

·光电探测·

基于视觉显著图的目标探测

何 鹤, 刘 伟, 张 健

(63956部队, 北京 100093)

摘 要:从人类视觉机制出发,围绕方向、亮度和色彩等特征,介绍了计算机视觉显著性实现方式和视觉模型。成功通过人为和天然刺激相结合的方式,将该模型充分应用于目标探测。利用模型对数据库中包含车辆的背景图片进行观察实验发现,在75%的图像中,计算机模型要比观察员先发现目标。结果表明该模型更适用于显著目标的探测。

关键词:显著图;目标探测;WTA

中图分类号:TP391.41

文献标识码:A

文章编号:1673-1255(2012)03-0033-04

Target Detection Based on Visual Saliency Image

HE Hu, LIU Wei, ZHANG Jian

(63956 Army Unit, Beijing 100093, China)

Abstract: According to the visual mechanism of the human, the implementing method and the visual model of the visual saliency for the computer are introduced from the direction, brightness and color. This visual model can be used for the target detection by combining with the artificial and natural methods. The background images of the vehicles in the database are observed. The experiment shows that in 75% of the images, the computer model detects the target faster than the human, and this model is more suitable for detecting the obvious target.

Key words: saliency image; target detection; WTA

人的视觉机制主要有两种形式,即自底向上和自上而下形式。自底向上,这是一种原始快速的机制,使人对自己认为显著的目标有一定偏重;自上而下,是一种指导认知机制,会在意识的控制下关注焦点。文中在第一种自底向上注意模式的基础上,利用计算机模拟人类视觉系统结构。将亮度、方向和颜色等信息合并成单一的显著图。由于面对现实、杂乱图像时,视觉算法经常失效,因此文中研究了搜索算法,对高分辨率的图片进行目标搜索。结果表明,同观察员相比,搜索算法的性能优越。

1 模型

模型的输入是数字彩色图像,利用高斯金字塔生成不同的空间比例(金字塔通过低通滤波对输入图像进行二次抽样);高斯金字塔包含9个等级,提供

水平和垂直图像的换算系数:从1:1(等级0)到1:256(等级8)。利用中心-边缘(center-surround)算法计算每个特征,与视觉感受野相类似,采用细致的和粗糙的标度来实现:感受野中心的像素相当于金字塔比例 $C = \{2, 3, 4\}$, 而周围区域的比例为 $s = c + d$, $d = \{3, 4\}$, 这样每种特征产生了6幅特征图像。图1为模型实现过程框图。

1.1 初级视觉特征提取

利用 r, g, b 代表输入图像的红、绿和蓝色通道,亮度图像 $I = \frac{r+g+b}{3}$ 。利用高斯金字塔将亮度图像变为 $I(s)$ 。在最大亮度10%以上的区域,利用 I 对 r, g, b 通道进行规范化。为了将色调从亮度中独立出来,生成4种基本颜色通道(其中 Y 代表黄色通道)

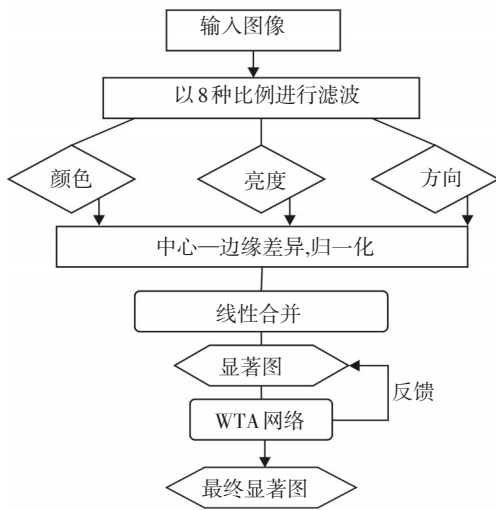


图1 模型实现过程框图

$$\begin{cases} R = r - \frac{g+b}{2} \\ G = g - \frac{r+b}{2} \\ B = b - \frac{r+g}{2} \\ Y = \frac{r+g}{2} - \frac{|r-g|}{2} - b \end{cases} \quad (1)$$

利用这4个颜色通道可以生成4种高斯金字塔颜色通道: $R(s)$, $G(s)$, $B(s)$, $Y(s)$; 利用高斯滤波从 0° 、 45° 、 90° 和 135° 对图像进行选择方向滤波。由于标准不同(中心采用细致标度和周围采用粗糙标度)使亮度对比、红绿色、蓝黄色和4个方向各产生6幅特征图像。一共产生42幅特征图像。

1.2 显著图像

显著图像的目的是计算图像可视区域每个位置的显著程度,并控制对当前位置的选择。

1.2.1 信息融合

将不同特征图像结合起来的难点是:它们代表着原始的不可比较的特征,有着不同的动态量程和提取机制。同时由于要将42幅图像进行合并,显著目标可能仅在几幅图像中非常明显,可能被其他大量图像噪声或次显著目标所掩盖。

前面给出了最简单的合并方案:将每幅特征图都归一化至固定的区间,然后将所有图像进行加法操作,这种方法在复杂背景显著目标识别中效果不是很明显。一种可行的改进方法是通过显著目标识

别案例学习,采用线性图像和权重的方式。这种方法的缺点是:当图像发生较大改动时,需要学习大量的目标识别案例。

文中提议一种简单易行的方法,全面推进包含小部分非常显著区域图像的显示,同时全面抑制大部分区域中包含类似峰值响应的特征图像。这种内部特征比较方式与电生理学中的抑制相互影响类似。

模型中的相互作用机理可以做如下概括:每幅特征图像首先归一化至一定的范围内(0~1),从而消除采用不同提取方式所产生的取决于特征的幅度差。然后将每幅图像进行DoG卷积操作。DoG滤波器的部分功能如图2所示,它在每个视觉位置都产生了强烈的影响,周围位置对该位置产生较强的抑制和抵抗。在每次反复中,给定的特征图像都会从先前注意的特征提取阶段收到输入数据(加入了DoG卷积的结果)。所有的消极因素都矫正至0,使迭代过程呈现高度非线性。此过程重复10次。

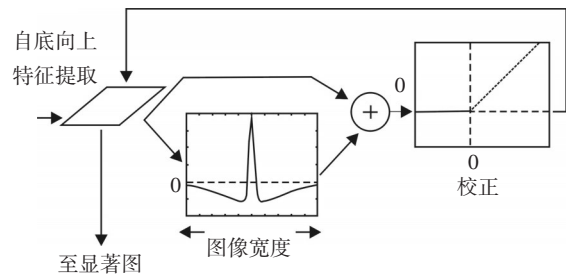


图2 DoG滤波

下面两个例子显示了这一过程的演化(如图3所示),并说明使用10次迭代能够使两幅图片产生足够的区别。正如所料,具有很多相似显著特征的图像由于相互影响而受到抑制,而拥有一个或几个最初强峰的图片得到增强。需要注意的是,此功能内的空间竞争机制,类似于一个“赢家通吃”(winner-take-all,简称WTA)的网络,它允许整个视觉场景内存在少数分布的赢家。

当一个(或几个)位置引起较强的反应时,它们会抑制比它们弱的其他地点;几次反复后的净结果是增强了最初的位置(S),抑制了较弱的位置。当没有显著位置时,所有地点发送和接收大约相同数量的抑制;这种情况下的最终结果是,逐步成为对所有地点的抑制(即全局抑制)。经过标准化、图像的亮度、颜色、方向特征映射归纳为3个独立的“显著图”,即亮度图、颜色图 and 方向图。

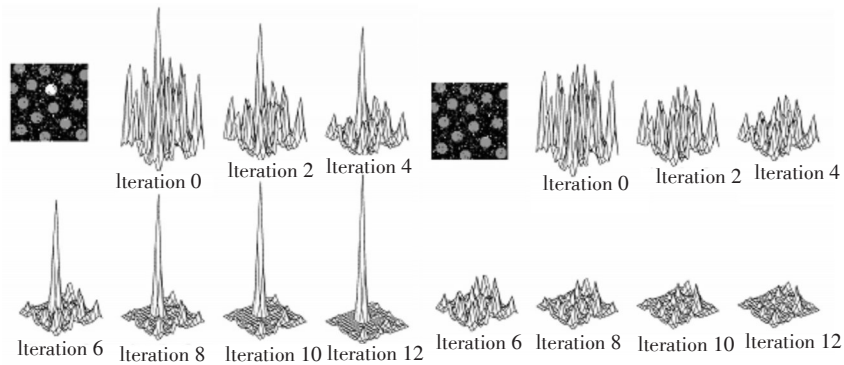


图3 迭代图例

1.2.2 WTA网络

根据定义,在任何时候,最大显著性的神经活动是在图像最突出的位置,即关注的焦点(FOA)。源于生物机理的赢家通吃(WTA)网络能够探测到FOA。WTA是一个比显著图更快的整合和预警2D层神经元。为了创建FOA的动态变化,而非每次稳定降低最突出的位置,有必要对显著图中目前的空间位置进行瞬时抑制。这也防止FOA立即回到以前所在的显著位置。人类也存在这样一个“返回抑制”机制。

由于没有建立基于自上而下机制的模型,FOA简单地代表一个圆盘,其半径是固定的,为输入图像长或宽(其中较小者)的1/12。选择时间常数、电导、模拟神经元的阈值,使FOA从一个突出的位置跳跃至下一个,仅需要30~70 ms(模拟时间),原来位置受到500~900 ms的抑制(这是一个生理物理学现象)。

2 结果分析

44幅原始图片和数据来自于DISSTAF(distributed interactive simulation, search and target acquisition fidelity)。每幅图的拍摄视角为 $6.9 \times 4.6^\circ$ 。每个场景都包含任务目标,观察距离从860~5 822 m。展示给观察者的图像大小为6 144×4 096像素。62位观察者年龄在18~45岁之间,视力为1.2以上。在测试之前,先让观察者对包含类似军事目标的图片进行观察和分析。图片展示过程中,当观察者发现目标后会马上按下按钮,并指出在哪个位置发现的目标(图像被划分成 10×10 的网格状)。

模型的输入也是同样大小的图片,但是却不需要进行预先实验。处理这些高分辨率的图像时,模型构建了3亿个模拟神经元。利用高速处理器处理每幅图的时间大概为15 min。关注焦点(FOA)为半

径340像素的圆。利用FOA覆盖全图需要123次移动,而随机搜索则需要61.5次。图4~图6展示了目标搜索的例子。



(a)原始图

(b)模型搜索图

图4 目标搜索(来自Search2)

图4a中人眼观察发现目标时间均值为2.8 s;图4b中模型在瞬时发现目标。对模型进行时间尺度划分,平均每秒钟都会有2~4次的注意力转移,考虑到人眼1.5 s的反应滞后,模型应该在2.2 s内发现目标。



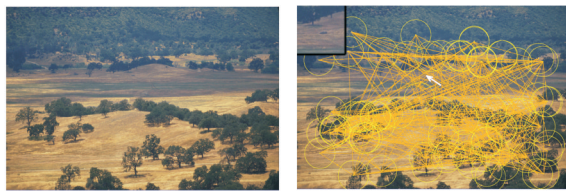
(a)原始图

(b)模型搜索图

图5 难以发现目标的例子(Search2数据集)

图5a中人类发现目标的平均时间为12.3 s;图5b为模型预测,由于背景对比度较低,目标要比建筑物的显著度低很多。模型开始序列搜索并发现目标,时间大概为4.9 s(时间尺度同图3)。

图6中观察员发现目标的平均时间为8.0 s;而该模型未能找到目标,其位置由白色箭头指出。仔细观察图片可以发现目标的特征尺寸与图像其他部分非常相似,如灌木和矮树。因此目标一点也不突出。



(a) 原始图 (b) 模型搜索图

图6 模型没有找到目标的图例

模型通过第一注意发现了44幅图片中的7幅;很快又在23幅图片中发现目标(FOA移动数不到20);在11幅图片中发现目标时所用FOA移动数超过20;在3幅图片中未发现目标。总体来说,模型在发现目标方面表现相当好,在很多搜索过程中FOA移动数远远小于61.5(除了6幅图像),但是这6幅图片中的目标相当小而且不显著,非常难发现。

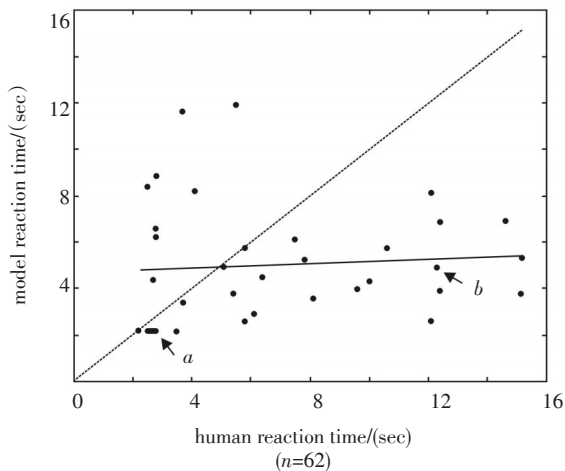


图7 观察员发现目标时间和文中算法的平均反应时间

下列分析生成如图7所示的曲线:首先,当模型在2000ms模拟时间(大概是40~80个FOA移动)内尚未发现目标时,或者超过一半的观察员未发现目标时,部分局外图像就会被丢弃,文中共有8幅图片被舍弃。模型模拟过程大概是每40ms移动一次,而人类是每秒钟移动3次,从而将模型的时间尺度和真实时间统一。另外由于人类的观察有时间滞后,约为1.5s。有了上述的标准化过程,在对36幅图像的分析中,人类和模型的最快反应时间均为2s左右,最慢为15s。图7中箭头a表示图4中的显著目标,而箭头b表示图5中较难发现的目标。

从图7中很难看到人类和模型在搜索时间方面的联系。然而,模型在发现目标方面要强于人的观

察,大概在75%的图像中,模型均比人要先发现目标。为了使模型发现较快的比例降到50%,必须假设人类每秒钟注视点的移动次数不多于两次,在快速搜索的情况下,这种速度有些不切实际的慢。

3 结 论

结果表明:相对简单的处理机制(基于早期视觉皮层架构的关键组织原则)与单一显著图相结合,在复杂场景和人为场景目标探测中效果非常好。

模型的关键是使用返回抑制和提取独立的显著性特征尺寸、及其经典搜索行为与人类心理物理学特征吻合。使用合理的模型,以人类的时间尺度,模型在36幅研究图像中的75%都比人类更快找到目标。这种优越性能的解释可能是:在自然场景的观察中,自上而下模式的影响发挥了重要作用。自上而下模式按照逐步构建整个场景的心理表征使人类注意力发生偏移。文中的模型,缺乏高层次的知识和以一个纯粹的自下而上的方式运作。这表明,某种情况下,这种高层次的知识会干扰最佳性能。例如,人类观察员经常趋向于观察道路或其他结构,或者有意识的决定彻底检查显著建筑物的周围地域,而车辆目标可能就在另一个领域或者在森林中。

虽然模型的最初设计不是用于检测车辆,研究结果表明,模型观察到的车辆目标相当“突出”,具有进一步的研究价值。

参考文献

- [1] Guo bei-yuan, fang wei-ning. Physiological fatigue determination based on percentage of eyelid closure with eye tracking system[J]. Chinese Journal of clinical Rehabilitation, 2005,9(7):26.
- [2] Itti L, Koch C, Niebur E. A model of saliency-based visual attention for rapid scene analysis[C]//IEEE Trans Patt Anal Mach Intel, 1998,20:1254-1262.
- [3] Sillito A M, Grieve K L, Jones H E, et al. Visual cortical mechanisms detecting focal orientation discontinuities[J]. Nature, 1995,378:492-497.
- [4] 靳薇,张建奇,张翔.基于视觉注意力模型的红外目标检测[J]. 红外技术, 2007(12):720-723.
- [5] 桑农,李正龙,张天序,等.人类视觉注意机制在目标检测中的应用[J]. 红外与激光工程, 2004,33(1):38-41.

(下转第58页)

照度均匀,效果也最好,因此在生产过程中采用此种阵列方式可以减少不必要的材料浪费,降低产品的生产成本,提高LED灯具的性价比,为LED照明的普及提供一个方面的借鉴。

参考文献

- [1] 公文礼.大功率LED在照明灯具设计中需要解决的问题[J].灯与照明,2008,12:21-24.
- [2] 郑代顺,钱可元.功率型白光LED研究进展[J].中国照明电器,2006(3):1-7.
- [3] 杨光.大功率发光二极管(LED)路灯配光方案[J].光源与照明,2009,12:17-21.
- [4] 张哲娟,杨介信,张燕萍,等.白光发光二极管及其集成光源模块的研究[J].半导体光电,2007,28(4):482-486.
- [5] 吴海彬,何素梅,王昌铃.多LED组合照明设计的关键技术[J].液晶与显示,2009(3):453-457.

(上接第32页)

4 结 论

以SPM探测器作为光子探测器件,对其光电探测特性进行分析,并搭建了光子探测实验系统,对SPM探测器的光子探测性能进行实验研究。从实验得出的数据可以看出,SPM探测器是一种性能优良的光子探测器件,与目前光电探测设备中所采用的探测器件(如PIN、APD等)相比较,在探测灵敏度方面具有极大的优势。随着光电探测领域研究的不断深入,可以预见这种器件将会在弱光探测领域发挥更大的作用。

参考文献

- [1] 刘伟.近红外单光子探测[J].物理,2010,39(12):825-831.
- [2] Amnon Yariv. Optical Electronics in Modern Communica-

(上接第36页)

- [6] Itti L, Gold C, Koch C. Visual attention and target detection in cluttered natural scenes[J]. Optical Engineering,2001,40(9):1784-1793.
- [7] Sun Y, Fisher R. Object-based visual attention for computer vision[J]. Artificial Intelligence,2003,146(1):77-123.
- [8] Itti L, Koch C. A comparison of feature combination strategies for saliency-based visual attention[C]//in SPIE Human Vision and Electronic Imaging IV, San Jose, CA, 1999 (in press).
- [9] Tsotsos J K, Culhane S M. Modeling visual-attention via se-

- [6] 夏勋力,余彬海,脉镇强.近朗伯光型LED透镜的光学设计[J].光电技术应用,2010,25(1):22-25.
- [7] D Wood. Optoelectronic semiconductor devices [M]. UK: Prentice Hall International, 1994.
- [8] Bahaa E A Saleh, Malvin Carl Teich. Fundamentals of Photonics [M]. US: John Wiley & Sons, Inc., 1991.
- [9] 胡海蕾,赖爱光,谢树森.基于道路照明的LED阵列光照度分布研究[J].照明工程学报,2009,20(1):78-80.
- [10] Reynolds George O, Develis John B, Thompson Brian J, et al. The new physical optics notebook [M]. San Francisco: SPIE Press, 1989.
- [11] M A Gennert, N Wittels, G L Leatherman. Uniform frontal illumination of planar surfaces: where to place the lamps [J]. Opt. Engng, 1993, 32: 1261-1271.
- [12] 杨光.照明灯具中白光LED阵列的排列及配光[J].灯与照明,2008,32(1):1-4.
- [13] 同济大学应用数学系.高等数学[M].5版.北京:高等教育出版社,2004.

tions[M]. Beijing: Publishing House of Electronics Industry, 2004.

- [3] 赵勋杰.光子计数成像原理及其应用[J].红外与激光工程,2003,32(1):42-45.
- [4] Dolgoshein B, Balagura, Buzhan P, et al. Status report on silicon photomultiplier development and its applications [J]. Nucl. Instr. and Meth. A.,2006,563:368-376.
- [5] SPM datasheet. <http://www.ams-technologies.com/>.
- [6] 赵帅.多像素光子计数器在单光子探测中的应用[J].光学精密工程,2011,19(5):972-976.
- [7] Cova S, Ghioni M, Lacaita M, et al. Avalanche photodiodes quenching circuits for single-photon detection[J]. Applied Optics,1996,35(12):1956-1976.
- [8] Gulinatti P, Maccagnani P, Rech I, et al. 35ps time resolution at room temperature with large area single photon avalanche diodes[J]. Electron.Lett.,2005,41:272-273.

lective tuning[C]//Artif Intel, 1995,78:507-551.

- [10] Barun J, Julesz B. Withdrawing attention at little or no cost: detection and discrimination tasks[C]//Percept Psycho, 1998,60:1-23.
- [11] Koch C, Ullman S. Shifts in selective visual attention: towards the underlying neural circuit[C]// Hum Neurobiol, 1985,4:219-235.
- [12] Treisman A. Features and objects: the fourteenth Barlett memorial lecture[C]//Q J Exp Psychos, 1998,40:201-237.