

多尺度的图像显著性检测方法

贾 宁¹, 柳先辉¹, 陈宇飞¹, 赵卫东¹, 邢尚文²

(1. 同济大学 电子与信息工程学院, 上海 201804; 2. 国家电网集团, 山东 济南 250000)

摘要: 为了提高显著性检测算法的准确性与鲁棒性, 提出了一种基于多尺度融合的对象显著性检测方法. 首先对图像进行平滑处理, 过滤掉图像中的高频噪声特征, 然后对图像进行尺度划分并分别采用不同的方法对不同尺度上的图像检测其显著性, 最后根据条件随机场理论对不同尺度上的显著性检测结果进行加权融合, 得到最终的显著性检测结果. 在两种公共数据集上与多种经典算法进行定性、量化比较, 结果表明该算法具有更好的表现.

关键词: 显著性; 多尺度融合; 条件随机场

中图分类号: TP391

文献标志码: A

A Multiscale Image Saliency Detection Method

JIA Ning¹, LIU Xianhui¹, CHEN Yufei¹, ZHAO Weidong¹, XING Shangwen²

(1. College of Electronic and Information Engineering, Tongji University, Shanghai 201804, China; 2. State Grid Corporation of China, Jinan 250000, China)

Abstract: In order to improve the accuracy and robustness of the saliency detection algorithm, this paper proposed a multiscale image saliency detection method. First, the smoothing algorithm was adopted to filter out the noise characteristics in the image. Then, the multiscale representation of an image was performed and saliency maps were computed at different scales. Finally, according to the conditional random field theory, the saliency detection results at different scales were weighted together to get the final results. Extensive experiments in which the proposed method was compared with 9 existing state-of-the-art methods on five benchmark data sets, ECSSD and MSRA10K, show that the proposed method performs better in terms of various evaluation metrics.

Key words: saliency; multiscale fusion; conditional random field

人的眼睛具有快速、准确发现视野中感兴趣物体的能力^[1]. 在机器视觉领域识别感兴趣物体的过程被定义为显著性检测. 显著性检测是机器视觉领域的基础性工作, 其主要目的就是让计算设备模仿人的视觉系统使其具有快速发现图像中需要重点关注区域的能力. 显著性检测理论已经大量应用于图像压缩^[2]、目标识别^[3]、图像分割^[4]等领域, 并且取得了显著的效果. 虽然显著性检测理论已经在多个领域内得到广泛应用, 但是在显著性检测研究中依然存在有许多难点, 其中一个重要的难点就是显著性检测算法的准确率与鲁棒性问题. 由于不同的显著性检测算法的侧重点不同, 导致算法在一个数据集上表现较好, 而在其他数据集上表现较差.

随着机器视觉的发展, 近年来人们提出了很多视觉显著性检测的算法和框架来解决上述问题, 这些模型和框架主要分为两类^[5], 一类是受数据驱动的自下而上的显著性检测方法, 一类是由任务驱动的自上而下的显著性检测方法. 基于数据驱动的方法依靠数据本身的特征来进行显著性检测; 基于任务驱动的方法依靠任务对象的属性来进行显著性检测. 这两类显著性检测方法都存在自身的优点与不足: 自下而上的显著性检测方法具有生物学以及神经学理论的支撑, 对于背景复杂的场景较为有效, 缺点是这类方法仅仅是利用图像的底层特征, 难以刻画图像对象的多义性; 自上而下的显著性检测方法根据显著对象的语义特征来定位显著性的区域, 同时利用图像的场景信息来提高显著性检测的可靠性与准确性, 这类算法能够充分表达图像内容的对象性以及轮廓特征, 但这类方法没有生物学理论的支撑, 在进行显著性检测的时候需要预先给出具体的检测任务来确定检测目标. 近年来随着机器学习以及深度学习技术的发展, 人们开始使用基于深度学习的方法对图像进行处理, 这类方法往往可以取得

收稿日期: 2018-04-19

基金项目: 国家重点研发计划(2017YFB0304102); 上海市科技创新行动计划(18511107400); 中央高校基本科研业务费专项资金

第一作者: 贾 宁(1989—), 男, 博士生, 主要研究方向为机器视觉、图像处理、机器学习. E-mail: 1510501@tongji.edu.cn

通信作者: 柳先辉(1979—), 男, 工学博士, 副研究员, 主要研究方向为机器学习、数据分析、图像处理. E-mail: lxh@tongji.edu.cn

常规方法难以达到的效果.但是基于深度学习的方法在前期的训练过程中需要大量的图片信息,而在某些特定的环境下是无法满足这个前提的,所以常规方法仍有其存在的价值.

本文提出了一种结合图像的底层特征和高层特征的多尺度的显著性检测算法.首先对图像进行滤波处理,将与显著性检测的非相关的细节信息进行平滑^[6].其次利用像素的空间分布先验^[7]和颜色先验^[8],获得像素尺度上的显著图.再次,利用 SLIC (simple linear iterative clustering) 对图像进行分割^[9],抽取图像边界区域的超像素以及待检测超像素的邻接超像素和二邻阶超像素作为待检测超像素的显著支撑区域,计算待测超像素与显著支撑区域在 Lab 颜色空间以及 LM(Leung-Malik)最大响应值上的颜色空间距离来获取超像素尺度上的显著图.第四,利用图割算法^[10]对图像进行多次区域分割,利用先验信息、图像块在 Lab 颜色空间、Hue 等特征进行局部对比获得图像在每一分割尺度下的显著图,然后将不同尺度下的显著图进行合并获得图像块尺度上的显著图.第五,由于不同尺度上的显著图对于图像最终的显著性计算的贡献不同,不同于文献[11]中依靠经验进行加权赋值的方法,本文利用 CRF(conditional random field)框架对每一部分的贡献进行加权,得到理论上更好的显著图.

将本文算法与多种典型算法在两个公开数据集(ECSSD、MSRA 1000)上进行实验对比,发现本文算法在准确性和鲁棒性方面都有较好的表现.

1 相关工作

由于对于图像中显著对象的大小缺乏先验知识,所以在一个尺度上无论采用何种基于对比的显著性检测方法,都不可能取得令人满意的效果.文献[12-14]将图像划分为多个尺度,分别获取每一个尺度上的显著性检测结果,最后对多个结果进行加权融合.文献[13]利用图割算法,通过不断调整图割算法的参数将图片划分为 10 个不同的尺度,利用局部对比方法获得每一尺度上的显著图,最后对各个显著图通过求均值的方式获得最终的显著性检测结果.文献[12]利用超像素分割方法将图像分割为多个超像素,根据相邻超像素特征的相似性对邻接进行合并,将每一张图像划分为四个尺度,然后根据局部对比方法获取每一尺度上的显著图,最后通过加权求和的方式获得最终的显著图.文献[14]对图像

进行超像素分割,利用全局对比方法获得每一个超像素的显著值,并计算相邻超像素的特征差异,根据超像素的显著值与特征差异对超像素进行区域合并,通过不断的重复合并,最终将一张图像表达为一个多尺度的树结构,最后对树结构进行加权求和得到最终的显著图.

目前多尺度融合的方式主要有两种:基于概率论的方法和基于经验的方法.文献[15]采用贝叶斯方法将基于前景种子和基于背景种子获得的显著结果进行加权融合,该方法将其中一个结果作为先验来对另一结果进行概率结算,最后将结果相加得到最终的结果.文献[7]采用学习的方式对多个尺度上的显著图进行融合,同一个测试样本集上将不同方法获得的显著图与真值进行对比,准确度高的方法获得较大的权重.文献[16]以及文献[11]采用最简单的相加和相乘的方式对多个显著图进行融合,而没有衡量不同尺度上显著性检测结果对最终生成结果的影响不同.

图像多尺度的划分大部分是从超像素尺度上开始的,这种方法忽略了图像底层特征的差异^[17],同时显著性检测结果严重依赖于图像超像素分割的结果,如果超像素划分的结果不佳,那么将可能得不到令人满意的显著性检测结果.为了克服上述问题,本文提出的多尺度显著性检测方法将像素划分为最底层的尺度,来进行显著性检测,同时为了克服单一方法的局限性,本文在不同尺度上采用不同的显著性检测方法,这样可以更好地检测出具有不同显著性属性的显著性区域.

2 多尺度图像显著性检测算法步骤

图 1 为本文算法的框架结构图,从中可以看出算法主要分为两大部分:①多层次初始显著性检测结果的获取;②融合多尺度初始显著性检测结果,获得最终的显著图.图 2 为不同尺度上的显著性检测的结果.

2.1 图像平滑

心理学实验表明,人的视觉系统对图像中的中频特征比高频特征更加敏感,图像中的高频特征不仅不会对图像的显著性检测提供帮助,而且还会在一些显著性模型中被当作噪声处理.同时图像大量存在的细小纹理变化,也会给图像的显著性检测造成不利的影 响,所以为减少高频特征对算法的性能的影响,首先对原始输入图像进行平滑处理,实现对

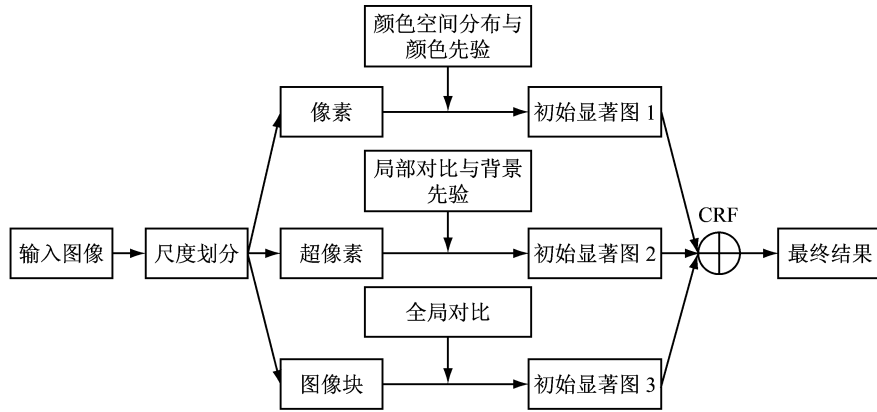


图 1 算法框架图

Fig. 1 Diagram of proposed method

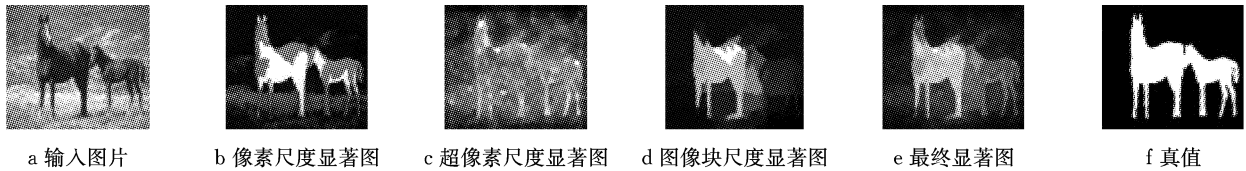


图 2 不同尺度上的显著图

Fig. 2 Saliency maps at different scales

图像中高频特征的抑制. 本文采用文献[6]中的方法对图像进行平滑处理. 平滑处理的目标函数为

$$\arg \min_S \sum_p (S_p - I_p)^2 + \lambda \left(\frac{D_x(p)}{L_x(p) + \epsilon} + \frac{D_y(p)}{L_y(p) + \epsilon} \right) \quad (1)$$

式中: I 为原始图像; S 为平滑后的图像; p 为图像像素; $\lambda=0.07$ 为平滑参数, 决定模型对图像的平滑程度; $\epsilon=2$ 为防止分母为 0 的一个小正数; $D_x(p)$ 与 $D_y(p)$ 分别是以点 p 为中心的矩形框中, 所有像素点沿着 X 轴方向与 Y 轴方向的总变分; $L_x(p)$ 和 $L_y(p)$ 作为区分纹理主要结构的内在变分. 图 3 为图像平滑前后的对比图.

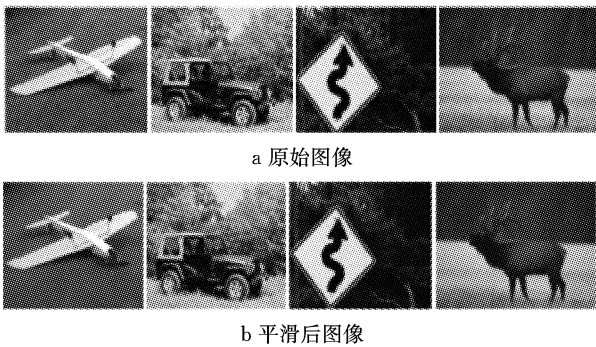


图 3 图像平滑处理
Fig. 3 Smoothed image

2.2 像素尺度

2.2.1 颜色先验

多种颜色通道共同作用使得图像呈现出不同的

色彩, 以 RGB 颜色空间为例, 理论上一幅图像最多含有 255^3 种颜色, 这些颜色根据色彩上的相似性可以简单分为十几类, 颜色类与类之间的色彩差异比较大, 类内的差异比较小. 本文认为一种颜色在图像中的空间分布越广, 空间分布方差越大, 该颜色属于背景的可能性越大. 根据这个前提, 本文利用 GMM (Gaussian mixed model) 模型来表示每个像素点.

$$p(c|I_{(x,y)}) = \frac{\omega_c N(I_{(x,y)} | \mu_c, \sigma_c)}{\sum_c \omega_c N(I_{(x,y)} | \mu_c, \sigma_c)} \quad (2)$$

式中: $I_{(x,y)}$ 为图像中坐标为 (x, y) 的像素点; c 为第 c 个颜色分量; μ_c 和 σ_c 分别为第 c 个分量的均值和方差; $N(\cdot)$ 为高斯模型; ω_c 为第 c 个分量权重. 本文采用最大估计法对式(2)求解.

$$c_{I(x,y)}^* = \arg \max_{c \in C} p(c | I(x, y)) \quad (3)$$

式中: C 为所有的颜色分量, 每个颜色分量的空间坐标位置可以简化为

$$\begin{cases} v_h(c^*) = \frac{1}{N} \sum_{I_{(x,y)} \in c^*} |x - \mu_x(c^*)|^2 \\ v_v(c^*) = \frac{1}{N} \sum_{I_{(x,y)} \in c^*} |y - \mu_y(c^*)|^2 \end{cases} \quad (4)$$

式中: $v_h(c^*)$ 、 $v_v(c^*)$ 分别为第 c^* 颜色分量中所有像素点在 X 方向和 Y 方向上的坐标均值; N 为颜色分量 c^* 中具有最大似然值的个数. 每个颜色分量的空间分布方差定义为

$$V(c^*) = \max(v_h(c^*), v_v(c^*)) \quad (5)$$

最后根据公式: $S_c(I_{(x,y)}) = V(c_{I_{(x,y)}}^*)$ 计算每个像素点的显著值。

通过大量实验发现,当颜色的分量的总数取 4 或者 5 时实验的效果最好,本文选择的颜色分量的个数为 4,图 4 为基于颜色空间分布得出的显著图。

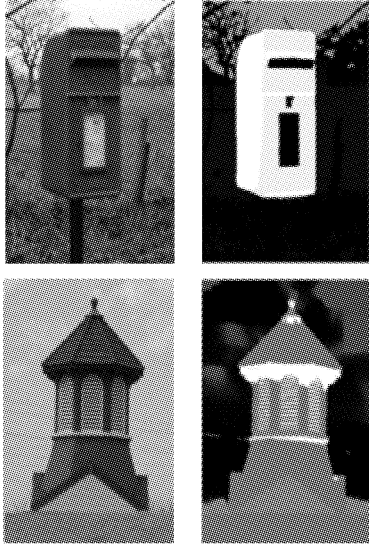


图 4 基于颜色空间分布的显著图

Fig. 4 Saliency maps based on color spatial distribution

2.2.2 空间分布先验

根据文献[11]中的描述,暖色调的物体要比冷色调的物体更能吸引人的注意力,也就是说暖色调的物体比冷色调的物体显著值更高。根据这一结论,本文构建了基于颜色先验的视觉显著性检测方法。在 Lab 颜色空间中, a 通道值的大小代表了颜色从绿色到红色的过度程度,而 b 通道值的大小代表了颜色从蓝色到黄色的过度程度。由于绿色与蓝色属于冷色调,而红色与黄色代表暖色调,所以 a 通道与 b 通道之和衡量了该像素颜色的冷暖程度,也就决定了该颜色显著值的大小。

根据上面的分析,本文采用文献[8]中提出的方法对像素点的显著性进行度量。首先将上述两个颜色通道进行归一化处理。

$$\begin{cases} f_{an}(x,y) = [0,1], & f_{bn}(x,y) = [0,1] \\ f_{an}(x,y) = \frac{f_{an}(x,y) - \min a}{\max a - \min a} \\ f_{bn}(x,y) = \frac{f_{bn}(x,y) - \min b}{\max b - \min b} \end{cases} \quad (6)$$

式中: $f_{an}(x,y)=[0,1]$ 与 $f_{bn}(x,y)=[0,1]$ 分别表示坐标为 (x,y) 的像素 n 在 a 通道与 b 通道的颜色值; $\max a$ 和 $\max b$ 分别表示所有像素点在 a 通道和 b 通道的最大值; $\min a$ 和 $\min b$ 分别表示所有像素点在 a 通道和 b 通道的最小值; $f_{an}(x,y)$ 与

$f_{bn}(x,y)$ 的和反映了像素 n 的显著性。最后通过公式(7)对每个像素点的显著性进行赋值。

$$S_f(x,y) = 1 - \exp\left(-\frac{f_{an}^2(x,y) + f_{bn}^2(x,y)}{\sigma^2}\right) \quad (7)$$

式中: σ 为调节参数,令 $\sigma^2=0.25$ 。结合图像颜色先验,最终像素尺度上的显著值定义为

$$S_p(x,y) = S_f(x,y) + S_c(I_{(x,y)}) \quad (8)$$

图 5 为基于颜色先验获得的显著性检测结果。

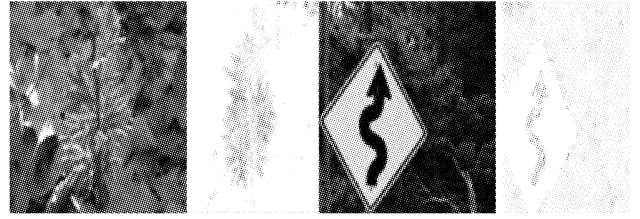


图 5 基于颜色先验的显著图

Fig. 5 Saliency detection based on color-prior

2.3 超像素尺度

基于文献[18]得出两个结论:①距离超像素越近的区域对该超像素的显著性检测结果影响越大;②图像的显著性区域一般处于远离图像边界的区域。

利用 SLIC 算法获得输入图像 I 的超像素表示为 $S=(s_{p_1}, s_{p_2}, \dots, s_{p_N}) \in \mathbf{R}^{D \times N}$ (N 为超像素的个数, D 为特征空间的维数)。然后构建基于超像素的图 $G=(V, E)$, V 为图的顶点(超像素), E 为顶点之间的边。根据图 G 构建起二值邻接矩阵 $A=[a_{ij}]^{n \times n}$, $a_{ij}=1$ 为超像素 i 和 j 邻接, $a_{ij}=0$ 表示超像素 i 和 j 不相交。与中心超像素邻接的所有超像素定义为一阶邻接区域 C_1 , 与一阶超像素邻接的超像素定义为二阶邻接区域 C_2 。基于结论①远离且不在边界区域的超像素对待测超像素的显著性影响基本上可以忽略。基于结论②,提取图像边界区域的所有超像素构建初始背景模板 B_1 , 由于 B_1 中可能存在少量的前景像素,所以利用公式(9)剔除初始背景模板 B_1 中对比度较高的超像素,得到图像的边界超像素集 B , 以其作为背景模板。

$$d(s_{p_{ik}}) = \sum_{s_{p_{ij}} \in B_1} d(s_{p_{ik}}, s_{p_{ij}}) \quad (9)$$

式中: $d(\cdot)$ 为两个超像素在 Lab 颜色空间的欧氏距离。

超像素显著支撑区域的提取见图 6。

将中心超像素的一阶邻接区域、二阶邻接区域以及背景模板 B 总称为超像素的显著性支撑区域。利用文献[19]中的方法提取所有超像素的 Lab 平均

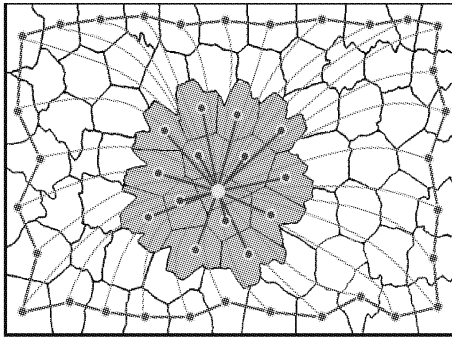


图 6 超像素显著支撑区域的提取

Fig. 6 Salient support region of center superpixel

颜色特征 v_{Lab} 、Lab 直方图特征 h_{Lab} 以及 LM 最大值特征 h_{tex} 构建超像素的特征向量, 两个超像素间的特征差异定义为

$$d(s_{p_i}, s_{p_j}) = \lambda_1 \|v_{Lab,i} - v_{Lab,j}\| + \lambda_2 \chi^2(h_{Lab,i}, h_{Lab,j}) + \lambda_3 \chi^2(h_{tex,i}, h_{tex,j}) \quad (10)$$

式中: $\chi^2(\cdot)$ 为两个超像素之间的卡方距离; λ 为权重参数.

最后超像素的显著性计算公式定义如下:

$$S_{sp}(s_{p_i}) = \omega_1 \sum_{m \in c_1} d(s_{p_i}, s_{p_m}) + \omega_2 \sum_{n \in c_2} d(s_{p_i}, s_{p_n}) + \omega_3 \sum_{j \in B} d(s_{p_i}, s_{p_j}) \quad (11)$$

式中: ω 为权重参数.

图 7 为超像素尺度上图像显著性检测的结果.

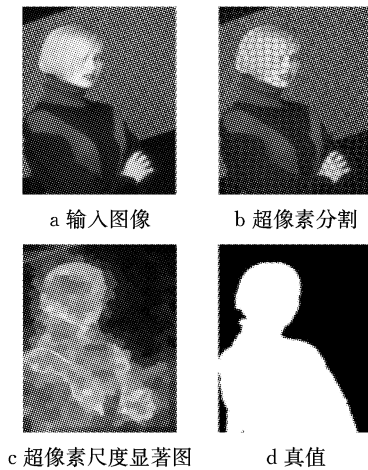


图 7 超像素尺度上的显著图

Fig. 7 Saliency maps on superpixel scales

2.4 图像块尺度

图像的语义信息对于对象的显著性检测来说意义重大. 由于像素和超像素的尺度太小, 导致像素和超像素基本不会携带有高层语义信息. 为了获取图像的语义信息用于显著性检测, 需要对图像进行更大尺度的划分. 首先利用图割算法对图像进行块(区

域)分割, 由于无法事先获取到显著对象的尺寸信息, 所以需要图像多次块分割. 图像 I 的块表示为 $\{r_k^{(n)}\}_{k=1}^{R(n)}$, 式中 n 为划分的尺度(本文中 $n=2$)、 $R(n)$ 为 n 尺度下分块的个数. 在尺度 n 下块 $r_i^{(n)}$ 的显著性计算公式定义如下:

$$S(r_i^{(n)}) = -\eta_i^{(n)} \lg\left(1 - \sum_{k=1}^{K(n)} \alpha_{ik}^{(n)} d(r_i^{(n)}, r_k^{(n)})\right) \quad (12)$$

式中: $d(r_i^{(n)}, r_k^{(n)})$ 为块 $r_i^{(n)}$ 和 $r_k^{(n)}$ 在 Lab 颜色空间上的卡方颜色距离; $\alpha_{ik}^{(n)}$ 为块 $r_i^{(n)}$ 面积占所有相邻分块面积总和的比例; 由于图像中越靠近图像空间中心的区域显著值越高, η_i 定义为分块 r_i 的空间分布权重, 参考文献[19], 本文定义 $\eta_i^{(n)} = \exp(-9(dx_i^n)^2 / \omega^2 - 9(dy_i^n)^2 / h^2)$, 融合多个块尺度显著图, 最终图像的显著计算公式定义为

$$S_r(p) = \frac{\sum_{n=1}^N \sum_{i=1}^{R(n)} S(r_i^{(n)}) (\|I_p - c_i^{(n)}\| + \epsilon)^{-1} \delta(p \in r_i^{(n)})}{\sum_{n=1}^N \sum_{i=1}^{R(n)} (\|I_p - c_i^{(n)}\| + \epsilon)^{-1} \delta(p \in r_i^{(n)})} \quad (13)$$

式中: N 为分割的尺度; $\|I_p - c_i^{(n)}\|$ 为分块中像素到分块中心的颜色距离; ϵ 为一个非常小的常数; $\delta(\cdot)$ 为指示函数. 通过计算, 图像块尺度上的显著图如图 8 所示.

2.5 多尺度融合

目前多尺度融合过程中的权重参数主要采用经验优先或者求均值的方式获取, 这两种方法不仅浪费资源而且对于得到的结果是否最优缺乏理论支撑. 本文采用 CRF 框架来融合多尺度显著性检测结果, 该方法不仅可以避免大量的常识性计算还能获得理论上更好的效果. 条件随机场是一种判别式概率模型, 其形式是根据模型的特征输入(观察值)以及随机变量之间的相互关系来预测输出的概率分布, 近年来条件随机场主要应用于序列标注以及状态预测, 并在机器学习与大数据分析方面表现突出.

2.5.1 CRF 融合

图像显著区域一般位于一个相对集中的区域, 因此可以通过像素周围区域的像素的显著性判断中心像素是否显著, 这样显著性检测方法更具鲁棒性. 但是这种方法过度依赖于其周围像素, 当像素位于显著区域的边缘位置时, 该方法便不再有效. 为了弥补这个不足, 本文采用基于 CRF 的多尺度融合方法, 利用像素之间的相互关系, 同时采用增加观测值的方法, 克服了上述方法的不足. 本文将不同尺度上的显著性检测结果作为条件随机场模型的特征输入, 相邻像素之间的特征值作为随机变量之间的相

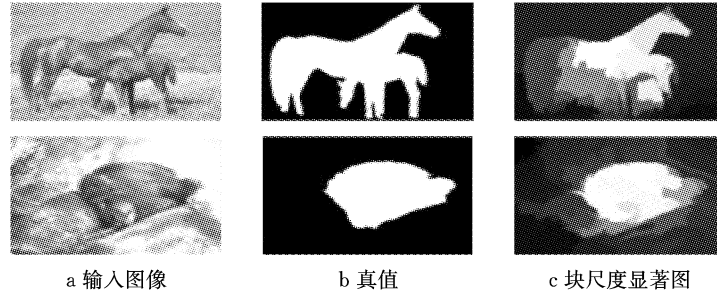


图 8 图像块尺度上的显著图

Fig.8 Saliency maps on region scales

互关系,通过训练学习,获得模型的参数组合,对图像的显著区域进行预测.根据条件随机场理论的一般模型,本文中的 CRF 模型为

$$P(A/I) = \frac{1}{z} \exp(-E(A/I)) \quad (14)$$

式中: I 为输入图像; z 为归一化因子.

图像 I 中的显著区域表示为一个二值模板 $A = \{a_x\}$, $a_x = 1$, 说明像素 x 属于显著区域; $a_x = 0$, 则说明像素 x 属于背景区域.

$E(A/I)$ 定义为

$$E(A/I) = \sum_x \sum_{k=1}^K \gamma_k F_k(a_x, I) + \sum_{x, x'} S(a_x, a_{x'}, I) \quad (15)$$

式中:等式右边第一项为输入特征(观察值);第二项为随机变量的相互关系; (x, x') 为邻接像素; γ_k 为第 k 个特征权值,其中

$$F_k(a_x, I) = \begin{cases} f_k(x, I) & a_x = 0 \\ 1 - f_k(x, I) & a_x = 1 \end{cases} \quad (16)$$

式中: $f_k(x, I)$ 为像素 x 是属于显著区域的概率(像素的显著值).

$$S(a_x, a_{x'}, I) = |a_x - a_{x'}| \exp(-\beta d(x, x')) \quad (17)$$

式中: $d(x, x')$ 为两个邻接像素在 Lab 颜色空间的距离; $\beta = (\langle \|I_x - I_{x'}\|^2 \rangle)^{-1}$ 为权值参数.

2.5.2 CRF 训练过程

本文 CRF 模型的特征输入(观察值)以及随机变量之间的相互关系,可以通过多尺度显著图来获得,未知变量仅有模型参数 γ . 本文通过对模型进行训练获得该参数组合.从 MSRA10K 中抽取一千张图片作为训练集,对 CRF 模型求解参数.

$$\gamma^* = \arg \max_{\gamma} \sum_n \lg P(A^n \setminus I^n; \gamma) \quad (18)$$

式中: (A, I) 为训练的图像样本.参考文献[20], 本文使用最大似然估计的方法对式(18)求解,获得最优的权值.

3 实验与分析

将本文提出的算法与 9 种经典算法在两个公共数据集上进行实验对比,通过实验结果来定性、定量检验本文模型的实际效果.

3.1 数据集与实验设置

实验在 MSRA10K 和 ECSSD 数据集上进行. MSRA10K 数据集包含有 10 000 张内容丰富且含有确定显著性对象的图像数据, ECSSD 数据集包含有 1 000 张含有确定显著性对象的数据集. ECSSD 数据集中的图片在区分前景与背景的困难程度明显高于 MSRA1000 数据集.

对比算法包括: CRF^[7]、AC^[21]、GB^[22]、FT^[23]、CA^[20]、PCA^[18]、CB^[13]、RC^[19]、DSR^[17]. 所有实验都是基于如下软、硬件基础进行:操作系统 WIN 7, 开发平台 Matlab 2015(b); Intel I 56500(3.2 GHz), 内存 8 G.

3.2 定性比较

图 9、10 分别为本文提出的方法与 9 种经典算法在数据集 ECSSD 和 MSRA 1000 上的实验结果.从图中可以看出,算法 AC 和 CA 不能提取显著对象全部轮廓信息,仅仅是对图像底层特征差异比较明显的区域敏感;算法 CRF 和 FT 可以比较完整地检测到显著区域,但是这种方法将大量的非显著区域错误地检测为显著区域,算法的错误率比较高.算法 PCA 和 GB 在实际的测试中将大量非显著区域错误地检测为显著区域.算法 RC、DSR、CB 在显著区域集中、明确且背景简单的图片上与本文提出的显著性检测方法可以达到类似的效果,但是对于背景复杂、前景背景差异比较小的图片,效果不及本文提出的方法.

3.3 定量分析

3.3.1 评价标准的选择

对于实验效果的评估需要一定的量化标准和评

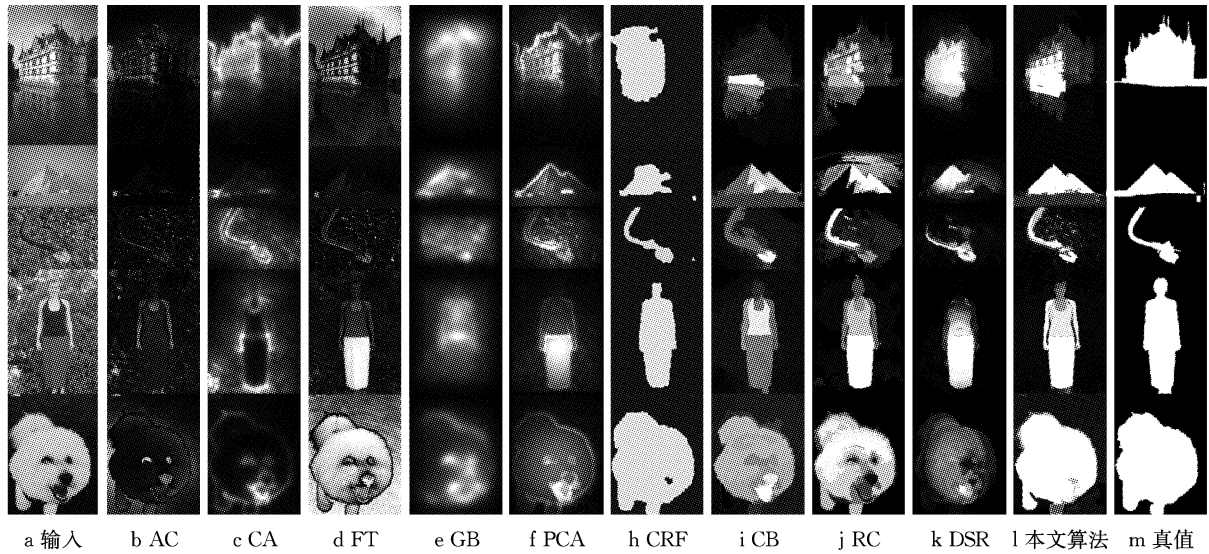


图 9 不同算法在 ECSSD 数据集上的定性对比结果

Fig.9 Qualitative comparison of different saliency maps on ECSSD dataset

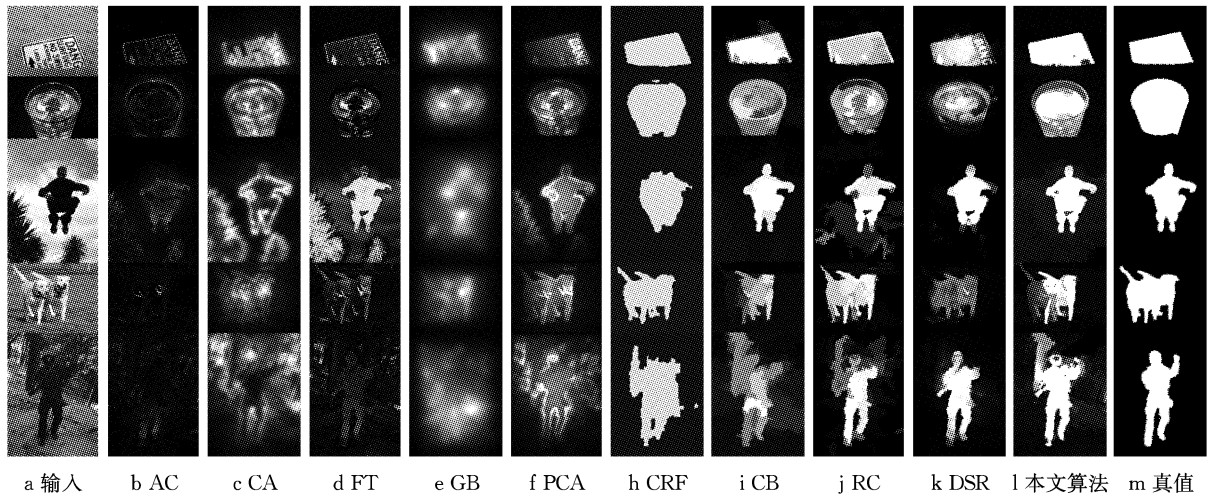


图 10 不同算法在 MSRA 1000 数据集上的定性对比结果

Fig.10 Qualitative comparison of different saliency maps on MSRA 1000 dataset

价方法,本部分采用 $P-R$ (precious-recall) 曲线、自适应阈值 F 以及平均误差 M 这 3 个评价指标作为度量标准,以固定阈值法和自适应阈值法为评价方法. 3 个度量标准的具体描述如下:

(1) $P-R$ 曲线. $P-R$ 曲线反应的是模型得到的显著值与真值之间的差别. $P-R$ 曲线越好,其反应的是模型越符合其模拟的系统. $P-R$ 曲线由两部分组成,分别是准确率 P (precious) 和召回率 R (recall),准确率反映的是模型产生显著区域与真值之间相同的部分占模型显著区域的百分比,召回率反映的是模型产生显著区域与真值之间相同的部分占真值的百分比. 具体的公式表述如下:

$$P = \frac{|B \cap G_T|}{|B|}, R = \frac{|B \cap G_T|}{|G_T|} \quad (19)$$

式中: B 为模型输出的显著区域; G_T 为图像的显著性真值.

(2) 自适应阈值 F . 当使用 $P-R$ 曲线来评价一些模型的效果时,独立地考量准确率或召回率会出现相互矛盾的情况,为了调和两者的矛盾需要对两者进行综合的考量. 定义自适应阈值 F 为

$$F = \frac{(1 + \beta^2)PR}{\beta^2 P + R} \quad (20)$$

参考文献[6]设定式(21)中参数 $\beta^2 = 0.3$.

(3) 平均误差 M . 在评估模型的实际效果时虽然 $P-R$ 曲线可以获得比 ROC (receiver operating characteristic) 曲线更好的表示,但是由于 $P-R$ 曲线本身所具有的局限性,仅仅使用 $P-R$ 曲线对模型有效性进行刻画还是比较片面的,因此引入平均误差

M 对模型的效果进行进一步的评估.

$$M = \frac{1}{wh} \sum_{i=1}^w \sum_{j=1}^h |S(i,j) - G_T(i,j)| \quad (21)$$

式中: w 、 h 分别为图像的宽度与长度; $S(i,j)$ 、 $G_T(i,j)$ 分别为坐标 (i,j) 处模型输出的显著值与真值.

3.3.2 评价方法

(1) 固定阈值法

所谓的固定阈值法就是设定一个固定的阈值来对实验结果进行二值分割, 超过阈值的部分定义为显著区域, 低于阈值的区域认为是背景区域. 通过得到的二值结果与 G_T 进行对比得到算法的召回率、准确率等量化数据. 在进行实验时, 阈值一般设 0~255 顺序变化, 就可以得到完整的 P-R 曲线, 以便评估在不同的阈值下算法的具体表现.

(2) 自适应阈值法

与固定阈值法不同, 自适应阈值法对于每张图像都有一个不同的阈值来决定显著图中哪些部分属于前景区域, 哪些部分属于背景区域. 阈值的设定遵循一定的惯例, 一般将阈值设置为平均显著值的两倍. 具体的公式描述见式(22).

$$T_a = \frac{2}{wh} \sum_{i=1}^w \sum_{j=1}^h |S(i,j)| \quad (22)$$

3.3.3 定量分析与比较

图 11、12 分别为本文提出的算法与 9 种经典算法在 ECSSD 和 MSRA 1000 数据集上定量比较的结果. 从图中可以看出, 本文的算法无论在精确度、召回率还是自适应阈值方面都明显优于 CRF、FT、GB、HC、PCA、AC、CA 和 CB 算法, 和 DSR 算法以及 RC 算法的表现处于同一水平. 由于本文算法在超像素尺度上的 M 较高导致算法最终的 M 要稍高于 DSR 算法, 但是在精确度以及召回率方面明显好于其他算法. 这说明本文算法不仅能精确地定位显著区域所在的位置, 同时还能相对比较完整的检测出显著区域.

3.4 单一尺度检测结果与采用不同融合方法检测结果对比

图 13 中展示了本文算法中不同成分以及采用不同优化方法融合后的定量分析结果. 由图 13 可以看出: 采用单一的显著性检测方法不能有效检测出完整的显著对象; 经过融合后的算法相对于每一个单一组成都更加有效; 采用 CRF 对算法不同组成部分进行有效性衡量是有必要的. 图中①为本文算法, ②为多尺度显著图线性相乘的结果, ③为多尺度显著图线性相加的结果, ④为像素尺度显著性检测结

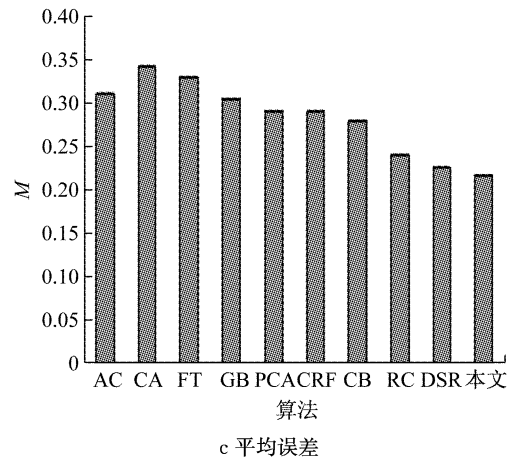
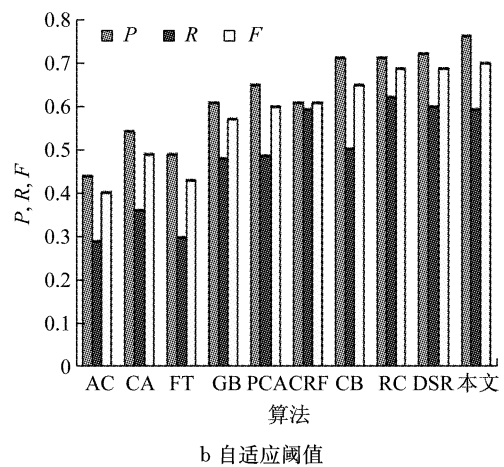
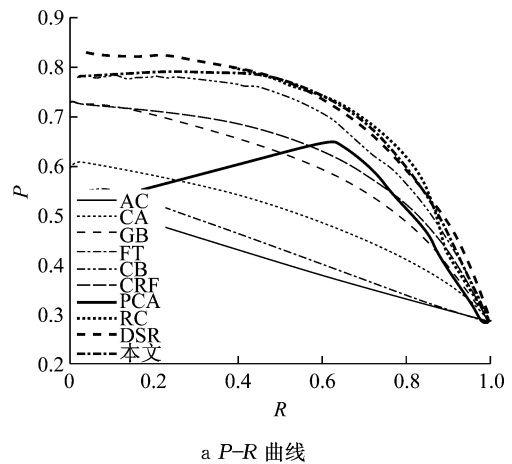
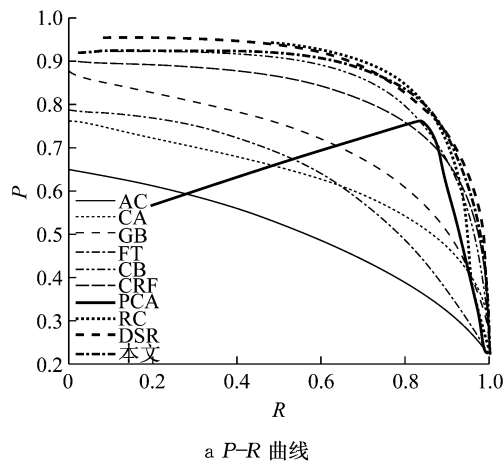


图 11 多个显著性检测算法在 ECSSD 数据集的量化比较
Fig. 11 Quantitative results of different methods on ECSSD dataset

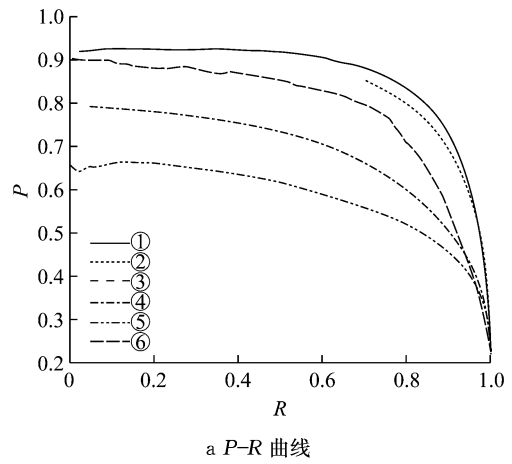
果, ⑤为超像素尺度显著性检测结果, ⑥为图像块尺度显著性检测结果.

3.5 算法运行效率

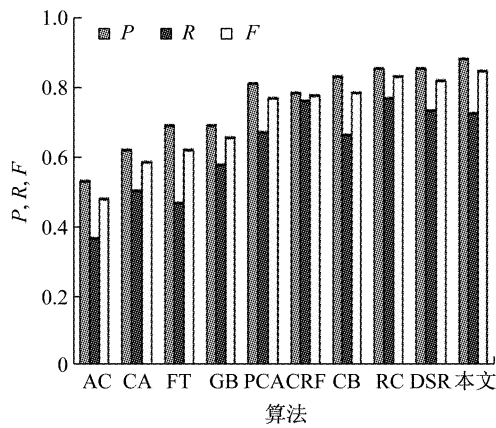
准确率与召回率并不是检验一个算法的唯一标准, 对于图像处理算法而言, 算法的运行效率也是一个重要的评价指标. 抽取 100 张图像对本文提出的算法与经典算法在同一平台上进行效率测试, 测试结果见表 1.



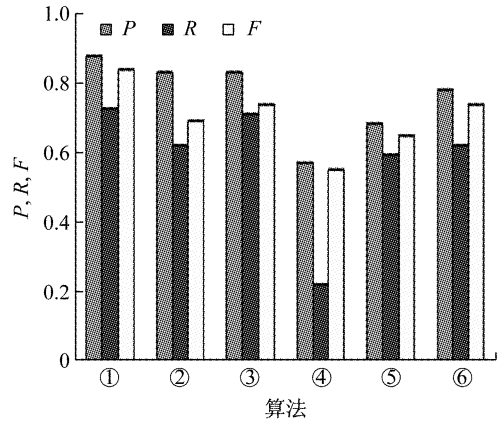
a P-R 曲线



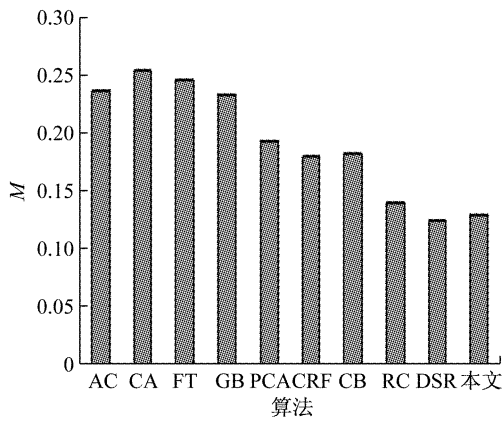
a P-R 曲线



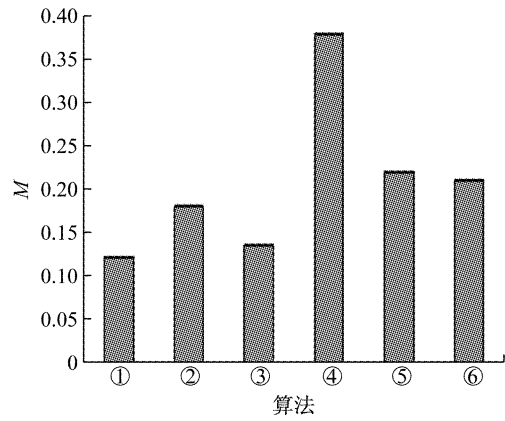
b 自适应阈值



b 自适应阈值



c 平均误差



c 平均误差

图 12 多个显著性检测算法在 MSRA1000 数据集的量化比较

Fig.12 Quantitative results of different methods on MSRA 1000 dataset

图 13 不同尺度显著性检测结果在 MSRA1000 数据集的量化比较

Fig.13 Quantitative results of different scales on MSRA 1000 dataset

表 1 算法效率对比

Tab.1 Comparison of execution time

算法	AC	CA	FT	GB	PCA	CRF	CB	RC	DSR	本文
运行时间/s	0.129	49.000	0.072	0.735	4.340	2.350	2.240	0.136	10.200	2.700

从表 1 中可以看出,相对于早期的显著性检测算法,本文的算法效率相对落后;与目前比较高效的 DSR 算法相比,本文的算法具有良好的表现. 本文算

法由于需要抽取多个层次,因此占用了大量的计算时间.

4 结论

本文提出的多尺度的显著性检测方法,通过结构提取方法对图像进行预处理;通过颜色空间分布以及颜色先验两种方法获得像素尺度的显著图;将图像进行超像素分割,利用超像素与显著支撑区域的特征差异求得超像素尺度上的显著图;使用图割算法对图像进行多尺度分割,利用局部对比方法求得图像块尺度上的显著图;利用 CRF 框架对多尺度显著图进行加权融合得到最终的显著图.与经典算法在两个公开数据集上进行性能的比较和分析表明,本文算法具有优越性.

参考文献:

- [1] WANG Q S, ZHENG W, PIRAMUTHU R. GraB: visual saliency via novel graph model and background priors [C]// CVPR 2016. Las Vegas: IEEE, 2016: 535-543.
- [2] GUO C, ZHANG L. A novel multiresolution spatio temporal saliency detection model and its applications in image and video compression[J]. IEEE Transactions on Image Processing, 2010, 19: 185-198.
- [3] ZHANG T, LIU S, AHUJA N, *et al.* Robust visual tracking via consistent low-rank sparse learning[J]. International Journal of Computer Vision, 2014, 111(2): 171.
- [4] LI A, SHE X C, SUN Q Z. Color image quality assessment combining saliency and FSIM[C]// Proceedings of the 5th International Conference on Digital Image Processing (ICDIP 2013). Beijing: Society of Photo-Optical Instrumentation Engineers, 2013: 8878: 1-5.
- [5] XIE Y, LU H C, YANG M H. Bayesian saliency via low and mid level cues[J]. IEEE Transactions on Image Processing 2013, 22: 1689.
- [6] ZHOU Q Q, ZHAO W D, ZHANG L, *et al.* Salient region detection by fusing foreground and background cues extracted from single image[J]. Mathematical Problems in Engineering, 2016, 7(2): 1.
- [7] LIU T, SUN J, ZHENG N N. Learning to detect a salient object [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 2: 353.
- [8] SHEN X H, WU Y. A unified approach to salient object detection via low rank matrix recovery [C]// CVPR2012. Providence: IEEE, 2012: 853-860.
- [9] ACHANATA R, SHAJI A, SMITH, SLIC superpixels compared to state-of-the-art superpixel methods[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34: 2274.
- [10] FELZENSWAL P F, HUTTENLOCHER D P. Efficient graph-based image segmentation [J]. International Journal of Computer Vision, 2004, 59(2): 167.
- [11] ZHANG L, GU Z Y, LI H Y. SDSF: a novel saliency detection method by combining simply priors[C]// IEEE International Conference on Image Processing. Melbourne: IEEE, 2013: 171-175.
- [12] LIU Z, ZOU W B, MEUR O L. Saliency tree: a novel saliency detection framework [J]. IEEE Transactions on Image Processing, 2014, 23: 1937.
- [13] JIANG H Z, WANG J D, YUAN Z J, *et al.* Automatic salient object segmentation based on context and shape prior[C]// Proceedings of the British Machine Vision Conference (BMVC' 11). Dundee: BMVA Press, 2011: 110. 1-110. 12.
- [14] SHI J P, YAN Q, XU L, *et al.* Hierarchical image saliency detection on extended CSSD[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38: 717.
- [15] LI X, LU H C, ZHANG L, *et al.* Saliency detection via dense and sparse reconstruction [C]// Proceedings of The IEEE International Conference on Computer Vision and Pattern Recognition. Portland: IEEE, 2013: 2976-2983.
- [16] YANG C, ZHANG L H, LU H C, *et al.* Saliency detection via graph-based manifold ranking[C]// Proceedings of The IEEE International Conference on Computer Vision and Pattern Recognition. Portland: IEEE, 2013: 3166-3173.
- [17] YUAN Y C, LI C Y, KIM J M, *et al.* Reversion correction and regularized random walk ranking for saliency detection[J]. IEEE Transactions on Image Processing, 2018, 27: 1311.
- [18] MARGOLIN R, ZELNIK-MANOR L, TAL A. What makes a patch distinct? [C]// Proceedings of The IEEE International Conference on Computer Vision and Pattern Recognition. Portland: IEEE, 2013: 1139-1146.
- [19] CHENG M M, MITRA N J, HUANG X, *et al.* Global contrast based salient region detection[J]. IEEE Trans Pattern Anal Mach Intell, 2015, 37(3): 569.
- [20] GOFERMAN S, ZELNIK-MANOR L, TAL A. Context-aware saliency detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34: 1915.
- [21] ACHANATA R, ESTRADA F, WILSON P, *et al.* Salient region detection and segmentation[C]// Proceedings of International Conference on Computer Vision Systems. Santorini: Springer, 2008: 66-75.
- [22] HAREL J, KOCH C, PERONA P. Graph-based visual saliency [C]// Proceedings of the Advances in Neural Information Processing Systems. Vancouver: NIPS, 2006: 545-552.
- [23] ACHANATA R, SHEILA H S, FRANCISCO E. Frequency-tuned salient region detection [C]// Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition. Miami: IEEE, 2009: 1597-1604.