

多目标显著性区域提取算法

孟 瑒, 陈妹雅
(东北大学 信息科学与工程学院, 辽宁 沈阳 110819)

摘 要: 结合对象估计和超像素分割, 提出面向多目标的显著性区域提取算法. 首先, 应用对象估计对图像中的多目标作初步检测, 得到若干个显著性区域的初步结果; 然后, 再将这些显著性区域与超像素分割的结果作图像拼接, 完善这些显著性区域; 最后, 将图像拼接的结果二值化, 作为多目标显著性区域提取的最终结果. 结果表明: 所提算法可实现面向多目标的显著性区域提取. 与 3 个经典算法的比较结果表明: 所提算法在面向多目标显著性区域提取时更优.

关 键 词: 多目标; 显著性区域; 对象估计; 超像素分割; 图像处理

中图分类号: TP 391 **文献标志码:** A **文章编号:** 1005-3026(2018)10-1380-05

Salient Region Extraction Algorithm for Multi-target

MENG Lu, CHEN Mei-ya
(School of Information Science & Engineering, Northeastern University, Shenyang 110819, China. Corresponding author: MENG Lu, E-mail: menglu@ise.neu.edu.cn)

Abstract: Combining object estimation and super-pixel segmentation, a salient region extraction algorithm for multi-target was proposed. First, object estimation was used to make a preliminary extraction of multi-target in image, and the preliminary results of several salient regions were obtained. Then, these several salient regions were concatenated with the results of super-pixel segmentation to complete the saliency extraction. Finally, the concatenated regions were binarized as the final results of salient region for multi-target. The results showed that the proposed algorithm can achieve multi-target salient region extraction. The comparison with three classical algorithms indicated that the proposed algorithm is better when it is faced with salient region extraction for multi-target.

Key words: multi-target; salient region; object estimation; super-pixel segmentation; image processing

人类的视觉系统具有主动选择机制, 可减少接收到的视觉信息数量, 具有忽略大量不感兴趣信息的能力, 常把这种能力称为视觉注意力机制^[1]. 对于一幅图像, 人眼可以很快关注到图片中感兴趣的部分, 把图像中这些能引起人们关注和能够体现图像信息的关键部分称为显著性区域. 将人类视觉注意机制引入到机器视觉领域, 就可以建立一种显著性区域提取算法^[2].

显著性区域提取算法分为三类: 基于特征的对比(局部或全局), 基于频谱的分析及二者的混合^[3]. 基于局部对比的显著区域检测算法通过计算每个图像子区域或像素同其周围一个小的局部领域的对比度, 度量该图像子区域或像素的显著性, 代表性算法为 Itti 等^[4]提出的 IT 算法. 基于全局对比的显著性检测方法的主要思想是利用整个图像的平均像素强度值来对比图像的像素, 代表性算法为 Cheng 等^[5-6]提出的基于区域对比的 RC 算法. 基于频谱域的显著性检测方法的主要思想是强调图像中频谱域中的不同频率, 在基于频域分析的显著性检测算法中最具代表性的是 SR 模型^[3]. 混合显著性检测方法考虑多尺度、颜色、亮度、频率、图像对比等, 代表性算法为 Jiang 等^[7]提出的判别式多尺度局部显著性检测算法.

现有的显著性提取算法大多是面向单目标图

像,无法对多目标图像进行检测,没有对图像中存在多目标时的底层特征进行分析,有较大的局限性.具体问题:只能识别众多多目标中的一个;把多个目标识别成一个;只能识别目标区域的一部分.针对上述问题,本文结合对象估计和超像素分割,提出了一种多目标显著性区域的提取算法.

1 方 法

本文提出多目标显著性区域提取算法,核心思想是利用对象估计方法对原始图像作一个初步处理,确定图像中各个目标区域的初步范围,从而将多目标显著性区域的提取转化为多个单目标显著性区域的提取.但是,对象估计方法往往只能得到目标区域的一部分,因此,本文采用超像素分割方法来弥补缺失的目标区域,再在各个弥补后的目标区域中计算显著性图,最后将各个目标区域拼接回原图中,从而实现多目标显著性区域提取.算法流程图如图 1 所示.

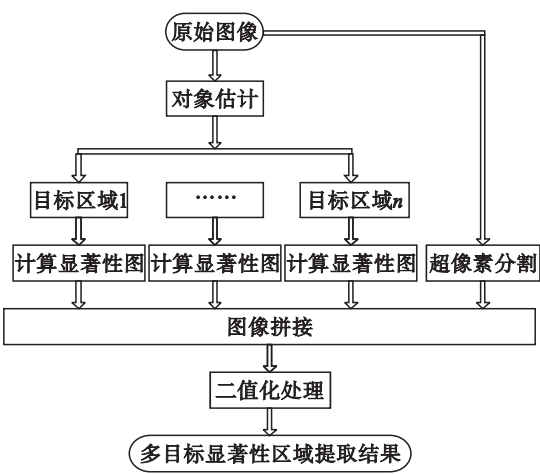


图 1 多目标显著性区域提取算法流程图
Fig. 1 Flow chart of salient region extraction algorithm for multi-target

1.1 对象估计

Alexe 等^[8]提出对象的特征:①具有一个边界,且该边界可以将对象完整包围;②与背景相比,对象在某些特征上体现了差异性,如像素值、纹理、梯度等.按照这一定义,对象估计算法会在一幅图像上生成一系列候选窗口,并给各个窗口赋予一个权值,该权值表示窗口中包含物体的可能性,权值越大,可能性越大.应用该方法,可排除背景区域,并按照权值的大小得到一系列可能包含对象的窗口.对象估计算法应做到以下 4 点^[9]:①实现高目标检测率;②产生少量候选窗口以减少后续探测器的计算时间;③拥有较高的计算效

率可应用于各种应用;④对不可见对象类别具有较好的推广能力.

Cheng 等^[9]发现在固定尺寸的窗口下,目标与背景的梯度特征有所不同,这是因为图像中背景区域往往呈现出同质化的特性,而目标则不会.基于以上分析,本文选择基于二值归一化梯度(binarized normed gradients, BING)的方法来实现对象估计^[9-10],具体步骤如下.

- 1) 在原图上设置若干窗口,这些窗口是随机摆放的,且互相可以重叠;
 - 2) 将原图作多尺度变换,并计算各个尺度下的梯度;
 - 3) 将各个尺度下窗口的尺寸归一化为固定尺寸 8×8 ,用模板 $[1, 0, 1]$ 在窗口的水平和垂直方向计算各个像素的梯度 g_x 和 g_y :
- $$g_x(i, j) = I(i + 1, j) - I(i - 1, j), \quad (1)$$
- $$g_y(i, j) = I(i, j + 1) - I(i, j - 1). \quad (2)$$
- 式中, $I(i, j)$ 表示坐标 (i, j) 处的像素值.
- 4) 计算 8×8 窗口中各个像素点的 NG 特征:

$$NG = \min(|g_x| + |g_y|, 255). \quad (3)$$

- 5) 使用一个 8×8 的分类矩阵,该矩阵是通过大量含有多目标显著性区域的图片训练得到的^[9-10].

$$s = \langle w, g \rangle. \quad (4)$$

式中, w 和 g 都是 64 维向量,二者作内积得到 s . s 值越大表示窗口包括的范围是目标的可能性越大.

1.2 计算显著性图

通过 1.1 节的对象估计,可以得到包含目标区域的若干窗口.在各个窗口中计算显著性图,具体方法如下^[5-6]:首先通过分水岭方法将窗口中的图像分成若干区域,然后对每个区域 $r_k (k = 0, 1, 2, \dots)$ 计算其与其他区域 $r_i (i \neq k)$ 的颜色对比度值:

$$S(r_k) = \sum_{r_k \neq r_i} w(r_k) D(r_k, r_i). \quad (5)$$

式中: $w(r_k)$ 为区域 r_k 的权值,为该区域像素的个数占整个区域的权值(用于区分大区域与小区域); $D(*, *)$ 为 2 个区域的颜色距离度量:

$$D(r_1, r_2) = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} f(c_{1,i}) f(c_{2,j}) D(c_{1,i}, c_{2,j}). \quad (6)$$

其中: n_1, n_2 分别表示区域中的颜色总数; $f(c_{1,i})$ 和 $f(c_{2,j})$ 分别表示指定区域中第 i 和第 j 种颜色出现的概率; $D(c_{1,i}, c_{2,j})$ 表示不同区域之间的颜色距离.除了区域颜色距离,再引入区域空间距

离,将其定义为两个区域中心点之间的欧氏距离.

在定义单个区域的显著性值时,综合考虑了区域颜色距离和区域空间距离:

$$S(r_k) = \sum_{r_k \neq r_i} \exp(-\frac{D(r_k, r_i)}{\sigma_s^2}) w(r_i) D_s(r_k, r_i).$$
 (7)

式中: $D_s(r_k, r_i)$ 为区域 r_k, r_i 的空间距离; σ_s 为控制权值强度,越大则影响越小. 该区域中所有像素的显著性值都是 $S(r_k)$.

1.3 超像素分割

在面向多目标显著性区域提取的时候,对象估计方法建立的窗口尺寸往往太小,无法将图像中的待测目标完全包含,导致显著性目标的识别不完整,如图 2 所示. 本文利用超像素分割方法来解决这一问题,超像素分割就是将图像细分为多个超像素(图像子区域),这些超像素能够保留后续进行进一步图像处理的有效信息,并且不会破坏图像中物体的边界信息,因此可以解决本文因对象估计方法而导致的显著性目标的识别不完整问题. 常见的超像素分割方法有: simple linear iterative cluster (SLIC), 图论法, 分水岭法, meanshift 等. 本文在进行超像素分割时更注重图像的边缘检测效果和后期的合并^[11], 因此选择了 meanshift.

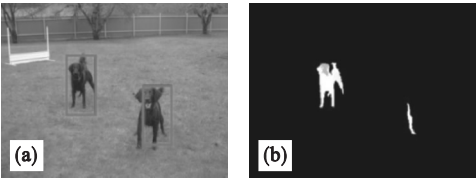


图 2 对象估计方法在面向多目标显著性区域提取时候的缺陷
Fig. 2 Defects in salient region extraction for multi-target by using object estimation method
(a)—原图; (b)—有缺陷的结果.

1.4 图像拼接

经过 meanshift 超像素分割后,原图像被划分为含有相似特征的像素块,每个像素块都有以下特性:不会跨越目标的边缘,要么完全在目标内部,要么完全属于背景^[12]. 因此,每个像素块应具有相同的显著性值.

对象估计方法可以标记出目标和背景,但存在目标区域缺失的问题. 在进行对象估计时,属于目标的像素块经常被窗口切割开,这使得窗口内部的部分被检测为目标,而窗口外部的部分被错认为是背景. 各个像素块和对象估计所得到的窗口之间的相互位置关系有 3 种:①像素块完全位

于窗口之内;②像素块完全位于窗口之外;③像素块一部分位于窗口之内,一部分位于窗口之外. 第三种情况正是需要弥补缺失的目标区域. 为此,本文提出一种非线性插值方法:

1) 首先对超像素分割得到的所有像素块进行编号.

2) 计算每一个像素块的平均显著性值:

$$I_a = \frac{\sum_{i=1}^N k_i}{N}.$$
 (8)

式中: N 表示像素块中的像素总数; k 表示各个像素的显著性值. 像素若位于窗口外部则其显著性值为 0,若位于窗口内部则其显著性值通过式(7)计算得到.

3) 对 I_a 进行非线性增益:

$$I = \begin{cases} 0, & I_a < 6; \\ 255 \times \frac{\lg(I_a + 1)}{\lg 256}, & I_a \geq 6. \end{cases}$$
 (9)

4) 将整个像素块中每个像素的显著性值都赋值为 I ,可有效弥补对象估计所缺失的目标区域.

5) 对图像拼接后的结果作二值化处理,可得到多目标显著性区域提取结果.

2 实 验

将本文算法与 3 种国际上较常见的显著性区域提取算法作比较,这 3 种算法分别为 HC 算法^[5-6]、RC 算法^[6]和 DRFI 算法^[7]. 比较过程分为定性对比和定量对比.

2.1 实验环境

- 1) 硬件环境. CPU: Intel(R) Core(TM) i3 - 3110M CPU@ 2.4 GHz; 内存: 6 GB.
- 2) 软件环境. 操作系统: Windows 7 旗舰版, 64 位; 编译器: Microsoft Visual Studio 2013, OpenCV2. 4. 13.
- 3) 数据集. PASCAL VOC 2007^[13].

2.2 定性对比

将本文算法、HC 算法、RC 算法、DRFI 算法的显著性区域检测结果通过人眼观察进行定性对比,对比过程中选择了 3 类图像: 普通的多目标图像、背景复杂的多目标图像及多个目标间颜色差异较大的图像.

1) 普通的多目标图像. 4 种算法的显著性区域检测结果如图 3 所示, HC 算法表现较差, RC 算法与 DRFI 基本能够正确检测全部目标, 正确区分背景和背景, 但 RC 算法依然还是存在较多

的背景区域被错误地检测到,DRFI 算法表现相对较好.而本文提出的算法基本没有错误地提取到背景区域,而且目标区域也赋予了较大的显著值,具有更多显著值为 255 的像素,因此相对检测效果更好.

2) 背景复杂的多目标图像.4 种算法的显著性区域检测结果如图 4 所示,HC 算法检测到的只是零散的亮点,没有一致高亮的区域;RC 算法表现也较差,几乎完全错误地划分了目标和背景区域;DRFI 算法尽管能大致检测出目标区域,但检测结果只能突出一小部分目标区域,较多的背景区域被检测到,因此效果较差;本文方法在背景复杂的多目标图像中,虽然不能完全正确分割出完整的物体轮廓,但清晰地提取出了目标区域,只有很少的背景区域被提取到,相对于其他 3 种方法具有明显优势.

3) 多个目标间颜色差异较大的图像.4 种算法的显著性区域检测结果如图 5 所示,HC 算法表现较差,无法正确区分目标和背景;RC 算法受

到目标间颜色差异的影响,只能检测到多目标中的 1 个目标,并且对于目标内部颜色特征变化较大的目标,检测效果尤其不佳,例如人脸和衣服;DRFI 算法也受到目标间颜色差异的影响,无法给出多个目标相似的显著性值,同时也存在无法检测到所有目标;本文算法能正确检测到所有目标,且能给各个目标区域赋予相似的显著性值,偶尔存在检测区域不完整的问题.

综上,针对多目标显著性区域提取,本文算法比 HC 算法、RC 算法、DRFI 算法提取到的显著图更好,尤其是在面向背景复杂的多目标图像及多个目标间颜色差异较大图像的时候,HC、RC、DRFI 3 种算法都难以正确提取显著性区域,而本文算法可以得到很好的结果.

2.3 定量对比

定性的评估算法具有较大的局限性,因人为的观察角度和思维方式都有很大区别.本文对 HC 算法、RC 算法、DRFI 算法及本文算法的多目标显著性区域提取结果作了定量对比.

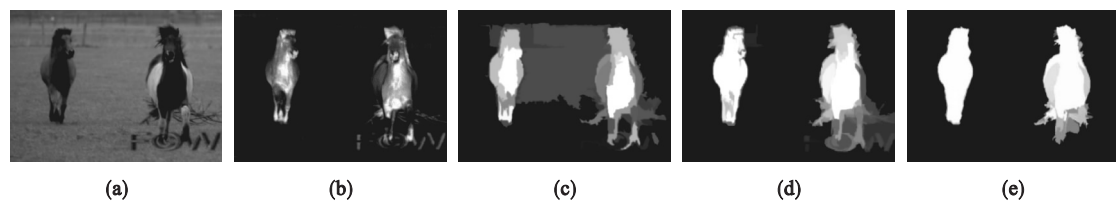


图 3 在面向普通的多目标图像时 4 种算法的显著性区域检测结果

Fig. 3 Salient region extraction results of four algorithms for multi-target image with regular background
(a)—原图; (b)—HC 结果; (c)—RC 结果; (d)—DRFI 结果; (e)—本文结果.

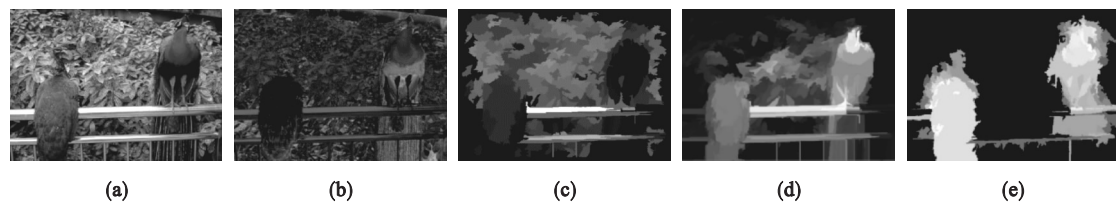


图 4 在面向背景复杂的多目标图像时 4 种算法的显著性区域检测结果

Fig. 4 Salient region extraction results of four algorithms for multi-target image with complicated background
(a)—原图; (b)—HC 结果; (c)—RC 结果; (d)—DRFI 结果; (e)—本文结果.



图 5 在面向多个目标间颜色差异较大的图像时 4 种算法的显著性区域检测结果

Fig. 5 Salient region extraction results of four algorithms for multi-target image with big differences between objects
(a)—原图; (b)—HC 结果; (c)—RC 结果; (d)—DRFI 结果; (e)—本文结果.

定量对比中,选择的指标为准确率和召回率, 并且以 VOC2007 数据集中手工标注的结果作为

金标准:

$$\left. \begin{aligned} \text{召回率} &= \frac{A}{A+C}, \\ \text{准确率} &= \frac{A}{A+B}. \end{aligned} \right\} \quad (10)$$

式中: A 表示算法检测结果和手工标注结果都认定为目标的像素集; B 表示算法检测结果认定为目标但手工标注结果认定为背景的像素集; C 表示算法检测结果认定为背景但手工标注结果认定为目标的像素集.

根据式 (10) 可得到各个算法的准确率 - 召回率曲线,如图 6 所示. 通过 4 种算法的定量对比,可以看出在面向多目标显著性区域提取的时候,本文算法要优于其他 3 种算法.

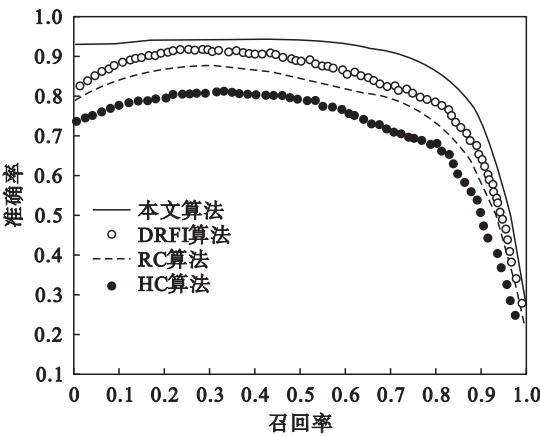


图 6 准确率 - 召回率曲线
Fig. 6 Accuracy rate-recall rate curves

3 结 论

本文结合对象估计方法和超像素分割方法,提出了面向多目标的显著性区域提取算法,该方法解决了其他算法无法从图像中提取多个显著性目标的问题. 定性、定量对比实验表明:当图像中存在多个显著性目标时,本文算法可准确提取到多个显著性区域,且提取结果要优于 HC 算法、RC 算法及 DRFI 算法.

参考文献:

[1] Liang Y, Yu J, Lang C Y, et al. Research on salient region extraction technology [J]. *Computer Science*, 2016 (3): 27 - 32.

[2] 陈美奂. 基于底层特征与高层先验的显著性区域检测算法 [D]. 秦皇岛:燕山大学,2015:1 - 61.
(Chen Mei-huan. Salient region detection based on low-level features and high-prior [D]. Qinhuangdao: Yanshan University, 2015:1 - 61.)

[3] Jing H Y, Han Q, Niu X M. Survey of salient region detection algorithms [J]. *Intelligent Computer and Applications*, 2014(1):38 - 39, 44.

[4] Itti L, Koch C. Computational modeling of visual attention [J]. *Nature Reviews Neuroscience*, 2001, 2(3):194 - 230.

[5] Cheng M M, Zhang G X, Mitra N J, et al. Global contrast based salient region detection [C]//IEEE Conference on Computer Vision and Pattern Recognition. Colorado Springs, 2011:409 - 416.

[6] Cheng M M, Mitra N J, Huang X, et al. Global contrast based salient region detection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(3):569 - 582.

[7] Jiang H, Wang J, Yuan Z, et al. Salient object detection: a discriminative regional feature integration approach [C]//IEEE Conference on Computer Vision and Pattern Recognition. Portland, 2013:2083 - 2090.

[8] Alexe B, Deselaers T, Ferrari V. What is an object [C]//IEEE Conference on Computer Vision and Pattern Recognition. San Francisco, 2010:73 - 80.

[9] Cheng M M, Zhang Z, Lin W Y, et al. BING: binarized normed gradients for objectness estimation at 300fps [C]//IEEE Conference on Computer Vision and Pattern Recognition. Columbus, 2014:3286 - 3293.

[10] 赵丹凤. 基于通用对象估计的目标检测与模糊车牌识别算法研究 [D]. 南京:南京邮电大学,2016:1 - 84.
(Zhao Dan-feng. Research on object detection based on objectness measure and fuzzy license plate recognition [D]. Nanjing: Nanjing University of Posts and Telecommunications, 2016:1 - 84.)

[11] 宋熙煜. 基于超像素的图像分割技术研究 [D]. 郑州:中国人民解放军信息工程大学,2015:1 - 74.
(Song Xi-yu. Research on image segmentation methods based on superpixel [D]. Zhengzhou: The PLA Information Engineering University, 2015:1 - 74.)

[12] Oh K, Lee M, Kim G, et al. Detection of multiple salient objects through the integration of estimated foreground clues [J]. *Image and Vision Computing*, 2016, 54:31 - 44.

[13] Li Y, Hou X, Koch C, et al. The secrets of salient object segmentation [C]//IEEE Conference on Computer Vision and Pattern Recognition. Columbus, 2014:280 - 287.