

7-31-2020

Image Saliency Analysis Based on Grey Relational Computation and Prior Combination

Qiangqiang Zhou

1. CAD Research Center of College of Electronics & Information, TongJi University, Shanghai 201804, China; ;2. College of Computer & Information, JiangXi Agriculture Universtiy, Nanchang 330027, China;

Zhicheng Wang

1. CAD Research Center of College of Electronics & Information, TongJi University, Shanghai 201804, China; ;

Weidong Zhao

1. CAD Research Center of College of Electronics & Information, TongJi University, Shanghai 201804, China; ;

Yufei Chen

1. CAD Research Center of College of Electronics & Information, TongJi University, Shanghai 201804, China; ;

See next page for additional authors

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the [Artificial Intelligence and Robotics Commons](#), [Computer Engineering Commons](#), [Numerical Analysis and Scientific Computing Commons](#), [Operations Research](#), [Systems Engineering and Industrial Engineering Commons](#), and the [Systems Science Commons](#)

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

Image Saliency Analysis Based on Grey Relational Computation and Prior Combination

Abstract

Abstract: An image saliency analysis method based on *the combination of grey relational computation and prior knowledge*, which is a saliency computation and integration based on three different *priori of boundary background, global color distribution and top salient information*. The SLIC method segment the image into superpixels, then selecting the boundary superpixels on four image sides respectively to construct the reference sequence, and the same method used to construct the comparative sequence of the rest superpixels. Then four initial saliency images was derived based on boundary background priori through a *grey relational computation*, to make weighted and improved. Furtherly integrated the global color spatial distribution information and the top priori information to produce the final saliency image. An experiment were made on the open MSRA-1000 datasets and compared the results with several existing classic methods, conclusions shows the method has a better precision and recall, even under a more complex background.

Keywords

image saliency, grey relational analysis, regional characteristics descriptor, prior knowledge

Authors

Qiangqiang Zhou, Zhicheng Wang, Weidong Zhao, Yufei Chen, and Wang Gang

Recommended Citation

Zhou Qiangqiang, Wang Zhicheng, Zhao Weidong, Chen Yufei, Wang Gang. Image Saliency Analysis Based on Grey Relational Computation and Prior Combination[J]. Journal of System Simulation, 2015, 27(7): 1511-1519.

联合灰色关联度和先验的图像显著性分析

周强强^{1,2}, 王志成¹, 赵卫东¹, 陈宇飞¹, 王刚¹

(1. 同济大学电子与信息工程学院 CAD 研究中心, 上海 201804; 2. 江西农业大学计算机与信息工程学院, 南昌 330027)

摘要: 在基于边界背景先验, 全局显著性先验和高层先验这三种信息基础上, 提出了一种联合灰色关联度计算和先验知识的图像显著性分析方法。用 SLIC 方法对图像进行超像素分割, 再分别选取图像 4 个边界的超像素构建参考序列, 对图像剩余区域构建比较序列。进行灰色关联度计算, 得到 4 个基于边界背景先验的初始显著性图, 进行加权融合优化。进一步综合全局的颜色空间分布信息和高层先验信息, 生成最终显著图。在公开的 MSRA-1000 数据集上进行试验, 并与现有的几个经典方法进行了比较, 显示有更好的准确率和召回率, 且在较复杂背景下也有好的显著性检测效果。

关键词: 图像显著性, 灰色关联分析, 区域特征描述子, 先验知识

中图分类号: TP391

文献标识码: A

文章编号: 1004-731X (2015) 07-1511-10

Image Saliency Analysis Based on Grey Relational Computation and Prior Combination

Zhou Qiangqiang^{1,2}, Wang Zhicheng¹, Zhao Weidong¹, Chen Yufei¹, Wang Gang¹

(1. CAD Research Center of College of Electronics & Information, Tongji University, Shanghai 201804, China;

2. College of Computer & Information, JiangXi Agriculture University, Nanchang 330027, China)

Abstract: An image saliency analysis method based on the combination of grey relational computation and prior knowledge, which is a saliency computation and integration based on three different priori of boundary background, global color distribution and top salient information. The SLIC method segment the image into superpixels, then selecting the boundary superpixels on four image sides respectively to construct the reference sequence, and the same method used to construct the comparative sequence of the rest superpixels. Then four initial saliency images was derived based on boundary background priori through a grey relational computation, to make weighted and improved. Furtherly integrated the global color spatial distribution information and the top priori information to produce the final saliency image. An experiment were made on the open MSRA-1000 datasets and compared the results with several existing classic methods, conclusions shows the method has a better precision and recall, even under a more complex background.

Keywords: image saliency; grey relational analysis; regional characteristics descriptor; prior knowledge

引言

心理生物学家通过对人类视觉注意选择机制



收稿日期: 2014-06-19

修回日期: 2014-10-05;

基金项目: 国家自然科学基金项目资助(61103070);

作者简介: 周强强(1979-), 男, 江西永新, 博士生, 研究方向为计算机视觉; 王志成(1975-), 男, 博士, 副研究员, 研究方向为模式识别。

的研究, 逐渐形成了视觉显著性理论。该理论认为: 人类视觉系统具有选择性注意特性, 只重点处理图像的某些局部信息, 而对其余部分进行忽略或抑制, 能在未知场景中不受复杂背景的影响, 快速而准确的检测到显著目标。这种被视觉系统重点关注处理的局部区域通常被称为显著性区域(saliency region)。显著性检测在信号处理, 计算机视觉, 机

器学习甚至生物学当中都有广泛的应用研究,比如自动图像剪切,图像/视频压缩,图像分类,内容感知,图像预览等。

视觉显著性检测可分为 2 类:(1)自底向上,数据驱动的显著性提取;(2)自顶向下,任务驱动的显著性提取。但基于运算复杂性和通用性方面的考虑,目前大多数的显著性检测算法都是自底向上的。

自从 Itti 等人^[1]提出了具有开拓性的显著性计算模型以来,目前已经涌现了大量各种类型的算法,有基于贝叶斯模型的^[2],基于决策理论模型的^[3],以及基于图模型、谱分析等的检测算法^[4]。近些年来受一些机器学习和数学模型的启发,压缩感知^[5]、流形学习^[6]、运筹学^[7]、偏微分方程^[8]等也运用进来检测显著图像特征。比如文献[6]中通过基于图的流形排序来对有前景信息或背景信息的图像元素(像素或区域)进行相似性排序,图像元素的显著性基于它们与给定种子或查询的相关性来定义。

以上所列自底向上的显著目标检测算法大多数都是基于局部或全局的对比度,只依赖当前观测图像的信息来衡量显著性的,因此当背景较杂乱时会出现检测结果不够理想,不能均匀的突出目标等情况。由于缺少高层的理解知识,自底向上的显著性目标检测就成了一个病态(ill-posed problem)问题。而利用合理的先验知识有助于病态问题的解决。文献[2]指出,人们长久以来积累的先验知识在显著性检测中也发挥着重要的作用,并用大量的人们经常会看到的自然场景图像来代表先验知识。文献[9]中提出,除了先前研究中经常利用的基于局部或全局的目标/背景的对比度先验,还发现更具一般性的背景先验:图像的边界大多数是背景,大部分背景区域容易与图像边界连通。并因此而提出了测地显著性的概念。将区域(图像块或者超像素)的显著性定义为该区域到图像四周的最短距离。算法的主要思想是首先检查出背景,进而得到目标区域。文献[10]中,利用边界背景先验,定义“伪背景”区域为图像的 15 像素宽的窄边界区域,并提取“伪背景”区域作参考,来为每个图像区

域计算背景描述子,构造特征向量。文献[6]的算法中也利用到了文献[9]中的背景先验来提取显著目标信息。

受文献[6,9-10]的启发,本文提出了一种联合灰色关联度计算和先验知识的图像显著性分析方法,并与几种较为典型的算法在公开数据集上进行了实验对比,证明了本文算法的有效性。本文的主要贡献归纳如下:(1)提出了一种新的基于灰色关联度预测的显著区域检测模型;(2)融合进了多种高层先验信息;(3)不同于现有大多数显著性检测模型中多特征的简单加权,本文中是构建一个具有判决性的特征向量输入,结合先验信息,通过阈值选择得到期望的显著性图。

1 系统框架

本文构造的系统模块框图如图 1 所示,主要是基于:边界背景先验,全局显著性先验和高层先验这 3 种信息基础上的显著性计算和融合。首先用 SLIC 方法^[11]对图像 I 进行超像素分割(设定超像素的数量 $N=200$),这是因为使用超像素能有效地降低计算复杂度,过分割得到的超像素能够有效的保留图像的结构信息,而且使用超像素进行显著度检测能够有效防止目标像素的漏检。然后分别选取图像 4 个边界的超像素,选择适当的图像特征构造参考序列,进行灰色关联度计算,得到 4 个基于边界背景先验的初始显著性图,进行加权融合,优化改善。再进一步综合全局的颜色空间分布信息和高层先验信息,以生成最终的多尺度融合显著图。

2 联合灰色关联度计算和背景先验的图像显著性分析

2.1 灰色关联分析

灰色关联分析法是一种按照待分析系统的各特征参量序列构成的发展趋势接近程度来度量待分析系统与参照系统之间相似性的一种指标,待分析系统与参照系统越贴近,则关联度越大。假设有 2 个数列,设参考序列为 $X_0=\{x_0(1), x_0(2), \dots, x_0(n)\}$,

比较序列为 $X_i = \{x_i(1), x_i(2), \dots, x_i(n)\} (i=1, 2, \dots, m)$, 则在同一时刻它们间的关联度定义为^[12]:

$$\gamma_{0i} = \frac{1}{n} \sum_{k=1}^n \gamma_{0i}(k), \quad i=1, 2, \dots, m; \quad \text{其中,}$$

$$\gamma_{0i}(k) = \frac{\min |x_0(k) - x_i(k)| + 0.5 \max |x_0(k) - x_i(k)|}{|x_0(k) - x_i(k)| + 0.5 \max |x_0(k) - x_i(k)|}$$

式中: $\gamma_{0i}(k)$ 为关联系数。

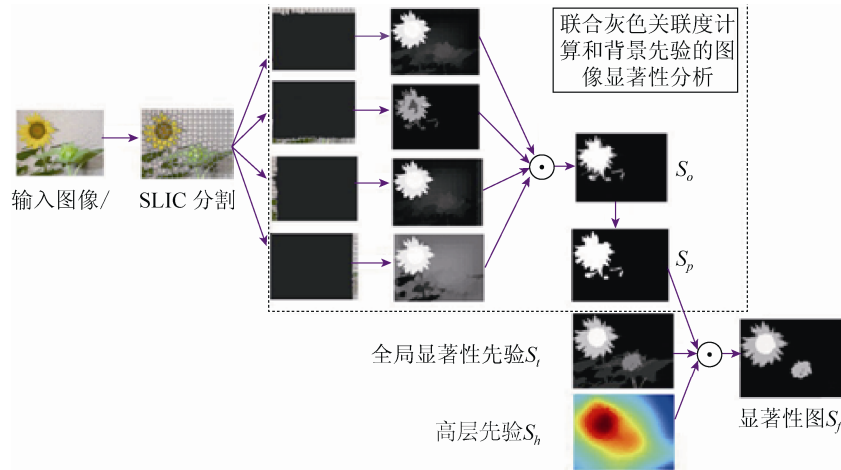


图 1 系统框图

为了规范各序列的数据量纲, 增强数据的可比性, 计算灰色关联度前一般要进行一个数据规范化处理。然后对得到的灰色关联度进行排序, 完成关联分析。

本文在将灰色关联度应用于图像显著性计算时, 首先在图像 SLIC 分割所产生的超像素节点集合 R_l 中选取合适的节点, 挑选一些具有判定性的图像特征来构造参考序列 X_0 , 然后其它每个超像素 R_i 用同样的特征参数做比较序列 X_i , 计算关联度 γ_{0i} , 并归一化到 $[0, 1]$ 范围得 γ_{0i}^* 。同时, 定义一个超像素 R_i 的显著性值为:

(1) 当 γ_{0i} 是根据具有明显的显著目标特性的参考序列计算而得时, 则 R_i 显著性值为 γ_{0i}^* ;

(2) 当 γ_{0i} 是根据具有明显的图像背景特性的参考序列计算而得时, 则 R_i 显著性值为 $1 - \gamma_{0i}^*$ 。

2.2 图像特征的序列构建和显著性计算

(1) 参考序列和比较序列的构建

选取具有指标性意义的图像特征来构建参考序列和比较序列, 是涉及基于灰色关联度计算的关键一步。这里, 我们选用显著性计算中具有判定性的最重要的 20 个区域特征^[10], 如下表 1 所示, 其中有 3 种类型的区域描述子, 分别是:

表1 具有判定性的20个区域特征

definition	dim	definition	dim
b_1 $x^2(h_1^R, h_1^S)$	1	b_5 $x^2(h_4^R, h_4^S)$	1
p_1 the 90 th percentile of the normalized x coordinates	1	b_6 $d(a_g^R, a_g^S)$	1
p_2 the 10 th percentile of the normalized y coordinates	1	b_7 $d(a_{a^*}^R, a_{a^*}^S)$	1
b_2 $x^2(h_2^R, h_2^S)$	1	c_1 $x^2(h_1^R, h_1^S)$	1
p_3 the aspect ratio of the bounding box	1	b_8 $d(a_{b^*}^R, a_{b^*}^S)$	1
p_4 the average normalized y coordinates	1	p_7 the variance of the G values	1
p_5 the 10 th percentile of the normalized x coordinates	1	p_8 the variance of the a* values	1
b_3 $x^2(h_3^R, h_3^S)$	1	p_9 the variance of the B values	1
p_6 the 90 th percentile of the normalized y coordinates	1	b_9 $d(a_{L^*}^R, a_{L^*}^S)$	1
b_4 $d(a_r^R, a_r^S)$	1	b_{10} $d(a_b^R, a_b^S)$	1

1) 区域对比度描述子: c_1

$$\chi^2(h_i^R, h_i^S) = \sum_{i=1}^b \frac{2(h_i^R - h_i^S)^2}{h_i^R + h_i^S}, \text{ 这里的 } b$$

是直方图中 bin 的个数, h_i 表示的是 $L*a*b$ 直方图; R 是某个超像素区域; S 是 R 的所有直接近邻;

2) 区域背景描述子: $b_1 \sim b_{10}$

文献[10]中定义了一个“伪背景区域”, 是图像的只具有 15 个像素宽的窄边界区域, 并在实验中发现图像边界 98% 左右的像素都属于背景, 从而证明了这个假设的存在。因此, 提取出这个“伪背景区域”来作参照, 以计算每个区域的背景描述子。表中 R 是某个超像素区域, S 是伪背景区域。

$$\chi^2(h_k^R, h_k^S) = \sum_{i=1}^b \frac{2(h_{ki}^R - h_{ki}^S)^2}{h_{ki}^R + h_{ki}^S},$$

其中 $k=1, 2, 3, 4$ 。这里的 h_1 为 $L*a*b$ 直方图; h_2 为 hue 直方图; h_3 为 saturation 直方图; h_4 为纹理元直方图。

$d(v_1, v_2) = (|x_{11} - x_{21}|, \dots, |x_{1k} - x_{2k}|)$, k 是向量 v_1 和 v_2 中的元素个数。表 1 中的 a_r, a_g, a_b 分别是 RGB 颜色空间中的 r, g, b 颜色分量均值。 a_{L*}, a_{a*}, a_{b*} 分别是 $L*a*b$ 颜色空间中的 $L*, a*, b*$ 分量均值。

3) 区域属性描述子: $p_1 \sim p_9$

考虑的主要是区域的外观和几何特征。这里的外观特征指的是区域颜色和纹理的分布, 几何特征指的是区域的大小和位置。表中, average normalized coordinates 表示区域中规一化了的 x 或 y 坐标的平均值; aspect ratio 表示区域边界盒的纵横比; normalized area 是规一化了的区域面积; the 10th percentile of the normalized x coordinates 表示处于区域中所有像素规一化了的 x 坐标由小到大排序中的第 10 个百分位处的值; the variance of the $G(B \text{ 或 } a^*)$ values 是区域 RGB($L*a*b^*$) 颜色空间中 $G(B \text{ 或 } a^*)$ 分量值的方差。

因此, 我们选用这 20 维的特征向量来构造参考序列和比较序列, 即:

$$\mathbf{X}_0 = \{x_0(1), x_0(2), \dots, x_0(20)\} = \{x_0(b_1), x_0(p_1), x_0(p_2), x_0(b_2), x_0(p_3), x_0(p_4), x_0(p_5), x_0(b_3), x_0(p_6),$$

$$x_0(b_4), x_0(b_5), x_0(b_6), x_0(b_7), x_0(c_1), x_0(b_8), x_0(p_7), x_0(p_8), x_0(p_9), x_0(b_9), x_0(b_{10})\}$$

$$\mathbf{X}_i = \{x_i(1), x_i(2), \dots, x_i(20)\} = \{x_i(b_1), x_i(p_1), x_i(p_2), x_i(b_2), x_i(p_3), x_i(p_4), x_i(p_5), x_i(b_3), x_i(p_6), x_i(b_4), x_i(b_5), x_i(b_6), x_i(b_7), x_i(c_1), x_i(b_8), x_i(p_7), x_i(p_8), x_i(p_9), x_i(b_9), x_i(b_{10})\} \quad (i=1, 2, \dots, m)$$

(2) 基于边界背景先验的显著性计算

受文献[6,13]启发, 这里主要考虑背景特性。一般认为图像四周都是背景区域, 图像中目标一般不会与边界相交, 我们把这称作-边界先验。边界先验比常用的中央先验更具普遍性。尽管图像目标不会经常出现在正中央, 但也一般不会与图像边界相交, 这是摄影学中的“1/3 规则”即: 目标一般被放置在中心到边界的 1/3 处。这里, 我们采用“先分离再合并”的策略, 分别选取图像 4 个边界上的超像素(X_{0t} : 上边界超像素集合; X_{0b} : 下边界超像素集合; X_{0l} : 左边界超像素集合; X_{0r} : 右边界超像素集合)作为参考序列, 如图 2 所示。于是求得每个边界所有参考序列的平均值, 即 $\overline{X_{0i}}, \overline{X_{0b}}, \overline{X_{0l}}, \overline{X_{0r}}$, 依次作为基于边界背景先验的参考序列, 来计算图像中其它所有超像素节点的灰色关联度并进行归一化得 γ_{0i}^* , 同时将超像素节点的显著性值定义为 $1-\gamma_{0i}^*$ 。如此可得到 4 幅图像的显著性图, S_t, S_b, S_l 和 S_r , 进行自定义的加权运算合并后可得到初始显著图 S_0 :

$$S_0 = S_t \odot S_b \odot S_l \odot S_r$$

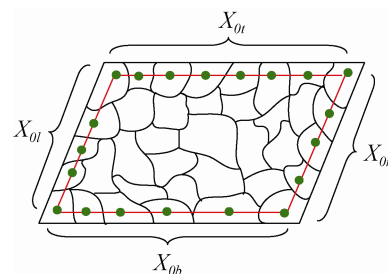


图 2 图像四边界超像素集合

上面采用“先分离再合并”的方法主要是基于如下考虑: 图像四边界超像素节点的特征差异一般较大, 比如上边界是蓝天, 下边界是沙漠。如果

将 4 个边界的超像素节点视为一体, 同时作为参考序列 X_0 , 那么参考序列内在的差异性将影响其它超像素节点与它的关联度计算, 也使得检测的 S_0 效果不理想。而且发现, 单个边界上超像素节点特征的变化一般比较小, 并且边界背景是互相连通的。因此, 将 4 个边界分别进行处理再合并, 要比 4 个边界同时处理更加合理。虽然以每个边界为参考序列所得到的显著图只能抑制与对应边界特征相似且位置临近的局部背景区域, 但最终得到的 4 个显著图一起, 就能较好地抑制所有背景区域, 凸显前景目标。

(3) 基于前景目标的显著性计算

尽管上面的基于边界背景先验的参考序列可以检测出大部分的显著目标, 并且从背景出发可以减少杂乱的背景噪声, 得到更加准确的前景检测结果。但由于背景中仍可能有“奇异点”的存在, 或当有显著目标贴着边界时(如图 3 所示), 以上基于边界构建的参考序列就会包含一些目标特征, 导致部分背景超像素节点不能够被很好的抑制, 影响显著性的检测效果。为了减轻这个问题和提升当目标挨靠边界时的检测效果, 我们通过以下方法进一步改善。



图 3 显著目标贴着边界时的情况

对初始显著性图 S_0 选定一个阈值, 进行二值分类以得到前景目标。这里我们可将阈值定义为显著图的均值。这其实是一种自适应阈值, 期望通过该方式, 得到更多真实显著性的目标节点, 也就是更高的召回率。不同的显著图, 显著值的范围变化很大, 因此自适应阈值要优于固定阈值的方法。

类似前面基于边界背景先验的参考序列, 先对二值分类所得到的前景目标构建参考序列 X_0 , 再对图像中的所有超像素节点(包括前景显著目标所

覆盖的超像素)进行第 2 次关联度的计算, 得到每个超像素节点 i 的关联度值 γ_{0i} , 亦即得到了进一步改善的显著性图 $S_p = \gamma_{0i}^* (i=1, 2, \dots, N)$ 。

3 全局显著性先验

上面的分析主要是从图像的区域特征来进行关联和判断的, 那么有没有存在一种与显著目标有关的全局特征呢? 从大量图像数据集中可观察到, 图像 I 中的颜色分布范围越宽广, 显著性目标中包含该颜色的可能性也越小, 即该种颜色是显著目标的可能性越低。因此, 特定颜色的全局空间分布可以用来描述一个目标的显著性, 这其实也是另一种图像先验性信息。要描述某个特定颜色的空间分布, 最简单的方法是计算该颜色的空间方差。

首先, 图像中所有颜色可以用混合高斯模型 $GMM\{w_c, \mu_c, \Sigma_c\}_{c=1}^C$ 建模来表示, 看作是若干个 GMM 分量的加权联结^[14]。这里 $\{w_c, \mu_c, \Sigma_c\}$ 分别表示第 c 个分量的权值, 平均颜色和协方差矩阵。每个超像素块 R_x 的颜色均值以下面的概率赋予所对应的颜色分量 c :

$$p(c | R_x) = \frac{\omega_c N(R_x | \mu_c, \Sigma_c)}{\sum_c \omega_c N(R_x | \mu_c, \Sigma_c)}.$$

其中, 每一颜色分量 c 的空间位置的水平方差 $V_h(c)$ 是:

$$V_h(c) = \frac{1}{|X|_c} \sum_x p(c | R_x) \cdot |x_h - M_h(c)|^2, \text{ 其中, } M_h(c) = \frac{1}{|X|_c} \sum_x p(c | R_x) \cdot x_h.$$

公式中 x_h 表示超像素 R_x 的 x -坐标, $|X|_c = \sum_x p(c | R_x)$, 垂直方差 $V_v(c)$ 也是类似定义。于是, 分量 c 的空间方差 $V(c) = V_h(c) + V_v(c)$ 。通过公式 $\frac{V(c) - \min_c V(c)}{\max_c V(c) - \min_c V(c)}$, 把 $\{V(c)\}$ 归一化到区间 $[0, 1]$ 。最后, 图像在超像素尺度上的颜色空间分布特征 $f_s(R_x, I)$ 定义为一加权和:

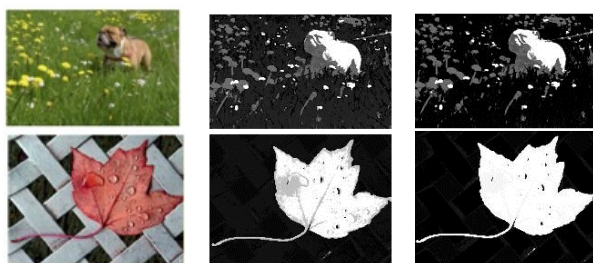
$$f_s(R_x, I) \propto \sum_c p(c | R_x) \cdot (1 - V(c)).$$

特征图 $f_s(\cdot, I)$ 也归一化到区间 $[0, 1]$ 。图 4(b) 中给出了 2 个图例的颜色空间分布特征图, 全局特

征都较好的覆盖了这些显著目标。但注意到图像角落或边界上的颜色空间分布也可能较小, 方差小, 显著性就强。这是因为图像是从一个完整的场景中分割出来所造成的。为了减少这种影响, 重新定义了一个有中央-加权, 空间-方差的特征为:

$$f_s(R_x, I) \propto \sum_c p(c | R_x) \cdot (1 - V(c)) \cdot (1 - D(c)),$$

其中 $D(c) = \sum_x p(c | R_x) d_x$ 是一个赋予图像边界附近的颜色有较小重要度的权值, 类似于 $V(c)$, 也把它归一化到 $[0, 1]$ 。 d_x 是从超像素 R_x 中央位置到图像中央的距离。如图 4(c) 所示, 中央-加权, 颜色空间方差显示了每种颜色显著性的较好预测。



(a) 输入图像 (b) 颜色空间方差特征图 (c) 中央-加权了的颜色空间方差特征图

图 4 全局显著性先验

4 高层先验

高层先验一般是基于人们感知的认识, 目前也在一些显著图研究中得到了应用^[11,13]。把高层先验融合到显著图的增强计算中来, 这样可以面向更加复杂环境背景下的显著性目标检测。这里我们选择结合:

(1) 中央先验: 靠近图像中央的目标更吸引人。因此, 可以用一基于像素到图像中央距离的高斯分布来生成一个先验图, $p_l(x) = \exp(-d^2(x, c)/\sigma_1)$, 这里的 σ_1 定为 $(2 \langle d^2(x, c) \rangle)^{-1}$, 其中 $\langle \cdot \rangle$ 表示所有“点对”距离的期望值;

(2) 语义先验: 人们经常甚至在没有特定目的的情况下, 对某些有特定语义的目标也很注意, 比如图片中的人脸(有时候还包括汽车、文字等)。因此, 这里我们把检测到有人脸的区域 Λ , 如果

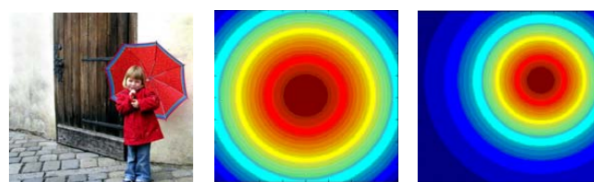
$x \in \Lambda$, 则赋以有能生成人脸先验图 $p_f(x) = \sigma_2$ 的更高概率, 否则, $p_f(x) = 0$ 。 σ_2 是一个可以通过人脸识别算法的输出阈值化获取的常量;

(3) 颜色先验: 根据人们的日常经验, 诸如红色和黄色这些的暖色调一般更醒目。为了利用这个信息, 我们对标定了显著目标的 MSRA 数据集, 在规则化了的 r - g 颜色空间中 ($r = R/(R+G+B)$, $g = G/(R+G+B)$), 获得一 2-D 的直方图分布 $H(S)$ 。类似地, 表示背景颜色分布的直方图 $H(B)$ 也可以得到。对于每个量化颜色, 我们从这两个直方图中获取它们的值, 即 h_S 和 h_B 。因此, 设定显著性的颜色先验为 $p_c(x) = \exp((h_S(c_x) - h_B(c_x))/\sigma_3)$ 。这里 c_x 表示位置 x 处的颜色, σ_3 取 0.02。

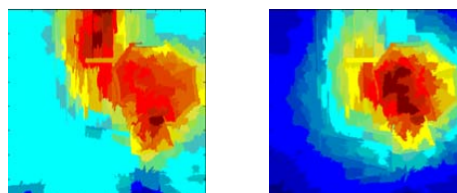
上面得到的这几个先验图 $p_l(x)$, $p_f(x)$, $p_c(x)$ 可以通过下式

$$p(x) = \frac{(p(x) - \min_x(p(x)))}{\max_x(p(x)) - \min_x(p(x))}$$

归一化到 $[0, 1]$ 。然后再通过均值化和空间平滑(一般通过加权平均, 权值是用高斯核计算超像素块间距离生成的)以生成最终联合的高层先验图 $P_H = [p^{(1)}, p^{(2)}, \dots, p^{(N)}]$, 这里 $p^{(i)} \in [0, 1]$, N 是图像的超像素个数。图 5 中给出了一些高层先验图的示例。对图像中的每个超像素块, 我们使用其中像素先验值的平均值, 当作它的高层先验值。



(a) 输入图 (b) 位置先验图 (c) 人脸识别先验图



(d) 颜色先验图 (e) 最终融合图

图 5 高层先验图

5 多尺度多显著图融合

前面得到的 3 种显著图信息不是相互排斥的, 而是互补使得目标整体更加显著。这里, 我们采用如下方式将上述基于边界背景先验得到的显著图 S_p 、基于全局空间分布先验的显著图 S_l , 以及根据高层先验生成的显著图 S_h 这 3 种信息进行融合得 S_f :

$$S_f = S_p \cdot e^{-S_l/\beta} \cdot S_h$$

其中, β 是平衡因子参数, 用于调整基于背景的先验图与基于全局颜色分布的先验图之间的权重。最后, 也要消除掉融合后图像中那些小的连通区域, 以得到显著图 S_f 。

但以上生成的显著图是按照图 1 中的框架在单一尺度下生成的, 图像的单一尺度表示在图像处理中不一定能有预期效果, 因此现在往往是采用多尺度的图像表示来进行弥补。本文将对 3 个尺度-分别是原始图像, 和下采样获得的 1/2 原始图像、1/4 原始图像这 3 种尺度下生成的显著性图 S_f 进行再次融合, 得到最终显著图 Sal 。具体公式如下所示:

$$Sal(x) = (\zeta_1 * S_{f1}(x) + \zeta_2 * S_{f2}(x)) * (\zeta_3 * S_{f3}(x))$$

其中: $S_{f1}(x)$, $S_{f2}(x)$, $S_{f3}(x)$ 分别代表原始图像、1/2 尺度、1/4 尺度下的显著性图中 x 位置的显著性值。

为了融合不同尺度下的显著性图, 需将它们重新调整为原始图像的大小后再进行融合。根据对不同尺度的显著性图的分析可知, 在原始图像的下采样和小尺度显著性图插值还原过程中图像均会产生失真, 尺度变化越大, 失真越大。因此原始尺度下的显著性图能更好地反映出图像显著性区域的细节, 融合过程中给原始图像的显著性图赋予较大的权值。而最小尺度的显著性图能更好地突出图像显著性区域的位置信息, 因此本文将其作为融合过程中的位置权值以突出图像中显著性区域的位置信息。实验中发现, 当权值 $\zeta_1=1/2$, $\zeta_2=1/4$, $\zeta_3=1$ 时可较完整突出图像中的显著性目标, 并抑制非显著性区域, 得到较好效果。

6 实验比较

为了客观的评估本文算法的有效性和准确度, 我们将本文算法与 4 种较主流算法在公开的数据库 MSRA-1 000 上进行了主观和定量的实验结果对比以及相同条件下的运算时间对比。MSRA-1 000 是 MSRA 数据集的子集, 包含有 1 000 幅人工标注了的可用作算法参照评价的 Ground truth 图像, 该数据库是本领域目前在评价视觉显著性算法上使用较多的数据库。参与比较的 5 个算法是: IT^[1] (本领域引用次数最多的经典算法), SR^[15] (基于频域分析的经典算法), GB^[16] (基于图论的显著性检测算法), LR^[13] (目前效果较好的算法之一, 利用了图像的低秩和稀疏信息, 并结合了一些高层知识), 这 4 种算法作者都提供有实验源码。

6.1 定量对比

为了客观地评价本文算法在 MSRA-1000 数据集上的总体表现, 我们采用本领域目前常用的 3 个指标-准确率 (Precision)、召回率 (Recall) 和 F-Measure, 在 2 种方式下-固定阈值的显著区域分割和自适应阈值的显著区域分割, 来定量地进行本文算法和另外 4 个主流算法的性能对比。

准确率: 定义为正确检测的显著像素个数比上检测到的所有显著像素数;

召回率: 定义为正确检测的显著像素个数比上真值的显著像素个数;

F-measure: 是正确率和召回率的调和加权平均值, 定义为:

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

这里设定 $\beta^2=0.3$ 以强调准确率的重要性。

(1) 固定阈值的显著区域分割

先将显著性图的值量化到 [0, 255] 范围, 并对数据集中的所有图像使用固定阈值的方法进行分割, 阈值从 0~255 顺序选取。给定一个阈值 T_p , 如果某位置点处的显著性值大于该阈值, 则认为此像素点属于显著目标, 设置为 1, 否则认为是背景像

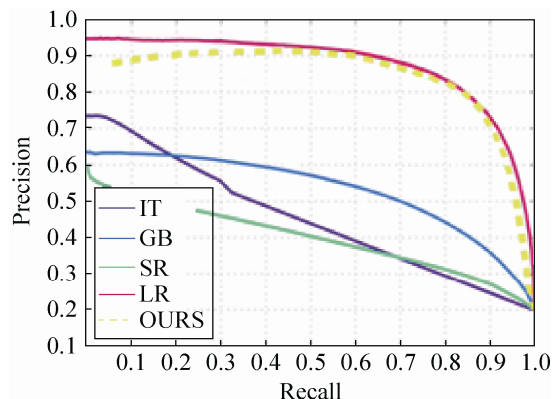
素, 设置为 0, 如此使得前景显著目标的二值掩码图。然后将该图与数据集中的 Ground truth 图进行比较, 就得到准确率和召回率的值。对数据集中的 1 000 幅图像在特定阈值下计算得到准确率和召回率, 再计算平均值即得到了当前阈值下的准确率和召回率。将阈值从 0~255 滑动, 共 256 个阈值, 得到 256 组准确率和召回率。最后, 以横坐标为召回率, 纵坐标为准确率进行描点连线, 便得到最终的平均准确率-召回率曲线, 如图 6(1)所示。该曲线给出了一个各种显著图能多好地突显图像中显著区域的可靠比较。可以发现, Itti 方法在一个很低召回率时(当 $T_f > 240$), 有高的准确率, 之后准确率就急剧下降。这是因为, 这种方法中的显著像素较好的落在了显著区域, 并有近乎一致的值, 但并没有覆盖到整个显著目标。阈值 $T=0$ 时, 召回率为最大的 100%, 所有方法有相同很低的查准率, 因为这时整个图像都被视为显著图, 因此有最低的准确率。从图中可以看到, 在固定阈值时我们的算法和 LR 算法效果相当, 有较好的表现。

(2) 自适应阈值的显著区域分割

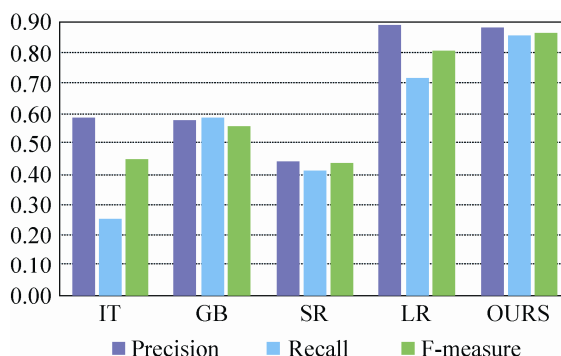
在 Achanta 的文中^[17], 将自适应的阈值设定为整幅图像显著性值均值 m 的 2 倍, 即 $T_f = 2m$ 。显著图更进一步的分析表明, 显著性值的分布是遵循一个非显著性值有比显著性值更大概率的高斯混合模型。因此, 我们可以重新定义自适应阈值为:

$$T_f = m + \delta$$

这里 m 表示显著均值, δ 是标准方差。使用该二值分类阈值后, 就得到了一个目标图。但这种阈值化也可能导致噪声分割, 存在一些分散的小连通区域。因此为了提升显著区域分割方法的健壮性, 我们可以在上述自适应阈值生成二值化图像后, 再在这种显著性指导下用 GrabCut^[18]方法对结果进行改进, 从而明显提升了我们这种方法的准确率, 召回率及 F -measure。 F -measure 的计算如上式 $F\beta$ 所述, 先得到每幅图像的 F -measure, 再在整个数据集上计算它们的平均值。图 6(2)结果表明自适应阈值下, 我们的算法总体上要优于其它 4 个方法。



(1) 固定阈值 P-R 图



(2) 自适应阈值 F-measure 图

图 6 定量对比

6.2 主观对比

如图 7 所示, 是我们 MSRA-1000 数据集实验中有代表性的 4 幅图像, 从检测比较效果上来看, 本文算法是其中最好的。IT 和 GB 方法由于算法本身的局限, 只能得到较模糊显著区域的边界特征, 而忽视了目标内部区域部分; SR 方法得到的显著图只保留了一些非常显著的点, 但丢失了图像中的大部分结构信息, 不能体现显著区域的整体性特征; LR 方法在一些背景相对复杂或当目标与背景颜色相近的情况下检测效果不甚满意; 相比而言, 本文方法由于融合了图像的全局与局部显著特征, 以及一些上层信息, 让检测的结果不仅能够把目标对象的整体显著性凸显出来, 而且还较好的保留了显著对象的边缘信息, 从而得到了相对更准确的检测结果。

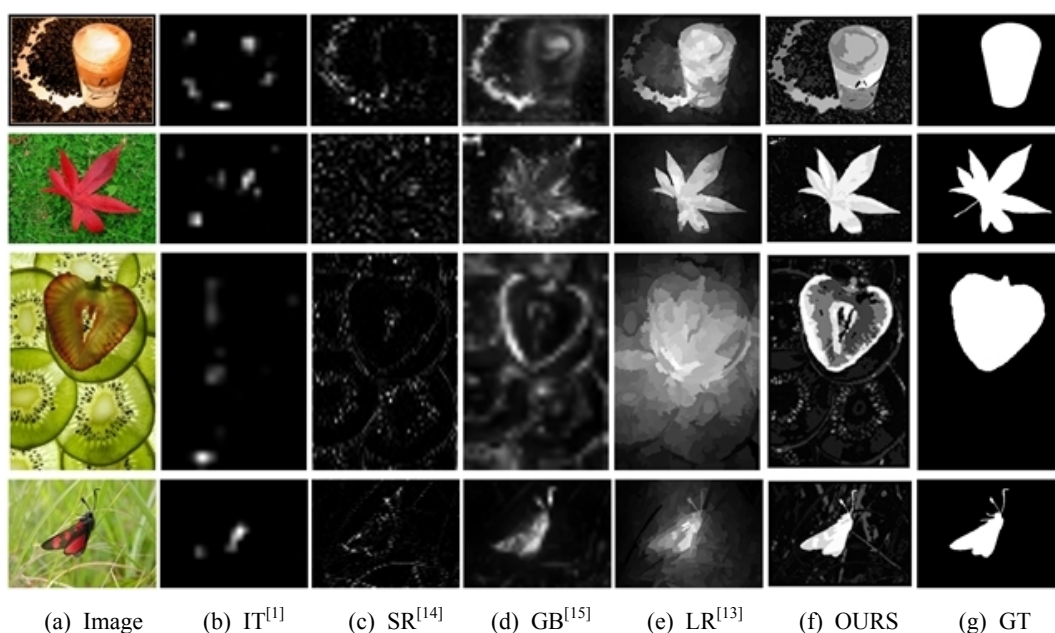


图 7 主观对比

7 结论

本文提出了一种联合灰色关联度计算和先验知识的图像显著性分析方法,这其实是基于边界背景先验,全局显著性先验和高层先验这 3 种信息基础上的显著性计算和融合。实验结果表明,相比于其他几种具有代表性的算法,本文所提算法的检测结果更加准确合理,能较好的抑制复杂背景的干扰,提高图像显著性检测的能力,但是本文模型也还需在运算的复杂度上作进一步的改进。

参考文献:

- [1] 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(S0162-8828), 1998, 20(11): 1254-1259.
- [2] Zhang L, Tong M, Marks T, *et al.* SUN: a Bayesian framework for saliency using natural statistics [J]. Journal of Vision (S1534-7362), 2008, 8(7): 32: 1-20.
- [3] Gao D, Han S, Vasconcelos N. Discriminant saliency, the detection of suspicious coincidences, and applications to visual recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence (S0162-8828), 2009, 31(6): 989-1005.
- [4] Borji A, Itti L. State-of-the-art in visual attention modeling [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence(S0162-8828), 2013, 35(1): 185-207.
- [5] Shen X, Wu Y. A unified approach to salient object detection via low rank matrix recovery [C]// 2012 IEEE Conference on. USA: IEEE, 2012: 853-860.
- [6] Yang C, Zhang L, Lu H, *et al.* Saliency detection via graph-based manifold ranking [C]// Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on. USA: IEEE, 2013: 3166-3173.
- [7] Jiang Z, Davis L S. Submodular Salient Region Detection [C]// Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on. USA: IEEE, 2013: 2043-2050.
- [8] Liu R, Cao J, Lin Z, *et al.* Adaptive partial differential equation learning for visual saliency detection[C]// Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on. IEEE, 2014: 3866-3873.
- [9] Wei Y, Wen F, Zhu W, *et al.* Geodesic saliency using background priors [M]// Computer Vision—ECCV 2012. Germany: Springer Berlin Heidelberg, 2012: 29-42.
- [10] Jiang H, Wang J, Yuan Z, *et al.* Salient object detection: A discriminative regional feature integration approach [C]// Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on. USA: IEEE, 2013: 2083-2090.
- [11] 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(S0162-8828), 2012, 34(11): 2274-2282.
- [12] 马苗, 张艳宁, 赵健. 灰色理论及其在图像工程中的应用[M]. 清华大学出版社, 2011.

(下转第 1555 页)