基于视觉注意机制的彩色图像显著性区域提取*

孟 琭
(东北大学 信息科学与工程学院, 沈阳 110004)

摘 要: 图像显著性区域提取是计算机视觉处理的重要步骤。结合人类视觉心理、生理模型,提出一种基于视觉注意机制的彩色图像显著性区域提取模型。通过改进的分水岭算法对彩色图像进行预分割,从而将原图像分成若干子区域,在此基础上运用提出的区域化空间注意力模型对各个子区域进行显著图计算,得到最终的显著性区域提取结果。实验结果表明,提出的显著性区域提取算法可以很好地从彩色图像中得到与视觉注意机制相一致的结果,且满足实时性要求,与传统方法相比,算法提取的区域更完整、更准确。

关键词: 显著性区域提取; 视觉注意机制; 分水岭; 区域化空间注意力模型
中图分类号: TP391.41 **文献标志码**: A **文章编号**: 1001-3695(2013)10-3159-03
doi:10.3969/j.issn.1001-3695.2013.10.069

Saliency detection for color image based on visual attention mechanism

MENG Lu
(*School of Information Science & Engineering*, *Northeastern University*, *Shenyang* 110004, *China*)

Abstract: Image saliency detection is an important step for computer vision processing. This paper presented a novel saliency detection algorithm for color image based on visual attention mechanism from human's psychology and physiology. Firstly, it used improved watershed segmentation method to divide the original image into several subdomains. Afterwards, it used regional spatial attention model to compute saliency map for each subdomain to get final saliency detection result. Experiment results show that the presented algorithm can accurately get saliency detection results in short time, and compared with traditional ones, the presented algorithm's results are more intact and more accurate.

Key words: saliency detection; visual attention mechanism; watershed; regional spatial attention model

0 引言

对于一幅图像来说,并不是所有的区域都可以使观看者产生同样的视觉或心理感受,不同的图像区域对人们所产生的感觉有着不同的贡献,人们的感受可能仅仅或很大程度上由图像的某一部分区域决定,从而使得人们在观看一幅图像的时候往往对图像的某一部分产生该区域很显著的感觉。人类视觉系统(human visual system, HVS)的这种处理特点被称为视觉注意机制,被关注的区域被称为注意焦点(focus of attention, FOA)。经由大量心理、生理学家的整理,视觉注意机制主要包括以下准则[1~5]:

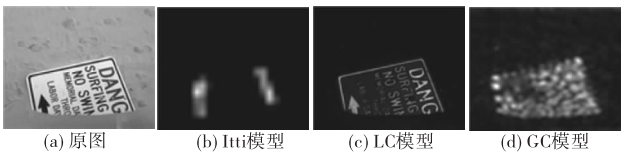
a)对局部突出的图像特征进行关注,如对比度、颜色、梯度等;

b)对全局中频繁、大量、反复出现的图像特征进行抑制,而对个别的突出特征进行关注;

c)视觉组织原则,即图像中某些部分之间的关系看上去比其他部分更加密切。

将 HVS 中的视觉注意机制引入到计算机图像分析过程中是非常必要的,它可以提供容易引起观察者注意的图像区域信息,帮助制定合理的计算资源分配方案,从而极大地提高现有图像分析系统的工作效率。显著性区域检测正是在这种思想的基础上提出并发展起来的[6]。

显著性区域提取算法由 Itti 等人于 1998 年提出[7,8],该方法提取多个低层视觉特征,如颜色、亮度、方向等,并将这些特征通过高斯金字塔和中心—周边算子合成为显著图,效果如图 1(b)所示。在此之后,该领域受到了国内外科研人员的持续关注,经过多年的研究,目前的显著性区域提取算法主要分为两类: local-feature contrast (LC) 模型和 global-feature contrast (GC)模型,而这两类方法的区别,主要源于对视觉注意机制所关注的侧重点的不同。文献[9,10]所提出的显著性区域提取算法(LC 模型),主要关注视觉注意机制的准则 a),其核心思想就是考虑图像中各个像素与邻域的对比,从而计算其显著图,效果如图 1(c)所示。文献[11,12]所提出的显著性区域提取算法(GC 模型),主要关注视觉注意机制的准则 b),其核心思想就是考虑图像中各个像素与所有像素的整体对比,从而计算其显著图,效果如图 1(d)所示。



(a) 原图 (b) Itti模型 (c) LC模型 (d) GC模型

图1 目前较热门的彩色图像显著性区域提取算法的运行效果

无论是 LC 模型还是 GC 模型,都仅仅关注了视觉注意机制中的某一条准则,而没有从整体上全面考虑,故此其提取的图像显著性区域难免有失偏颇。鉴于此,本文基于视觉注意机制的全部三条准则,提出一个区域化空间注意力模型,并结合分水岭算法对彩色图像进行显著图计算,进而提取显著性

收稿日期: 2012-12-24; **修回日期**: 2013-02-04 **基金项目**: 国家自然科学基金资助项目(61101057)
作者简介: 孟琭(1982-),男,辽宁沈阳人,博士,主要研究方向为模式识别、计算机视觉(menglu1982@gmail.com).

区域。

1　改进的分水岭算法

分水岭算法是一种基于数学形态学的图像分割方法，核心思想来源于地理学，其分割结果具有定位准确、区域封闭等优点，但该算法最大的缺点在于对图像的过分割问题。

本文首先利用分水岭算法对彩色图像进行预分割，从而将原图像分成若干子区域，以便于接下来应用本文提出的区域化空间注意力模型进行显著性区域提取。为此，本文对传统的分水岭算法作出改进，通过定义并计算子区域间的度量，从而对子区域进行合并来克服其过分割的问题；同时，也定义了中止规则来保证一定数量的子区域作为下一步的输入。例如，本文通过大量实验表明，对于分辨率 800 × 600 的彩色图像，通过分水岭得到的子区域数量在 30 ~ 80 个为最佳。

改进的彩色图像分水岭分割算法流程如图 2 所示。

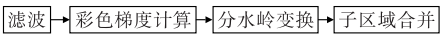


图2　改进的彩色图像分水岭分割算法流程

a) 应用高斯低通滤波器对原图像进行平滑处理，既可以达到去噪的效果，同时又可以实现对图像中无效细节的剔除，可以为后续的分水岭变换有效抑制过分割现象。

b) 运用下式计算彩色图像的梯度 $I(\theta)$。

$$I(\theta)=\left\{\frac{1}{2}\left[(g_{xx}+g_{yy})+(g_{xx}-g_{yy})\cos 2\theta+2g_{xy}\sin 2\theta\right]\right\}^{\frac{1}{2}} \tag{1}$$

其中：

$$\theta=\frac{1}{2}\arctan\left[\frac{2g_{xy}}{g_{xx}-g_{yy}}\right] \tag{2}$$

$$\begin{aligned} g_{xx}&=u\cdot u\\ g_{yy}&=v\cdot v\\ g_{xy}&=u\cdot v \end{aligned} \tag{3}$$

$$\begin{aligned} u&=\frac{\partial R}{\partial x}r+\frac{\partial G}{\partial x}g+\frac{\partial B}{\partial x}b\\ v&=\frac{\partial R}{\partial y}r+\frac{\partial G}{\partial y}g+\frac{\partial B}{\partial y}b \end{aligned} \tag{4}$$

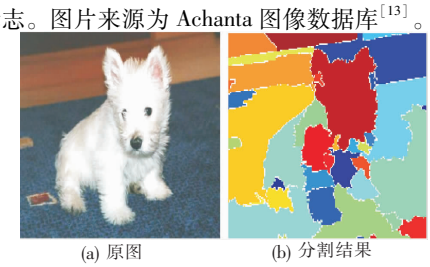
其中：r、g、b 分别表示彩色空间中 R、G、B 轴的单位向量；u、v 分别表示水平、垂直方向的彩色梯度。

c) 子区域合并规则主要依据各个子区域颜色均值向量间的欧式距离来确定。令 $d_{i,j}$ 表示子区域 i、j 间的距离，当某个 $d_{i,j}$ 小于指定阈值 T_x 时且子区域 i、j 相邻域时，将这两个子区域合并；子区域合并过程迭代进行，即对合并得到的中间结果进行进一步的子区域合并，直到所有 $d_{i,j}$ 都小于指定阈值 T_y（为防止图像中所有子区域合并为一个整体）或者迭代次数达到上限（为防止迭代过程陷入死循环）时，合并中止。

$$d_{i,j}=\|\mu_i-\mu_j\| \tag{5}$$

其中：μ_i、μ_j 分别表示子区域 i、j 的颜色均值向量。

本文分水岭算法的分割结果如图 3 所示，各个子区域通过伪彩色标志。图片来源为 Achanta 图像数据库[13]。



(a) 原图　　(b) 分割结果

图 3　本文分水岭算法分割结果

2　区域化空间注意力模型

经过分水岭算法的预分割后，原图像被分为若干子区域，这些子区域内的像素具备颜色相近、材质相似的特点。以分水岭分割结果为输入，基于视觉注意机制，本文提出一个区域化空间注意力模型，对各个子区域进行显著图的计算。从视觉注意机制的准则 a) 角度出发，取各个子区域内所有像素的 R、G、B 颜色通道数值，与其他所有子区域进行比较；从视觉注意机制的准则 b) 角度出发，取各个子区域的面积（即像素数），来表示某一图像特征（如颜色、材质等）在全局中反复出现的情况，并且该图像特征出现越频繁，则对其显著图抑制程度越大；视觉注意机制的准则 c)，已经通过分水岭算法中的子区域合并步骤得到了体现。同时，该模型还考虑了各个子区域间的距离因素，令距离近的子区域之间的影响较大，距离远的子区域之间的影响较小。上述即为本文基于视觉注意机制的彩色图像显著性区域提取算法的核心思想。

各个子区域的显著值通过式(6)计算，该值即为显著性区域提取的依据，即显著值越大的子区域，受到人类视觉关注的程度越大。

$$\mathrm{Sal}(R_i)=\sum_{R_i\neq R_j}D_{\mathrm{color}}(R_i,R_j)D_{\mathrm{space}}(R_i,R_j) \tag{6}$$

其中：R_i 表示第 i 个子区域；$D_{\mathrm{color}}(R_i,R_j)$ 表示子区域 i、j 之间色彩距离；$D_{\mathrm{space}}(R_i,R_j)$ 表示子区域 i、j 之间的空间距离。

$$D_{\mathrm{space}}(R_i,R_j)=\mathrm{e}^{\frac{\sigma_1}{d_1(f_i,f_j)}} \tag{7}$$

其中：$\sigma_1>0$，表示调节系数；f_i、f_j 分别表示子区域 i、j 的质心；$d_1(\cdot)$ 则用于计算两质心间的欧氏距离，该表达形式使得距离近的子区域之间的影响较大，距离远的子区域之间的影响较小。

$$D_{\mathrm{color}}(R_i,R_j)=\sum_{R_i\neq R_j}\omega(R_i)d_2(R_i,R_j) \tag{8}$$

其中：$\omega(R_i)$ 表示子区域 i 的面积系数，通过式(9)计算；$d_2(R_i,R_j)$ 表示子区域 i、j 之间色彩距离，通过式(10)计算。

$$\omega(R_i)=\mathrm{e}^{\frac{\sigma_2}{K_i}} \tag{9}$$

其中：$\sigma_2>0$，表示调节系数；K_i 表示子区域 i 内的像素数目，该表达形式使得子区域中图像特征出现得越频繁，则对该子区域的显著值抑制程度越大。

$$d_2(R_i,R_j)=\sum_{P_i\in R_i}\sum_{P_j\in R_j}\sum_{\mathrm{color}\in R,G,B}\|P_i-P_j\| \tag{10}$$

其中：P_i、P_j 分别表示子区域 i、j 中的任一像素；$\|\cdot\|$ 表示子区域 i、j 中两个像素间某一颜色通道的色彩距离，如图 4 所示。图中，横、纵坐标分别表示两个像素某一颜色通道的数值。由图可知，颜色差异越大，得到的色彩距离越大。该式分别在 R、G、B 三个颜色通道中进行色彩距离的计算并累加。

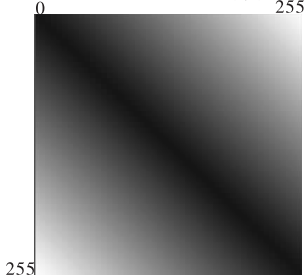


图4　颜色通道色彩距离示意图

3　实验结果与分析

实验中,将本文算法与 Itti 模型[7,8]、LC 模型[9,10]、GC 模型[11,12]的彩色图像感兴趣区域提取结果作比较。测试图片来源为 Achanta 图像数据库[13],这是近年来该领域最热门也是最受国内外同行认可的数据库之一,并且该数据库提供了准确的观察者手绘的图像显著性区域。本文将该手绘结果作为验证各个算法查准率(accuracy rate)、查全率(recall rate)的标准。

$$\text{accuracy rate} = \frac{T_p}{T_p + F_p}$$

$$\text{recall rate} = \frac{T_p}{T_p + F_n} \tag{11}$$

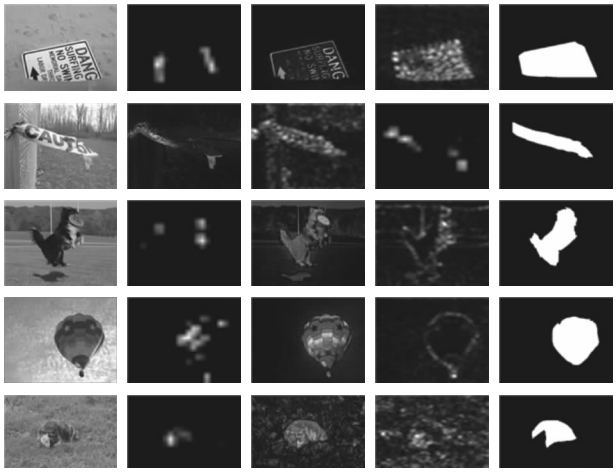
其中:T_p 表示真阳性,F_p 表示假阳性,F_n 表示假阴性。

首先比较本文算法与 Itti 模型、LC 模型、GC 模型的执行效率,其中本文算法由笔者编写,使用 C ++ 语言,其他三个模型的提取结果由 Cheng 提供[14,15]。Achanta 图像数据库共有 1 000张图片,图片最大分辨率为 400 × 400,最小分辨率为 400 × 224。分别使用四种模型对全部 1 000 张图片进行处理,并计算其平均执行时间,比较结果如表 1 所示。通过比较可知,本文算法可以实现彩色图像显著性区域提取的实时性要求。

表 1　四种图像显著性区域提取算法模型的执行效率对比

	模型			
	Itti	LC	GC	本文
平均执行时间/s	0.743	0.266	0.158	0.241

四种模型的彩色图像显著性区域提取结果如图 5 所示。通过比较可知,Itti 模型的提取结果仅为显著性区域的一部分,并不完整;LC 模型存在一些误提取的部分;GC 模型大多数关注了显著性区域的边缘;本文算法在显著性区域提取的完整性、准确性方面均优于以上三种模型。



(a) 原图　(b) Itti模型　(c) LC模型　(d) GC模型　(e) 本文模型

图5　Itti、LC、GC模型以及本文模型的显著性区域提取结果对比

图 6 给出了四种彩色图像显著性区域提取模型的 ROC 曲线对比,通过对比可知,本文模型优于其他三种。

4　结束语

本文基于人类视觉心理、生理模型,提出一种彩色图像显著性区域提取模型。通过对分水岭分割算法进行改进,并结合视觉注意机制的三条准则,设计了符合人类视觉特性的显著性区域提取模型。实验结果表明,本文提出的显著性区域提取模型可以很好地从彩色图像中得到与视觉注意机制相一致的结果,且满足实时性要求,与传统方法相比,本文算法提取的区域更完整、更准确。

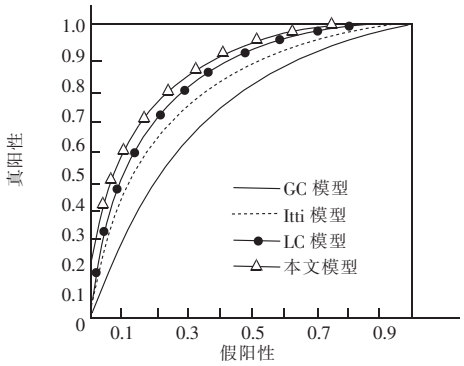


图 6　四种彩色图像显著性区域提取模型的 ROC 曲线对比

参考文献:

[1] KOCH C, POGGIO T. Predicting the visual world: silence is golden [J]. Nature Neuroscience, 1999, 2(1): 79-87.

[2] KOFFKA K. Principles of gestalt psychology [M]. London: Lund Humphries, 1999.

[3] TREISMAN A, GELADE G. A feature-integration theory of attention [J]. Cognitive Psychology, 1980, 12(1): 97-136.

[4] WOLFE J. A revised model of visual search [J]. Psychonomic Bulletin & Review, 1994, 1(2): 202-238.

[5] GOFERMAN S, ZELNIK-MANOR L, TAL A. Context-aware saliency detection [C]//Proc of IEEE Conference on Computer Vision and Pattern Recognition. 2010: 2376-2383.

[6] 张鹏,王润生. 静态图像中的感兴趣区域检测技术[J]. 中国图象图形学报, 2005, 10(2): 142-148.

[7] 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.

[8] ITTI L, KOCH C. Computational modelling of visual attention [J]. Nature Reviews Neuroscience, 2001, 2(3): 194-203.

[9] BRUCE N, TSOTSOS J. Saliency based on information maximization [C]//Advances in Neural Information Processing Systems. 2005.

[10] HAERL J, KOCH C, PERONA P. Graph-based visual saliency[C]// Advances in Neural Information Processing Systems. 2007.

[11] GUO Chen-lei, MA Qi, ZHANG Li-ming. Spatio-temporal saliency detection using phase spectrum of quaternion Fourier transform [C]// Proc of IEEE Conference on Computer Vision and Pattern Recognition. 2008: 1-8.

[12] HOU Xiao-di, ZHANG Li-qiang. Saliency detection: a spectral residual approach[C]//Proc of IEEE Conference on Computer Vision and Pattern Recognition. 2007: 1-8.

[13] ACHANTA R, HEMAMI S, ESTRADA F. Frequency-tuned salient region detection[C]//Proc of International Conference on Computer Vision and Pattern Recognition. 2009: 1597-1604.

[14] http://cg.cs.tsinghua.edu.cn/people/~cmm/[EB/OC].

[15] CHENG Ming-ming, ZHANG Guo-xin, MITRA N J, *et al*. Global contrast based salient region detection [C]//Proc of IEEE Conference on Computer Vision and Pattern Recognition. Washington DC: IEEE Computer Society, 2011: 409-416.