

文章编号:1007-2780(2015)01-0120-06

基于纹理抑制和连续分布估计的 显著性目标检测方法

邓 丹*, 吴 谨, 朱 磊, 刘 劲

(武汉科技大学 信息科学与工程学院, 湖北 武汉 430081)

摘要:为了取得更好的显著性检测结果,针对传统的显著性检测方法易造成边界模糊以及应用中央-周边差进行图像检测时,感兴趣目标的内部纹理会破坏目标的整体性的问题,提出了一种基于纹理抑制和连续分布估计的显著性检测方法。先对图像进行双边滤波的预处理,以平滑目标以及背景区域内部的纹理扰动,保留目标与背景之间的主要边缘。再采用 SLIC 超像素分割算法,对图像中具有相同特征的像素进行分组,通过多维正态分布提取分割区域的特征,利用二范数 Wasserstein 距离计算区域相似度;结合局部显著性检测以及全局显著性检测实现目标区域的提取。实验结果表明,本文的方法能够较好地提取显著性目标区域。

关键词:显著性检测;背景先验;Wasserstein 距离;正态分布

中图分类号:TP391.4

文献标识码:A

doi:10.3788/YJYXS20153001.0120

Significant target detection method based on texture inhibition and continuous distribution estimation

DENG Dan*, WU Jin, ZHU Lei, LIU Jing

(College of Information Science and Engineering, Wuhan University,
of Sciences & Technology, Wuhan 430081, China)

Abstract: The traditional methods of saliency detection tend to blur the border and internal texture of interested target, which usually destroys the integrity of the target by applying the center-surround differences. As to this problem, this paper proposes a saliency detection method based on texture inhibition and continuous distribute estimation in order to obtain better saliency detection results. This paper proposes a saliency detection method based on texture inhibition and continuous distribute estimation. Firstly, bilateral filter is used to smooth texture disturbances of the target and background region, while retaining the major edges between the target and background. Then SLIC superpixel image segmentation algorithms are used to divide the pixels into groups which have lots of pixels with the same characteristics. Then features of segment region are extracted by multidimensional normal distribution, and the similarity of feature saliency is calculated by two norm Wasserstein distance. At last, saliency region can be extracted by the local and global saliency detection. The experiment result shows that the experiment method in this paper can extract saliency region effectively. The experiment results show that our method can obtain better detection result.

收稿日期:2014-03-04; **修订日期:**2014-06-06.

基金项目:湖北省自然科学基金(No. 2013CFB333);高等学校博士学科点专项科研基金(No. 20124219120002);湖北省教育厅科研计划(No. Q20131110)

* 通信联系人, E-mail: 527693165@qq.com

Key words: saliency detection; background prior; Wasserstein distance; normal distribution

1 引言

人类视觉系统提取出视觉场景中感兴趣的区域的过程包含两个阶段。首先是一个自底而上的,快速的,由数据驱动的过程。接着是一个自顶而下的,慢速的,由知识驱动的过程。在第一个阶段中,由于整个处理只与观察的场景有关,具有较高的客观性,近几年来在图像处理领域被广泛地研究。

图像显著性研究方法主要有3类:像素、固定块和区域。2007年,Hou等人^[1]提出了显著性检测的一个谱残差的方法,该方法主要是通过计算增量编码长度来测量信息熵增益,但是这种结构的前提是图像必须具备相对复杂的背景。同年Harel等人^[2]提出一种基于图的视觉显著性,该方法主要是在显著性区域检测中应用马尔可夫链的平稳分布作为图像的显著性度量。2009年Achanta^[3]提出一种调频显著性检测方法并于2010年对该方法进行了改进,该方法主要是通过每一个像素上计算当前像素的颜色以及亮度与全图像的平均颜色及亮度的差异来去除低频部分(模糊部分)。Goferman等人^[4]在2011年提出一种基于上下文的显著性检测方法,该方法使用局部对比,结合显著区域的上下文信息检测出现显著对象。在众多的显著性检测方法中以Itti算法^[5]最为典型,该方法通过中央-邻域差异算子(Center-surround differences)进行多尺度显著性度量并将度量结果合并成一幅显著图。

中心-邻域差异性被证明是目标之所以显著的首要原因,受到该原理的启发,很多文献提出了他们的显著性检测方法^[6-7]。这种测量通常是在图像的一个局部区域内完成的,在自然场景下不可避免地受到目标内部纹理扰动的影响。本文首先通过对图像进行双边滤波处理,在保留目标和背景之间的主要边缘的同时,平滑目标或者背景区域内部的扰动。

中心-邻域的差异性仅考虑一个像素或区域在局部范围内的显著性,并未考虑到目标在整幅图像中空间分布。在此基础上,本文结合被广泛使用的背景先验知识来强化自然图像中目标在空间分布上所具备的普遍特征。

2 本文提出的显著性检测方法

2.1 目标及背景区域内部的纹理抑制

双边滤波(Bilateral Filtering)最早是由Tomasi和Manduchi提出的^[8]。传统低通滤波器,在像素空间完成滤波以后,导致图像的边缘细节丢失。双边滤波器可以在抑制目标以及内部纹理的同时很好地保留目标边缘。这主要是因为双边滤波在处理相邻各像素时,不仅考虑到了灰度的空间邻近关系,同时考虑到其相似关系。

对于图像 $f(x)$,在空间范围内得到低通滤波的图像为 $I(x)$,则

$$I(x) = k_d^{-1}(x) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\epsilon) c(\epsilon, x) d\epsilon, \quad (1)$$

其中: $c(\epsilon, x)$ 代表中心点与其邻近点 ϵ 的空间邻近度,设 $x = (x_1, x_2)$, $\epsilon = (\epsilon_1, \epsilon_2)$ 为图像的空间坐标,那么

$$k_d(x) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} c(\epsilon, x) d\epsilon. \quad (2)$$

同理,在灰度范围内进行滤波得到的图像表示为:

$$I(x) = k_r^{-1}(x) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\epsilon) s(f(\epsilon), f(x)) d\epsilon, \quad (3)$$

其中: $s(f(\epsilon), f(x))$ 代表了中心点 x 与它的邻近点 ϵ 的灰度相似度,此时

$$k_r(x) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(f(\epsilon), f(x)) d\epsilon. \quad (4)$$

与空间邻近度不同的是,灰度相似度是由图像间两点灰度的差值来决定的。根据式(1)与式(3),将空间邻近度与灰度相似度相结合进行滤波,这样得到的输出图像为:

$$I(x) = k^{-1}(x) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\epsilon) c(\epsilon, x) s(f(\epsilon), f(x)) d\epsilon, \quad (5)$$

其中:

$$k(x) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} c(\epsilon, x) s(f(\epsilon), f(x)) d\epsilon. \quad (6)$$

目前对图像的处理大都以像素为单位,用二维矩阵来表示一张图像,但是这样没有考虑像素之间的空间组织关系,在很大程度上会降低算法处理效率。2003年Ren^[9]等提出超像素的概念,并在2010年利用图像超像素(superpixel)提取来

获取一致性片段,然后对超像素进行 GMM 模糊聚类,最后利用 Google 开发的 PageRank 算法来计算每个超像素的显著性。

超像素是指具有相似纹理、颜色和亮度等特征的相邻像素构成的图像块。超像素分割算法主要利用像素之间特征的相似度对像素分组以获取图像的冗余信息,从而降低后续图像处理的复杂度。Achanta 等^[10]提出一种简单线性迭代聚类 (SLIC)算法,该算法是在 LAB 空间使用聚类,由 LAB 颜色空间的亮度分量和两个颜色分量以及像素坐标来计算。

2.2 基于连续分布的区域相似性度量

本文在进行相似性度量之前,先对滤波后图像进行 SLIC 超像素分割的预处理。图像预处理之后,分割成很多在颜色空间上具有相似特征的区域,我们假设每一个超像素内部的像素集合在颜色空间内符合多维正态分布,那么可以通过多维正态分布来挤去超像素的特征。那么第 i 个超像素集合可表示如下:

$$S_i = \left\{ N_s(\mu, \Sigma), C_s = \begin{pmatrix} x \\ y \end{pmatrix} \right\}_i, \quad (7)$$

其中: N_{s_i} 表示均值为 μ , 协方差为 Σ 的多维正态分布, C_{s_i} 表示该超像素的中心位置, x, y 是超像素的坐标。

有许多用于区域对比的检测技术均被用来辨

识局部对比。最近, Klein 和 Frintrop^[11]统计了以每个像素为中心的两个不同尺度的高斯加权矩形区域内像素在颜色空间内的分布,并利用二范数约束下的 Wasserstein 距离来计算该像素中心-邻域差异。在实数空间中,欧式范数约束下的 Wasserstein 距离定义如下:

$$w_2(\chi, \nu) = \left(\inf_{\gamma \in \Gamma(\chi, \nu)} \int_{\mathbf{R}^n \times \mathbf{R}^n} \|X - Y\|^2 d\gamma(X, Y) \right)^{\frac{1}{2}}, \quad (8)$$

其中: χ 和 μ 是在测度空间 $(\mathbf{R}^n, L^2)$ 上的概率测度, $\Gamma(\chi, \mu)$ 表示与 χ 和 μ 有关的在 $\mathbf{R}^n \times \mathbf{R}^n$ 上的所有测度。

简单的说, Wasserstein 距离计算了从一个分布变换到另外一个分布的最小代价,它不仅累计了在两个分布中对应每个点在测度空间中的个体变化,同时也考虑从一个分布到另一个分布的距离。在机器视觉中,离散化的 W_1 距离通常被称为推土机距离,该距离也经常用于比较直方图。

一般地,对于任意连续分布,二范数 Wasserstein 距离 (W_2 距离) 没有解析的计算方法。幸运的是,针对于两个正态分布,文献[12]给出了其 W_2 距离的解析解。对于任意两个经过 SLIC 超像素分割的区域 R_i 和 R_j 均服从多维正态分布,那么这两个区域的 W_2 距离可以由公式(9)得到:

$$W_2(R_i, R_j) = \sqrt{\|\mu_i - \mu_j\|^2 + \text{tr}(\Sigma_i) + \text{tr}(\Sigma_j) - 2\text{tr}(\sqrt{\Sigma_i \Sigma_j})}. \quad (9)$$

其中: μ_i 和 μ_j 分别是区域 R_i 和 R_j 的均值, Σ_i 和 Σ_j 分别是区域 R_i 和 R_j 的协方差矩阵, $\text{tr}(\Sigma_i)$ 和 $\text{tr}(\Sigma_j)$ 是协方差矩阵的迹。

2.3 基于中央一周边差的局部显著性检测

传统的中心-邻域差异性计算首先将粗尺度插值转换到相应的细尺度图像,然后细尺度逐个像素进行减法运算。计算公式如(10)所示。

$$I(c, s) = |I(c) \ominus I(s)|, \quad (10)$$

其中: I 为亮度通道的尺度图,用来产生一个高斯金字塔 $I(\sigma)$, $\sigma \in [0 \cdots 8]$, 表示高斯金字塔尺度值,中心 $c \in \{2, 3, 4\}$, $s \in \{4, 5, 6\}$, $\ominus$ 是差分算子。

传统的中心-邻域差异计算没有考虑像素的空间相似性,易降低算法处理率,因此本文先对图像进行 SLIC 超像素分割,然后提取超像素特征,

在不进行多尺度插值计算下采用上文提到的 W_2 距离进行局部显著性检测。

2.4 基于背景先验的全局显著性检测

根据背景先验知识,背景往往集中在图像边缘,那么可以通过提取边界区域特征进行显著性检测得到全局显著图。本文整体考虑所有边缘在特征空间的相关性,根据边缘在特征空间中的潜在流形分布结构,采用边界区域的流形排序方法^[13-15],对边界细节进行排序。

在 SLIC 超像素分割的基础上,采用闭环图模型来模拟边界数据的流形结构^[16-17]。以超像素为节点,计算包空间的近邻点,连接超像素和近邻点作为加权图的边。

假定超像素集合由 $X = \{x_1, x_2 \cdots x_n\} \in \mathbf{R}^{m \times n}$

构成, m 为特征空间的维数, n 为超像素的个数。由集合中的 n 个节点构建图, 节点之间的边权值矩阵为:

$$w_{ij} = \exp(-d(x_i, x_j)^2 / (2\sigma)^2), \quad (11)$$

其中: σ 为给定的参数, $d(x_i, x_j)$ 为超像素 x_i, x_j 在特征空间的距离, 由前文提到的 W_2 距离计算得到, 为了防止自相似性, 设 $w_{ii} = 0$, 据 $S = D^{-1/2}WD^{-1/2}$ 归一化 W 。其中 D 为对角阵, 满足:

$D = \text{diag}\{d_{11}, \dots, d_{nn}\}, d_{ii} = \sum_{j=1}^n w_{ij}$ 。

定义 f 为向量 $f = \{f_1, f_2, \dots, f_n\}$, 其中 f_i 为像素 x_i 的排序值, $i \in \{1, 2, \dots, n\}$ 。同时定义向量 y , 满足 $y = \{y_1, y_2, \dots, y_n\}$, 如果 x_i 是查询元素, 那么 $y_i = 1$, 反之为 0。

定义 f^* 为序列 $\{f_i(t)\}$ 的极限, 假定 $f(t)$ 收敛于 f^* , 文献[13]证明区域最优查询排序结果可由公式(12)得到:

$$f^* = \operatorname{argmin} \frac{1}{2} \left(\sum_{i,j=1}^n w_{ij} \left\| \frac{f_i}{\sqrt{d_{ii}}} - \frac{f_j}{\sqrt{d_{jj}}} \right\|^2 + \alpha \sum_{i=1}^n \|f_i - y_i\|^2 \right). \quad (12)$$

在衡量超像素的相似性中, 每一个比较后的像素作为下一个查询元素, 依次传递下去, 所得到的排序评分记为第一个查询元素的评分大小, 那么在 $t+1$ 时刻, 边界区域的排序评分为:

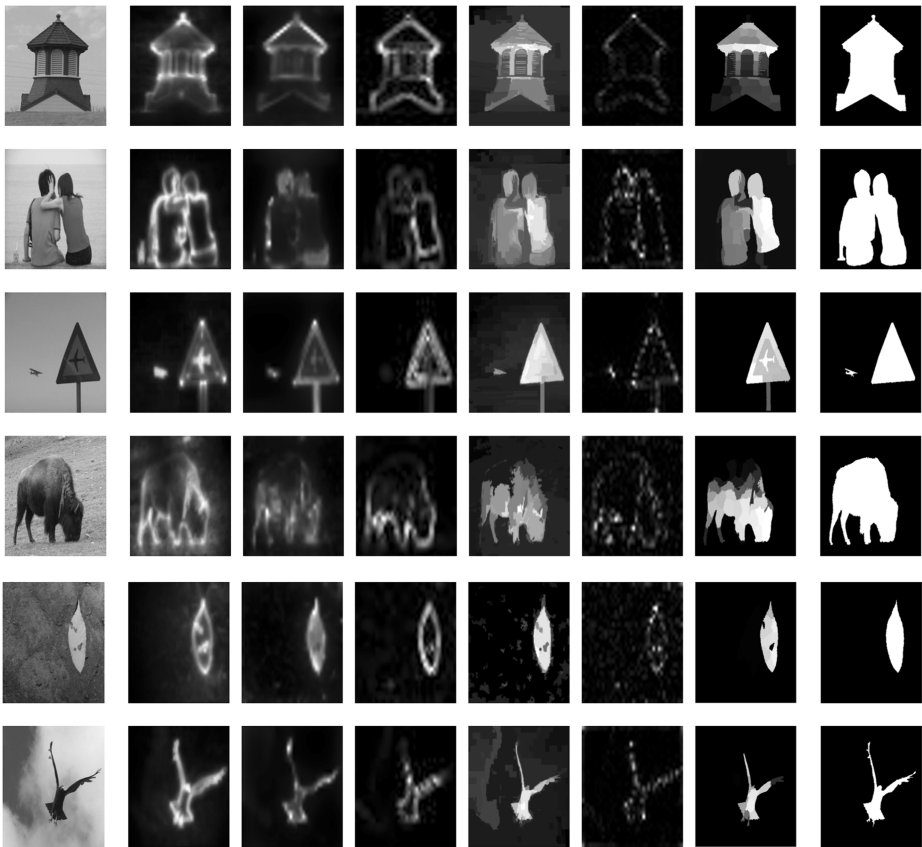
$$f(t+1) = \alpha S f(t) + (1-\alpha)y, \quad (13)$$

其中: α 为参数并且满足 $\alpha \in [0, 1)$, 它定义了

在排序评分传递过程中, 某区域所获得的排序评分的来源。定义 f^* 为序列 $\{f_i(t)\}$ 的极限, 假定 $f(t)$ 收敛于 f^* , 那么对于公式(13)可有:

$$f^* = \alpha S f^* + (1-\alpha)y. \quad (14)$$

整理后可以得到:

$$f^* = (1-\alpha)(1-\alpha S)^{-1}y. \quad (15)$$


(a)原图 (b)CA (c)GB (d)MZ (e)RC (f)SR (g)本文算法 (h)GT

图 1 显著性结果对比图
Fig. 1 Significant results of comparison chart

3 实验结果分析

本文通过双边滤波和 SLIC 超像素分割对图像进行预处理,应用 W_2 距离度量特征相似度,结合背景先验知识进行显著性检测。为了验证本文算法的有效性,我们采用 MSRA 图像库应用 MATLABR2010a 实现算法有效性,并将其与经典的 5 种显著性算法 CA^[4]、GB^[2]、MZ^[12]、RC^[18]、SR^[1] 进行对比,最后将由我们的算法得到的显著性图对比 ground truth 数据库中图片得到精准-查全率曲线(precision-recall curves)。

MSRA 显著性目标图像库中包含 1 000 幅可以用来准确提取 ground truth 的二进制结果的图片。这些图片每幅均含有一个显著性目标且经由众多研究者认同才被收纳到 MASR 中的。Ground truth 是由许多学者通过提取那些他们觉得是显著性的图像边框得到的标准图片集。据我们所知, MASR 是被认可的具有精准 ground truth 图片集中最大的图像集。实验对比结果如图 1 所示。

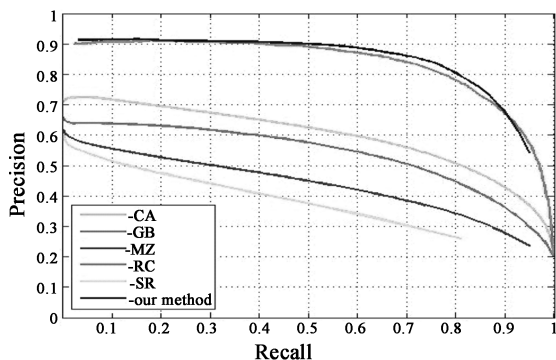


图 2 精准-查全率曲线

Fig. 2 Precision-recall curves

由图 1 可以看到本文算法得到的显著性图最接近 ground truth 的结果。本文算法有效地抑制了目标内部纹理,比其他算法更好地保留了目标与背景之间的边界,更加突出了显著性目标在背景中的整体轮廓。

将本文提出的显著性检测方法和 CA、GB、MZ、RC、SR 算法分别对 MSRA 数据库 1 000 幅图片进行显著图提取,再对比显著图和 ground truth 二值图像计算精度和查全率,得到图 2 精度-查全率曲线。

精度-查全率曲线提取显著性目标最简单的方法就是在一定的灰度值范围内根据固定阈值分割显著图。我们将所有的显著图归一化到 $[0, 255]$ 上,并采用 256 个变化阈值对其进行分割。对于每一幅图像,通过比较计算该图的显著图和数据库的 ground truth 中二值图像来得到精度和查全率。最后,我们将数据库中所有图片的精度和查全率平均化来得到选择基准的全部评估。

假设显著图 I 经固定阈值分割得到的图形中共有 S 个区域,其中有 M 个是前景区域,也就是显著区域。Ground truth 中 I 的前景区域个数为 N ,定义 true positive 为两者的交集,表示为:

$$T_p = M \cap N. \quad (16)$$

那么精度 $P(\text{presition})$ 可表示为:

$$P = \frac{T_p}{M}. \quad (17)$$

查全率 $R(\text{recall})$ 可表示为:

$$R = \frac{T_p}{N}. \quad (18)$$

从图 2 也可以看出,本文算法的精度和查全率明显高于 CA、GB、MZ、SR 4 种方法,说明本文算法相对于这 4 种算法检测效果更好。

当查全率为 $0.1 \sim 0.5$ 时,本文算法的精度-查全率曲线和 RC 算法有部分是重合的,但是从曲线上可以看到,对应的精度达到 0.9,这说明本文和 RC 算法都能够取得不错的检测效果。当查全率为 $0.5 \sim 0.9$ 时,本文算法的精度略高于 RC 算法,表明本文算法可以获得更精确的显著图。

4 结 论

提出了一种应用双边滤波抑制图像内部纹理,通过多维正态分布提取区域特征,结合中央-周边差异性和背景先验知识,应用二范数约束下的 Wasserstein 距离进行显著性检测的方法。首先对图像进行双边滤波,以平滑目标以及背景区域内部纹理的同时较好地保留图像边界。显著图提取建立在 SLIC 超像素的分割基础上,通过多维正态分布提取区域特征,应用 W_2 距离对超像素进行相似性度量,分别得到局部显著图和全局显著图,最后两者像素相乘得到图像显著图。实验结果表明采用本文的方法获得了较好的检测效果。

参 考 文 献:

- [1] Hou X, Zhang L. Saliency detection: A spectral residual approach [C]//*Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2007: 1-8.
- [2] Hare J, Koch C, Perona P. Graph-based visual saliency [C]. *The 20th Annual Conference on Neural Information Processing Systems*, 2006: 545-552.
- [3] Achanta R, Hemami S, Estrada F, et al. Frequency-tuned salient region detection [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2009: 1597-1604.
- [4] Goferman S, Zelnik-Manor L, Tal A. Context-aware saliency detection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(10): 1915-1926.
- [5] 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.
- [6] 董宇星, 刘伟宁. 基于灰度特性的海天背景小目标检测[J]. *中国光学与应用光学*, 2010, 15(3): 252-256.
Dong Y X, Liu W N. Small target detection based on the characteristics of the gray sky background [J]. *Chinese Optics and Applied Optics*, 2010, 15 (3): 252-256. (in Chinese)
- [7] 宋建中. 图像处理智能化的发展趋势[J]. *中国光学*, 2011, 4(5): 431-440.
Song J Z. Development trend of intelligent image processing [J]. *Chinese Optics*, 2011, 4 (5): 431-440. (in Chinese)
- [8] Zhang X N, Liu G Q, Hu Y, et al. A new two-sided coupling channel drop filter based on a two-dimensional photonic crystal [C]. *The 2nd International Conference on Opto-Electronics Engineering and Materials Research*, 2013: 417-420.
- [9] Ren Z, Hu Y, Chia L T, et al. Improved saliency detection based on superpixel clustering and saliency propagation [C]. *The 18th Annual ACM International Conference on Multimedia*, 2010.
- [10] Achanta R, Shaji A, Smith K, et al. SLIC superpixels compared to state-of-the-art superpixel methods [C]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012: 2274-2281.
- [11] Klein D A, Frintrop S. Salient pattern detection using w_2 on multivariate normal distributions [C]//*The 8th Symposium on Multispectral Image Processing and Pattern Recognition*, 2012: 246-255.
- [12] Givens C R, Shortt R M. A class of wasserstein metrics for probability distributions [J]. *Michigan Math.*, 1984, 31(2): 33-40.
- [13] Wei Y, Wen F, Zhu W, et al. Geodesic saliency using background priors [C]. *The 12th European Conference on Computer Vision*, 2012: 29-42.
- [14] Wan X, Yang J, Xiao J. Manifold-ranking based topic-focused multi-document summarization [C]. *The 20th International Joint Conference on Artificial Intelligence*, 2007: 2903-2908.
- [15] Belkin M, Niyogi P. Towards a theoretical foundation for laplacian based manifold methods [C]. *The 18th Annual Conference on Learning Theory*, 2007: 486-500.
- [16] Yang C, Zhang L, Lu H, et al. Saliency detection via graph-based manifold ranking [C]//*Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2013: 3166-3173.
- [17] Cheng M M, Zhang G X, Mitra N J, et al. Global contrast based salient region detection [C]//*Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2011: 409-416.
- [18] Ma Y F, Zhang H J. Contrast-based image attention analysis by using fuzzy growing [C]. *The 11th Annual ACM International Conference on Multimedia*, 2003: 374-381.

作者简介: 邓丹(1988—), 女, 湖北仙桃人, 硕士研究生, 从事图像处理与计算机视觉研究工作。E-mail: 527693165@qq.com

吴谨(1967—), 女, 安徽芜湖人, 博士生导师, 从事图像处理与模式识别研究工作。

朱磊(1982—), 男, 湖北武汉人, 博士生, 从事图像处理与机器学习研究工作。

刘劲(1982—), 男, 湖南娄底人, 讲师, 从事航空航天研究工作。