

# 基于多图流形排序的图像显著性检测

于明<sup>1</sup> 李博昭<sup>1</sup> 于洋<sup>1</sup> 刘依<sup>1</sup>

**摘要** 针对现有图像显著性检测算法中显著目标检测不完整和显著目标内部不均匀的问题, 本文提出了一种基于多图流形排序的图像显著性检测算法. 该算法以超像素为节点构造 KNN 图 (K nearest neighbor graph) 模型和 K 正则图 (K regular graph) 模型, 分别在两种图模型上利用流形排序算法计算超像素节点的显著性值, 并将每个图模型中超像素节点的显著值加权融合得到最终的显著图. 在公开的 MSRA-10K、SED2 和 ECSSD 三个数据集上, 将本文提出的算法与当前流行的 14 种算法进行对比, 实验结果显示本文算法能够完整地检测出显著目标, 并且显著目标内部均匀光滑.

**关键词** 图像显著性检测, 多图模型, 流形排序, 超像素节点

**引用格式** 于明, 李博昭, 于洋, 刘依. 基于多图流形排序的图像显著性检测. 自动化学报, 2019, 45(3): 577–592

**DOI** 10.16383/j.aas.2018.c170441

## Image Saliency Detection With Multi-graph Model and Manifold Ranking

YU Ming<sup>1</sup> LI Bo-Zhao<sup>1</sup> YU Yang<sup>1</sup> LIU Yi<sup>1</sup>

**Abstract** To resolve the incompleteness and non-uniformity of salient object detection, this paper proposes an image salient object detection algorithm with multi-graph model and manifold ranking. The algorithm uses superpixels as nodes to construct KNN graph model and K regular graph model. For each model, manifold ranking algorithm is used to calculate saliency values of superpixel nodes. The saliency values of the nodes obtained from the two graph models are fused together with different weights to form the image saliency map. On three public available databases, MSRA-10K, SED2 and ECSSD, the proposed algorithm is compared with fourteen state-of-art algorithms. Experimental results show that the proposed algorithm can detect a salient object completely, yet the object is uniform and smooth inside.

**Key words** Image saliency detection, multi-graph model, manifold ranking, superpixel node

**Citation** Yu Ming, Li Bo-Zhao, Yu Yang, Liu Yi. Image saliency detection with multi-graph model and manifold ranking. *Acta Automatica Sinica*, 2019, 45(3): 577–592

图像显著性检测是计算机视觉领域的关键技术之一, 有着广泛的应用领域, 包括目标识别、图像分类<sup>[1]</sup>、图像检索<sup>[2]</sup>、基于内容感知的图像压缩等<sup>[3]</sup>. 人类视觉注意机制可以从自然图像中提取出感兴趣的区域供给大脑进行后续分析, 如何使计算机能够更好地模拟人类的视觉注意机制, 准确、快速地提取图像中的感兴趣区域, 是图像显著性检测研究的主要内容.

图像显著性检测方法从驱动方式上可以分为两大类: 自底向上的基于数据驱动的方法和自顶向下的基于任务驱动的方法<sup>[4]</sup>. 自底向上的方法主要利用图像的底层特征, 如对比度、颜色、亮度、纹理和

边缘等, 提取图像中的显著目标. 自顶向下方法是根据具体任务对自底向上的检测结果进行调整或者根据任务样本自动建立显著性模型, 这类方法需要大量的带有类别标记的示例数据集进行训练, 且需人为控制, 并依赖于所需的任务. 由于自底向上的显著性检测方法多从图像特征入手, 不依赖于特定的任务, 因此目前大多数研究都是基于自底向上的方式进行视觉显著性检测.

对比度是引起人类视觉注意的最大因素, Goferman 将自底向上的基于对比度的图像显著性检测算法分成以下三类<sup>[5]</sup>: 考虑局部特征, 考虑整体特征, 以及融合整体特征和局部特征的方法. 基于局部对比度的算法通常通过计算图像的子区域或者像素与其周围一个小的局部邻域的对比如来衡量该图像子区域或者像素的显著性, 代表算法有 Achanta 等<sup>[6]</sup>于 2008 年提出的 AC (Salient region detection and segmentation) 算法和 Yan 等<sup>[7]</sup>于 2013 年提出的 HS (Hierarchical saliency) 算法. AC 算法通

收稿日期 2017-08-02 录用日期 2018-02-07  
Manuscript received August 2, 2017; accepted February 7, 2018  
天津市科技计划 (14RCGFGX00846, 15ZCZDNC00130, 17ZLZDZF00040), 河北省自然科学基金 (F2015202239) 资助

Supported by Tianjin Science and Technology Planning Project (14RCGFGX00846, 15ZCZDNC00130, 17ZLZDZF00040) and Hebei Province Natural Science Foundation (F2015202239)

本文责任编辑 杨健

Recommended by Associate Editor YANG Jian

1. 河北工业大学计算机科学与软件学院 天津 300401

1. School of Computer Science and Engineering, Hebei University of Technology, Tianjin 300401

过计算像素块在不同邻域上的局部对比度实现多尺度显著性计算. HS 算法对图像进行多级分割, 在每个层级上利用局部对比度计算显著图, 并引入等级机制对每个层级上的显著图进行融合得到最终的显著图. 基于全局对比度的算法特点是对于每个子区域或像素, 将整个图像作为对比区域计算其显著值, 代表算法有 Cheng 等<sup>[8]</sup>于 2015 年提出的 HC (Histogram based contrast) 算法和 RC (Region based contrast) 算法. HC 算法将一个像素的显著性定义为该像素与其他像素的颜色对比度之和, 并使用直方图进行高效处理. 针对 HC 算法仅考虑了像素的颜色特征, RC 算法在此基础上加入了位置信息. 融合局部和全局对比度的方法在计算图像显著性时分别考虑局部对比度和全局对比度, 并将二者通过一定的方式结合起来. Borji 等<sup>[9]</sup>提出的 LGPR (Local and global patch rarities) 算法利用 LAB 空间的三个通道分别计算图像块之间的欧氏距离和图像块直方图的概率分布, 作为局部和全局对比度, 并进行融合得到最后的显著图.

通常, 在空域中计算像素 (或像素块) 的显著性时间复杂度高, 而在频域中计算显著性可以提高算法的速度. 如 Achanta 等于 2009 年提出的 FT (Frequency-tuned) 算法<sup>[10]</sup>和 2010 年提出的 MSS (Maximum symmetric surround) 算法<sup>[11]</sup>, 还有 Hou 等<sup>[12]</sup>提出的 SR (Spectral residual) 算法都是基于频域的显著性检测方法.

Wang 等<sup>[13]</sup>于 2015 年提出了 BFSS (Background and foreground seed selection) 算法, 通过对原始图像进行超像素分割并选择背景超像素生成背景模板, 计算每个超像素与背景模板中超像素块的对数之和, 生成初始显著图; 在二值化的初始显著图中选择前景超像素生成前景模板, 计算每个超像素与前景模板中超像素块的对数之和, 生成最终的显著图. 该算法中生成背景模板的机制耗时多, 算法的时间复杂度较高. 除了上述基于对比度的算法, 基于阈值分割的方法也被用于进行图像显著性检测. Kim 等<sup>[14]</sup>于 2014 年提出了 HDCT (High-dimensional color transform) 算法, 首先利用基于 Otsu 的多级自适应阈值算法, 计算两个阈值, 对现有算法生成的初始显著图进行阈值分割, 得到三值图像, 提取三值图像中非中间值区域像素的颜色特征, 这些像素将被用作显著区域和背景区域的初始颜色样本, 然后将低维颜色空间映射到高维颜色空间并以颜色样本为约束, 根据颜色样本的约束估计颜色通道的最优线性组合, 最后结合高维颜色空间中的颜色值获得显著性图. Zhang 等<sup>[15]</sup>于 2016 年提出的 BMA (Boolean map approach) 算法将输入图像利用多个阈值进行分割得到多幅二值

图像, 通过二值图像间的拓扑分析计算显著图. 基于阈值分割的显著性检测方法对阈值的依赖性强, 阈值选择不好往往会使用算法失效.

基于图模型的显著性检测算法利用图的不同连接方式直观地刻画数据节点之间多样的结构特征, 通过构造不同的关联矩阵 (Affinity matrix) 反映图像区域之间的相似程度, 由于其检测准确率高且算法复杂度低而引起更多的关注. 在 IT 模型<sup>[16]</sup>的基础上, Harel 等<sup>[17]</sup>于 2006 年提出 GBVS (Graph based visual saliency) 算法, 利用马尔科夫随机场构建二维图像的马尔科夫链, 通过求其稳定状态得到显著图. Wei 等<sup>[18]</sup>于 2012 年提出了 GS (Geodesic saliency) 算法, 以超像素为节点构造图模型, 并将节点的显著性值定义为该节点与所有背景节点之间的测地距离之和, 该方法原理简单, 能够较完整地检测出显著目标. Li 等<sup>[19]</sup>于 2013 年提出了 CHM (Contextual hypergraph modeling) 算法将超图模型引入到图像显著性检测中, 通过对原始图像进行多级分割, 在每一级上建立超图模型并结合边缘特征计算显著图. Li 等<sup>[20]</sup>于 2015 年提出 RW (Random walk) 算法, 该算法融合显式视觉线索和潜在视觉相关性, 同时考虑图像区域本身的信息性及不同图像区域之间潜在的相关性, 构建有向图模型, 在该图模型上利用随机游走计算各区域的显著性值. Yuan 等<sup>[21]</sup>于 2018 年提出的 RR (Regularized random) 算法中引入了正则化的随机游走排序模型 RRWR (Regularized random walk ranking) 和反转修正模型 RC (Reversion correction), 该算法同时利用区域和像素特征, 将先验显著性估计引入图像中的每个像素, 从而生成基于像素细节特征和超像素区域特征的显著图.

基于图模型的显著性检测算法普遍需要解决以下问题: 如何构建图模型的节点集、边集以及边的权值选择; 如何利用图模型计算节点的显著性值. Yang 等<sup>[22]</sup>提出的 MR (Manifold ranking) 算法利用图模型刻画图像区域之间内在的关系, 并把流形排序的概念引入到显著性检测中, 将图像的显著性检测视为对超像素节点的二分类问题. 通过基于图模型的流形排序算法将局部分组融入到节点的分类问题中, 与直接利用图模型上节点之间的测地距离计算显著图的 GS 算法相比, MR 算法得到的显著目标内部更加均匀光滑<sup>[23]</sup>. 然而 MR 算法在构造图模型时仅将位置相邻的节点进行连接, 没有考虑颜色相似的节点之间的连接性, 这样检测到的显著目标存在不完整的问题, 有时也会错误地突出非显著区域.

针对以上显著性检测算法中存在的问题, 例如显著目标不完整, 目标内部不能一致均匀突出, 本

文提出了一个基于多图流形排序的显著性检测算法. 与仅利用单一图模型的显著性检测算法不同, 本文引入多个图模型, 分别以超像素为节点构造基于超像素颜色特征的 KNN 图模型和基于空间位置特征的 K 正则图模型. 在 KNN 图模型中将颜色相近的节点进行连接, 在 K 正则图模型中将位置近邻的节点进行连接. 本文算法的主要贡献是: 1) 在两种图模型中分别利用流形排序算法计算节点的显著性值, 并且在利用流形排序算法计算节点的显著性值时使用全局前景假设, 一定程度上解决了显著目标位于边界时 MR 算法失效的问题; 2) 将两种图模型下得到的节点显著性值在节点级上进行加权融合, 得到节点最终的显著性值, 在一定程度上弥补了单一图模型下生成的显著图中噪声较多和显著目标不完整的问题.

其余章节组织如下: 在第 1 节中介绍提出的基于多图流形排序的图像显著检测算法, 在第 2 节中将本文算法与经典的显著性检测算法进行对比实验, 第 3 节中给出总结和展望.

## 1 基于多图流形排序的图像显著性检测算法

本文提出的算法是在超像素基础上进行的, 超像素是利用像素之间特征的相似性将颜色相近的像素分组, 这样就可以用少量的超像素代替大量的像素表示图像, 从而为后续的图像处理降低复杂度. 本文利用 SLIC 算法<sup>[24]</sup> 对图像进行超像素分割, 对得到的超像素块提取颜色特征和空间位置特征, 以超像素为节点分别构造基于颜色特征和空间位置特征的两类图模型  $G_1$  和  $G_2$ . 在这两类图模型上分别利用流形排序算法计算所有节点的显著性值  $\mathbf{f}_1 = \{f_{1,1}, f_{1,2}, \dots, f_{1,m}\}$  和  $\mathbf{f}_2 = \{f_{2,1}, f_{2,2}, \dots, f_{2,n}\}$ , 其中  $f_{1,i}$  表示超像素节点  $i$  在  $G_1$  模型下计算得到的显著性值,  $f_{2,i}$  表示节点  $i$  在  $G_2$  模型下计算得到的显著性值. 将两个图模型上得到的节点的显著性值进行加权融合, 利用融合后节点的显著性值生成初始显著图, 并根据中心先验和边界抑制对初始显著图进行优化得到最终的显著图. 在此基础上将两种图模型下生成的显著图进行融合, 能够使得检测到的显著目标更加完整均匀, 并且能够在一定程度上抑制单一图模型下由于节点连接错误, 在显著图中产生的噪声. 本文算法的流程伪码如算法 1 所示:

**算法 1. 基于多图流形排序的图像显著性检测算法**

输入: 图像  $I_n$ ;

- 1) 利用 SLIC 算法对图像  $I_n$  进行超像素分割;
- 2) 以超像素为节点构造图模型  $G_1$ , 计算图模型  $G_1$  的关联矩阵  $W_1$  和度矩阵  $D_1$ ;

- 3) 计算图模型  $G_1$  下节点的显著性值  $\mathbf{f}_1 = (D_1 - \alpha W_1)^{-1} \mathbf{y}$ ;

- 4) 以超像素为节点构造图模型  $G_2$ , 计算图模型  $G_2$  的关联矩阵  $W_2$  和度矩阵  $D_2$ ;

- 5) 计算图模型  $G_2$  下节点的显著性值  $\mathbf{f}_2 = (D_2 - \alpha W_2)^{-1} \mathbf{y}$ ;

- 6) 加权融合节点显著性值  $\mathbf{f}_1$  和  $\mathbf{f}_2$  得到显著性值  $\mathbf{f}$ , 生成初始显著图  $S_{co}$ ;

- 7) 对初始显著图进行优化, 得到最终的显著图  $S$ .

输出: 显著图  $S$ .

### 1.1 基于图模型流形排序的显著性计算

本文涉及的流形排序问题又称为半监督回归问题<sup>[25]</sup>, 是基于半监督学习分类领域的一个分支. 流形排序可以简单描述为: 给定正样本 (或负样本), 通过排序值的大小, 判定某个样本与正样本 (或负样本) 的相似性, 依此对该样本进行分类. 基于图模型的流形排序算法引进图模型刻画数据之间的关系, 使得数据的表达更加直观.

基于图模型流形排序的图像显著性计算以超像素为节点构造图模型, 在图模型的基础上通过求目标函数的最小值计算每个节点相对于查询节点的排序值, 并将这个排序值定义为节点的显著性值. 这里定义查询节点为已标记好的超像素节点, 例如将某个超像素节点标记为前景 (或背景) 节点.

设图像中超像素个数为  $n$ , 构造数据集  $X = \{x_1, x_2, \dots, x_n\}$ ,  $X \in \mathbf{R}^{m \times n}$ , 其中  $m$  表示超像素节点特征的维数,  $n$  表示超像素节点个数, 其中一些节点被标记为查询节点. 定义指示向量  $\mathbf{y} = [y_1, y_2, \dots, y_n]^T \in \mathbf{R}^n$ , 若  $x_i$  为查询节点, 则  $y_i = 1$ , 否则  $y_i = 0$ . 定义排序函数  $F: X \rightarrow \mathbf{f} = [f_1, \dots, f_n]^T \in \mathbf{R}^n$ , 排序函数  $F$  将  $m \times n$  的数据集矩阵  $X$  映射到  $n$  维向量  $\mathbf{f}$ , 其中  $f_i$  表示节点  $i$  相对于查询节点的排序值, 即节点的显著值. 在数据集  $X$  上构建一个图模型  $G = (X, E)$ , 其中数据集  $X$  是节点集, 边集  $E$  是连接节点的边构成的集合, 边的权值由关联矩阵  $W = [w_{ij}]_{n \times n}$  表示. 定义  $D = \text{diag} \{d_{11}, d_{22}, \dots, d_{nn}\}$  为图模型  $G$  对应的度矩阵 (Degree matrix), 其中  $d_{ii} = \sum_j w_{ij}$ ,  $j = 1, \dots, n$ . 根据流形排序的思想, 排序值  $\mathbf{f}$  的最优解  $\mathbf{f}^*$  通过以下的最优化问题得到<sup>[22]</sup>:

$$\mathbf{f}^* = \arg \min_{\mathbf{f}} \frac{1}{2} \left( \sum_{i,j=1}^n w_{ij} \left\| \frac{f_i}{\sqrt{d_{ii}}} - \frac{f_j}{\sqrt{d_{jj}}} \right\|^2 \right) +$$

$$\mu \sum_{i=1}^n \|f_i - y_i\|^2 \quad (1)$$

其中参数  $\mu$  用来平衡平滑项 (第一项) 和拟合项 (第二项). 根据式 (1) 求得  $\mathbf{f}$  的最优解为:

$$\mathbf{f}^* = (I - \alpha S)^{-1} \mathbf{y} \quad (2)$$

其中  $I$  为单位矩阵, 参数  $\alpha = 1/(1 + \mu)$ ,  $S$  是归一化的拉普拉斯矩阵  $S = D^{-1/2} W D^{-1/2}$ . 利用非归一化的拉普拉斯矩阵, 得到  $\mathbf{f}^*$  的另一种形式:

$$\mathbf{f}^* = (D - \alpha W)^{-1} \mathbf{y} \quad (3)$$

文献 [22] 中通过实验验证, 式 (3) 比式 (2) 得到的显著图更准确. 本文利用式 (3) 计算每个节点的排序值即显著性值  $f_i$ . 令  $A = (D - \alpha W)^{-1}$ , 则  $\mathbf{f}^* = A\mathbf{y}$ , 当计算  $\mathbf{f}^*$  时, 需要将  $A$  的对角线元素置为 0, 否则得到的  $\mathbf{f}^*$  中将包含节点到自身的排序值, 这是一个无意义的值并且会影响最终的排序结果.

基于图模型流形排序的显著性检测算法中, 查询节点的选择是一个关键的问题. MR 算法利用背景先验, 假设图像边界为背景区域, 将图像边界节点设置为查询节点, 利用式 (3) 计算每个节点  $i$  相对于背景节点的排序值  $f_i^*$ , 则节点的显著性值为  $1 - f_i^*$ . 目前基于背景先验的显著性检测算法大多利用边界假设, 即假设图像边界为背景区域, 生成背景模板, 并根据背景模板计算图像的显著图, 这些算法大都会遇到背景鉴别这个具有挑战性的问题. 其中一些算法通过不同的方法优化背景模板, RBD (Robust background detection) 算法<sup>[26]</sup> 通过定义节点的边界连接性判定边界节点属于背景的可能性大小, 边界连接性越大的节点, 其属于背景的可能性越大. BFSS 算法<sup>[13]</sup> 则是通过定义节点的边界概率选择背景节点, 节点的边界概率越大, 越有可能属于背景. UILSG (Updating initial labels from spectral

graph) 算法<sup>[27]</sup> 通过一个边界节点分类机制筛选真正的背景节点. 这些算法在一定程度上解决了背景鉴别的问题, 但是仍然是基于边界假设的思想, 当显著目标位于图像边界时, 算法会失效. 为了解决这个问题, 本文采用全局前景假设的思想, 将所有的节点作为查询节点, 即在利用式 (3) 计算节点的显著性值时, 假设所有的节点为前景节点, 设置节点指示向量  $\mathbf{y} = [1, \dots, 1]^T$ , 将  $\mathbf{y}$  代入式 (3) 计算节点的显著性值. 将所有的节点设置为查询节点, 避免了背景鉴别的问题, 检测到的显著目标更加完整. 图 1 显示了 MR 算法和本文算法分别利用边界假设和全局前景假设生成的显著图, 其中图 1 (a) 为原图, 图 1 (b) 是人工标注图 Ground-truth, 简称 GT, 图 1 (c) 为 MR 算法中利用边界假设生成的显著图, 图 1 (d) 为 MR 算法中利用全局前景假设生成的显著图, 图 1 (e) 为本文算法中利用边界假设生成的显著图, 图 1 (f) 为本文算法中利用全局前景假设生成的显著图. 从图中可以看出利用全局前景假设得到的显著目标更加完整.

## 1.2 多图模型构造

在基于图模型的显著性检测算法中, 图模型的构造对显著图的计算至关重要. 现有的基于图模型的显著性检测算法大多利用单一的图模型描述图像中不同区域之间的关系, 例如 GS 和 MR 算法都是以超像素为节点, 将空间相邻的节点进行连接构造图模型. 然而在自然图像中, 不同区域之间往往具有复杂的关系, 仅使用单一的图模型并不能够准确地描述图像区域之间的关系. Li 等<sup>[28]</sup> 提出的 MGL (Multi-graph learning) 算法利用优化的多图模型学习进行显著目标检测, 该算法融合多个图模型刻画数据节点之间的关系, 但是这些图模型中节点之间的连接关系并没有变化, 只是边的权值

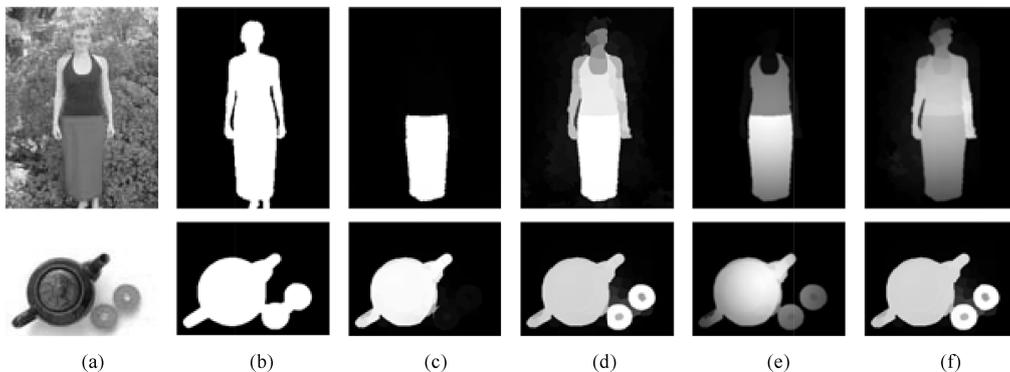


图 1 全局前景假设和边界先验的比较

Fig. 1 Comparison of global foreground prior and background prior

不同, 仅改变图模型中边的权值并不能反映数据本身的特性, 在不同的特征下, 节点之间的连接性也是不同的. Zhang 等<sup>[29]</sup> 提出的 GBOF (Graph-based optimization framework) 算法通过构造两个图模型并计算不同图模型下的权值矩阵, 从不同的角度刻画节点之间的连接性. GBOF 的图模型分别是利用阈值法查找邻接点的颜色特征图模型和基于 K 正则图模型的位置特征图模型, 颜色特征图模型利用设定的阈值寻找每个节点的邻接节点, 当阈值设置不合适时, 会产生孤立节点, 影响显著目标的检测结果. 位置图模型与 MR 算法一致, 两类图模型融合时利用图模型的权值矩阵, 通过最小化能量函数得到节点相对于查询节点的排序值, 进而生成显著图.

在基于图模型的显著性检测算法中, 图模型需要能够体现数据节点之间的连接性. 常用的构图方法有 K 正则图、KNN 图、全连接图和  $\varepsilon$ -图. K 正则图构造方式简单, 且具有对称性, 可以用于描述超像素节点的空间位置特性. KNN 图与  $\varepsilon$ -图都可以用于刻画节点之间基于颜色特征的连接性, KNN 图模型是根据节点之间的颜色差异寻找每个节点的  $K$  个邻接节点,  $\varepsilon$ -图是根据阈值寻找每个节点的近邻节点, 但是如果阈值设置不准确,  $\varepsilon$ -图构造的图模型中会产生孤立点, 影响算法的性能, 所以本文使用 KNN 图模型刻画节点之间基于颜色特性的连接性. 全连接图指的是将图模型中所有的节点两两相连, 不能准确地体现不同节点之间的连接关系.

### 1.2.1 KNN 图模型

颜色是自然图像中的重要信息, 颜色相似的区域很可能属于同一个物体, 本文基于颜色特征构造 KNN 图模型作为图模型  $G_1 = (V_1, E_1)$ . 在 KNN 图中每个节点只与它在某个距离测度下的  $K$  个近邻节点相连. KNN 图模型能够自适应地反映节点在特征空间中的分布特性, 在高密度区域, KNN 图的半径变小; 在低密度区域, KNN 图的半径变大<sup>[30]</sup>, 这种能够自动适应节点分布特性的优点, 使得 KNN 图能够很好地拟合节点在颜色特征下的分布特性, 因而可以刻画节点在颜色特征下的连接性.

本文提取 CIELAB 颜色空间中超像素块所包含的像素点的颜色特征均值作为节点的颜色特征  $\mathbf{c} = (l, a, b)^T$ , 提取超像素块所包含的像素点的空间位置特征均值作为节点的空间位置特征  $\mathbf{z} = (x, y)^T$ . 定义  $dist(\mathbf{c}_i, \mathbf{c}_j)$  为超像素  $i$  和  $j$  之间的颜色距离, 其中  $\mathbf{c}_i$  和  $\mathbf{c}_j$  表示超像素  $i$  和  $j$  的颜色特征; 定义  $dist(\mathbf{z}_i, \mathbf{z}_j)$  为超像素  $i$  和  $j$  之间的空间位置距离, 其中  $\mathbf{z}_i$  和  $\mathbf{z}_j$  表示超像素  $i$  和  $j$  的空间位置特征. KNN 图模型中, 如果仅使用颜色距离  $dist(\mathbf{c}_i, \mathbf{c}_j)$  寻找每个节点的近邻节点, 而不考虑相邻节点之间

的连接性, 这样构造出的  $G_1$  图模型生成的显著图中显著目标内部不均匀; 如果仅使用空间位置距离  $dist(\mathbf{z}_i, \mathbf{z}_j)$  寻找每个节点的近邻节点, 构造出的  $G_1$  图模型生成的显著图中背景与显著目标边界不明确, 并且会错误地突出背景区域, 所以本文在寻找近邻节点时利用式 (4) 的距离度量方式:

$$dist(i, j) = \lambda_z dist(\mathbf{z}_i, \mathbf{z}_j) + \lambda_c dist(\mathbf{c}_i, \mathbf{c}_j) \quad (4)$$

式 (4) 中利用欧氏距离度量两节点之间的空间位置距离  $dist(\mathbf{z}_i, \mathbf{z}_j)$  和颜色距离  $dist(\mathbf{c}_i, \mathbf{c}_j)$ , 其中  $\lambda_z$  和  $\lambda_c$  表示权重, 用来调节空间位置距离  $dist(\mathbf{z}_i, \mathbf{z}_j)$  和颜色距离  $dist(\mathbf{c}_i, \mathbf{c}_j)$  对总距离  $dist(i, j)$  的影响,  $dist(\mathbf{z}_i, \mathbf{z}_j)$  和  $dist(\mathbf{c}_i, \mathbf{c}_j)$  的定义如下:

$$\begin{aligned} dist(\mathbf{z}_i, \mathbf{z}_j) &= \|\mathbf{z}_i - \mathbf{z}_j\|_2 \\ dist(\mathbf{c}_i, \mathbf{c}_j) &= \|\mathbf{c}_i - \mathbf{c}_j\|_2 \end{aligned} \quad (5)$$

计算超像素  $i$  与其他超像素之间的距离  $dist(i, j)$ ,  $j = 1, 2, \dots, n, i \neq j$ , 根据  $dist$  值寻找超像素  $i$  的  $K$  个近邻超像素. 利用式 (4) 搜索得到近邻节点, 既考虑了节点之间颜色特征的相似性又考虑了节点之间空间位置的连接性, 不仅能够限制颜色相近但位置相距较远的节点之间的连接性, 又能够提高颜色差异较大但是位置相距较近的节点之间的连接性.

这里考虑节点位于图像的內部和位于图像边界区域两种情况, 当节点位于图像内部时, 在节点  $i$  与它的  $K$  个近邻节点  $j$  之间构造一条有向边  $e_{ij} \in E_{1, \text{Inner}}$ ,  $E_{1, \text{Inner}}$  为有向边集合, 有向边的权值定义为:

$$W_{1, \text{Inner}}(e_{ij}) = \exp\left(\frac{-dist(\mathbf{c}_i - \mathbf{c}_j)}{2\sigma^2}\right), \quad (6)$$

$$e_{ij} \in E_{1, \text{Inner}}$$

其中  $\sigma$  是一个常数, 用来控制权值的强度.

由于图像的边界区域大多属于背景区域, 在 KNN 图模型的基础上, 本文将边界节点两两连接. 边界节点指的是当超像素中包含图像边界像素时, 则该超像素节点被定义为边界节点. 将图像的任意两个边界节点  $i$  和  $j$  之间构造两条有向边, 分别为从节点  $i$  指向节点  $j$  的有向边  $e_{ij}$  和从节点  $j$  指向节点  $i$  的有向边  $e_{ji}$ , 其中  $e_{ij}, e_{ji} \in E_{1, \text{Boundary}}$ ,  $E_{1, \text{Boundary}}$  为有向边集合, 集合  $E_{1, \text{Boundary}}$  中的有向边上的权值定义为:

$$W_{1, \text{Boundary}}(e_{ij}) = \exp\left(\frac{-dist(\mathbf{z}_i - \mathbf{z}_j)}{2\sigma^2}\right), \quad (7)$$

$$e_{ij} \in E_{1, \text{Boundary}}$$

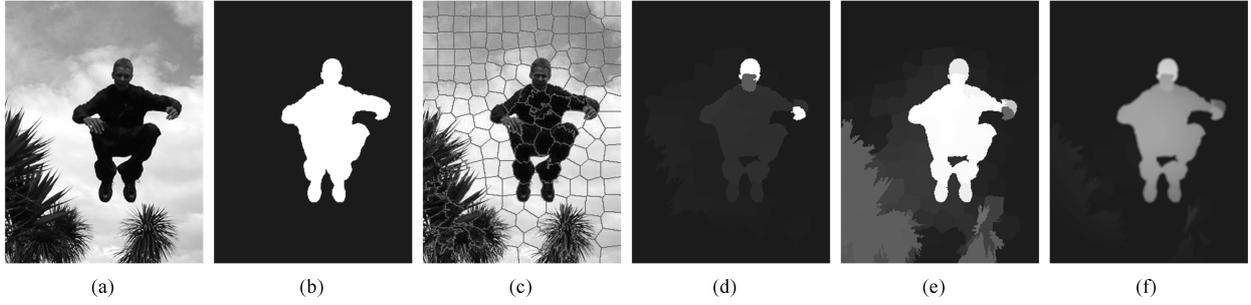


图 2 不同图模型生成的显著图比较

Fig. 2 Comparison of saliency maps obtained by different graphs

与式 (5) 中将  $dist(\mathbf{z}_i, \mathbf{z}_j)$  定义为欧氏距离的方式不同, 这里根据文献 [29], 将式 (7) 中两节点之间的空间位置距离  $dist(\mathbf{z}_i, \mathbf{z}_j)$  重新定义为:

$$dist(\mathbf{z}_i, \mathbf{z}_j) = \sqrt{\sin^2(\pi |x_i - x_j|) + \sin^2(\pi |y_i - y_j|)} \quad (8)$$

其中  $\mathbf{z}_i = (x_i, y_i)$ ,  $\mathbf{z}_j = (x_j, y_j)$ , 通过式 (8) 能够控制两节点空间位置距离  $dist(\mathbf{z}_i, \mathbf{z}_j)$  的上界. 在式 (7) 中如果依然利用式 (5) 中定义  $dist(\mathbf{z}_i, \mathbf{z}_j)$  的方式, 当两个节点在空间上相距较远, 但是颜色相似时, 利用式 (5) 计算得到的  $dist(\mathbf{z}_i, \mathbf{z}_j)$  较大, 这两个节点很有可能被分到不同的类中 (前景或背景). 实际上图像边界上的节点大多属于同一类 (背景类), 例如天空, 草地等. 采用式 (8) 计算这两个节点之间的空间位置距离  $dist(\mathbf{z}_i, \mathbf{z}_j)$  会得到一个较小的值, 将式 (8) 代入式 (7) 中能够得到一个较大的值, 意味着这两个节点之间有较大的相似性, 那么这两个节点被分到同一类的概率增大.

综上所述,  $G_1$  图模型中边的权值为:

$$W_1(e_{ij}) = \begin{cases} W_{1,Inner}(e_{ij}), e_{ij} \in E_{1,Inner} \\ W_{1,Boundary}(e_{ij}), e_{ij} \in E_{1,Boundary} \end{cases} \quad (9)$$

### 1.2.2 K 正则图模型

在自然图像中位置相近的图像区域属于同一类区域 (前景或背景区域) 的概率很大. 与文献 [22] 中构造图模型的方法相同, 这里将空间位置相近的节点进行连接, 利用 K 正则图 (K regular graph) 构造  $G_2 = (V_2, E_2)$  图模型. 在  $G_2$  图模型中每个节点不仅与它的相邻节点连接, 并且和与它相邻节点有共同边界的节点进行连接. 根据中心偏置原理, 显著目标大多位于图像中心, 图像的边界区域很可能属于同一区域 (背景区域), 故与  $G_1$  图模型的构造方式类似, 在  $G_2$  中本文也将图像的边界节点两两连接, 加强边界节点之间的连接性.  $G_2$  图模型中边的

权值为:

$$W_2(e_{ij}) = \begin{cases} \exp\left(\frac{-dist(\mathbf{c}_i - \mathbf{c}_j)}{2\sigma^2}\right), e_{ij} \in E_2 \\ 0, e_{ij} \notin E_2 \end{cases} \quad (10)$$

### 1.3 显著性融合

图模型  $G_1$  将颜色相近的节点连接起来, 得到的显著图能够比较完整地突出显著目标, 但是存在显著目标内部不均匀或者显著目标不能一致高亮的问题;  $G_2$  模型中将位置相近的节点连接起来, 得到的显著图能够使显著目标的内部一致高亮, 但是显著图中存在突出背景区域, 显著目标与背景区域边界不清晰的问题. 如图 2 所示, 图 2(a) 为原图, 图 2(b) 为 GT, 图 2(c) 为超像素分割图像, 图 2(d) 为仅利用  $G_1$  图模型生成的显著图, 图 2(e) 为仅使用  $G_2$  图模型生成的显著图, 图 2(f) 为  $G_1$  和  $G_2$  融合生成的显著图. 从图 2(d) 中可以看出检测到的显著目标虽然比较完整且边界清晰, 但其内部没有均匀一致高亮; 图 2(e) 中尽管显著目标内部均匀一致高亮但是背景噪声明显. 故本文提出了一种融合方式, 将这两种图模型得到的节点的显著性值在节点级上进行融合, 即在两种图模型中, 分别利用第 1.1 节中的式 (3) 计算节点的显著性值, 记为  $\mathbf{f}_1$  和  $\mathbf{f}_2$ , 将  $\mathbf{f}_1$  和  $\mathbf{f}_2$  进行加权融合, 得到超像素最终的显著性值:

$$\mathbf{f} = \mu_1 \times \mathbf{f}_1 + \mu_2 \times \mathbf{f}_2, \text{ 且 } \mu_1 + \mu_2 = 1 \quad (11)$$

其中  $\mathbf{f} = (f_1, f_2, \dots, f_n)^T$ ,  $f_i$  表示超像素  $i$  的显著性值,  $i = 1, 2, \dots, n$ . 从图 2(f) 中可以看出检测到的显著目标内部不仅一致均匀突出, 并且滤除了部分背景噪声.

### 1.4 显著图优化

利用第 1.3 节中式 (11) 计算超像素的显著性值  $\mathbf{f} = (f_1, f_2, \dots, f_n)^T$ , 其中  $f_i$  表示第  $i$  个超像素的显著性值,  $i = 1, 2, \dots, n$ , 将显著性值赋给超像素  $i$  所包含的所有像素作为每个像素的显著性值, 这样得到一个全分辨率的与原图像大小相等的初始显著图  $S_{co}$ .

由于显著目标大多处于图像的中央, 并且位于图像中央的区域更能引起人们的关注, 故本文引进中心先验图加强初始显著图  $S_{co}$  中显著目标的亮度. 由于初始显著图  $S_{co}$  的边界中存在一些噪声, 为了进一步滤除背景噪声, 本文引入边界抑制操作对初始显著图  $S_{co}$  进行优化. 首先, 根据中心先验思想对图像中心进行加强, 利用关于图像中心的各向同性的对称高斯模型模拟图像的中心先验图<sup>[31]</sup>:

$$S_{ce} = \exp\left(-\frac{(x_i - x_0)^2}{2\sigma_x^2} - \frac{(y_i - y_0)^2}{2\sigma_y^2}\right) \quad (12)$$

其中  $x_i$  和  $y_i$  是像素的水平和垂直坐标,  $x_0$  和  $y_0$  是图像中心像素的水平和垂直坐标,  $\sigma_x$  和  $\sigma_y$  分别代表水平方向和垂直方向上的方差,  $\sigma_x$  和  $\sigma_y$  取值相同. 将中心先验图  $S_{ce}$  与初始显著图  $S_{co}$  进行融合得到优化后的显著图:

$$S = S_{ce} \cdot S_{co} \quad (13)$$

式(13)中“ $\cdot$ ”表示点乘, 即将  $S_{co}$  与  $S_{ce}$  两幅显著图中对应像素点的显著值逐个相乘, 得到显著图  $S$  中对应像素的显著值.

其次, 在显著图  $S$  的基础上, 通过 MATLAB 中的 `imclearborder` 函数的边界抑制操作抑制  $S$  中与边界相连的高亮区域的亮度, 即将与边界相连的且面积较小的高亮区域的亮度降低, 以进一步去除背景噪声, 得到优化的显著图.

## 2 实验结果与分析

### 2.1 数据库及评价指标

为了验证本文方法的有效性, 将本文提出的方法在 MSRA-10K<sup>[32]</sup>、SED2<sup>[33]</sup> 和 ECSSD<sup>[34]</sup> 这三个公开的数据集上进行测试. MSRA-10K 数据库是目前该领域最常用的一个数据库, 该库中共包含 10000 张图像. 数据库中的图像显著目标的种类繁多, 但是大多数只包含一个显著目标, 并且图像的背景相对简单. SED2 数据库包含 100 张图像, 每张图像中包含两个显著目标. ECSSD 数据库包含 1000 幅内容丰富且背景复杂的图像. 这三个数据库都提供了人工标记的二值图像作为 GT, 以便于评价显著图的性能.

对于本文方法所涉及的参数, 由于数据库中的图像多为  $400 \times 300$  左右, 为了保证算法时间效率的同时又不影响检测结果, 本文将初始超像素个数设置为 200. 在构造 KNN 图模型时, 如果每个节点的邻接节点的个数  $K$  的取值太小会影响目标内部的完整性,  $K$  取值太大会检测到较多的噪声, 实

验中选取了多种  $K$  值进行对比实验, 结果显示当  $K = 5$  时, 实验结果最好. 在式(4)中,  $\lambda_z$  和  $\lambda_c$  分别用来控制位置特征与颜色特征在距离公式中所占的权重, 当  $\lambda_z$  过大时, 检测到的显著目标边界不明确, 当  $\lambda_z$  过小时, 显著目标内部不完整, 实验中设置多组  $\lambda_z$  与  $\lambda_c$  的值, 当  $\lambda_z = 0.5$ ,  $\lambda_c = 1$  实验结果最好, 故本文中设置  $\lambda_z = 0.5$ ,  $\lambda_c = 1$ . 式(6)、(7)和(10)中的  $\sigma^2$  均设置为经验值 0.05. 式(11)中的  $\mu_1$  和  $\mu_2$  分别用来控制不同图模型下超像素块的显著性值融合时的权重, 实验中选取了多组  $\mu_1$  和  $\mu_2$  的值, 当  $\mu_1 = \mu_2 = 0.5$  时实验结果最好. 式(12)中  $\sigma_x^2$  和  $\sigma_y^2$ , 实验中选择多组参数值进行对比实验, 当  $\sigma_x^2 = \sigma_y^2 = 141$  实验结果最好.

图像显著性检测方法很多, 本文选取 14 种常用的显著性检测算法进行对比, 这 14 种算法分别是基于全局对比度的 HC 和 RC 算法, 基于局部对比度的 AC 算法, 基于分层显著性融合的 HS 算法, 基于频域的 SR 算法、FT 算法和 MSS 算法, 基于背景先验的 GS 算法, MR 算法和 BFSS 算法, 基于随机游走的 RW 算法, 基于高维颜色转化的 HDCT 算法, 基于二值图拓扑分析的 BMA 算法和基于正则随机游走排序的 RR 算法. 所用对比算法源码链接<sup>1</sup> 和本文算法源码<sup>2</sup> 可见相关网页. MR 和 GS 算法都是基于图模型的显著性检测算法, MR 是利用基于图模型的流形排序算法进行显著性检测, GS 是基于图模型上的测地距离计算显著性. 本文引入四种常用的评价指标评测提出的算法, 这四种指标分别是 P-R (Precision-recall) 曲线、ROC 曲线、F-measure 和 AUC 值<sup>[34]</sup>.

#### 1) P-R 曲线

P-R 曲线是最常用的评价显著性检测算法优劣的指标. 对于一幅显著图  $S$ , 将  $S$  中像素值归一化到  $[0, 1]$ , 然后从 0 到 1 依次选取阈值对  $S$  进行二值化, 得到二值图像  $M$ ,  $GT$  表示 Ground-truth. 根据式(14) 计算查准率 (Precision) 和查全率 (Recall):

$$\begin{aligned} \text{Precision} &= \frac{|M \cap GT|}{|M|} \\ \text{Recall} &= \frac{|M \cap GT|}{|GT|} \end{aligned} \quad (14)$$

#### 2) ROC 曲线

ROC 曲线的绘制方式与 P-R 曲线类似, 对于显著图  $S$ , 将  $S$  调整到  $[0, 1]$ , 然后从 0 到 1 依次选取阈值对  $S$  进行二值化, 得到二值图像  $M$ ,  $GT$  表示 Ground-truth,  $\overline{GT}$  表示  $GT$  取反. 根据式(15) 计算真阳性率 (TPR) 和假阳性率 (FPR):

<sup>1</sup><https://github.com/libozhao/saliency-detection-methods/blob/master/README.md>

<sup>2</sup><https://github.com/libozhao/saliency>

$$\begin{aligned} \text{FPR} &= \frac{|M \cap GT|}{|M \cap GT| + |\overline{M} \cap \overline{GT}|} \\ \text{TPR} &= \frac{|M \cap GT|}{|GT|} \end{aligned} \quad (15)$$

相同真阳性率下, 假阳性率越小, 显著性检测算法越好。

### 3) F-measure

F-measure 用来综合衡量查全率和查准率:

$$F_{\beta} = \frac{(1 + \beta^2)\text{Precision} \times \text{Recall}}{\beta^2\text{Precision} + \text{Recall}} \quad (16)$$

式(16)中,  $\beta$  用来决定查全率和查准率的影响程度, 为了强调准确率的重要性, 通常  $\beta^2$  设置为 0.3。F 值越大表明显著性检测算法的性能越好。

### 4) AUC 值

AUC 值被定义为 ROC 曲线下面积, 即 ROC 曲线与 X 轴包围的面积。ROC 曲线有较小的假阳性率和较大的真阳性率是最优的, 所以 AUC 值越大则表示显著性算法性能越好。

本文首先使用一系列的固定阈值对数据库中所有图像在某个算法下的显著图进行二值化, 利用得到的二值图像与  $GT$  进行比较, 计算每幅显著图在不同阈值下的 Precision 和 Recall, 然后在每个阈值下分别计算所有图像的 Precision 和 Recall 的平均值, 作为该阈值下的 Precision 和 Recall 绘制 P-R 曲线。ROC 曲线的绘制方式与 P-R 曲线类似, 也是利用一系列的固定阈值。与 Precision 和 Recall 的计算方式不同, F 值的计算方式是利用自适应阈值对每幅显著图进行二值化, 然后与  $GT$  进行比较, 计算每幅图像的 F, 然后计算所有图像 F 的平均值, 作为该方法在该数据库上的 F。

## 2.2 多图模型有效性实验比较

为了验证多图模型的有效性, 本文分别在数据库 MSRA-10K、SED2 和 ECSSD 上进行了三组实验。这三组实验分别是仅利用 K 正则图模型生成显著图, 仅利用 KNN 图模型生成显著图, 以及融合两种模型计算得到的显著图, 并计算这三组实验的

P、R 和 F 值。表 1 是在数据库 MSRA-10K、SED2 和 ECSSD 上的对比实验结果。表 1 中 K 正则表示仅利用 K 正则图模型生成显著图, KNN 表示仅利用 KNN 图模型生成显著图, K 正则 + KNN 表示将 K 正则图模型下生成的显著图与 KNN 模型下生成的显著图进行融合。

从表 1 的第 2~4 列和第 8~10 列可以看出在 MSRA-10K 和 ECSSD 数据库上, 与仅利用 K 正则图和 KNN 图模型的实验结果相比, 将 K 正则图和 KNN 图模型进行融合的实验结果优于仅利用单一图模型的实验结果。在 SED2 数据库中, 由于存在前景与背景颜色非常相似的图片, 在构造 KNN 图模型时, 会错误地将背景节点与前景节点连接起来, 故 KNN 图模型下生成的显著图准确率低于 K 正则图模型下生成的显著图。但是将 K 正则图模型与 KNN 图模型下生成的显著图进行融合, 结果优于仅利用 KNN 图模型生成的显著图, 如表 1 第 5~7 列所示。故将两个图模型下生成的显著图进行融合能够改善实验结果。

## 2.3 全局前景假设有效性实验比较

为了验证本文中全局前景假设的有效性, 本文在数据库 MSRA-10K、SED2 和 ECSSD 上分别进行对比实验, 在这三个数据库上分别利用边界假设生成显著图, 利用全局前景假设生成显著图, 并计算这两类显著图的 P、R 和 F 值记录在表 2 中。表 2 中“边界”表示利用边界假设生成的显著图。“全局”表示利用全局前景假设生成的显著图。

从表 2 可以看出, 利用全局前景假设生成的显著图在三个数据库上的 R 值平均提升 31.0%, F 值平均提升 8.6%, 优于利用边界假设生成的显著图。

## 2.4 优化策略有效性实验比较

为了验证本文所用优化策略的有效性, 分别在数据库 MSRA-10K、SED2 和 ECSSD 上计算初始显著图与优化显著图, 并计算相应的 P、R 和 F 值将结果记录在表 3 中。表 3 中的“初始图”表示初始显著图, “优化图”表示优化后的显著图。

从表 3 第 2~4 列和第 8~10 列可以看出, 在 MSRA-10K 和 ECSSD 数据库上, 与初始显著图相

表 1 K 正则图模型、KNN 图模型和 K 正则图模型 + KNN 图模型比较

Table 1 The comparison of K regular graph, KNN graph, and K regular graph + KNN graph

	MSRA-10K			SED2			ECSSD		
	P	R	F	P	R	F	P	R	F
K 正则	0.8557	0.7705	0.7948	0.8293	0.6990	0.7524	0.7053	0.6950	0.6344
KNN	0.8627	0.7468	0.7897	0.7918	0.6034	0.6837	0.7231	0.7229	0.6677
K 正则 + KNN	0.8723	0.7962	0.8228	0.8080	0.7253	0.7504	0.7279	0.7476	0.6828

表 2 边界假设和全局前景假设比较

Table 2 The comparison of boundary assumption and global foreground assumption

	MSRA-10K			SED2			ECSSD		
	P	R	F	P	R	F	P	R	F
边界	0.8815	0.6670	0.7846	0.7267	0.4883	0.6156	0.7927	0.5830	0.6768
全局	0.8723	0.7962	0.8228	0.8080	0.7253	0.7504	0.7279	0.7476	0.6828

表 3 初始显著图和优化后显著图的比较

Table 3 The comparison of original saliency maps and refined saliency maps

	MSRA-10K			SED2			ECSSD		
	P	R	F	P	R	F	P	R	F
初始图	0.8723	0.7962	0.8228	0.8080	0.7253	0.7504	0.7279	0.7476	0.6828
优化图	0.8796	0.7880	0.8327	0.8210	0.6729	0.7456	0.7587	0.7152	0.7034

比, 优化后的显著图在 P 值和 F 值上均有提升. 数据库 SED2 中, 图片数量小并且存在大量显著目标靠近边界的图片, 在该数据库上优化后的显著图比初始显著图在 F 值上有所下降. 但是在 MSRA-10K 和 ECSSD 上, 加入中心先验和边界抑制进行优化后, 实验结果中 F 值和 P 值均有所提升. 由于大多数图像的显著目标处于图像的中央, 并且位于图像中央的区域更能引起人们的关注, 故本文引入优化策略对初始显著图进行优化.

## 2.5 MSRA-10K 数据库上的实验结果

本文算法与其他的 14 种算法在 MSRA-10K 数据库上的部分实验结果的对比图见图 3. 从图 3 中可以看出, 基于对比度的 HC 和 AC 算法, 在背景与前景颜色相似时, 会错误地突出背景区域. FT 和 MSS 算法, 虽然保留了比较完整的图像信息, 但是没有一致高亮显著目标. SR 算法没有保留足够的高频信息, 导致显著目标的边界不清晰. 当显著目标之间存在较大的差异时, MR 算法会出现漏检的现象, 如图 3 中 MR 算法的 (a)、(b) 所示; 当显著目标与边界相连或者与边界颜色相似时, 该算法会错误地突出背景区域, 如图 3 中 MR 算法的 (c)、(d) 图像所示. RC 和 HS 的检测结果相似, 基本能够准确地定位显著目标, 但是从图 3 中的 (d)、(e) 可以

看出 RC 和 HS 算法都存在错误突出背景区域的问题. 当显著目标与背景颜色相似时, HDCT 算法生成的显著图中不能正确地检测出显著目标, 如图 3 中 HDCT 算法的 (c)、(e) 所示. BFSS 算法也是基于边界假设的一种算法, 当显著目标与边界颜色相似时会错误地突出背景区域, 例如 BFSS 算法中的 (d)、(e) 所示. RW 和 BMA 算法生成的显著图中显著目标的边界不清晰. RR 算法生成的显著图中存在较多的背景噪声, 如图 3 中的 RR 算法 (d) 和 (e) 所示, 当显著目标之间存在较大颜色差异时, RR 算法存在漏检的现象, 如图 3 中 RR 算法 (a) 和 (b) 所示. 本文算法不仅能够准确地定位显著目标, 并且显著目标内部更加光滑, 边界更加清晰, 并且有效减少背景噪声.

为了更加客观地进行比较, 本文中利用 P-R 曲线、ROC 曲线和 F 值进行定量对比. 图 4 是 MSRA-10K 数据库上各种算法的 P-R 曲线、ROC 曲线和 F 值, 其中图 4(a) 和图 4(b) 为 P-R 曲线, 图 4(c) 和图 4(d) 为 ROC 曲线, 图 4(e) 和图 4(f) 为 F 值. 图 4 中第一列表示本文算法与 FT、GS、AC、MSS、HS、MR、RC、HC 和 SR 算法的对比, 图 4 中第二列表示本文所提算法与 BFSS、RW、HDCT、BMA 和 RR 算法的对比. 在后文中图 6 和图 8 中的排列方式同图 4. 从图 4(a)

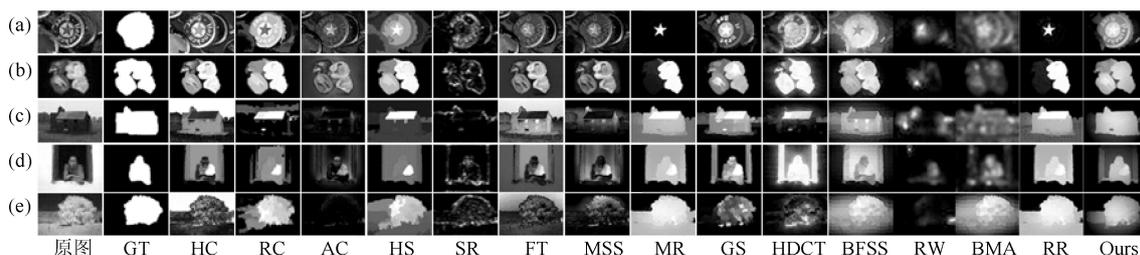


图 3 MSRA-10K 数据库上实验结果

Fig. 3 Experimental results on the MSRA-10K database

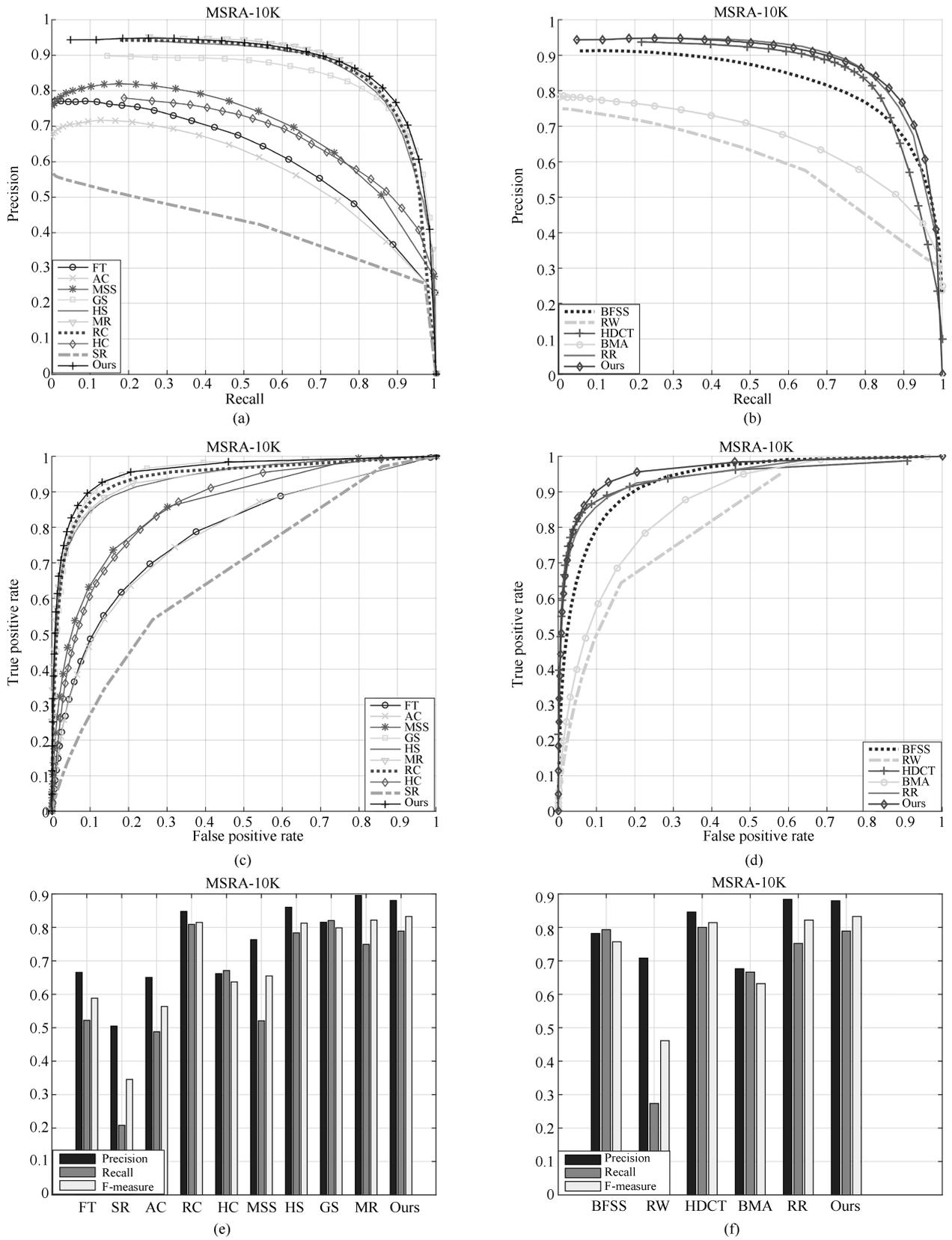


图4 MSRA-10K 数据库上的 P-R 曲线、ROC 曲线和 F 值  
Fig.4 P-R curves, ROC curves, and F values on the MSRA-10K database

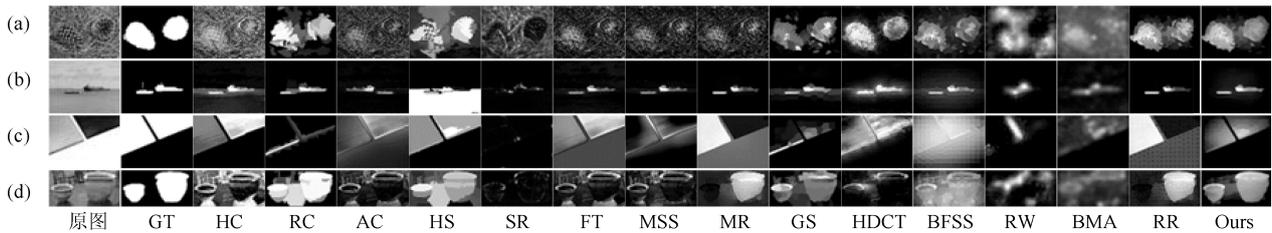


图5 SED2 数据库上实验结果

Fig.5 Experimental results on the SED2 database

和 (b) 中可以看出本文算法的 P-R 曲线与 MR、RC、HS 和 RR 算法基本持平且优于其他算法. 图 4(c) 和 (d) 中显示本文算法的 ROC 曲线优于其他算法. 图 4(e) 中显示本文算法的查全率 (Recall) 高于 MR 算法低于 GS 算法, 查准率 (Precision) 略低于 MR 算法但高于 GS 算法. 说明本文算法在保持较高查准率的同时, 能够比较完整地检测到图像的显著目标.

## 2.6 SED2 数据库上的实验结果

SED2 数据库中的图像大多含有两个显著目标, 并且有的显著目标靠近图像边界, 相比 MSRA-10K 数据库中只包含单个显著目标, SED2 更具挑战性, 但是该数据库中的图片背景比较简单. 图 5 是 SED2 数据库上的部分实验结果对比图. 从图 5 中可以看出, HC 和 RC 算法生成的显著图背景噪声较多, 显著目标的边界不明确. AC、SR、FT 和 MSS 算法没有突出显著目标. 基于背景先验的 MR 算法, 当显著目标靠近边界时, 会将显著目标错误的检测为背景, 并且存在漏检的现象. GS 算法也存在与 MR 相似的问题, 并且生成的显著图中背景噪声较多. HDCT 算法基本能够准确地定位显著目标, 但是当显著目标比较大时, 算法会失效, 如图 5 中 HDCT 算法的 (d) 所示. 当显著目标与边界相连或者靠近边界时, BFSS 算法会错误地突出背景区域. BMA 算法生成的显著图没有突出显著目标, 如图 5 中 BMA 算法所示. RR 算法检测到的显著目标不完整, 如图 5 中的 RR 算法 (c) 和 (d) 所示. 本文算法得到的显著目标不仅一致均匀, 并且显著图中背景噪声相对较少.

图 6 是 SED2 数据库上的 P-R、ROC 曲线和 F 值. 图 6(a) 中显示, 在该数据库上, 除了 SR 的 P-R 曲线比较低以外, 其他 9 种算法的 P-R 曲线比较接近, HS 的 P-R 曲线最高, 本文算法略低于 HS. 图 6(b) 中显示, 本文算法的 P-R 曲线与 BFSS、RW、HDCT 算法基本持平. 图 6(c) 显示, 除了 SR 的 ROC 曲线比较低以外, 其他的曲线比较接近. 图 6(d) 显示本文算法的 ROC 曲线于其他 5 个算法基本持平. 图 6(e) 显示本文算法的查准率略低于最高的 MSS 算法; 图 6(f) 显示本文算法的查准

率略低于 HDCT 算法. 由于该数据库中存在显著目标非常小的图片, 本文算法计算该种图片的显著图时, 在利用 SLIC 算法对原始图像进行超像素分割时可能会将显著目标与背景区域分割到同一超像素块中, 故本文算法在该数据库中的查全率低于基于像素点对比度的 HC 算法, 但是由于引进了基于颜色特征的 KNN 图模型, 本文算法的查全率高于 MR 算法. 通过比较发现本文算法在该数据库上保持高查准率的同时能够较完整地突出显著目标.

## 2.7 ECSSD 数据库上的实验结果

ECSSD 数据库中的图片不仅背景复杂, 并且显著目标的纹理也很复杂. 比 MSRA-10K 和 SED2 更具挑战性. 图 7 是 ECSSD 数据库上实验结果的对比图. 图 7 中第 (a) 行的原始图像中显著目标与背景中的深色草地部分颜色相似, 除了 GS 和本文算法外, 其他的算法都突出了部分背景区域或者没有完整突出显著目标. 与 GS 相比, 本文算法的显著目标内部更加均匀; 图 7 第 (b) 行中, MR 算法只突出了部分显著目标; 图 7 第 (c) 行中显著目标纹理比较复杂, 只有 MR、GS 和本文算法能够比较均匀地突出显著目标, 但是本文算法的显著目标内部更加光滑完整; 图 7 第 (d) 行中显著目标颜色与背景相近, 只有 RC、MR 和本文算法能够较好地突出显著目标; 图 7 第 (e) 行的原始图像中显著目标各部分颜色差异较大, 本文算法相对完整地突出了显著目标.

图 8 是该数据库上不同算法的 P-R 曲线、ROC 曲线和 F 值. 从图 8(a)、(b) 和 (c)、(d) 可以看出本文算法的 P-R 曲线和 ROC 曲线都是该数据库上最高的. 无论从 P-R 曲线还是 ROC 曲线都可以看出, 本文算法优于当前流行的 14 种算法. 图 8(e) 和 (f) 可以看出本文算法的 F 值是该数据库上最高的, 表明本文算法能够比较准确地检测显著目标, 并且显著目标更加完整.

表 4 是各种算法在不同数据库上的 AUC 值和 F 值, 从表 4 第 2 列和第 5 列可以看出本文算法在 MSRA-10K 数据库上的 AUC 值为 0.9532, F 值为 0.8327, 这两个指标都是该数据库上最优的. 从表 4

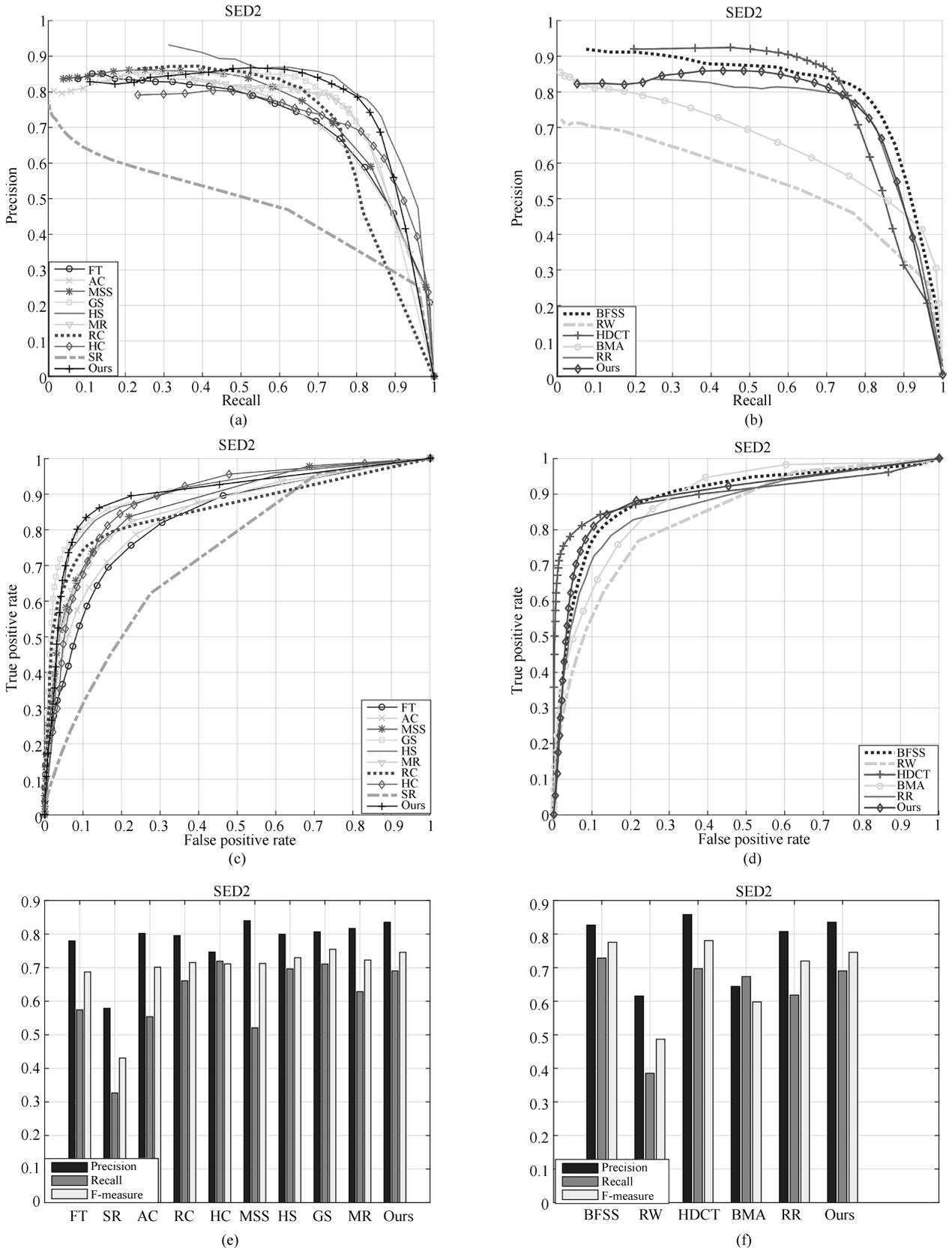


图6 SED2 数据库上的 P-R 曲线、ROC 曲线和 F 值  
Fig.6 P-R curves, ROC curves, and F values on the SED2 database

第3列和第6列可以看出本文算法在 SED2 数据库上的 AUC 值达到 0.8937, F 值达到 0.7456. 从表 4 第4列和第7列可以看出本文算法在 ECSSD 数据库上的 AUC 值达到 0.9114, F 值达到 0.7034, 这两个指标都是该数据库上最优的. 通过以上的比较发现, 本文算法能够准确地检测出显著目标, 并且能够一致均匀地突出显著目标.

本文算法采用 MATLAB R2014b 编程实现, 系统环境为 Windows10, CPU 频率为 3.6 GHz, 8 GB 内存. 表 5 为不同算法在该实验环境下运行数据库 ECSSD、SED2 上所有图片和 MSRA-10K 上前 1000 张图片所用时间的平均值. 从表 5 中可以看出本文算法的平均运行时间为 1.30 秒, 结合本文算法的检测效果, 本文提出的算法具有优越性.

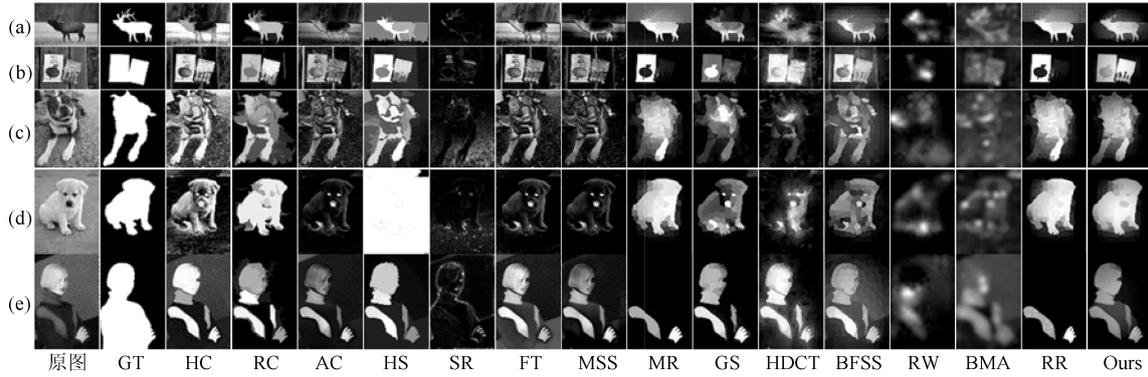


图7 ECSSD 数据库上实验结果

Fig. 7 Experimental results on the ECSSD database

表4 各种方法在不同数据库上的 AUC 值和 F 值

Table 4 The AUC and F values of the various methods on different databases

	AUC 值			F 值		
	MSRA-10K	SED2	ECSSD	MSRA-10K	SED2	ECSSD
HC <sup>[8]</sup>	0.8589	0.8839	0.7022	0.6370	0.7113	0.4186
RC <sup>[8]</sup>	0.9406	0.8580	0.8919	0.8150	0.7151	0.6802
AC <sup>[6]</sup>	0.7767	0.8430	0.6750	0.5636	0.7012	0.4026
HS <sup>[7]</sup>	0.9353	0.8960	0.8833	0.8128	0.7298	0.6657
SR <sup>[12]</sup>	0.6716	0.7290	0.6054	0.3454	0.4307	0.3001
FT <sup>[10]</sup>	0.7797	0.8310	0.6591	0.5884	0.6872	0.3959
MSS <sup>[11]</sup>	0.8607	0.8768	0.7678	0.6549	0.7125	0.4918
MR <sup>[22]</sup>	0.9404	0.8618	0.8858	0.8219	0.7227	0.6868
GS <sup>[18]</sup>	0.9527	0.9055	0.8834	0.7987	0.7546	0.6268
BFSS <sup>[13]</sup>	0.8807	0.8932	0.8625	0.7570	0.7755	0.6080
RW <sup>[20]</sup>	0.3976	0.8364	0.7252	0.4614	0.4868	0.3746
HDCT <sup>[14]</sup>	0.8483	0.9036	0.8644	0.8143	0.7804	0.6673
BMA <sup>[15]</sup>	0.8605	0.8825	0.7762	0.6318	0.5980	0.4927
RR <sup>[21]</sup>	0.9423	0.8674	0.8887	0.8219	0.7197	0.6882
本文算法	0.9532	0.8937	0.9114	0.8327	0.7456	0.7034

表5 各种方法平均运行时间

Table 5 The average runtimes of different methods

	SR	FT	MSS	MR	GS	BFSS	RW	BMA	RR	HS	HC	RC	AC	HDCT	Ours
MATLAB	0.02	0.04	1.04	1.10	0.21	7.25	1.33	0.002	1.70	—	—	—	—	—	1.30
C++	—	—	—	—	—	—	—	—	—	0.39	0.01	0.25	0.18	4.12	—

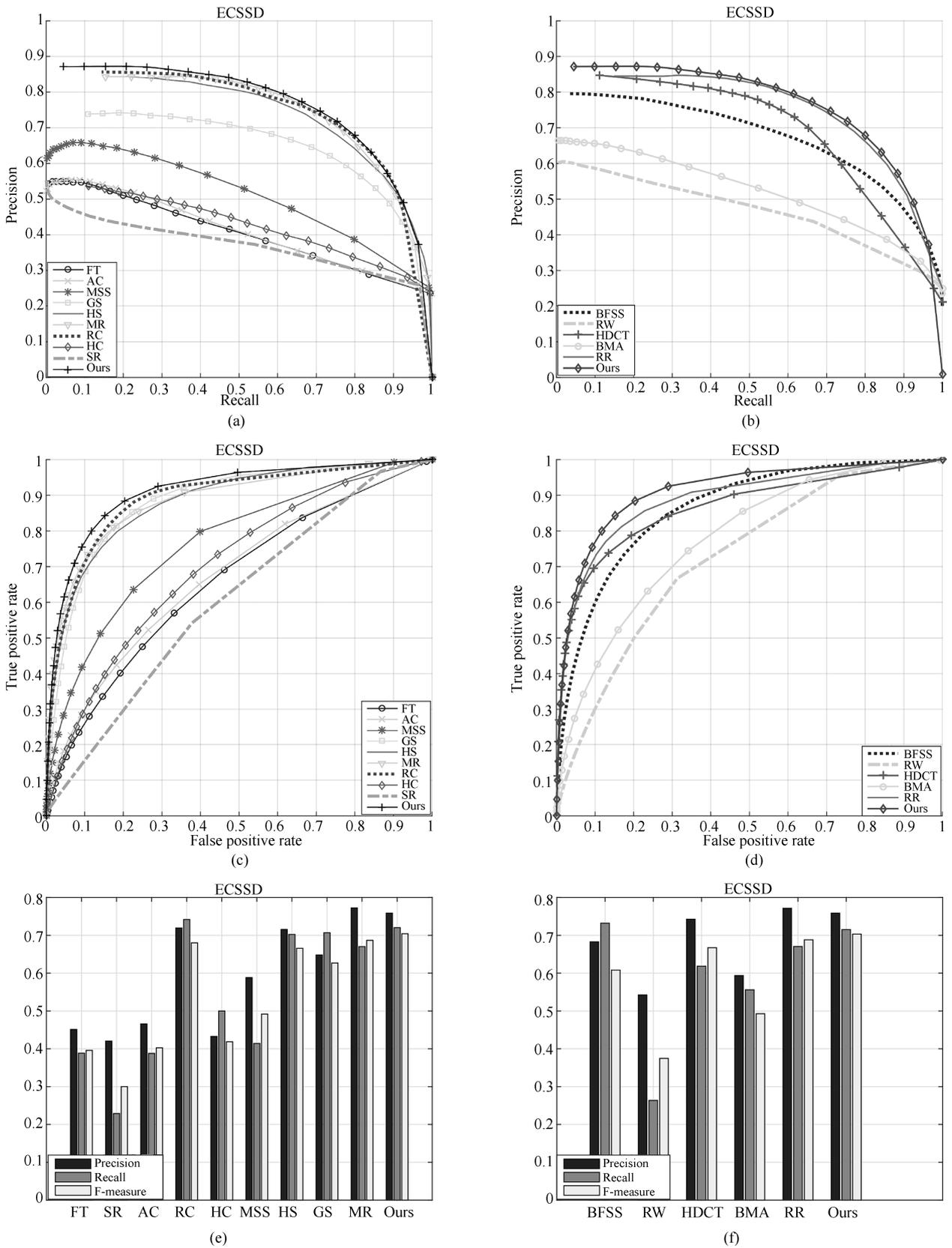


图8 ECSSD 数据库上的 P-R 曲线、ROC 曲线和 F 值

Fig.8 P-R curves, ROC curves, and F values on the ECSSD database

### 3 结论

本文提出了一个基于多图流形排序的图像显著性检测算法. 该算法分别利用超像素的位置特征和颜色特征, 基于 K 正则图和 KNN 图构造两类图模型, 分别在两种图模型上利用流形排序算法计算超像素的显著性值, 并将每个图模型下得到的超像素的显著性值进行加权融合得到超像素最终的显著性值. 本文将提出的算法与当前流行的 14 种算法在 MSRA-10K、SED2 和 ECSSD 数据库上进行了对比, 无论是在视觉效果上还是在定量指标上, 本文算法都具有明显优势. 针对背景与前景颜色相似的图片, 本文算法的检测效果需要进一步提高, 未来考虑提取更多的特征比如超像素的形状特征, 并结合流形排序算法提取更加准确的显著目标.

### References

- Luo Jian-Hao, Wu Jian-Xin. A survey on fine-grained image categorization using deep convolutional features. *Acta Automatica Sinica*, 2017, **43**(8): 1306–1318  
(罗建豪, 吴建鑫. 基于深度卷积特征的细粒度图像分类研究综述. 自动化学报, 2017, **43**(8): 1306–1318)
- Wang Rui-Xia, Peng Guo-Hua. An image retrieval method with sparse coding based on riemannian manifold. *Acta Automatica Sinica*, 2017, **43**(5): 778–788  
(王瑞霞, 彭国华. 基于黎曼流形稀疏编码的图像检索算法. 自动化学报, 2017, **43**(5): 778–788)
- Yang Sai, Zhao Chun-Xia, Xu Wei. A novel salient object detection method using bag-of-features. *Acta Automatica Sinica*, 2016, **42**(8): 1259–1273  
(杨赛, 赵春霞, 徐威. 一种基于词袋模型的新的显著性目标检测方法. 自动化学报, 2016, **42**(8): 1259–1273)
- Xu M, Jiang L, Ye Z T, Wang Z L. Bottom-up saliency detection with sparse representation of learnt texture atoms. *Pattern Recognition*, 2016, **60**: 348–360
- Goferman S, Zelnikmanor L, Tal A. Context-aware saliency detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, **34**(10): 1915–1926
- Achanta R, Estrada F, Wils P, Süsstrunk S. Salient region detection and segmentation. In: Proceedings of Computer Vision Systems, Lecture Notes in Computer Science, Vol. 5008. Berlin, Heidelberg: Springer, 2008. 66–75
- Yan Q, Xu L, Shi J P, Jia J Y. Hierarchical saliency detection. In: Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition. Portland, OR, USA: IEEE, 2013. 1155–1162
- Cheng M M, Mitra N J, Huang X L, Torr P H S, Hu S M. Global contrast based salient region detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, **37**(3): 569–582
- Borji A, Itti L. Exploiting local and global patch rarities for saliency detection. In: Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence, RI, USA: IEEE, 2012. 478–485
- Achanta R, Hemami S, Estrada F, Süsstrunk S. Frequency-tuned salient region detection. In: Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami, FL, USA: IEEE, 2009. 1597–1604
- Achanta R, Süsstrunk S. Saliency detection using maximum symmetric surround. In: Proceedings of the 17th IEEE International Conference on Image Processing. Hong Kong, China: IEEE, 2010. 2653–2656
- Hou X D, Zhang L Q. Saliency detection: a spectral residual approach. In: Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis, Minnesota, USA: IEEE, 2007. 1–8
- Wang J P, Lu H C, Li X H, Tong N, Liu W. Saliency detection via background and foreground seed selection. *Neurocomputing*, 2015, **152**: 359–368
- Kim J, Han D, Tai Y W, Kim J. Salient region detection via high-dimensional color transform. In: Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA: IEEE, 2014. 883–890
- Zhang J M, Sclaroff S. Exploiting surroundedness for saliency detection: a boolean map approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, **38**(5): 889–902
- Itti L, Koch C, Niebur E. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1998, **20**(11): 1254–1259
- Harel J, Koch C, Perona P. Graph-based visual saliency. In: Proceedings of the 19th International Conference on Neural Information Processing Systems. Canada: ACM, 2006. 545–552
- Wei Y C, Wen F, Zhu W J, Sun J. Geodesic saliency using background priors. In: Proceeding of the 2012 European Conference on Computer Vision, Lecture Notes in Computer Science, Vol. 7574. Berlin, Heidelberg: Springer-Verlag, 2012. 29–42
- Li X, Li Y, Shen C H, Dick A, Van Den Hengel A. Contextual hypergraph modeling for salient object detection. In: Proceeding of the 2013 IEEE International Conference on Computer Vision. Sydney, Australia: IEEE, 2013. 3328–3335
- Li J, Fang S, Tian Y H, Huang T J, Chen X W. Image saliency estimation via random walk guided by informativeness and latent signal correlations. *Signal Processing: Image Communication*, 2015, **38**: 3–14
- Yuan Y C, Li C Y, Kim J, Cai W D, Feng D D. Reversion correction and regularized random walk ranking for saliency detection. *IEEE Transactions on Image Processing*, 2018, **27**(3): 1311–1322
- Yang C, Zhang L H, Lu H C, Ruan X, Yang M H. Saliency detection via graph-based manifold ranking. In: Proceeding of the 2013 IEEE Conference on Computer Vision and Pattern Recognition. Portland, USA: IEEE, 2013. 3166–3173
- Tao D P, Cheng J, Song M L, Lin X. Manifold ranking-based matrix factorization for saliency detection. *IEEE Transactions on Neural Networks and Learning Systems*, 2016, **27**(6): 1122–1134
- Achanta R, Shaji A, Smith K, Lucchi A, Fua P, Süsstrunk S. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, **34**(11): 2274–2281

- 25 Zhao M Y, Jiao L C, Ma W P, Liu H Y, Yang S Y. Classification and saliency detection by semi-supervised low-rank representation. *Pattern Recognition*, 2016, **51**: 281–294
- 26 Zhu W J, Liang S, Wei Y C, Sun J. Saliency optimization from robust background detection. In: Proceeding of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA: IEEE, 2014. 2814–2821
- 27 Chen J Z, Ma B P, Cao H, Chen J, Fan Y B, Li R, et al. Updating initial labels from spectral graph by manifold regularization for saliency detection. *Neurocomputing*, 2017, **266**: 79–90
- 28 Li S Q, Zeng C, Fu Y, Liu S P. Optimizing multi-graph learning based salient object detection. *Signal Processing: Image Communication*, 2017, **55**: 93–105
- 29 Zhang J X, Ehinger K A, Wei H K, Zhang K J, Yang J Y. A novel graph-based optimization framework for salient object detection. *Pattern Recognition*, 2017, **64**: 39–50
- 30 Lan Y D, Deng H F, Chen T. A new method of distance measure for graph-based semi-supervised learning. In: Processing of the 2011 International Conference on Machine Learning and Cybernetics. Guilin, China: IEEE, 2011. 1444–1448
- 31 Jiang H Z, Wang J D, Yuan Z J, Liu T, Zheng N N. Automatic salient object segmentation based on context and shape prior. In: Proceedings of the 22nd British Machine Vision Conference. Dundee, Scotland, British: BMVA Press, 2011. 110.1–110.12
- 32 Liu T, Yuan Z J, Sun J, Wang J D, Zheng N N, Tang X O, et al. Learning to detect a salient object. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, **33**(2): 353–367
- 33 Alpert S, Galun M, Brandt A, Basri R. Image segmentation by probabilistic bottom-up aggregation and cue integration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, **34**(2): 315–327
- 34 Guo Ying-Chun, Yuan Hao-Jie, Wu Peng. Image saliency detection based on local and regional features. *Acta Automatica Sinica*, 2013, **39**(8): 1214–1224  
(郭迎春, 袁浩杰, 吴鹏. 基于 Local 特征和 Regional 特征的图像显著性检测. 自动化学报, 2013, **39**(8): 1214–1224)



于 明 河北工业大学计算机科学与软件学院教授. 1999 年于北京理工大学获得通信与信息系统专业博士学位. 主要研究方向为语音与图像视觉信息融合的生物特征识别, 图像数学变换、图像与视频编码的高效算法.

E-mail: yuming@hebut.edu.cn

(YU Ming Professor at School of Computer Science and Engineering, Hebei University of

Technology. He obtained his doctoral degree in communication and information systems from Beijing Institute of Technology in 1999. His research interest covers biometrics of voice and image vision information fusion, image mathematical transformation, efficient algorithm of image and video coding.)



李博昭 河北工业大学计算科学与软件学院硕士研究生. 2015 年获得河北工业大学信息与计算科学专业学士学位. 主要研究方向为图像显著性检测.

E-mail: 201522102010@hebut.edu.cn

(LI Bo-Zhao Master student at School of Computer Science and Engineering, Hebei University of Technol-

ogy. She received her bachelor degree from Hebei University of Technology in 2015. Her main research interest is image saliency detection.)



于 洋 河北工业大学计算机科学与软件学院讲师. 2012 年获得河北工业大学微电子学与固体电子学博士学位. 主要研究方向为图像处理与分析、模式识别、计算机视觉, 智能交通系统. 本文通信作者. E-mail: yuyang@scse.hebut.edu.cn

(YU Yang Lecturer at School of Computer Science and Engineering,

Hebei University of Technology. He received his Ph.D. degree in microelectronics and solid state electronics from Hebei University of Technology in 2012. His research interest covers image processing and analysis, pattern recognition, computer vision and intelligent transportation system. Corresponding author of this paper.)



刘 依 河北工业大学计算机科学与软件学院讲师. 2003 年获得河北工业大学计算机应用技术专业硕士学位. 主要研究方向为图像处理与识别、图像显著性检测. E-mail: liuyi@scse.hebut.edu.cn

(LIU Yi Lecturer at School of Computer Science and Engineering, Hebei University of Technology. She received

her master degree in computer applied technology from Hebei University of Technology in 2011. Her research interest covers image processing and recognition, image saliency detection.)