

基于封闭先验的图像显著度检测*

陈加忠¹, 马丙鹏², 范晔斌¹, 李榕¹, 曹华¹

¹(华中科技大学 计算机科学与技术学院, 湖北 武汉 430074)

²(中国科学院大学 计算机与控制工程学院, 北京 100190)

通讯作者: 陈加忠, E-mail: chenjz70@163.com

摘要: 颜色对比度是图像关注区域检测的重要线索, 准确地提取反映图像不同颜色特征的区域, 非常有助于计算各个区域的对比度。为了得到有效的对比图, 首先利用封闭先验通过检测位平面中的连通性来提取具有不同颜色特征的封闭区域。其次, 利用背景先验消除与图像边界连通的封闭区域并得到封闭区域掩膜。然后利用对比度先验与封闭先验, 提出某区域在各个位平面中表现为封闭的次数越多越有可能是关注区域的假设, 并通过封闭区域掩膜的叠加计算各个封闭区域的对比度。同时, 结合人眼对小面积的封闭区域与封闭轮廓的感知特性, 以及对关注区域视觉资源的分配特性, 在获取对比度图的关键环节采取形态学滤波和高斯模糊, 最终实现面向凝视点估计的图像显著度检测。与多种经典的检测模型相比, 提出的方法取得了较好的性能。

关键词: 显著度检测; 对比度; 封闭先验; 位平面; 视觉凝视点

中文引用格式: 陈加忠, 马丙鹏, 范晔斌, 李榕, 曹华. 基于封闭先验的图像显著度检测. 软件学报, 2015, 26(Suppl. (2)): 208-217. <http://www.jos.org.cn/1000-9825/15031.htm>

英文引用格式: Chen JZ, Ma BP, Fang YB, Li R, Cao H. Image saliency detection based on closure prior. Ruan Jian Xue Bao/Journal of Software, 2015, 26(Suppl. (2)): 208-217 (in Chinese). <http://www.jos.org.cn/1000-9825/15031.htm>

Image Saliency Detection Based on Closure Prior

CHEN Jia-Zhong¹, MA Bing-Peng², FANG Ye-Bin¹, LI Rong¹, CAO Hua¹

¹(School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China)

²(School of Computer and Control Engineering, University of Chinese Academy Sciences, Beijing 100190, China)

Abstract: Color contrast is an important cue for image attention region detection. Extracting image regions that contain distinguishing color features is very helpful for computing the contrast of each image region. To obtain an efficient contrast map, the closure prior is firstly exploited to pick up the image regions containing distinguishing color features via connectivity detection in layered bit-planes. Secondly, the background prior is used to remove closed regions that touch image boundaries, and obtain closed region masks, in which the elements of closed regions are labeled with "1". Thirdly, a hypothesis, that a region should have big chance to be an attention region if it appears more times as a closed region in layered bit-planes, is proposed based on the contrast and closure priors. Further, the closed region masks of all bit-planes are accumulated to obtain the contrast of each connected region. Meanwhile, by taking account of the characteristics of human visual system with respect to the perception for small attention region, and visual resource allocation, several morphological filtering technologies are adopted to the key steps of contrast computing. Finally, the saliency map oriented to visual fixation estimation is generated. The experimental results show the presented detection method achieves acceptable performance compared with several state-of-the-art models.

Key words: saliency detection; contrast; closure prior; bit-plane; visual fixation

视觉关注是人类视觉系统 HVS 的一个决定性的特征, 而视觉显著度则把测量到的 HVS 感知的底层激励作

* 基金项目: 国家自然科学基金(61300140, 61402430)

收稿时间: 2015-05-15; 定稿时间: 2015-10-12

为视觉关注^[1]。已有的显著度检测模型大致可以分成两类:对比度先验和背景先验^[2]。对比度先验以 Itti 等人提出的 Center-Surround 方法为代表,在多个空间尺度上提取 Center-Surround 对比度来发现重要的区域^[3],它包含两个部分:特征的提取和特征的融合。特征的提取需要考虑图像中什么样的特征会吸引人的视觉关注。目前,颜色、方向、运动、闪烁等底层特征被认为最能吸引人的关注。这些特征通过 Center-Surround 机制进一步处理以提取对比度,从而得到关于单个特征分布的显著度图,再通过线性或者非线性融合得到最终的显著度图。林丽莉等人进一步把韦伯定律与 Center-Surround 机制结合起来,把中心区域和图像背景之间的全局对比度用于计算图像显著度^[4]。

对于自然图像而言,背景先验基于背景是连通的而图像边界又属于背景的假设。第 1 个假设来自于图像中反映真实世界中背景的表现特征,这意味着背景区域都比较大且是同构的,即背景中的大多数图像块可以很容易地相互连通。第 2 个假设来自于图像构造的基本准则,即大多数摄影者都不会让关注区域接触到图像的边界,因此大多数的图像边界属于背景^[2]。Yang 等人提出的基于流形排序的显著度检测是背景先验运用的一个成功例子^[5]。

轮廓是重要的图像底层特征,HVS通过目标的轮廓实现图像内容的感知与识别^[6]。吕肖庆等人发现一些图像区域的边界即便存在明显的缺口,但并不影响人眼感知到完整的轮廓^[7]。由位平面分解的实验可知,在图像的位平面中分布着形状、轮廓等视觉信息^[8]。图1中列出了4幅图像及其凝视点分布密度叠加图,可以看出视觉关注区域基本符合对比度先验和背景先验,即它们的背景是大片连通的,且与其他区域具有较大的对比度。还可以看出关注区域都是一些相对封闭的区域,如犬的嘴巴、熊猫的鼻子以及人脸。基于关注区域由相对封闭的区域组成的观察,以及人眼对封闭区域轮廓视觉感知的特性,在上述两个先验之外,本文提出封闭先验:如果某个区域是封闭的,则它可能是值得关注的区域,如果它在各个位平面中表现为封闭的次数越多,则越有可能是关注的。理由是在各个量化层的位平面中,如果某个区域表现为封闭区域的次数越多,则表明该区域和周围区域的对比度越大。



图1 热图表示的视觉关注区域,颜色越暖越受关注

基于上述3个先验,本文提出一种基于位平面中封闭区域提取的显著度检测方法。该方法先对图像做分层量化,以得到每个量化层位平面的二值图;再对每层的二值图做连通与邻接分析提取图像的封闭区域,并利用背景先验消除与边界连通的封闭区域以生成封闭区域掩膜;然后通过各量化层封闭区域掩膜的叠加,得到体现图像不同区域颜色特征的对比度。提出的方法由于合理利用了对比度先验、背景先验以及封闭先验,和近几年的多种经典的视觉关注区域检测方法相比,取得了较好的Shuffled AUC性能。

本文第1节描述如何利用上述3个先验得到图像的对比度图。第2节描述如何在得到对比度图的各个环节中施加形态学滤波。第3节为实验与分析。第4节作全文总结。

1 基于封闭先验的对比度提取

在本节中,封闭区域中的元素用 1 表示,非封闭区域中的元素用 0 表示,并把这样的二值图称作封闭区域掩膜。进而通过封闭区域掩膜的叠加,得到体现图像中各个区域不同颜色特征的对比度。基于封闭先验的对比度提取框架如图 2 所示。

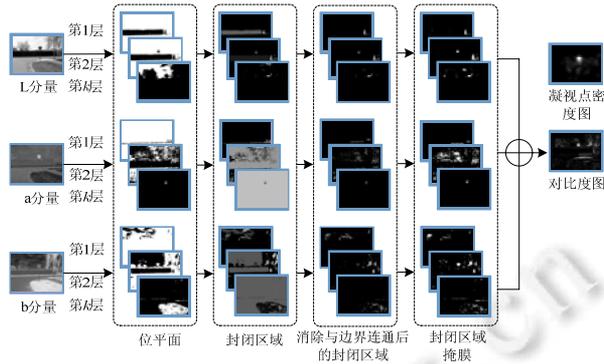


图 2 基于封闭先验的对比度计算框图

1.1 多层次量化得到位平面

在图像分层编码中,一般采用分层量化的方法得到分层的位平面,并让较大的量化步长对应的位平面优先编码和传输.在逐层解码中,随着较小量化步长得到位平面的叠加,重建的图像质量会增加,图像中各个区域的轮廓也会越来越清晰.基于这种经验,我们首先对图像各个颜色通道做分层量化,得到各量化层的位平面.在面向逐层编码的分层量化中,如果某像素值大于该层的阈值,则该像素值要减去该层的阈值,但在我们的应用中,并不做这样的处理,而是根据式(1)得到位平面.

$$y_i = \begin{cases} 1, & x_i > \delta_k \\ 0, & x_i \leq \delta_k \end{cases} \quad (1)$$

在式(1)中, δ_k 为第 k 个量化层的量化步长, x_i 为某像素值, y_i 为它在位平面上 b 的值.假设初始的量化步长为 Δ ,在图像分层编码中,后续量化层的量化步长一般选为 $\delta_k = 2^{-k+1}\Delta$, k 为大于 1 的递增整数.为了更加精细地提取像素在不同取值区间内的对比度,在我们的应用中,初始的量化步长较小,后续量化层的量化步长按 $\delta_k = k\Delta$ 选取.在实验中,首先把输入图像的 RGB 颜色分量依据最大-最小准则归一化到 0~1 之间,再转换到 Lab 颜色空间. Δ 取 8,若 Lab 各颜色分量的最大值和最小值之差为 100 左右,3 个颜色通道量化后可以得到大约 3×12 个位平面.

1.2 位平面中的封闭区域提取

位平面可以看成元素为 0 或 1 的多个集合.它具有邻接和连通的基本集合特征.假设 A 和 B 是位平面中的子集,如果 A 中至少有一点在 B 内,则说明 A 和 B 邻接.假设 S 是位平面中的一个子集, p 和 q 为 S 中的元素,如果从 p 到 q 存在一个全部点都在 S 内的路径,则认为 p 与 q 在 S 内是连通的.我们可以利用位平面的邻接和连通特征,提取位平面中的封闭区域.为了更好地提取封闭区域的轮廓,我们采用 8 邻点路径.在实验中,我们采用 Matlab 中的 bwlabel 函数. bwlabel 函数对位平面中的 0 元素不做处理,而对位平面中连通的 1 元素集合根据出现的顺序依次标号.因此,为提取位平面中 0 元素组成的封闭区域,需要对位平面取反,再调用 bwlabel 函数,获取位平面中 0 元素集合构成的封闭区域的标签.设位平面为 b ,根据式(2)得到由 1 元素和 0 元素构成的封闭区域的标签集合 l_1 与 l_0 :

$$\begin{cases} l_1 = \text{bwlabel}(b) \\ l_0 = \text{bwlabel}(1-b) \end{cases} \quad (2)$$

再通过 $l = l_0 + l_1$ 实现封闭区域提取,即 l 中既含有 1 元素构成的封闭区域的标签,也含有 0 元素构成的封闭区域的标签.为了避免 l_1 与 l_0 间的标签冲突,在求取 0 元素构成的封闭区域的标签后,需对 l_0 中的非 0 标签加上一个足够大的数如 $h \times w$, h 与 w 为输入图像的高与宽.如果实验环境为 OpenCV,可以调用 FloodFill 函数实现封闭区域的提取.

1.3 边界处封闭区域的消除

如果某个封闭区域中有元素处于图像的边界位置,根据背景先验,则认为该区域属于背景.在实验中,先采集标签集合在 4 个边界上的标签,然后让标签集合 l 中封闭区域的标签与边界上的标签作比较,如果相等,则封闭区域内所有的标签置 0,实现边界处封闭区域的消除.相比于其他方法对背景先验的利用,如文献[5]把处于边界位置的聚类标记为背景,文献[2]认为,离边界的距离越大,则越有可能是关注的区域,我们的方法只在某个量化层的位平面上利用背景先验,若某个区域在某个量化层上被判断为属于边界,但并不表示该区域在其它量化层上也被判断为边界的一部分.它是否确实属于边界或者属于关注区域,还需要看它在所有量化层上的表现.

1.4 封闭区域掩膜叠加得到对比度

根据封闭先验的假设,如果某个区域是封闭的,则有可能是值得关注的,而并不用关心这个区域出现的顺序.机器内部的像素是大于阈值还是小于阈值.因此对边界封闭区域消除后的标签集合 l 做进一步的处理,即如果 l 中的元素为非 0,则置 1,从而得到封闭区域掩膜 $m^{c,k}$:

$$m^{c,k} = \begin{cases} 1, & \text{label} \in l \geq 1 \\ 0, & \text{label} \in l = 0 \end{cases} \quad (3)$$

其中, c 为颜色通道的编号, k 为量化层的编号.如果某个图像区域的子集在多个量化层上的位平面表现为封闭区域,则说明它与周围图像内容具有不同的对比度特征.根据对比度先验,则认为该区域是值得关注的区域.因此,可以把所有颜色通道中所有量化层的封闭区域掩膜按式(4)叠加,来衡量图像某个区域和它周围图像区域之间的对比度,从而得到对比度图.从图 2 可以看出,经过上述几个步骤处理后,得到的对比度图能够在视觉凝视点密度最大的区域取较大的值.

$$Contrast = \sum_c \sum_k m^{c,k} \quad (4)$$

2 面向封闭区域的形态学滤波

采用面向封闭区域的形态学滤波的原因是:虽然图像的关注区域一般不与边界接触,但在各个量化层的位平面中,关注区域可能通过比较狭窄的缝隙与边界连通.为了提升一些尺寸比较小的关注对象亮度,需要对得到的封闭区域掩膜作二范数的归一化处理,但这种处理容易在面积较大的关注区域的突出处带来负面的额外提升.视觉凝视点除了集中在关注区域外,还会在关注区域邻近的区域分配少量的视觉凝视点.因此需要对第 1 节中的几个关键步骤的输出做相应的形态学滤波等处理.处理的强度与图像尺寸有密切的关系,因此假定图像较长的边为 400 个像素.

2.1 位平面的膨胀与腐蚀

虽然封闭区域总体上不会接触到图像的边界,但在各量化层的位平面中,关注区域有可能接触到图像的边界,从而被当成是边界处的封闭区域而被消除,以致降低关注区域对封闭区域掩膜的贡献.为了缓解这种现象的影响,在提取封闭区域之前,对位平面做膨胀与腐蚀操作:膨胀可以填补位平面中关注区域与边界区域之间连通的细缝,或者可以使接近边界却又与边界断开的非关注区域与边界连通;腐蚀可以消除关注区域边界上的小突起,并降低极小面积的非关注区域对视觉关注的贡献.如实验环境为 Matlab 或者 OpenCV,可以直接调用 `imdilate` 与 `imerode` 函数实现这两个操作.膨胀和腐蚀操作的模式都是 `disk` 型,膨胀半径 $r_{d1} = 3$.腐蚀在膨胀后执行,半径 $r_{e1} = 1$.从图 3(b)第 1 行中可以看出,膨胀操作使得左边的树枝和边界的树枝连通,从而被判断为背景.膨胀填补了第 2 行中红辣椒边沿区域在某些位平面中与边界连通的细缝,使得图 3(c)中的红辣椒区域比图 3(b)更亮.而腐蚀使得红辣椒区域的右下突出处变暗,如图 3(d)所示.

2.2 封闭区域掩膜的腐蚀与膨胀

有些图像中关注区域面积较小,这些小面积的关注区域在后续的高斯模糊处理中很容易被平滑掉.因此需要针对这种情况做膨胀以增加关注区域的面积,并做二范数的归一化处理提升关注区域的亮度.二范数归一化

的方法如式(5):

$$\begin{aligned} m_r &= \text{Normalize}(m, \text{norm2}, \text{row}) \\ m_c &= \text{Normalize}(m, \text{norm2}, \text{column}) \end{aligned} \quad (5)$$

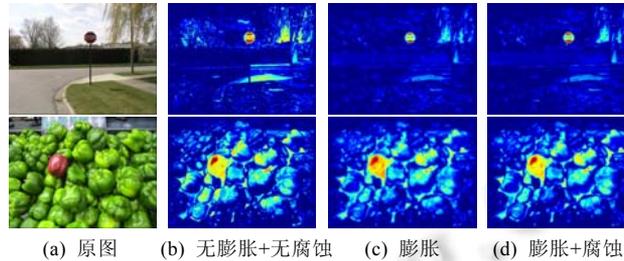


图3 位平面中有无膨胀与腐蚀操作的对比度图比较

在式(5)中, m 表示封闭区域掩膜, m_r 和 m_c 分别表示按列与按行归一化得到的两个矩阵.然后对 m_r 和 m_c 中的每个元素先取平方,再把结果相加,得到归一化后的封闭区域掩膜:

$$\bar{m} = m_r.^2 + m_c.^2 \quad (6)$$

在(6)式中, \bar{m} 表示归一化后的封闭区域掩膜,“ $.^2$ ”表示矩阵中的元素取平方.然而,对于面积较大的关注区域边沿上的小突起,会因为二范数归一化而特别亮,从而影响了最后的检测效果.因此,在膨胀与归一化之前需要对封闭区域掩膜做腐蚀处理来磨掉这些小突起.从图4第1行可以看出,虽然关注区域面积较小,但是对封闭区域掩膜做膨胀与归一化后,其亮度与面积得到了较大的提升,并有效地抑制了背景,而对对比度图基本上没有受到腐蚀的影响.图4的第2行,以水平方向观察钟的右上角,由于该区域关注像素较少,如图4(c)归一化后在钟的右上角带来了额外的亮度提升.而在膨胀和归一化之前施以腐蚀处理,可以缓解这种额外的亮度提升,如图4(d)所示.在本步骤中,腐蚀和膨胀操作的模式都是disk型,腐蚀半径 $r_{e2}=1$,膨胀半径 $r_{d2}=5$.

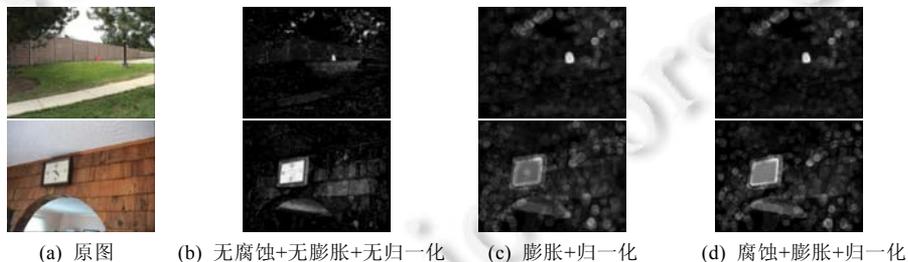


图4 封闭区域掩膜有无腐蚀操作的对比度图比较

2.3 对比度图的膨胀与高斯模糊

对于显著对象检测而言,希望能准确地分割关注对象,以满足计算机图形学等处理的需要.对于凝视点检测而言,人眼的视点不仅会落在关注的区域上,而且还会落在关注区域的附近.如果一幅图像内有两个以上的关注区域,这些关注区域之间还会有凝视点的过渡区域.换句话说,显著对象检测的计算主要由图像对象本身的底层特征决定,而面向凝视点估计的关注区域检测,还必须考虑人眼的视觉资源的分配特性.为了使得检测结果体现人眼的这种视觉资源分配特点,需要对对比度进一步做膨胀和高斯模糊处理.实际上,在视觉关注区域研究中作为参考的凝视点密度图,就是对采集离散的凝视点做高斯模糊后得到的,而这些离散的凝视点则是计算检测性能时的Ground truth.

本步骤中膨胀的模式为disk型,半径 $r_{d3}=3$.模糊处理采用标准方差为 $\sigma = (W + 1)/4$ 的高斯函数作为滤波器,其窗口大小为 $W \times W = 43 \times 43$.在Matlab环境中,把fspecial函数的滤波器类型参数设置为gaussian,并把边界延拓方式设置为symmetric.

高斯模糊后对矩阵中的每个元素取平方,然后对整个矩阵采用最小-最大值方式归一化到 0~255 之间,得到最后的关注区域检测图.此外,为了提高检测精度,在实验中不但对每层的位平面提取封闭区域掩膜,而且还对每层的位平面按 $1-b$ 取反,再提取封闭区域掩膜.在计算对比度图时,需要叠加上根据 $1-b$ 得到的封闭区域掩码.可以用算法 1 描述基于封闭先验的图像关注区域检测.

算法 1. 利用封闭先验的图像关注区域检测.

初始化: $\Delta = 8$; $b(1:h,1:w) = 0$; $Contrast(1:h,1:w) = 0$

for $ColorChannel = 1:3$

 计算当前颜色通道的最小值与最大值 $MinValue$, $MaxValue$

 for $Qstep = MinValue + 1 : \Delta : MaxValue$

$b(x(:, :, ColorChannel) > Qstep) = 1$; $ib = 1 - b$

$b = imdilate(b)$; $b = imerode(b)$; $m_b = getMask(b)$

$m_b = imerode(m_b)$; $m_b = imdilate(m_b)$

$ib = imdilate(ib)$; $ib = imerode(ib)$; $m_{ib} = getMask(ib)$

$m_{ib} = imerode(m_{ib})$; $m_{ib} = imdilate(m_{ib})$

$m_b = Normalize(m_b, norm2)$; $m_{ib} = Normalize(m_{ib}, norm2)$

$Contrast = Contrast + m_b + m_{ib}$

 end for

end for

$Contrast = imdilate(Contrast)$

$Contrast = GaussianBlurring(Contrast)$

$SaliencyMap = Normalize(Contrast)$

3 实验结果

实验针对 5 个数据集:MIT^[9]、Toronto^[10]、Kootstra^[11]、Cerf^[12]与 ImgSal^[13].MIT 含有 1 003 张图像,主题主要是室内外的日常生活.Toronto 含有 120 张图像,大多数图像中感兴趣的区域面积都较小,甚至有些关注区域并不处于镜头的聚焦处.Kootstra 含有 100 张图像,分成 6 个主题:动物、汽车、街道、建筑物、花与植物、自然景色.Cerf 含有 181 张图像,其主题为人脸、电话、玩具等.ImgSal 含有 235 张图像,主题涵盖风景、人类户外活动、动物、昆虫、鲜花、水果等.凝视点检测与显著对象度检测如 Cheng 等提出的 RC^[14]的目标并不一样,而提出的算法主要针对凝视点的检测,因此主要和其他 11 种经典的视觉凝视点检测算法做比较,它们是 Itti 等人提出的 IT^[3],Judd 等人提出的 LP^[9],Bruce 等人要提出的 AIM^[10],Li 等人提出的 HFT^[13],Zhang 等人提出的 BMS^[15],Harel 等人提出的 GBVS^[16],Goferman 等人提出的 CA^[17],Schauerte 提出的 $\Delta QDCT$ ^[18],Hou 等人提出的 SigSal^[19],Borji 等人提出的 LG^[20],以及 Li 等人提出的 AWS^[21].

3.1 实验指标的选择

当几条 ROC(receiver operating characteristic)曲线很接近以致难分高下时,就用 ROC 和纵横坐标轴之间的封闭区域的面积作为性能指标,即 AUC(area under ROC curve).然而,由于感兴趣区域靠近图像中心区域所造成的中心偏向,AUC 指标受 Edge effect 影响极大^[15].Zhang 等人指出,一个从图像中心发散的高斯函数状圆盘,在 Toronto 数据集上能取得 0.8 左右的 AUC 性能,超过了大多数的视觉凝视点检测模型^[22].GBVS 方法在找到大致的关注区域外,还在图像的大部分区域叠加上一层矩形状的薄雾,在 Toronto 数据集上的 AUC 性能达到 0.815,但输出并不能很好地体现人眼的关注机制.为了消除由“Edge effect”带来的干扰,Zhang 采用了改进的评估工艺来计算 AUC^[22],即 Shuffled AUC.因此,我们采用由 Borji^[20]提供的代码计算 Shuffled AUC(以下简称 S-AUC).

表 1 在不同数据集上取得最高 S-AUC 性能的参数

数据集	r_{d1}	r_{e1}	r_{e2}	r_{d2}	r_{d3}	W	S-AUC
Cerf	3	1	1	5	3	43	0.743 4
Toronto	3	1	1	8	5	43	0.728 1
ImgSal	3	1	1	8	3	43	0.700 7
MIT	3	1	1	5	3	43	0.694 0
Kootstra	3	1	1	8	7	43	0.625 5

3.2 实验参数的选取

实验中,当量化步长 $\Delta < 8$ 时,性能基本没有大的增长,而当 W 增加到 51 时,平均 S-AUC 可以提高约 0.000 1,但却会增加额外的计算量.因此,量化步长选 8, W 选 43.膨胀和腐蚀的半径对实验结果有较大的影响.对于不同的数据集,取得最优性能的关键参数见表 1.当所有数据集按照适合 ImgSal 数据集选择最优参数时,提出的方法在各个数据集上的平均 S-AUC 最高,达到 0.697 3,但此时 MIT 上的 S-AUC 为 0.691 9,低于 BMS 在该数据集上的性能.为了在所有的数据集上都取得比其他方法高的 S-AUC,我们采用适合 MIT 数据集的参数,此时各个数据集上的平均 S-AUC 为 0.697 0.

3.3 背景先验对性能的影响

从图 5 可以看出,由于对背景先验的合理利用,消除了和图像边界接触的位平面中的小的封闭区域,有效抑制了显著度图中背景亮度,并且提升了关注区域的亮度.如第 1 列图中的拱形门框、第 2 列图中纹理十分丰富的桌布、第 3 列图中的栏杆、第五列图中的灌木墙,都由于和边界连通而被准确判断成了背景.测试表明,背景先验能带来约 0.002 的 S-AUC 性能提升.

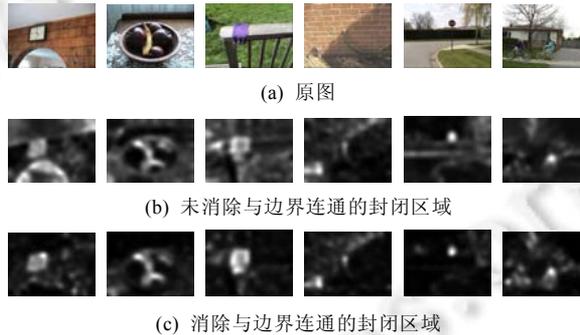


图 5 利用背景先验的显著度图比较

3.4 S-AUC性能比较

表 2 中列出了提出的方法和其他 11 种方法的 S-AUC 性能.为了公平的比较,我们仿照文献[15],对由不同检测方法得到的显著度图在测试 S-AUC 性能前用高斯函数做一个额外的模糊处理.用于不同方法得到的显著度图的高斯函数的标准方差各不相同,而且还会随着数据集的不同而变化.为了便于表达,该标准方差记作 σ_p .从表 2 中可以看出,所有的方法在 Cerf 数据集上都取得了较好的效果,除了 HFT、GBVS 与 IT 低于 0.7 外,其他方法都高于 0.7,而我们的方法达到了 0.743 4,高于第 2 名 BMS 近 0.08.对于 ImgSal 数据集,相比于 BMS 方法,提出的方法性能改善相对较小,只有 0.000 2.提出的方法在 5 个数据集上的平均 S-AUC 高于其他的 11 种方法,平均 S-AUC 比性能第 2 名 BMS 高出约 0.003,且在各个数据集上提出方法的 S-AUC 性能也均高于其它方法.需要指出的是,文献[15]提供的 BMS 取得的 S-AUC 性能并不准确,比如该文列出 BMS 在 ImgSal 与 MIT 数据集上的 S-AUC 分别达到了 0.771 2 与 0.701 7,而我们根据作者提供的显著度图的测试结果各为 0.700 2 与 0.693 0.

表 2 与 11 种不同方法在 5 个数据集上的 Shuffled AUC 性能比较

数据集	提出的	BMS	Δ QDCT	SigSal	AIM	LP	CA	AWS	LG	HFT	GBVS	IT
Cerf	0.743 4	0.735 5	0.728 2	0.728 0	0.724 2	0.714 6	0.714 0	0.709 1	0.702 9	0.699 2	0.680 3	0.676 1
σ_p	0.05	0.01	0.03	0.01	0.03	0.04	0.04	0.01	0.06	0.03	0.01	0.05
Toronto	0.724 0	0.720 6	0.716 3	0.705 6	0.689 8	0.689 5	0.695 5	0.711 8	0.698 4	0.689 5	0.639 2	0.657 3
σ_p	0.05	0.03	0.03	0.00	0.04	0.05	0.04	0.01	0.05	0.02	0.02	0.03
ImgSal	0.700 4	0.700 2	0.676 4	0.668 4	0.651 3	0.666 0	0.685 3	0.693 6	0.662 8	0.671 1	0.638 4	0.639 8
σ_p	0.06	0.08	0.10	0.10	0.10	0.10	0.06	0.08	0.12	0.10	0.06	0.10
MIT	0.694 0	0.693 0	0.680 8	0.675 6	0.678 7	0.672 6	0.680 3	0.691 5	0.686 8	0.660 6	0.651 8	0.655 9
σ_p	0.05	0.05	0.04	0.04	0.06	0.05	0.05	0.01	0.07	0.1	0.01	0.06
Kootstra	0.623 4	0.620 4	0.601 4	0.600 2	0.590 5	0.593 5	0.600 1	0.616 7	0.605 4	0.587 4	0.554 3	0.578 0
σ_p	0.03	0.00	0.00	0.00	0.01	0.03	0.03	0.00	0.04	0.01	0.01	0.00
Avg.	0.697 0	0.694 0	0.680 6	0.675 6	0.666 9	0.667 2	0.675 0	0.684 5	0.671 3	0.661 6	0.632 8	0.641 4

图 6 列出了 15 幅图像的显著度图,原图分别来自于以上 5 个数据集,每个数据集各取 3 幅.图 6 中的其他方法的显著度图由[15]提供.可以看出,提出的方法主要利用颜色对比度特征,而不考虑结构频率特征,因此不像 Δ QDCT 与 HFT 等易受背景中复杂纹理的干扰,如第 1、7、8 行.在列出的图像中,人脸的面积相对较小,如第 4、8、10 与 15 行,由于在提出的方法中采用形态学滤波与二范数归一化,人脸区域得到了有效的检测.提出的方法也不像 LG 那样受全局对比度的干扰,如第 4 行中的人脸,背景颜色与人脸颜色非常接近,采用全局对比度的 LG 很难有效检测到关注区域,但提出的方法由于对背景先验与封闭先验的合理利用,有效滤除了颜色与人脸接近的背景区域.

3.5 时间开销

在 2.50 GHz 的 Intel i5 4200M CPU 以及 4GB RAM 的机器上运行提出的算法,运行环境为 Matlab R2012a.对于图像维度为 681 \times 511 的 Toronto 数据集,每幅图像的平均处理时间为 2.906s.其中,分层量化占 0.024s,位平面的膨胀与腐蚀占 0.898s,提取封闭区域掩膜及归一化各占 0.351s 与 0.705s,封闭区域掩膜的腐蚀与膨胀占 0.872s,对比度图的膨胀与高斯模糊各占 0.01s 与 0.046s.

4 结束语

基于关注区域是封闭的假设,本文通过各量化层位平面的封闭区域掩膜叠加,有效利用了对比度先验.通过消除与边界连通的封闭区域,有效利用了背景先验.并结合人类视觉系统对封闭区域的面积与封闭区域轮廓的感知特点,以及人眼视觉资源的分配特点,灵活运用数学形态学滤波,较好地检测出了图像中的关注区域.实验结果也验证了本文提出的封闭先验的合理性与有效性.

图 6 中可以看出,当多个关注区域具有相同的底层特征时,人眼倾向于把有限的视觉资源分配于接近图像中央的关注区域.因此,把今后的工作可以针对具有多个关注区域的位置分布做加权处理,以便体现人眼的视觉资源分配的位置特性.此外,分层编码中的位平面和本文方法中的位平面具有确定的对应关系,通过分层编码中的位平面可以得到本文方法所需要的位平面,因此可以利用分层编码中的位平面实现视觉关注区域的提取,进而为重要区域的优先处理提供依据.

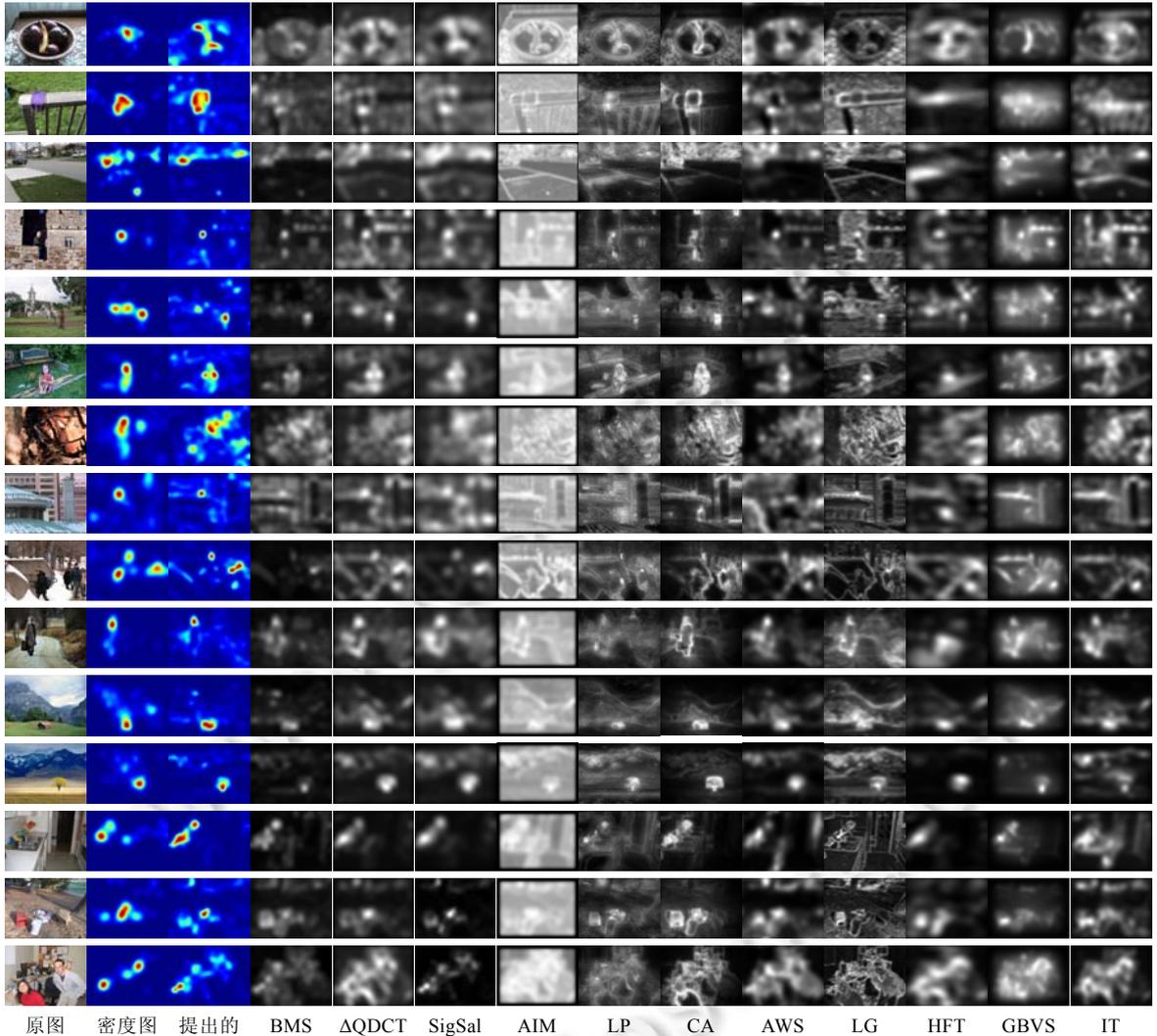


图 6 与 11 种经典方法得到的显著度图作比较

References:

- [1] Itti L, Koch C. Computational modeling of visual attention. *Nature Reviews Neuroscience*, 2001,2(3):194–203.
- [2] Wei YC, Wen F, Zhu WJ, Sun J. Geodesic saliency using background priors. In: *Proc. of the European Conf. on Computer Vision 2012*. 2012. 29–32.
- [3] Itti L, Koch C, Niebur E. A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1998,20(11):1254–1259.
- [4] Lin LL, Zhou WH. Image saliency detection model with Weber's law based on center-surround structure. *Journal of Image and Graphics*, 2012,17(10):1261–1267 (in Chinese with English abstract).
- [5] Yang C, Zhang LH, Lu HC, Ruan X, Yang MH. Saliency detection via graph-based manifold ranking. In: *Proc. of the 2013 IEEE Conf. on Computer Vision and Pattern Recognition*. 2013. 3166–3173.
- [6] Shi W, Zhu XF. Image retrieval based on contour reconstruction and feature point chord length. *Ruan Jian Xue Bao/Journal of Software*, 2014,25(7):1557–1569 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/4443.htm> [doi: 10.13328/j.cnki.jos.004443]
- [7] Lü XQ, Fu SP, Song JG, Gao LC, Tang Z. A gestalt-based method of envelope extraction for multi-object shapes. *Chinese Journal of Computers*, 2013,36(10):2070–2077 (in Chinese with English abstract).
- [8] Tian H, Yang SG. Color image retrieval algorithm based on significant bit-planes. *Journal of Computer-Aided Design and Computer Graphics*, 2010,22(2):279–286 (in Chinese with English abstract).

- [9] Judd T, Ehinger K, Durand F, Torralba A. Learning to predict where humans look. In: Proc. of the IEEE Conf. Computer Vision 2009. 2009. 2106–2113.
- [10] Bruce NDB, Tsotsos JK. Saliency based on information maximization. In: Proc. of the Neural Information Processing System 2005. 2005. 155–162.
- [11] Kootstra G, Nederveen A, De Boer B. Paying attention to symmetry. In: Proc. of the British Machine Vision Conf. 2008. 2008. 1–10.
- [12] Cerf M, Harel J, Einhäuser W, Koch C. Predicting human gaze using low-level saliency combined with face detection. In: Proc. of the Neural Information Processing System 2008. 2008. 1–8.
- [13] Li J, Levine MD, An XJ, Xu X, He HG. Visual saliency based on scale-space analysis in the frequency domain. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2013,35(4):996–1010.
- [14] Cheng MM, Zhang GX, Mitra NJ, Huang XL, Hu SM. Global contrast based salient region detection. Proc. of the IEEE Computer Vision and Pattern Recognition 2011. 2011. 409–416.
- [15] Zhang JM, Sclaroff S. Saliency detection: A Boolean map approach. In: Proc. of the IEEE Computer Vision and Pattern Recognition 2013. 2013. 153–160.
- [16] Harel J, Koch C, Perona P. Graph-Based visual saliency. In: Proc. of the Neural Information Processing System 2006. 2006. 545–552.
- [17] Goferman S, Zelnik-Manor L, Tal A. Context-Aware saliency detection. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2012,34(10):1915–1925.
- [18] Schauerte B, Stiefelwagen R. Quaternion-Based spectral saliency detection for eye fixation prediction. In: Proc. of the European Conf. on Computer Vision 2012. 2012. 116–129.
- [19] Hou X, Harel J, Koch C. Image signature: Highlighting sparse salient regions. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2012,34(1):194–201.
- [20] Borji A, Itti L. Exploiting local and global patch rarities for saliency detection. In: Proc. of the IEEE Computer Vision and Pattern Recognition 2012. 2012. 478–485.
- [21] García-Díaz A, Vidal X, Pardo X, Dosil R. Saliency from hierarchical adaptation through decorrelation and variance normalization. Image and Vision Computing, 2012,30:51–64.
- [22] Zhang L, Tong M, Marks T, Shan H, and Cottrell G. SUN: A Bayesian framework for saliency using natural statistics. Journal of Vision, 2008,8(7):1–20.

附中中文参考文献:

- [4] 林丽莉,周文晖.韦伯-中心环绕结构的图像显著性检测模型.中国图象图形学报,2012,17(10):1261–1267.
- [6] 师文,朱雪芳.基于轮廓重构和特征点弦长的图像检索.软件学报,2014,25(7):1557–1569. <http://www.jos.org.cn/1000-9825/4443.htm> [doi: 10.13328/j.cnki.jos.004443]
- [7] 吕肖庆,符松平,宋建国,高良才,汤帆.一种基于 Gestalt 原理的多连通分量形状外包络提取方法.计算机学报,2013,36(10):2070–2077.
- [8] 田宏,杨树刚.基于重要位平面的真彩色图像检索算法.计算机辅助设计与图形学学报,2010,22(2):279–286.



陈加忠(1970—),男,浙江嘉善人,博士,副教授,CCF 高级会员,主要研究领域为图像和视频处理,计算机视觉,多媒体数据分布式计算.



李榕(1971—),男,博士,讲师,主要研究领域为计算机视觉,图像与视频处理,多媒体通信.



马丙鹏(1978—),男,博士,副教授,主要研究领域为计算机视觉和模式识别,人脸识别,人体身份再确认,跨模态检索.



曹华(1971—),男,博士,讲师,主要研究领域为数字图像水印,视觉跟踪.



范晔斌(1976—),男,博士,讲师,主要研究领域为机器学习,移动环境中的视觉信号处理.