

一种自适应空间邻域的显著图获取方法研究*

李海洋^a, 何东健^{b†}

(西北农林科技大学 a. 信息工程学院; b. 机械与电子工程学院, 陕西 杨凌 712100)

摘要: 针对复杂背景下显著图提取精准度不高的问题,提出了一套基于自适应空间邻域的获取方案。该方案考虑人眼神经元感受野的同心圆结构,计算自适应圆形空间邻域;然后结合二维正态分布的显著权值计算空间邻域内每个像素点的显著值,获取图像的显著图,再利用简单的阈值分割算法提取二值图像;最后通过在两个自然图像集进行实验,并与三种经典算法进行比较。实验结果表明,该方法可以在复杂背景下有效地获取精确的显著图。

关键词: 显著图; 空间邻域; 人眼神经元; 正态分布; 阈值分割

中图分类号: TP393.4 **文献标志码:** A **文章编号:** 1001-3695(2013)06-1882-03

doi:10.3969/j.issn.1001-3695.2013.06.074

Salient region detection using adaptive circular spatial surround

LI Hai-yang^a, HE Dong-jian^{b†}

(a. College of Information Engineering, b. College of Mechanical & Electronic Engineering, Northwest A&F University, Yangling Shaanxi 712100, China)

Abstract: In order to improve the situation of low accuracy of image retrieval in complex backgrounds, this paper proposed a retrieval scheme based on adaptive circular spatial surround. Considering the concentric structure of the human eye reception field, it adopted circular spatial surround. Then it calculated the salient values of pixels in that surround through weights generated by a two-dimensional normal distribution. Finally, it employed a simple threshold segmentation process on the saliency to generate a binary image. Experiments were carried out on 2 natural image sets. The results show that, compared with three state-of-the-art salient region detection methods, the proposed method gives improved accuracy in acquiring the saliency.

Key words: saliency; spatial surround; human eye reception field; normal distribution; threshold segmentation

人们观察图像的时候,对图像中的目标对象的关注度不同,关注度高的目标对象作为图像的显著图,显著图在图像压缩、目标检测、图像分类和高层语义理解中具有广泛的应用价值。

为高效率、精确地获取图像的显著图,国内外学者开展了大量的研究。Itti等人基于特征综合理论^[1,2]提出了第一个可计算的视觉注意力模型。依据显著目标的整体效应原则和中心偏好原则,邵俊明^[3]提出了一种基于Itti模型的区域生长获取显著图的方法,沈兰荪等人^[4]提出了一种基于Itti模型和进化规划的显著区域检测模型。但基于Itti模型的方法计算复杂度比较高,而且在复杂背景与噪声的干扰下,显著图的提取不够精确。Ma等人^[5]提出了一种基于分块局部对比度的方法来生成显著图,并通过模糊增长的方法获得图像的显著目标。Hou等人^[6]在基于傅里叶变换幅度残余谱的启发下,采用频域处理的方法获取图像的显著区域,但仅利用了图像的灰度特征,对色彩对比度不鲜明的图像和显著图边界的提取不太准确。Achanta等人^[7]提出一种基于CIE Lab色彩空间的颜色、亮度特征,计算中心周边距的FSRD算法,对于显著图的边界检测较为精确,但对于显著目标较大的图像精确度较低。

为提高复杂背景下显著图的提取精度,本文在Achanta提出的FSRD模型基础上,研究结合人眼神经元感受野的同心圆

结构,改进了空间邻域的显著图检测模型,通过这种缩小邻域为合适大小的算法,弥补了在复杂背景、显著目标大的问题下的缺陷,以有效解决低效率、显著图边界模糊、计算复杂度高和显著图获取不精确等问题。

1 空间邻域模型

获取图像显著区域有基于图像特征、基于频域特征、基于生物视觉注意模型等多种方法,其中,空间邻域的方法对图像边界的分割效果较好,且计算效率高,效果精确^[7]。利用空间邻域提取显著图的模型如图1所示。

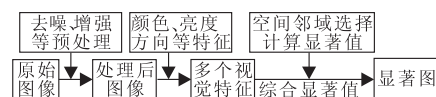


图1 空间邻域模型

本文依据空间邻域模型,a)由于外界环境中各种噪声及相机感光面的成像噪声以及采样和量化误差等众多因素的影响,必须对原始图像去噪处理,采取DoG滤波器选择人眼关注的频率范围;b)提取颜色、亮度人类视觉的基本特征;c)针对图像特征对图像每个像素点选取合适的空间邻域;d)对整合每个像素点的空间邻域,计算每个像素点的显著值;e)将图像

收稿日期: 2012-09-18; 修回日期: 2012-10-31 基金项目: 国家自然科学基金资助项目(60975007)

作者简介: 李海洋(1988-),女,黑龙江双鸭山人,硕士研究生,主要研究方向为图像分析和机器视觉;何东健(1957-),男(通信作者),陕西汉中人,教授,博导,博士,主要研究方向为图像分析与识别、智能化检测与控制及虚拟现实技术应用等(hdj168@nwsuaf.edu.cn)。

特征整合,从而获取图像的显著图。

2 自适应空间邻域的计算

邻域的选择是空间邻域法的关键技术之一,FSRD 算法将空间邻域范围设定为整幅图像,算法实现简单。对于一些显著目标较小和显著目标与背景区别明显的图像,显著区域获取效果较好,但在复杂背景、显著目标较大的条件下,提取的显著图不完整。

针对上述问题,本文结合神经元感受野的同心圆结构^[8],将空间邻域设定为圆形邻域,即针对每个像素点,将其邻域设定为以当前像素点为中心的圆形邻域。为了提高算法效率以及显著图精确度,首先设定两个自适应参数。针对一些背景复杂、显著目标较大的图像,显著图的边界与图像边界过于接近,从而导致显著图的边界提取不准确。为了弥补该问题,本文对距离图像边界过小的中心点,为邻域面积设定一个最大的自适应参数 $\gamma_{\min}$,并计算此时的半径为 $r_{\min}$;为提高算法效率,减少冗余计算量,本文为空间邻域设定一个最大的自适应参数 $\gamma_{\max}$,并计算此时的半径为 $r_{\max}$ 。

本文邻域选择方法如图2所示。将当前计算的像素点设为中心点,分别计算中心点与图像四个边界的距离,选择最小距离作为 r 的取值。根据 r 的取值与 $r_{\min}$ 和 $r_{\max}$ 的比较将中心点分为三类,即一般中心点、内部中心点和边界中心点。

对于一般中心点,邻域范围直接选取以中心点为圆心、半径为 r 的圆形邻域,如图2中的一般中心点空间邻域;对于内部中心点,则将圆形邻域缩小至半径为 $r_{\max}$ 的邻域,以有效避免不必要的计算,如图2中的内部中心点空间邻域;而对于边界中心点,将圆形邻域扩大为半径为 $r_{\min}$ 的圆形区域与整幅图像的交集,如图2中的外部中心点空间邻域。

计算邻域的方法如下:

$$r_{\text{cen}} = \begin{cases} r_{\min} & r < r_{\min} \\ r & r_{\min} \leq r \leq r_{\max} \\ r_{\max} & r_{\max} < r \end{cases} \quad (1)$$

$$r = \min(r_{\text{width}}, r_{\text{height}}) \quad (2)$$

$$r_{\text{width}} = \min(\text{cen}_w, \text{width} - \text{cen}_w) \quad (3)$$

$$r_{\text{height}} = \min(\text{cen}_h, \text{height} - \text{cen}_h) \quad (4)$$

$$r_{\min} = \sqrt{\text{area}_{\text{img}} \times \gamma_{\min} / \pi} \quad (5)$$

$$r_{\max} = \sqrt{\text{area}_{\text{img}} \times \gamma_{\max} / \pi} \quad (6)$$

其中: r_{cen} 为图像当前像素点的圆形邻域半径; r 为中心点与图像边界计算的最小距离; $r_{\min}$ 、 $r_{\max}$ 分别为设定的圆形邻域的最小和最大半径,用来判定是否需要修改 r_{cen} 的值; $\gamma_{\min}$ 、 $\gamma_{\max}$ 分别为设定的最小和最大的自适应参数,用来计算 $r_{\min}$ 、 $r_{\max}$ 的值,本文取 $\gamma_{\min} = 0.25$, $\gamma_{\max} = 0.9$ 。

3 显著权值的计算

人们拍摄图像的时候会自觉不自觉地目标物体放在图像中心区域,具有中心偏好原则^[3,10],且一个图像的显著区域具有连续性。Judd 等人^[9]由15人对LabelMe图像库选出1003幅图像进行实验,结果表明40%的注意焦点落在整幅图

像中心的11%区域内,70%的注意焦点落在图像的整幅图像中心的25%区域内;而实际人眼实验结果表明,60%的注意焦点落在整幅图像中心的5%区域内,90%的注意焦点落在整幅图像中心的20%区域内。结果均表明人眼对图像中心区域的关注高于周边区域,注意焦点分布图如图3所示。

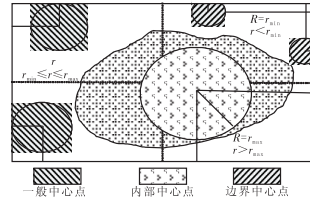


图2 空间邻域选择方法

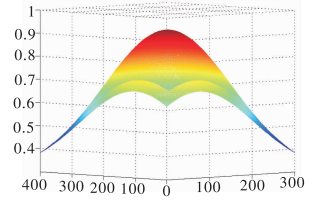


图3 注意焦点分布图

因此,本文通过将人眼对图像的注意焦点赋予一个符合二维正态分布的显著权值,其取值依照式(7)(8)。

$$\omega_{\text{cen}} = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\right) \times \omega_{\text{imp}} \quad (7)$$

$$\omega_{\text{imp}} = \frac{(\text{cen}_w - \text{mid}_w)^2}{\sigma_1^2} - \frac{2\rho(\text{cen}_w - \text{mid}_w)}{\sigma_1\sigma_2} \times \frac{(\text{cen}_h - \text{mid}_h)}{\sigma_1\sigma_2} + \frac{(\text{cen}_h - \text{mid}_h)^2}{\sigma_2^2} \quad (8)$$

对图像来说,吸引人类注意力的是图像特征的差异对比度^[6],也就是中心点与其空间邻域的特征值的差异,差异越大,则像素点的显著性越高,也就是该像素点在图像中的显著性越高,则其为显著图的概率越高。本文结合上述公式给出的权值得出式(9),重新计算图像中每个像素点的显著图。

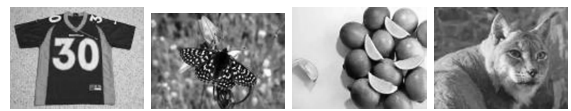
$$\text{Im } g_{\text{sal}} = \|\text{Im } g_{\text{cen}} - \text{Im } g_{\text{sur}}\| \times \omega_{\text{cen}} \quad (9)$$

$$\text{Im } g_{\text{sur}} = \frac{1}{\text{area}_{\text{cen}}} \sum_{\text{pos} \in \text{area}_{\text{cen}}} \text{Im } g(\text{pos}) \quad (10)$$

其中: $\text{Im } g_{\text{sal}}$ 为图像的显著值; $\text{Im } g_{\text{cen}}$ 为当前的像素值; $\text{Im } g_{\text{sur}}$ 为计算的空间邻域 area_{cen} 的像素平均值。

本文具体生成显著图的过程为:首先,将每个像素的特征值赋予对应的权值,得到新的显著值,然后对所得到的颜色亮度两个特征分别进行归一化合并操作,以保证各关注图中的稀疏的显著特征更加凸显^[1]。Itti 假设同一特征的显著性依靠DOG 竞争机制获得,而不同特征的显著性对最终显著图生成的贡献是独立、均等的,即每个关注图对最终显著图生成的贡献率相等。根据这个假设,最终显著图可用归一化后的颜色、亮度特征值线性综合计算如式(10)得到图像的显著图。

本文算法得到的显著图结果如图4所示。



(a) 原始图像



(b) 显著图

图4 提取出的显著图

4 实验结果及分析

实验采用 Hou 的62幅自然场景图像集 HOU^[6]和 Liu 建

立的 20 840 幅图像集 LIU^[11] 作为实验图像库。

实验方法如下:将每个图像库中的图像进行显著图提取,并与三种经典算法提取的显著图进行比较,图 5 是部分图像数据的实验结果。然后随机提取 LIU 图像库中的 1 000 幅图像,利用阈值分割获取图像的显著区域,再与图像的人工标注结果比较获得本文算法的精准程度,最后与三种经典的显著图获取算法的精准程度进行对比分析。

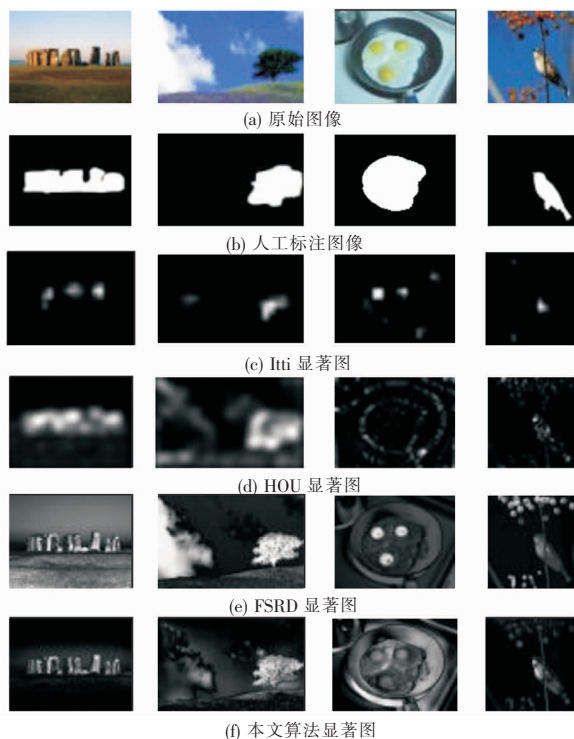


图 5 图像集上几种显著图结果比较

图像显著图获取的精准程度通常采用查全率与查准率的数值表示,通用的查全率 (precision)、查准率 (recall) 以及 F_α 公式^[7] 如式 (11) ~ (13) 所示。四种算法分别提取 LIU 中的 1 000 幅图像的显著图,采用相同的阈值分割^[12],提取出二值的显著目标图像,再与人工标注显著目标图像对比分析得到查全率、查准率的比较分析图 (图 6),确保查全率、查准率不受图像分割的影响。

$$\text{precision} = \frac{\sum_x g_x \alpha_x}{\sum_x \alpha_x} \quad (11)$$

$$\text{recall} = \frac{\sum_x g_x \alpha_x}{\sum_x g_x} \quad (12)$$

$$F_\alpha = \frac{(1 + \alpha) \times \text{precision} \times \text{recall}}{\alpha \times \text{precision} + \text{recall}} \quad (13)$$

上述公式中 x 为当前检测的图像的像素个数; α_x 为当前像素点是否属于检测到的显著图的二值参数, $\alpha_x \in \{0, 1\}$, 1 为显著图, 0 为非显著图; g_x 为经过实验验证计算得到的图像为真实显著图的概率, $g_x \in [0, 1]$; α 是非负常数参数,本文采用 $\alpha = 0.5$ 。

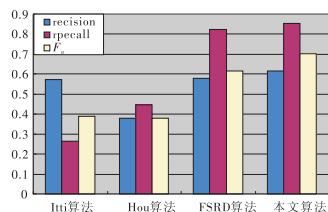


图 6 与其他算法的对比结果

图 6 是采用不同算法对 LIU 图像集中的 1 000 幅图像提取的显著图进行相同的阈值分割^[12]得到的二值图像进行综合评价的结果,分别计算出各个算法在 1 000 幅图像上的 precision、recall 和 F_α 的平均值。通过对比实验可知,查全率、查准率以及 F_α 的数值都要明显高于另外三种算法。观察发现,本文算法的精确度较好。

5 结 束 语

本文针对普通的显著图提取算法对图像背景复杂、显著目标较大的图像提取显著图精确度不高、显著图边界不明确等缺点,采用神经元感受野的同心圆结构的思想,将每个像素点的邻域范围设定为自适应的圆形邻域,然后利用自适应空间邻域的思想,对每个像素点的空间邻域自适应地调整。通过结合二维正态分布的显著权值计算每个像素点的显著值,最后综合各个特征生成显著图。实验结果表明,该算法获取的显著图具有很强的适用性和精确性。

参 考 文 献:

- [1] ITTI L, KOCH C, NIEBUR E. A model of saliency-based visual attention for rapid scene analysis[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 1998, 20(11): 1254-1259.
- [2] DAMIEN J, ITTI L, EGROIZARD P, et al. Left ventricle detection in radionuclide ventriculography by a model of neural network[C]//Proc of the 14th Annual International Conference on Engineering in Medicine and Biology Society. 1992: 994-995.
- [3] 邵俊明. 基于感兴趣区域的图像语义获取方法研究[D]. 杨凌: 西北农林科技大学, 2008.
- [4] 沈兰荪, 张菁, 李晓光. 图像检索与压缩域处理技术的研究[M]. 北京: 人民邮电出版社, 2008: 1-113.
- [5] MA Yu-fei, ZHANG Hong-jiang. Contrast-based image attention analysis by using fuzzy growing[C]//Proc of the 11th ACM International Conference on Multi-media. New York: ACM Press, 2003: 374-381.
- [6] HOU Xiao-di, ZHANG Li-qing. Saliency detection: a spectral residual approach[C]//Proc of IEEE Conference on Computer Vision and Pattern Recognition. 2007: 1-8.
- [7] ACHANTA R, HEMAMI S, ESTRADA F, et al. Frequency tuned saliency region detection[C]//Proc of IEEE Conference on Computer Vision and Pattern Recognition. 2009: 1597-1604.
- [8] 曾明, 孟庆浩, 王湘晖, 等. 视觉注意机制在图像增强中的应用研究[J]. 光子学报, 2009, 38(5): 1283-1287.
- [9] JUDD T, EHINGER K, DURAND F, et al. Learning to predict where humans look[C]//Proc of IEEE Conference on Computer Vision. 2009: 2106-2113.
- [10] 刘俊涛, 刘文予, 吴彩华, 等. 视觉注意原理局部特征的行人检测[J]. 中国图象图形学报, 2012, 17(3): 370-379.
- [11] LIU Tie, SUN Jian, ZHENG Nan-ning, et al. Learning to detect a salient object[C]//Proc of IEEE Conference on Computer Vision and Pattern Recognition. 2011: 353-367.
- [12] 何东健, 耿楠, 张义宽. 数字图像处理[M]. 西安: 西安电子科技大学出版社, 2008.