

doi: 10.3969/j.issn.1674-8425(z).2015.09.016

基于对比度和局部结构特征的显著性检测

曾祥鑫 李 颢 刘 坤

(国防科学技术大学 ATR 重点实验室 ,长沙 410073)

摘 要: 依据图像底层的颜色对比度特征 ,提出一种自底向上、数据驱动的视觉显著性检测方法 ,从全局对比度考虑 ,提取显著性目标的全分辨率显著图。首先通过对图像局部结构特征的分析得到关于目标和背景的先验分布信息;在分布信息的基础上分别提取图像的全局颜色对比度特征和空间位置关系特征 ,以空间关系权重优化显著性检测结果;进一步融合频域谱残差显著图 ,降低背景冗余及弱小显著目标对全局显著性检测结果的影响。在国际公开的显著性测试数据集 MSRA-1000 上进行实验 ,结果表明:该方法由于抑制了非显著区域的干扰 ,相对于现有的一些方法更能突出复杂背景下的显著目标。

关 键 词: 显著性; 结构特征; 图像对比度; 目标检测

中图分类号: TP391.4

文献标识码: A

文章编号: 1674-8425(2015)09-0093-05

Visual Saliency Detection Based on Contrast and Local Structure Feature

ZENG Xiang-xin , LI Biao , LIU Kun

(ATR Laboratory , National University of Defense Technology , Changsha 410073 , China)

Abstract: Based on the low-level features of the image color contrast , we proposed a bottom-up data driven algorithm of visual saliency detection. According to global contrast , our model generated a full-resolution saliency map. First of all , by analyzing the characteristics of the local structure of the image , the distribution prior knowledge of the object and the background was obtained. On the basis of the distribution information , we extracted the global color contrast and spatial position features that could be used to significantly optimize the detection performance. Furthermore , we integrated the Spectral Residual (SR) saliency map on the purpose of reducing the impact of background redundancy and small regions with high saliency. And due to the suppression of background , results of the experiments on public saliency detection database (MSRA-1000) showed that the presented method

收稿日期: 2015-04-30

基金项目: 国家自然科学基金资助项目(61103082)

作者简介: 曾祥鑫(1988—) ,男 ,硕士研究生 ,主要从事图像处理、目标识别方面的研究。

引用格式: 曾祥鑫 ,李颢 ,刘坤. 基于对比度和局部结构特征的显著性检测 [J]. 重庆理工大学学报: 自然科学版 ,2015 (9) : 93-97.

Citation format: ZENG Xiang-xin , LI Biao , LIU Kun. Visual Saliency Detection Based on Contrast and Local Structure Feature [J]. Journal of Chongqing University of Technology: Natural Science 2015(9) : 93-97.

could effectively highlight the salient object region in complex background compared to some existing methods.

Key words: visual saliency; structure feature; image contrast; object detection

视觉显著性的概念是在研究人类视觉感知机制的过程中提出的。针对人类视觉系统(human visual system, HVS)的研究发现,对于一幅图像,人眼可以快速地将图像中的目标和背景分离出来,并且通常对图像中部分区域的关注度较高^[1],这部分区域也称为感兴趣区域(regions of interest, ROI)。显著性检测的研究目的在于使机器可以模拟人类的视觉自动定位到图像中“显著”或“感兴趣”的目标对象,在图像自适应压缩、目标检测与分割、图像检索等多个方面都具有广泛的应用前景,是当前计算机视觉领域研究的一个热点。

视觉显著性的计算模型大致可以分为自底向上的数据驱动模型和自顶向下的任务驱动模型。任务驱动模型的思想是利用先验知识建模来计算显著性,以解决特定场景下的应用问题。在从底层特征出发的数据驱动模型中比较经典的有从生物视觉原理出发提出的Itti模型^[2],该模型通过提取图像在不同尺度下的颜色、亮度、方向特征,再根据特征融合策略生成最终的显著图。Hou等^[3]从频域出发,提出利用图像傅里叶变换后的幅度谱残差来计算显著度。文献[4]基于全局对比度,定义单个像素的显著度为该像素与图像中其他所有像素的距离总和,并采用直方图统计的方法优化计算。Cheng等^[5]在文献[4]的基础上进一步利用颜色信息,在RGB空间对颜色通道进行量化,以及采用颜色合并、平滑等操作来减少计算量,最后基于颜色直方图统计信息计算全局显著度。Achanta等在文献[6]中提出频率调谐方法,采用某个像素的颜色与平滑处理后的图像颜色均值之间的差异来定义该像素的显著性。在上述自底向上的显著性检测算法中,几乎都以目标和背景的差异为前提,从图像的局部或全局出发,以对比度来计算显著性。本文从基于全局对比度的方法^[4-8]中得到启发,结合对图像中显著目标的结构信息的先验知识分析,针对现有的算法进行改进,通过抑制背景的干扰和引入空间位置关系,取得了较好的检测效果。

1 显著性检测框架

1.1 基于局部结构特征的同质化区域检测

根据视觉特性,人眼通常对图像中的结构信息表现出较高的敏感性^[9]。图像中的显著部分也是人眼关注的感兴趣区域,其大多数情况下都具有较丰富的结构特征。因此,采用检测区域的结构特征也成为显著性检测的一种有效方式。Weickert等^[10]提出将结构张量(structure tensor)作为图像分析的有力工具,已成功应用到图像结构方向场计算、特征检测和图像去噪等领域。本文采用局部结构张量对图像局部几何结构特征进行分析。

图像梯度的变化可以有效地表明局部区域的内同质化程度。通常,图像中的背景部分较为平坦,同质化程度较高,显著度低,而目标部分结构化特征明显,显著度较高。采用基于图块的形式对区域的内同质性进行检测,考虑图像中的某一个像素 p 在该像素的 $N \times N$ 邻域计算局部梯度向量^[11]:

$$\varphi = \begin{bmatrix} \vdots & \vdots \\ \varphi_x(p) & \varphi_y(p) \\ \vdots & \vdots \end{bmatrix}, p \in N \times N \quad (1)$$

该图块的局部结构张量(也称为局部梯度协方差矩阵)定义为

$$T = G_\rho^* \begin{bmatrix} \sum \varphi_x^2(p) & \sum \varphi_x(p) \varphi_y(p) \\ \sum \varphi_y(p) \varphi_x(p) & \sum \varphi_y^2(p) \end{bmatrix} \quad (2)$$

式中: G_ρ 是方差为 ρ 的高斯函数,主要用来减少图像噪声对梯度运算的影响。

通过对局部结构张量矩阵的分析,可知矩阵的特征值能反映该区域结构的复杂度。在图像平坦区域,矩阵特征值 λ_1 和 λ_2 都较小,表明该区域的内同质化程度较高;在复杂结构区域中,特征值 λ_1 或 λ_2 相对较大,表明在相应特征方向上的灰度梯度变化较大。对图像进行分块预处理,通过设定

一个经验阈值 λ , 可以大致区分图像的结构区域和非结构区域^[12] 结果如图 1 所示。

$$S_{\text{mask}}(p) = \begin{cases} 1, & \max(\lambda_1, \lambda_2) \geq \lambda \\ 0, & \max(\lambda_1, \lambda_2) < \lambda \end{cases} \quad (3)$$



图 1 结构区域和非结构区域检测结果

1.2 基于全局颜色对比度的显著性计算

作为一种重要的底层特征, 颜色信息可以直观地描述自然图像的特征分布, 人们很容易注意到与背景形成鲜明颜色对比的区域。基于全局对比度的显著性检测模型可以对图像中相似的区域分配一个相近的显著性值, 均匀地突出目标。在图像结构化区域检测中, 通过局部结构张量分析得到了关于背景和前景目标的粗略分布信息。以此为基础, 对基于全局颜色对比度的显著性检测主要考虑两个方面: 根据已获得的背景先验知识, 以全局颜色对比度计算显著度; 根据显著性目标的空间分布信息, 一致性突出显著目标。

1.2.1 全局对比度计算

考虑到 CIELab 颜色模型与人眼视觉的相似性, 分别提取图像对应到该颜色空间上的 l 、 a 和 b 分量作为图像颜色特征值。与 Achanta 等提出的 FT 算法^[5] 利用图像整体平滑后的颜色均值作为参考值不同的是, 本文利用得到的目标和背景的先验信息生成背景掩膜, 计算背景部分的所有像素的颜色均值, 以图像中的所有像素与该均值的差值定义全局显著度:

$$d_{\text{color}}(p) = \|C_p - C_m\| \quad (4)$$

其中: $C_p = (l_p, a_p, b_p)$ 为第 p 个像素的颜色特征矢量; $C_m = (l_m, a_m, b_m)$ 为背景颜色均值矢量; $\|\cdot\|$ 表示欧式距离。

1.2.2 空间位置关系

在从全局计算每个像素的颜色特征相对于背

景颜色均值的对比度时, 还应注意到目标在图像中通常是集中分布的特性, 即需要考虑以目标的空间位置为参考的空间关系^[13]。同样, 根据背景掩膜的分布情况, 可以定义图像中目标区域的中心位置 $P_m = (x_m, y_m)$, x_m 和 y_m 为掩膜中前景目标的质心坐标。对其进行归一化, 以此定义空间位置关系度和空间关系权重系数(见图 2(b)):

$$d_{\text{position}}(p) = \|P_p - P_m\| \quad (5)$$

$$w(p) = \exp\left(-\frac{1}{2\sigma d_{\text{position}}(p)}\right) \quad (6)$$

其中: σ 为空间关系权重因子, 用来衡量空间距离权重对显著度的影响程度。

计算全局对比度显著性, 在考虑颜色对比度对显著性贡献时, 以空间位置关系作为全局对比度显著性计算的权重参数, 定义基于颜色的全局对比度显著性为

$$S_{\text{global}}(p) = \log(1 + w(p) d_{\text{color}}(p)) \quad (7)$$

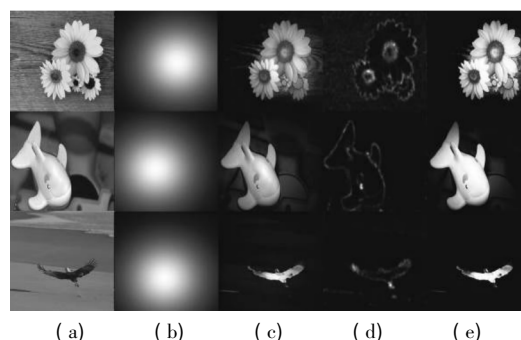


图 2 显著性检测各步骤对应的显著图

2 显著性模型的优化

全局颜色对比度的显著性检测由于是基于背景抑制的方式和通过参考空间分布的先验信息, 因而得到的显著图具有全分辨率, 更突出体现在具有较高对比度的目标上。通常由于目标颜色分布的均匀性, 目标处的显著度一致性较好, 对单一背景下的目标检测的结果更加验证了这一点(见图 2(c))。其不足之处是局部结构检测只得到关于目标和背景的粗略分布信息, 故在背景杂波较多、纹理复杂的情况下会在背景中对比度较高处也表现出较大的显著度。

文献[3]基于频域谱残差(SR)的方法,利用对灰度图像傅里叶变换后的幅度谱均值滤波计算残差谱,再结合相位谱逆变换生成显著图(见图2(d))。由于SR方法对冗余背景的滤波效果较好,均匀纹理的背景所呈现的显著值较低,故通过一定的规范化处理使其满足像素级的输出。综合考虑上述两种显著图的优势,可以采用下面的非线性融合方式对基于颜色对比度的显著图进行优化。

$$\begin{cases} S_1(p) = (\lambda_{11}S_{SR}(p) + \lambda_{12}S_{global}(p))^2 \\ S_2(p) = (\lambda_{21}S_{global}(p) + \lambda_{22}S_{SR}(p))^2 \\ S(p) = S_1(p) \exp(S_1(p) - S_2(p)) \end{cases} \quad (8)$$

式中: S_{SR} 和 S_{global} 分别代表基于谱残差法的显著性图像和基于全局颜色对比度的显著性图像。通过归一化处理,在像素级上完成两种显著性检测模型的优化整合。 $\lambda_{11} = 0.1$, $\lambda_{12} = 0.9$, $\lambda_{21} = 0.8$, $\lambda_{22} = 0.2$ 为多次实验得到的权值参数。 S_1 强调显著性目标的完整性, S_2 突出显著性目标的边缘。经过融合后的显著图(见图2(e))明显增强了显著性目标区域,抑制了背景以及可能存在的弱小显著目标。

3 实验结果与分析

使用Achanta等提出的公共数据集MSRA-1000^[6]对本文算法进行分析评价。该数据集共包含1000幅自然图像,均来自MSRA显著目标数据库,并且每幅图都提供了与显著区域相对应的人工标注的二值基准图像(ground truth,GT)。图3直观地给出了本文算法(Ours)与应用较多的几种显著性检测算法($SR^{[3]}$, $GB^{[14]}$, $AC^{[15]}$, $FT^{[6]}$, $HC^{[5]}$)在上述数据集上的检测比较结果,其中(a)和(b)分别为原始图像和GT人工分割图。可以看出:以上几种算法均能在一定程度上达到符合人眼主观感觉的显著性目标检测水平的要求,其中FT算法、HC算法与本文算法所检测的结果相对更优,并且可以看出本文算法在突出目标区域、抑制背景方面取得了更明显的效果。

为了客观地比较不同的显著性检测方法对高亮显著目标区域的效果,设定两组实验进行分析。实验1以准确率和召回率(precision-recall,PR)曲线作为衡量显著区域检测算法性能的评价指标。准确率反映了检测算法的有效性,即算法所检测到的显著性像素同时是人工标注显著区域像素与整个显著检测区域的比值。召回率则是人工标注的参考区域中能被正确检测的像素比值,体现了算法检测的完整性。通过设定固定阈值 $T \in [0, 255]$ 来二值化显著性图像,并绘制PR曲线。图4(a)给出了本文所提出的算法与其他几种算法的准确率和召回率的计算结果。实验2对显著图进行自适应阈值分割,设定自适应阈值 \bar{T} 为自身显著图均值的2倍,引入F-measure评价方法,即在综合平均准确率和平均召回率的基础上评价显著性区域的检测效果,结果见图4(b)。自适应分割阈值 \bar{T} 和F-measure指标的计算公式如下:

$$\bar{T} = \frac{2}{H \times W} \cdot \sum_{i=1}^H \sum_{j=1}^W S(i, j) \quad (9)$$

$$F_{\beta} = \frac{(1 + \beta^2) \text{precision} \times \text{recall}}{\beta^2 \times \text{precision} + \text{recall}} \quad (10)$$

其中: H 和 W 为显著图的高和宽; $S(i, j)$ 为坐标 (i, j) 处的像素显著值; β 为权衡平均准确率和平均召回率的影响程度,设定 $\beta^2 = 0.3$ 使得平均准确率的权重高于平均召回率。由图4可以看出:本文算法取得了较好的检测准确率和召回率,与本文算法思路较接近的FT算法相比性能提升明显。



图3 几种显著性检测算法显著图的直观比较

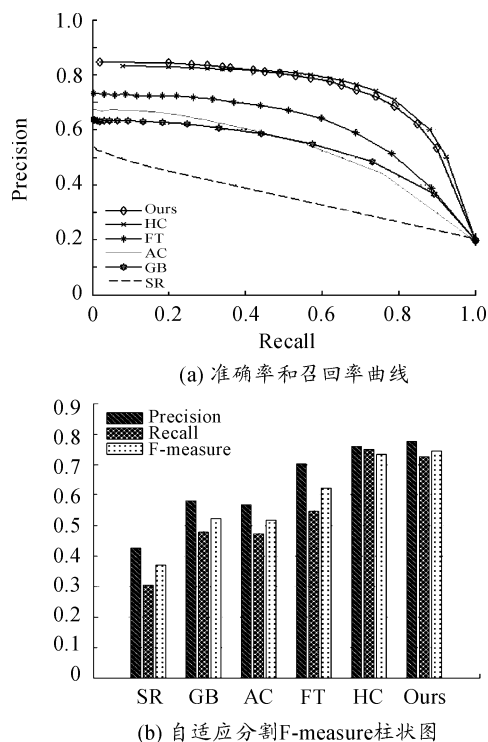


图4 显著性检测算法客观性能比较分析

4 结束语

通过对图像中显著目标的结构特征的分析,为基于图像全局颜色对比度特征的显著性检测模型提供了更多的先验信息,使检测结果更加突出目标区域,抑制背景的干扰和影响。本文算法在较复杂的自然场景下对显著目标定位准确,所检测的目标与背景的显著度差异明显。后续工作将进一步解决目标区域非均匀情况下的显著区域一致高亮的问题。

参考文献:

- [1] Jiang P, Ling H, Yu J et al. Salient Region Detection by UFO: Uniqueness, Focusness and Objectness [C]//IEEE International Conference on Computer Vision. USA: [s. n.], 2013: 1976 – 1983.
- [2] 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, 1998, 20(11): 1254 – 1259.
- [3] Hou X, Zhang L. Saliency detection: A spectral residual approach [C]//IEEE Conference on Computer Vision and Pattern Recognition. USA: [s. n.], 2007: 1 – 8.
- [4] Zhai Y, Shah M. Visual attention detection in video sequences using spatiotemporal cues [C]//ACM International Conference on Multimedia. USA: [s. n.], 2005: 815 – 824.
- [5] Cheng M, Zhang G, Mitra N et al. Global contrast based salient region detection [C]//IEEE Conference on Computer Vision and Pattern Recognition. USA: [s. n.], 2011: 409 – 416.
- [6] Achanta R, Hemami S, Estrada F et al. Frequency-tuned salient region detection [C]//IEEE Conference on Computer Vision and Pattern Recognition. USA: [s. n.], 2009: 1597 – 1604.
- [7] Perazzi F, Krahenbuhl P, Pritch Y et al. Saliency filters: Contrast based filtering for salient region detection [C]//IEEE Conference on Computer Vision and Pattern Recognition. USA: [s. n.], 2012: 733 – 740.
- [8] Jiang Z, Davis L S. Submodular salient region detection [C]//IEEE Conference on Computer Vision and Pattern Recognition. USA: [s. n.], 2013: 2043 – 2050.
- [9] 李崇飞, 高颖慧, 卢凯, 等. 基于结构相似度的视觉显著性检测方法[J]. 计算机工程与科学, 2013, 35(10): 181 – 185.
- [10] Weickert J. Anisotropic Diffusion in Image Processing [M]. Germany: Teubner, Stuttgart, 1988.
- [11] 邵宇, 孙富春, 刘莹. 基于局部结构张量的无参考型图像质量评价方法[J]. 电子与信息学报, 2012, 34(8): 1779 – 1785.
- [12] Junq C, Kim C. A unified spectral-domain approach for saliency detection and its application to automatic object segmentation [J]. IEEE Transactions on Image Processing, 2012, 21(3): 1272 – 1283.
- [13] Zhou L, Yang Z. Salient region detection based on spatial and background priors [C]//IEEE International Conference on Information and Automation. USA: [s. n.], 2014: 262 – 266.
- [14] Harel J, Koch C, Perona P. Graph-based visual saliency [J]. Advances in Neural Information Processing Systems, 2006, 19: 545 – 552.
- [15] Achanta R, Estrada F, Wils P et al. Salient region detection and segmentation [C]//IEEE Conference on Computer Vision Systems. USA: [s. n.], 2008: 66 – 75.

(责任编辑 杨黎丽)