

基于超像素的物体似然概率计算模型研究

纪超,刘慧英,孙景峰,贺胜,黄民主

(西北工业大学 自动化学院,陕西 西安 710072)

摘要: 提出基于超像素建立物体似然概率模型来检测图像的显著区域。首先根据显著性原理和物体的自然属性分析影响物体显著度大小的因素;然后使用 SLIC 算法把图像分成 K 个超像素,并根据纹理、颜色、梯度特征信息建立不同准则下显著物体概率的计算模型;包括类内紧凑性、颜色空间分布估计以及边缘连续性;再结合细胞调节和指数函数的特征对每个准则下的显著物体概率组合得到物体的似然概率;最后利用该算法在较复杂的场景中对显著区域进行提取实验,证明该算法比其他算法更高效。

关键词: 机器视觉; 显著性提取; SLIC 算法; 特征组合

中图分类号: TP911.73 **文献标志码:** A **文章编号:** 1007-2276(2013)11-3156-07

Calculating probability of objectness likelihood model based on superpixels

Ji Chao, Liu Huiying, Sun Jingfeng, He Sheng, Huang Minzhu

(School of Automation, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: Establishing calculation model on the probability of saliency objectness likelihood based on superpixels was introduced to detect image saliency. At first, factors which affected the size of saliency were analyzed according to the principle of saliency and natural characteristics; And then the SLIC algorithm was used to divide image into K superpixels; Next, according to the texture, color and gradient feature information, calculation models were established on probability saliency object under different rules: including compactness in class, color spatial distribution estimation and edge continuity; Moreover, integrating the probability of saliency object under each rule to get the probability of objectness likelihood according to the characteristics of combining with activity in cells responding to stimuli and exponential function; Finally some experiments extracting regions of interest from complex scenes prove that the proposed algorithm is more efficient than other algorithms.

Key words: machine vision; saliency detection; SLIC algorithm; feature combination

收稿日期: 2013-03-10; 修订日期: 2013-04-25

基金项目: 航空基金资助课题(2012ZC53042)

作者简介: 纪超(1987-), 男, 博士生, 主要从事机器视觉与图像处理方面的研究。Email: dachao9898@163.com

导师简介: 刘慧英(1956-), 女, 教授, 博士生导师, 博士, 主要从事机器视觉与图像处理, 计算机控制及网络控制等方面研究。

Email: lhy2005@nwpu.edu.cn

0 引言

视觉注意机制是指视觉系统顶层的观察者对视觉场景中显著区域的选择和处理能力,是建立在复杂的生理结构基础上,涉及神经学、心理学、生物仿生学等多个领域。注意机制在建模时应重点解决两类问题:一类是“最吸引注意的区域具有什么样的特征”;另一类是“如何快速准确地衡量区域的显著度”。模仿并构造人类视觉注意机制的系统,挖掘图像的重要信息,可广泛应用于对感兴趣目标的浏览、内容感知、图像分割、以及图像编辑等方面。

国内外研究显著性检测算法中,Itti 模型^[1]被认为是最具有代表性的算法;Goferman 等人提出的 CA 算法^[2]是根据图像中新颖的特征具有高显著度,通过模仿视觉组织的工作原理建模来突出显著特征;GBVS 算法^[3]是基于图的显著区域提取方法,是在图上建立马尔科夫链,链的平衡状态反映了随机游走在每个节点上所花时间的积累,积累程度与视觉显著度等价;Jian 等人提出的 HFT 算法^[4]认为幅度频域同样包含着重要的显著信息,幅度频域中的峰值信息对应图像中的非显著信息,通过抑制非显著信息得到显著特征;Achanta 等人提出的 AC 算法^[5]属于图论-切割算法,是基于图像的先验信息,通过大量工作试图从图像中分割出显著物体,但此方法对于背景较复杂的图像而言,所得结果的鲁棒性较差;2011 年, Ming ming chen 提出的 HC 算法^[6]是依据像素之间的色彩对比差异来分配像素的显著值,并以此产生高分辨率的显著特征图。

单纯依赖图像中像素间对比的方法,倾向于在显著区域边缘产生高显著值;单纯基于图像整体分布特征的模型,则不易确定和突出显著物体完整的内部信息;文中根据图像中显著特征分布在信息量较大和相位信号改变频繁的区域,也是图像中非周期性的纹理、颜色和边缘等特征变化剧烈的部分的原理,提出基于超像素建立物体似然概率模型来检测图像的显著特征。

1 算法分析

1.1 超像素表示

超像素是把一幅图像分成许多个小区域,区域

内的像素具有极为相似的特征信息。超像素既保持了图像中目标原有的边界信息,又包含比单个像素更多的统计信息,因此,超像素具有表示的集中性和计算的高效性。

传统产生超像素的方法,通常不能控制超像素的大小和数量,也不考虑空间因素的影响。文中基于 SLIC(Simple Linear Iterative Clustering)算法^[8],综合考虑像素间颜色相似性和空间位置距离进行初步聚类,产生期望数量的超像素。

SLIC 算法是用一组 5 维向量 $\{l, a, b, x, y\}$ 作为一个像素点处的特征向量,其中 $\{l, a, b\}$ 为 Lab 空间的色度向量, $\{x, y\}$ 表示该像素的空间位置。设图像中共有 N 个像素,需要分成 K 个超像素,每个超像素约含有 N/K 个像素,每隔 $\sqrt{N/K}$ 个像素点会有一个聚类中心。在 SLIC 算法中衡量相似性的方法是分别计算一个像素与聚类中心的颜色特征之间的距离 L_{lab} 和空间位置距离 L_{xy} , 然后根据不同的权重叠加得到相似度 W_x , 如公式(2)、(3)、(4)所示。

$$L_{lab} = \sqrt{(l_i - l_k)^2 + (a_i - a_k)^2 + (b_i - b_k)^2} \quad (2)$$

$$L_{xy} = \sqrt{(x_i - x_k)^2 + (y_i - y_k)^2} \quad (3)$$

$$W_x = \frac{L_{lab} + \delta L_{xy}}{\sqrt{1 + \delta^2}} \quad (4)$$

式中: k 代表第 k 个超像素的聚类中心; i 代表以该聚类中心为中心的搜索范围内的第 i 个像素; δ 为调节参数, δ 越大,空间距离在衡量两个像素相似性时发挥的作用就越大。文中实验取 δ 为 0.2,图 1 为不同 K 值下的超像素分布图。



图 1 超像素表示图

Fig.1 Superpixels diagram

1.2 显著物体概率计算

聚类只负责把图像分成组,对于每个组内是什么并不关心,因此视觉系统在接收特征刺激的同时需要积极评估每个目标点的显著概率。根据高斯区

域内的特征分布是高斯型,文中提出结合物体自然属性,采用高斯混合模型(GMM)对各超像素建模,通过计算物体似然概率来反映显著物体特征。显著物体的概率是通过类内紧凑性、颜色空间分布估计和边缘连续性三个准则计算得到。

1.2.1 类内紧凑性计算

文中根据显著目标的纹理特征在空间位置上分布是相对紧凑的。采用一组 6 维的特征向量 S 代表超像素的特征,如公式(5)所示。

$$S=[I_x', I_y', I_{xx}'', I_{yy}'', \lambda x, \lambda y] \quad (5)$$

式中: x 和 y 代表超像素在图像中的空间位置; I_x', I_y' 分别代表图像的灰度值在水平和垂直方向上的一阶导数;二阶导数 I_{xx}'', I_{yy}'' 代表了图像的纹理信息; λ 为调节纹理特征与空间位置特征的权重比例参数。由于协方差矩阵可作为融合多种可能相关特征的代表方法,其中对角元元素代表某种特征在区域内的方差,而非对角元元素则代表不同特征之间的关联性^[9],因此文中对每个超像素建立 6×6 维的特征协方差矩阵 Σ ,如公式(6)所示:

$$\Sigma = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{16} \\ \cdots & & \\ \sigma_{61} & \cdots & \sigma_{66} \end{pmatrix} \quad (6)$$

$$\sigma_i = \frac{1}{M-1} \sum_{m=1}^M (f_i^m - u_a)(f_j^m - u_b) \quad (7)$$

式中: σ_{ij} 的计算如公式(7)所示; M 为一个超像素中含有像素的个数; f_i^m, f_j^m 分别代表超像素 a, b 内第 m 个像素的第 i, j 个特征; u_a, u_b 分别代表超像素和中特征的平均值。通过距离度量方法衡量两个协方差矩阵 Σ_a 和 Σ_b 的相似性,如公式(8)所示;并将相似的超像素合并形成 n 个类。

$$\Delta(\Sigma_a, \Sigma_b) = \sqrt{\sum_{i=1}^6 \ln^2(\Sigma_a, \Sigma_b)} \quad (8)$$

文中采用高斯混合模 $\{\pi_c, \mu_c, \sigma_c^2\}_{c=1}^6$ 型^[10] 计算类内相似超像素的紧凑性,类内的超像素 (x_i, y_i) 的纹理特征 f_i 被赋予一个紧凑性概率,如公式(9)所示:

$$p(c|f_i) = \frac{\pi_c N(f_i|\mu_c, \sigma_c^2)}{\sum_{c=1}^6 \pi_c N(f_i|\mu_c, \sigma_c^2)} \quad (9)$$

每个类内图可以看做是一系列纹理的合成。类内超像素总的方差计算如公式(10)所示:

$$\bar{x}_m = \frac{\sum_{m=1}^n p(c|f_i) x_m}{\sum_{m=1}^n p(c|f_i)}, \quad \bar{y}_m = \frac{\sum_{m=1}^n p(c|f_i) y_m}{\sum_{m=1}^n p(c|f_i)} \quad (10)$$

整幅图像的超像素方差计算如公式(11)所示:

$$\bar{x} = \frac{\sum_{k=1}^K p(c|f_i) x_k}{\sum_{k=1}^K p(c|f_i)}, \quad \bar{y} = \frac{\sum_{k=1}^K p(c|f_i) y_k}{\sum_{k=1}^K p(c|f_i)} \quad (11)$$

$$CP = \sum_{k=1}^K [(x_m - \bar{x})^2 + (y_m - \bar{y})^2] \quad (12)$$

CP 表示类内纹理的显著信息量,若信息量越大,则该类就越显著,最后将纹理显著特征叠加为显著特征图 S_r 。

1.2.2 颜色空间分布估计

图像中分布越集中的颜色在整个颜色空间中所占的显著性权重就越大。文中首先确定显著区域的大致位置,将少数不显眼的颜色排除,大大减少计算量和随机性对显著值造成的影响,最后通过估计显著区域中的颜色空间分布确定显著区域的准确信息。

首先通过比较区域信息量筛选出信息量较大的区域。通过公式(13)计算出区域信息量 SL,再通过公式(14)计算出区域内信息量的平均值 \overline{SL} 。

$$SL = \sum_{x,y \in \Omega_{\text{area}}} (\beta E(x,y) + (1-\beta)Q(x,y)) \quad (13)$$

$$\overline{SL} = SL / \Omega_{\text{area}} \quad (14)$$

式中: $E(x,y)$ 和 $Q(x,y)$ 表示熵和能量; β 为平衡图像复杂度中熵和能量的参数,文中实验选取 β 值为 0.3。

由于区域显著度的大小还和区域的位置有关,所以区域的位置越处于图像的中央,则显著度就越大;反之,则越小。因此需要根据该像素到中心像素的距离确定位置权重 PW,计算方法如公式(15)所示:其中 (Cx, Cy) 表示图像的中心位置, (W, H) 表示图像的宽和高:

$$PW = 1 - \frac{2\sqrt{(x-Cx)^2 + (y-Cy)^2}}{\sqrt{W^2 + H^2}} \quad (15)$$

局部特征对比是计算图像显著信息的重要方法。文中采用“中心-周围”领域加权算子 Θ_l 计算,能良好地提取显著物体的整体信息。方法如公式(16)、(17)所示:

$$\Theta_l(x,y) = \sum_{x',y' \in \Omega_{area}} \left\{ \exp \left(\frac{\|SL(x,y) - \bar{SL}(x',y')\|_2}{2\sigma^2} \right) - 1 \right\} PW \quad (16)$$

$$LC(x,y) = \sum_{l=1}^n \Theta_l(x,y) \quad (17)$$

式中: Ω_{area} 是 (x,y) 的领域; l 为尺度变量。文中实验选 Ω_{area} 的大小是 4×4 , 尺度大小为 5。通过叠加所有局部显著特征得到大致显著区域。

然后采用高斯混合模型(GMM)估计区域颜色空间分布。首先,用 $\{\mu_c, w_a, \Sigma_a\}_{a=1}^A$ 分别表示图像中第 a 个颜色成分的均值,权重和协方差,每个像素以公式(18)分配得到一个颜色成分^[10]。

$$p(a|x) = \frac{w_a \mu(l_x | \mu_c, \Sigma_a)}{\sum_a w_a \mu(l_x | \mu_c, \Sigma_a)} \quad (18)$$

每个颜色成分 a 空间位置的水平方差 $V_h(a)$ 如公式(19)所示:其中 x_h 是像素 x 的 x 轴坐标, $M_h(a) = \frac{1}{|X|_a} \sum_x p(a|x) x_h$ 。

$$V_h(a) = \sum_x p(a|x) x_h \cdot |x_h - M_h(a)|^2 \quad (19)$$

垂直方差 $V_v(a)$ 可用类似的方法定义。因此第 a 个颜色成分空间方差可表示为 $V(a) = (V_h(a) + V_v(a)) / p(a|x)$, 颜色空间分布特征 $F_a(x,l)$ 可由公式(20)表示为加权的形式:

$$F_a(x,l) = \sum_a p(a|x) (1 - V(a)) \quad (20)$$

颜色空间方差加强了对显著区域颜色更准确的表达。最后叠加显著特征得到显著特征图 S_g 。

1.2.3 边缘连续性

连续性是表明初级特征的内部一致性,由物体的边界分布来衡量,连续的物体边缘可增大显著度。文中结合概率密度梯度的定义,提出一种基于类的边缘检测方法,给出图像边缘在概率密度梯度场中的描述,从而确定区域显著度。

首先对类中各超像素 (x_i, y_i) 的概率密度 P 通过核函数表示^[10,12], 如公式(21)所示:

$$P = \sum_W \frac{C}{w_s^2 w_l w_{ab}} g \left(\left\| \frac{x_i - x, y_i - y}{w_s} \right\| \right) \cdot g \left(\left\| \frac{l_i - l}{w_l} \right\| \right) \cdot g \left(\left\| \frac{a_i - a, b_i - b}{w_{ab}} \right\| \right) \quad (21)$$

式中: C 为常量; (x_i, y_i) 为以 x 为中心的窗口 W 中的一点; (l, a, b) 为中的均值; g 为高斯核函数; w_s 图像平面空间中的带宽; w_l 和 w_{ab} 为图像空间中的亮度和色差带宽。根据公式(22)、(23)可以推导出各像素点在概率密度场中的梯度 $D(x,y)$ 和散度 $S(x,y)$ 。

$$D(x,y) = (P'_x, P'_y) \quad (22)$$

$$S(x,y) = \frac{dP'_x}{dx} + \frac{dP'_y}{dy} \quad (23)$$

概率密度梯度场中的点 (x,y) 在某直线方向 θ 上的内积响应 R 可通过公式(24)计算得到。

$$R = \{D(x + \cos(\theta), y + \sin(\theta)) \cdot D(x - \cos(\theta), y - \sin(\theta))\} \quad (24)$$

根据边缘检测定义,在图像的概率密度梯度场中寻找散度大于零且内积响应小于零的点作为边缘点,组成边缘区域。由于噪声一般属于小区域,而小区域无法提供显著信息,因此采用对各连续区域的面积或周长大小进行比较来设置阈值去除无效区。文中选取周长大于连续区域边缘总和平均值的区域作为显著区域 S_r 。

2 特征组合

Ansger 等人^[13]提出“V1 假设理论”,认为视觉系统的 V1 细胞对特征之间的相互作用有调节功能; Itti 等人^[1]认为各个特征(颜色,灰度和方向)对于视觉刺激的贡献相同,采用等权值的方法进行特征组合,但对于其他特征,现有技术无法确定其对刺激的响应强度;参考文献[14-15]用实验证明颜色特征与方向特征之间、运动特征与颜色特征之间存在相互作用的关系。显著物体或者区域可能在某些特征上具有较强的反应而在其他特征上没有反应,如果将所有特征直接相加则会带来强烈的噪声,那么真正的显著特征被淹没。为解决此问题,文中基于高斯混合模型,结合细胞调节和指数函数的特性,提出将超像素在各个准则下得到的显著特征图 S_r, S_g 和 S_f 进行概率组合得到物体的似然概率。

首先建立每个超像素的物体似然概率,它是在

每个准则下得到的显著概率的组合模型,如公式(25)所示:

$$p(o_i=1|F_i)=\frac{1}{Z}\sum_{j=1}^K p(o_i=1|\gamma_i^j) \quad (25)$$

式中: Z 为归一化因子。

然后采用每个成分的似然性来推测每个超像素的似然概率。通过高斯混合模型对每个超像素建模,模型中每个成分作为物体似然计算的基本单位。通过公式(26)实现:

$$p(o_i=1|\gamma_i^j)=\sum_{c=1}^6 p(o_i=1|c)p(c|\gamma_i^j) \quad (26)$$

采用高斯混合模型中的各个成分来估计像素特征的分布,如公式(27)所示:

$$p(\gamma_i^j)=\sum_{c=1}^6 \pi_c N(\gamma_i^j|\mu_c, \sigma_c^2) \quad (27)$$

式中:参数 π_c 、 μ_c 、 σ_c 可以由期望最大化算法计算得到, c 为高斯成分变量。

$p(o_i=1|c)$ 表示在各准则下每个成分属于物体的

概率。 $p(o_i=1|c)$ 可以通过公式(28)采用指数函数分布来计算:

$$p(o_i=1|c)=\exp\left(-\frac{S_c^2}{\lambda}\right) \quad (28)$$

根据细胞调节中特征独立和联合作用的方式将 S_r 、 S_g 和 S_f 组合形成显著特征 S' 。具体计算公式如公式(29)所示:

$$S'^2=\frac{1}{N}(S_r^2+S_g^2+S_f^2+S_r*S_f+S_g*S_f+S_r*S_g) \quad (29)$$

式中: N 为归一化因子。

最后整个物体似然概率计算如公式(30)所示:

$$p(o_i=1|F_i)=\frac{1}{Z}\sum_{j=1}^K\sum_{c=1}^6\exp\left(-\frac{S_c^2}{\lambda}\right)\cdot\left[\frac{\pi_c N(\gamma_i^j|\mu_c, \sigma_c^2)}{\sum_{c=1}^6 \pi_c N(\gamma_i^j|\mu_c, \sigma_c^2)}\right] \quad (30)$$

参数 λ 是尺度控制因子,控制着指数函数的形状,图2为不同 λ 对概率组合结果的影响。从图中可

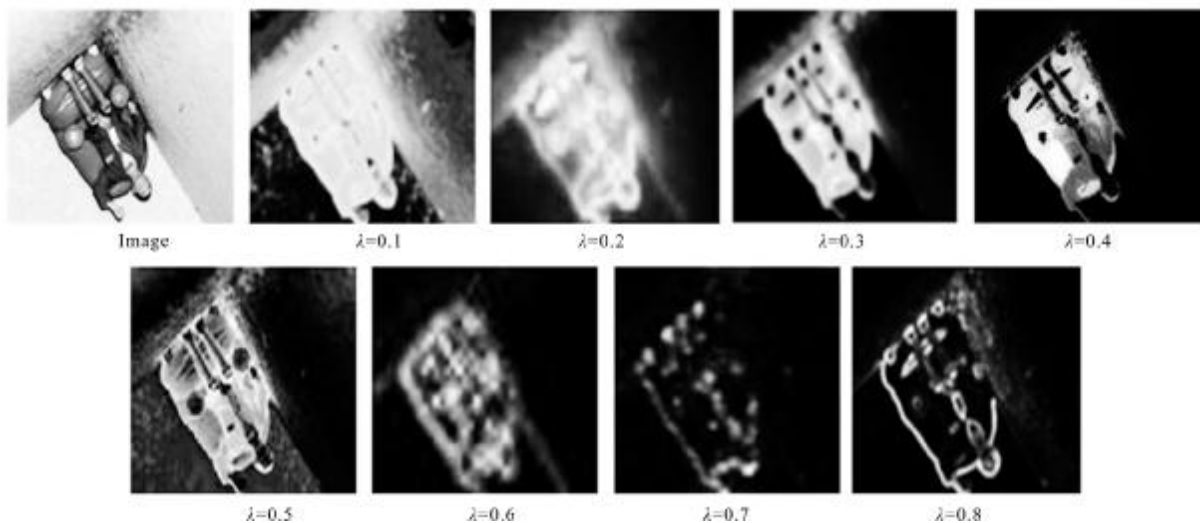


图2 λ 对概率组合结果的影响

Fig.2 Effect of λ on probability

以看出,采用较小的 λ 值,能够产生分布紧凑的显著性图,但包含较大的噪声;较大的 λ 值通常能产生清晰的轮廓,但失去较多的内部信息。因此实验中选择 $\lambda=0.4$ 。

3 实验结果与分析

为验证算法的高效性,文中分别对图像进行显著区域检测的准确度和复杂度对比实验。实验设

备采用的处理器为 Intel 酷睿 i5 处理器,主频 3 GHz,内存 2 GHz 的配置,环境为 Matlab 2010 a。实验选取的原图来自 MSRA 图库。从实验结果对比看出:文中算法比 Itti 算法^[1]、CA 算法^[2]以及 HFT 算法^[4]能更完整清晰地提取显著物体内部信息;比 GVBS 算法^[3]和 HC 算法^[6]能更好地突出显著物体的边缘连续性和准确性;比 AC 算法^[5]能更好地去除背景信息,如图 3 所示。

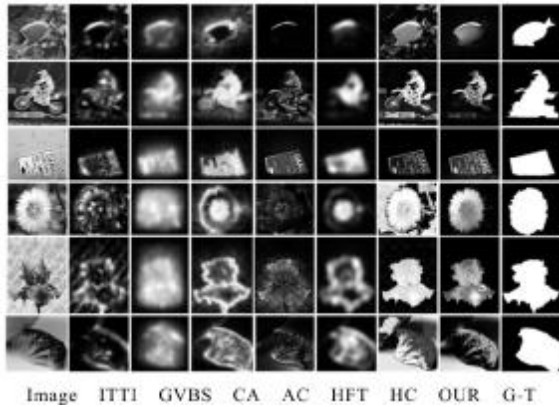


图 3 不同的 Saliency 算法处理的结果图

Fig.3 Detection result by different saliency algorithms

算法准确度对比也可用 ROC 曲线来衡量, 图 4 和图 5 为不同算法的 ROC 曲线和对应的曲线面积图。比较结果得出: 文中算法能更准确地检测出显著区域。

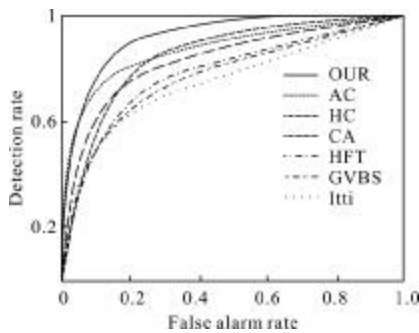


图 4 不同方法的 ROC 曲线

Fig.4 ROC curves of different algorithms

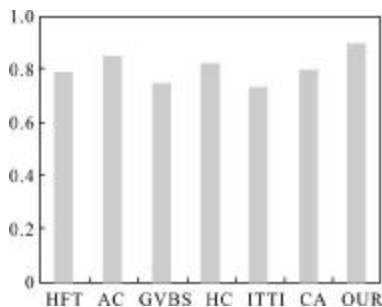


图 5 不同算法 ROC 曲线下的面积

Fig.5 Areas under ROC curves of different algorithms

复杂度对比实验是从 MSRA 库中选取 500 幅图像平均分成 5 组进行处理, 分别计算每组处理方法所用的平均时间, 实验结果如表 1 所示。可以看出: 文中算法比其他算法所用的时间更少。

表 1 不同方法对 100 张图像平均计算时间

Tab.1 Average time of different methods on 100 images

| Method | AC | CA | Itti | GVBS | HFT | HC | OUR |
|--------|--------|--------|------|--------|--------|-----|--------|
| Time/s | 86.9 | 9.2 | 5.9 | 20.6 | 13.7 | 8.6 | 5.2 |
| Code | Matlab | Matlab | C++ | Matlab | Matlab | C++ | Matlab |

4 结 论

文中根据灵长类动物的视觉注意机制, 基于显著性原理和物体的自然属性, 分析影响物体显著度的因素, 最后组合显著物体的概率计算出物体的似然概率。实验证明该算法能更准确更快地检测出图像中显著物体, 更好地提取显著物体的整体信息并确定其边界; 若将此算法应用于相关领域中, 可更有效地提取图像的有效区, 为后续处理提供可靠的信息。因此文中算法具有广泛的应用场合和可观的经济价值。

参考文献:

- [1] 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.
- [2] Goferman S, Zelnik -Manor L, Tal A. Context -aware saliency detection[C]//IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2010.
- [3] Harel J, Koch C, Perona P. Graph-based visual Saliency[C]// Advances in Neural Information Processing Systems (NIPS), 2007.
- [4] Jian L, Levine M D, Xiangjing A, et al. Saliency detection based on frequency and spatial domain analysis [C]// Conference of British Machine Vision Conference, 2011: 1-15.
- [5] Achanta R, Estrada F. Salient region detection and segmentation[J]. Computer Vision Systems, 2008: 66-75.
- [6] Cheng M M, Zhang G X, Mitra N J. Global contrast based salient region detection [C]//IEEE Conference on Computer Vision and Pattern Recognition, 2011: 409-416.
- [7] Bruce N, Tsotsos J K. Saliency, attention, and visual Search: an information theoretic approach [J]. Journal of Vision, 2009, 9(3): 1-24.
- [8] Radhakrishna Achanta, Appu Shaji, Kevin Smith, et al. SLIC

- superpixels[R]. EPFL, 2010: 1-14.
- [9] Xie Y, Lu H. Vision saliency detection based on bayesian model [C]//IEEE International Conference on Image Processing (ICIP), 2011: 653 -656.
- [10] Liu Z, Xue Y, Yan H, et al. Efficient Saliency detection based on Gaussian models[J]. IET Image Processing, 2011, 5(2): 122-131.
- [11] Meer P, Georgescu B. Edge diction with embedded confidence [C]//IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23(12): 1351-1365.
- [12] Jin Ting, Zhou Fugen, Bai Xiangzhi. Moving object detection in airborne video using kenel density estimation [J]. Infrared and laser Engineering, 2011, 40(1): 153-158. (in Chinese)
金挺, 周付根, 白相志. 利用和密度估计的空基视频运动目标检测[J]. 红外与激光工程, 2011, 40(1): 153-158.
- [13] Koene A R, Zhaoping L. Feature -specific interactions in salience from combined feature contrasts: evidence for a bottom-up saliency map in V1 [J]. Journal of Vision, 2007, 7(7): 1-14.
- [14] Feng J, Wei Y, Tao L, et al. Salient object detection by composition [C]//IEEE International Conference on Computer Vision, 2011: 2113-2120.
- [15] Long Yunli, Tang hongbin, Xu Hui, et al. Spatial -temporal fused suppressing of infrared clutter based on adaptive regression Markov model[J]. Infrared and laser Engineering, 2012, 41(3): 791-795. (in Chinese)
龙云利, 唐宏斌, 徐晖, 等. 采用马尔可夫自回归模型的红外杂波时空域融合抑制 [J]. 红外与激光工程, 2012, 41(3): 791-795.