

# 基于 Local 特征和 Regional 特征的图像显著性检测

郭迎春<sup>1</sup> 袁浩杰<sup>1</sup> 吴鹏<sup>1</sup>

**摘要** 提出了一种基于颜色空间的 Local 特征和 Regional 特征的自然图像显著性检测方法. 该方法将图像分成  $8 \times 8$  的子块, 计算多个尺度下每一个子块的 Local 特征和 Regional 特征, 并将其加权组合来确定子块的显著程度, 从而得到整个图像的显著特征. 此外, 通过计算 4 个颜色通道上的色度对比度, 获得显著物体的边缘. 将图像的显著特征与显著物体的边缘综合后得到图像中的显著目标. 实验结果显示, 本文提出的方法能够快速、清晰而准确地提取出图像中的显著性目标.

**关键词** 显著性检测, Local 特征, Regional 特征, 多尺度

**引用格式** 郭迎春, 袁浩杰, 吴鹏. 基于 Local 特征和 Regional 特征的图像显著性检测. 自动化学报, 2013, 39(8): 1214–1224

**DOI** 10.3724/SP.J.1004.2013.01214

## Image Saliency Detection Based on Local and Regional Features

GUO Ying-Chun<sup>1</sup> YUAN Hao-Jie<sup>1</sup> WU Peng<sup>1</sup>

**Abstract** This paper proposes a model of salient region detection based on local and regional features in color space. Firstly, the image is divided into  $8 \times 8$  sub-blocks, in each sub-block, then the multi-scale local and regional features are calculated and combined by weighted summation as the sub-block's salient value; secondly, the salient edge is calculated by the color contrast of the four color channels; finally, the salient map can be extracted by combining the salient features and salient edge together. The experiment results show that our model can extract salient objects in images fast and exactly.

**Key words** Saliency detection, local feature, regional feature, multi-scale

**Citation** Guo Ying-Chun, Yuan Hao-Jie, Wu Peng. Image saliency detection based on local and regional features. *Acta Automatica Sinica*, 2013, 39(8): 1214–1224

在数字图像处理的诸多领域内, 显著区域的检测有着广泛的应用, 如基于内容的图像检索<sup>[1]</sup>、图像自动剪裁<sup>[2]</sup>、图像视频压缩<sup>[3]</sup>, 以及在小的显示设备上调节图像的大小<sup>[4]</sup>等.

显著区域检测的本质是一种视觉注意模型 (Visual attention model, VAM). VAM 是依据视觉注意机制而建立的模型, 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<sup>[5]</sup>. 研究表明, 人们产生注意力有两方面的原因: 一方面是视觉刺激引起的注意, 另一方面是基于任务驱动的主观注意. 由此 VAM 分为两种类型: 自底向上 (Bottom-up) 的数据驱动注意模式和自顶向下 (Top-down) 的任务驱动注意模式<sup>[5]</sup>. 自底向上的注意模型<sup>[6–8]</sup>分为两个主要的步骤: 底层视觉特征的提取和综合提取出的特征生成显著图.

自顶向下的注意模型<sup>[9–10]</sup>主要是通过调整选择准则, 以适应外界命令的要求来获得显著图.

在显著性检测中, 自底向上的模型是比较常见的. 其中经典的 Itti 模型<sup>[6]</sup>, 根据人眼视觉特性, 采用高斯金字塔进行非均匀采样生成 9 个尺度图像, 对 9 个尺度图的颜色、亮度、方向等底层特征进行计算, 根据中央周边差计算各个尺度上的关注图, 最后对关注图进行融合得到显著图. 以 Itti 模型为基础, 文献 [11] 在考虑视觉焦点转移和选择时, 分析了显式注意和隐式注意两种情况, 从而提高对视觉焦点的准确定位; 而文献 [12] 在分析了环境上下文、返回抑制等 5 种计算模型后提取显著目标. 但上述模型在提取图像显著特性时仅仅考虑了图像的局部特征, 不能全局地考虑图像显著特性. Jiao 等<sup>[13]</sup>提出的自底向上的检测模型, 在多尺度的基础上, 通过计算全局特征来确定区域的视觉刺激程度, 将较为显著的区域节点传递给上层进行计算, 其余的区域被合并在一起, 通过区域合并最终生成图像的显著区域, 该方法中多尺度提取特征的计算相对简单, 较采用局部特征计算显著性的模型有更好的计算精度. 但是, 该计算模型容易忽略掉图像中较小的显著物体, 此外, 该方法获得的显著图中的物体没有明显的

收稿日期 2011-07-25 录用日期 2012-07-25  
Manuscript received July 25, 2011; accepted July 25, 2012  
国家自然科学基金 (60302018), 河北省科技支撑计划项目 (11213518D) 资助  
Supported by National Natural Science Foundation of China (60302018) and Hebei Science and Technology Support Program (11213518D)  
本文责任编辑 戴琼海  
Recommended by Associate Editor DAI Qiong-Hai  
1. 河北工业大学计算机科学与软件学院 天津 300401  
1. School of Computer Science and Engineering, Hebei University of Technology, Tianjin 300401

边缘, 仅在某一区域内有较大的亮度值. 文献 [14] 通过计算不同尺度图像上的亮度、颜色、纹理等底层对比特征后, 将这些对比特征进行插值迭代生成三幅特征图, 最后, 将三幅特征图综合, 生成最终的显著区域. 该方法在提取底层的特征时局限在局部范围内, 仅在归一化时考虑了全局特征, 由于特征提取的范围较小, 导致了较大显著物体的整体性较差, 同时也不利于复杂背景下显著物体的整体表达. 文献 [15] 在多尺度提取图像的底层特征的基础上, 在频域范围内分析了特征图的幅度谱和相位谱, 在空间域上计算显著图, 以此检测出图像中的显著区域. 这种方法处理速度较快, 主要缺点是没有考虑局部区域的特征, 显著物体的边缘不够明确. 自顶向下注意模式涵盖了自底向上注意模式的内容, 由于高层知识的参与, 自顶向下的模型比自底向上的模型更加复杂, 研究成果相对较少<sup>[5]</sup>. 文献 [16] 融合了自底向上和自顶向下两种思路, 提取了图像的多尺度对比、中央周边直方图等特征确定图像的显著区域, 但该方法计算比较复杂, 时间和空间的耗费较大.

度量和计算图像显著性的时候, 不同大小的区域可以反映出不同的显著特性, 提取不同区域内的特征, 获得综合的图像显著性信息, 可以更加准确地描述图像的显著性目标. 基于此, 本文提出了一种基于 Local 特征和 Regional 特征的图像显著性检测方法. 一幅图像中很少有小于  $8 \times 8$  像素块的单一显著物体, 本文在显著性检测时, 以大小为  $8 \times 8$  像素块为单位, 分别计算多尺度下综合 Local 特征和 Regional 特征的显著图, 以及显著物体的边缘, 确定图像中的显著物体.

本文组织结构如下: 第 1 节主要介绍 Local 特征提取的相关概念及过程; 第 2 节介绍 Regional 特征提取的相关步骤; 第 3 节主要描述了 Local 特征和 Regional 特征的特点以及两种特征在单一尺度下的综合和多尺度下的综合, 最终提取出图像中的显著特征; 第 4 节引入了提取图像中显著物体边缘的方法, 显著物体边缘可以更好地描述图像中的显著物体; 第 5 节综合了图像中显著物体以及边缘特征, 最终生成显著图像; 第 6 节和第 7 节的内容分别为实验结果及相关算法的比较和结论.

## 1 Local 特征的提取

### 1.1 Local 特征提取的本质

以块为单位进行显著性检测既符合人类的视觉特性, 又利于显著性检测模型的建立. 图像中一个块的显著程度是通过其与周围区域颜色的突变与反差所体现的, 突变越大对人眼的刺激就越强, 越能引起人眼的注意. 这种变化往往表现在较大显著物体的边缘或较小的显著物体, 尽管人眼可以很快地定位

到一个显著物体的整体而非边缘, 但边缘的突变是引起这种注意的根本原因. Local 特征的提取本质上是上述过程的模拟. 在进行 Local 特征提取的时候, 图像的最大边像素被调整到 384 像素, 另一边的大小是按照原比例调节设定的, 这里最大边 384 像素是考虑到运算的速度和实际得到的显著图的效果来设定的.

### 1.2 提取 Local 特征的区域与步骤

Local 区域本身是一个模糊的概念, 在视觉系统中没有明确的界线, 在建立特征提取模型的时候, 选取以块  $B(i, j)$  为中心的  $l \times l$  块的区域为 Local 区域, 如图 1 所示. 如果 Local 区域选取太大, 计算时就需要考虑距离的因素, 会使计算变得复杂, 同时也不符合 Local 特征提取的本质, Local 特征提取的本质是宏观地检测图像中最为显著的区域, 该特征的提取不考虑图像中显著物体的整体性, 这种非整体的检测并非是不够的. 由于在 Regional 特征提取的时候, 一些较小的显著物体在较大的子块邻域内计算时会被淡化, 这就需要通过 Local 特征的提取来保留图像中显著物体信息的完整性. 在本文中, 取  $l$  的值为 3.

$B(i-1, j-1)$	$B(i-1, j)$	$B(i-1, j+1)$
$B(i, j-1)$	$B(i, j)$	$B(i, j+1)$
$B(i+1, j-1)$	$B(i+1, j)$	$B(i+1, j+1)$

图 1 Local 特征提取的范围

Fig. 1 The region of local feature

根据人眼视觉特性可知, 物体的显著程度是通过颜色的对比与反差体现的, 在度量某一块的显著程度时, 计算与该块相邻区域颜色的对比与反差是确定该块显著程度的重要途径. 这种对比与反差在数学上可以用平方距离表示, 平方距离越大, 颜色的变化越大. 也可以用颜色向量的相关性来表示, 向量相关性越低, 颜色的变化越大. 由于相关性计算相对复杂, 故在本文中选择了平方距离进行度量. 块  $B(i, j)$  的 Local 特征提取如下:

1) 计算  $B(i, j)$  与其相邻子块 (即以  $B(i, j)$  为中心的  $3 \times 3$  块的区域  $B(m, n)$ ,  $m = i-1, i, i+1$ ;  $n = j-1, j, j+1$ ) 在  $R$ 、 $G$ 、 $B$  三个分量上的绝对差值. Local 特征中, 块  $B(i, j)$  与  $B(m, n)$  之间的颜色变化值  $Dis_{RGB}(B(i, j), B(m, n))$  表示为

$$Dis_{RGB}(B(i, j), B(m, n)) = (R_{i,j} - R_{m,n})^2 + (G_{i,j} - G_{m,n})^2 + (B_{i,j} - B_{m,n})^2 \quad (1)$$

如果式 (1) 的值小于阈值  $T$ , 则表示颜色变化不大, 不足以刺激人的眼睛, 在显著性检测时应该抑制淡化这种轻微的变化, 反映在数学形式上即置为 0.

如果式 (1) 的值大于阈值  $T$ , 则表明该变化能够引起视觉的注意, 应该正向地放大这种变化, 其中,  $\alpha$  为大于 1 的放大因子. Local 特征中, 块  $B(i, j)$  与块  $B(i + m, j + n)$  间的显著值为

$$Dis_{Loc-rgb}(B(i, j), B(m, n)) = \begin{cases} 0, & Dis_{rgb}(B(i, j), B(m, n)) \leq T \\ \alpha \times Dis_{rgb}(B(i, j), B(m, n)), & Dis_{rgb}(B(i, j), B(m, n)) > T \end{cases} \quad (2)$$

2) 将与  $B(i, j)$  相邻块的颜色显著值相加, 取平均值为该块的显著值, 显著值可以反映出该块的显著程度. Local 特征中, 块  $B(i, j)$  的显著程度为

$$S_{Loc}(i, j) = \frac{1}{9} \sum_{m=i-1}^{i+1} \sum_{n=j-1}^{j+1} Dis_{Loc-rgb}(B(i, j), B(m, n)) \quad (3)$$

3) 重复步骤 1) 和 2), 依次计算图像内每一个块的显著值. 由于显著值是以块为单位进行计算的, 故每一个块的显著值在表示图像显著特性的灰度图中实际上代表了  $8 \times 8$  个像素点的值, 如图 2 所示. 64 个等值的像素点会使得显著图出现方块化效应, 如图 3 所示. 为了消除这种人为的块效应使显著图变得平滑, 需要对显著图进行平滑滤波处理. 本文中使用了高斯滤波将显著图进行平滑操作.

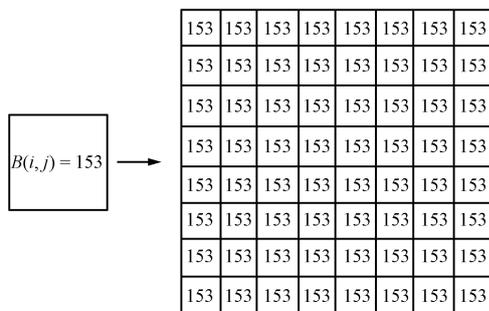


图 2 块  $B(i, j)$  的显著值与显著图中某子块灰度值的对应关系

Fig. 2 The relationship between the salient value of block  $B(i, j)$  and the gray values of its corresponding sub-block

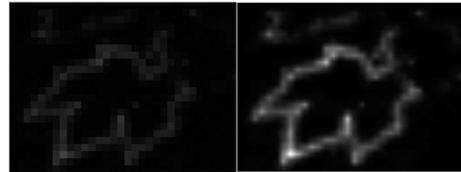


图 3 Local 特征下具有方块化效应的显著图和平滑后结果  
Fig. 3 Saliency map with blocking artifact and its smooth result with local feature

这里将 Local 特征下的显著图表示为  $S_{Local}$ .

### 1.3 Local 特征的结果与分析

在 Local 特征提取时, 设定的阈值  $T$  为 20070400 (即 4480 的平方),  $\alpha$  为 1.2, 这两个值的确定是在计算和分析文献 [16] 提供的数据库中用作显著性实验的 1000 幅图像<sup>[17]</sup> 中随机选定的 500 幅图像获得的统计经验值. 图 4 和图 5 分别为原始的输入图像和 Local 特征下的显著图, 通过观察发现  $S_{Local}$  有以下特点:

1) 针对较大的显著区域, Local 特征可以客观准确地刻画出显著区域的轮廓, 此时的显著物体是一种镂空的状态, 如图 5 (a) 中飞鸟的内部; 图 5 (b) 中花朵的内部; 图 5 (c) 中花朵的内部和图 5 (d) 中鹰的内部.

2) 针对较小的显著区域, Local 特征可以一定程度上反映出该区域的显著特性, 由于其采用了阈值处理并且特征提取的范围相对较小, 所以会在一定程度上丢失一些细节的特征, 如图 5 (b) 中花朵的中间部分.

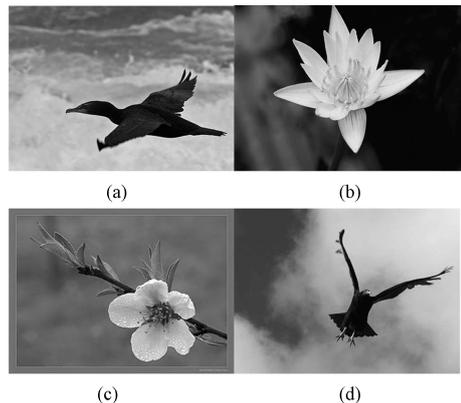


图 4 输入图像  
Fig. 4 Input images

造成以上不足的原因是度量 Local 特征时选取的区域较小, 不能整体刻画显著物体的总体特性, 通过放大 Local 特征的提取区域弥补以上缺点是不可行的, 相对一个显著物体的整体而言, Local 特征的表述是有限的. 为改善 Local 特征中丢失细节的问题, 图像的显著特性的检测需要在一个较大的范围

内进行, 采用 Regional 特征可以解决这一问题.

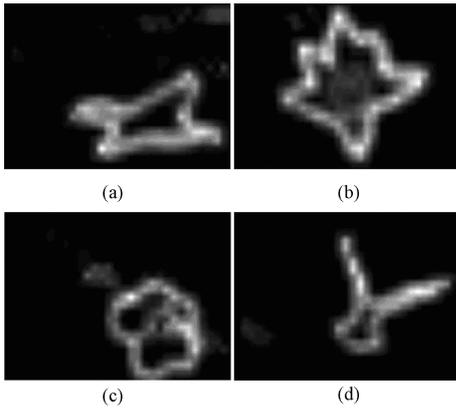


图 5 Local 特征下的显著图 ( $S\_Local$ )

Fig. 5 Saliency maps with local feature ( $S\_Local$ )

## 2 Regional 特征的提取

### 2.1 Regional 特征提取的意义与特征提取的区域

Local 特征的提取可以获得图像显著区域的局部特征. 由于该特征仅能反映小区域的显著特性, 不能够在整体上刻画物体的显著性, 所以物体的整体显著特性应该在一个相对较大的区域内提取, 即 Regional 特征的提取.

Regional 特征提取的范围是影响图像 Regional 特征的一个重要因素, 如果计算区域太大, 就会有与块  $B(i, j)$  显著性无关的区域参与到显著性的计算中, 不仅使计算变得复杂, 而且计算的精度也会受到影响. 如果计算区域太小就会丢失与块  $B(i, j)$  显著性相关的信息, 显著物体的整体性会被破坏, 从而影响计算结果准确性.

Regional 特征提取的区域与输入的原始图像大小是成正比的, 因为图像越大显著的物体越大. 故在显著性检测前, 应对图像大小进行调整, 在本文中图像是按照比例调整的, 最大边被调整到 384 像素. 图像大小被调整后, Regional 特征提取的区域就可以通过经验值来确定, 大量的对比实验结果表明, 在以块  $B(i, j)$  为中心的  $15 \times 15$  块的区域内计算 Regional 特征结果是最好的, 这里不仅考虑了图像中显著物体的大小, 同时考虑了计算的复杂程度. 故本文中确定 Regional 特征提取的范围为  $15 \times 15$  块大小, 即 120 像素  $\times$  120 像素的区域.

### 2.2 Regional 特征的提取步骤

不同于 Local 特征的提取, Regional 特征的提取是在 CIELab 颜色空间进行的, 因为 CIELab 颜色空间更符合人眼的感知特性<sup>[18-19]</sup>. 需要说明的是 Local 特征的提取没有在该颜色空间进行是因为其在显著性提取的时候要抑制不明显的变化, 这样

就导致 CIELab 颜色空间的感知优势不能体现.

1) 在提取 Regional 特征时, 颜色的变化仍然是确定块  $B(i, j)$  显著性的重要依据, 块间的颜色变化通过欧几里德距离进行度量. 在 CIELab 颜色空间内, 块  $B(i, j)$  与块  $B(p, q)$  的颜色变化表示为

$$Dis_{CIE}(B(i, j), B(p, q)) = \sqrt{(L_{i,j} - L_{p,q})^2 + (a_{i,j} - a_{p,q})^2 + (b_{i,j} - b_{p,q})^2} \quad (4)$$

2) 图像的高频分量包含了图像的细节部分, 在显著性检测中, 将高频分量的信息保留是重要的, 为了凸显这些高频分量, 将对每一个块进行离散余弦变换和逆变换,  $Table$  为 DCT 变换的量化表,  $Table^{-1}$  为反量化表, 这样两个块间的高频距离为

$$Dis_{HF}(B(i, j), B(p, q)) = Table^{-1} \times |Table \times B(i, j) - Table \times B(p, q)| \quad (5)$$

3) 两个子块间的位置距离是影响显著性的一个重要因素, 块间的位置距离越大, 显著的关联性越小, 反之则越大. 文献 [20] 在建模的时候也考虑了距离这一因素, 子块间的相异性正比于颜色变化, 反比于位置间的距离. 定义块  $B(i, j)$  和块  $B(p, q)$  间的位置距离为

$$Dis_{pos}(B(i, j), B(p, q)), \quad p \in i - 7, \dots, i + 7, q \in j - 7, \dots, j + 7 \quad (6)$$

分两种情况考虑  $Dis_{pos}(B(i, j), B(p, q))$ , 分别为中心位置距离和外围位置距离. 定义中心位置距离  $Dis_{pos\_cen}$  为

$$Dis_{pos\_cen}(B(i, j), B(p, q)) = \sqrt{(i - p)^2 + (j - q)^2}, \quad p \in i - 4, \dots, i + 4, q \in j - 4, \dots, j + 4 \quad (7)$$

中心位置距离计算的区域如图 6 所示, 图 6 中每个块的大小为 8 像素  $\times$  8 像素. 块  $C$  表示 Regional 特征提取区域的中心块, 提取的特征值表示该块的显著程度. 此外, 图 6 中还标出了 Local 特征提取的区域.

为了避免距离太大而减小外围块的影响力, 应主观地保留外围块的特征信息, 抑制位置距离的快速增长, 定义外围位置距离为

$$Dis_{pos\_out}(B(i, j), B(p, q)) = \max\{Dis_{pos\_cen}(B(i, j), B(p, q))\}, \quad p \in i - 7, \dots, i - 5, i + 5, \dots, i + 7, \quad q \in j - 7, \dots, j - 5, j + 5, \dots, j + 7 \quad (8)$$

外围位置计算区域见图 6.

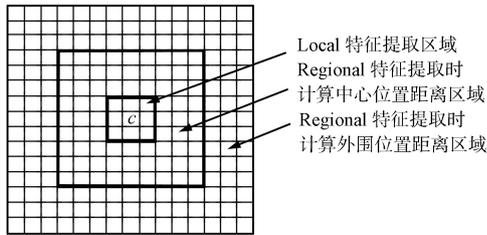


图6 Regional 特征提取时计算中心位置和外围位置的区域  
Fig. 6 Center and outer regions of regional feature

在计算上面的特征参数后, 块  $B(i, j)$  在 Regional 区域的显著性数值为

$$S\_Reg(i, j) = \frac{Dis_{CIE}(B(i, j), B(p, q)) \times Dis_{HF}(B(i, j), B(p, q))}{1 + c \times Dis_{pos}(B(i, j), B(p, q))},$$

$$p \in i - 7, \dots, i + 7, q \in j - 7, \dots, j + 7 \quad (9)$$

为了避免除 0 错误以及显著性数值太大的情况, 需要对位置距离进行调整. 分别对应于式 (9) 中分母加 1 和乘  $c$ .

按照上述方法, 依次计算图像内每一个块在 Regional 特征的显著值, 其中参数  $c$  取值为  $3^{[20]}$ . 对计算结果进行高斯平滑滤波操作. 以 Regional 特征为基础的显著图记为  $S\_Regional$ , 结果如图 7 所示.

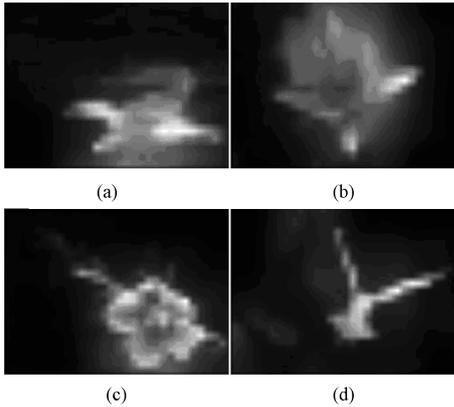


图7 Regional 特征下的显著图  
Fig. 7 Saliency map with regional feature

### 2.3 Regional 特征的特点与分析

相对而言, Regional 特征提取的范围较广, 在分析某一个子块显著特性时计算空间较大, 这就更能把握显著物体的整体特性, 相对 Local 特征下的显著图而言, 显著物体的镂空得到改善, 如图 7(a) 中飞鸟的内部, 图 7(c) 中花朵的内部以及图 7(d) 中鹰的身体内部. 但是, 较大的计算空间也会给显著图带来负面的影响, 使得显著物体中较大的图像变得杂乱粗糙. 如图 7(b) 中的花朵, 这种负面影响是不

可避免的, 是解决 Local 特征下显著图像镂空的必然产物.

## 3 多尺度下 Local 特征和 Regional 特征的综合

### 3.1 Local 特征和 Regional 特征的综合

Local 特征和 Regional 特征分别代表了不同空间下的图像显著性特征, Local 特征反映了子块的局部特征, 而 Regional 特征反映了子块的区域特征. 为了全局地体现出图像中物体的显著特性, 需要将两种特征进行综合. 本文通过加权的方法将两种特征综合, 从而得到图像原始尺度下的显著特性图. 由于 Regional 特征提取时考虑的范围较大, 更接近于全局, 故在综合 Local 特征和 Regional 特征时, Regional 特征的权值应大于 Local 特征的权值.

将综合了 Local 特征和 Regional 特征的显著图记为

$$S\_LocReg = w_1 \times S\_Local + w_2 \times S\_Regional,$$

$$w_1 + w_2 = 1, w_1 < w_2 \quad (10)$$

综合了 Local 特征和 Regional 特征的显著图如图 8 所示.

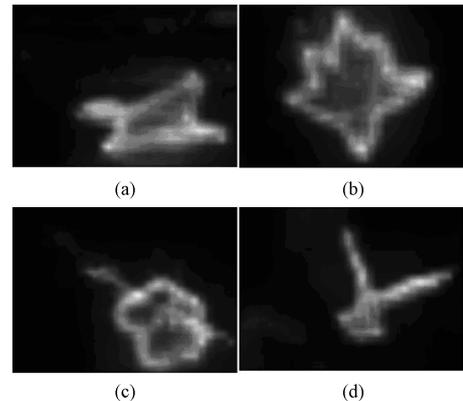


图8 综合 Local 特征和 Regional 特征的显著图  
Fig. 8 Saliency map combined with local feature and regional feature

通过比较 Local 和 Regional 特征下的显著图, 以及综合了两种特征的显著图可以发现: 综合两种特征的显著图可以清晰地描述出图像中物体的显著性. 由于 Local 特征的提取范围较小, 以及其对比部分受到阈值的调节, 这就使得 Local 特征下的显著图更能体现出一幅图像中最为显著的点和区域, 相对较小的显著物体而言, 其结果比较明显, 如图 5(d) 中的鹰, 但其不易检测出物体的总体特性, 如图 5(a) 中的飞鸟的镂空. 在综合了相对全局的 Regional 特征后, 会使得图像的镂空有所缓解, 易于显

著物体的整体描述. 单一尺度下综合了 Local 特征和 Regional 特征的显著图, 会丢失显著物体的一些细节特征, 而显著物体的镂空就是细节丢失的一种表现, 尽管 Regional 特征会缓解这一现象, 还需要在多尺度下进行图像显著性的提取.

### 3.2 多尺度下提取图像显著性特征

多尺度下提取图像中显著物体的依据是尺度空间思想, 其基本内容是在图像处理模型中引入描述图像尺度的参数, 通过图像尺度参数的变化, 提取不同尺度下图像的信息. 就观察者而言, 不同尺度下图像的信息是不同的, 分析和综合这些不同尺度下的图像信息, 可以更全面地体现出一幅图像的本质特征.

图像显著性特征的提取是在多尺度下进行的. 因为图像的尺度对显著区域的检测是有影响的. 当图像的尺度较小时, 图像中一些相对较大的显著区域容易被检测出来, 而较小的显著区域比较容易被忽略; 相反, 当图像的尺度较大时, 较小的显著区域容易被检测出来而较大的区域容易被忽略. 为了尽可能客观地检测图像中的显著物体, 显著性特征的提取将在多个尺度下进行.

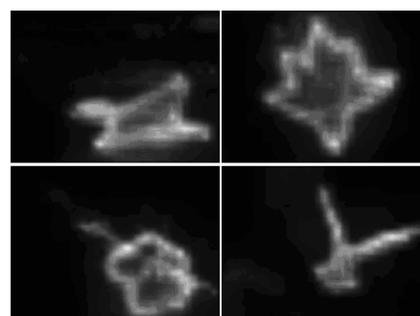
多尺度的提取有很多方法, 本文对图像进行下采样获得了 4 个尺度的图像. 其中, 尺度 1 为原始图像的大小, 尺度 2 的大小为原始图像的 80%, 尺度 3 的大小为原始图像的 60%, 尺度 4 的大小为原始图像的 40%. 需要说明的是最小尺度不能太小, 如果太小, 则插值放大后的图像会严重变形, 不符合提取特征的初始目的.

按照上述 Local 特征和 Regional 特征的提取方法, 分别计算 4 个尺度下的显著图. 由于 4 个尺度下显著图像的大小互不相等, 在将 4 个尺度下的特征进行综合前, 需要将尺度 2、尺度 3、尺度 4 下的显著图进行最近邻插值放大到原始图像的大小. 插值放大后, 4 个尺度下的显著图分别记为:  $Scale1$ ,  $Scale2$ ,  $Scale3$ ,  $Scale4$ . 4 个尺度下, 融合了 Local 特征和 Regional 特征的多尺度显著图如图 9 所示.

根据实验结果, 原始尺度下的显著图是最能反映出物体的显著特征的. 尺度越大, 插值还原到原图像大小时造成的显著图失真就越严重, 这容易导致显著图的整体性受到影响. 所以综合多尺度特征时, 应为它们赋予不同的权值  $\{w_3, w_4, w_5, w_6\}$ . 则综合 4 个尺度的显著图表示为

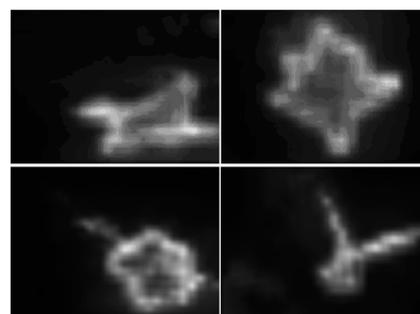
$$Saliency\_Object = w_3 \times Scale1 + w_4 \times Scale2 + w_5 \times Scale3 + w_6 \times Scale4 \quad (11)$$

在具体的实验过程中,  $w_3, w_4, w_5, w_6$  的取值分别为: 0.4, 0.3, 0.2, 0.1.



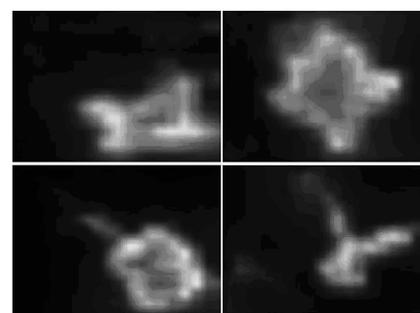
(a) 尺度 1 下显著图

(a) Saliency map of  $Scale1$



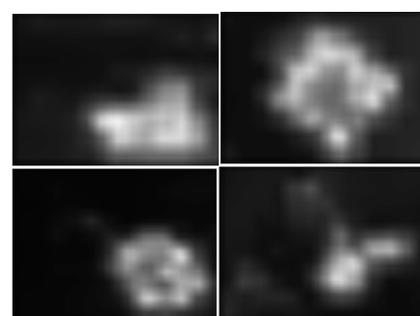
(b) 尺度 2 下显著图

(b) Saliency map of  $Scale2$



(c) 尺度 3 下显著图

(c) Saliency map of  $Scale3$



(d) 尺度 4 下显著图

(d) Saliency map of  $Scale4$

图 9 综合了 Local 特征和 Regional 特征的 4 个尺度下显著图

Fig. 9 Salient maps of combining local and regional features in four scales

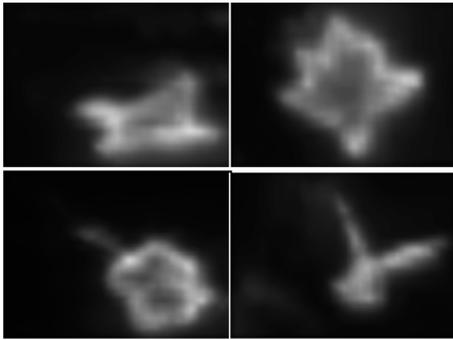


图 10 综合 4 个尺度的显著图

Fig. 10 Saliency maps of *Saliency\_Object*

图 10 表明: Local 特征和 Regional 特征下, 综合了 4 个尺度的显著图在表示显著物体整体特性时, 较单一尺度有明显的提高, 显著物体的区域更加明显.

## 4 显著物体的边缘提取

### 4.1 显著物体边缘提取的目的

显著性检测的最终目的是对显著性物体进行分割, 从而提取特征进行后续的相关操作. 在描述显著物体及其提取与分割的时候, 显著物体清晰的边缘是很重要的. 而此前的显著性方法中, 很少有提取这一特征的. 本节将通过计算 4 个颜色通道上的色度特征来检测图像中显著物体的边缘.

### 4.2 显著物体边缘提取的过程与结果

显著物体边缘是依据色度特征来进行计算, 分别以  $R$ 、 $G$ 、 $B$ 、 $Y$  这 4 个颜色通道的值来确定的: 红色的色度特征:  $R = r - 0.5(g + b)$ ; 绿色的色度特征:  $G = g - 0.5(r + b)$ ; 蓝色的色度特征:  $B = b - 0.5(r + g)$ ; 黄色的色度特征:  $Y = 0.5(r + g) + 0.5|r - g| - b$ ; 其中,  $r$ ,  $g$ ,  $b$  分别为图像中红、绿、蓝颜色分量值.

在进行显著物体边缘的提取前, 需对图像进行高斯滤波, 去掉图像中的噪声和部分颜色突变, 而滤波对显著物体边缘的剧烈突变的影响是有限的.

显著物体的边缘提取是以像素为单位的, 提取的范围是以某像素为中心的  $5 \times 5$  邻域. 分别计算某像素点与其邻域内像素的色度差值的绝对值之和, 当其中的一个或多个分量大于设定的阈值  $t_1$ , 或者 4 个分量的和大于阈值  $t_2$  时, 就认为该像素点处的颜色变化剧烈, 处在原始图像中物体的边缘位置.

在完成了以上操作后, 需要对图像进行第二次高斯滤波, 目的是使显著物体的边缘变宽并且平滑, 有利于突出显著图中的边缘部分, 从而使显著图的表现更加突出.

在实验过程中, 经过统计标准图像<sup>[16]</sup>中 500 幅随机图像获得经验阈值  $t_1$  为 250,  $t_2$  为 500, 利用这组数据对另外 500 幅图像进行测试, 均获得有效的边缘. 显著物体的边缘特征记为:  $S\_Edge$ , 结果如图 11 所示.

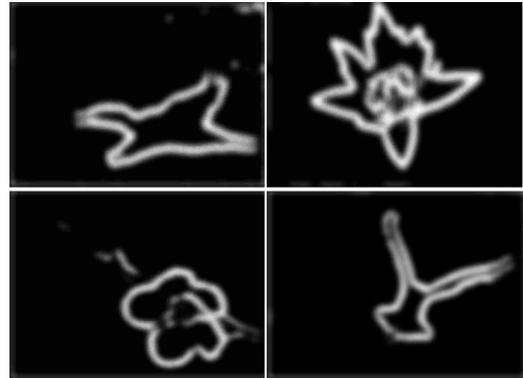


图 11 显著物体边缘图

Fig. 11 Edge maps of saliency objects

## 5 显著图的生成

在获得多尺度下 Local 和 Regional 特征的显著图和图像中显著物体的边缘信息后, 将这两种特征信息进行融合, 就可以得到图像中显著物体的特征, 从而完成图像的显著性检测. 特征显著图和显著物体的边缘是通过加权的方式进行综合的. 但在综合之前需要进行噪声抑制, 剔出非边缘区域. 虽然在显著性物体边缘提取的过程中, 进行了滤波消除噪声的处理, 但是该方法对一些非边缘的突变还是比较敏感的, 为了抑制这种非边缘区域造成的影响, 需要根据融合 Local 特征和 Regional 特征的显著图调节边缘灰度图的数值. 因为 Local 特征和 Regional 特征描述的是一个全局的显著信息, 通过分析其灰度直方图, 发现当其灰度值小于 40 的时候, 很少有区域在显著物体的边缘位置. 这样, 当融合 Local 特征和 Regional 特征的显著图的某个像素点的灰度值小于 40 时, 滤掉显著物体边缘图像对应位置处的值, 可以消除显著物体边缘图像中非边缘区域的干扰, 从而使得显著物体的边缘更加准确. 灰度值 40 的确定是通过统计文献 [16] 中随机 500 幅图像的 *Saliency\_Object* 对应的灰度直方图确定的, 是一个经验统计值.

在具体的实验过程中, 以像素为单位, 依次扫描融合 Local 特征和 Regional 特征的显著图, 当其灰度值小于 40 的时候, 则修改显著物体边缘图中对应点的灰度值为 0, 修改后的显著物体边缘图记为:  $Saliency\_Edge$ .

在完成了上述的调整后, 通过线性加权的方法将描述显著物体特征和显著物体边缘特征进行综合:

$$\begin{aligned} \text{Saliency\_Map} = \\ w_7 \times \text{Saliency\_Object} + w_8 \times \text{Saliency\_Edge} \end{aligned} \quad (12)$$

在描述图像显著特性的时候, 融合了 Local 特征和 Regional 特征的显著物体的整体性是至关重要的, 显著物体边缘图像的提取是为了更好地描述物体的整体, 是一种辅助性的提高, 因此,  $w_7 > w_8$ . 通过对比多组  $w_7$  和  $w_8$  的值, 当  $w_7 = 0.7143$ ,  $w_8 = 0.2857$  时, 得到的显著图结果是最优的, 结果如图 12 所示.

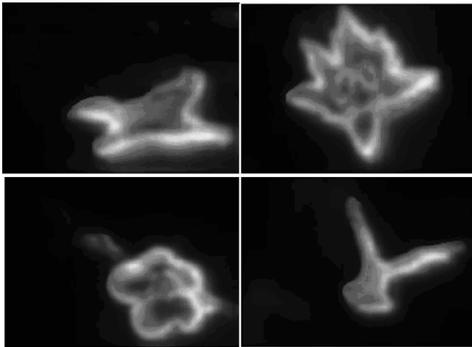


图 12 显著图

Fig. 12 Saliency map

## 6 实验结果与分析

实验的预处理包括图像的大小调整和图像的延拓处理. 图像的大小是按照其原始比例进行缩放的, 最大维的值为 384. 调整图像的大小基于两方面的考虑: 1) 计算量的大小, 图像太大会使计算变得复杂; 2) Regional 特征的提取是在  $15 \times 15$  的范围内进行的, 如果图像太大或太小, 图像的整体性就不能体现. 对图像进行延拓的目的是确保图像边缘信息的完整. 需要特别说明的是, 在实验过程中, 块间距离度量距离等计算需要归一化到  $[0, 1]$ , 各显著图的灰度值范围归一化到  $[0, 255]$ .

为了评价本文提出的显著目标检测方法, 这里比较了 5 种主流的显著性检测模型: Walther 模型<sup>[21]</sup>、Ma 模型<sup>[22]</sup>、Harel 模型<sup>[23]</sup>、Hou 模型<sup>[24]</sup>、Achanta 模型<sup>[25]</sup> 与本文模型在图像数据库<sup>[16]</sup> 中的表现, Achanta 从该数据库选出 1000 幅图像组成一个显著性检测的标准图像库, 图像库中每幅图像均有被大部分人接受的单一显著物体存在. 5 种主流模型的 Matlab 程序可以从相关网站上下载.

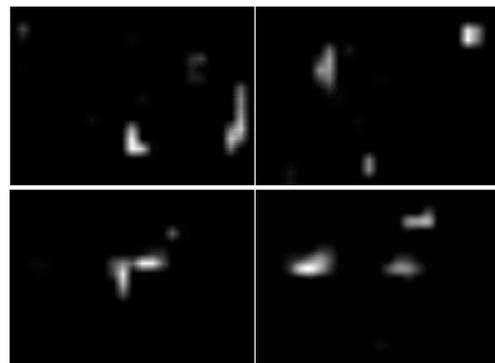
一些具有代表性的测试图像及实验结果见图 13(a) ~ 13(g). 从图 13 中可以看到本文方法产生

的显著图能够客观地反映出图像中显著物体的特性, 特别是图像的边缘信息比较完整与清晰, 显著物体的整体性也比较完整. 结果与人的主观感觉比较一致. 而 Walther 显著图只显示一些亮点; Ma、Harel 和 Hou 的显著图边缘模糊, 且对于复杂背景的图像不能有效地抑制背景噪声; 在 Achanta 的显著图中能够检测到显著物体的边缘, 但是显著物体与背景区域的对比不如本文模型得到的显著图明显, 因为本文模型充分利用了 Local、Regional 和边缘特征, 能够有效弥补上述方法的缺陷.



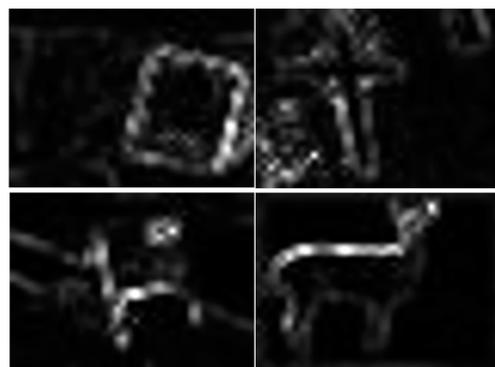
(a) 输入图像

(a) Input image



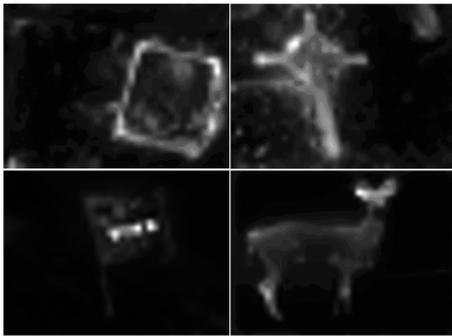
(b) 使用 Walther 方法

(b) Using Walther model



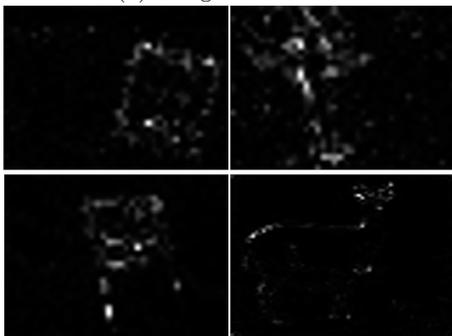
(c) 使用 Ma 模型

(c) Using Ma model



(d) 使用 Harel 模型

(d) Using Harel model



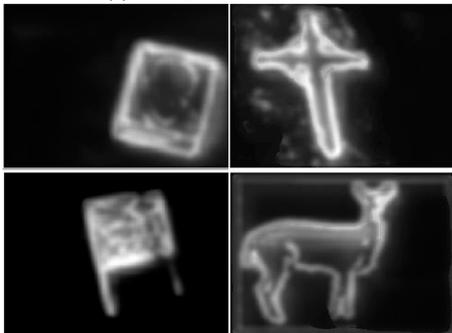
(e) 使用 Hou 模型

(e) Using Hou model



(f) 使用 Achanta 模型

(f) Using Achanta model



(g) 本文方法

(g) Method in this paper

图 13 6 种显著性检测方法的显著性图  
Fig. 13 Saliency map of the six saliency detection methods

除了上述的主观评价外, 本文还计算了几种不同显著性检测方法的 ROC (Receiver operating characteristic) 曲线. ROC 曲线是一种分析判断和决策性能的定量方法, 可以动态地、客观地评价一种分类方法的优劣. 客观评价显著性结果的本质是一个二分问题, 即算法计算的结果是否与统一接受的主观认定相一致. 一个二分问题会出现 4 种情况: 实际是正类并且被预测为正类, 即真正类 (True positive, TP); 实际是负类被预测成正类, 即假正类 (False positive, FP); 实际是负类被预测成负类, 即真负类 (True negative, TN); 实际是正类被预测成负类, 即假负类 (False negative, FN). 则有:

$$\text{True positive rate} = \frac{TP}{TP + FN}$$

$$\text{False positive rate} = \frac{FP}{FP + TN}$$

ROC 曲线的横坐标为 False positive rate, 代表错误的命中率; 纵坐标为 True positive rate, 代表正确的肯定率. 取灰度值为阈值, 显著图的 False positive rate 和 True positive rate 是正相关的, 我们期望的结果是: 较小的 False positive rate 对应较大的 True positive rate.

本文中绘制 ROC 曲线的数据库图像由文献 [24] 提供, 该数据库共包含 62 幅自然场景的图像, 图像中显著物体的大小占整幅图像的大小各不相同, 且图像背景比较复杂. 图 14 中的 ROC 曲线表明本文提出的方法在较小的 False positive rate 时, 有较大的 True positive rate, 其结果优于文献 [6, 17, 21–23] 中提出的方法.

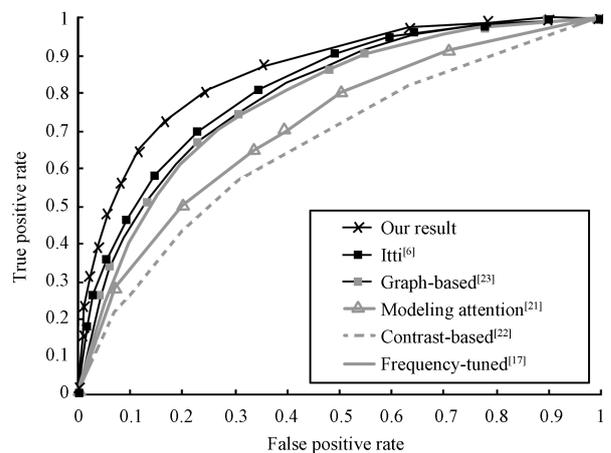


图 14 ROC 曲线

Fig. 14 ROC curves

本算法运行的软件环境为 Matlab 7.4.0 (R2007a), 系统环境为 Windows XP SP2, 内存为 2 GB, CPU 主频为 2.8 GHz, 运行计算一幅图像显著性的平均耗时为 30.889 秒, 单纯从时间的角度看,

本文的方法并不占优势, 结合显著性检测的效果, 本文提出方法的表现是优秀的.

## 7 结论及展望

本文主要从两个区域 Local 和 Regional 提取了图像子块的显著特征, 综合两种特征可以更加准确地刻画出整幅图像中物体的显著性.

Local 特征的提取在一个较小的区域内进行, 能够客观地反映图像的局部特征, 通过抑制与强化不同的相关特征, 获得图像中显著物体的较突出特征. 该区域提取出的特征是图像中最为显著的区域. 在阈值的控制下, 显著区域的背景比较单一, 这有利于显著图的表示.

Regional 特征的提取在一个较大的区域进行, 能够客观地反映图像的全局特征, 是模拟显著物体整体表示的提取过程.

图像中显著目标的特征提取是在多个尺度下进行的, 这样就可以把大小不同的显著性物体检测出来, 同时多尺度还能够获得在 Local 特征提取时忽略的细节, 使得显著区域的确定会更加稳定准确. Local 特征和 Regional 特征各有侧重, 同时又能互补, 主要体现在 Local 特征下的显著图可以提供较为单一的图像背景, Regional 特征下的显著图不只表现物体的边缘轮廓, 而且把物体作为一个整体来表现. 将两种特征综合, 可以更好地确定物体的显著性, 客观地反映出图像中显著物体.

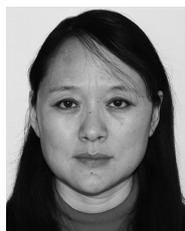
显著物体的边缘轮廓对描述图像显著性具有一定的作用, 它不仅可以体现图像中显著物体的全局, 同时为显著性提取的后续操作——图像分割提供了便利. 而这一点在同类显著性检测中少有提及.

本文提出的显著性检测方法主要是基于颜色对比来获取图像显著特性的. 接下来的工作将把图像的纹理特征用于显著性检测, 以使显著性检测方法更加完善.

## References

- Jian M W, Dong J Y, Ma J. Image retrieval using wavelet-based salient regions. *The Imaging Science Journal*, 2011, **59**(4): 219–231
- Hua Shun-Gang, Chen Guo-Peng, Shi Shu-Sheng. Image resizing algorithm based on similarity criterion. *Computer Engineering*, 2012, **38**(4): 191–193  
(华顺刚, 陈国鹏, 时树胜. 基于相似性判据的图像尺寸调整算法. 计算机工程, 2012, **38**(4): 191–193)
- Gupta R, Chaudhury S. A scheme for attentional video compression. *Pattern Recognition and Machine Intelligence*, 2011, **6744**: 458–465
- Kim W, Kim C. A novel image importance model for content-aware image resizing. In: Proceedings of the 18th IEEE International Conference on Image. Brussels, Belgium: IEEE, 2011. 2469–2472
- Shen Lan-Sun, Zhang Jing, Li Xiao-Guang. *Image Retrieval and Compressed Domain Processing*. Beijing: Posts and Telecom Press, 2008. 102–103  
(沈兰荪, 张菁, 李晓光. 图像检索与压缩域处理技术的研究. 北京: 人民邮电出版社, 2008. 102–103)
- Itti L, Koch C, Niebur E. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1998, **20**(11): 1254–1259
- Zhang X L, Li Z P, Zhou T G, Fang F. Neural activities in V1 create a bottom-up saliency map. *Neuron*, 2012, **73**(1): 183–192
- Sun J Y, Chen R F, He J. A modified GBVS method with entropy for extracting bottom-up attention information. *Lecture Notes in Electrical Engineering*, 2012, **121**: 765–770
- Hansen L K, Karadogan S, Marchegiani L. What to measure next to improve decision making? On top-down task driven feature saliency. In: 2011 IEEE Symposium on Computational Intelligence, Cognitive Algorithms, Mind, and Brain. Paris, France: IEEE, 2011. 86–87
- Baluch F, Itti L. Mechanisms of top-down attention. *Trends in Neurosciences*, 2011, **34**(4): 210–224
- Itti L, Koch C. A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, 2000, **40**(6): 1489–1506
- Itti L, Koch C. Computational modelling of visual attention. *Nature Reviews Neuroscience*, 2001, **2**(3): 194–230
- Jiao W, Peng Q M, Lv W X, Huang W J. Multiscale detection of salient regions. In: Proceedings of the 4th IEEE Conference on Industrial Electronics and Applications. Xi'an, China: IEEE, 2009. 2408–2411
- Zhang J, Sun J D, Liu J, Yang C X, Yan H. Visual attention model based on multi-scale local contrast of low-level features. In: Proceedings of the 10th IEEE International Conference on Signal Processing (ICSP). Beijing, China, 2010. 902–905
- Zhang Qiao-Rong, Gu Guo-Chang, Liu Hai-Bo, Xiao Hui-Min. Salient region detection using multi-scale analysis in the frequency domain. *Journal of Harbin Engineering University*, 2010, **31**(3): 361–365  
(张巧荣, 顾国昌, 刘海波, 肖会敏. 利用多尺度频域分析的图像显著区域检测. 哈尔滨工程大学学报, 2010, **31**(3): 361–365)
- Liu T, Yuan Z J, Sun J, Wang J D, Zheng N N, Tang X O, Shun H Y. Learning to detect a salient object. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, **33**(2): 353–367
- Achanta R, Hemami S, Estrada F, Süsstrunk S. Frequency-tuned salient region detection. In: Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Miami Beach, Florida: IEEE, 2009. 1579–1604

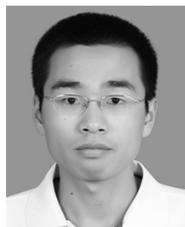
- 18 Yao Jun-Cai. Color image compression technology based on chromatic aberration. *Chinese Journal of Liquid Crystal and Displays*, 2012, **27**(3): 391–395  
(姚军财. 基于颜色色差的彩色图像压缩技术研究. 液晶与显示, 2012, **27**(3): 391–395)
- 19 Wang Xiang-Yang, Yang Hong-Ying, Zheng Hong-Liang, Wu Jun-Feng. A color block-histogram image retrieval based on visual weight. *Acta Automatica Sinica*, 2010, **36**(10): 1489–1492  
(王向阳, 杨红颖, 郑宏亮, 吴俊峰. 基于视觉权值的分块颜色直方图图像检索算法. 自动化学报, 2010, **36**(10): 1489–1492)
- 20 Goferman S, Zelnik-Manor L, Tal A. Context-aware saliency detection. In: *Proceedings of the 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. San Francisco, CA: IEEE, 2010. 2376–2383
- 21 Walther D, Koch C. Modeling attention to salient proto-objects. *Neural Networks*, 2006, **19**(9): 1395–1407
- 22 Ma Y F, Zhang H J. Contrast-based image attention analysis by using fuzzy growing. In: *Proceedings of the 11th ACM International Conference on Multimedia*. New York: ACM, 2003. 374–381
- 23 Harel J, Koch C, Perona P. Graph-based visual saliency. *Advances in Neural Information Processing Systems*, 2007, **19**: 545–552
- 24 Hou X D, Zhang L Q. Saliency detection: a spectral residual approach. In: *Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition*. Minneapolis, MN: IEEE, 2007. 1–8
- 25 Achanta R, Estrada F, Wils P, Süsstrunk S. Saliency region detection and segmentation. In: *Proceedings of the 6th International Conference on Computer Vision Systems*. Berlin, Heidelberg: Springer-Verlag, 2008, **5008**: 66–75



**郭迎春** 河北工业大学计算机科学与软件学院副教授. 2006 年获得天津大学信号与信息处理博士学位. 主要研究方向为数字图像处理与视频压缩技术. 本文通信作者.

E-mail: springguo2001@gmail.com

(**GUO Ying-Chun** Associate professor at Hebei University of Technology. She received her Ph.D. degree from Tianjin University in 2006. Her research interest covers image processing and video coding. Corresponding author of this paper.)



**袁浩杰** 河北工业大学计算机科学与软件学院硕士研究生. 主要研究方向为数字图像处理, 图像检测.

E-mail: yuanhaojie9@163.com

(**YUAN Hao-Jie** Master student at the School of Computer Science and Engineering, Hebei University of Technology. His research interest covers digital image processing and image detection.)



**吴鹏** 河北工业大学计算机科学与软件学院硕士研究生. 主要研究方向为数字图像处理, 图像检测.

E-mail: wpsnake@163.com

(**WU Peng** Master student at the School of Computer Science and Engineering, Hebei University of Technology. His research interest covers digital image processing and image detection.)