

中图法分类号: TP391 文献标志码: A 文章编号: 1006-8961(2011)05-0767-06

论文索引信息: 张巧荣, 景丽, 肖会敏, 刘海波. 利用视觉显著性的图像分割方法 [J]. 中国图象图形学报, 2011, 16(5): 767-772

利用视觉显著性的图像分割方法

张巧荣¹⁾, 景丽¹⁾, 肖会敏²⁾, 刘海波³⁾

¹⁾ (河南财经学院计算机与信息工程学院, 郑州 450001) ²⁾ (河南财经学院数学与信息系, 郑州 450002)

³⁾ (哈尔滨工程大学计算机科学技术学院, 哈尔滨 150001)

摘要: 提出一种利用视觉显著性对图像进行分割的方法。首先提取图像的底层视觉特征, 从局部显著性、全局显著性和稀少性3个方面计算各特征图像中各像素的视觉显著性, 得到各特征显著图; 对各特征显著图进行综合, 生成最终的综合显著图。然后对综合显著图进行阈值分割, 得到二值图像, 将二值图像与原始图像叠加, 将前景和背景分离, 得到图像分割结果。在多幅自然图像上进行实验验证, 并给出相应的实验结果和分析。实验结果表明, 该方法正确有效, 具有和人类视觉特性相符合的分割效果。

关键词: 图像分割; 视觉注意; 显著图; 阈值

Image segmentation based on visual saliency

Zhang Qiaorong¹⁾, Jing Li¹⁾, Xiao Huimin²⁾, Liu Haibo³⁾

¹⁾ (College of Computer and Information Engineering, Henan University of Finance and Economics, Zhengzhou 450001 China)

²⁾ (Department of Mathematics and Information, Henan University of Finance and Economics, Zhengzhou 450002 China)

³⁾ (College of Computer Science and Technology, Harbin Engineering University, Harbin 150001 China)

Abstract: An approach for image segmentation based on visual saliency is proposed in this paper. First low-level visual features of the image are extracted. Local saliency, global saliency and rarity saliency are computed for each feature map to get the feature conspicuity maps. Then these conspicuity maps are integrated to generate the saliency map. The saliency map is segmented using a threshold and a binary mask map is obtained. Finally the foreground and background of the original image are separated by adding the binary map to the original image. The proposed model has been tested on many natural images. Experimental results show that the proposed approach is valid and the segmentation results are consistent with human visual system.

Keywords: image segmentation; visual attention; saliency map; threshold

0 引言

在图像分析和处理中, 人们往往只对图像中的某些部分感兴趣, 这些部分通常称为目标或前景, 一般对应于图像中具有特定性质的区域。为此, 需要将目标从图像中分离出来, 才能够做进一步的分析

和处理。图像分割就是将一幅图像划分成若干个具有某种均匀一致性的区域, 从而将人们感兴趣的区域从复杂的场景中提取出来的技术。

视觉注意机制是人类及其他灵长目动物一个重要的内在属性, 能够帮助人类在大量视觉信息中迅速地找到显著的或感兴趣的物体, 并忽略其他不重要的内容, 降低了信息处理的计算量。因此, 如果将

收稿日期: 2010-05-18; 修回日期: 2010-07-16

基金项目: 国家自然科学基金项目(60774041); 河南省科技攻关项目(102102210398); 中央高校基本科研业务费专项资金(HEUCF100604); 国家教育部博士点专项基金(20092304120013)。

第一作者简介: 张巧荣(1978—), 女, 副教授, 博士。研究方向为图像处理。E-mail: zhangqiaorong@sina.com。

视觉注意机制引入图像处理中,则可以显著地提高图像处理的效率。

目前已有一些利用视觉注意机制进行图像分割的方法,首先需要生成显著图。比较有代表性的生成显著图的方法有文献[1]中提出的生物学启发方法,文献[2-4]提出的基于特征对比度的方法和文献[5]提出的谱残差方法。但采用这些方法得到的显著图有时不能正确反映出图像中前景目标和背景区域的显著性,不能直接根据显著图进行图像分割,只能根据生成的显著图结果选取种子点,再结合其他算法比如区域增长等进行图像分割^[6-7],计算量较大。

为此,采用一种新的生成显著图的方法,从局部显著性、全局显著性和稀少性 3 个方面计算图像中各像素的视觉显著性。对生成的显著图直接进行简单的阈值分割,即可将图像的前景目标和背景区域分离,不需要结合其他算法,大大地减少了计算量。

1 模型

图 1 是本文给出的利用视觉显著性的图像分割方法示意图,主要包括视觉显著性度量和图像分割两部分。

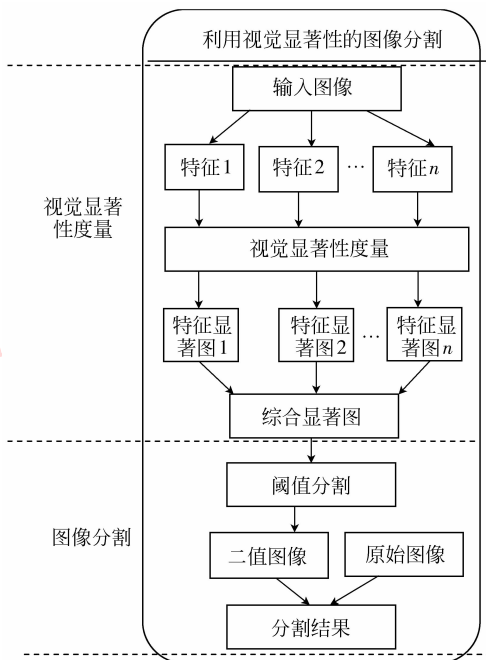


图 1 利用视觉显著性的图像分割方法

Fig. 1 Image segmentation based on visual saliency

2 视觉显著性度量

视觉显著性度量是基于视觉显著性的图像分割方法中的关键环节,主要计算图像中各像素的显著性,结果用与原始输入图像大小相等的一幅灰度图像来表示,称为显著图。其中每一个像素值代表了原始图像中对应位置像素的显著值,值越大说明该像素在原始图像中越显著,越容易获得观察者的注意。

2.1 底层特征提取

图像中一个区域的显著性依赖于它自身的特征与周围环境的差异,如果图像中一个区域为显著区域,则该区域至少有一种特征与其周围环境不同。不同图像中,同一特征对视觉显著性的影响是不同的。有的图像中亮度为显著特征,有的图像中颜色为显著特征,因此需要提取图像的各种不同早期视觉特征。文献[1]提取亮度、颜色和方向等特征,但通过实验发现,方向特征在自然图像中的作用并不明显,却增加了计算的复杂度,因此本文只考虑亮度和颜色特征。

因为 HSI 颜色空间用色调 (hue)、饱和度 (saturation) 和亮度 (intensity) 来描述颜色,比 RGB 颜色空间更符合人的视觉特性。因此,首先利用式(1)将输入的图像从 RGB 颜色空间转换到 HSI 颜色空间,得到亮度和颜色特征图。

$$\begin{cases} H = \frac{1}{360} \left[90 - \arctan\left(\frac{F}{\sqrt{3}}\right) + \begin{cases} 0 & G > B \\ 180 & G < B \end{cases} \right] \\ S = 1 - \left[\frac{\min(R, G, B)}{I} \right] \\ I = \frac{R + G + B}{3} \\ F = \frac{2R - G - B}{G - B} \end{cases} \quad (1)$$

2.2 生成显著图

本文主要从计算局部显著性、全局显著性以及稀少性等几个方面来生成显著图。

1) 局部显著性度量

图像中一个像素的显著性不在于该像素特征值的大小,而在于该像素与其周围像素的对比度,对比度越大,该像素越显著。文献[2-3]通过分析图像中各像素与其周围邻域的对比度来计算显著性,但是邻域大小不容易确定,而且计算量较大。本文从

频域分析像素的局部显著性,在图像的频域特征中,幅度谱和相位谱的作用和含义不同。相位谱包含了图像的结构特征,能够反映出图像中像素特征值的变换情况;而幅度谱则包含了图像中各像素特征值的大小。

实验证明,幅度谱和相位谱在图像重构中的作用不同^[8-9]。仅利用相位谱对图像进行重构,可以得到与原始图像结构相似的重构结果;而仅利用幅度谱对图像进行重构,结果与原始图像差距很大。

因此,本文利用式(2)对各特征图像进行局部显著性度量。首先对图像进行傅里叶变换,提取幅度谱和相位谱;然后仅利用相位谱对图像进行重构,得到各特征图像的局部显著性图。

$$\begin{cases} F(u,v) = \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N f(x,y) e^{-\frac{j2\pi ux}{M}} e^{-\frac{j2\pi vy}{N}} = \\ |F(u,v)| e^{j\phi(u,v)} \\ |F(u,v)| = [R^2(u,v) + I^2(u,v)]^{1/2} \\ \phi(u,v) = \arctan\left(\frac{I(u,v)}{R(u,v)}\right) \\ S_{\text{Local}}(x,y) = \frac{1}{MN} \sum_{u=1}^M \sum_{v=1}^N \phi(u,v) e^{-\frac{j2\pi ux}{M}} e^{-\frac{j2\pi vy}{N}} \end{cases} \quad (2)$$

式中, f 表示像素的特征值, MN 为图像大小。 $|F|$ 为傅里叶变换后图像的幅度谱, ϕ 为图像的相位谱。 S_{Local} 为得到的图像中各像素的局部显著性值。

2) 全局显著性度量

如果只考虑局部显著性,则图像中变化比较剧烈的边缘或复杂的背景区域显著性较高,而比较平滑的目标内部显著性较低,因此还需要考虑全局显著性。像素的全局显著性是指该像素相对于整幅图像(而不是其某个邻域)来说其显著程度,本文利用式(3)来生成各特征图像的全局显著性图。

$$\begin{cases} S_{\text{Global}}(x,y) = e^{\frac{1f(x,y) - f_{\text{avg}}(x,y)}{f_{\text{avg}}(x,y)}} \\ f_{\text{avg}}(x,y) = \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N f(x,y) \end{cases} \quad (3)$$

式中, $S_{\text{Global}}(x,y)$ 为图像中各像素的全局显著性值。

3) 稀少性度量

稀少性意味着某个特征值在图像中出现的次数越少,具有该特征值的像素越与众不同,则该像素的显著性值就可能越高。本文利用下式来衡量图像中各像素所具有的特征的稀少性。

$$S_{\text{Rarity}}(x,y) = \frac{1}{\text{hist}(f(x,y))} \quad (4)$$

式中, $\text{hist}(\cdot)$ 为图像的特征直方图, $S_{\text{Rarity}}(x,y)$ 为

该像素的稀少性度量值。

4) 特征显著图

利用式(5)对计算得到的某个特征图的局部显著性、全局显著性和稀少性度量结果进行综合,得到最终的特征显著图。

$$\begin{cases} V = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N \left| f(x,y) - \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N f(x,y) \right| \\ w_i = \frac{V_i}{\sum_{i=1}^3 V_i} \\ S_f = w_1 S_{\text{Local}} + w_2 S_{\text{Global}} + w_3 S_{\text{Rarity}} \end{cases} \quad (5)$$

具体的特征显著图生成的示例如图2所示。

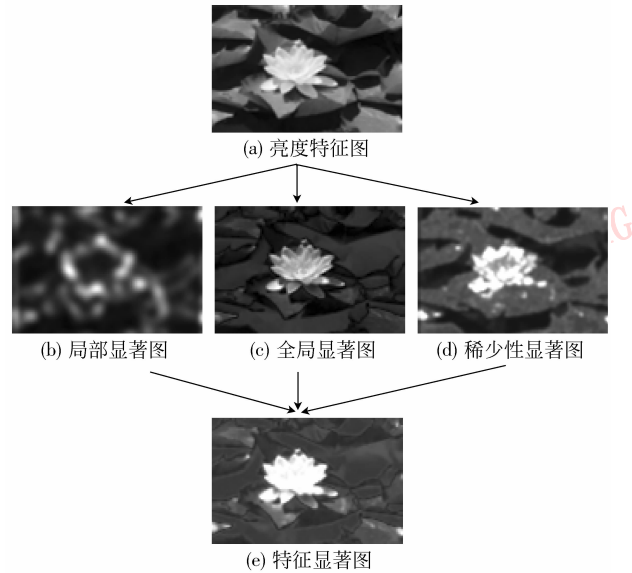


图2 生成特征显著图示例

Fig.2 Example of computing feature saliency map

5) 特征整合

从图2中可以看出,不同的特征显著图的贡献是不同的,有的特征显著图能够有效地指示显著区域,而有的则不能。因此,需要一个合理的特征整合策略,对得到的多个特征显著图进行组合,生成最终的显著图。本文从显著点位置、个数以及分布情况等几个因素来动态选择特征和加权。

首先需要对各特征显著图计算阈值,提取显著值大于阈值的作为显著点。根据稀少性原则,一个特征图显著图中的显著点越多,则该特征显著图对最终显著图的贡献越小。为此定义权值 W_{area} 为显著点的个数,即

$$W_{\text{area}} = \text{Num}_{\text{salient}} \quad (6)$$

人们面对一幅图像时,注意力很容易集中在图

像的中心位置。即图像中心的区域更容易成为显著区域。为此,定义权值 $W_{location}$ 为各显著点距离图像中心的平均距离,即

$$W_{location} = \frac{1}{N} \sum_{i=1}^N Dist(sp_i, center) \quad (7)$$

式中, N 为特征显著图中显著点的个数, sp_i 为显著点, $center$ 为图像的中心位置。

如果特征显著图中各显著点不集中,而是比较分散,则该特征显著图对最终显著图的贡献不大。为此定义权值 $W_{distribution}$ 为各显著点之间的平均距离, $centroid$ 为显著点的中心位置,即

$$W_{distribution} = \frac{1}{N} \sum_{i=1}^N Dist(sp_i, centroid) \quad (8)$$

根据显著点个数、位置和分布 3 个因素,利用式(9)计算各特征显著图的权值,对各特征显著图进行整合,得到最终的显著图,如图 3 所示。

$$\begin{cases} S = \sum_{i=1}^m W_i S_{fi} \\ W_i = \frac{1}{\sum_{i=1}^m \frac{1}{W_{fi}}} \\ W_{fi} = W_{area}^i + W_{location}^i + W_{distribution}^i \end{cases} \quad (9)$$

式中, S_{fi} 为第 i 个特征显著图, W_{area}^i 、 $W_{location}^i$ 和 $W_{distribution}^i$ 分别为第 i 个特征显著图的面积因素的权值、位置权值和分布因素的权值, W_{fi} 为第 i 个特征显著图上述 3 个因素权值的综合, W_i 为第 i 个特征显著图的最终权值, S 为最终显著图。

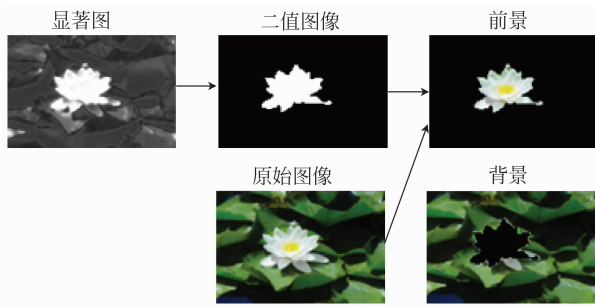


图 3 图像分割过程示例

Fig. 3 Example of image segmentation

3 图像分割

生成最终显著图之后,利用式(10)选择合适的阈值对显著图进行阈值分割,得到二值黑白图像,其

中白色区域代表了原始图像中的前景目标,而黑色区域则表示原始图像中的背景部分。

$$\begin{cases} E_B = \frac{-\sum_{i=1}^L p_i}{\sum_{j=1}^L p_j} \times \ln \frac{p_i}{\sum_{j=1}^L p_j} \\ E_F = \frac{-\sum_{i=L+1}^L p_i}{1 - \sum_{j=1}^L p_j} \times \ln \frac{p_i}{1 - \sum_{j=1}^L p_j} \\ T = \arg \max_i (E_F + E_B) \end{cases} \quad (10)$$

式中, L 为显著图中像素灰度值的最大值, p_i 为显著图中灰度值 i 出现的概率。 E_F 为前景图像的信息熵, E_B 为背景图像的信息熵。使得这两部分的信息熵之和最大的灰度值 T 即为图像分割的最佳阈值。

$$B(x, y) = \begin{cases} 1 & S(x, y) \geq T \\ 0 & S(x, y) < T \end{cases} \quad (11)$$

对阈值分割后得到的二值图像进行简单的形态学操作,去除一些孤立的不连续的白色区域,得到最终的二值图像。将二值图像叠加到原始图像上,即可实现前景与背景的分割。具体的过程示例如图 3 所示。

4 实验结果与分析

为了验证本文给出的图像分割方法的正确性和有效性,选取大量自然图像进行实验。实验环境为 3.0 GHz, Pentium4 处理器,内存 512 MB。分割结果示例如图 4 所示。



图 4 图像分割结果示例

Fig. 4 Results of image segmentation

为了进一步验证本文图像分割结果的正确性, 本文将分割结果与人工分割结果 (ground truth) 进行了比较^[10], 如图 5 所示, 其中白色区域表示前景, 黑色区域表示背景部分。



图 5 与 ground truth 对比结果

Fig. 5 Comparison results with ground truth

通常用查全率 (precision) 和查准率 (recall) 作为衡量分割结果正确性的准则^[10], 其定义为

$$\left\{ \begin{aligned} \text{precision} &= \frac{\sum_{(x,y)} G(x,y)B(x,y)}{\sum_{(x,y)} B(x,y)} \\ \text{recall} &= \frac{\sum_{(x,y)} G(x,y)B(x,y)}{\sum_{(x,y)} G(x,y)} \end{aligned} \right. \quad (12)$$

式中, G 为 ground truth 图像, B 为显著图阈值分割后得到的二进制图像。

图 6 是分别对利用本文给出的算法与其他算法生成的显著图进行阈值分割得到的二进制图像计算平均查全率和查准率的对比结果。

从图 6 中可以看出, 本文的分割结果要优于其他算法的结果, 具有较高的查全率和查准率。影响图像分割结果的因素主要有两个: 1) 视觉显著性度量结果。如果生成的显著图不正确, 则分割结果就不正确。文献[2,5]仅考虑局部显著性, 文献[10]仅考虑全局显著性, 因此生成的显著图不能正确反映图像各部分的视觉显著性; 而本文综合考虑局部、全局和稀少性等因素, 生成的显著图更合理。2) 阈

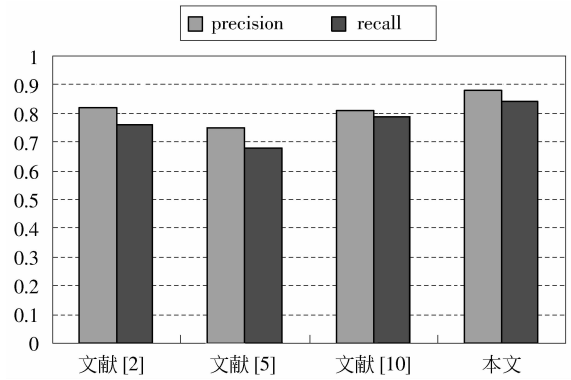


图 6 与其他算法对比结果

Fig. 6 Comparison results with other methods

值的选择。阈值选择太小, 容易将背景区域作为前景目标; 阈值选择过大, 容易造成分割的前景目标不完整, 因此需要选择合适的阈值。文献[5,10]选取显著图均值作为阈值, 本文根据信息熵选取阈值, 分割效果更合理。文献[2]利用模糊增长方法进行分割, 效果较好, 但计算复杂度高。

5 结 论

提出一种利用视觉显著性的图像前景、背景分割方法。该方法首先根据原始图像的各种早期视觉特征, 计算相应的显著图; 然后根据显著图动态选择阈值, 进行简单的阈值分割, 得到二进制图像, 将二进制图像与原始图像叠加, 即可实现前景与背景的分

离。与其他基于视觉注意机制的图像分割方法相比, 用本文给出的视觉显著性度量方法生成的显著图, 更能准确有效地表示出图像中各像素的视觉显著性, 直接根据显著图进行简单的阈值分割即可对图像前景和背景进行分离, 不需要结合其他图像分割方法, 从而降低了计算复杂度, 图像分割的结果也更符合人类的视觉特性。本文算法对于包含显著前景目标的图像可以得到很好的分割效果, 但对于一些背景比较复杂且前景目标不明确的图像, 分割效果较差。因此, 如何能够结合图像自身的特点, 设计出更合理的视觉注意计算方法, 适合于更复杂的图像是本文进一步努力的方向。

参考文献 (References)

[1] Itti L, Kouch C. Computational modeling of visual attention[J]. Nature Reviews Neuroscience, 2001, 2(3):194-230.

- [2] Ma Yufei, Zhang Hongjiang. Contrast-based image attention analysis by using fuzzy growing [C] // Proceedings of the 11th ACM International Conference on Multimedia. New York, NY: Association for Computing Machinery, 2003: 374-381.
- [3] Radhakrishna Achanta, Francisco Estrada, Patricia Wils, et al. Salient region detection and segmentation [C] // Proceedings of International Conference on Computer Vision Systems. Berlin Heidelberg: Springer-Verlag, 2008: 66-75.
- [4] Zhang Peng, Wang Runsheng. Detection salient regions based on location shift and extent trace [J]. Journal of Software, 2004, 15(6): 891-898. [张鹏, 王润生. 基于视点转移和视区追踪的图像显著区域检测 [J]. 软件学报, 2004, 15(6): 891-898.]
- [5] Xiaodi Hou, Liqing Zhang. Saliency detection: a spectral residual approach [C] // Proceedings of 2007 IEEE Conference on Computer Vision and Pattern Recognition. Florida, USA: IEEE, 2007: 1-8.
- [6] Luo Tong, Chen Yuquan. Visual attention guidance and region competition for medical image segmentation [J]. Journal of Zhejiang University: Engineering Science, 2007, 41(11): 1797-1800. [罗彤, 陈裕泉. 视觉注意引导和区域竞争控制的医学图像分割 [J]. 浙江大学学报: 工学版, 2007, 41(11): 1797-1800.]
- [7] Fu Yu, Cheng Jian, Li Zhenglong, et al. Saliency cuts: an automatic approach to object segmentation [C] // Proceedings of the 19th International Conference on Pattern Recognition. Florida, USA: IEEE, 2008: 1-4.
- [8] Ni Xuelei, Huo Xiaoming. Statistical interpretation of the importance of phase information in signal and image reconstruction [J]. Statistics & Probability Letters, 2007, 77(4): 447-454.
- [9] Peter J Bex, Walter Makous. Spatial frequency, phase, and the contrast of natural images [J]. Journal of Optical Society, 2002, 19(6): 1096-1106.
- [10] Achanta R, Hemami S, Estrada F, et al. Frequency-tuned salient region detection [C] // Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition. Florida, USA: IEEE, 2009: 1597-1604.