

基于自适应半径搜索的图像感兴趣区域检测

张立保 李浩

(北京师范大学信息科学与技术学院, 北京 100875)

摘要 感兴趣区域是图像中重要性较高并被优先关注的部分。传统视觉关注模型利用半径固定的圆描述感兴趣区轮廓,无法获得感兴趣区域的精确描述。提出一种基于自适应半径搜索的图像感兴趣区自动检测方法。提取图像的颜色、亮度和方向特征并生成多尺度的视觉显著图;通过计算显著图的全局显著度阈值获得视觉注意焦点搜索结束的条件;利用基于显著比的自适应半径搜索策略获取感兴趣区的精确描述信息。实验结果表明,新方法不仅能够提高对图像感兴趣区的自动检测精度,而且更符合人眼视觉系统的特点。对今后基于图像感兴趣区的目标自动识别具有重要价值。

关键词 图像处理;感兴趣区域;视觉注意焦点;自适应半径搜索

中图分类号 TN919.8 **文献标识码** A **doi**: 10.3788/CJL201340.0714001

Detection of Interest Image Region Based on Adaptive Radius Search

Zhang Libao Li Hao

(College of Information Science and Technology, Beijing Normal University, Beijing 100875, China)

Abstract The regions of an interest image are the parts of priority attention and have more significance. The traditional visual attention model describes the information of regions of interest using fixed-size circles and can't accurately express the outline of the regions of interest. A new automatic detection algorithm of the regions of interest based on adaptive radius search (ARS) is proposed. The new algorithm extracts color, intensity and orientation features of the image to generate a multi-scale saliency map. The global saliency threshold is calculated, which can get the end condition of searching the focus of attention. The adaptive radius search mechanism based on saliency ratio is proposed in the description of regions of interest to acquire the accurate information of regions of interest. The experimental results show that the new algorithm not only can effectively improve the detection precision of regions of interest, but also is more suitable to the features of human visual system. It has important value for the automatic target recognition of regions of interest in the future.

Key words image processing; region of interest; focus of attention; adaptive radius search

OCIS codes 100.2000; 100.3008; 110.2960; 110.2970

1 引言

图像的感兴趣区域通常是指一幅图像中重要性较高并被观察者优先关注的部分^[1-2]。大量视觉注意理论方面的研究发现:当面对复杂场景时,人眼趋向于快速准确地选择少数感兴趣区域并优先提供给大脑处理^[3-4]。所选感兴趣区域的中心被称为视觉注意焦点(FOA)。

构建视觉注意模型是提取感兴趣区的核心环

节。目前研究最为广泛的视觉注意模型是 Navalpakkam 等^[5-7]提出的自底向上的数据驱动型视觉注意模型(以下简称“Itti 模型”)。与传统方法利用图像分割或模板匹配获取感兴趣区有所不同,Itti 模型依据人眼对图像的感知过程,采用高斯差分滤波器、Gabor 滤波器与高斯金字塔提取图像在多个尺度上的视觉显著特征图,利用中央-周边差与跨尺度合并策略将多个特征图转化为一幅显著图,

收稿日期: 2013-01-11; 收到修改稿日期: 2013-03-11

基金项目: 国家自然科学基金(60602035, 61071103)、中央高校基本科研业务费专项资金(2012LYB50)

作者简介: 张立保(1977—),男,博士,副教授,主要从事遥感影像压缩与目标识别等方面的研究。

E-mail: libaozhang@163.com

运用“胜者为王”(WTA)与返回抑制机制获得多个FOA。最终通过以FOA为中心、半径固定的圆提取图像的兴趣区域。

Itti模型一方面结合人眼观察图像的特点,另一方面充分利用图像在多尺度及多特征上存在的差异,通过构建视觉显著图来提取感兴趣区域,在图像压缩及目标快速检测等领域得到广泛关注。文献[8]结合Itti模型提出一种基于空间降维的高分辨率遥感影像感兴趣区域快速检测,降低了Itti模型在高空间分辨率遥感图像应用中的计算复杂度。文献[9]直接将Itti模型应用于海上舰船目标的自动检测与识别中,获得了较好效果。文献[10]将区域增长策略与Itti模型结合,实现了对遥感影像视觉显著目标的快速检测。

Itti模型的一个突出问题是,以FOA为中心使用半径固定的圆提取感兴趣区,导致所检测到的区域与人眼的视觉注意区域并不完全吻合。当感兴趣区小于圆面积时,所提取的区域包含过多背景信息,从而降低检测精度;当感兴趣区大于圆面积时,由于无法被单个圆全部包含,需要用多个圆描述感兴趣区,增加了计算复杂度,甚至会将单一感兴趣区拆分成若干部分,违背人眼视觉的注意特点。此外,如果感兴趣区数量较多,Itti模型所选区域几乎覆盖整幅图像,失去了感兴趣区提取的意义。本文提出一

种新的感兴趣区自动检测算法。利用全局显著度阈值与自适应半径搜索获得感兴趣区的精确信息,在充分保留Itti模型优点的基础上提高了感兴趣区的检测精度。

2 基于自适应半径搜索的感兴趣区检测算法

2.1 基于自适应半径搜索的感兴趣区检测算法概述

针对Itti模型的缺点,提出了一个新的解决方法。新方法的显著图生成过程与Itti模型基本相同,但是,在获得显著图后,没有直接利用WTA与返回抑制机制获得FOA,而是首先引入全局显著度阈值计算对FOA结束搜索的条件,然后利用基于显著比的自适应半径搜索策略获取对感兴趣区更为精确有效的描述。

图1给出了Itti模型的基本框架。图2给出了采用Itti模型对三幅自然景物图像的兴趣区检测效果。可以看出,Itti模型虽然能够将视觉感兴趣的部分逐一检测出来,但是所得到的区域面积过大,从而影响了对真正感兴趣信息的提取精度。图3展示了所提新模型的兴趣区检测过程。与图1相比,增加了全局显著度阈值计算与自适应半径搜索两个环节。

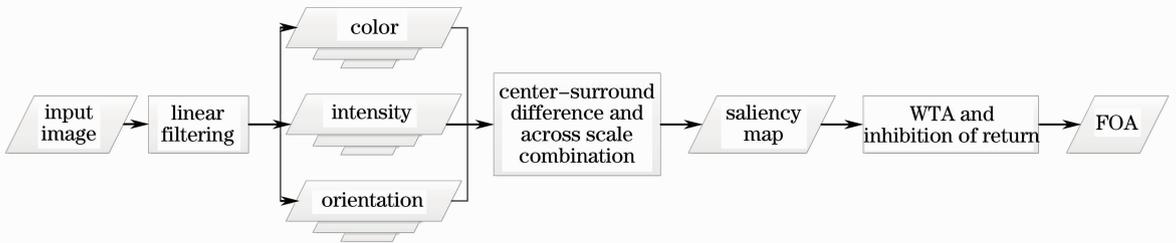


图1 Itti的传统视觉注意模型

Fig. 1 Traditional visual attention model by Itti

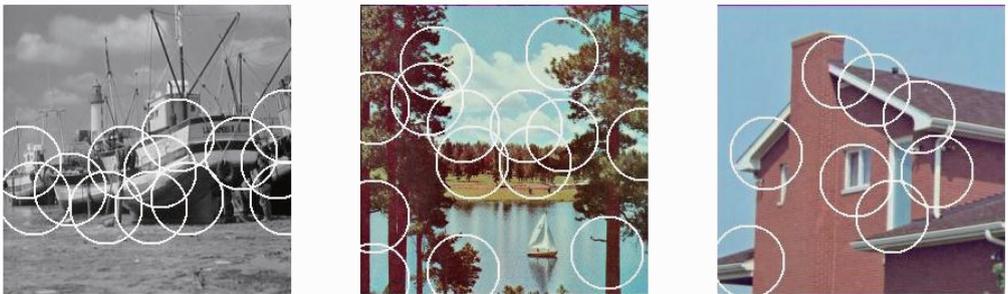


图2 Itti模型的兴趣区域检测结果

Fig. 2 Detection results of regions of interest based on Itti's model

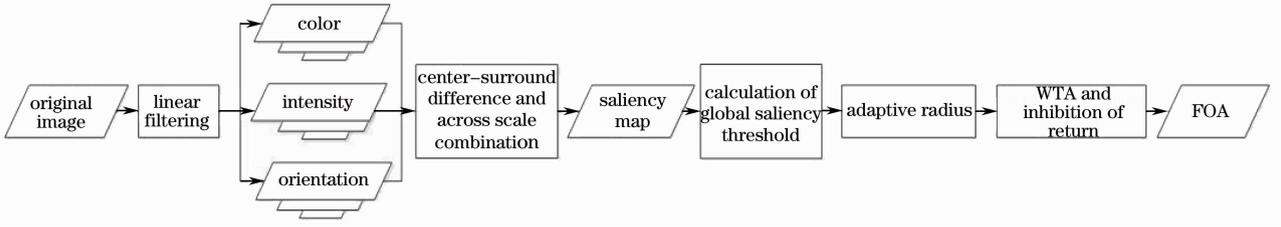


图3 基于自适应半径搜索的感兴趣区域检测模型

Fig. 3 Detection model of regions of interest based on adaptive radius search

2.2 特征提取

与 Itti 模型相似,新算法首先要提取图像的亮度、颜色和方向特征。对于灰度图像,图像本身就代表了其亮度特征,对于彩色图像,其亮度特征为

$$I = (r + g + b)/3, \quad (1)$$

利用亮度特征 I 生成高斯金字塔 $I(\sigma)$, 其中 $\sigma \in [0, 1, \dots, 6]$ 为尺度。对于灰度图像,可以省去对颜色特征的提取。对于彩色图像,利用亮度特征 I 对图像的 r, g, b 三个颜色通道进行标准化,生成4个颜色通道 R, G, B 和 Y , 表达为

$$R = r - (g + b)/2, \quad (2)$$

$$G = g - (r + b)/2, \quad (3)$$

$$B = b - (r + g)/2, \quad (4)$$

$$Y = (r + g)/2 - |r - g|/2 - b. \quad (5)$$

图像的方向特征 $O(\sigma, \theta)$ 则通过 Gabor 方向金字塔得到, 其中 $\sigma \in [0, 1, \dots, 6]$ 为尺度, $\theta \in [0^\circ, 45^\circ, 90^\circ, 135^\circ]$ 为滤波器的滤波方向, 分别生成四个颜色通道的六阶高斯金字塔 $R(\sigma), G(\sigma), B(\sigma), Y(\sigma)$ 。得到图像特征金字塔之后, 模拟人眼视觉的中央-周边差,

$$I(c, s) = |I(c)\Theta I(s)|, \quad (6)$$

$$f_{RG}(c, s) = |[R(c) - G(c)]\Theta[G(s) - R(s)]|, \quad (7)$$

$$f_{BY}(c, s) = |[B(c) - Y(c)]\Theta[Y(s) - B(s)]|, \quad (8)$$

$$O(c, s, \theta) = |O(c, \theta)\Theta O(s, \theta)|. \quad (9)$$

式中 $s = c + \delta, \delta \in (2, 3), c \in (1, 2, 3), \Theta$ 为中央-周边差算子。

2.3 基于全局增强的显著图生成

利用中央-周边差算子可以获得各尺度上亮度、颜色和方向的特征图,但是感兴趣区检测是在一张图上进行,需要将各尺度上的特征图合并为一幅显著图。为了克服各特征图间的相互干扰,采用全局增强的合并策略,具体算法由三个步骤构成:

1) 将各特征图归一化到同一个范围 $[0, 1, \dots, M]$ 之内;

2) 找出每幅特征图的全局最大值 M 以及除 M 外其余局部最大值的平均值 \bar{m} ;

3) 将每一幅特征图乘以加强因子 $(M - \bar{m})^2$ 。

用算子 $N(\cdot)$ 代表全局增强。一幅特征图的全局极大值与其他活跃区的平均水平差别越大就能获得越大的权值,使显著性被放大,特征图被突出。反之,差别越小,权值就越小,该特征图就被抑制。这样最终合并得到的显著图就能够在突出注意目标的同时抑制干扰。全局增强可表达为

$$\bar{I} = \bigoplus_{c=2s=c+3}^4 \bigoplus_{c=2s=c+3}^{c+4} N[I(c, s)], \quad (10)$$

$$\bar{C} = \bigoplus_{c=2s=c+3}^4 \bigoplus_{c=2s=c+3}^{c+4} \{N[f_{RG}(c, s)] + N[f_{BY}(c, s)]\}, \quad (11)$$

$$\bar{O} = \sum_{\theta \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}} N\left\{\bigoplus_{c=2s=c+3}^4 \bigoplus_{c=2s=c+3}^{c+4} N[O(c, s, \theta)]\right\}, \quad (12)$$

式中 \bar{I} 代表最终的亮度特征图, \bar{C} 代表最终的颜色特征图, \bar{O} 代表最终的方向特征图。 \oplus 为跨尺度求和算子,具体做法是将相应特征图经插值达到高斯金字塔第3层大小再逐像素相加。彩色图像的最终显著图可表达为

$$S = \frac{1}{3} [N(\bar{I}) + N(\bar{O}) + N(\bar{C})], \quad (13)$$

而灰度图像,由于缺少颜色信息,合成显著图时不考虑颜色特征,最终显著图可表达为

$$S = \frac{1}{2} [N(\bar{I}) + N(\bar{O})]. \quad (14)$$

2.4 全局显著度阈值计算

与 Itti 模型不同,这里算法利用全局显著度阈值为 FOA 的搜索过程设置了结束标志。该阈值的选取主要依据显著图中像素的灰度值,因为只有当图像中某些部分的显著度达到该阈值时才能得到人眼关注,因此当全局极大值达到这个阈值时,FOA 的搜索过程就结束了。

图4给出了一个显著图直方图的示例。图中 A 点为显著图的灰度均值, B 点为显著图灰度均值与标准差的和, C 点为灰度均值与标准差的差, D 点为显著图的灰度最大值,大部分像素的灰度值都在 B

和 C 之间,显著度比较高的区域具有明显较高的灰度等级,而且在整幅图中所占的比例也较少,对应应在直方图中就应该位于 C 和 D 之间,因此,全局显著度阈值应该设定在 C 和 D 之间的某一点,所提模型设定全局显著度阈值 T 的依据为

$$T = \frac{1}{5}[u + \sigma + (M - u - \sigma)], \quad (15)$$

式中 u 为图像灰度均值, σ 为图像灰度标准差, M 为图像灰度最大值。

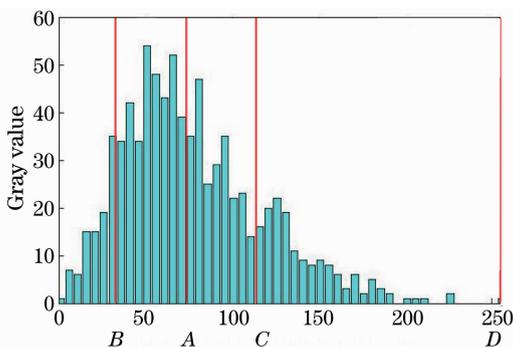


图 4 视觉显著图的直方图示例

Fig. 4 Histogram of the visual saliency map

2.5 自适应半径搜索

自适应半径搜索的实质是自动设置以 FOA 为中心的圆的半径。首先需计算显著比 $k(r, T)$

$$k(r, T) = \frac{\text{Num}[P(x, y) \mid P(x, y) \geq T, P(x, y) \in R(r)]}{\text{Num}[P(x, y) \mid P(x, y) \in R(r)]}, \quad (16)$$

式中 $\text{Num}(\cdot)$ 代表符合括号中条件的像素个数, $R(r)$ 表示半径为 r 的圆形区域, $P(x, y)$ 表示像素点. $k(r, T)$ 的取值范围为 $0 \sim 1$. $k(r, T)$ 代表了圆形区域内像素值在全局显著度阈值 T 之上的像素比例. 当 $k(r, T) = 1$ 时, $R(r)$ 内所有像素的灰度值均在全局显著度阈值 T 之上, 而当 $0 < k(r, T) < 1$ 时, $R(r)$ 内并非所有像素的灰度值都在全局显著度阈

值 T 之上, 会包含一定比例的非显著区域. 自适应半径搜索算法的流程如图 5. 首先将半径 r 设为一个较大值, 一般设为图像较小边长的 $1/2$, 并设定显著比阈值 K . 计算当前半径下的显著比 $k(r, T)$, 如果 $k(r, T)$ 小于 K , 则将半径 r 减小 1, 重复上述过程直到 $k(r, T)$ 达到显著比阈值 K 为止, 此时的半径 r 就是最终确定的半径.

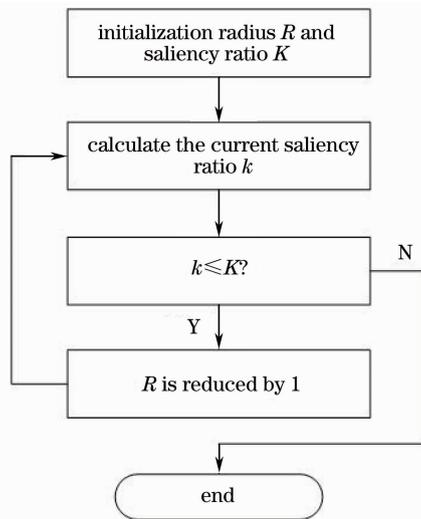


图 5 自适应半径搜索算法

Fig. 5 Algorithm of adaptive radius search

显然, 当 $k(r, T)$ 增大时半径 r 会减小. 因此在显著图中, 一旦显著度阈值 T 被确定, 大于阈值 T 的像素数量也就确定了, 当单个 FOA 的半径较小时, 就需要更多的 FOA 来表示这些区域. 相反, 当 $k(r, T)$ 减少时, 相应的半径 r 就会增大, 只需要较少的 FOA 就能够包含所有区域.

3 实验结果与分析

为了验证所提算法对感兴趣区检测有效性, 选用美国南加州大学的 USI-SIPI image database 提供的彩色与灰度图像, 图像分辨率均为 $512 \text{ pixel} \times 512 \text{ pixel}$.

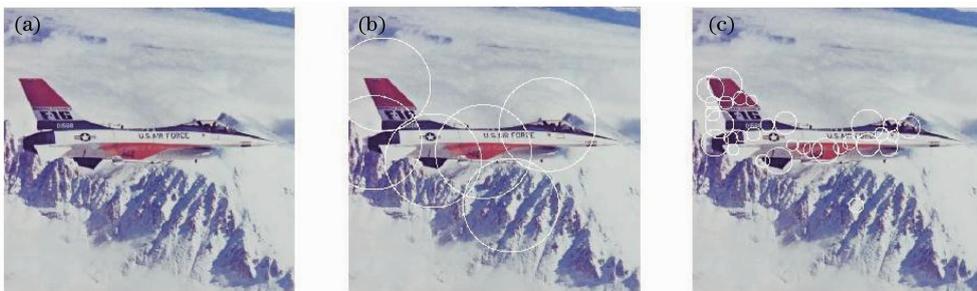


图 6 所提模型与 Itti 模型感兴趣区检测结果的比较(彩色图像)。(a) 原始图像; (b) Itti 模型; (c) 所提模型

Fig. 6 Comparison of detection results between Itti's model and the proposed model for airplane (color image).

(a) Original image; (b) Itti's model; (c) proposed model

图 6 给出了针对彩色图像 Airplane 的感兴趣区域搜索对比。图 7 给出了针对灰度图像 Truck and Tank 的感兴趣区域搜索对比。将所提算法与 Itti 模型所得实验结果对比后不难发现,新算法中提取出的感兴趣区域在整幅图中所占的比例明显减少,

但是准确程度得到提高,基本能够准确定位和覆盖人们所关注的对象,Itti 模型则只能对感兴趣区域进行大概描述,而且当感兴趣区域数量较多时几乎覆盖了整幅图像,失去了提取感兴趣区域的意义。

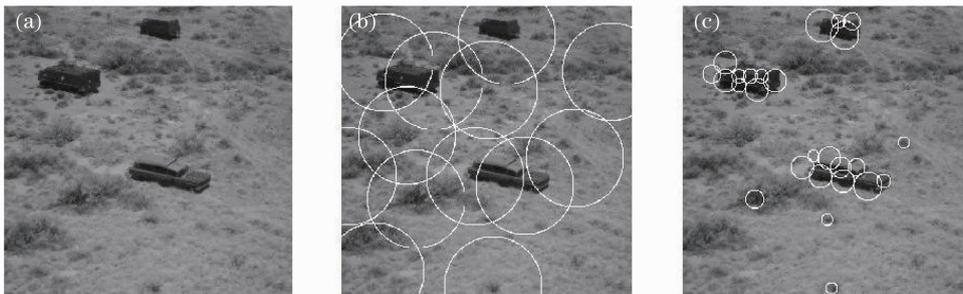


图 7 所提模型与 Itti 模型感兴趣区检测结果的比较(灰色图像)。(a) 原始图像;(b) Itti 模型;(c) 所提模型

Fig. 7 Comparison of detection results between Itti's model and the proposed model (gray scale image).

(a) Original image; (b) Itti's model; (c) proposed model

4 结 论

本文提出一种新的基于自适应半径搜索的图像感兴趣区域自动检测算法。新方法的最大创新之处在于提出全局显著度阈值算法与自适应半径搜索机制,从而获得了图像中感兴趣区域的准确信息,不仅能够提高图像感兴趣区域的检测精度,而且更为符合人眼视觉系统的特点,同时极大减少后续处理的数据量,对今后的图像感兴趣目标自动检测具有重要理论意义与实际价值。

参 考 文 献

- 1 A Signoroni, F Lazzaroni, R Leonardi. Exploitation and extension of the region-of-interest coding functionalities in JPEG2000 [J]. IEEE Transactions on Consumer Electronics, 2003, 49(4): 818-823.
- 2 L Liu, G Fan. A new JPEG 2000 region of interest image coding method: partial significant bitplanes shift [J]. IEEE Signal Processing Letters, 2003, 10(2): 35-38.
- 3 L Itti, C Koch. Computational modeling of visual attention[J]. Nature Reviews Neuroscience, 2001, 2 (3): 194-202.
- 4 O Le Meur, P Le Callet, D Barba, *et al.*. A coherent

computational approach to model bottom-up visual attention[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(5): 802-817.

- 5 V Navalpakkam, L Itti. Modeling the influence of task on attention[J]. Visual Research, 2005, 45(2): 205-231.
- 6 L Itti, C Koch, E Niebur. 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.
- 7 D Lee, L Itti, C Koch. Attention activates winner-take-all competition among visual filters[J]. Nature Neuroscience, 1999, 2(4): 375-381.
- 8 Zhang Libao, Wang Pengfei. Fast detection of regions of interest in high resolution remote sensing image [J]. Chinese J Lasers, 2012, 39(7): 0714001.
张立保,王鹏飞. 高分辨率遥感影像感兴趣区域快速检测[J]. 中国激光, 2012, 39(7): 0714001.
- 9 R Xin, R Lei. Search aid system based on machine vision and its visual attention model for rescue target detection [C]. Second WRI Global Congress on Intelligent Systems, 2011, 2: 149-152.
- 10 Zhang Libao. Fast detection of visual saliency regions in remote sensing image based on region growing [J]. Chinese J Lasers, 2012, 39(11): 1114001.
张立保. 基于区域增长的遥感影像视觉显著目标快速检测[J]. 中国激光, 2012, 39(11): 1114001.

栏目编辑:何卓铭