

文章编号:1004-9037(2014)05-0801-08

基于全局和局部特征融合的显著性提取方法

王红艳¹ 高尚兵²

(1. 南京财经大学经济管理实验教学中心, 南京, 210046; 2. 淮阴工学院计算机工程学院, 淮安, 223001)

摘要:显著性提取方法在图像处理、计算机视觉领域有着广泛的应用。然而,基于全局特征和基于局部特征的显著性区域提取算法存在各自的缺点,为此本文提出了一种融合全局和局部特征的显著性提取算法。首先,对图像进行不重叠地分块,当每个图像块经过主成分分析(Principle component analysis, PCA)映射到高维空间后,根据孤立的特征点对应显著性区域的规律得到基于全局特征的显著图;其次,根据邻域内中心块与其他块的颜色不相似性得到基于局部特征的显著图;最后,按照贝叶斯理论将这两个显著图融合为最终的显著图。在公认的三个图像数据库上的仿真实验验证了所提算法在显著性提取和目标分割上比其他先进算法更有效。

关键词:视觉显著性;主成分分析;特征提取;显著图

中图分类号:TP391

文献标志码:A

Saliency Detection Based on Fusion of Global and Local Features

Wang Hongyan¹, Gao Shangbing²

(1. Experimental Teaching Center for Economics and Management, Nanjing University of Finance and Economics, Nanjing, 210046, China;

2. Faculty of Computer Engineering, Huaiyin Institute of Technology, Huai'an, 223001, China)

Abstract: The saliency detection methods have been widely used in the field of image processing and computer vision. However, the saliency detection algorithms via global feature and local feature extraction have shortcomings. Therefore, a significant saliency detection algorithm is proposed based on fusion of global and local features. Firstly, an image is partitioned to non-overlapped blocks. When each image block is mapped to high dimensional space by principle component analysis(PCA) method, according to the law that the isolated feature points correspond to the salient regions, the saliency map based on the global features is obtained; Secondly, based on the color dissimilarities between center block and its neighborhoods, the saliency map via the local features is obtained; Lastly, based on the Bayes theory, the two obtained saliency maps are fused to the final saliency map. The simulation results on three public image database verify that the proposed algorithm can combine the significant advantages of the global and the local saliency detection algorithms, and it is more effective on saliency detection and object segmentation compared with other state-of-art algorithms.

Key words: vision saliency; principle component analysis(PCA); feature extraction; saliency map

引 言

视觉显著性是为生物和人工视觉系统提供解决方案的一种概念,它是一个检测与周围与众不同的图像区域的过程(通常称之为自下而上的显著性)。

最近,对视觉显著性建模已经引起理论和应用中很大的兴趣^[1]。大多数模型都遵循特征融合理论^[2]。文献[3]提出这个理论的计算框架,文献[4]提出了该框架的一种算法,主要思想是在并行计算关于若干特征中的某一个显著性(比如颜色、灰度、方向,显著性就是一个区域和它周围的相对差值),然后融合

它们到一个标量图,该标量就称之为“显著图”。

根据人类的视觉特征,对目标的感知包括整体感知和局部感知。全局特征体现了图像整体的环境信息,而局部特征表现了局部区域与其邻域的特征的不同。人们选择性地通过整体感知和局部感知来认识世界。现有的显著性检测文献大致分为 3 类:基于局部特征的显著性目标提取^[5]和基于全局特征的显著性目标提取^[6]以及结合局部和全局特征的显著性目标提取方法^[7]。基于局部特征的显著性目标提取方法一般通过采用局部 Center-surround 操作计算图像的显著性,当处于中心位置的图像块与它周边的图像块差异非常小,但该局部区域在整个图像中属于显著性区域时,局部显著性度量不能检测到全部的显著区域,这将导致显著性目标边界是显著的,而显著性目标中心存在“黑洞”。基于全局特征的显著性检测则通过在整个场景中计算特征的稀少性来衡量目标的显著性。图像中存在的噪声作为稀有特征,容易影响显著性目标的检测。结合局部和全局特征的显著性目标提取方法在计算图像显著性时分别考虑局部特征和全局特征,并将两者有机地结合在一起。

本文提出一种新的基于局部和全局特征的显著性检测方法(Local and global features based saliency detection method, LG)。该算法对图像进行不重叠地分块后,对这些不重叠的分块分别采用主成分分析(Principle component analysis, PCA)降维得到全局显著图 S_G 和中心像素块与邻域像素块的颜色不相似性得到局部显著图 S_L ,然后基于贝叶斯理论融合 S_G 和 S_L 形成最终的显著图 S_{LG} 。最后,在不同的数据库上进行测试,验证了文中算法的优越性能,并与 Grabcut 方法^[8]结合对显著图进行分割,分割的准确性优于目前流行算法。

1 全局显著性

基于图像块的显著性计算属于全局显著性特征提取方法,它通过对图像进行分块,然后通过 PCA 映射到高维空间里,孤立的特征点一般对应着原始图像的显著性区域。

1.1 图像块表示

对于维数为 $H \times W$ 的图像 I ,算法将其分解为一系列大小为 $n \times n$ 的图像块。为简便起见,所提算法不考虑图像块的重叠问题。在图像边界部分,对于不足图像块大小的图像块则直接丢弃。图像块的总数为 $L = \lfloor H/n \rfloor \cdot \lfloor W/n \rfloor$ 。对于任一图像块 $p_i, i=1, 2, \dots, L$,将其表示成向量形式 x_i ,最终得到一个表示图像块的矩阵 $\mathbf{X} = [x_1, x_2, \dots, x_L]$ 。

1.2 特征抽取

为了去除图像块中的噪声和节省计算时间,所提算法使用 PCA^[9]抽取特征。对图像块矩阵 $\mathbf{X} = [x_1, x_2, \dots, x_L]$,计算 $\bar{x} = \sum_{i=1}^L x_i / L$, $\hat{\mathbf{X}} = [x_1 - \bar{x}, x_2 - \bar{x}, \dots, x_L - \bar{x}]$,对应的协方差矩阵为

$$\mathbf{A} = (\hat{\mathbf{X}}^T \hat{\mathbf{X}}) / L^2 \quad (1)$$

对 \mathbf{A} 进行特征值分解,选择前 d 个最大特征值对应的特征向量 $\mathbf{U} = [u_1, u_2, \dots, u_d]^T$,其中 \mathbf{U} 的维度为 $d \times L$ 。

图像分块后映射到高维空间里,如图 1 所示。根据图 1 发现:(1)数量占绝大多数的背景块因为具有相同的颜色、亮度、纹理等,通过映射后聚集在一起;(2)显著性目标对应的特征点一般是野点(或孤立点),远离高维空间中背景块对应的特征点的中心。

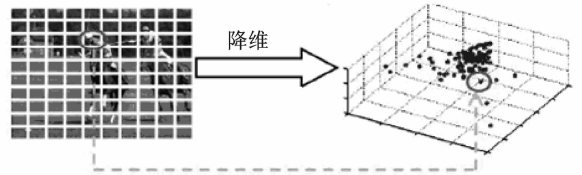


图 1 图像分块映射到高维空间

Fig. 1 Image blocks mapping to high dimensional space

1.3 图像块的全局显著性

在高维空间里的 N 个特征点用 p_x 表示,那么中心点 p_A 在 L_1 范式下表示为

$$p_A = \frac{1}{N} \sum_{x=1}^N p_x \quad (2)$$

如果特征点 p_x 离 p_A 越远,亦即 p_x 不在以 p_A 为中心的聚类里的可能性越大,那么该特征点显著的可能性就越大。因此,计算每个特征点到中心点的 L_2 距离,然后排序,选择到中心点距离最大的前 25% 的特征点。

由于没有显著目标尺寸大小的先验信息,因此为了提高准确率需要在多尺度情况下计算,为此对图像采取 3 种分辨率:100%, 50% 和 25%,然后求平均得到全局显著图 S_G 。

2 局部显著性

图 2 表示图像块局部显著性度量的邻域选择的方式,文中模型的 S_L 是中心像素块 i (图 2 中小方块表示) 和它的矩形邻域(图 2 中大矩形表示)里的 K ($K=24$) 个图像块的不相似度的平均值来衡量,局部显著性式

$$S_L(x_i) = \frac{1}{K} \sum_{j=1}^K W_{ij}^{-1} d(x_i, x_j) \quad (3)$$

式中: W_{ij} 是图像块 x_i 和 x_j 之间的欧氏距离, $d(x_i, x_j)$ 表示 x_i 和 x_j 的不相似性。

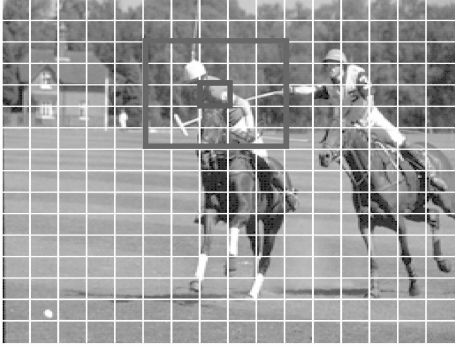


图 2 图像块局部显著性度量的邻域选择

Fig. 2 Neighborhood selection by local saliency metric of image blocks

本文用颜色显著性来衡量不相似性,因为颜色是显著性的至关重要的因素,尤其对有些图像,只根据颜色显著性就可以找到其显著区域,如图 3 所示。



(a) 原图 (a) Original image (b) 颜色显著图 (b) Color saliency map

图 3 颜色显著性

Fig. 3 Color saliency

特别地,考虑到图像像素 I_a 和 I_b ,不相似性可以定义为

$$d(I_a, I_b) = D(I_a, I_b) \quad (4)$$

式中: $D(I_a, I_b)$ 为像素 I_a 和 I_b 在 $L * a * b$ 空间的颜色距离。假设 $I_a \in x_i, x_j$ 为邻近块,那么 I_a 和 x_j 的非相似性如下

$$d(I_a, x_j) = \frac{1}{|x_j|} \sum_{I_b \in x_j} D(I_a, I_b) = \frac{1}{n_j^2} \sum_{I_b \in x_j} D(I_a, I_b) \quad (5)$$

式中: $|x_j|$ 表示块 x_j 的像素的个数。该定义可以很容易理解为在忽略空间关系的情况下,具有相同颜色值的像素具有相同的非相似值。从而可以将式(5)中具有相同颜色值 c_j 的项合并,得到每个颜色的非相似值

$$d(I_a, x_j) = d(c_a) = \sum_{b=1}^{n_j} p_b \times D(c_a, c_b) \quad (6)$$

式中: c_a 为像素 I_a 的颜色值, n_j 表示区域 x_j 包含

的不同的颜色值的个数, p_b 为区域 x_j 中颜色 c_b 出现的概率。如果 $I_a \in x_i$,那么区域 x_i 和 x_j 之间的非相似性可以写成

$$d(x_i, x_j) = \sum_{I_a \in x_i} \sum_{I_b \in x_j} D(I_a, I_b) = \sum_{a=1}^{n_i} \sum_{b=1}^{n_j} p(c_{i,a}) p(c_{j,b}) D(c_{i,a}, c_{j,b}) \quad (7)$$

式中: $p(c_{s,a})$ 表示第 a 个颜色 $c_{s,a}$ 在第 s 个区域 x_s 对应的所有 n_s 个颜色中发生的概率, $s = \{i, j\}$ 。由于颜色直方图比其他特征描述算子更鲁棒,那么根据式(7)可以很容易地使用颜色直方图来表示区域 x_i 和 x_j 。区域 x_i 和 x_j 的非相似性可以由对应的颜色直方图的差值计算得到。

因此,根据式(3)就可以得到局部显著图 S_L 。

3 融合全局和局部显著性

上述的全局和局部显著性方法各有优点和缺点,如图 1,3 所示。每个方法对一些图像提取显著性区域效果不错,但是不会对所有图像效果都好,这两种方法经常互补。因此,对全局和局部显著性融合后的结果很可能比其中一种显著图效果更好。

给定一幅图像 I 的全局显著图 S_G 和局部显著图 S_L ,融合后的显著图在像素 p 的显著值 $S(p)$ 可以表示为 $P(y_p = 1 | S_G(p), S_L(p))$ 。

因为 S_G 和 S_L 是相互独立的,所以

$$P(y_p = 1 | S_G(p), S_L(p)) \propto \frac{1}{Z} P(y_p = 1 | S_G(p)) P(y_p = 1 | S_L(p))$$

式中: $S_G(P)$ 和 $S_L(P)$ 分别代表显著图 S_G 和 S_L 的对应像素 p 的显著值; y_p 为二值随机变量,当 p 是显著像素时, $y_p = 1$, 否则为 $y_p = 0$; $P(y_p)$ 表示图像像素 p 属于显著性目标的概率, $P(y_p = 1 | S_G(p))$ 表示全局显著图给定时 p 是显著像素的概率, Z 是常数。

$P(y_p = 1 | S_G(p))$ 与 $P(y_p = 1 | S_L(p))$ 是相乘的关系,那么其中一个显著图中取值小会抑制另一个,因此需要考虑使用下面的线性组合, S_{LG} 代表融合全局和局部的显著图

$$S_{LG}(p) = P(y_p = 1 | S_G(p), S_L(p)) \propto \frac{1}{Z} (\zeta(S_G(P)) + \zeta(S_L(P))) \quad (8)$$

函数 ζ 有 3 种选择: $\zeta_1(x) = x$, $\zeta_2(x) = \exp(x)$, $\zeta_3(x) = \frac{-1}{\log(x)}$ 。选择 ζ_2 和 ζ_3 是为了避免负值,对更显著的区域赋予更大的权重。

4 实验结果与分析

实验在 Intel Dual Core E2160 1.8 GHz CPU

和 2 GB RAM 的机器上运行。实验所用数据库包括: Bruce&Tsotsos 数据库^[6], MIT 数据库^[10] 和 Achanta 数据库^[11]。

4.1 Bruce&Tsotsos 数据库

为验证本文算法在 Bruce&Tsotsos 数据库上有效性,将 LG 算法在 Bruce&Tsotsos 数据库上运行,并且将实验结果与本文算法 LG 以及其他经典算法进行比较,如 AIM^[6],SUN^[12],ICL^[13],GB-

VS^[5],SR^[14],SF^[15]。

为了评价在 Bruce&Tsotsos 数据库上的整体性能,采用接受者操作特征(Receiver operating characteristic,ROC)曲线与坐标轴围成的区域的面积(Area under the curve,AUC)给出上述所有算法的统计结果。表 1 中给出了不同算法的 AUC 值。从表 1 中可以看出,算法 LG 要优于其他算法。

图 4,5 为不同数据库不同算法得到的显著图。

表 1 Bruce&Tsotsos 数据库上不同算法的性能比较

Table 1 Performance comparison of different algorithms on Bruce&Tsotsos database

算法	Human	LG	AIM	SUN	ICL	GBVS	SR	SF
AUC	0.913	0.789	0.764	0.664	0.696	0.684	0.712	0.721



(a) 原图 (b) 人工 (c) LG (d) AIM (e) SUN (f) ICL (g) GBVS (h) SR (i) SF
(a) Original image (b) Human (c) LG (d) AIM (e) SUN (f) ICL (g) GBVS (h) SR (i) SF

图 4 Bruce&Tsotsos 数据库中不同算法得到的显著图

Fig. 4 Saliency images by different algorithms on Bruce&Tsotsos database

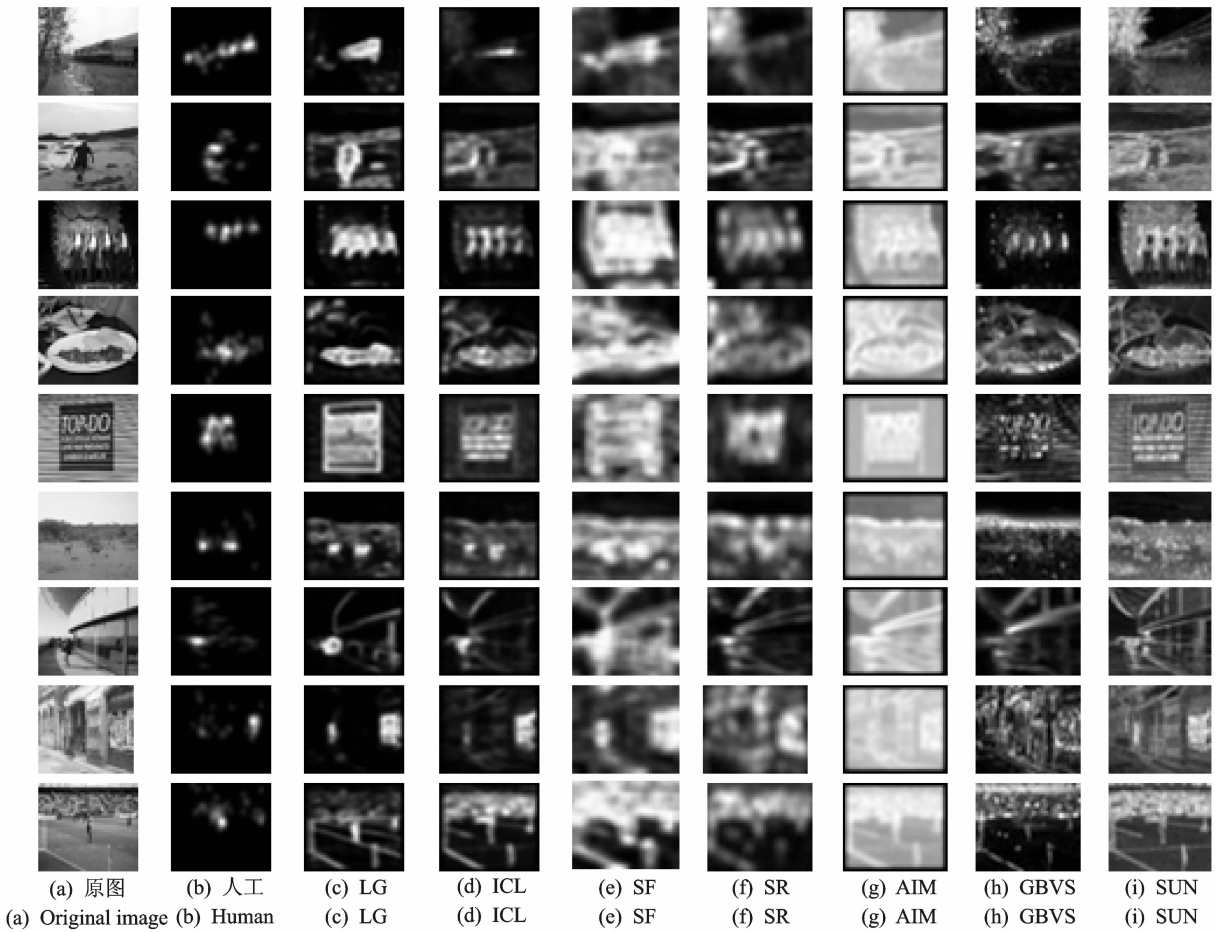


图 5 MIT 数据库上不同算法的显著图

Fig. 5 Saliency images by different algorithms on MIT database

从图 4 中可以看到,AIM 算法能够检测到图像中的显著性目标,但是受噪声影响,许多非显著性目标同样被检测出来。SUN,ICL,GBVS 和 SR 算法能够将图像中大部分显著性目标检测出来,对于图 5 中第 8 行到第 10 行图像中的行人和汽车却检测失效。SF 算法要优于其他经典算法,对于背景单一的图像,SF 算法能够有效地检测其中的显著性目标;当图像背景比较复杂,背景噪声特征与目标类似时,该算法虽然能检测到显著性目标,但是背景中的噪声同样也被认定为显著性目标而被检测出来。

4.2 MIT 数据库

该数据集来源于 Flickr 图像数据库和 LabelMe 图像数据库,包含 1 003 张图像,其中有 779

张景观图像和 228 张肖像图像,每幅图像的最大维度为 1 024,另一维度则在 405 到 1 024 之间变化。15 名受试者通过自由观看图片 1~3 s,对观看的图片进行标定,得到人工标定的数据集。图 5 中第 1 列给出了 MIT 数据库中的部分图像,为便于显示,将图像均缩放至同一维度。

在 MIT 数据库上,ICL 得到的结果比在 Bruce&Tsotsos 数据库上得到的结果要准确。相对于其他算法,算法 LG 检测得到的显著区域更接近人工标定数据。而且从表 2 可以看出,LG 算法的 AUC 值最高,表明它的性能也最好。

4.3 Achanta 图像数据库

图 6 中为不同算法在 Achanta 图像数据库上

表 2 不同算法在 MIT 数据库上的性能比较

Table 2 Performance comparison of different algorithms on MIT database

算法	Human	LG	ICL	SF	SR	AIM	GBVS	SUN
AUC	0.944	0.842	0.743	0.702	0.718	0.765	0.683	0.680

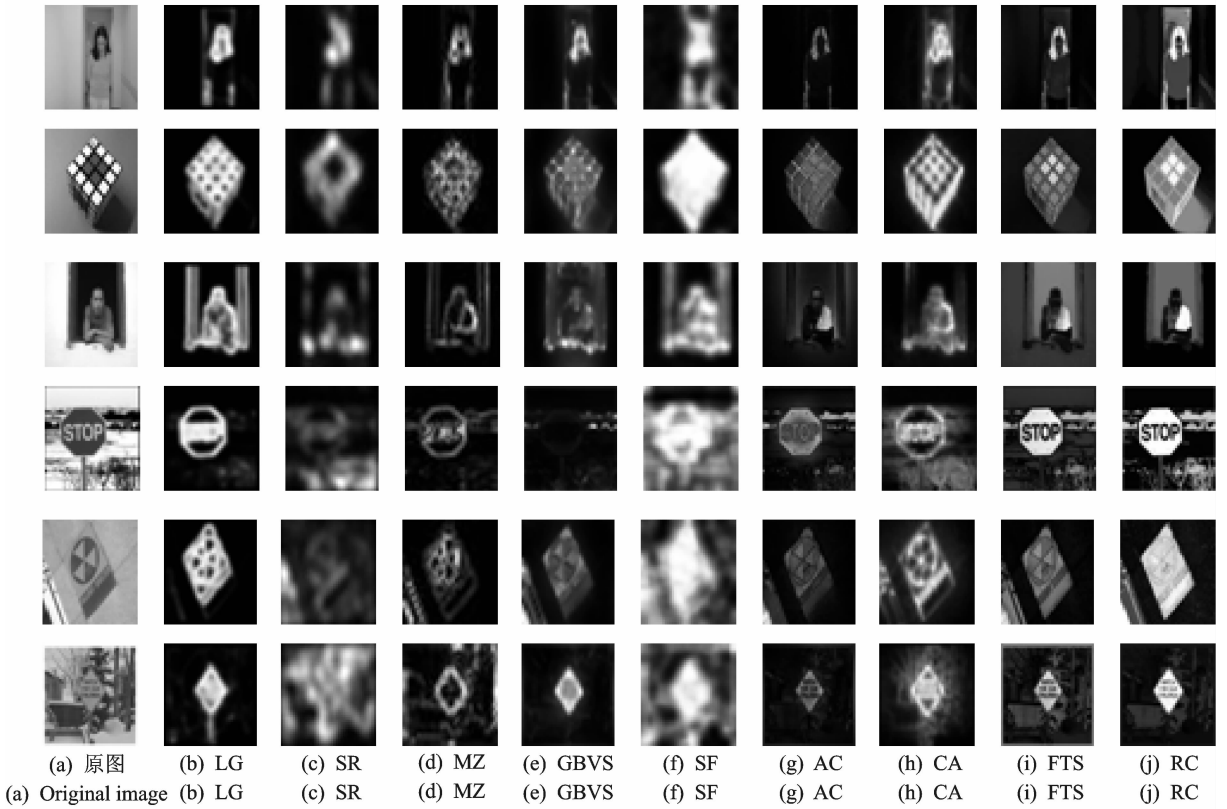


图 6 不同算法在 Achanta 图像数据库上的显著图

Fig. 6 Saliency images by different algorithms on Achanta database

的显著性提取结果,从左至右分别为原图, LG, SR, MZ^[16], GBVS, SF, AC^[17], CA^[18], FTS^[11], RC^[8]得到的显著图。可以看出, LG 显著图效果最好。

为了验证 LG 算法的运行效率,将该算法与其他算法在 Achanta 图像数据库上进行测试。表 3 给出了每个方法的计算一幅显著图的平均运行时间。算法 LG, SR, SF, CA 编程使用 Matlab, 其他算法使用 C++。可以看出, LG 算法计算复杂度并不高。

因为 Achanta 图像数据库一般测试的是显著性算法目标的分割性能, 所以为了验证 LG 的分割

性能,将在 Achanta 图像数据库上进行验证,将 LG 与 Grabcut 算法结合(简称 LGC),并与基于区域对比度的分割(Region contrast cut, RCC)算法^[8]进行比较,两者都使用 Grabcut 算法。LGC 与 RCC 算法在 Achanta 图像数据库上分割结果比较如图 7,8 所示。从图 7 可以看出, LGC 基本上只分割出显著目标,而 RCC 把背景也分割出来了。

另外,本文采用查准率、查全率和 F(F-measure)指标来衡量算法的优劣, LGC 算法得到的指标(查准率 = 0.925, 查全率 = 0.921, F 指标 = 0.924)整体上要优于 RCC 算法(0.9, 0.9, 0.9)^[8]。

表 3 各个算法的平均用时比较

Table 3 Comparison for the average time of each algorithm

算法	SR	LG	MZ	SF	AC	CA	FTS	RC
时间/s	0.064	0.743	0.070	0.902	0.109	53.1	0.016	0.253
代码	Matlab	Matlab	C++	Matlab	C++	Matlab	C++	C++

5 结束语

本文提出一种新的显著性目标检测方法 LG。

它首先对图像进行分块,其次利用基于局部显著性特征提取和全局显著特征提取方法对图像块提取特征得到不同的显著图;最后利用贝叶斯理论融合得到的全局和局部显著图,得到最终显著图。在 3 种

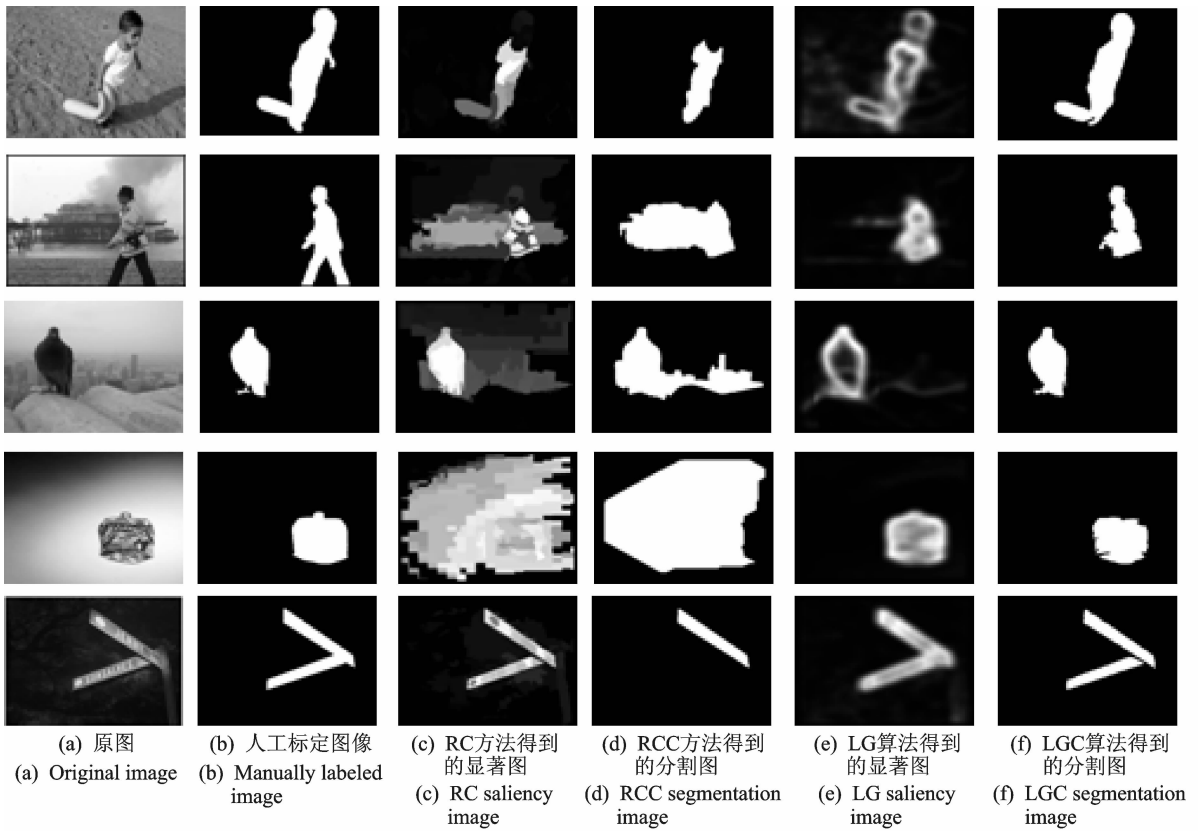


图 7 LG 与 RCC 算法分割图比较

Fig. 7 Comparison of segmentation images by LG and RCC

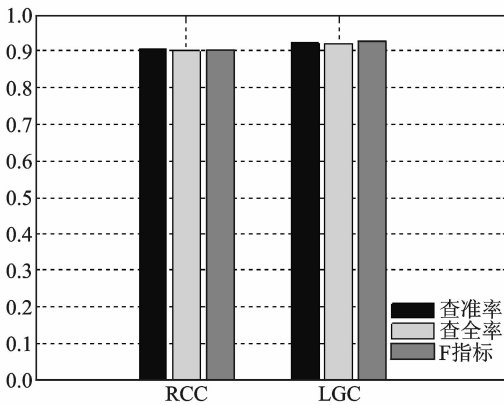


图 8 LGC 与 RCC 算法在 Achanta 图像数据库上的查准率、查全率与 F 指标比较

Fig. 8 Comparison of the precision, recall and F between the LCC method and the RCC method

典型的图像数据库的实验表明, 算法 LG 检测显著目标的准确率要高于目前其他先进的显著性提取算法。

参考文献:

[1] Toet A. Computational versus psychophysical image saliency: A comparative evaluation study[J]. IEEE

Transactions on Pattern Analysis and Machine Intelligence, 2011,33(11):2131-2146.

[2] Treisman A M. Gelade G. A feature integration theory of attention[J]. Cognitive Psychology, 1980,12(1):97-13.

[3] Koch C, Ullman S. Shifts in selective visual attention: Towards the underlying neural circuitry [J]. Human Neurobiology, 1985,4(4):219-227.

[4] 曾孝平, 谢春兰, 刘国金. 基于视觉注意的目标预检测模型[J]. 数据采集与处理, 2010,25(4):469-473.

Zeng Xiaoping, Xie Chunlan, Liu Guojin. Target pre-detection model based on visual attention [J]. Journal of Data Acquisition and Processing, 2010,25(4):469-473.

[5] Harel J, Koch C, Perona P. Graph-based visual saliency[C]//Proceedings of Advances in Neural Information Processing Systems. Vancouver, Canada: MIT Press, 2003: 545-55.

[6] Bruce N, Tsotsos J. Saliency based on information maximization[C]//Proceedings of Advances in Neural Information Processing Systems. Vancouver, Canada: MIT Press, 2006: 155-162.

[7] Liu T. Learning to detect a salient object[J]. IEEE

- Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(2):353-367.
- [8] Cheng M. Global contrast based salient region detection[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Providence, USA: IEEE, 2011: 409-416.
- [9] 文巧钧, 宋执环, 苗爱敏. 基于主元空间动态模型的故障检测方法[J]. 南京航空航天大学学报, 2011, 43(S1):159-162.
- Wen Qiaojun, Song Zhihuan, Miao Aimin. Fault detection approach based on linear dynamical system in principal component space[J]. Journal of Nanjing University of Aeronautics & Astronautics, 2011, 43(S1):159-162.
- [10] Judd T, Ehinger K, Durand F, et al. Learning to predict where humans look[C]// Proceedings of International Conference of Computer Vision Systems. Piscataway, USA: Springer-Verlay, 2009: 8-20.
- [11] Achanta R, Hemami S, Estrada F, et al. Frequency-tuned salient region detection[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Miami, USA: IEEE, 2009: 1597-1604.
- [12] Zhang L, Tong M, Marks T, et al. SUN: A bayesian framework for saliency using natural statistics [J]. Journal of Vision, 2008, 8(7):1-20.
- [13] Hou X, Zhang L. Dynamic visual attention: Searching for coding length increments[C]// Proceedings of Advances in Neural Information Processing Systems. Vancouver, Canada: MIT Press, 2008: 681-688.
- [14] Hou X, Zhang L. Saliency detection: A spectral residual approach[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Anchorage, USA: IEEE, 2008: 1-8.
- [15] Perazzi F. Saliency filters: Contrast based filtering for salient region detection[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Providence, USA: IEEE, 2012: 733-740.
- [16] Ma Y F, Zhang H J. Contrast-based image attention analysis by using fuzzy growing[C]// Proceedings of 11th ACM International Conference on Multimedia. Berkeley, USA: ACM, 2003:374-381.
- [17] Achanta R, Estrada F, Wils P, et al. Salient region detection and segmentation[C]// Proceedings of International Conference of Computer Vision Systems. Santorini, Greece: Springer, 2008:66-75.
- [18] Goferman S, Zelnik-Manor L, Tal A. Context-aware saliency detection[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR). San Francisco, USA: IEEE, 2010: 2376-2383.

作者简介:王红艳(1979-)女,讲师,研究方向:图像处理、模式识别,E-mail:whydhL99@126.com;高尚兵(1981-),男,博士,研究方向:图像处理、模式识别。

