

基于双目视觉的显著性目标检测方法

李庆武^{1,2}, 周亚琴¹, 马云鹏¹, 邢俊¹, 许金鑫¹

¹ 河海大学物联网工程学院, 江苏 常州 213022;

² 常州市传感网与环境感知重点实验室, 江苏 常州 213022

摘要 针对现有的显著性目标检测算法在受到相似背景干扰时, 易出现目标检测准确度低、稳定性差的问题, 提出一种基于双目视觉的显著性目标检测方法。受人眼视觉特性启发, 将双目视觉模型感知的深度信息作为显著性特征与多特征聚类分割结果进行协同处理, 定量分析图像区域级的深度显著性, 再将全局显著性与区域深度显著性进行加权融合, 突出目标区域, 根据融合结果的区域分布进行背景抑制, 完成显著性目标的检测。实验结果表明, 与现有的显著性目标检测算法相比, 该算法有效地抑制了相似背景的干扰, 并且准确度高、稳定性好。

关键词 图像处理; 目标检测; 双目视觉; 多特征融合; 区域分割; 视觉显著性

中图分类号 TP242.6

文献标识码 A

doi: 10.3788/AOS201838.0315002

Salient Object Detection Method Based on Binocular Vision

Li Qingwu^{1,2}, Zhou Yaqin¹, Ma Yunpeng¹, Xing Jun¹, Xu Jinxin¹

¹ College of Internet of Things Engineering, Hohai University, Changzhou, Jiangsu 213022, China;

² Changzhou Key Laboratory of Sensor Networks and Environmental Sensing, Changzhou, Jiangsu 213022, China

Abstract Aiming at the problem that the existing salient object detection algorithms suffers from the similar background interference, the detection accuracy of the target is low and the stability is poor. We propose a salient object detection method based on binocular vision. Firstly, inspired by the visual characteristics of the human eye, we consider the depth information acquired by binocular vision model as the salient features based on human visual characteristics. Secondly, we use the depth information and the result of region segmentation based on multi-feature fusion clustering to analyze the regional level depth saliency of image quantitatively. Thirdly, we make the weighted fusion of the global saliency map and regional level depth saliency map to highlight the objection area. Finally, we suppress the background to complete salient object detection based on the regional distribution of fusion results. The results show that compared with the existing methods, the proposed method can effectively suppress the interference of similar background with high accuracy and stability simultaneously.

Key words image processing; object detection; binocular vision; multi-feature fusion; region segmentation; visual saliency

OCIS codes 150.1135; 150.3040; 150.6910; 150.3045

1 引 言

人眼视觉系统能在复杂场景中选择性地处理感兴趣目标, 这种自适应选择的能力被称为视觉注意机制。计算机视觉领域将模仿人眼的注意机制定义为图像显著性检测技术, 用于辅助计算机快速提取图像目标区域, 实现资源合理分配, 提高处理准确率与效率。该技术已被广泛应用于目标检测与识别^[1-3]、图像分割^[4]、目标跟踪^[5]以及图像检索^[6]等

领域。

Borji 等^[7]根据目的差异将显著性检测方法分为人眼注意点位置预测和显著性目标检测。本文算法以目标检测为内驱力, 属于显著性目标检测范畴。根据检测特征的不同, 现存显著性目标检测方法可分为基于对比度信息的显著性目标检测与基于先验信息的显著性目标检测。

在人类注意力模型的简化过程中发现, 对比度是视觉注意机制的主要驱动因素。大量基于对比度

收稿日期: 2017-08-09; 收到修改稿日期: 2017-10-08

基金项目: 江苏省重点研发计划(BE2016071, BE2017648)

作者简介: 李庆武(1964—), 男, 博士, 教授, 主要从事智能感知与图像处理方面的研究。E-mail: li_qingwu@163.com

信息的显著性目标检测算法被相继提出, Ma 等^[8]在 LUV 颜色空间计算邻域像素的颜色欧氏距离, 利用局部颜色对比信息度量图像显著性, 但该方法邻域尺寸固定, 难以满足不同尺寸的显著性目标检测。Achanta 等^[9]提出利用可变尺寸感知单元的对比度信息实现多尺度显著性目标检测, 该方法更加注重图像细节, 但易受背景噪声的影响, 从而增大误检率。Zhai 等^[10]提出以像素点灰度特征的全局对比度作为显著性度量依据, 该方法运算速度快, 但忽略了图像空间关系。Achanta 等^[11]采用图像颜色模型(Lab)均值作为全局信息, 根据图像每个像素点的颜色距离得到显著图, 但该方法仅考虑了一阶平均颜色, 不满足复杂背景下显著目标的检测要求。Cheng 等^[12]提出基于直方图对比度的方法, 根据每一个像素与图像中其他像素的颜色差异确定显著值, 该方法考虑了区域的空间相关性, 从而提高了检测准确率, 但在与目标具有相似颜色或具有复杂纹理的背景下检测效果不佳。综上所述, 对比度信息能在显著性检测中取得良好的效果, 但仅依靠图像对比度进行检测, 难免将图像中具有稀有性或独特性的背景区域误检为目标。因此研究人员根据显著性目标空间位置约束关系, 将一些先验条件引入视觉显著性检测模型中, 从而降低误检率。

先验信息可大致分为两类: 中心先验和背景先验。中心先验通常假设显著性目标位于图像中心, 如文献^[13]以中心矩形构图为先验信息, 构建基于中心矩形线上的超像素特征, 根据超像素流行排序获取显著图。中心先验方法有一定主观性, 易将显著性检测的注意点引入背景范围, 从而忽略了显著性区域。与中心先验不同, 背景先验则假设图像边界大多为背景区域。Li 等^[14]通过超像素提取图像边界作为背景模板, 计算密集与稀疏重构误差, 利用多尺度重建误差融合计算像素级显著性, 该算法能有效突出显著性目标, 但易造成边界目标的漏检, 且同一区域中的显著值不连续。Zhu 等^[15]提出以边界连通度作为背景定义准则, 解决了目标与图像边界连接的问题, 但在目标区域与背景区域颜色相似情况下效果不佳。近年来, 很多研究人员将两种先验信息相结合, 提出新的融合算法。周帅骏等^[16]利用边缘超像素为吸收节点计算背景先验值, 再将其与中心先验图融合得到视觉显著图, 与单一先验信息方法相比, 该方法充分利用了图像信息, 能够高亮地凸显图像显著目标。林华锋等^[17]利用基于目标中心优先与背景模板抑制的空间先验方法获得空间

先验显著图, 该方法能够准确有效地检测出位于图像任意位置的显著性物体。但是上述两种融合算法在显著性目标与背景模板特征相似的情况下, 均不能完整地检测出显著性区域。

人眼视觉由视觉注意力机制和双目视觉系统共同构建。视觉注意力机制通过双目视觉系统接收视觉信号, 并根据视觉信息协同处理结果快速提取目标, 帮助人类完成重要目标的快速感知。双目视觉系统是在人类的发展进化过程中逐渐形成的, 是为感知和适应环境而自我进化的信息获取工具, 不仅可以助人脑感知画面色彩、亮度等基本信息, 还可以根据图像深度分布特征, 获取场景深度分布信息, 从而更准确地获得外界物体形状、方位以及距离等描述信息, 从而判断自身与客观环境间的位置关系, 是人类视觉注意力机制感知信息的重要工具。该过程表明深度信息分布也是一种重要的视觉感知信息, 将其作为显著性特征与其他现有显著性特征协同处理, 可以更加准确地完成相似背景中的目标显著性检测。

受上述人眼视觉特性的启发, 本文提出基于双目视觉的显著性目标检测算法, 将深度信息作为视觉显著性特征与其他显著性特征协同处理, 克服了显著性目标检测易受相似背景干扰的问题, 并且具有显著性目标检测准确度高、通用性好等优点。

2 基于双目视觉的显著性目标检测

本文算法处理过程如图 1 所示, 为方便与其他算法进行比较, 本文目标检测对象均设为左目图像, 除视差获取步骤, 其余步骤均对左目图像进行操作。首先利用双目视觉系统获取待测物体的左右目图像信息, 再利用改进 SURF 算法在左右目图像中获取稀疏视差点, 并使用线性迭代聚类(SLIC)算法获取左目图像的超像素, 然后利用多特征融合聚类合并相邻且相似的超像素完成区域分割, 统计稀疏视差点在分割区域中的投影分布, 根据每个区域中匹配点视差的中值构建稀疏视差图获取深度显著性, 最后将深度显著性与全局显著性进行加权融合与背景抑制, 获取最终显著图。

2.1 基于多特征融合聚类的区域分割

人类的视觉观察系统以图像区域为基本单位, 基于区域进行显著性检测符合人眼视觉特征。超像素技术作为像素区域的构建方法在计算机视觉领域中得到了广泛应用^[18]。超像素能将图像快速分割为具有一定语义的子区域, 有利于局部特征的提取

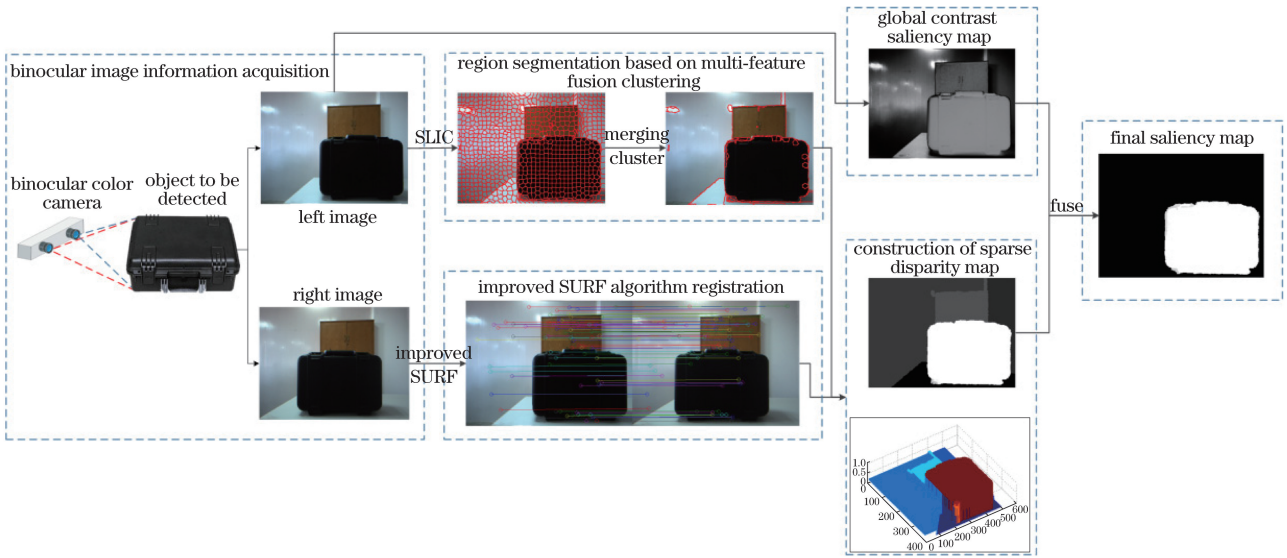


图 1 本文算法流程

Fig. 1 Flow chart of the proposed algorithm

与结构信息的表达,满足人眼视觉系统区域观察特性^[19]。在众多超像素构建方法中,SLIC 算法在边缘贴合度和紧密度两方面取得了较好的平衡,具备较高的计算效率,是目前综合性能较为优异的超像素分割算法^[20]。本文采用 SLIC 算法在左目图像中进行超像素块分割,并利用多特征融合机制制定超像素聚类策略,从而完成整幅图像的区域分割。

将左目图像转换到 CIE Lab 颜色空间,通过 SLIC 方法将图像分割为 k 个超像素,每个超像素赋予一个独有的标识符 i ,且所有标识符为 1 到 k 的连续正整数,提取每个超像素区域 S_{P_i} 的 7 维特征作为超像素区域的度量属性。

$$S_{P_i} = \{L_i, a_i, b_i, C_{coni}, C_{cori}, E_i, H_i\},$$

$$i = 1, 2, 3, \dots, k, \quad (1)$$

式中 L_i, a_i, b_i 分别为每个超像素区域的 L, a, b 颜

色分量的均值, $C_{coni}, C_{cori}, E_i, H_i$ 分别为每个超像素区域灰度共生矩阵特性中的对比度均值、互相关均值、能量均值,以及齐次性(同质性)均值。

超像素能够完成局部相似区域的聚类,但不满足大区域分割的需求,为此本文提出利用颜色、纹理特征融合机制,对相邻超像素的差异性进行度量,完成超像素区域的自动合并,实现大区域聚类,从而满足多尺寸目标检测的要求。

利用超像素区域颜色、纹理以及空间位置信息对相似超像素进行合并,实现图像的区域划分。图 2 为多特征融合区域分割算法示意图,图像被分为若干个超像素图像块,每个图像块拥有一个连续的标识符,经过合并处理后,图像被分为 3 个完整区域,且每个新的区域被赋予新的标识符,最后根据标识符找出图像中的 3 个完整区域,完成区域分割。

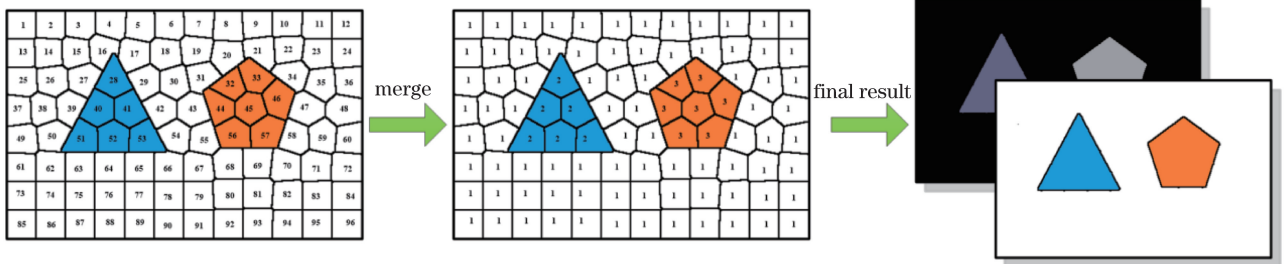


图 2 基于多特征融合聚类的区域分割示意图

Fig. 2 Schematic of region segmentation based on multi-feature fusion clustering

图 2 的合并策略算法如下:

1) 首先获取所有超像素的相邻信息,找出所有相邻的超像素对,

$$S_{P_{ij}} = \{(S_{P_i}, S_{P_j})\}, i \in [1, k], j \in [1, k],$$

$$i \neq j, \quad (2)$$

式中 $S_{P_{ij}}$ 为标识符为 i, j 的超像素组成的超像素对,

k 为整幅图的超像素个数, S_{P_i} 和 S_{P_j} 为一组相邻超像素的两个超像素 7 维特征参数, 每幅图像中相邻超像素对的个数由 SLIC 超像素分割情况决定。

2) 以超像素相邻作为合并的前提条件, 利用颜色、纹理特征计算所有相邻超像素对 $S_{P_{ij}}$ 之间的差异性,

$$D_{ij} = \omega_1 \cdot \sqrt{\frac{\epsilon + (L_i - L_j)^2 + (a_i - a_j)^2 + (b_i - b_j)^2}{3}} + \omega_2 \cdot \sqrt{\frac{\epsilon + (C_{cori} - C_{corj})^2 + (C_{coni} - C_{conj})^2 + (E_i - E_j)^2 + (H_i - H_j)^2}{4}}, \quad (3)$$

式中 i, j 为相邻超像素的标识符值。 ϵ 为接近于 0 的正数, 用于保证数值有效性, 本文取值为 10^{-4} 。 ω_1 和 ω_2 分别为颜色特征与纹理特征的权重, 用于调节两种特征的比值关系。 图像中某种特征数据集

的离散程度越大, 则表明该特征差异性及其对图像的影响力越大。 均方差能够有效表征数据之间的差异程度, 故本文利用图像全局颜色、纹理均方差调节两类特征的权重。

$$\omega_1 = \frac{\sqrt{\sum_i^H \sum_j^W (L_{ij} - \bar{L})^2} + \sqrt{\sum_i^H \sum_j^W (a_{ij} - \bar{a})^2} + \sqrt{\sum_i^H \sum_j^W (b_{ij} - \bar{b})^2}}{3\sqrt{WH}}$$

$$\omega_2 = \frac{\sqrt{\sum_i^H \sum_j^W (C_{corij} - \bar{C}_{cor})^2} + \sqrt{\sum_i^H \sum_j^W (C_{conij} - \bar{C}_{con})^2} + \sqrt{\sum_i^H \sum_j^W (E_{ij} - \bar{E})^2} + \sqrt{\sum_i^H \sum_j^W (H_{ij} - \bar{H})^2}}{4\sqrt{WH}}, \quad (4)$$

式中 (i, j) 为图像 i 行 j 列的像素点, H 和 W 为图像的高与宽, $\bar{L}, \bar{a}, \bar{b}$ 分别为全局 L, a, b 颜色分量均值。 相似地, $\bar{C}_{cor}, \bar{C}_{con}, \bar{E}, \bar{H}$ 分别为全局灰度共生矩阵特性中的对比度均值、互相关均值、能量均值和齐次性(同质性)均值。

3) 若相邻超像素之间的差异性小于一定阈值 T_h , 将认为该相邻超像素对可合并。

$$T_h = \sqrt{\frac{\omega_1 \cdot (\bar{L} + \bar{a} + \bar{b})}{3}} + \sqrt{\frac{\omega_2 \cdot (\bar{C}_{cor} + \bar{C}_{con} + \bar{E} + \bar{H})}{4}}。 \quad (5)$$

利用(5)式找出所有相邻且相似的超像素对 $S'_{P_{ij}} | D_{ij} < T_h$, 如图 3 所示, 从图像左上角起始点开始, 按照从左到右, 从上到下顺序进行合并, 下文中的“点”均以超像素为单位。

1) 以起始点为聚类起点, 寻找所有与其相邻且相似的超像素归为一类, 同时对寻找到的每个同类超像素继续寻找其相邻且相似的超像素归入该类中, 以此类推, 直至找到满足条件的所有超像素, 则该类超像素的聚类完成, 将已归类超像素标记为已处理状态, 不再参与其他超像素的聚类, 更新聚类起始点。

2) 对新的生长起始点进行同样的归类处理, 直至所有超像素均被标记为已处理状态, 则停止聚类, 获取所有分类情况。

此处可将超像素合并过程视为“tree”的生长, 以每个途经超像素为节点, 不断向下寻找满足相邻且相似条件的子节点, 直到没有满足条件的子节点, 则停止该 tree 的生长。 遍历图像中所有超像素, 找出每个节点所有的父节点与子节点, 最后生长形成若干 tree, 每棵 tree 上的所有节点对应的超像素均可合并形成具有相似特征的大区域。

2.2 构建稀疏视差图

双目能够同时从两个不同的角度观察同一个物体, 由于视角不同而产生的方向上的位置差异叫做视差。 根据视差大小即可获取物体的远近关系, 从而找出距离最近的物体区域, 利用空间位置信息, 将最远的区域作为背景区域, 从而有效抑制相似背景区域、亮斑以及噪声的影响。

在构建稀疏视差图的过程中, 首先利用改进 SURF 算法获取稀疏匹配点。 因传统 SURF 算法匹配错误率高, 极大地降低了视差计算的准确度, 故本文利用几何特征对匹配特征点进行约束, 提出基于斜率主导性的改进特征点匹配策略, 有效提升了匹配精度。 该匹配策略步骤如下:

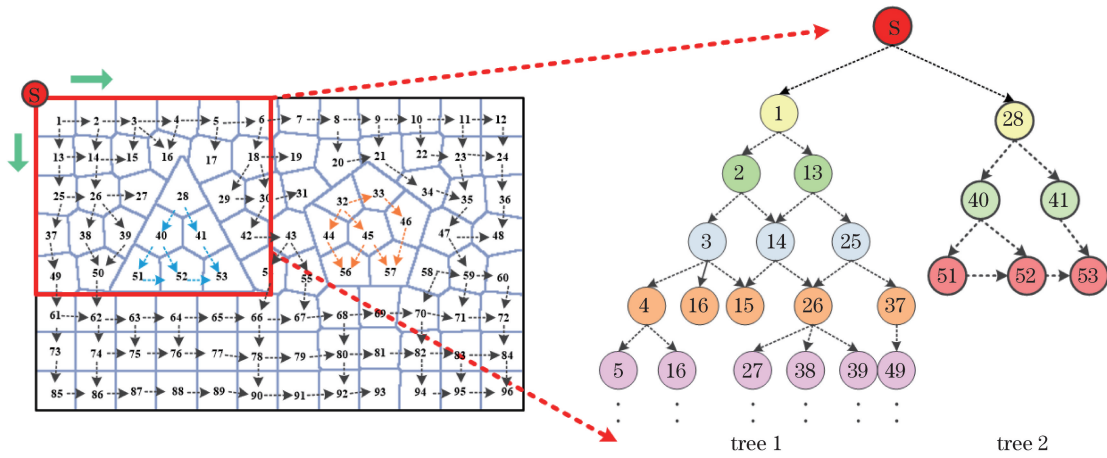


图 3 超像素合并策略示意图

Fig. 3 Schematic of superpixel merging strategy

1) 利用 SURF 算法获得左目图像和右目图像特征点坐标及所有特征点的 64 维描述符,左目图像和右目图像的特征点定义为

$$\begin{cases} P_{os1} = \{(x'_1, y'_1), (x'_2, y'_2), \dots, (x'_i, y'_i), \dots, \\ \quad (x'_m, y'_m)\}, 1 \leq i \leq m \\ P_{os2} = \{(x_1, y_1), (x_2, y_2), \dots, (x'_j, y'_j), \dots, \\ \quad (x_n, y_n)\}, 1 \leq j \leq n \end{cases} \quad (6)$$

式中 P_{os1} 为左目图像的特征点参数, P_{os2} 为右目图像特征点参数, m, n 分别为左目图像和右目图像中特征点的数目, i, j 分别为左目图像和右目图像特征点的下标, (x'_m, y'_m) 为左目图像第 m 个特征点坐标, (x_n, y_n) 为右目图像第 n 个特征点坐标。

2) 计算左目图像的特征点参数 P_{os1} 和右目图像特征点参数 P_{os2} 中所有点的欧氏距离,选择欧氏距离最小的点作为粗略匹配点,按欧氏距离升序对粗略匹配点进行排序,删除异常点,选择前 k 个匹配点,定义为

$$P_{os_k} = \{ \{(x'_1, y'_1), (x_1, y_1)\}, \dots, \{(x'_i, y'_i), (x_i, y_i)\}, \dots, \{(x'_k, y'_k), (x_k, y_k)\} \}, 1 \leq i \leq k. \quad (7)$$

3) 根据 k 个匹配点 P_{os_k} 中相应点的斜率筛选匹配点,计算所有粗略匹配点对的斜率,以 10^{-2} 数量级保留所有斜率值,计算所有斜率的出现频率,选取出现频率最大的斜率作为主导斜率,删除其他异常斜率对应的匹配点,更新得到 H 组准确匹配点对 $P_{os_k_new}$,公式如下:

$$P_{os_k_new} = \{ \{(x_{z1}, y_{z1}), (x_{y1}, y_{y1})\}, \dots, \{(x_{zi}, y_{zi}), (x_{yi}, y_{yi})\}, \dots, \{(x_{zH}, y_{zH}), (x_{yH}, y_{yH})\} \}, 1 \leq i \leq H, \quad (8)$$

式中 (x_{zi}, y_{zi}) 和 (x_{yi}, y_{yi}) 分别为一组匹配点对中

左目图像和右目图像的特征点坐标。

图 4 为匹配策略示意图,获取图像的左、右目图像的特征点后,左、右目图像误匹配率高,经过描述子欧氏距离排序筛选,剔除部分异常点得到粗匹配点对,再通过对匹配点对斜率的出现频率进行排序,保留与出现频率最高的斜率一致的匹配点对,获取最终匹配结果。由图 4 可知,匹配正确率得到了极大提高。

获取的匹配点仅能得到离散的视差点图,不能反映整幅图像的视差信息,获取区域级的视差图还需要进一步的处理。在研究过程中发现,在检测目标时,同一物体上的视差值相近,可视为相同视差。结合分割得到的图像区域,提出基于区域构建图像稀疏视差图,如图 5 所示。

图 5(b)中以合并区域为单位,统计每个区域中视差点的投影情况 $\{p_1, p_2, \dots, p_N\}$,对落入该区域的所有视差点按视差值进行升序排序,若该区域内投影视差点个数 $N \geq 3$,则取视差值中值作为区域视差值 $P_{R,c}$,若其 $0 < N < 3$,取最小值作为区域视差值,若该区域内无视差点,则视差值取 0。区域视差值越大,则表明该区域深度显著性越高。

$$P_{R,c} = \begin{cases} \text{median}\{p_1, p_2, \dots, p_N\}, & N \geq 3 \\ \min\{p_1, p_2\}, & 0 < N < 3, \\ 0, & N = 0 \end{cases} \quad c \in [1, M], d \in [1, N], \quad (9)$$

式中 c 为合并区域的标识符值, M 为合并区域数目, N 为落入该区域内的视差点的总数, p_d 为该区域内第 d 个视差点的视差值。

将图 5 中的黑色箱子作为图中距离相机最近的完整目标,在图 5(c)中视差值达到最大,而后面距

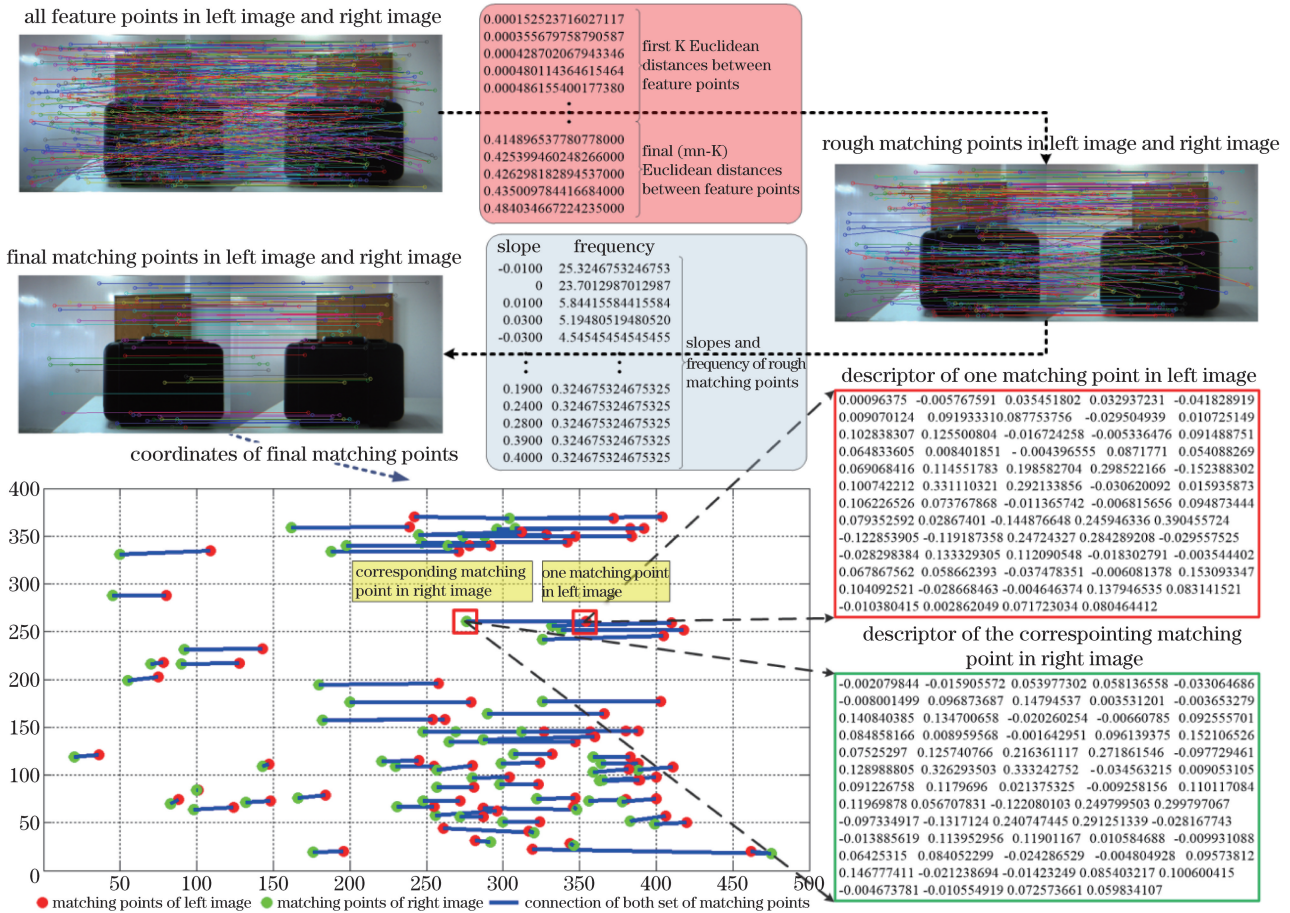


图 4 匹配策略示意图

Fig. 4 Schematic of matching strategy

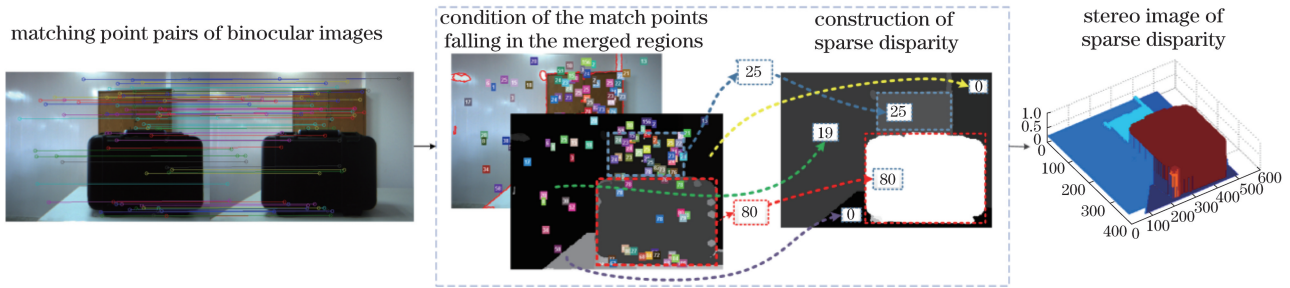


图 5 稀疏视差构建示意图。(a)匹配点对;(b)合并区域匹配点;(c)稀疏视差图;(d)立体视差图

Fig. 5 Schematic of sparse disparity construction. (a) Matching point pairs; (b) match points in the merged regions; (c) sparse disparity; (d) stereo image of sparse disparity

离相机较远的棕色柜子视差则减小。稀疏视差图中的亮度与视差大小成正比，视差越大，该区域亮度值越高。从图 5(d)中也可以看出，黑色箱子区域明显突起，表明该区域与相机镜头距离很近，后方背景中的柜子与墙壁的视差则随着与镜头距离的增大而减小。

2.3 显著图融合背景抑制

根据视差信息可以获取深度显著区域，该区域不仅包含待测物体，也可能包含承载物体的平台等

其他背景信息。如图 6 所示，由于物体承载平台同样距离相机镜头较近，部分区域甚至比待测物体区域更接近相机镜头，仅利用视差深度显著信息难以完全避免该区域影响。

本文利用全局显著信息对该类背景进行抑制，去除距离相机过近的非目标区域信息。Achanta^[11]提出的快速显著物体检测算法(FT)计算简单，能有效检测出显著目标区域，弱化背景区域，弥补深度显著性检测的缺陷。将深度显著性检测的稀疏视差图

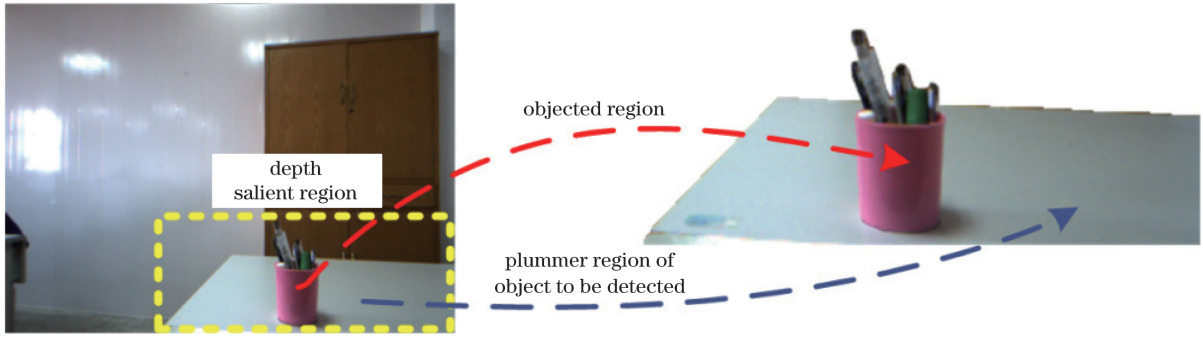


图 6 深度显著性检测局限性

Fig. 6 Limitation of depth saliency detection

与 FT 全局显著图按比例进行融合,得到融合显著图 S_{fuse} :

$$\begin{cases} S_{fuse}(x,y) = \omega_1 \cdot S_{parallax}(x,y) + \omega_2 \cdot S'_{FT}(x,y) \\ S_{FT}(x,y) = \|\lceil \bar{L} \quad \bar{a} \quad \bar{b} \rceil - I_{whc}(x,y)\| \\ S'_{FT}(i) = \frac{1}{n} \sum_{j=1}^n S_{FT}(x,y) \end{cases}, \quad (10)$$

式中 $\lceil \bar{L} \quad \bar{a} \quad \bar{b} \rceil$ 为图像 L, a, b 的 3 个分量的算术平均值, $I_{whc}(x,y)$ 为对原图像进行的高斯模糊后的平滑图像。 $\|\cdot\|$ 为欧氏距离, $S_{parallax}$ 为稀疏视差图, S'_{FT} 为 FT 方法检测的全局显著图在每个分割区域中的均值显著图, i 为第 i 个合并区域, j 为该区域内第 j 个像素点, n 为该区域内的像素点的个数。 ω_1 和 ω_2 分别为其权重, 由于稀疏视差图的贡献较全局显著图更大, 本文取值分别为 0.7 和 0.3。

最后对融合后的所有区域的显著值进行排序, 取其中值的 80% 作为阈值, 若显著值低于阈值, 则认为该区域为背景区域, 将该区域显著值置 0, 降低背景区域的影响, 得到最终显著图:

$$S_{final} = \begin{cases} S_{fuse}, & S_{fuse} > 0.8 \times \text{median}(S_{fuse}) \\ 0, & \text{otherwise} \end{cases}. \quad (11)$$

图 7 为显著图融合及背景抑制前后对比示意图, 由图 7 可知, 深度显著图中部分桌子区域也具有较高的显著性, 易被误判为目标, FT 显著图中该部分显著性较小, 通过融合, 全局显著性有效弥补了深度显著性造成的误检情况, 降低了桌子部分的显著性, 最后通过抑制背景突出目标区域, 获取最终的显著结果。

3 实验结果与分析

本文算法的运行环境为 Intel Core I5, 2.60 GHz CPU, 8 GB 内存, 编译环境为 MATLAB R2013a, 仿真操作界面如图 8 所示, 成像设备采用 Bumblebee 2

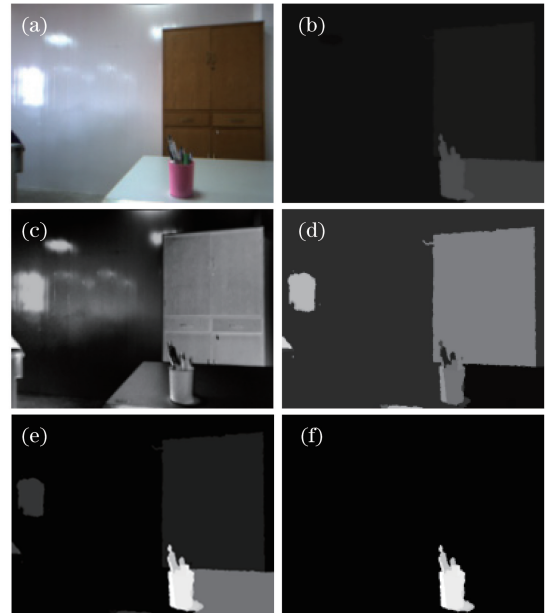


图 7 显著图融合背景抑制结果比较示例。

(a) 原左目图像; (b) 稀疏视差图; (c) FT 全局显著图; (d) FT 区域均值显著图; (e) 融合结果显著图; (f) 背景抑制结果显著图

Fig. 7 Comparison of saliency map fusion and background suppression results. (a) Original left image;

(b) sparse disparity map; (c) FT global saliency map; (d) FT regional mean saliency map; (e) fusing saliency map; (f) saliency map of results after background interference

双目立体视觉设备, 基线距离为 120 mm, 分辨率为 1024 pixel \times 768 pixel。

3.1 评价准则

本文采用了查准率-查全率 ($P-R$) 图、 F 值、AUC 以及平均绝对误差值 (MAE) 这 4 个指标综合评价算法性能。

查准率与查全率的曲线 ($P-R$ 曲线) 计算公式为

$$P = \frac{\sum_{(x,y)} G_T(x,y) B_S(x,y)}{\sum_{(x,y)} B_S(x,y)}, \quad (12)$$

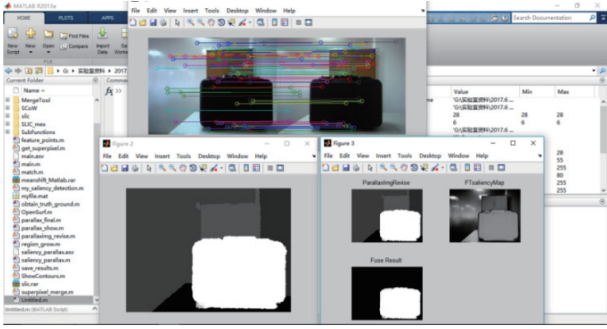


图 8 仿真软件操作界面

Fig. 8 Operation interface of simulation software

$$R = \frac{\sum_{(x,y)} G_T(x,y) B_S(x,y)}{\sum_{(x,y)} G_T(x,y)}, \quad (13)$$

式中 $G_T(x,y)$ 为人工标注图 (GT), $B_S(x,y)$ 为二值分割图, 二值分割阈值 $t = \{0, 5, 10, 15, \dots, 250, 255\}$, 将分割阈值 t 设定为 $0 \sim 255$, 步长为 5 的正整数, 对图像进行二值分割, 计算每个阈值下分割图像的查准率和查全率, 以查全率为横轴, 查准率为纵轴, 画出关于阈值 t 在图像中的 $P-R$ 曲线图。

由 (12) 和 (13) 式可知, 查准率与查全率成反比, 查全率越高, 查准率则会下降, 不足以定量地对算法性能进行评价。 F 值与 AUC 值作为综合二者指标的评估指标, 能够定量地对 $P-R$ 曲线图结果进行分析。

$$F = \frac{(1 + \beta^2) \cdot P \cdot R}{\beta^2 \cdot P + R}, \quad (14)$$

式中 $\beta = 0.3$ 。 F 值越大, 表示该算法性能越好。

由于不同阈值下的查全率与查准率会有不同的表现, 因此设置 AUC 能更全面地评价系统性能。系

统的性能越好, 其评价曲线就会尽可能地向上突出, 即曲线与坐标轴之间的面积越大。平均准确率 AUC 即 $P-R$ 曲线与坐标轴之间包含的面积, 其数学定义如下:

$$A_{AUC} = \int_0^1 P(R) dR. \quad (15)$$

为了避免二值分割阈值选取对算法评价的干扰, 可通过 MAE 评价算法性能。即计算图像中每个像素点的显著值 $S(x,y)$ 与人工标注图 $G_T(x,y)$ 之间绝对差的均值, MAE 值越小, 算法性能越优异。

$$M_{MAE} = \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N |S(x,y) - G_T(x,y)|. \quad (16)$$

3.2 结果分析

为全面评价本文算法, 将本文算法与 12 种显著性检测算法进行比较。为更加直观地进行对比, 将 12 种算法按提取特征分为 2 类: 1) 基于对比度信息算法, 包括 $LC^{[10]}$ 、 $AC^{[9]}$ 、 $FT^{[11]}$ 、 $CA^{[21]}$ 、 $SEG^{[22]}$ 、 $HC^{[12]}$ 、 PCA 算法^[23]; 2) 基于先验信息算法, 包括 $DSR^{[14]}$ 、 $GR^{[24]}$ 、 $RBD^{[15]}$ 、 $LPS^{[25]}$ 、 $MILPS$ 算法^[26]。

采用实验室随机拍摄的 100 组实景双目图像进行测试, 并对本文算法与上述 12 种算法测试的部分对比显著图像与所有测试图像处理结果进行数据分析, 为方便比较, 实验中本文算法与 12 种对比算法均对左目图像进行测试。

3.2.1 显著图视觉效果对比

图 9 为本文算法与全局对比度算法的显著图对比。第 1 行图像目标为放有一支马克笔的绿色笔筒, 该笔筒表面纹理复杂, 且光照不均, 背景为橙黄色

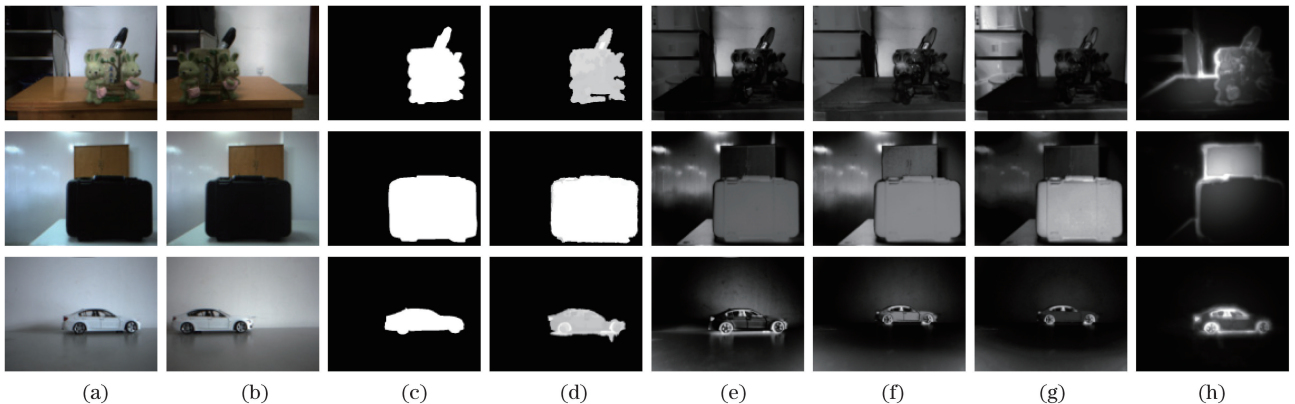


图 9 本文算法与全局对比度算法显著图对比。(a)原左目图;(b)原右目图;(c) GT 算法;(d)本文算法;(e) LC 算法;(f) FT 算法;(g) HC 算法;(h) PCA 算法

Fig. 9 Comparison of saliency maps generated by the proposed algorithm and global contrast algorithms.
(a) Original left image; (b) original right image; (c) GT algorithm; (d) proposed algorithm;
(e) LC algorithm; (f) FT algorithm; (g) HC algorithm; (h) PCA algorithm

平台、白墙以及浅灰色柜子,墙壁部分区域光线较亮,易对检测结果造成干扰。LC算法与FT算法能够保留完整的物体轮廓,且显著图较清晰,但是均将背景中的柜子与墙壁标记为显著部分,HC算法错误地将背景中的柜子标记为显著性区域,PCA算法能够检测出目标的部分边缘,但未能检测出目标完整区域,且存在误检情况。本文算法受到橙黄色平台阴影部分干扰,造成底部较暗区域的漏检,但是该算法能够检测并准确标注该笔筒的大部分区域,准确率较高。第2行检测图像目标物为置于最前方的黑色柜子,黑色为非显著性颜色,在显著性检测中具有一定难度,背景中存在明显光照不均造成的亮斑,是检测过程中的主要干扰信息。FT、LC、HC算法能够正确将黑色箱子标注为显著区域,但是受背景中亮斑的影响,存在误检情况。PCA算法误检率高,检测失败。本文算法能够准确检测目标区域,准确率高于95%。第3行图像为置于相似背景下的白色汽车模型,5种算法均能正确将该汽车区域标注为显著性区域,LC、HC、PCA算法检测结果中,仅汽车边缘部分显著度较高,汽车表面显著度则较低,而FT算法存在漏检情况。本文算法检测结果轮廓清晰,虽存在少量误检测部分,但大部分区域标注正确。

本文算法与局部对比度模型算法的显著图对比情况如图10所示。3幅测试图像背景相似,均为白色墙壁、浅灰色柜子以及橙黄色放置平台。第1行测试图像中目标物为黑色带花纹陶瓷杯,该目标颜色较暗且表面光照不均,检测难度大。AC算法和CA算法均检测失败,SEG算法检测结果误检率高,本文算法存在少量误检区域,但仍能够正确标注出大部分显著目标区域,且轮廓清晰,效果明显优于其他算法。第2行检测目标为放置有木质勺子的绿底波点陶瓷杯。AC算法无法检测,CA算法仅检测出物体边界,误检率高,SEG算法将大量背景区域错误标记为高显著区域,效果较差,本文算法能够正确检测目标区域中的陶瓷杯部分,由于木质勺子颜色差异明显,且表面光滑,难以进行匹配,故未能准确检测。第3行图像中目标物为灰色升降模型,其中红色承载台颜色鲜艳,较易检测,但灰色升降台与墙壁颜色相似,是检测难点。AC、CA、SEG算法仅标注出模型中颜色较鲜明的小部分区域。由于深度信息显著性的局限性,本文算法检测结果中,距离拍摄镜头较近的部分平台区域被标注为显著区域,但该算法能够成功检测出模型区域,较其他算法结果突出。

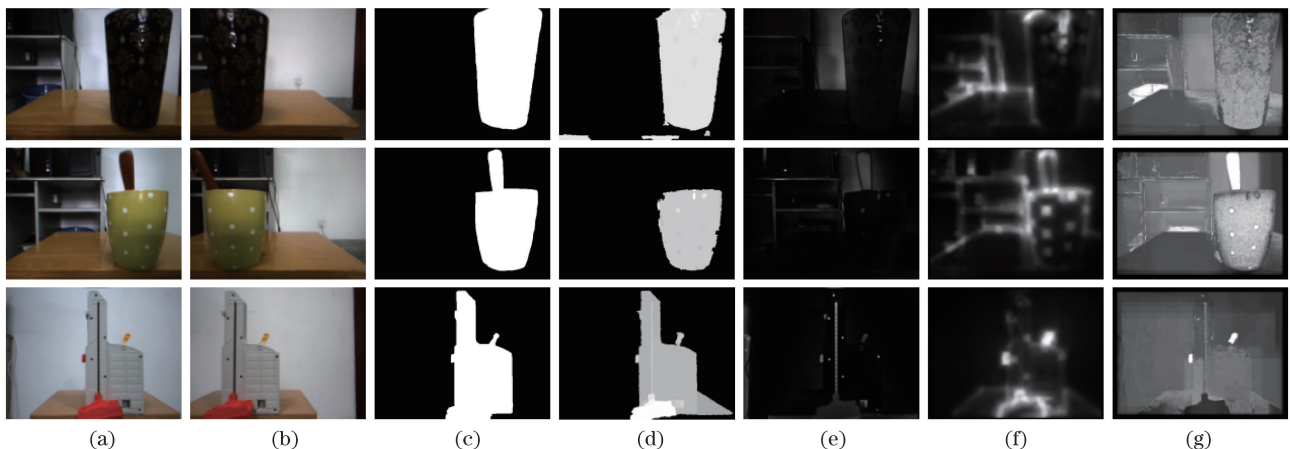


图10 本文算法与局部对比度算法显著图对比 (a)原左目图;(b)原右目图;(c) GT算法;
(d)本文算法;(e) AC算法;(f) CA算法;(g) SEG算法

Fig. 10 Comparison of saliency maps generated by the proposed algorithm and local contrast algorithms.

(a) Original left image; (b) original right image; (c) GT algorithm; (d) proposed algorithm;
(e) AC algorithm; (f) CA algorithm; (g) SEG algorithm

本文算法与先验信息算法显著图的对比情况如图11所示。前3行测试图像的背景均为白墙木纹柜,目标为放置于浅灰色平台上的较小物体,图像背景中墙壁反光现象严重,易对检测造成干扰,木纹柜颜色鲜明、体积大,且目标体积较小,检测难度大幅增加。5种算法在这3幅图像的检测中均不能避免

背景中柜子的干扰,仅本文算法能够成功检测出目标区域,且检测区域相对完整,检测结果较其他算法优异。第4行图像中的目标物为最前景的白色汽车模型,该组图像背景复杂,且光照不均现象严重,极易导致误检测。RBD、LPS、MILPS算法中都将亮斑区域误检为显著区域,DSR算法和GR算法漏检

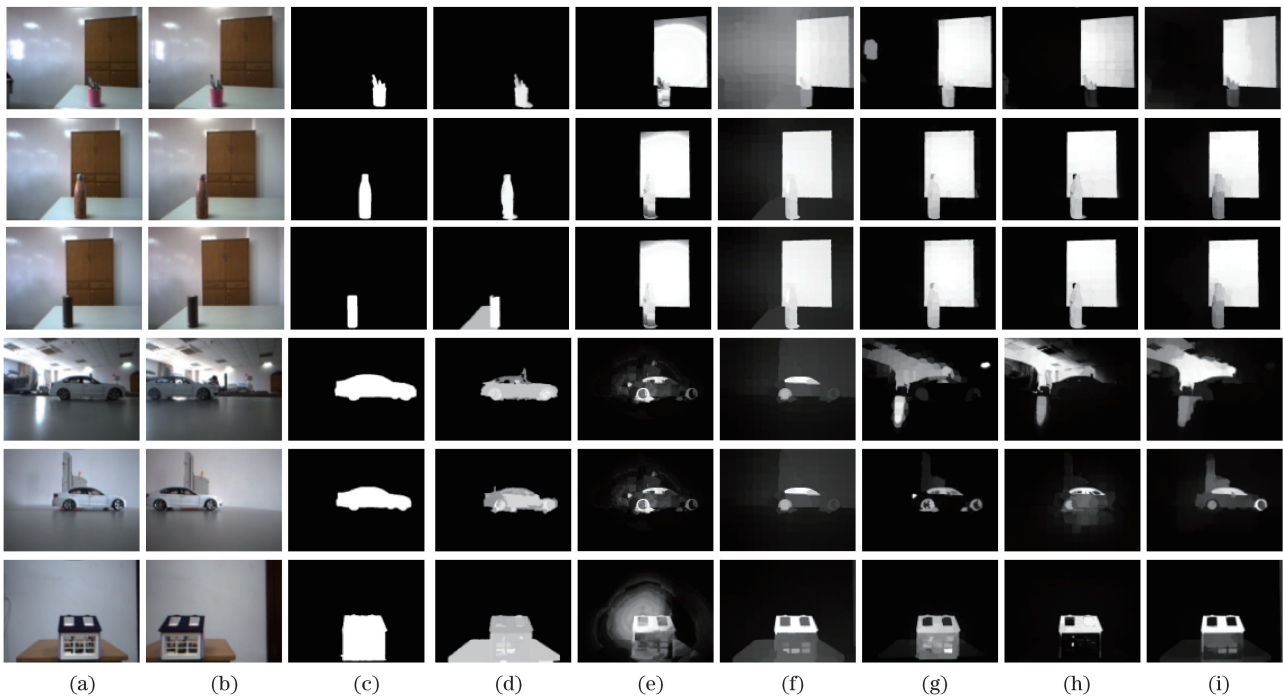


图 11 本文算法与先验信息算法显著图对比。(a)原左目图;(b)原右目图;(c) GT 算法;(d)本文算法;
(e) DSR 算法;(f) GR 算法;(g) RBD 算法;(h) LPS 算法;(i) MILPS 算法

Fig. 11 Comparison of saliency maps generated by the proposed algorithm and prior information algorithms.

(a) Original left image; (b) original right image; (c) GT algorithm; (d) proposed algorithm; (e) DSR algorithm;
(f) GR algorithm; (g) RBD algorithm; (h) LPS algorithm; (i) MILPS algorithm

情况严重,本文算法能够有效抑制背景亮斑干扰,将目标区域正确标注为显著性区域。第 5 行图像中背景与汽车模型目标颜色相似,其中浅灰色升降台增大了干扰力度。在 RBD 算法和 MILPS 算法中,均不可避免受到该升降台的影响。DSR、GR、LPS 算法标注区域缺乏目标完整性,但本文算法能够较好地抑制该升降台的干扰,并能够准确标注目标区域。第 6 行图像中目标为房屋模型,背景简单,但模型表面纹理复杂,橙黄色承载平台易与模型混为一体。DSR、GR、LPS、MILPS 算法仅能将屋顶检测为高显著区域,RBD 算法能够完整检测出房屋模型区域,但存在漏检情况,未保证目标完整性。由于承载平台是距离相机最近的区域之一,本文算法未能将其与房屋区分开,但本文算法能够将房屋区域完整标注,查全率较高。

3.2.2 定量分析

为对本文算法性能进行定量评估,列出了本文算法与 12 种对比算法对所采集的 100 组双目图像进行测试的数据分析图,包括 $P-R$ 均值曲线图、MAE、AUC、 F 值柱状图。

将测试图像从 $0 \sim 255$,步长为 5 取阈值进行二值分割,得到图 12 所示的 $P-R$ 曲线对比图。由

图 12 可知,本文算法在同一查全率下,查准率均高于其他算法,本文算法曲线图类似折线图,与其他算法曲线图形态差别大,这是由于本文算法是基于区域赋予显著值,同一区域显著值相同,故阈值选取影响范围呈区域化,在曲线中则体现为曲线折线化。

图 13 为本文算法与 12 种对比算法对 100 组测试图像进行实验所获得的 MAE、AUC、 F 值的柱状图。由图 13 可知,本文算法的 MAE 值低于 0.1,远小于 12 种对比算法测试值,表明本文算法分割结果与人工标注图相比误差小,准确度高。通过对准确率和召回率进行综合考量,本文算法 AUC 值为 0.75, F 值为 0.71,明显超过其他测试模型,表明本文算法能在保证检测目标完整性的同时保持较高的准确率,性能明显优于其他算法。

4 结 论

受人眼视觉特性的启发,提出一种基于双目视觉的显著性目标检测算法。该算法首先利用双目视觉模型进行深度感知,再将感知深度与多特征融合聚类结果进行协同处理,获取区域级的深度显著性,最后利用全局显著性与深度显著性的加权融合结果进行背景抑制,完成显著性目标检测。实验结果表

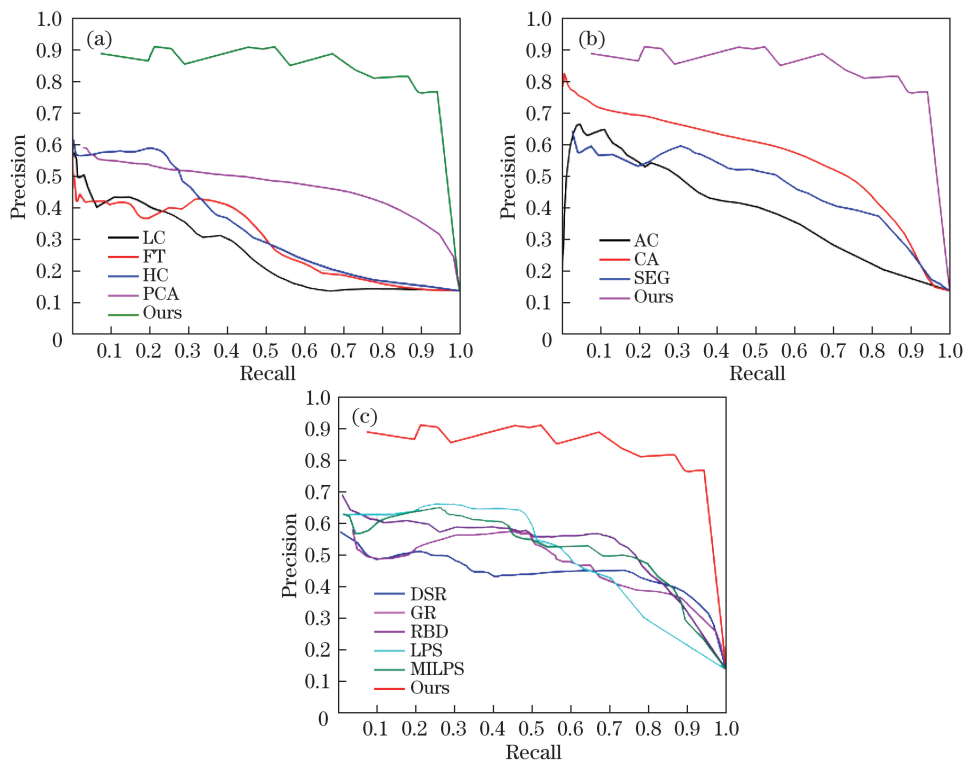


图 12 本文算法与其他算法 $P-R$ 曲线图。(a)与全局对比度算法对比;(b)与局部对比度算法对比;(c)与先验信息算法对比
 Fig. 12 $P-R$ curves of the proposed method and other algorithms. (a) Comparison with the global contrast algorithm; (b) comparison with the local contrast algorithm; (c) comparison with the prior information algorithm

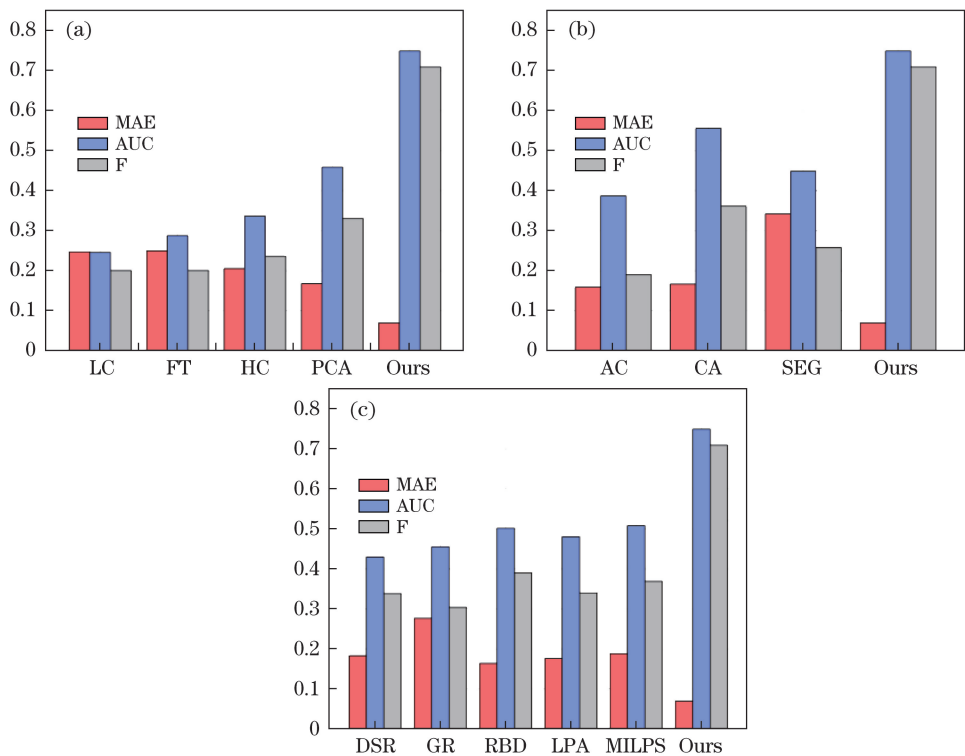


图 13 本文算法与其他算法的 MAE、AUC、F 值柱状图。(a)与全局对比度算法对比;
 (b)与局部对比度算法对比;(c)与先验信息算法对比

Fig. 13 MAE, AUC, F value histograms of the proposed method and other algorithms. (a) Comparison with the global contrast algorithm; (b) comparison with the local contrast algorithm; (c) comparison with the prior information algorithm

明,本文算法能有效抑制相似背景的干扰,较其他算法准确率更高,性能更好。但该算法不能有效抑制具有一定深度显著性的连续背景区域,针对该算法的不足,本文下一步的研究方向是抑制前后分布的连续背景,更有效地突出目标区域。

参 考 文 献

- [1] Ren Z X, Gao S H, Chia L T, *et al.* Region-based saliency detection and its application in object recognition[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2014, 24(5): 769-779.
- [2] Yang L N, An W, Lin Z P, *et al.* Small target detection based on visual saliency improved by spatial distance[J]. Acta Optica Sinica, 2015, 35(7): 0715004.
杨林娜, 安玮, 林再平, 等. 基于空间距离改进的视觉显著性弱小目标检测[J]. 光学学报, 2015, 35(7): 0715004.
- [3] Wang B, Su Y M, Wan L, *et al.* Sea sky line detection method of unmanned surface vehicle based on gradient saliency[J]. Acta Optica Sinica, 2016, 36(5): 0511002.
王博, 苏玉民, 万磊, 等. 基于梯度显著性的水面无人艇的海天线检测方法[J]. 光学学报, 2016, 36(5): 0511002.
- [4] Zhu J Y, Wu J J, Xu Y, *et al.* Unsupervised object class discovery via saliency-guided multiple class learning[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(4): 862-875.
- [5] Jiang X L, Li C H, Li X Z. Saliencybased tracking method for abrupt motions via two-stage sampling[J]. Acta Automatica Sinica, 2014, 40(6): 1098-1107.
江晓莲, 李翠华, 李雄宗. 基于视觉显著性的两阶段采样突变目标跟踪算法[J]. 自动化学报, 2014, 40(6): 1098-1107.
- [6] Yang X Y, Qian X M, Xue Y. Scalable mobile image retrieval by exploring contextual saliency[J]. IEEE Transactions on Image Processing, 2015, 24(6): 1709-1721.
- [7] Borji A, Cheng M M, Jiang H Z, *et al.* Salient object detection: a benchmark[J]. IEEE Transactions on Image Processing, 2015, 24(12): 5706-5722.
- [8] Ma Y F, Zhang H J. Contrast-based image attention analysis by using fuzzy growing[C]//Proceedings of the eleventh ACM international conference on Multimedia, 2003: 374-381.
- [9] Achanta R, Estrada F, Wils P, *et al.* Salient region detection and segmentation[C]//Proceedings of the 6th international conference on Computer vision systems, 2008: 66-75.
- [10] Zhai Y, Shah M. Visual attention detection in video sequences using spatiotemporal cues[C]//Proceedings of the 14th ACM international conference on Multimedia, 2006: 815-824.
- [11] Achanta R, Hemami S, Estrada F, *et al.* Frequency-tuned salient region detection[J]. IEEE International Conference on Computer Vision and Pattern Recognition, 2009, 22(9/10): 1597-1604.
- [12] Cheng M M, Mitra N J, Huang X L, *et al.* Global contrast based salient region detection[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2011, 37(3): 569-582.
- [13] Song T F, Liu Z Y. Saliency detection based on center rectangle composition prior[J]. Journal of Image and Graphics, 2017, 22(3): 315-326.
宋腾飞, 刘政怡. 中心矩形构图先验的显著目标检测[J]. 中国图象图形学报, 2017, 22(3): 315-326.
- [14] Li X H, Lu H C, Zhang L H, *et al.* Saliency detection via dense and sparse reconstruction[C]. IEEE International Conference on Computer Vision, 2013: 2976-2983.
- [15] Zhu W J, Liang S, Wei Y C, *et al.* Saliency optimization from robust background detection[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2014: 1049/1050(8): 2814-2821.
- [16] Zhou S J, Ren F J, Du J, *et al.* Salient region detection based on the integration of background-bias prior and center-bias prior[J]. Journal of Image and Graphics, 2017, 22(5): 584-595.
周帅骏, 任福继, 堵俊, 等. 融合背景先验与中心先验的显著性目标检测[J]. 中国图象图形学报, 2017, 22(5): 584-595.
- [17] Lin H F, Li J, Liu G D, *et al.* Saliency detection method using adaptive background template and spatial prior[J]. Acta Automatica Sinica, 2017, 43(10): 1736-1748.
林华锋, 李静, 刘国栋, 等. 基于自适应背景模板与空间先验的显著性物体检测方法[J]. 自动化学报, 2017, 43(10): 1736-1748.
- [18] Chen H Y, Qie L Z, Yang D D, *et al.* Visual background extraction algorithm based on superpixel information feedback[J]. Acta Optica Sinica, 2017, 37(7): 0715001.
陈海永, 郗丽忠, 杨德东, 等. 基于超像素信息反馈的视觉背景提取算法[J]. 光学学报, 2017, 37(7): 0715001.
- [19] Achanta R, Shaji A, Smith K, *et al.* SLIC superpixels compared to state-of-the-art superpixel methods[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(11): 2274-2282.
- [20] Song X Y, Zhou L L, Li Z G, *et al.* Review on

- superpixel methods in image segmentation[J]. *Journal of Image and Graphics*, 2015, 20(5): 599-608.
- 宋熙煜, 周利莉, 李中国, 等. 图像分割中的超像素方法研究综述 [J]. *中国图象图形学报*, 2015, 20(5): 599-608.
- [21] 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.
- [22] Rahtu E, Kannala J, Salo M, *et al.* Segmenting salient objects from images and videos[J]. *European Conference on Computer Vision*, 2010, 6315: 366-379.
- [23] Margolin R, Tal A, Zelnik-Manor L. What makes a patch distinct?[J]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2013, 9(4): 1139-1146.
- [24] Yang C, Zhang L H, Lu H C. Graph-regularized saliency detection with convex-hull-based center prior[J]. *IEEE Signal Processing Letters*, 2013, 20(7): 637-640.
- [25] Li H Y, Lu H C, Lin Z, *et al.* Inner and inter label propagation: salient object detection in the wild[J]. *IEEE Transactions on Image Processing*, 2015, 24(10): 3176-3186.
- [26] Huang F, Qi J Q, Lu H C, *et al.* Salient object detection via multiple instance learning[J]. *IEEE Transactions on Image Processing*, 2017, 26(4): 1911-1922.