

基于中低层结合的图像感兴趣区域标注

周 洁 王士同

(江南大学数字媒体学院, 无锡, 214122)

摘 要: 图像感兴趣区域标注是近年来图像处理领域的重要研究课题之一。利用中低层次信息相结合的方式确保中低层信息相互补充, 可以得到可靠结果。中层次显著图由改进的 Harris 角点形成的凸包区域与 GBR 超像素聚类结果相结合得到, 低层次信息由不同权重的高斯差分滤波器对图像进行处理得到。最后通过加权融合两个层次显著图得到最终结果。本文利用微软亚洲研究院公开数据库对实验结果进行验证, 并选取其他前沿方法进行对比, 从主观和客观角度对实验结果进行判断, 本文方法结果较好, 可准确定位显著度区域并高亮表示, 同时可有效消除背景噪声。

关键词: 感兴趣区域; 显著图; GBR; 改进 FT

中图分类号: TP181 **文献标志码:** A

Region of Interest Marked by Low and Middle Levels

Zhou Jie, Wang Shitong

(School of Digital Media, Jiangnan University, Wuxi, 214122, China)

Abstract: Image marking of the region of interest is an important research topic in image processing in recent years. The combination of low and middle levels can ensure the result has both of their information. First, we get the middle-level coarse saliency map by using the boosting Harris to make a convex hull and superpixels clustered by GBR. And then we weight different Gaussian filters to get the low-level saliency map. The final saliency map is combined by middle-level saliency map and low-level saliency map. Experiments on the public databases coming from Microsoft Research Asia show that the proposed algorithm performs better than state-of-art algorithms not only on subjective evaluation but also on objective evaluation, and it is effective at the eliminate of background noise and outstanding at making the saliency regions high light.

Key words: region of interest; saliency map; GBR; improved FT

引 言

对于人类来说, 视觉是最重要的接受信息的途径之一, 人类对于视觉所接受的信息在头脑中的处理方式便是形成图像。伴随着科学技术的不断发展进步, 信息也逐步进入了高速发展的大爆炸时期。人类每秒接受的信息流为 $(10^8 \sim 10^9)^{[1-2]}$, 十分丰富且庞大。因此, 如何使机器能够像人类对于图像数据

的处理一样快速和精准便成为了现今重要的研究课题之一。研究发现在人类的视觉识别系统中能够高效准确地定位到复杂场景中的一个或几个区域,而忽略其他不重要的较多重复的背景信息。这个看似人人都会的简单的过程,实际并不简单,甚至较为复杂。人类视觉的识别系统通过对图像进行特征分析来锁定图像的某个或某几个区域,并将该区域挪动到视网膜的正中央使得大脑对图像的观察更加清晰,以便对其进行进一步的分析观察处理。人类的这种选择性注意机制在感兴趣区域的选定上十分有效,也是现今感兴趣区域标注的关键技术之一。因此从人类的视觉研究入手,能够方便人类对于感兴趣区域算法的不断改进。尽管如此,迄今为止,人类仍然没有开发出能和人类视觉的识别系统相抗衡的算法。因此,又提出了基于数学计算的模型来实现对人类感兴趣区域的标注,不断对图像感兴趣区域标注的算法进行补充和改进。

在图像处理中,通过对图像感兴趣区域的锁定,能够更好地把握图像中有意义的信息,去除图像冗余,使后续的处理工作更加简单,能够在图像处理领域加快后续处理的速率,提高算法的效率,节省时间。因此,图像感兴趣区域标注在图像处理领域意义非凡。感兴趣区域的标注能够对图像进行预处理,以便在后续的图像分类、图像压缩、定位图像中目标物体、图像匹配、视频跟踪、图像有效编码^[3]等领域中对图像进行处理和应用。图像感兴趣区域提取能够帮助更好地理解、处理和分析图像,是一个结合人类视觉系统、数学方法、特征提取、图像处理的综合研究过程,对于图像研究和分析的应用意义重大。

近年来,国内外越来越多的学者和机构在图像感兴趣区域标注领域投入了更多的研究。1998年,Itti等人提出了一种基于低层的视觉特性的计算机模型^[4]。该模型利用自底而上的模型对图像进行检测,从局部的特征着手,通过中间-周边(Center-surround)算子计算图像的低层信息,如有灰度、颜色、方位等特征,并将其显著度的度量结果合成为最终的显著图。Itti方法通过不同尺度上的显著度计算,需要对原图进行缩放和平滑,导致特征显著图清晰度降低,所形成的显著度结果图轮廓不准确,结果图成块。2006年提出的基于图的视觉显著度方法(Graph-based visual saliency, GB)^[5]在Itti方法上进行了改进,在通过视觉系统进行判断的基础上加入了纯数学计算的方法。GB算法也是从底层细节信息出发,保持了Itti模型中的特征和提取方式。但在特征的显著度计算和特征显著图的结合方式上不同。融入了马尔科夫链,但对于马尔科夫链的平衡态的计算十分复杂,降低了算法的整体效率。在结果上比Itti方法稍好,但显著区域边缘不清晰,仍然成块。2008年,Achanta等人针对Itti方法中对于图像不同尺度的处理进行改进,不再对图像进行缩放处理,而是通过改变计算特征提取单元的大小来实现原本Itti算法中不同尺度的高斯金字塔处理方式,即AC算法^[6]。AC算法也通过对图像的底层信息亮度和颜色变化强烈的部分来模拟人眼对场景的感知,以此来确定图像的感兴趣区域。通过多尺度提取单元的运算使得算法的精确度和容错性较高。但是,显著度结果图显著度区域不够突出,背景噪声较多。光谱残差法(Spectral residual approach, SR)^[7]是从频域的角度出发对图像的显著度进行标注。Barlow提出去除输入信息中的冗余信息就是高效编码^[8]。在人类的视觉系统中,对图像中频繁出现的特征不敏感,对偏离正态的特征较为敏感,即高频抑制原理^[9]。利用傅里叶变换和反变换来对图像进行处理。SR算法效率很高,但是显著图效果一般,只能对显著性区域进行粗略定位,对显著物体标注不明显。Xie等人所提出的贝叶斯算法^[10]通过对图像的灰度、对比度、颜色等方面的低层次信息和包含超像素及聚类结果的中层次信息进行结合,得到的显著图有明确边缘,内部高亮。但效率较低,背景图中噪声较多。频率调谐显著度区域检测算法(Frequency-tuned saliency region detection, FT)^[11]是利用高斯差分滤波器对图像进行处理来检测图像感兴趣区域的。优点是显著度区域具有清晰的边缘,但对部分显著度区域有中心抑制,显著度区域亮度较低,结果较差。文献[12]提出了一种结合区域和边界信息的图像显著度检测,该方法提出了对不同特征下的显著度信息统一度量标准,利用Beltrami滤波器对图像的边界信息进行检测利用,再与图像区域结合,较之前的Itti, AC, GB方法有很大提高,但所得到的结果仍旧成块,没有明确边缘。

在现阶段的研究中,已有算法的不足主要是显著性区域定位不够精准,无清晰边缘,显著度区域不够高亮及背景中包含噪声。本文中采用超像素聚类与凸包相结合的中层次图像信息和图像不同频率不同权重差分滤波的低层次图像信息相结合的方法来对图像的显著度进行标注。中层次信息为显著性区域提供了精准的定位和粗略图,低层次信息为显著区域补充边缘及内部信息。利用微软亚洲研究院提供的公开数据库进行人工标注的 1 000 张图对本文的方法进行验证,实验结果表明本文方法能够有效定位显著度区域,对背景中的噪声有抑制作用,与其他方法实验结果相比具有更高亮的显著度区域,且边缘清晰。

1 结合中、低层信息的显著区域标注

近年来,图像的目标区域检测不再局限于单一层面的检测,而是采用高中低三个层次进行结合的方式来得到更加精确的结果。高层次图像信息主要是从整体出发,利用先验知识和语义信息将图像映射到抽象的语义信息中。中层次信息是指对图像结构信息的定位处理,以及对图像的频域颜色等进行分析。低层次的信息是从图像本身出发,对算法进行底层设计和细节实现。本文即采用中层次信息与低层次信息相结合的方式来对显著度区域标注。

1.1 中层次显著信息标注

贝叶斯算法从中层次和低层次出发,对显著度区域进行标注。其中在高层次采用超像素聚类和凸包相结合的方法得到先验概率图,再通过贝叶斯公式求出低层次的信息,对于显著图的定位较好,且准确率较高。但其超像素聚类(拉普拉斯子空间聚类)无法自主确定聚类个数,且时间复杂度较高、迭代寻优时间较长、算法效率较低。同时低层次的图像不够清晰,有些图像显著度区域亮度较低且包含噪声。故本文算法的中层次采用贝叶斯算法获取先验概率图的方法,通过基于图的松弛聚类算法(Graph-based relaxed clustering,GBR)^[13]对超像素聚类方法进行替换,保证算法的运算效率及自适应性,下面对其进行具体阐述。

首先在图像中使用彩色增强的 Harris 角点检测算子^[14]对图像的彩色角点进行检测。彩色增强的 Harris 角点检测算子在传统的 Harris 角点检测算子^[15]中加入了颜色特征,能够更好地排除背景中的干扰因素,对显著区域的定位效果更好。彩色增强的 Harris 角点检测算法通过对输入图像进行颜色分析,计算得到彩色增强矩阵 M ,再将图像通过 M 矩阵进行颜色增强,计算增强后图像的 Harris 角点能量函数,再选取该函数形成的能量图中能量值最大的点集,并剔除边界处的点,得到最终显著点。图 1 中展示了彩色增强的 Harris 角点检测算子结果与传统的 Harris 角点检测算子的结果比较图。

将刚刚通过改进 Harris 角点检测的角点图用凸包进行包围。为了减少边缘对于图像凸包形成的影响,应去除图像边界处的角点。图 2 中显示了不同 Harris 角点形成凸包区域与真值图的对比情况。

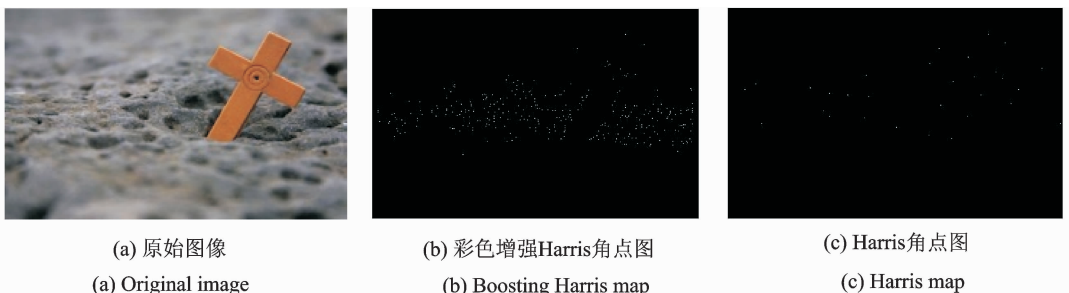


图 1 角点检测图像

Fig. 1 Comparison of different corner detection

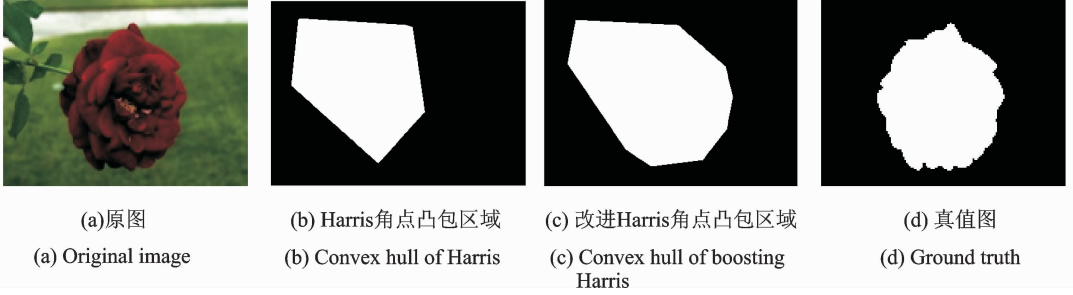


图 2 凸包区域对比图

Fig. 2 Comparison of different convex hull

真值图,即由 $\{0,1\}$ 组成的显著度区域的人工标注图像。由图 2 可以看出改进的 Harris 角点所形成的凸包区域对于图像显著区域的定位更为精准。

1.1.1 超像素及超像素聚类

超像素能够从整体上对图像的像素点进行表达。将一系列距离相近且颜色、亮度等特征相似的像素点合并为一个超像素。超像素既能保留图像的边界,又能使图像的后续操作更加简便。简单线性迭代聚类算法(Simple linear iterative clustering, SLIC)^[16]能够很好地将图像聚类成超像素。SLIC 算法能够在边缘贴合度和紧密度上取得较好的平衡且计算效率较高,数量可控,是目前综合性能较好的一种超像素方法。因此本文利用 SLIC 算法将图像像素聚类形成超像素。图 3 显示了图像的原图与 SLIC 算法的超像素图,其中 SLIC 聚类个数设定为 200。

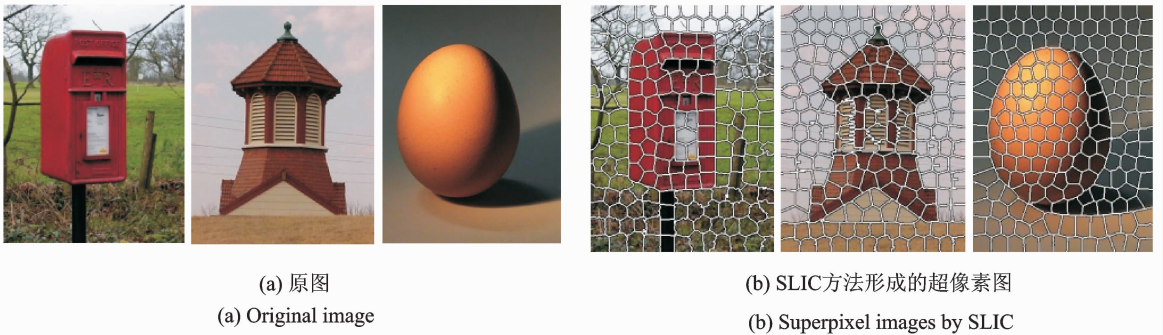


图 3 SLIC 聚类超像素图

Fig. 3 Pictures of SLIC superpixel algorithm

然后,将处理过的超像素图再利用 GBR 方法进行聚类。GBR 算法是基于谱聚类(Spectral clustering)算法的一项最新研究成果。不同于传统谱聚类需要求解特征值及特征向量的复杂方式,该方法无需对特征值、特征向量进行求解,且其可以自主确定聚类个数,是一种简单、快速、自适应的聚类方法。相比 Xie 等人^[10]所采用的拉普拉斯子空间聚类方法,该方法可以一次划分出所有的类别,并且时间复杂度较低,效率较高。

GBR 算法中, Lee 等人^[17]在 Normalized cuts 的求解方法中采用了半正定松弛的方法。

$$\min \mathbf{y}^T \mathbf{L} \mathbf{y} \quad \text{s. t.} \quad \|\mathbf{y}\| = 1, \quad \mathbf{e}^T \mathbf{y} = 0 \quad (1)$$

式中: $\mathbf{e}^T = (1, 1, \dots, 1)$, \mathbf{L} 为拉普拉斯矩阵。计算公式如下

$$D(i, i) = \sum_1^n W(i, j) \quad (2)$$

$$W(i, j) = \exp\left(\frac{-\delta(v_i - v_j)^2}{2\sigma^2}\right) \quad \sigma > 0 \quad (3)$$

$$L = D - W \quad (4)$$

式中: W 为图 $G=(V, E)$ 各个顶点的相似度矩阵; D 为对角矩阵,称为 W 的度矩阵。

由于式(1)主要解决两类问题,为了将其应用到多分类上, Lee 对式(1)限制条件进行修改,提出了 GBR 算法。

$$\min \mathbf{y}^T L \mathbf{y} \quad \text{s. t.} \quad A \mathbf{y} = \varphi, \varphi \neq 0 \quad (5)$$

式中 $A=e^T$ 。式(5)所求解出的最优解 \mathbf{y}^* 可直接表示出聚类结果,可直观地通过 \mathbf{y}^* 来观察所有样本的分类情况及每个样本的类别。Lee 通过引入拉格朗日乘子 λ ,并分别对 \mathbf{y} 和 λ 求偏导后得到结果 \mathbf{y}^* 。

$$\mathbf{y}^* = L^{-1} A^T (A L^{-1} A^T)^{-1} \varphi \quad (6)$$

1.1.2 中层次显著图

结合凸包区域和超像素聚类结果,由式(7)得出中层次的概率图。

$$I_m = \frac{|n \cap h|}{|n|} \quad (7)$$

式中: $|n|$ 表示聚类一个类内的像素数量, $|n \cap h|$ 表示既在凸包区域内又在该类内的像素数量。两者的比值确定了该像素的显著度概率,即中层次的显著图。数据库中部分结果如图 4 所示。

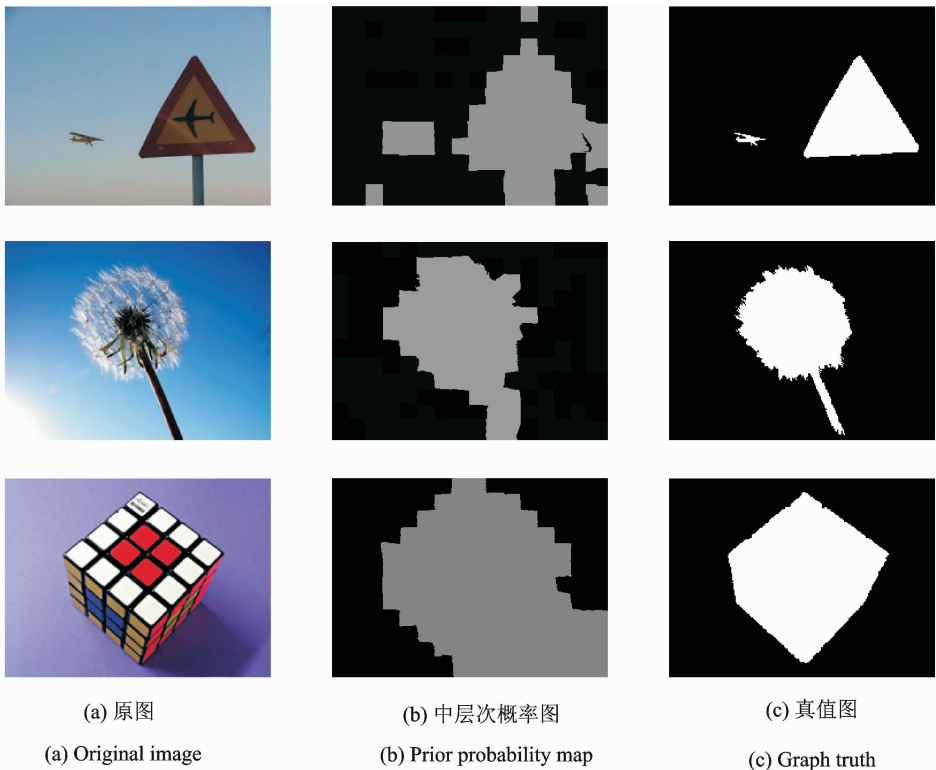


图 4 中层次显著图

Fig. 4 Saliency map of middle level

1.2 低层次显著信息标注

1.2.1 FT 算法

由于中层次的显著图仅能够定位显著度区域。内部较为粗略,缺少底层的细节及边缘信息,故利用低层次的信息对中层次显著图进行补充,更好地对图像的显著区域进行标注。

FT 算法是一种利用多组不同尺度的高斯差分滤波器对图像进行处理得到显著度图的方法。FT 算法在频率域内对图像进行处理。图像在频域中表示分为高频和低频,高频代表图像的细节信息;低频则主要反映图像的整体概况。通过 FT 算法控制图像的频率输出所得到显著度图像能够很好地体现图像显著区域的轮廓及内部细节。假设图像的低频截止频率为 ω_c , 高频的截止频率为 ω_{hc} 。

FT 算法主要采用高斯差分(Difference of Gaussian, DoG)滤波器对图像进行高斯差分滤波。DoG 的公式如下

$$\text{DoG}(x, y) = G(x, y, \sigma_1^2) - G(x, y, \sigma_2^2) \quad (8)$$

$$G(x, y, \sigma_i^2) = \frac{1}{2\pi\sigma_i^2} \exp\left(-\frac{(x^2 + y^2)}{2\sigma_i^2}\right) \quad i = 1, 2 \quad (9)$$

式中 σ_1, σ_2 是高斯标准差,且 $\sigma_1 : \sigma_2 = 1.6$ 。

由于单个 DoG 滤波器的带宽为标准差 $\rho = \sigma_1 : \sigma_2$ 较窄,无法满足实际需求。因此利用多个窄带的 DoG 滤波器结合的方式来适应实际需求。

$$F_N(x, y) = \sum_{n=0}^{N-1} G(x, y, \rho^{n+1}) - G(x, y, \rho^n) = G(x, y, \rho^N) - G(x, y, \rho) \quad (10)$$

式中 N 表示 DoG 滤波器数量。当 $N \geq 0$ 时,式(10)可以简化为两个高斯函数差 $K = \rho^N$ (因为除第一项和最后一项,其他项可以相互抵消)。这样就可以选择一个尽可能大的 K 值来代替多个 DoG 滤波器组合的结果。

由于 $\sigma_1 > \sigma_2$, 因此最高频率取决于 σ_1 , 最低频率取决于 σ_2 。因此可以将 σ_1 趋于无穷大。即利用一个较小的高斯核对图像进行滤波来达到多个 DoG 滤波的效果。

经过滤波的图像 I_f , 转化到 CIELab 空间,得到 3 个分量 l, a, b 。得到显著图像 I_{Sal} 为

$$I_{\text{Sal}} = \| I_u - I_f \| \quad (11)$$

式中: I_u 是图像在 CIELab 空间的平均值, I_f 是图像经过滤波后消除了一定的高频噪声的图像, $\| \cdot \|$ 表示 L_2 范式,用欧式距离来计算。

1.2.2 改进的 FT 算法

传统的 FT 算法有诸多不足。在传统的 FT 算法中,虽然对于高斯差分滤波器采用了不同的频率进行滤波,但对于不同尺度的 DoG 滤波器的权重相同,使得高频信息和低频信息的比重相同。但在人的视觉系统中,低频的整体信息比高频信息更容易被观测,且高频信息中可能包含噪声,使得显著图检测依赖于物体的情况,结果不佳。因此应该在 FT 算法中多突出频率较低的部分,使得图像中显著物体的区域高亮。

由于低频信息更能突出图像的整体部分。故在改进的 FT 方法中,采取增加低频信息权重的方法来抑制高频信息,得到更好的图像显著度结果。于是,在式(10)的基础上进行改进,使得 DoG 滤波器中频率越低的部分权重越大,如式(12)所示

$$F_N(x, y) = \sum_{n=0}^{N-1} W_n [G(x, y, \rho^{n+1}) - G(x, y, \rho^n)] \quad (12)$$

式中 W_n 是权重,且为了使低频信息权重增高,则 $W_0 < W_1 < W_2 < \dots < W_{N-1}$ 。在本文中为了简化处理,将 W_n 设置为 $(n+1)$,使得在频率逐渐降低情况下权重呈线性增长,可将式(12)推导成式(13)的形式。

$$F_N(x, y) = \sum_{n=0}^{N-1} (n+1) [G(x, y, \rho^{n+1}) - G(x, y, \rho^n)] \quad (13)$$

通过 FT 算法中的式(12)可以类似地推导出改进 FT 算法求解显著度图像的式(14)。

$$\mathbf{I}_{\text{sal}}^W = \sum_{n=0}^{N-1} \|\mathbf{I}_u - \mathbf{I}_{\text{ohc}}^n\| \quad (14)$$

式中: \mathbf{I}_u 为图像在 CIELab 空间各个变量的平均值的特征向量, $\mathbf{I}_{\text{ohc}}^n$ 表示图像经过相应高斯滤波后的图像的特征向量。在本文中,为了简化计算并保证较好的检测效果,分别选取 3×3 , 7×7 , 13×13 , 23×23 , 39×39 五个高斯滤波模板对图像进行处理。

使用 FT 算法和改进 FT 算法得到的显著图比较结果如图 5 所示。通过对比可以发现由于改进 FT 加重了低频信息的权重,使图像清晰度有所下降,但是影响较小,边缘依旧清晰,且能够更多忽略背景中的细节和噪声,将显著度区域高亮表示,具有更好的检测效果。



图 5 FT 与改进 FT 对比图

Fig. 5 Comparison of FT and the improved FT

1.3 中低层次显著度图结合

在不同层次下提取出的不同显著度图代表了不同层次下的信息,中层定位显著度区域,低层补充细节和边缘,通过两者的结合能够更好地对显著度区域进行检测。那么可以用式(15)将两个层次的显著度图像进行融合。

$$S = \alpha \times I_m + (1 - \alpha) \times I_{sal}^W \quad (15)$$

式中参数 α 代表了中层次显著度图在结合图中的权重。对于参数 α 的选择,阅读相关文献[18]可知在 α 取值为 0.5 时融合效果最好。故在本文中 $\alpha = 0.5$ 。

2 实验结果与分析

本文采用微软亚洲研究院提供的显著性检测数据集 MSRA 作为实验数据集,该数据集包含 1 000 张由志愿者人工标注了显著性区域的自然图像。本文使用 Itti, GB, AC, SR, 基于对比度模糊生长的图像注意分析(MZ)算法^[19]进行对比实验,比较结果如图 6 所示。从图中可以看出,本文提出方法较之前的方法更能精确定位感兴趣区域,且区域高亮、细节完整、边缘清晰、背景噪声较少,是一种有效的图像感兴趣区域提取方法。

本文采用准确率-回归率曲线(Precision-recall curve)来实现对本文显著度检测效果的定量衡量。



(a) 原图

(a) Original image



(b) Itti方法

(b) Itti method



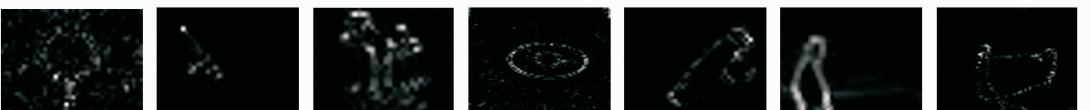
(c) GB方法

(c) GB method



(d) AC方法

(d) AC method



(e) SR方法

(e) SR method

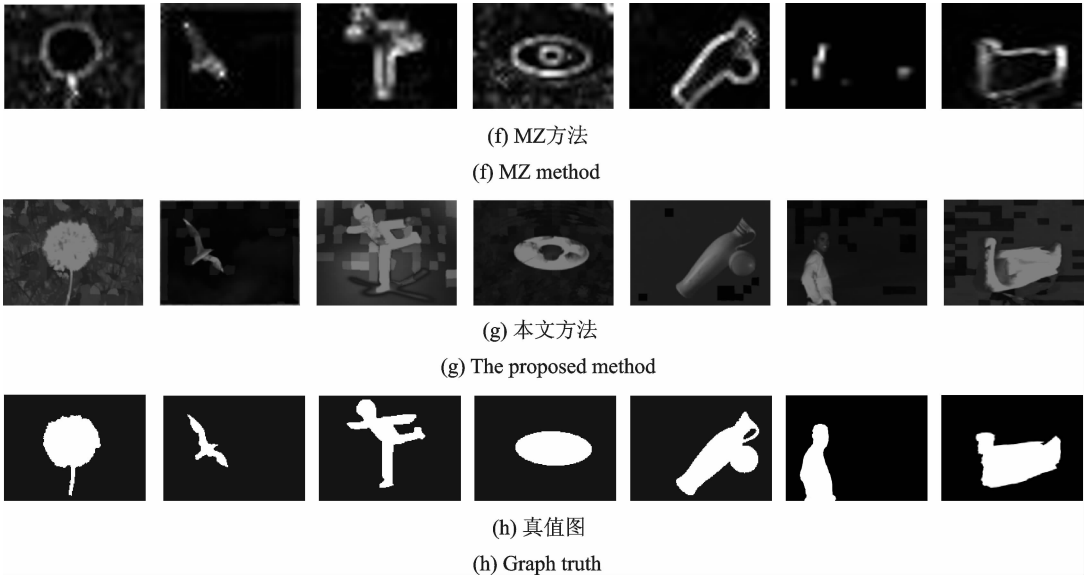


图 6 各种方法比较图

Fig. 6 Comparison of the proposed saliency map with five start-of-art methods

其中,准确率的求解方式为

$$P_i = \frac{1}{N} \sum_{n=1}^N f\left(\frac{|I_n^i \cap I_n^s|}{I_n^s} - \xi_i\right) \quad \xi_i = \{0, 1, \dots, 255\} \quad (16)$$

式中: N 为图像的总数,本文实验为 $N=1\,000$, $|I_n^i \cap I_n^s|$ 为检测到的感兴趣区域与图像真值图中所包含的区域的相交部分, I_n^s 表示检测到的感兴趣区域, ξ_i 表示阈值, $f(x)$ 表示二值函数,如式(17)所示。

$$f(x) = \begin{cases} 0 & x \leq 0 \\ 1 & x > 0 \end{cases} \quad (17)$$

回归率求解公式与准确率求解公式相似,如式(18)所示,用显著度图和图像真值图的相交区域与图像真值图所包含的区域的比值表示。

$$R_i = \frac{1}{N} \sum_{n=1}^N f\left(\frac{|I_n^i \cap I_n^s|}{I_n^i} - \xi_i\right) \quad \xi_i = \{0, 1, \dots, 255\} \quad (18)$$

准确率-回归率曲线能够精确反映算法的感兴趣区域检测性能。曲线越高,代表算法的准确率越高;曲线下降越慢,代表算法的精确率和查全率越高;曲线覆盖越全面说明算法的查全率越高。本文提出算法的准确率-回归率曲线如图 7 中 OUR 所示,从图中可看出本文提出的方法具有更高的精确率和查全率。在回归率值为 1 时,所有方法都具有相同的准确率值,这是因为检测图像已经全部被看作为感兴趣区域。

3 结束语

本文提出了一种中低层相结合的图像显著度检测方法。利用彩色增强的角点检测算子形成凸包与采用 GBR 方法对 SLIC 所形成的超像素进行聚类的结果相结合以确定中层次的显著区域,再利用多尺度不同权重的差分高斯滤波器求解

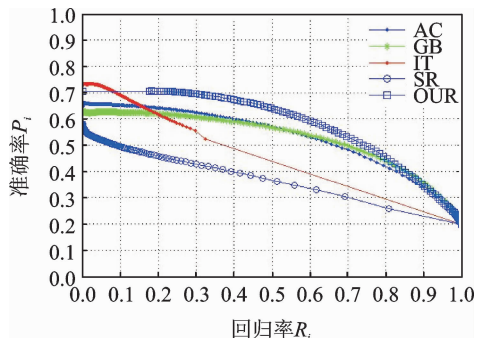


图 7 准确率-回归率曲线图

Fig. 7 Precision-recall curve of different methods

低层次的显著图,最后加权融合中低层次的显著度区域得到最终结果。中层次信息保证了图像显著区域的正确定位,多尺度不同权重的高斯滤波为显著区域添加细节,使显著区域更加完善。通过主观观察可以发现本文方法较其他方法具有更准确的定位,边缘更加清晰,内部细节更加完整,具有更好的图像显著度标注效果。在 MSRA 数据库上对本文方法进行定量分析,结果显示本文方法较其他比较算法具有更高的准确率和查全率,是一种较好的图像感兴趣区域标注方法。但 GBR 算法中的结果对于参数 σ 较敏感,且参数 σ 的取值范围较大,在后续的研究工作中将对参数 σ 进行分区间遍历选择来实现 GBR 算法的最好效果。

参考文献:

- [1] Itti L, Koch C. Computational modelling of visual attention[J]. *Nature Reviews Neuroscience*, 2001, 2(3): 194-203.
- [2] Koch K, McLean J, Segev R, et al. How much the eye tells the brain [J]. *Current Biology*, 2006, 16(14): 1428-1434.
- [3] 欧阳甸, 张伟华, 董寿, 等. 基于 SVAC 感兴趣区域的码率控制算法[J]. *数据采集与处理*, 2014, 29(1):101-107.
Ouyang Dian, Zhang Weihua, Dong Shou, et al. Rate control algorithm for region of interest based on SVAC[J]. *Journal of Data Acquisition and Processing*, 2014, 29(1):101-107.
- [4] 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 (11): 1254-1259.
- [5] Harel J, Koch C, Perona P. Graph-based visual saliency[C]//*Advances in Neural Information Processing Systems*. Vancouver, B. C., Canada: MIT Press, 2006: 545-552.
- [6] Achanta R, Estrada F, Wils P, et al. Salient region detection and segmentation[C]//*Computer Vision Systems*. Berlin Heidelberg: Springer, 2008: 66-75.
- [7] Hou X, Zhang L. Saliency detection: A spectral residual approach[C]//*CVPR'07*, IEEE Conference on Computer Vision and Pattern Recognition, 2007. Minneapolis, Minnesota, USA: IEEE, 2007: 1-8.
- [8] Franco L, Rolls E T, Aggelopoulos N C, et al. Neuronal selectivity, population sparseness, and ergodicity in the inferior temporal visual cortex[J]. *Biological Cybernetics*, 2007, 96(6):547-560.
- [9] Koch C, Poggio T. Predicting the visual world: Silence is golden[J]. *Nature Neuroscience*, 1999, 2: 9-10.
- [10] Xie Y, Lu H, Yang M H. Bayesian saliency via low and mid level cues[J]. *Image Processing*, IEEE Transactions on, 2013, 22(5):1689-1698.
- [11] Achanta R, Hemami S, Estrada F, et al. Frequency-tuned salient region detection[C]// *IEEE International Conference on Computer Vision and Pattern Recognition 2009*. Miami, FL, USA: IEEE, 2009:1597-1604.
- [12] 闯跃龙, 楼宋江, 张石清, 等. 结合区域和边界信息的图像显著度检测[J]. *中国图象图形学报*, 2016, 21(3):314-322.
Chuang Yuelong, Lou Songjiang, Zhang Shiqing, et al. Image saliency by regions and edges[J]. *Journal of Image and Graphics*, 2016, 21(3):314-322.
- [13] Qian P, Chung F L, Wang S, et al. Fast graph-based relaxed clustering for large data sets using minimal enclosing ball[J]. *Systems, Man, and Cybernetics, Part B: Cybernetics*, IEEE Transactions on, 2012, 42(3): 672-687.
- [14] Van de Weijer J, Gevers T, Bagdanov A D. Boosting color saliency in image feature detection[J]. *Pattern Analysis and Machine Intelligence*, IEEE Transactions on, 2006, 28(1): 150-156.
- [15] Harris C, Stephens M. A combined corner and edge detector[C]//*Alvey Vision Conference*. Manchester, UK: Organising Committee AVC 88 Press, 1988, 15: 50.
- [16] Achanta R, Shaji A, Smith K, et al. SLIC superpixels compared to state-of-the-art superpixel methods[J]. *Pattern Analysis and Machine Intelligence*, IEEE Transactions on, 2012, 34(11): 2274-2282.
- [17] Lee C H, Zaiane O R, Park H H, et al. Clustering high dimensional data: A graph-based relaxed optimization approach[J]. *Information Science*, 2008, 178(23):4501-4511.
- [18] 肖志涛, 王红, 张芳, 等. 复杂自然环境下感兴趣区域检测[J]. *中国图象图形学报*, 2015, 20(5):625-632.
Xiao Zhitao, Wang Hong, Zhang Fang, et al. ROI detection under the complicated natural environment[J]. *Journal of Image and Graphics*, 2015, 20(5):625-632.
- [19] Ma Y F, Zhang H J. Contrast-based image attention analysis by using fuzzy growing[C]//*Proceedings of the eleventh ACM international conference on Multimedia*. Berkeley, CA, USA: ACM, 2003: 374-381.

作者简介:



周洁(1992-),女,硕士研究生,研究方向:模式识别、目标检测, E-mail: 799489588@qq.com.



王士同(1964-),男,教授,研究方向:人工智能与模式识别, E-mail: wxwangst@aliyun.com.