

基于条件随机场和图像分割的显著性检测

钱生¹ 陈宗海¹ 林名强¹ 张陈斌¹

摘要 针对当前常见的显著性方法检测得到的显著性区域边界稀疏不明确、内部不均匀致密等问题,提出了一种基于条件随机场(Condition random field, CRF)和图像分割的显著性检测方法.该方法综合利用边界信息、局部信息以及全局信息,从图像中提取出多种显著性特征;在条件随机场框架下融合这些特征,通过显著性区域与背景区域的区域标注实现显著性区域的粗糙检测;结合区域标注结果和交互式图像分割方法实现显著性区域的精确检测.实验结果表明本文提出的方法能够清晰而准确地提取出图像中的显著性区域,有效提高显著性检测精度.

关键词 显著性检测,多特征融合,条件随机场,图像分割

引用格式 钱生,陈宗海,林名强,张陈斌.基于条件随机场和图像分割的显著性检测.自动化学报,2015,41(4):711-724

DOI 10.16383/j.aas.2015.c140328

Saliency Detection Based on Conditional Random Field and Image Segmentation

QIAN Sheng¹ CHEN Zong-Hai¹ LIN Ming-Qiang¹ ZHANG Chen-Bin¹

Abstract The problem of sparse and unclear boundary with uneven and non-compact interior existed in the saliency region detected by most saliency detection methods. In order to solve this problem, this paper proposes a saliency detection method based on the conditional random field (CRF) and image segmentation. This method comprehensively utilizes boundary information, local information and global information to extract a variety of salient features from an image. By fusing these features into the framework of conditional random field, a coarse detection for saliency region is realized based on region labeled of saliency region and background region, and then a fine detection for saliency region is achieved through combining the result of region labeled with an interactive image segmentation method. Experimental results show that the proposed approach can clearly and accurately extract saliency regions and improve the detection precision.

Key words Saliency detection, multi-feature fusion, conditional random field (CRF), image segmentation

Citation Qian Sheng, Chen Zong-Hai, Lin Ming-Qiang, Zhang Chen-Bin. Saliency detection based on conditional random field and image segmentation. *Acta Automatica Sinica*, 2015, 41(4): 711-724

在数字图像处理的诸多领域内,显著区域的检测有着广泛的应用,如基于内容的图像检索^[1-2]、图像自动剪裁^[3-4]、图像视频压缩^[5-6]、目标检测与识别^[7]等.视觉显著性与人类如何感知、处理视觉刺激紧密相关,并且正在被认知心理学^[8-9]、神经生物学^[10-11]等多个学科进行研究. Treisman等^[12]、Koch等^[13]的早期工作,以及Itti等^[14]的视觉注意理论研究表明,视觉注意机制可以归结为两个阶段:基于数据驱动的自底向上的视觉注意;基于目标驱动的自顶向下的视觉注意.显著区域检测的本质是一种依据视觉注意机制而建立的视觉注意模型.它利用视觉注意机制得到图像中最容易引起注

意的显著部分,并用一幅灰度图像表示其显著度.基于人类的视觉注意机制,研究人员将视觉注意模型分为两种类型:自底向上的注意模型和自顶向下的注意模型.其中,自底向上的注意模型^[14-15]分为两个主要的步骤:底层视觉特征的提取和综合提取出的特征生成显著图.自顶向下的注意模型^[16-17]主要是通过调整选择准则,以适应外界需求来获得显著图.它涵盖了自底向上注意模型的内容,由于高层知识的参与,比自底向上的模型更加复杂.在显著性检测中,自底向上的模型是比较常见的,本文主要关注与其相关的文献.这类方法或是基于生物学原理的,或是纯数学计算的,又或是两者皆融合,主要利用颜色、边缘、纹理等底层特征属性来度量图像区域与其周围图像区域的差异性.根据是否对图像进行频域变换,本文把这些方法宽泛地分为以下两类.

1) 空间域模型

其中经典的Itti模型^[14],根据人眼视觉特性,采用高斯金字塔进行非均匀采样生成多个尺度图像,对这些尺度图像的颜色、亮度、方向等底层特征进行

收稿日期 2014-05-09 录用日期 2014-10-27
Manuscript received May 9, 2014; accepted October 27, 2014
国家自然科学基金(61375079)资助
Supported by National Natural Science Foundation of China (61375079)
本文责任编辑 章毓晋
Recommended by Associate Editor ZHANG Yu-Jin
1. 中国科学技术大学自动化系 合肥 230027
1. Department of Automation, University of Science and Technology of China, Hefei 230027

计算, 根据中央周边差计算各个尺度上的关注图, 最后对关注图进行融合得到显著图. 而文献 [18] 在分析了环境上下文、返回抑制等 5 种计算模型后提取显著区域. 但上述模型在提取图像显著特性的时候仅仅考虑了图像的局部特征, 不能全局地考虑图像显著特性, 因而检测出的显著性区域比较粗糙, 且含有较多的背景区域. Harel 等^[15] 提出了基于图论的算法, 这种算法在特征提取的过程中类似 Itti 算法模拟视觉原理, 在显著图生成的过程中引入马尔科夫链, 用纯数学计算得到显著值. Judd 等^[19] 将显著性区域检测问题建模为两类分类问题, 从数据集中提取底层、中层、高层三个层次的特征, 以此作为输入训练线性支持向量机 (Support vector machine, SVM), 通过已训练的 SVM 实现显著性区域的检测. Liu 等^[20] 将显著性区域检测问题建模为两类标注问题, 从局部、区域和全局这三个层次分别提出了多尺度对比度、中央周边直方图和颜色空间分布这三种特征来描述图像的显著区域, 并在条件随机场框架下对基于上述特征所获得的显著图进行融合. Borji 等^[21] 提出了区域局部孤立性和区域全局孤立性两种特征, 并分别计算 RGB 空间和 Lab 空间中每个通道的孤立性特征. 与大部分直接使用颜色直方图的方式不同, 该方法利用从数据集中学习到的一个自然图像字典对图像块进行稀疏编码, 通过对图像块的稀疏编码系数进行特征分析, 按一种竞争互补的方式对多通道、局部、全局显著图进行融合. Jiang 等^[22] 基于图的过分割方法将图像进行多层次分割, 并对分割区域进行多特征描述, 利用随机森林回归的方法学习从特征描述到显著度的映射, 基于最小二乘的方法进行多层次上的显著性融合. 这几类基于统计学习模型的显著性检测方法一般都能较好地检测出显著性区域, 但是显著性区域的边界不够清晰明确. Shen 等^[23] 认为图像是由低秩结构和稀疏噪声构成的, 并使用低秩矩阵表示背景区域, 稀疏矩阵表示显著性区域. 该方法基于颜色和纹理等底层特征, 并结合中心先验、语义先验、颜色先验等高层特征, 利用低秩稀疏分解技术进行显著性区域检测. 当显著性区域在图像中占据较大区域时, 由于低秩稀疏的局限性, 该方法就不能将其检测出. Yan 等^[24] 考虑到小尺度、高对比的模式 (高纹理区域) 会干扰显著性检测, 提出了一种多尺度的分析方法, 通过不同大小的过分割图像块, 构成树形结构, 利用层次推断的方法得到最终的显著图. Yang 等^[25] 主要利用图像边界大多是背景的先验信息, 通过流形排序的方法将此先验扩散并增加, 得到前景的可靠的估计; 随后将前景的信息用类似的流形排序的方法加强, 从而得到最终的显著图. Li 等^[26] 充分

利用图像边界大多是背景的先验信息, 通过将边界区域作为背景模板进行稀疏分析和主成分分析, 以重构误差获得两幅初始显著图, 在贝叶斯框架下以一种交互方式将两者进行显著性融合. 由于对边界区域信息的依赖程度较大, 当较大面积的显著性区域位于边界区域时, 以上两类基于边界区域信息的方法就不能完整地检测出显著性区域.

2) 频域模型

基于频域的显著性模型是近年来逐渐兴起的显著区域检测模型, 其主要思想是首先把图像从空间域变换为频域, 然后通过分析频域信息, 找到其与显著性特征的关系. 代表方法有 Hu 等^[27] 提出的特征值极坐标变换方法, 此方法将图像从空间域变换为 $[0, 2\pi]$ 的线性子空间坐标内, 并对子空间内呈现线性的特征值进行分析并抽取, 分析该子空间内分布的线性部分, 进而找到显著性区域. Hou 等^[28] 提出了谱残差方法, 该方法认为图像中包含的信息可以分为显著性信息和冗余信息两部分, 通过在图像幅度谱上作对数运算, 利用平均滤波器与幅度谱进行卷积得到冗余信息, 以图像幅度谱与其卷积的差值来表征显著性信息, 并将其反变换到空间域上以获得显著性区域. 基于谱残差方法的启发, Guo 等^[3, 29] 提出了四元相位谱方法, 认为图像的显著性信息隐藏在图像的傅里叶频谱的相位谱中, 因而通过对图像的多特征分量进行相位谱提取可以得到显著区域, 相位谱方法速度更快, 但是其缺点是无法找到合适的心理物理学解释. 张巧荣等^[30] 在多尺度上提取图像的底层特征的基础上, 在频域范围内分析了特征图的频谱信息, 在空间域上计算显著图, 以此检测出图像中的显著性区域. 基于频域的方法处理速度较快, 其检测出的显著区域边界比较明确. 然而, 由于没有考虑局部区域的特征, 这使得该类方法不能获得内部均匀致密的显著性区域.

为了获得内部均匀致密且边界清晰明确的显著性区域, 为后序对目标的跟踪、识别、行为分析等高层应用提供更为精确的结果, 本文提出了一种基于条件随机场 (Condition random field, CRF) 和图像分割的显著性检测方法. 该方法综合利用边界信息、局部信息以及全局信息, 提取出稀疏重构误差、空域调频对比度以及颜色空间分布这三种显著性特征; 在条件随机场 CRF^[31] 框架下实现多种显著性特征的融合, 将显著性区域的粗糙检测建模为显著性区域与背景区域这两类区域的标注问题; 基于图像分割方法 GrabCut^[32], 结合图像区域标注结果实现显著性区域的精细检测. 本文的内容组织如下: 第 1 节主要介绍了基于条件随机场和图像分割的显著性区域检测模型; 第 2 节详细阐述了显著性特征

的相关概念及提取过程; 第 3 节重点介绍了实验设计和实验结果分析; 第 4 节对文章进行总结.

1 显著性检测方法

本文将显著性区域检测问题归结为两类区域的标注与分割问题, 提出了一种基于条件随机场和图像分割的显著性区域检测的方法. 首先, 在条件随机场 CRF 框架下, 通过进行图像标注将显著性区域与背景区域区分开来, 形成粗糙的初始显著性区域; 然后, 以初始显著性区域为输入, 迭代应用图像分割方法 GrabCut 来改善显著性区域的检测, 从而获得更为精确的显著性区域. 图 1 显示了显著性区域检测模型的框图. 在条件随机场 CRF 框架下, 对于尺寸为 $W \cdot H$ 的输入图像 $I(x, y)$, 图像标注 $A(x, y)$ 的概率分布可以建模为条件概率分布, 可表示为

$$p(A|I) = \frac{\exp(-E(A|I))}{Z} \quad (1)$$

其中, $E(A|I)$ 是能量函数, Z 是配分函数. 为了检测显著性区域, 需要定义合适的能量函数, 本文将定义为多种显著性特征函数和成对特征函数的线性组合, 可表示为

$$E(A|I) = \sum_{x=1}^W \sum_{y=1}^H (\lambda_0 + \sum_{m=1}^M \lambda_m \cdot Map_m(x, y) + \mu_0 + \sum_{m=1}^M \mu_m \cdot Pair_m(x, y)) \quad (2)$$

其中, 显著性特征函数 $Map_m(x, y)$ 可表示为

$$Map_m(x, y) = A(x, y) \cdot F_m(x, y) \quad (3)$$

成对特征函数 $Pair_m(x, y)$ 可表示为

$$Pair_m(x, y) = \sum_{x'=x-1}^{x+1} \sum_{y'=y-1}^{y+1} \|F_m(x, y) - F_m(x', y')\| \cdot (A(x, y)A(x', y')D((x, y), (x', y'))) \quad (4)$$

其中, $F_m(x, y)$ 为第 m 种显著性特征; M 表示显著性特征的总数目; 参数 $p_{\lambda\mu} = (\lambda_0, \lambda_1, \dots, \lambda_M, \mu_0, \mu_1, \dots, \mu_M)$ 是待学习的权重参数, 参数的具体学习方法在文献 [33] 中有详细介绍. 对于位于 (x, y) 处的像素 $p(x, y)$, 标注 $A(x, y) = +1$ 表示其属于显著性区域, $A(x, y) = -1$ 则表示其属于背景区域. $D((x, y), (x', y'))$ 表示颜色差异函数, 可表示为

$$D((x, y), (x', y')) = \exp\left(-\frac{\|(I(x, y) - I(x', y'))\|}{\beta}\right) \quad (5)$$

其中, $\beta = (\sum_{x=1}^W \sum_{y=1}^H (\sum_{x'=x-1}^{x+1} \sum_{y'=y-1}^{y+1} \|I(x, y) - I(x', y')\|)) / (W \cdot H)$ 表示整幅图像的 8-邻域梯度均值. 当邻域像素 $p(x', y')$ 与 $p(x, y)$ 颜色相似而标注不同时, $D(\cdot, \cdot)$ 作为惩罚因子越接近于 1. 基于 CRF 模型推断出的初始显著性区域一般比较粗糙, 其中仍可能包含一些误检测的显著性区域和背景区域. 为了提高显著性区域检测的精度, 本文迭代应用 GrabCut 来优化显著性区域的检测. GrabCut 本身使用高斯混合模型和 GrapCut^[32] 进行迭代, 从而改善每一步区域的分割效果, 靠近初始显著性区域的部分成为显著性区域的几率更大. 不同于原始的 GrabCut 是由人工选中矩形区域来进行初始化操作, 本文采用多边形逼近的方法, 使用一个包含初始显著性区域的最小多边形对 GrabCut 进行初始化, 可以使得初始化操作的区域更为精确. 借鉴文献 [34], 本文也进行多次迭代应用 GrabCut 来改善显著性分割结果, 即在每一次迭代后采用形态学操作所得的显著性分割结果作为下一次迭代的输入. 本文方法的创新点主要在于: 对图像区域显著性的多信息、多特征描述, CRF 图像标注理论和 CrabCut 图像分割技术的有机结合, 以及显著性区域由粗糙到精细的分层检测, 具体内容在相关章节中有详细介绍.

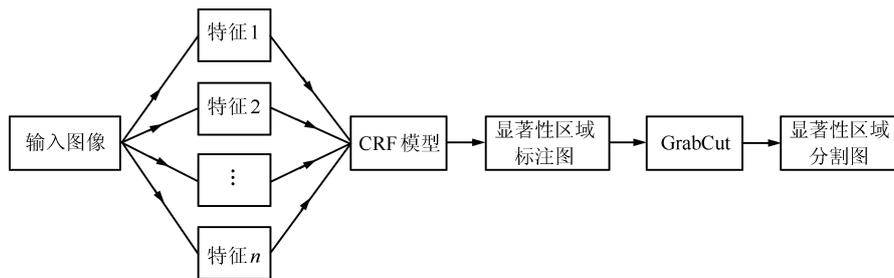


图 1 显著性检测方法框图

Fig. 1 Diagram of the saliency detection method

2 显著性特征

本节通过分析图像中边界信息、局部信息以及全局信息与显著性区域的关系,详细阐述了稀疏重构误差、空域调频对比度以及颜色空间分布这三种显著性特征的相关概念及提取过程,并分析了每种特征在显著性区域检测中的特点.

2.1 稀疏重构误差

一方面,相对于图像其他区域而言,边界区域可以作为图像背景模型中更好的视觉线索^[35-36];另一方面,显著性区域有很大的可能性出现在图像的中心区域^[21, 37].虽然这些假设不是总成立,但是边界区域仍为检测显著性区域提供了丰富有效的信息.

1) 背景模板

为了更好地捕获到图像中显著性区域的结构信息,首先采用文献 [38] 所提出的超像素分割方法 SLIC (Simple linear iterative clustering) 对输入图像 $I(x, y)$ 进行分割,获得组成该图像的超像素集合 $S = \{s_1, \dots, s_i, \dots, s_N\}$. 其中,第 i 个超像素 s_i 表示一块均匀致密的图像区域, N 为超像素的总数目.如文献 [22] 中所述,同时在 Lab 和 RGB 颜色空间中进行颜色特征处理可以有效提高显著性检测精度.本文进一步将空域信息融入特征中,以期获得更高精度的显著性检测结果.针对每一个超像素 s ,统计其像素的颜色均值和空间位置均值作为其特征 \mathbf{sf} ,可表示为

$$\mathbf{sf} = [\tilde{R}, \tilde{G}, \tilde{B}, \tilde{L}, \tilde{a}, \tilde{b}, \tilde{x}, \tilde{y}] \quad (6)$$

其中, \tilde{R} 、 \tilde{G} 、 \tilde{B} 、 \tilde{L} 、 \tilde{a} 、 \tilde{b} 、 \tilde{x} 、 \tilde{y} 分别表示特征中对应的颜色分量和空间位置分量.因而,整个图像的超像素特征集合 \mathbf{SF} 可表示为

$$\mathbf{SF} = [\mathbf{sf}_1, \dots, \mathbf{sf}_i, \dots, \mathbf{sf}_N] \in \mathbf{R}^{D \times N} \quad (7)$$

其中, \mathbf{sf}_i 表示 S 中超像素 s_i 的特征, D 为超像素特征的维度.根据超像素特征集合 \mathbf{SF} ,从中提取出

属于上、下、左、右 4 个边界区域的超像素特征集合 \mathbf{B}_j ,可表示为

$$\mathbf{B}_j = [b_1^j, \dots, b_i^j, \dots, b_{N_j}^j] \in \mathbf{R}^{D \times N} \quad (8)$$

其中, \mathbf{B}_j 表示第 j 个边界区域的超像素特征集合, b_i^j 表示 \mathbf{B}_j 中第 i 个超像素特征, N_j 表示 \mathbf{B}_j 中超像素特征的数目, $j \in \{1, 2, 3, 4\}$ 分别代表上、下、左、右 4 个边界.图 2 显示了某示例图像的超像素分割结果,以及从中提取的背景模板.其中,非背景模板区域以白色轮廓线区分.

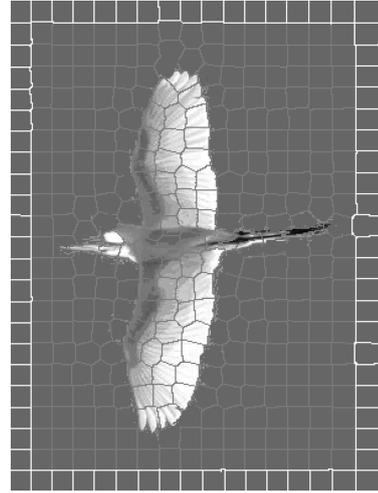


图 2 示例图像及其超像素分割结果
Fig. 2 A sample image and its result of super-pixel segmentation

不同于已有文献 [26] 使用边界区域的方式,本文通过组合不同的边界区域建立背景模板集合 $\mathbf{BTS} = \{\mathbf{BT}_1, \dots, \mathbf{BT}_k, \dots, \mathbf{BT}_{15}\}$,以期获得图像背景模型更多的视觉线索.表 1 给出了基于边界区域组合的背景模板集合.其中,编号 k 指示的列元素表示由 \mathbf{B}_j 组合成的一种背景模板 \mathbf{BT}_k , $b_j = 1$ 和 $b_j = 0$ 分别表示 $\mathbf{B}_j \in \mathbf{BT}_k$ 和 $\mathbf{B}_j \notin \mathbf{BT}_k$.

表 1 基于边界区域组合的背景模板集合

Table 1 A background template set based on the combination of boundary regions

编号	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
b_4	1	0	1	1	1	0	0	0	1	1	1	0	0	0	1
b_3	1	1	0	1	1	0	1	1	0	0	1	0	0	1	0
b_2	1	1	1	0	1	1	0	1	1	0	0	0	1	0	0
b_1	1	1	1	1	0	1	1	0	0	1	0	1	0	0	0
模板	\mathbf{BT}_1	\mathbf{BT}_2	\mathbf{BT}_3	\mathbf{BT}_4	\mathbf{BT}_5	\mathbf{BT}_6	\mathbf{BT}_7	\mathbf{BT}_8	\mathbf{BT}_9	\mathbf{BT}_{10}	\mathbf{BT}_{11}	\mathbf{BT}_{12}	\mathbf{BT}_{13}	\mathbf{BT}_{14}	\mathbf{BT}_{15}

2) 基于背景模板的稀疏分析

使用背景模板 \mathbf{BT}_k 作为稀疏分析的基向量, 对 s_i 进行重构, 可表示为

$$\alpha_i^k = \arg \min_{\alpha_i^k} \|\mathbf{sf}_i - \mathbf{BT}_k \cdot \alpha_i^k\|_2^2 + \lambda \|\alpha_i^k\|_1 \quad (9)$$

$$\varepsilon_i^k = \|\mathbf{sf}_i - \mathbf{BT}_k \cdot \alpha_i^k\|_2^2 \quad (10)$$

其中, α_i^k 和 ε_i^k 分别表示基于 \mathbf{BT}_k 对 s_i 进行稀疏重构时的重构系数和重构误差. 进一步对 ε_i^k 进行线性归一化, 使其位于 $[0, 1]$ 的范围内. ε_i^k 越大, 则表明 \mathbf{sf}_i 与 \mathbf{BT}_k 的相关性越小, 即 s_i 属于显著性区域的可能性越大. 为了获得全分辨率的显著图 (Sparse reconstruction error map, SREM) $SREM_k(x, y)$, 将 s_i 所包含像素的显著度设为 ε_i^k , 可表示为

$$SREM_k(x, y) = \varepsilon_i^k, \quad p(x, y) \in s_i \quad (11)$$

其中, $p(x, y) \in s_i$ 表示该像素 $p(x, y)$ 属于超像素 s_i .

3) 背景模板融合

由于显著性区域可能位于图像的中心区域或边界区域, 基于不同 \mathbf{BT}_k 获得的显著图 $SREM_k(x, y)$ 有很大的差异, 且它们与真实显著图的差异也不同. 因此, 如何以有效的方式融合不同的显著图, 这对提高显著性检测的精度有重要的影响. 本文提出使用 \mathbf{BT}_k 进行主成分分析, 并以重构误差作为权重, 按线性加权的方式对显著图进行融合. 使用 \mathbf{BT}_k 作为主成分分析的基向量, 对 s_i 进行重构, 可表示为

$$\beta_i^k = \mathbf{BT}_k^\top (\mathbf{sf}_i - \tilde{\mathbf{sf}}) \quad (12)$$

$$\eta_i^k = \|\mathbf{sf}_i - (\mathbf{BT}_k \cdot \beta_i^k + \tilde{\mathbf{sf}})\|_2^2 \quad (13)$$

其中, β_i^k 和 η_i^k 分别表示基于 \mathbf{BT}_k 对 s_i 进行主成分分析时的重构系数和重构误差, $\tilde{\mathbf{sf}}$ 是 \mathbf{sf}_i 的均值. 进一步对 η_i^k 进行线性归一化, 使其位于 $[0, 1]$ 的范围内. 根据 η_i^k 获得 $SREM_k(x, y)$ 的加权权重 ω_k , 可表示为

$$\omega_k = \sum_{i=1}^N (1 - (\eta_i^k)^2) \quad (14)$$

加权权重 ω_k 越大, 表明基于背景模板 \mathbf{BT}_k 进行主成分分析时, 背景模板 \mathbf{BT}_k 表征图像场景的能力越强, 因而基于背景模板进行稀疏分析时所得 $SREM_k(x, y)$ 的准确度越高. 通过背景模板的稀疏分析和线性加权融合, 定义图像的重构误差特征

$SREM(x, y)$, 可表示为

$$SREM(x, y) = \frac{\sum_{k=1}^{15} \omega_k \cdot SREM_k(x, y)}{\sum_{k=1}^{15} \omega_k} \quad (15)$$

尽管借鉴了 DSR^[26] 方法的思路, 但是本文方法与其在利用边界信息和所采用的技术上有以下区别:

1) 背景字典的构成: 在 DSR 方法中, 背景字典由图像四个边界区域的图像块集合组成; 而本文方法中背景字典由一组背景模板构成, 该组背景模板由四个边界区域的组合形式给出, 通过组合不同的边界区域, 以期获得图像背景模型更多的视觉线索. 2) 背景字典的分析方法: 尽管都应用了稀疏表示和主成分分析两种技术来分析背景字典, 但是对这两种技术的融合上仍存在以下区别. DSR 方法在第 1 节中有扼要叙述. 首先, 本文方法对单个背景模板分别应用稀疏表示和主成分分析技术对背景字典进行分析, 以稀疏重构误差反映图像中其他区域与边界区域之间的差异, 从而反过来衡量该区域不属于背景而是属于显著对象的可能性; 然后, 以主成分分析重构误差和来反映此背景模板表征整个图像的可能性, 即度量此背景模板作为真实背景区域的有效性; 最后, 按线性加权平均的方式对所有背景模板的分析结果进行融合, 并将最终结果作为原始图像的一种显著性特征描述.

2.2 空域调频对比度

在度量显著性的时候, 每个像素的显著度可以通过其与局部区域的对比度体现出来. 对比度越大, 对人眼的刺激就越强, 就越能引起人眼的注意. 不同大小的局部区域可以反映出不同的对比度. 通过分析每个像素与不同大小的局部区域的对比度, 可以获得不同的图像显著性信息, 从而更加准确地描述图像中的显著性区域. 这种对比度可以用颜色向量的欧氏距离或相关性表示. 文献 [39] 中详细分析了在空域内使用不同大小区域等价于利用了不同的通频带, 提出了一种空域调频的显著性检测方法. 在文献 [39] 的基础上, 文献 [40] 分析了待考察的像素与以其为中心的最大对称边界子区域的位置关系, 提出了一种基于最大对称局部区域的显著性检测方法. 本文基于韦伯定律^[41], 将上述两种空域内的对比度分析方法和多局部区域分析进行融合, 提出了一种空域调频对比度. 对于像素, 其颜色特征可表示为

$$p\mathbf{f}(x, y) = [R, G, B, L, a, b] \quad (16)$$

定义图像的空域调频对比度特征 (Spatial frequency-tuned contrast map, SFCM)

$SFCM(x, y)$ 如下:

$$SFCM(x, y) = \sum_R \frac{\|\mathbf{pf}(x, y) - \mathbf{pf}_R(x, y)\|}{M \cdot \|\mathbf{pf}_R(x, y)\|} \quad (17)$$

其中, M 表示要进行对比的周边区域总数目, $\mathbf{pf}_R(x, y)$ 表示以像素 $p(x, y)$ 为中心的局部区域 R 内所有像素的特征均值, 可表示为

$$\mathbf{pf}_R(x, y) = \sum_{i=x-x_\delta}^{x+x_\delta} \sum_{j=y-y_\delta}^{y+y_\delta} \frac{\mathbf{pf}(i, j)}{A_R} \quad (18)$$

其中, x_δ 和 y_δ 表示区域 R 的半宽高, A_R 为该区域的面积. 局部区域 R 可由 x_δ 、 y_δ 和 A_R 唯一确定. 本文要进行对比的局部区域包括以下区域: 整个图像区域、1/2 图像区域、1/3 图像区域、1/4 图像区域以及 1/5 图像区域, 可表示为

$$R_m = [x_\delta, y_\delta, A_R] = \left[\frac{W}{2\sqrt{m}}, \frac{H}{2\sqrt{m}}, \frac{WH}{m} \right] \quad (19)$$

其中, $m = 1 \sim M$, 文中 M 设为 5.

2.3 颜色的空间分布

如果一种颜色在图像中的分布越广, 那么显著性区域包含这种颜色的可能性就越小, 如图 3 所示. 因此, 颜色的全局空间分布情况可以作为一种特征, 被用来描述一个区域的显著性, 相关论述在文献 [20] 也有涉及.



图 3 示例图片

Fig. 3 Sample images

考虑到实际图像中颜色数目在一定范围之类, 相关论述在文献 [34] 也有涉及. 针对如何有效确定图像中颜色数目的问题, 本文提出了一种自适应确定图像中所含颜色数目的方法. 首先, 对输入图像采用最小方差量化算法^[42]进行颜色量化处理, 使其包含的颜色数目不多于 64 种; 然后, 采用 Otsu 算法^[43]对量化图像的颜色直方图进行自适应阈值分割; 最后, 根据归一化的分割阈值和方差确定颜色的

数目. 其中方差与颜色数目的映射公式如下:

$$C = \begin{cases} 3, & 0.85 \leq \sigma_{ostu} \leq 1 \\ 4, & 0.8 \leq \sigma_{ostu} < 0.85 \\ 5, & 0.75 \leq \sigma_{ostu} < 0.8 \\ 6, & 0 \leq \sigma_{ostu} < 0.75 \end{cases} \quad (20)$$

其中, C 表示图像中所含颜色的数目, σ_{ostu} 表示由 Otsu 算法所得的分割方差. 为了描述每种颜色的空间分布, 一种可行的方法就是计算这种颜色的空间分布方差. 首先, 使用混合高斯模型对图像中的颜色分布进行建模^[20], 可表示为

$$\{\omega_c, \mu_c, \Sigma_c\}_{c=1}^C \quad (21)$$

其中, ω_c 、 μ_c 、 Σ_c 分别表示第 c 种颜色的权重、均值和方差, 并使用 EM 算法对它们进行参数优化. 对于像素 $p(x, y)$, 它的颜色概率分布可表示为

$$p(c|\mathbf{pf}(x, y)) = \frac{\omega_c \cdot \mathcal{N}(\mathbf{pf}(x, y)|\mu_c, \Sigma_c)}{\sum_{\omega_c} \omega_c \cdot \mathcal{N}(\mathbf{pf}(x, y)|\mu_c, \Sigma_c)} \quad (22)$$

因而, 第 c 种颜色的空间分布方差 $SV(c)$ 可表示为

$$SV(c) = \sum_{x=1}^W \sum_{y=1}^H p(c|\mathbf{pf}(x, y)) \cdot \|(x, y) - (\tilde{x}, \tilde{y})\|_2^2 \quad (23)$$

其中, (\tilde{x}, \tilde{y}) 为第 c 种颜色的空间分布重心, 可表示为

$$(\tilde{x}, \tilde{y}) = \left(\sum_{x=1}^W x \cdot p(c|\mathbf{pf}(x, y)), \sum_{y=1}^H y \cdot p(c|\mathbf{pf}(x, y)) \right) \quad (24)$$

进一步对 $SV(c)$ 进行线性归一化, 使其位于 $[0, 1]$ 的范围之内. 根据空间分布方差, 定义图像的颜色空间分布特征 (Color spatial distribution map, CSDM) $CSDM(x, y)$, 可表示为

$$CSDM(x, y) = \sum_{c=1}^C p(c|\mathbf{pf}(x, y)) \cdot (1 - SV(c)) \quad (25)$$

第 c 种颜色的空间分布方差 $SV(c)$ 越大, 则表明该颜色对 $CSDM(x, y)$ 的影响就越小, 显著性区域包含该种颜色的概率也就越小. 图像的颜色数目计算和颜色的空间分布特征如图 4 所示. 其中图 4(a) 是输入图像, 图 4(b) 是颜色量化图像, 图 4(c) 是颜色量化图像的颜色直方图, 图 4(d) 是颜色的空间分布

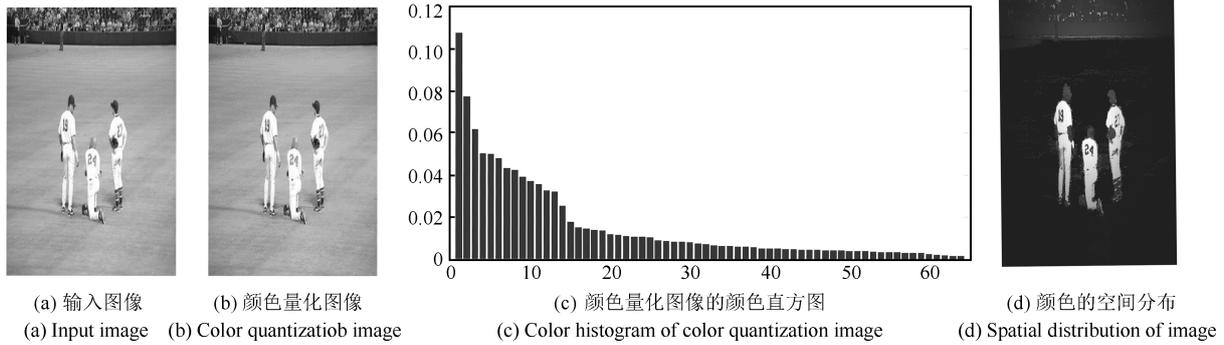


图4 颜色数目计算和颜色的空间分布特征提取示例

Fig. 4 An example of calculating the number of colors and color spatial-distribution

特征图. 在该示例中, 由 Otsu 算法计算所得的归一化自适应阈值和方差分别为 0.028、0.75, 根据分割方差与颜色数目的映射式 (20) 可得颜色数目 $C = 5$.

3 实验

本文以文献 [34] 中的数据集 THUS10000 作为测试数据集. 该数据集包含了 10000 张测试图片, 而且还提供了像素级的显著性区域标记图, 比文献 [39] 中提出的经典数据集更丰富. 在平均精度 (Precision)、召回率 (Recall)、以及 F 度量 (F -measure) 这三种评价指标下, 本文在该数据集上对所提方法以及其他显著性检测方法进行实验测试和结果分析. 其中, F 度量定义为

$$F_{\beta} = \frac{(1 + \beta^2) \cdot Precision \cdot Recall}{\beta^2 \cdot Precision + Recall} \quad (26)$$

为了强调在显著性检测中精度相对召回率的重要性, 同文献 [25, 39] 一致, 将 β^2 设为 0.3.

1) 特征融合对比实验

为了评估每种显著性特征的有效性, 本文随机从数据集中选择 1000 张图片建立训练数据集, 训练了如下 7 个 CRF 模型: 3 个只包含一种特征的 CRF 模型、3 个融合两种特征的 CRF 模型、1 个融合所有特征的 CRF 模型. 本文使用 CGSD (Saliency detection based on CRF and GrabCut) 简记所提出的显著性区域检测模型, 使用 sfc、csd、sre 分别标记空域调频对比度特征 (Spatial frequency-tuned contrast)、颜色空间分布特征 (Color spatial distribution)、稀疏重构误差特征 (Sparse reconstruction error), 使用 + 表示模型所使用的特征组合. 其中, CGSD + sfc、CGSD + csd、CGSD + sre、CGSD + sfc + csd、CGSD + sfc + sre、CGSD + csd + sre、CGSD + sfc + csd + sre 分别简记为 CGSD1 ~ CGSD7. 同时, 为了验证 CRF 模型在特征融合中的有效性, 本文还

按以下常用方法对多特征进行融合: 最大值、算术平均、几何平均、乘积, 并将对应的模型分别简记为 FMax、FAve、FGem、FMul. 对于按上述方式进行多特征融合的模型, 它们所给出的是区域的显著度, 在区间集合 $[0, 1]$ 中取值. 本文借鉴文献 [39] 中的方法, 对在区间集合 $[0, 1]$ 进行自适应阈值分割, 以获得对应的分类集合 $\{0, 1\}$, 以便进行对比实验. 在数据集上对以上模型进行对比实验, 表 2 和图 5 显示了这些模型所获得的性能. 本文还选取单张图片的实验结果进行示例说明, 如图 6 所示. 其中, 图 6(a) 是输入图片, 图 6(b)~6(d) 分别是 sfc、csd、sre 这 3 种显著性特征, 图 6(e)~6(k) 分别是模型 CGSD1~CGSD7 对应的显著性区域检测结果, 图 6(l)~6(o) 分别是模型 FMax、FAve、FGem、FMul 对应的显著性区域检测结果, 图 6(p) 是显著性区域的真实检测结果 GT. 对于基于单个特征的模型 CGSD + sfc、CGSD + csd、CGSD + sre, 从 Precision 指标看, 稀疏重构误差特征相对于其他两种特征在描述显著性时更为准确, 这表明边界区域能够为显著性区域的检测提供准确的信息. 从 Recall 指标看, 空域调频对比度特征和颜色空间分布特征相对于稀疏重构误差特征在描述显著性时更为全面, 这表明局部区域和全局区域能够为显著性区域的检测提供全面的信息. 从整体性能指标 F -measure 看, 颜色空间分布特征和稀疏重构误差特征更优于空域调频对比度特征. 对于基于两种特征的模型, 模型 CGSD + csd + sre 获得最优的 Precision 和 F -measure, 整体性能最好; 模型 CGSD + sfc + sre 获得次优的 Precision 和 F -measure, 整体性能中等; 模型 CGSD + sfc + csd 获得最差的 Precision 和 F -measure, 仅获得最优的 Recall, 整体性能最差. 这表明基于颜色空间分布特征和稀疏重构误差特征的组合能够有效提高显著性区域的检测精度和召回率. 在所有模型中, 模型 CGSD7 获得次优的 Precision、

次优的 Recall 和最优的 F -measure. 从整体性能指标 F -measure 看, 该模型整体性能最好. 这表明, 基于空域调频对比度特征、颜色空间分布特征、稀疏重构误差特征的特征组合能够有效描述显著性区域, 使得显著性区域检测模型能够获得最佳的整体性能. 在特征融合方法的对比中, 从 Precision 指标看, 仅模型 FMax 优于模型 CGSD7; 从 Recall 指标和整体性能指标 F -measure 看, 模型 CGSD7 明显优于模型 FMax、FAve、FGem、FMul. 实验结果

表明, 基于 CRF 的特征融合方法要优于常用的特征融合方法. 为进一步进行其他对比实验分析, 现将模型 CGSD + sfc + csd + sre 简记为 CGSD, 并给出模型 CGSD 训练所获得最优参数 $p_{\lambda\mu}^*$, 其值为 $p_{\lambda\mu}^* = (-0.1245, 0.8486, 0.5741, 0.7778, 0.8486, 0.5486, 0.3007, 0.5843)$.

2) 粗糙检测与精细检测的对比实验

本文方法针对显著性区域的检测主要包含粗糙检测和精确检测两步, 为了验证这种分步策略的有

表 2 模型性能统计表

Table 2 The statistical table of models' performance

模型	CGSD1	CGSD2	CGSD3	CGSD4	CGSD5	CGSD6	CGSD7	FMax	FAve	FGem	FMul
Precision	0.52	0.86	0.94	0.74	0.84	0.89	0.89	0.83	0.86	0.88	0.95
Recall	0.80	0.78	0.62	0.84	0.72	0.73	0.81	0.71	0.70	0.70	0.58
F -measure	0.57	0.84	0.84	0.76	0.81	0.84	0.87	0.80	0.82	0.83	0.83

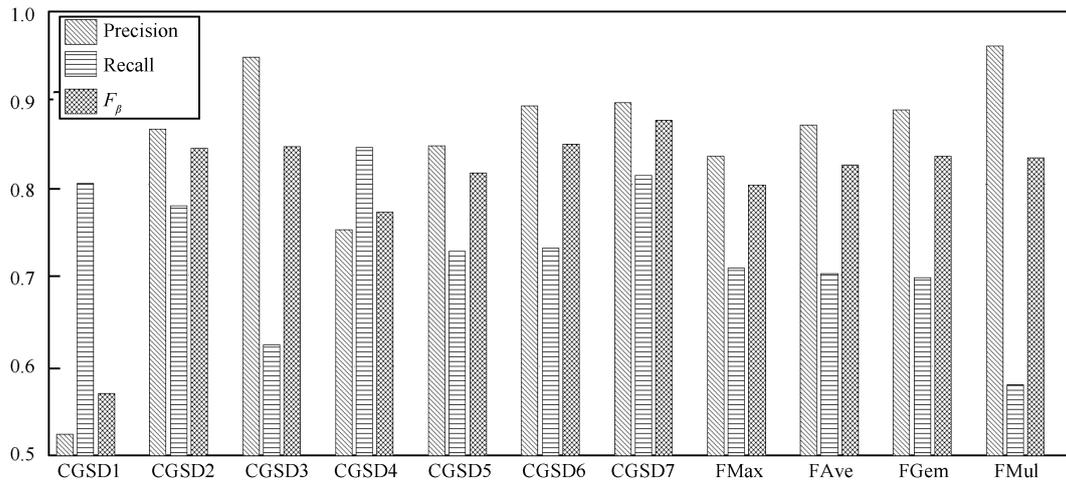


图 5 模型性能柱状图

Fig. 5 The bar graph of models' performance

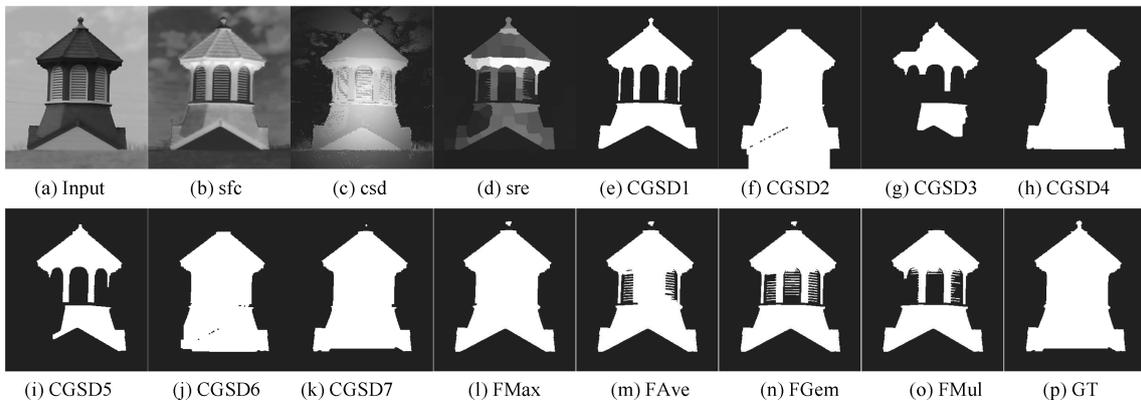


图 6 特征融合对比实验示例

Fig. 6 An example of the contrast experiment of multi-feature fusion

效性, 同时与已有的基于 CRF 的方法 LD^[20] 进行对比, 本文进行了粗糙检测与精细检测的分步对比实验. 在测试数据集上, 实验结果表明: 仅采用粗糙检测 CSD (Saliency detection based on CRF) 时, 其 Precision、Recall、 F -measure 分别为 0.86、0.79、0.84; 在粗糙检测的基础上, 进一步结合图割方法 GrabCut 进行精细检测 CGSD 时, 其 Precision、Recall、 F -measure 分别为 0.89、0.81、0.87. 相比而言, CGSD 方法比 CSD 在 Precision、Recall、 F -measure 上都有所提高, 分别约提高 3.5%、2.5%、3.6%. 与方法 LD 相比, 从使用的特征和显著性区域的检测过程的角度来看, 本文方法与其差异主要体现在以下方面: 对图像边界信息的利用; 由粗糙到精细的分步检测过程. 实验结果表明, 方法 LD 的 Precision、Recall、 F -measure 分别为 0.73、0.91、0.76. 在 Precision 和 F -measure 上, CSD、CGSD 显著优于 LD; 然而在 Recall 上, LD 显著优于 CSD、CGSD. 图 7 显示了这些模型所获得的性能. 部分示例的检测结果如图 8 所示. 其中, 图 8(a) 是原始图片, 图 8(b) 是 CSD 检测结果, 图 8(c) 是 CGSD 检测结果, 图 8(d) 是 LD 检测结果, 图 8(e) 是显著性区域的真实检测结果. 从图示对比可以看出, CGSD 方法比 CSD、LD 能更为清晰地检测出了均匀致密且边界清晰明确的显著性区域, 同时有效地剔除了一些误检测的背景区域.

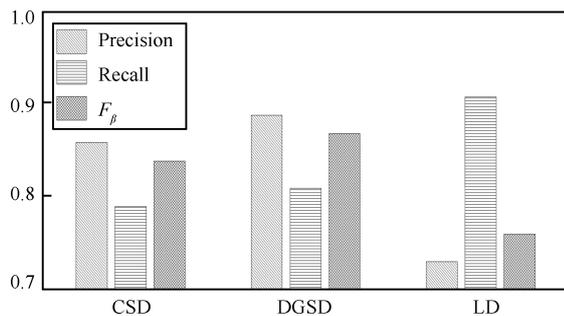


图 7 粗糙检测与精细检测的对比实验结果

Fig. 7 Results of the contrast experiment of coarse detection and fine detection

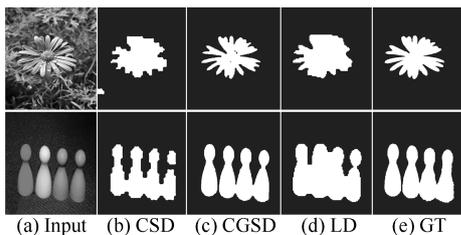


图 8 粗糙检测与精细检测的对比实验示例

Fig. 8 Examples of the contrast experiment of coarse detection and fine detection

3) 与其他算法的对比实验

本文在测试数据集上还与其他经典方法进行了比较. 按照文献 [39], 本文根据以下几个因素来选择进行比对的方法: 引用数 (IT^[14], GB^[15], SR^[28], FT^[39]), 较新的方法 (DSR^[26], RC^[34], CA^[44], CB^[45], SEG^[46]), 多种类 (IT 为生物驱动的, GB 为基于图模型的, SR 为基于频域处理的, DSR 为基于稀疏分析和主成分分析的, SEG 为基于条件随机场的). 由于本文方法将显著性区域的检测定义为一个二分类问题, 即分类显著性区域与背景区域, 给出的是区域的分类标记, 在分类集合 $\{0, 1\}$ 中取值. 其他对比方法给出的是区域的显著度, 在区间集合 $[0, 1]$ 中取值. 因而, 本文方法与其他方法在结果的表现形式上有一定的区别和差异. 为了比较本文方法与其他方法的性能, 本文主要借鉴文献 [39] 中的方法, 对区间集合 $[0, 1]$ 进行自适应阈值分割, 以获得其他方法对应的分类集合 $\{0, 1\}$. 与其他算法的对比实验结果如图 9 所示, 部分示例的检测结果如图 10 所示. 其中, 图 10(a) 是输入图片, 图 10(b)~图 10(j) 是算法 IT~DSR 的检测结果, 图 10(k) 是本文方法 CGSD 的检测结果, 图 10(l) 是显著性区域的真实检测结果. 从图 9 可以看出, 与算法 IT、GB、SR、FT、CA 相比, CGSD 在所有指标上至少都比这些算法高出 10%, 整体性能明显优于这些算法. 与算法 CGSD 相比, 算法 SEG 的 Precision、Recall、 F -measure 分别为 0.88、0.27、0.57, 在 Precision 上 CGSD 仅比其高 1%, 两者相差无几, 然而在 Recall、 F -measure 指标上 CGSD 明显优于 SEG; 算法 RC 的 Precision、Recall、 F -measure 分别为 0.90、0.77、0.87, 在 Precision 和 F -measure 上, CGSD 和 RC 相差无几, 然而在 Recall 上 CGSD 比 RC 高出 4%, 整体性能上 CGSD 要优于 RC; 算法 CB 和算法 DSR 的 Precision、Recall、 F -measure 分别为 0.90、0.67、0.83 和 0.88、0.79、0.85, 在 Precision 指标上, CGSD 和 CB、DSR 相差无几, 然而在 Recall 上 CGSD 比 CB 高出 14%, 在 F -measure 上 CGSD 比 DSR 高出 2%, 整体性能上 CGSD 要优于 CB 和 DSR. 由于融合了边界信息、局部信息和全局信息, CGSD 能更准确地提取显著区域, 更好地抑制非显著性区域. 这一优势不仅可直观地从图 10 中的显著性检测示例图上观测到, 而且从图 9 中所示的高精度和高召回率上也有较好的体现. 在图 10 中, 从区域的内部致密性来看, 相比其他算法, 算法 CGSD、RC、CB、CA、DSR 都取得了比较好的结果, 且 CGSD 所提取的显著性区域内部更加致密, 与真实的显著性区域最接近. 从区域的边界清晰度

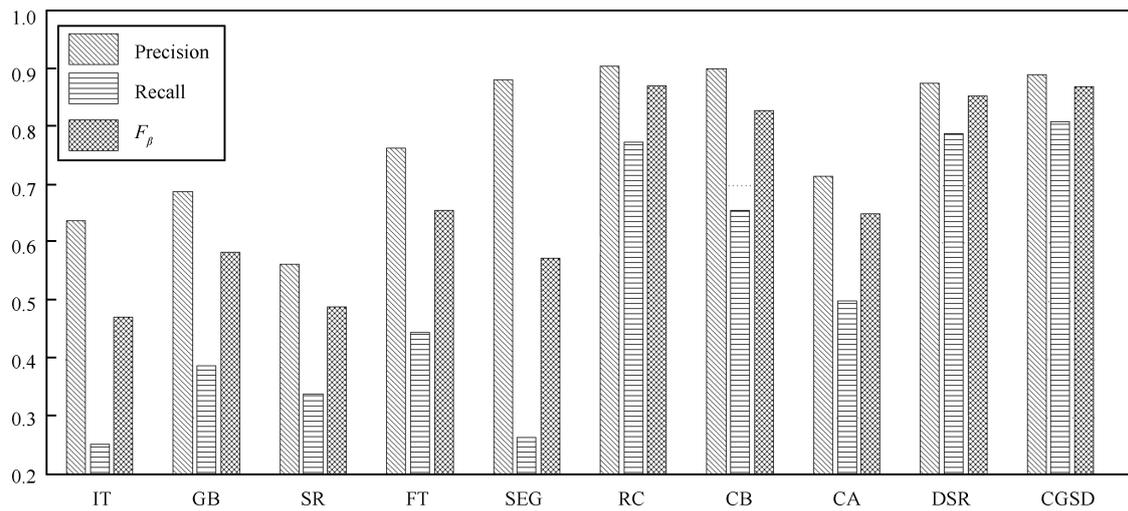


图9 显著性检测对比实验结果

Fig.9 Results of the contrast experiment of saliency detection

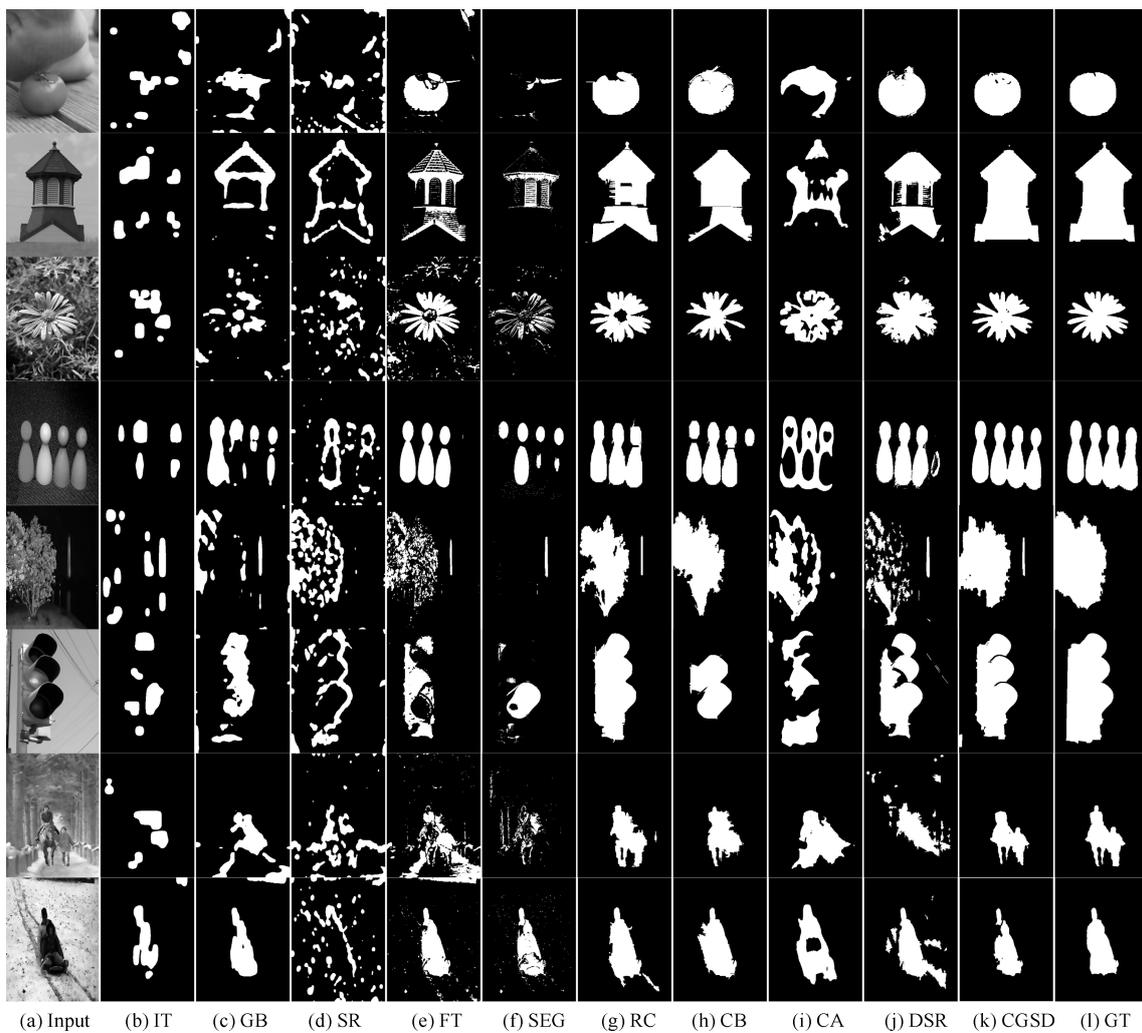


图10 显著性检测对比实验示例

Fig.10 Examples of the contrast experiment of saliency detection

来看, 相比其他算法, 算法 CGSD、SR、FT、RC、CB、CA、DSR 都取得了比较好的结果, 且 CGSD 所提取的显著性区域边界更加明确, 与真实的显著性区域最接近. 从显著性区域的检测精度来看, 算法 CGSD 提取的显著性区域最为接近真实显著性区域, 且包含的误检测区域最少.

进一步, 本文进行了显著性对象检测对比实验. 考虑到本文方法采用了分割方法 GrabCut, 因此在进行其他对比方法的显著性对象检测时, 本文在这些方法的基础上统一采用分割方法 GrabCut 进行处理. 首先, 本文以其他对比方法的显著图为输入, 利用文献 [40] 中基于图的分割方法获得一幅初始化掩码图; 然后, 以该初始化掩码图作为初始化输入, 利用分割方法 GrabCut 进行处理, 以获得显著性对象的检测结果. 与其他算法的显著性对象检测对比实验结果如图 11 所示, 部分示例的检测结果如图 12 所示. 从图 11 可以看出, 与算法 IT、GB、SR、SEG、FT、CA 相

比, CGSD 在所有指标上都高于这些算法, 整体性能明显优于这些算法. 与算法 CGSD 相比, 算法 RC 和算法 DSR 的 Precision、Recall、 F -measure 分别为 0.91、0.72、0.86 和 0.89、0.79、0.86, 在 Precision 和 F -measure 上三者相差无几, 然而在 Recall 上 CGSD 比 RC 高出 12.5%, 从整体上看 CGSD 性能要优于 RC、CGSD, 与 DSR 性能相当; 算法 CB 的 Precision、Recall、 F -measure 分别为 0.91、0.67、0.85, 在 Precision 和 F -measure 上 CGSD 与 CB 相差无几, 但在 Recall 上 CGSD 比 CB 高出 20%, 从整体上看 CGSD 性能要优于 CB. 在图 12 中, 从区域的内部致密性和区域的边界清晰度来看, 相比其他算法, 算法 CGSD、RC、CB 都取得了比较好的结果. 整体来说, CGSD 所提取的显著性区域内部更加致密, 边界更加明确, 与真实的显著性区域最接近, 且包含的误检测区域和漏检区域较少.

从整个数据集的角度来看, 本文方法的显著性

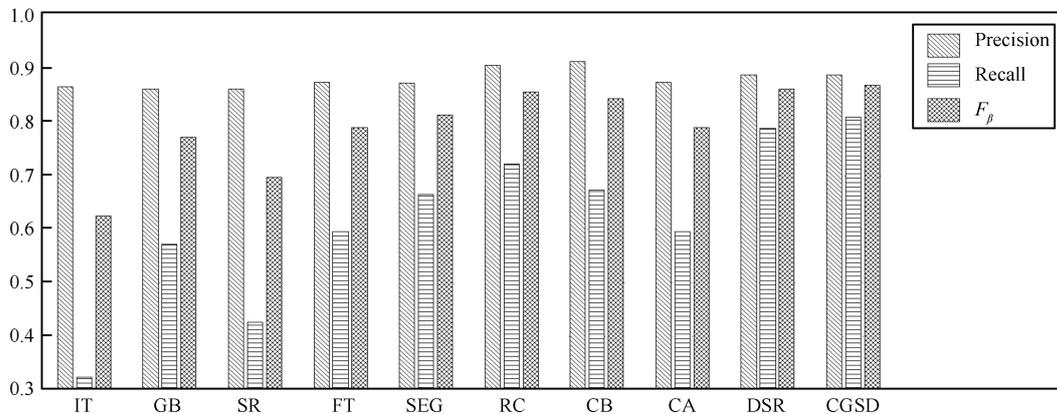


图 11 显著性对象检测对比实验结果

Fig. 11 Results of the contrast experiment of salient object detection

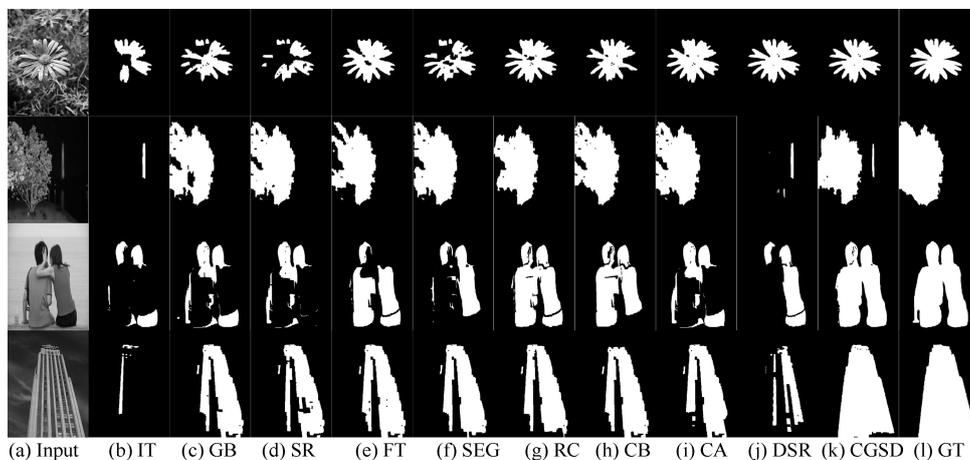


图 12 显著性对象检测对比实验示例

Fig. 12 Examples of the contrast experiment of salient object detection

区域检测取得了良好效果,但不可否认的是针对一些测试用例仍存在误检测或检测效果不佳的情形,如图 13 所示.其中,在图 13 中第 1 行是测试用例,第 2 行是本文方法 CGSD 的检测结果,第 3 行是显著性区域的真实检测结果.误检测示例的结果中存在包含背景区域、显著性区域检测不完整等情形.此外,本文方法还存在以下局限性:模型训练完成后,模型参数在测试过程中对所有测试用例是一样的,没有考虑到不同测试用例的特点;显著性特征的种类选取还不够充分,显著性特征的层次分类还比较粗糙.进一步的研究可以按以下内容展开:增加显著性特征种类;细化特征分层;结合场景分析方法分析测试用例的特点,并构建针对不同种类测试用例的训练数据集以获得更为有效灵活的模型参数.

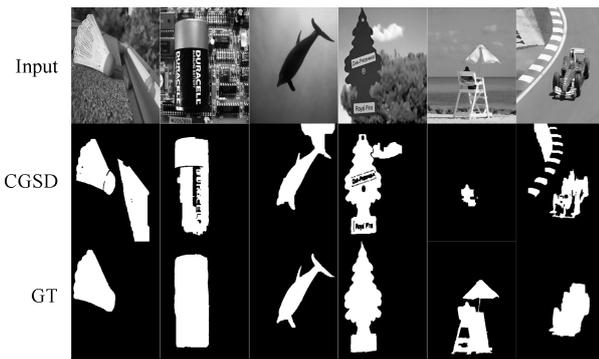


图 13 本文显著性检测方法的误检测示例

Fig. 13 False detection examples of the proposed method of salient object detection

4 结论

针对当前常见的显著性算法检测得到的显著性区域边界稀疏不明确、内部不均匀致密等问题,本文提出了一种基于条件随机场和图像分割的显著性检测方法.考虑了图像中边界信息、局部信息以及全局信息,该方法分别提取稀疏重构误差、空域调频对比度以及颜色空间分布这三种显著性特征.为了提高显著性检测的精度,该方法将显著性区域检测问题划分为两个步骤:显著性粗糙检测和精细检测.通过结合条件随机场模型与显著性特征,粗糙检测主要通过显著性区域与背景区域的区域标注实现;进一步结合图像区域标注结果,精细检测利用图像分割方法实现.针对数据集的多种对比实验表明,该方法能够有效提高显著性检测精度,清晰而准确地提取出图像中的显著性区域.

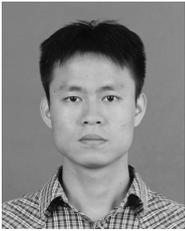
References

- 1 Loupiaz E, Sebe N, Bres S, Jolion J M. Wavelet-based salient points for image retrieval. In: Proceedings of the 2000 IEEE International Conference on Image Processing. Vancouver, British Columbia, Canada: IEEE, 2000, 2: 518–521

- 2 Jian M W, Dong J Y, Ma J. Image retrieval using wavelet-based salient regions. *The Imaging Science Journal*, 2011, 59(4): 219–231
- 3 Guo C L, Zhang L M. A novel multi-resolution spatiotemporal saliency detection model and its applications in image and video compression. *IEEE Transactions on Image Processing*, 2010, 19(1): 185–198
- 4 Kim W, Kim C. A novel image importance model for content-aware image resizing. In: Proceedings of the 18th IEEE International Conference on Image Processing. Brussels, Belgium: IEEE, 2011. 2469–2472
- 5 Gupta R, Chaudhury S. A scheme for attentional video compression. In: Proceedings of the 4th International Conference on Pattern Recognition and Machine Intelligence. Berlin, Heidelberg: Springer-Verlag, 2011. 458–465
- 6 Itti L. Automatic foveation for video compression using a neurobiological model of visual attention. *IEEE Transactions on Image Processing*, 2004, 13(10): 1304–1318
- 7 Kanan C, Cottrell G. Robust classification of objects, faces, and flowers using natural image statistics. In: Proceedings of the 2010 IEEE Conference on Computer Vision and Pattern Recognition. San Francisco, California, USA: IEEE, 2010. 2472–2479
- 8 Teuber H L. Physiological psychology. *Annual Review of Psychology*, 1955, 6: 267–296
- 9 Wolfe J M, Horowitz T S. What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, 2004, 5(6): 495–501
- 10 Desimone R, Duncan J. Neural mechanisms of visual selective attention. *Annual Review of Neuroscience*, 1995, 18(1): 193–222
- 11 Mannan S K, Kennard C, Husain M. The role of visual salience in directing eye movements in visual object agnosia. *Current Biology*, 2009, 19(6): 247–248
- 12 Treisman A M, Gelade G. A feature-integration theory of attention. *Cognitive Psychology*, 1980, 12(1): 97–136
- 13 Koch C, Ullman S. Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology*, 1985, 4(4): 219–227
- 14 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
- 15 Harel J, Koch C, Perona P. Graph-based visual saliency. In: Proceedings of the 2006 Advances in Neural Information Processing Systems. Vancouver, British Columbia, Canada: Bradford Book, 2006. 545–552
- 16 Hansen L K, Karadogan S, Marchegiani L. What to measure next to improve decision making? On top-down task driven feature saliency. In: Proceedings of the 2011 IEEE Symposium on Computational Intelligence, Cognitive Algorithms, Mind, and Brain. Paris, France: IEEE, 2011. 1–7

- 17 Baluch F, Itti L. Mechanisms of top-down attention. *Trends in Neurosciences*, 2011, **34**(4): 210–224
- 18 Itti L, Koch C. Computational modelling of visual attention. *Nature Reviews Neuroscience*, 2001, **2**(3): 194–203
- 19 Judd T, Ehinger K, Durand F, Torralba A. Learning to predict where humans look. In: Proceedings of the 12th IEEE International Conference on Computer Vision. Kyoto, Japan: IEEE, 2009. 2106–2113
- 20 Liu T, Yuan Z, Sun J, Wang J D, Zheng N N, Tang X O, Shun H Y. Learning to detect a salient object. *IEEE Transactions on Software Engineering*, 2011, **33**(2): 353–367
- 21 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, Rhode Island, USA: IEEE, 2012. 478–485
- 22 Jiang H Z, Wang J D, Yuan Z J, Wu Y, Zheng N N, Li S P. Salient object detection: a discriminative regional feature integration approach. In: Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition. Portland, Oregon, USA: IEEE, 2013. 2083–2090
- 23 Shen X, Wu Y. A unified approach to salient object detection via low rank matrix recovery. In: Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence, Rhode Island, USA: IEEE, 2012. 853–860
- 24 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, Oregon, USA: IEEE, 2013. 1155–1162
- 25 Yang C, Zhang L H, Lu H C, Ruan X, Yang M H. Saliency detection via Graph-Based manifold Ranking. In: Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition. Portland, Oregon, USA: IEEE, 2013. 3166–3173
- 26 Li X H, Lu H C, Zhang L H, Ruan X, Yang M H. Saliency detection via dense and sparse reconstruction. In: Proceedings of the 2013 IEEE International Conference on Computer Vision. Sydney, Australia: IEEE, 2013. 2976–2983
- 27 Hu Y Q, Rajan D, Chia L T. Robust subspace analysis for detecting visual attention regions in images. In: Proceedings of the 13th Annual ACM International Conference on Multimedia. Santa Fe, New Mexico, USA: ACM, 2005. 716–724
- 28 Hou X D, Zhang L. Saliency detection: a spectral residual approach. In: Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis, Minnesota State, USA: IEEE, 2007. 1–8
- 29 Guo C L, Ma Q, Zhang L M. Spatio-temporal saliency detection using phase spectrum of quaternion fourier transform. In: Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition. Anchorage, Alaska, USA: IEEE, 2008. 1–8
- 30 Zhang Qiao-Rong, Gu Guo-Chang, Liu Hai-Bo, Xiao Hui-Min. Salient region detection using multi-scale analysis in the frequency domain. *Journal of Harbin Engineering University*, 2010, **31**(3): 361–365
(张巧荣, 顾国昌, 刘海波, 肖会敏. 利用多尺度频域分析的图像显著区域检测. 哈尔滨工程大学学报, 2010, **31**(3): 361–365)
- 31 Lafferty J D, McCallum A, Pereira F C N. Conditional random fields: probabilistic models for segmenting and labeling sequence data. In: Proceedings of the 8th IEEE International Conference on Machine Learning. Williamstown, MA, USA: IEEE, 2001. 282–289
- 32 Rother C, Kolmogorov V, Blake A. Grabcut: interactive foreground extraction using iterated graph cuts. *ACM Transactions on Graphics*, 2004, **23**(3): 309–314
- 33 Kumar S, Hebert M. Discriminative random fields: a discriminative framework for contextual interaction in classification. In: Proceedings of the 9th IEEE International Conference on Computer Vision. Nice, France: IEEE, 2003, **2**: 1150–1157
- 34 Cheng M M, Zhang G X, Mitra N J, Huang X L, Hu S M. Global contrast based salient region detection. In: Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition. Colorado Springs, Colorado, USA: 2011 IEEE, 2011. 409–416
- 35 Sun J, Lu H C, Li S F. Saliency detection based on integration of boundary and soft-segmentation. In: Proceedings of the 19th IEEE International Conference on Image Processing. Orlando, Florida, USA: IEEE, 2012. 1085–1088
- 36 Wei Y C, Wen F, Zhu W J, Sun J. Geodesic saliency using background priors. In: Proceedings of the 12th European conference on Computer Vision-Volume Part III. Berlin, Heidelberg: Springer-Verlag, 2012. 29–42
- 37 Borji A, Sihite D N, Itti L. Salient object detection: A benchmark. In: Proceedings of the 12th European conference on Computer Vision-Volume Part II. Berlin, Heidelberg: Springer-Verlag, 2012. 414–429
- 38 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): 2274–2282
- 39 Achanta R, Hemami S, Estrada F, Susstrunk S. Frequency-tuned salient region detection. In: Proceedings of the 2009 IEEE International Conference on Computer Vision and Pattern Recognition. Miami Beach, Florida, USA: IEEE, 2009. 1597–1604
- 40 Achanta R, Susstrunk S. Saliency detection using maximum symmetric surround. In: Proceedings of the 2010 IEEE International Conference on Image Processing. Hong Kong, China: IEEE, 2010. 2653–2656
- 41 Jain, Anil K. *Fundamentals of Digital Image Processing*. Englewood Cliffs, New Jersey, USA: Prentice Hall, 1989. 51
- 42 Thomas S W. Efficient inverse color map computation. *Graphics Gems II*. Boston: Academic Press, 1991: 116–125
- 43 Otsu N. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 1979, **9**(1): 62–66

- 44 Goferman S, Zelnik-Manor L, Tal A. Context-aware saliency detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, **34**(10): 1915–1926
- 45 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 2011 British Machine Vision Conference*. Dundee, Scotland, UK: BMVA Press, 2011. 1–12
- 46 Rahtu E, Kannala J, Salo M, Heikkilä J. Segmenting salient objects from images and videos. In: *Proceedings of the 11th European conference on Computer Vision — Part V*. Berlin, Heidelberg: Springer-Verlag, 2010. 366–379



钱 生 2014 年获中国科学技术大学自动化系硕士学位。主要研究方向为计算机视觉, 模式识别和生物认知。

E-mail: qsheng@mail.ustc.edu.cn

(**QIAN Sheng** Received his master degree from the Department of Automation, University of Science and Technology of China (USTC) in 2014.

His research interest covers computer vision, pattern recognition, and biological intelligence.)

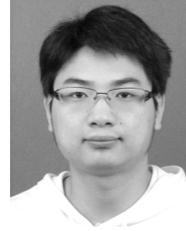


陈宗海 中国科学技术大学自动化系教授。1991 年获中国科学技术大学自动化系硕士学位。主要研究方向为复杂系统的建模, 仿真与控制, 智能机器人和量子信息控制。本文通信作者。

E-mail: chenzh@ustc.edu.cn

(**CHEN Zong-Hai** Professor in the Department of Automation, University

of Science and Technology of China (USTC). He received his master degree from USTC in 1991. His research interest covers modeling, simulation and control of complex system, intelligent robot, and quantum control. Corresponding author of this paper.)



林名强 中国科学技术大学自动化系博士研究生。2011 年获中国科学技术大学自动化系学士学位。主要研究方向为计算机视觉, 鲁棒估计和模式识别。

E-mail: kdmlmq@mail.ustc.edu.cn

(**LIN Ming-Qiang** Ph.D. candidate in the Department of Automation, University of Science and Technology of

China (USTC). He received his bachelor degree from USTC in 2011. His research interest covers computer vision, robust estimation, and pattern recognition.)



张陈斌 博士, 中国科学技术大学自动化系副教授。主要研究方向为复杂系统建模, 仿真与智能控制, 具体包括微电网系统的能源管理和控制, 新能源汽车动力能源系统的分析、设计和优化, 量子系统控制与仿真, 信息物理融合系统。

E-mail: zhangchb@ustc.edu.cn

(**ZHANG Chen-Bin** Ph.D., associate professor in the Department of Automation, University of Science and Technology of China (USTC). His research interest covers modeling, analysis and control of complex systems, such as energy management system of microgrids and electrical vehicles, mobile robots system, quantum control system, and cyber-physical system.)