC 曲线来衡量,图 8 和图 9 显示了不同算法的 ROC 曲线和对应的面积图。比较结果可以看出文中算法能更好地获取显著区域。

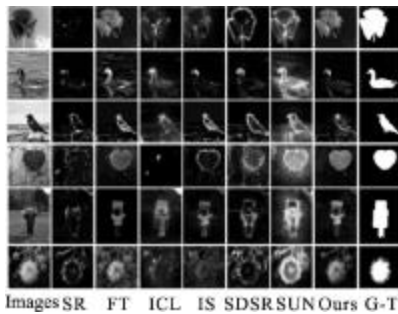


图 7 不同的 saliency 算法处理的结果图
Fig.7 Saliency detection by different algorithms

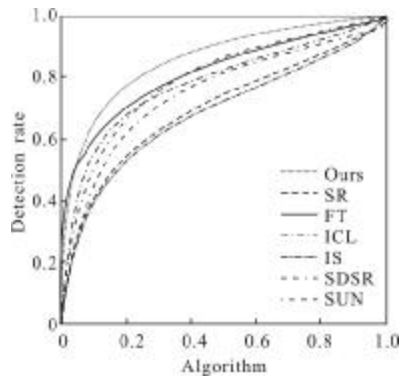


图 8 不同方法的 ROC 曲线
Fig.8 ROC curves of saliency detection by different algorithm

5 视觉注意模型在虚拟与现实融合中应用

虚拟与现实交互技术是一门新兴的综合技术体系,用户借助必要的设备以自然的方式与虚拟环境中的对象进行交互。其中虚拟与现实场景的融合是虚拟与现实交互中最关键的部分,比如在电视节目一些美轮美奂、惊心动魄的场景和特技,大都是通过软件制作出的虚拟场景与拍摄的真实前景进行融合所得,目前合成方式是检测前景中每一个像素点,如果像素值是图 10(a)中的灰色就换成对应虚拟场景的像素值,此方法会把前景中的虚景区比如图 10(a)左上方的污区当成有效区保留下,如图 10(b)所示;将图 10(a)经过文中算法后得到图 11(a),并与虚拟背景合成后得到图 11(b)。证明文中提出的显著性计算模型可运用于虚拟与现实融合。

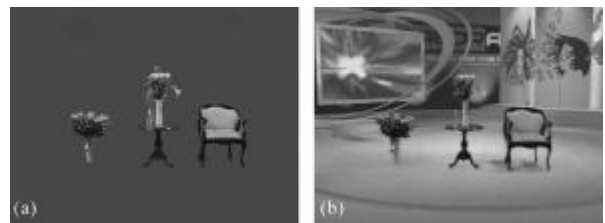


图 10 真实前景与其合成图
Fig.10 Real foreground and its composite diagram

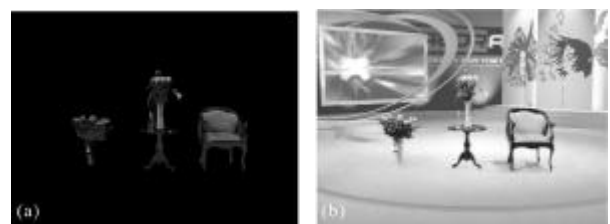


图 11 处理后前景与其合成图
Fig.11 Processed diagram and its composite diagram

6 总 结

文中从生物学规律出发,在稀疏基的基础上对显著图进行建模,加入更多的语义信息进行辅助处理,同时考虑了显著物体的全局性、独立性和空间相干性。与其他算法相比,文中算法的检测结果具有更高的精度,更符合人类视觉特性。

将文中算法运用于虚拟与现实融合中可很好地提取图像的有效区和剔除虚景区域,有广泛的应用场合和可观的经济价值。

参考文献:

- [1] Koene A R, Zhaoping L. Feature-specific interactions in Saliency from Combined Feature Contrasts: Evidence for a Bottom-Up Saliency Map in V1 [J]. *Journal of Vision*, 2007, 7(7): 1-14.
- [2] 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.
- [3] Achanta R, Hemami S, Estrada F, et al. Frequency-tuned salient region detection [C]//*IEEE Conference on Computer Vision and Pattern Recognition*, 2009: 1597-1604.
- [4] Goferaman S, Zelnik L, Tal A. Context aware saliency detection [C]//*IEEE Conference on Computer Vision and Pattern Recognition*, 2010: 2376-2383.
- [5] Hou X, Zhang L. Saliency Detection: A Spectral Residual Approach [C]//*IEEE Conference on Computer Vision and Pattern Recognition*, 2007: 1-8.
- [6] Wang W, Wang Y, Huang Q, et al. Measuring visual saliency by site entropy rate [C]//*IEEE Conference on Computer Vision and Pattern Recognition*, 2010: 135-143.
- [7] Hou X, Zhang L. Dynamic visual attention: searching for coding length increments[C]//*Advances in Neural Information Processing Systems(NIPS)*, 2008: 35-41.
- [8] Lin Y, Fang B, Yuanyan T. A computational model for saliency maps by using local entropy [C]//*Proceedings of Twenty-Fourth AAAI Conference of Artificial Intelligence*, 2010: 967-973.
- [9] Yang Yawei, Li Junshan, Yang Wei, et al. Multiclass and multiview object detection approach based on sparse biological vision features[J]. *Infrared and laser Engineering*, 2012, 41(1): 267-272. (in Chinese)
- 杨亚威, 李俊山, 杨威, 等. 利用稀疏化生物视觉特征的多类多视角目标检测方法[J]. *红外与激光工程*, 2012, 41(1): 267-272.
- [10] Cheng M M, Zhang G X, Mitra N J. Global contrast based salient region detection [C]//*IEEE Conference on Computer Vision and Pattern Recognition*, 2011: 409-416.
- [11] F Shahbaz Khan, R Muhammad Anwer, van de weijer, et al. Color attributes for object detection[C]//*IEEE Conference on Computer Vision and Pattern Recognition*, 2012.
- [12] Wang Z, Li B. A two-stage approach to saliency detection in images [C]//*Proceeding of IEEE International Conference on Acoustics Speech and Signal Processing*, 2008: 965-968.
- [13] He S, Han J, Hu X, et al. A Biologically inspired computational model for image saliency detection [C]//*MM' 11 Proceedings of the 19th ACM International Conference*, NY, USA, 2011: 1465-1468.
- [14] Bi Fukun, Gao Lining, Long Teng, et al, Remote sensing target detection based on visual saliency guidance and classifier fusion [J]. *Infrared and laser Engineering*, 2011, 40 (10): 2058-2064. (in Chinese)
- 毕福昆, 高立宁, 龙腾, 等. 结合视觉显著性引导与分类器融合的遥感目标检测 [J]. *红外与激光工程*, 2011, 40(10): 2058-2064.
- [15] Bruce N, Tsotsos J K. Saliency, attention, and visual search: an information theoretic approach [J]. *Journal of Vision*, 2009, 9(3): 1-24.
- [16] Hou X, Zhang L. Saliency detection: a spectral residual approach [C]//*IEEE Conference on Computer Vision and Pattern Recognition*, 2007: 1-8.
- [17] Olshausen B. Emergence of simple-cell receptive field properties by learning a sparse code for natural images [J]. *Nature*, 1996, 38(3): 607-609.
- [18] Hou X, Harel J, Koch C. Image signature: highlighting sparse salient regions [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(1): 194-201.
- [19] Seo H, Milanfar P. Nonparametric bottom-up saliency detection by self-resemblance [C]//*IEEE Conference of Computer Vision and Pattern Recognition*, 2009: 420-426.
- [20] Zhang L, Tong M H, Marks K T, et al. SUN: a Bayesian framework for saliency using natural statistics[J]. *Journal of Vision*, 2008, 8(7): 1-10.