

文章编号:1000-582X(2009)06-691-06

视觉注意的图像目标预检测

曾孝平, 陈 礼, 刘国金, 谢春兰, 朱 斌

(重庆大学 通信工程学院, 重庆 400030)

摘要:提出了融合自底向上和自顶向下模式的视觉注意机制的改进算法。该算法使用以完整轮廓信息作为物体识别标志的方式,几乎不需要先验知识,改进了自顶向下算法的需要大量先验知识的弱点;同时由于将物体轮廓作为一个识别和预检测的重要参考信息,改进原视觉机制中的仅使用三个显著图,没有整体意识,从而致使识别范围过大的弱点。实验证明该方法提升了搜索预检测目标的准确度。

关键词:目标检测;视觉注意;显著图

中图分类号:TP301.6

文献标志码:A

Target pre-detection based on visual attention

ZENG Xiao-ping, CHEN Li, LIU Guo-jin, XIE Chun-lan, ZHU Bin

(College of Communication Engineering, Chongqing University, Chongqing 400030, P. R. China)

Abstract: An improved integration model combining the bottom-up and top-down cues for visual attention is proposed. It uses an integrated outline as the symbol for object recognition and overcomes the shortcomings of the top-down visual attention method that needs excessive a priori knowledge. Furthermore, adopting the object outline as the main character for recognition and pre-detection also overcomes the drawbacks existing in the bottom-up visual attention model that only uses three saliency maps from the original picture, resulting in an overextended recognition region. Experimental results show that compared to the original algorithm, the proposed algorithm is more effective for target pre-detection.

Key words: target detection; visual attention; saliency map

人类能从复杂背景中快速准确地识别目标,主要依赖于人类具有复杂的神经生理结构和功能。基于此原理, Koch, Itti 等人提出了视觉感知计算模型^[1-8],选择注意机制是视觉感知信息处理系统的一种主动策略,具有选择性、竞争性和定向性等特点,它与学习、记忆模块等协同工作,完成将注意目标从背景中分离、注意焦点在多个目标间转移,注意目标与记忆中的模式匹配等任务。该模型涉及到两种不同的处理模式,分别是自下而上的,包括与任务无关、不受意识支配、自底向上的注意(bottom-up

visual attention)和与任务相关、受意识控制、由顶向下的视觉注意(top-down visual attention)。自底向上的感知模型先注意到图像的局部特征,而自顶向下的感知模型先注意到图像的整体特征。目前,自底向上的视觉方法没有融合任何的上层知识,而大多数融合了上层知识的注意算法又难以实现。提出了基于视觉注意机制的两步处理模型,用以解决图像中的目标预检测问题。该模型综合利用两种模式的优点,将自底向上的注意机制和由顶向下的视觉注意结合起来。在模式处理的过程中,加入目标的

收稿日期:2009-02-11

基金项目:教育部归国人员启动基金资助项目(20050466)、重庆市科委自然科学基金资助项目(CSTC,2008BB2313)

作者简介:曾孝平(1956-),男,重庆大学博士生导师、教授,主要从事信号处理,通信信息处理方向研究,

(Tel)13668000086;(E-mail)axp@ccee.cqu.edu.cn.

欢迎访问重庆大学期刊社网 <http://qks.cqu.edu.cn>

轮廓信息,实现了目标的正确检测。

1 视觉注意计算模型

Itti 的自底向上的注意模型,主要基于 Treisman 的特征整合理论,从输入图像提取多方面的特征,如颜色、方向、亮度等,形成各个特征维上的显著图。然后对这些显著图进行分析、融合得到兴趣图。兴趣图中可能含有多个待注意的候选目标,通过竞争机制选出一个注意目标。可通过多次迭代选出感兴趣目标。

对于一幅输入图像,Itti 模型在图像的不同尺度,分别计算出颜色、亮度、方向上的显著性特征图,再用 DOG 函数计算出显著图。之后用胜者为王(winner-take-all, WTA)神经网络扫描由显著性图合成的重要图像,从中产生最重要的图像的地址并转化为图像坐标。最后用返回禁止信号,产生一个围绕在重要区域,半径可变的圆形区间,以后 WTA 神经网络迭代计算的时候就会寻找次重要的区域,而不是刚才已经找过的区域。图 1 是视觉模型示意图。

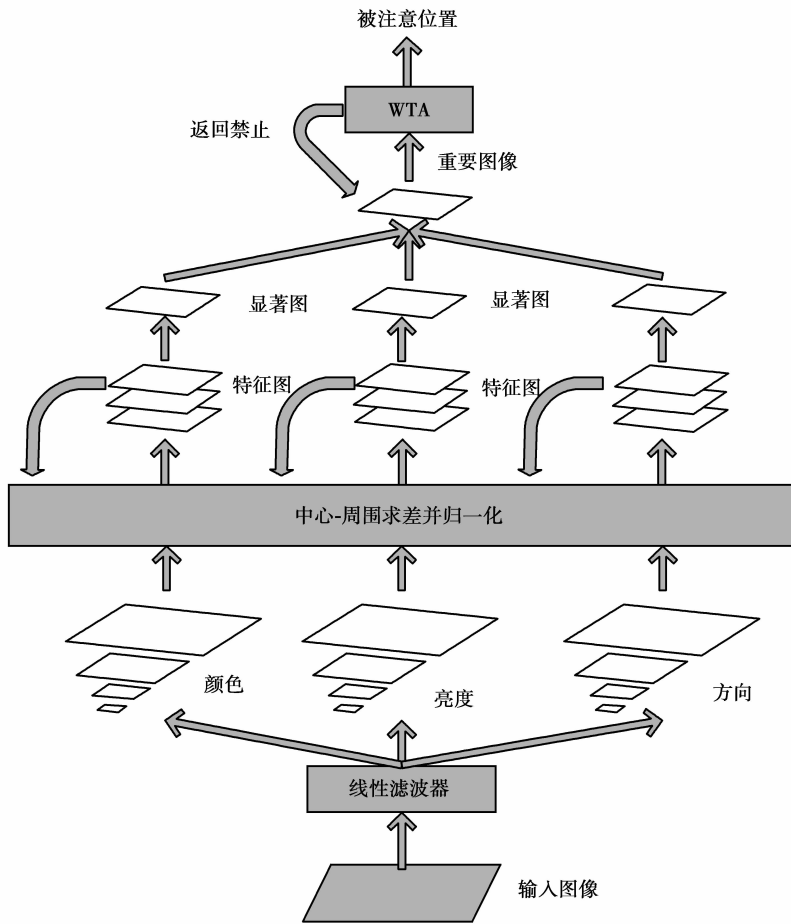


图 1 Itti 的自底向上的视觉注意模型

模型中分别对 3 种初级视觉特征进行处理。亮度没有方向性,使用对称的 Gaussian 金字塔(使用 Gaussian 滤波器)对其进行分解。亮度特征有 1 个量,它表示图像的亮度对比。对真彩色图像,则先用公式 $I=(r+g+b)/3$ 将其转换为灰度图像,其中 r 、 g 、 b 分别表示真彩色图像的红、绿、蓝通道。定义运算符 \triangleright 表示上述中心层和外周层之间的减操作,亮度对比映射图计算公式为 $I(c,s)=I(c)\triangleright I(s)$ 。

颜色也没有方向性,用对称的 Gaussian 金字塔处理。颜色特征有 2 个量,分别表示“红/绿”和“蓝/黄”色彩通道的对比。定义为 $M_{RG} = \frac{r-g}{\max(r,g,b)}$ 和 $M_{BY} = \frac{b-\min(r,g)}{\max(r,g,b)}$ 。为了避免在低亮度的情况下,颜色分量剧烈变化, M_{RG} 和 M_{BY} 在 $\max(r,g,b) < 1/10$ 时赋为 0,确保他们在 $[0,1]$ 的范围内。

模型采用 Gabor 金字塔处理方向特征(使用 Gabor 滤波器),该特征有 4 个量,表示 4 个局部方向 $\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ 对比。方向对比映射图计算公式为 $O(c, s, \theta) = O(c, \theta) \triangleright O(s, \theta)$ 。

目前视觉注意模型还存在如下缺点:不能从整体上注意全部目标,缺乏对图像中目标的整体性描述,如轮廓信息等。

针对此缺点,提出了一个改进的模型,综合利用颜色、幅度、方向、形状轮廓信息。在原来模型的基础上,提出综合模型,将与任务无关,不受意识支配、自底向上的注意模型和任务相关,受意识控制,自顶向下的视觉注意机制结合起来,融合上层注意的知识,在处理的过程中,加入目标的轮廓信息,对目标进行注意和识别。在原来的 3 个显著图的基础上,利用边缘检测算子提取图像中的轮廓信息,将此显著性图像中目标的整体信息加入自顶向下的机制中来,从而融合图像中局部和整体信息。

这种方法的好处有以下几方面:1)不需要具体的提供需要注意目标物的图片信息,或储存大量的物体照片的情况下,就可以提高注意正确率。2)模

型简单,自顶向下和自底向上的模型融合困难的原因之一就是 2 个模型都比较复杂,并且由于基于任务的自顶而下的方法大多需要训练的过程,因而对目标没有先验知识的情况下,会减少注意的成功率。3)减少扫描时间,减少了目标物的训练过程,仅仅以各个目标的外形轮廓代替,减少了算法复杂度,从而减少注意时间。模型如图 2 所示。

此模型将被初步注意到的图像范围(如图 2 中的 A 点)做适当增加,并将该增加的范围作为掩模模版 *Mask*,该模版为二值矩阵数据,注意范围内是 1,注意范围外为 0。图像作如下处理

$$I = Mask \cdot (I_{org} * G) + Mask \cdot I_{org},$$

其中 I_{org} 为原始图像; G 为高斯低通滤波器。

通过上式,对注意区域内的图像不做处理,对于范围外的图像,采用高斯低通滤波,对其进行虚化或柔化,弱化该部分的图像边缘信息。为保证 *Mask* 模版中被注意范围的边缘信息不要太突出,进而会被以后的处理识别出来,对注意边缘也进行弱化:

$$Edge(i_o + i, j_o + j) = I(i_o + i, j_o + j) * G, \\ i \in [-1, 1], j \in [-1, 1].$$

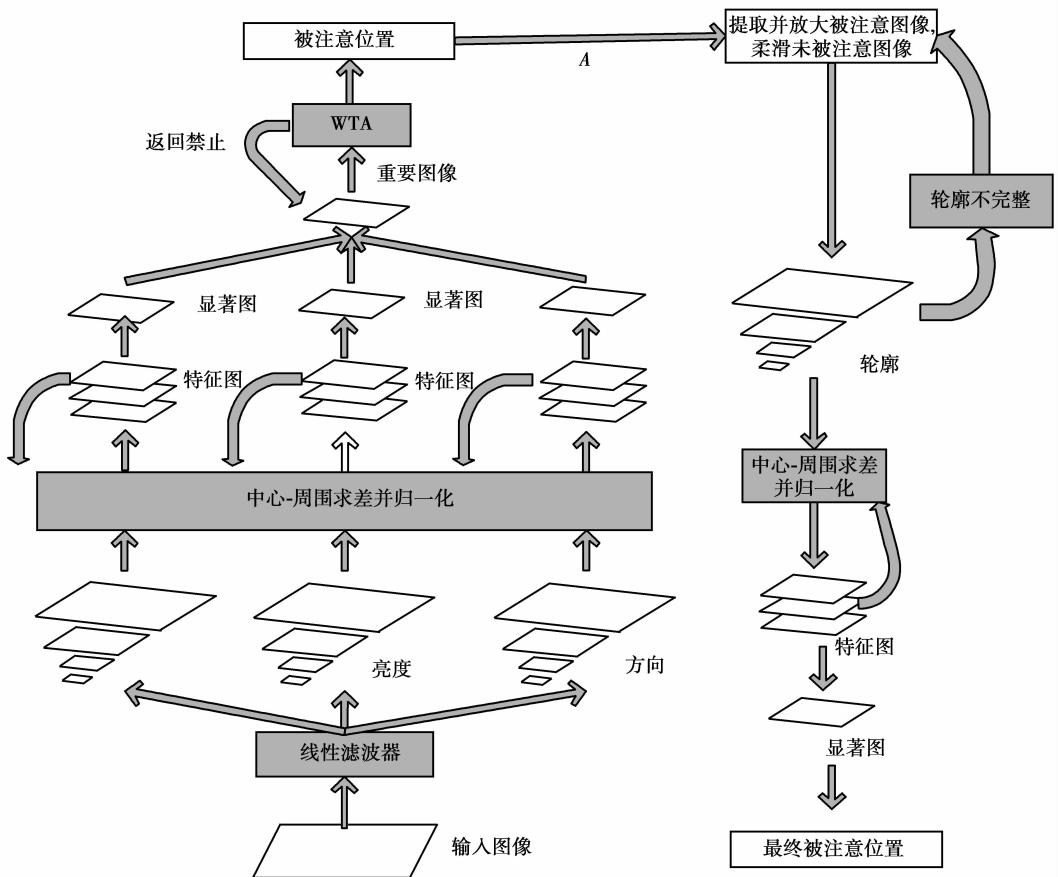


图 2 融合自底向上和自顶向下的视觉注意模型

其中, $Edge(i_o, j_o)$ 为注意范围的边缘; $Edge(i_o+i, j_o+j)$ 为 3×3 的矩阵。通过对组合以后的图像的模版边缘, 再次进行高低通滤波, 可以明显地减少模版边缘对边缘提取的影响。

对于掩模之后的图像, 可以进行如下处理: 通过边缘算子获得图像的边缘信息。使用对称高斯金字塔获得多分辨率的图像信息。在低分辨率的图像层中, 由于分辨率不够高, 在这里边缘信息比较集中, 很少出现完整图像边缘不完整的情况。轮廓信息可以通过连通域判断是否是一个完整的图像。同时为了防止边缘不完整, 除了使用 8-连通域外, 还在没有发现连通的地方跳 1~2 个象素点进行查找。由于在低分辨率图片中进行搜索, 因而跳 1~2 个象素点进行搜索满足需求。

如果轮廓不完整那么继续将注意范围扩大, 再次寻找。如 2 次结果相同, 停止搜索。只要轮廓完整, 那么就认定该轮廓是一个完整的物体或者是物体的主体部分。鉴于图像从整体认识, 在查找的时候加入了完整轮廓的意识, 而不单单是基于物体的简单特征, 也将这个称为自顶而下的注意。

对于被模糊的部分, 也同时获得图像的轮廓。由于是被低通滤波器滤除过的, 如果仍然有值得注意的轮廓, 那么该图像的边缘一定很清晰, 应该也是一个值得注意的物体, 那么也需要把它加入注意区域之中。

之后通过中心-周围求差并归一化, 获得每一层的特征图, 综合每层特征图获得显著图, 最终得到被注意的位置。

2 实验分析

根据 Koch、Itti 等人提出的机遇视觉生理和心理物理实验结果, 模拟生物体视觉注意机制的选择性注意模型, 对各种图像进行实验。

从图 3 可以看出, Itti 的模型能自动地寻找注意目标。与人眼的识别效果基本上一样。特别是图 3(b), 在如此复杂的图像里面也能成功地找到玩具小车和图书, 而没有被巨石, 花草所干扰, 证明该模型确实有一定的效果。

然而在对图 3(a) 的分析中, 可以看到, 近景的红色帆船和白色帆船, 被很好的注意到了。但是远景的白色帆船却没有被识别出来, 而是和远景一起被识别成了与众不同的颜色块。可以猜想这一识别通路和近景的红色帆船是一样的。为了证实这一论断, 现在用黑白图像进行视觉注意的实验。如图 4 所示: 进行了多次胜者为王(WTA)的神经网络对

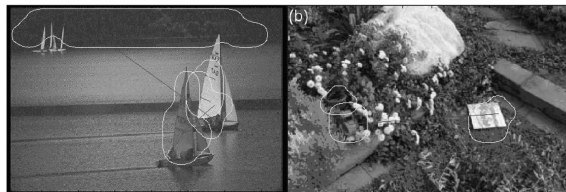


图 3 Itti 模型的彩色图像实验结果

图 4 进行扫描。可见由于亮度不突出, 方向性也不是很强, 所能注意到的近景的红色帆船和远景的帆船都没有被注意到, 反而是湖面的波光由于亮度被注意到了。但是即使是黑白图片, 正常人的视觉也能注意到近景的两条帆船和远景的几条帆船, 而对湖面情况应该不那么关心。

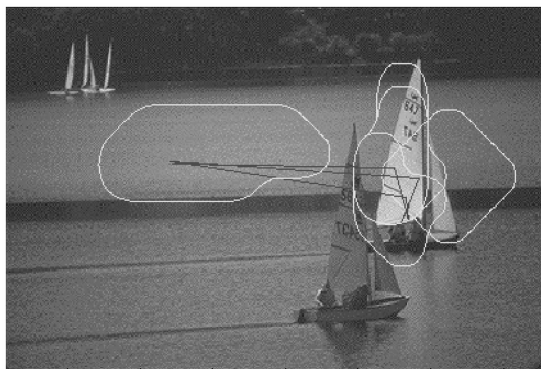


图 4 Itti 模型对黑白图像帆船的实验结果

对这种黑白图像该模型不能按人类注意的方式进行聚焦, 可见除了 Itti 模型所指出的人类视觉有从底到顶的视觉注意模型外, 也可能有从顶到底的视觉注意方式。也就是说在有一定先验知识的情况下, 对目标进行了二次注意和识别。或者更普遍的情况是, 人类已经有了一些图片库, 对图片库的检索对比, 找到需要注意的图像。当图片库没有这一图像的时候, 人类根据已有的经验可以按照图形构成的基本特性进行注意。最简单的就是首先注意到图像轮廓, 进而注意到图像的局部特征。

基于以上的假设, 对 Itti 的模型进行了一下修改, 融入整体的轮廓信息。同样对刚才的帆船图进行视觉注意实验。实验结果如图 5 所示。可以看出, 修正的模型能够很好地检测图像中的目标。为了更好地验证模型的性能, 用彩色图像对模型进行实验, 图 6、图 7 分别是该图像的最终实验结果和显著图。

可以看出, 修正的模型能够很好地检测到图像

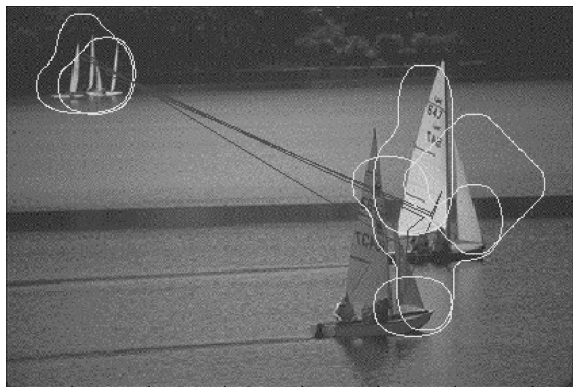


图 5 改进模型对黑白图像帆船的实验结果的目标。

在显著图中,加入了自顶向下的显著图信息。从图 7(d)可以看出,轮廓显著图可以检测出远处帆船的信息。

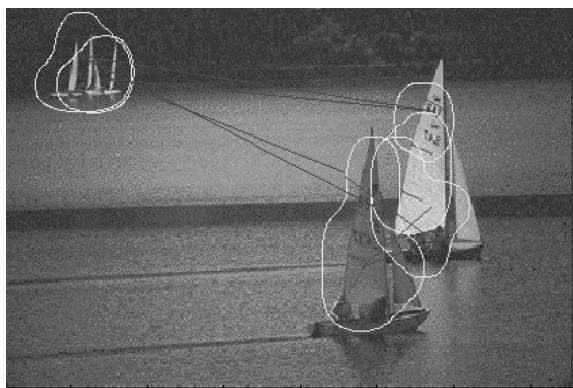


图 6 改进模型对彩色图像帆船的实验结果

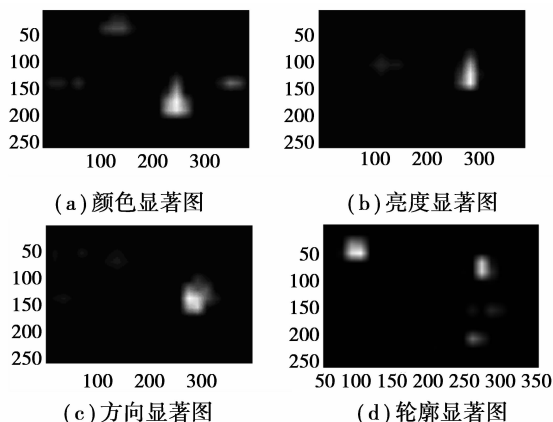
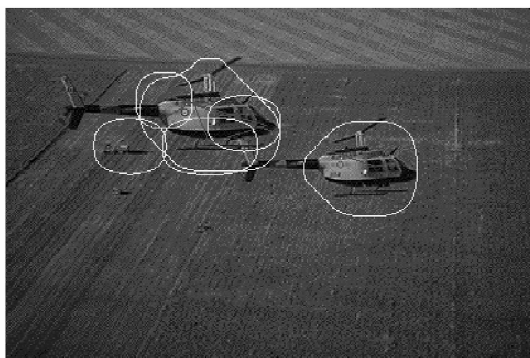


图 7 模型对彩色图像帆船的显著图

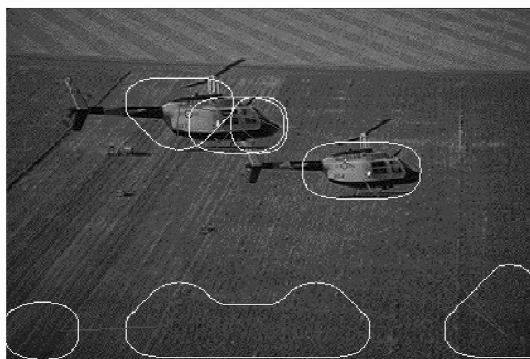
从图 8 中可以看到,模型可以注意到地面的房屋,而 Itti 的模型进行多次迭代之后也未能找到房屋,而是找到一些无关紧要的绿色耕地等。在此图像中改进模型较优。



(a) 改进模型找到直升机和地面的房子



(b) Itti模型经过3次迭代找到的直升机



(c) Itti模型经过多次迭代未能找到地面上的房子

图 8 改进模型与 Itti 模型观测对比

3 结 论

视觉注意是一个很复杂的问题。就目前的研究情况来看,单纯的说它是从低向上的方式或是从顶向下的方式完成视觉注意,都有不足,这是一个综合的行为。最初,人眼对光亮、颜色、方向等信息进行预处理,得到几个值得注意的点,然后再经过和大脑里面的已有信息比对,锁定目标。完成视觉注意这一个过程。也就是一个自底向上然后再自顶向下的过程。

根据上面的思想,提出了一个模型,综合利用 2 个模型,得到了很好的效果。

参考文献:

- [1] ITTI L, KOCH C. Computational modelling of visual attention[J]. *Nature Reviews Neuroscience*, 2001, 2(3):194-202.
- [2] KOCH C, ULLAMAN S. Shifts in selective visual attention; towards the underlying neural circuitry[J]. *Human Neurobiology*, 1985, 4(4):219-227.
- [3] 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.
- [4] ITTI L, KOCH C. A saliency-based search mechanism for overt and covert shifts of visual attention [J]. *Vision Research*, 2000, 40(10/12): 1489-1506.
- [5] KATO R, TAKAURA K, ISA T, et al. Saliency-based guidance of spontaneous saccades in monkeys with unilateral lesion of primary visual cortex[C/OL] // *Society for Neuroscience Annual Meeting (SFN'07)*, Nov 2007, Los Angeles CA, USA. Los Angeles CA: [s. n.], 2007. http://sns.ibr.neuroinf.jp/modules/xoonips/detail.php?item_id=17882&ml_lang=en-28k-
- [6] SIAGIAN C, ITTI L. Biologically-inspired robotics vision monte-carlo localization in the outdoor environment[C/OL] // *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Oct 2007, San Diego, CA, USA:[s. n.], 2007. http://www.citeulike.org/user/sylvain_chevallier/article/3199748.
- [7] EINHAEUSER W, MUNDHENK T N, BALDI P, et al. A bottom-up model of spatial attention predicts human error patterns in rapid scene recognition[J]. *Journal of Vision*, 2007, 7(10): 1-13.
- [8] PETERS R J, ITTI L. Beyond bottom-up: incorporating task-dependent influences into a computational model of spatial attention [C] // *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 17-22, 2007, Minneapolis, Minnesota, USA. Minnesota, USA: IEEE, 2007: 1-8.
- [9] GONZALEZ R C, RICHARD E. Digital image processing (second edition) [M]. Beijing: Publishing House of Electronics Industry, 2003.
- [10] NAVALPAKKAM V, ITTI L. An integrated model of top-down and bottom-up attention for optimizing detection speed computer[C] // *2006 IEEE Computer Society Conference on Vision and Pattern Recognition*, June 17-22, 2006, New York, USA. USA: IEEE, 2006:2049-2056.
- [11] ITTI L. Automatic foveation for video compression using a neurobiological model of visual attention[J]. *Image Processing*, *IEEE Transactions*, 2004, 13 (10): 1304-1318.
- [12] YOU Z, HO I L. A new weak edge detecting algorithm based on BMO and maximum function[C] // *IEEE/ICME International Conference on Complex Medical Engineering*. China: IEEE, 2007:748-751.
- [13] GOSHTASBY A, SATTER M. An adaptive window mechanism for image smoothing[J]. *Computer Vision and Image Understanding*, 2007, 111(2):155-169.
- [14] KUO Y H, LEE C S, LIU C C. A new fuzzy edge detection method for image enhancement [C] // *Proceedings of the Sixth IEEE International Conference on Fuzzy Systems*. Spain: IEEE, 1997:1069-1074.
- [15] BEZDEK J C, KERR D. Training edge detecting fuzzy neural networks with model-based examples [C] // *Proceedings of the third IEEE Conference on Fuzzy Systems*. Florida, USA: IEEE, 1994: 894-901.
- [16] EL-KHAMY S E, LOTFY M, YAMANY N. A modified fuzzy sobel edge detector [C] // *Seventeenth National Radio Science Conference*. Egypt: C32/1-C32/9.
- [17] WU Y X, HU D, WU M P, et al. A numerical-integration perspective on gaussian filters[J]. *IEEE Transactions on Signal Processing*, 2006, 54(8): 2910-2921.
- [18] SU Y, XU Z J, JIANG X Q. GPGPU-based gaussian filtering for surface metrological data processing[C] // *12th International Conference on Information Visualisation*. UK: IEEE, 2008: 94-99.
- [19] LAMPERT C H, WIRJADI O. Anisotropic gaussian filtering using fixed point arithmetic[C] // *2006 IEEE International Conference on Image Processing*. USA: IEEE, 2006: 1565-1568.

(编辑 侯 湘)