

## 基于局部特征与全局特征的图像显著性目标检测

陈振学<sup>1,2</sup>, 贺超<sup>1</sup>, 刘成云<sup>1,3</sup>

(1. 山东大学控制科学与工程学院, 济南 250061; 2. 南京信息工程大学江苏省大数据分析重点实验室, 南京 210044; 3. 浙江大学CAD & CG国家重点实验室, 杭州 310058)

**摘要:** 在采用图像谱残差分析方法获取全局特征显著性图像的基础上, 利用小波变换在时域和频域具有的局部特征信息表征能力, 通过对图像包含的不同特征信息进行小波变换, 去除各个特征图中的冗余信息, 得到图像局部特征显著部分, 对两种分析方法下获得的显著图进行融合分析, 获得最终的图像显著部分, 并利用视觉转移机制在原图中勾画出显著性目标. 实验结果分析表明, 改进后的方法提高了图像显著目标检测的准确率.

**关键词:** 谱残差; 小波变换; 局部特征; 全局特征; 目标检测

中图分类号: TP391.4

文献标志码: A

## Image saliency target detection based on global features and local features

CHEN Zhen-xue<sup>1,2</sup>, HE Chao<sup>1</sup>, LIU Cheng-yun<sup>1,3</sup>

(1. School of Control Science and Engineering, Shandong University, Ji'nan 250061, China; 2. Jiangsu Key Laboratory of Big Data Analysis Technology, Nanjing University of Information Science and Technology, Nanjing 210044, China; 3. State Key Laboratory of CAD & CG, Zhejiang University, Hangzhou 310058, China. Correspondent: CHEN Zhen-xue, E-mail: chenchenxue@sdu.edu.cn)

**Abstract:** The method of spectral residual analysis is adopted to get the global features' saliency map, and the wavelet transforms' local feature information representation ability in the time domain and frequency domain is used to remove redundant information and get local features highlights. The two ways are combined to get the final-saliency map and block out the saliency target. Experimental results show that the detection accuracy is improved by using the proposed method.

**Keywords:** spectral residual; wavelet transform; local features; global features; target detection

### 0 引言

实验心理学家 Treicher 做过两个著名的心理实验, 实验表明, 大部分人类日常生活中获取的信息 80% 以上是图像信息. 图像作为一种信息的载体, 可通过人的视觉神经系统被感知和理解<sup>[1]</sup>, 人类视觉机制能从复杂的自然环境中抑制冗余的信息, 获取简单但重要的视觉注意信息<sup>[2-5]</sup>. 伴随着计算机技术和多媒体技术的发展, 图像的处理显得愈发重要, 模拟人类视觉机制, 获取图像显著性部分已经成为图像处理应用的一个重要发展方向<sup>[6]</sup>. 图像显著性检测技术通过模拟人的视觉检测机制获取图像中的重要信息<sup>[7]</sup>, 对于智能机器人、智能交通以及其他智能领域具有重要意义. 经典的图像显著性检测理论方法包括: Itti

等<sup>[8]</sup>提出的模拟生物视觉显著模型, Harel 等<sup>[9]</sup>提出的基于图形的视觉显著性算法 (GBVS), Achanta 等<sup>[10-11]</sup>提出的基于频率通道显著性检测 (FT) 算法以及 Hou 等<sup>[12]</sup>提出的谱残差全局显著性检测方法. 他们分别从不同的角度和理论对图像显著性进行了研究, 并取得很大的进展. Hou 对图像进行谱残差分析, 是统计意义上的显著性检测, 属于全局特征的显著图像分析方法. 香农信息论编码思想认为, 一幅图像一般可以分为敏感部分和不敏感部分<sup>[13]</sup>, 也就是图像显著性理论中的图像显著性部分和不显著性部分. 小波变换能够有效地去除图像冗余信息, 通过对图像进行多种特征的小波分析可以得到一幅图像的敏感部分<sup>[14]</sup>.

本文研究方法是在采用谱残差图像显著性检测

收稿日期: 2015-08-09; 修回日期: 2015-10-27.

基金项目: 国家自然科学基金项目(61203261, 61273277); 山东省自然科学基金项目(ZR2012FQ003); 浙江大学CAD & CG国家重点实验室开放课题(A1514); 江苏省大数据分析技术重点实验室(南京信息工程大学)开放课题(KXK1404).

作者简介: 陈振学(1977—), 男, 副教授, 博士, 从事生物特征识别、显著性特征提取与目标检测识别、信息融合等研究; 贺超(1991—), 男, 硕士生, 从事模式识别与图像处理、显著性区域检测的研究.

方法获取全局显著特征<sup>[15]</sup>的基础上,利用小波在时域和频域上良好的局部特征表征能力,对图像小波进行分解,去除特征图像中的冗余信息获得局部特征显著图,并通过融合算法获得最终显著图。

## 1 图像显著性模型的建立

从图像统计学理论出发,图像中的冗余信息更接近自然环境,大量文献<sup>[16-17]</sup>表明,图像中的自然环境具有统计意义上的规律。Hou 谱残差及小波变换的检测方法就是在此理论的基础上建立的。HSV 颜色空间符合人类视觉观测机制,因此,采用 HSV 颜色空间对图像的多个特征进行提取并进行多尺度分析,能去除特征图像冗余信息,获得局部特征的显著目标。

首先,对输入图像的亮度、方向、纹理以及颜色进行二维小波分解,即

$$H_i(x) = WT[I_i(x)]. \quad (1)$$

通过式(1)可获取各自的特征图像。

然后,利用二维的小波重构算法对分解后的图像进行重构,从而获得各自特征的显著图像,即

$$S_i(x) = g(x) \times WT^{-1}[c(x) \times H_i(x)]. \quad (2)$$

其中:  $g(x)$  为高斯滤波器,  $c(x)$  为均值滤波器。

最后,通过式(3)对4类特征显著图进行组合运算,获得图像最终的局部特征显著图(LsMap),即

$$LsMap = \frac{\sum_{i=1}^n k_i \times S_i}{\max\{S_i, 0.5\}} - \gamma. \quad (3)$$

其中: LsMap 为局部特征显著图,  $k_i$  为权值系数,  $S_i$  为各个特征的显著图,  $\gamma$  为调节系数(通过对大量图像处理发现其取值范围在 0.2 ~ 0.5 之间时检测效果较好,将在后文的实验结果中进行验证)。谱残差方法是一种普遍意义上的检测方法,其核心思想是抑制频繁出现的特征,同时对非常规特征保持敏感。Hou 在文献<sup>[12]</sup>中指出,自然图像的振幅谱  $A(f)$  进行傅里叶变换后的平均值和频率遵循正比关系,即

$$E\{A(f)\} \propto \frac{1}{f}.$$

通过对大量图像的振幅进行对数变换,结果表明,图像的平均振幅谱的对数呈现出局部线性,用一幅图片的振幅谱减去平均振幅谱的对数,就是图片的显著部分,这是 Hou 提出的谱残差提取图像显著部分理论。谱残差计算公式为

$$R(f) = \log[A(f)] - h_n \times \log[A(f)]. \quad (4)$$

其中:  $h_n$  为  $n \times n$  的均值滤波器,本文中  $n$  取值为 3。图像的谱残差中包含了图像的显著部分信息,对其进行傅里叶反变换和重构即可得到显著图像,具体的计算步骤如下。

Step 1: 对输入图像进行傅里叶变换,并对变换后的图像求取振幅谱  $A(f)$ , 即

$$A(f) = \Re(F[I(x)]); \quad (5)$$

Step 2: 计算图像相位谱

$$P(f) = \Gamma(F[I(x)]); \quad (6)$$

Step 3: 利用式(4)计算谱残差  $R(f)$ ;

Step 4: 对谱残差及相位谱进行反傅里叶变换,并进行重构获得显著图像

$$GsMap(x) = g(x) \times F^{-1}[\exp(R(f) + P(f))]^{-2}, \quad (7)$$

其中  $g(x)$  为高斯滤波器。

Hou 的谱残差图像显著性检测方法与大多数其他检测方法相比,该算法更加简单,且在计算速度上具有明显的优势,对工程上的实际应用具有重要意义。

通过基于小波变换的检测方法对图像的多个特征进行多尺度分析,利用小波局部特征表征能力去除图像冗余信息,获取局部特征显著图,其中多特征多尺度分析降低了该方法的检测速度,且由于其细节检测能力较强,导致部分非显著目标被检测。而 Hou 的方法从全局出发,对多数图像具有较好的检测效果,具有普遍意义上的应用,但是对于部分图像的检测效果并不理想。本文将 LsMap 与全局特征显著图(GsMap)的显著性检测方法结合,通过下式的计算方法获取最终的显著图(FsMap),使显著目标更加准确:

$$FsMap(x) = \frac{LsMap + GsMap}{2} - \eta, \quad (8)$$

其中  $\eta$  为调节系数,显著图流程如图 1 所示。

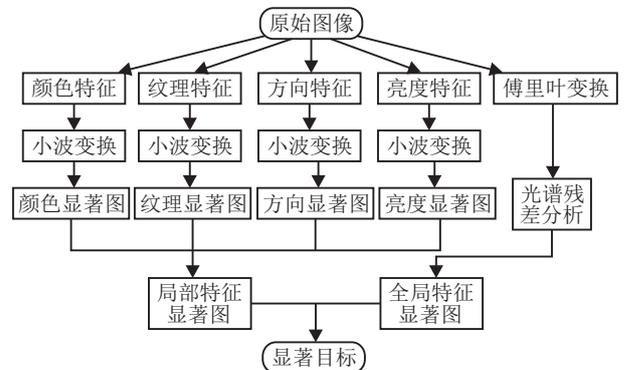


图 1 图像显著性

## 2 实验结果分析

各特征的显著图如图 2 所示。在图 2 中: Local Saliencymap 为小波变换与重构下的图像局部特征显著图,并通过式(3)对颜色、亮度、方向以及纹理 4 种特征的局部特征显著图进行非线性组合运算。本文在利用小波变换进行图像显著性检测时,发现在对颜色背景单一、目标明确的图像进行检测时,显著性效果较好,如图 2(a),但是对于部分颜色鲜艳、纹理特征复杂的图像,其检测效果较差,如图 2(b)。就图 2 的 Global Saliencymap 的实验结果而言, Hou 的方法能够较好地解决小波变换中存在的复杂纹理特征图像显著目标检测问题,如图 2(b);但是,对于部分简单背景

图像反而有更多的冗余信息被保留, 如图 2(a). 因此, 本文利用式 (8) 将两种显著图融合, 其检测效果如图 2 的 Final Sliencymap. 就检测结果而言, 改进后的方法有效地抑制了冗余信息, 突出了显著目标, 增加了检测精度.



图 2 各特征显著图

实验中, 对每一幅输入图像进行手动标记, 获得二值化图像  $O_x$ , “1”作为目标, “0”作为背景,  $S_x$  为显著图像, 则目标检测的命中率 HR 与误检率 FAR 可以通过下式获得<sup>[7]</sup>:

$$HR = E(O_x \times S_x), \quad (9)$$

$$FAR = E((1 - O_x) \times S_x). \quad (10)$$

不同  $\gamma$  值下的显著目标命中率及误检率如表 1 所示.

表 1 不同  $\gamma$  值下的显著目标命中率及误检率

HR/FAR - $\gamma$	$\gamma$ 值						
	0.0	0.1	0.2	0.3	0.4	0.5	0.6
HR (FAR=0.021)	0.055	0.081	0.179	0.485	1.638	5.796	$\infty$
FAR (HR=0.485)	0.187	0.126	0.057	0.021	0.006	0.002	0

不同  $\gamma$  值的显著图如图 3 所示. 在式 (3) 中,  $\gamma$  的取值影响了局部特征的检测效果, 对同一幅图像显著目标的命中率和误检率具有较大的影响, 如图 3 的第 1 幅图为原始图像, 其余图像为  $\gamma$  依次取  $\{0.0, 0.1, \dots, 1.0\}$  时的局部特征显著性图.

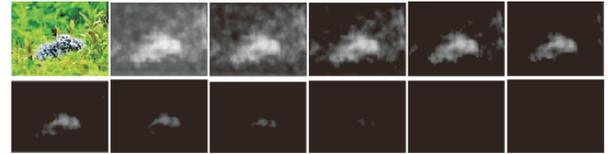


图 3 不同  $\gamma$  值的显著图

从图 3 可以看出: 当  $\gamma$  取值在 0.3~0.5 范围内时, 既有效去除了冗余信息, 又较好地保证了显著目标的完整性; 当  $\gamma$  取值低于 0.3 时, 冗余信息过多; 当  $\gamma$  取值大于 0.5 时, 显著目标的信息已经丢失.

图 4 给出了不同  $\gamma$  值下的显著目标的命中率与误检率的原始数据 HF 曲线.

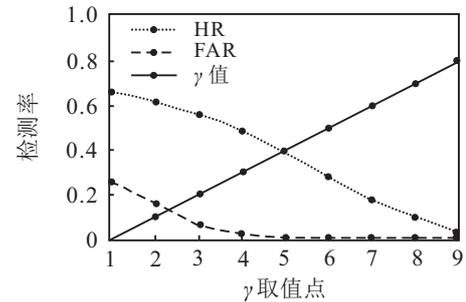
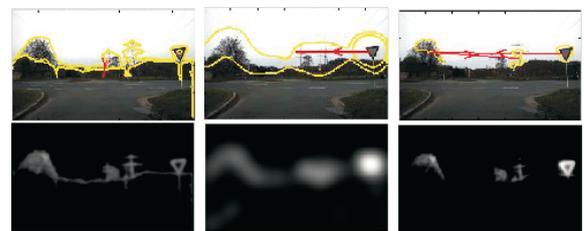


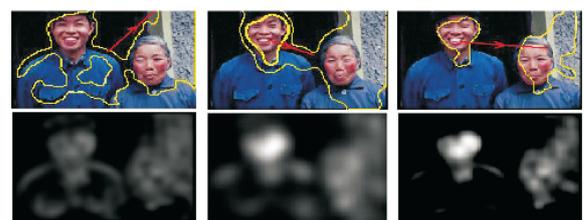
图 4 不同  $\gamma$  值的 HF 曲线

从图 4 可以看出, HR 和 FAR 都随着  $\gamma$  值的增大而减小. 结合表 1 进行分析可知: 在误检率相同的情况下, 目标命中率随着  $\gamma$  值的增大而增大; 在命中率相同时, 误检率随着  $\gamma$  值的增大而减小. 考虑到重要信息不能丢失, 且对多幅图像进行分析可知,  $\gamma$  取值在 0.3~0.5 时效果好. 同理, 在式 (8) 中, 对  $\eta$  值采用相同的方法进行验证, 发现  $\eta$  取值在 0.2~0.4 之间时检测效果较佳.

为了更好地体现特征融合方法的优势, 采用视觉注意转移机制对一幅图像中的多显著目标进行提取, 检测结果如图 5 所示.



(a) 自然场景目标检测



(b) 人工背景目标检测

图 5 视觉注意转移及显著图

在图5中,从左向右分别对应局部特征、全局特征以及特征融合之后的目标检测及相应的显著效果图。当单独采用一种方法进行图像显著性检测时,利用视觉转移机制分析单独的局部特征和全局特征检测方法都会导致大量背景信息的融入,不能很好地将图像中最显著的标志牌以及人脸单独地显示出来。而通过将两种方法结合对图像进行处理可以看出,标志牌和人脸是作为第1显著目标出现的,并且能将背景很好地分离开,比起前面两种方法,其检测结果具有较为明显的优势。

本文分别对图像数为10, 20, ..., 70时的命中率和误检率进行检验,并对图像数为70的数据进行归一化处理,得到表2。

表2 3种检测方法的实验结果对比

	全局特征显著图	局部特征检测	特征融合检测
HR(FAR = 0.017)	0.085	0.184	0.277
FAR(HR = 0.481)	0.044	0.096	0.030

由表2可以看出:当3种检测方法具有相同的误检率时,本文的改进方法命中率更高;当命中率相同时,本文方法误检率最低。综合分析,本文提出的改进方法具有更好的检测性能。本文改进的检测方法主要贡献是进一步增加了显著性目标的检测精度,但同时多个特征的融合会导致检测速度降低。本文采用有形目标图像,对于点目标和小目标图像,其检测效果并不理想,未来会对此类图像显著性检测作进一步研究。

### 3 结 论

小波检测在时域和频域上的局部信息表征能力使其在图像处理领域得到了广泛的应用,本文采用图像冗余信息分析方法,对图像的多个特征进行小波分解与重构,以获取局部特征显著图像;通过对大量图像进行小波和谱残差分析发现,小波处理的局部特征显著性检测以及谱残差全局显著性检测方法对自然背景具有很好的滤除效果,在自然背景下能够获得较好的图像显著目标。

### 参考文献(References)

[1] 卢佩. 基于类生物视觉特性的目标识别[D]. 成都: 中国科学院光电技术研究所, 2014.  
(Lu P. Target recognition based on biological-like vision identity[D]. Chengdu: The Institute of Optics and Electronics, The Chinese Academy of Science, 2014.)

[2] Liu Z. Saliency tree: A novel saliency detection framework[J]. IEEE Trans Image Processing, 2014, 23(5): 1937-1952.

[3] Luo S, Li L. Efficient saliency detection using regional color and spatial information[C]. Visual Information Processing. Paris: Praeger Publishers, 2013: 184-189.

[4] Ren Z. Region-based saliency detection and its application in object recognition[J]. IEEE Trans on Circuits and Systems for Video Technology, 2013, 24(5): 769-779.

[5] Xie Y, Lu H. Bayesian saliency via low and mid level cues[J]. IEEE Trans on Image Processing, 2012, 22(5): 1689-1698.

[6] Cao C, Liu J. Automatic image cropping via the novel saliency detection algorithm[C]. IEEE Trans on Software Engineering and Service Science. Beijing: IEEE Press, 2013: 955-958.

[7] Song M. Low-level and high-level prior learning for visual saliency estimation[C]. Information Science. New York: North-Holland, 2014: 573-585.

[8] Itti L. Models of bottom-up and top-down visual attention[D]. Pasadena: California Institute of Technology, 2000.

[9] Harel J, Koch C, Perona P. Graph-based visual saliency[C]. Advances in Neural Information Processing Systems. Massachusetts: MIT Press, 2007: 545-552.

[10] Achanta R, Estrada F, Wils P, et al. Salient region detection and segmentation[C]. Int Conf on Computer Vision Systems. Germany Prince: Springer, 2008: 66-75.

[11] Achanta R, Hemami S, Estrada F, et al. Frequency-tuned salient region detection[C]. IEEE Int Conf Computer Vision and Pattern Recognition. Prince: Springer, 2009: 1597-1604.

[12] Hou X, Zhang L. Saliency detection: A spectral residual approach[C]. Proc IEEE Int Conf Computer Vision and Pattern Recognition. Prince: Springer, 2007: 1-8.

[13] 涂小坡. 图像显著性算法和评价研究[D]. 南京: 南京航空航天大学理学院, 2011.  
(Tu X B. Image significant algorithm and evaluation research[D]. Nanjing: College of Science, Nanjing Aeronautics and Astronautics University, 2011.)

[14] Imamoglu N, Lin W, Fang Y. A saliency detection model using low-level features based on wavelet transform[J]. IEEE Trans on Multimedia, 2012, 15(1): 96-105.

[15] Goferman S, Zelnik-manor L, Tal A. Context-aware saliency detection[C]. Proc IEEE Int Conf Computer Vision and Pattern Recognition. Prince: Springer, 2010: 2376-2383.

[16] Gluckman J. Order whitening of natural images[C]. IEEE Int Conf on Computer Vision and Pattern Recognition. Prince: Springer, 2005: 354-360.

[17] Martin D. Learning to detect natural image boundaries using local brightness, color and texture cues[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2004, 26(5): 3385-3395.