

融合深度信息的室内 RGB 图像视觉显著物体快速检测方法

林 昌 何炳蔚 董升升

(福州大学机械工程及自动化学院, 福建 福州 350108)

摘要 针对传统视觉注意机制在室内三原色 (RGB) 图像视觉显著物体检测中存在的运算复杂、检测精度低等缺点, 提出了一种融合深度信息的室内 RGB 图像视觉显著物体快速检测方法。对室内 RGB 图像进行降采样和金字塔量化处理, 从而降低图片的空间分辨率和计算复杂度。利用亮度、红绿以及黄蓝三通道的多特征视觉注意机制显著性检测模型以获得室内 RGB 图像的显著图。在显著图分析中提出显著区域生长策略, 从而获得视觉显著区域的精确轮廓。融合深度信息获取视觉显著区域内显著物体数目以及显著物体相互之间的位置关系。通过室内场景实验, 验证了方法的可行性和有效性。

关键词 机器视觉; 注意机制; 显著区域; 检测

中图分类号 O436 文献标识码 A doi: 10.3788/CJL201441.1108005

An Indoor Object Fast Detection Method Based on Visual Attention Mechanism of Fusion Depth Information in RGB Image

Lin Chang He Bingwei Dong Shengsheng

(School of Mechanical Engineering and Automation, Fuzhou University, Fuzhou, Fujian 350108, China)

Abstract The traditional visual attention mechanism is complex and rough-detection for visual saliency detection indoor red-green-blue (RGB) image. In order to overcome these defects, a new fast visual saliency object detection method based on fusion depth information on indoor RGB image is proposed. A certain scale image is obtained by sub-sampling and pyramid-quantization to reduce the spatial resolution of the images so as to decrease the computational complexity. The intensity, red-green and yellow-blue three-channel features visual attention mechanism significant detection model is proposed to acquire saliency map. The saliency growing strategy is proposed to acquire the precise saliency region in the saliency analysis. The fusion depth information is utilized to detect the objects in salient region. The feasibility and effectiveness of the algorithm is verified in indoor detection experiments.

Key words machine vision; attention mechanism; significant region; detection

OCIS codes 100.3008; 150.1135; 100.1160

1 引 言

移动机器人通过检测场景中的感兴趣物体并对其定位, 再控制机器人向物体靠近, 最后对感兴趣物体进行下一步的分类或识别。所以移动机器人进行室内物体搜索的前提是知道环境中“是否存在物体”

以及“物体在哪里”。

目前视觉显著物体快速检测方法主要从两个角度研究。一个是基于高级视觉特征, 与任务、上层知识等相关的模仿人类自顶向下的视觉显著物体检测模型^[1-4], 这种方法是基于辨别显著性的检测方法,

收稿日期: 2014-04-14; 收到修改稿日期: 2014-06-08

基金项目: 国家自然科学基金(51175087)、福建省杰出青年基金(2013J06013)、福建省海外留学基金

作者简介: 林 昌(1989—), 男, 硕士研究生, 主要从事机器视觉、机器人技术等方面的研究。E-mail: linchangpt@163.com

导师简介: 何炳蔚(1973—), 男, 博士, 教授, 主要从事机器视觉、目标跟踪、逆向工程等方面的研究。

E-mail: mebwe@fzu.edu.cn

本文电子版彩色效果请详见中国光学期刊网 www.opticsjournal.net

将提取的像素块特征聚类,并建立模型匹配库作为先验知识来模拟人眼对不同物体的辨别能力。Ekvall等^[2]提出了基于感知域直方图(RFCH)的目标检测方法。该方法利用RFCH建立物体的外观模型,综合利用归一化颜色、梯度幅值及拉普拉斯等描述器,对尺度、光照以及物体遮挡等具有一定的适应性。另一个是基于低级视觉特征,直接利用场景中信息进行检测的自底向上的视觉显著物体检测模型^[5-12]。Itti等^[7-8]提出基于空间域的视觉显著物体计算模型,基本思路是依据灰度、颜色和方向等视觉特征,形成各个特征维上的特征图,然后分析、合并得到最终的显著图。文献^[13]提出情境感知(CA)显著检测方法,该模型同时考虑图像的局部特征和全局特征。克服了显著区域范围是固定模型以及区域只考虑到前景图形,忽视含有信息量的背景做法,能提取出显著区域轮廓,利于后续处理,但需要计算图像中每个像素相对于局部区域的显著性,计算量较大。

本文基于三原色深度传感器(RGB-D)获取室内场景图,提出了一种融合深度信息的室内三原色(RGB)图像视觉显著物体快速检测方法。该方法有效降低了传统视觉注意机制对图像显著性检测计算的复杂度,提高了检测精度,实现了室内显著物体快速、准确的检测目的。

2 多特征的视觉注意机制显著性检测模型

2.1 单特征视觉注意机制显著性检测模型

文献^[14]提出了基于光谱残差和傅里叶反变换的显著物体检测方法。该模型主要通过检测图像中在强度上具有突变性的区域并消除冗余的背景信息来获取显著注意区域。该模型在频域中将图像的对数谱 $L(f)$ 定义为表示图像的整体信息 $H(I_m)$ 。采用均值模板 $h_n(f)$ 平滑对数谱 $L(f)$ 得到平均谱 $M[L(f)]$,即

$$M[L(f)] = h_n(f) * L(f), \quad (1)$$

式中 $h_n(f)$ 为 $n \times n$ 的均值滤波模板,*表示卷积运算。

将对数谱和平均谱相减 $S(f) = L(f) - M[L(f)]$ 得到谱残差 $S(f)$,根据(2)式,对谱残差 $S(f)$ 和相位谱 $P(f)$ 的和值进行傅里叶反变换得到最终显著图 $S(x)$,即

$$S(x) = g(x) * F_{ITTF} \{ \exp[S(f) + P(f)] \}^2, \quad (2)$$

式中 $g(x)(\sigma=8)$ 是高斯滤波器。

2.2 多特征视觉注意机制显著性检测模型

针对文献^[14]中模型得到的显著图倾向于在边缘产生高显著性值的缺点,提出多特征视觉注意机制显著性检测模型,计算过程如图1所示。该模型主要将图像中色彩对比度大和亮度变化大的区域检测出来,获得视觉注意机制显著图。

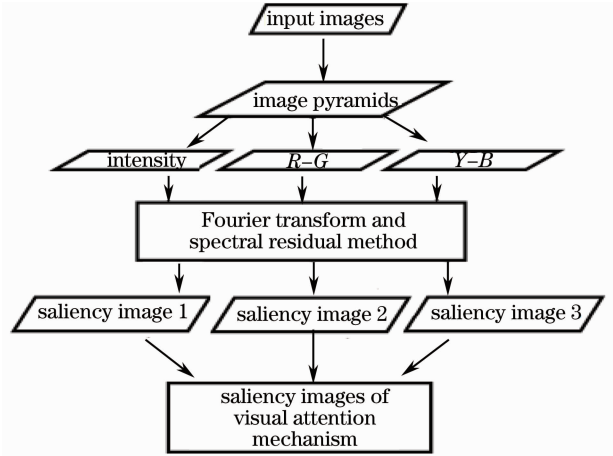


图1 显著图获得流程图

Fig.1 Flow chart of saliency map

为获得每个像素点的亮度、红绿以及黄蓝对立颜色通道特征图谱,将输入的源图像从RGB空间转换到四原色(RGBY)空间^[8]。然后根据每个像素点的 R 、 G 、 B 、 Y 颜色特征分量,计算三个通道特征图谱,即

$$I_i = (r + g + b) / 3, \quad (3)$$

$$D_{RG_i} = R_i - G_i, \quad (4)$$

$$D_{YB_i} = Y_i - B_i, \quad (5)$$

式中 (r, g, b) 为单个像素点的彩色信息,每个点的颜色特征分量: $R_i = \frac{r - (g + b) / 2}{I_i}$, $G_i = \frac{g - (r + b) / 2}{I_i}$, $B_i = \frac{b - (r + g) / 2}{I_i}$, $Y_i = \frac{r + g - 2(|r - g|) + b}{I_i}$ 。

对得到的特征图谱进行谱残差计算和傅里叶变换,得到不同显著子图 $S_1(x)$, $S_{RG}(x)$, $S_{YB}(x)$ 。最后对不同显著子图采用归一化线性合并,得到视觉注意机制显著图,即

$$S(x) = \omega_1 \times S_1(x) + \omega_2 \times S_{RG}(x) + \omega_3 \times S_{YB}(x), \quad (6)$$

式中权值 $\omega_1 + \omega_2 + \omega_3 = 1$,通常取 $\omega_1 = \omega_2 = \omega_3 = 1/3$ 。

3 视觉注意机制显著区域生长策略

传统视觉注意机制对显著图进行简单的阈值分割得到二值图,然后直接根据二值化图像获得显著目标。这种方法计算简单,但存在显著区域提取不准确和漏分割的缺点。因此,提出显著区域生长策略以实现准确分割显著区域的目的。

对显著图采用自动双阈值 Canny 边缘检测^[15],得到边缘检测图。选取显著值最高的像素点作为区域生长种子点 a ,对周围显著值小于或等于种子点的像素点进行吸收,并将吸收进来的点做为新的种子点 a_i ,当 a_i 点的显著值 S_i 满足显著区域生长终

止条件(7)式时停止生长。重复显著区域生长,直到所有显著区域被分割为止。

$$S_i \leq T, i = 1, 2, \dots, \quad (7)$$

式中 $T = \frac{\sum_{i=1}^N S_e^i(x)}{N}$, $S_e^i(x)$ 为边缘点的显著值, N 为边缘点的总个数。

按照上述方法,对图 2(b)显著图进行显著区域生长。对比传统的显著区域分割方法[图 2(c)],本文所提方法可以有效提取显著区域,并准确分割出两个物体显著区域(图 2 中红色椭圆区域)。

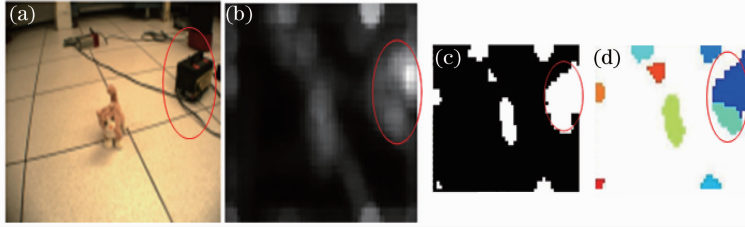


图 2 显著区域生长结果比较图。(a) 原始图; (b) 显著图; (c) 二值化图像; (d) 显著区域生长策略分割图像

Fig. 2 Comparison results of significant regional segmentation, (a) Original map; (b) saliency map; (c) binary image; (d) region segmentation image based on the proposed approach

4 融合深度信息的显著区域内物体检测方法

从图 2 可以看出,在显著区域被成功分割后,物体的轮廓、数目以及物体间的相互关系依然很难确定。因此,提出融合深度信息的显著区域内物体检测方法,实现室内显著物体快速、准确检测的目的。

根据阈值分割原理^[16],定义本文的显著物体分割判据,即(8)式。根据(8)式分割出的显著物体范围,求取显著物体像素坐标平均值,即显著物体中心坐标点 $M_i(\bar{x}_i, \bar{y}_i, \bar{z}_i)$ ($i = 1, 2$),则显著物体间关系

如(9)式。对显著物体求取深度平均值 \bar{d}_i ($i = 1, 2$),计算显著物体间的距离 $D = |\bar{d}_1 - \bar{d}_2|$ 。

$$Q = \begin{cases} 1, & |D(i) - T_d| > t \\ 2, & |D(i) - T_d| \leq t \end{cases}, \quad (8)$$

式中 $T_d = \frac{\sum_{i=1}^M D(i)}{M}$, $D(i)$ 为显著区域内像素点的深度值, M 为显著区域像素点总数, Q 为物体序号, t 为深度测量误差阈值, $t = 35$ mm。

$$\begin{cases} \bar{x}_1 > \bar{x}_2, & \bar{y}_1 > \bar{y}_2, & \text{object 2 is at the upper left of object 1} \\ \bar{x}_1 < \bar{x}_2, & \bar{y}_1 > \bar{y}_2, & \text{object 2 is at the upper right of object 1} \\ \bar{x}_1 > \bar{x}_2, & \bar{y}_1 < \bar{y}_2, & \text{object 2 is at the lower left of object 1} \\ \bar{x}_1 < \bar{x}_2, & \bar{y}_1 < \bar{y}_2, & \text{object 2 is at the lower right of object 1} \end{cases}, \quad (9)$$

其中,RGB 图像左上角起始点为坐标原点, x 轴水平向右, y 轴竖直向下。

将文献[13]提出的显著检测方法与本文的检测方法进行比较(图 3)。由图 3 可知,前者仅能获得

显著物体所在区域,而本文方法不仅在物体轮廓的获取上具有明显优势,而且还能确定出物体数目以及物体间的相互关系(指显著物体中心点间的方向关系以及距离大小),具体结果如表 1 所示。

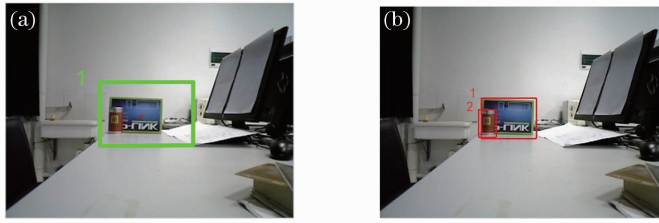


图 3 显著物体检测结果对比图。(a) 利用文献[13]方法的检测结果；(b) 本文方法检测结果
Fig. 3 Comparison of saliency target detection results by the two methods. (a) Detection results by the method in Ref. [13]; (b) saliency detection results by the proposed approach

表 1 图 3 中显著物体的相互关系
Table 1 Relationship between the saliency objects in Fig. 3

Region	Direction relations	Calculation distance /mm	Actual distance /mm	Absolute error /mm	Relative error /%
1	Object 2 at the lower right of object 1	55.6	60	4.4	7.3

5 实验与结果讨论

为验证所提方法的合理性和有效性,在 Intel Core i5-3470 CPU,内存 4G 的微机 and Windows 7 操作系统下,采用 Matlab2010a 编程。实验采用华硕 XtionProLive 型号的 RGB-D 传感器,随机选取

室内 4 个场景并分别获取 RGB 图像和深度图像,图像分辨率为 640 pixel×480 pixel,将本文方法与文献[13]方法进行比较。

5.1 显著物体检测对比实验

两种方法实验结果比较如图 4 所示。

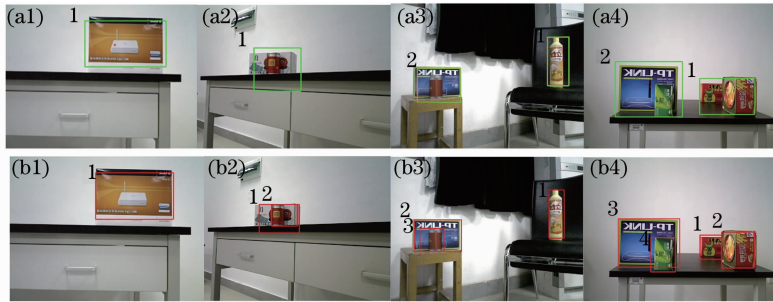


图 4 显著物体检测对比实验结果。(a1)~(a4)文献[13]方法；(b1)~(b4)本文方法；(a1)(b1)室内场景 1；
(a2)(b2)室内场景 2；(a3)(b3)室内场景 3；(a4)(b4)室内场景 4

Fig. 4 Comparison of saliency object detection results by the two approaches. (a1)~(a4) Method in Ref. [13];(b1)~(b4) proposed method; (a1)(b1) indoor scene 1; (a2)(b2) indoor scene 2; (a3)(b3) indoor scene 3; (a4)(b4) indoor scene 4

图 4 中物体数目以及相互关系的检测结果如表 2 所示,两种方法检测时间如表 3 所示。

表 2 4 个场景内显著物体间的相互关系

Table 2 Relationships between significant objects in the four scenes

Region	Direction relations	Calculation distance /mm	Actual distance /mm	Absolute error /mm	Relative error /%
2	Object 2 at the upper right of object 1	401.3	455.0	53.7	11.8
3	Object 3 at the at lower left object 2	58.5	60.0	1.5	2.50
4	Object 2 at the bottom right of object 1	356.5	400.0	43.5	10.9
	Object 4 at the bottom right of object 3	117.0	120.0	3.0	2.50

表 3 两种方法复杂度对比结果

Table 3 Comparison results of different method complexity

Method	Testing time				Average occupy memory space /MB
	Scene 1 /s	Scene 2 /s	Scene 3 /s	Scene 4 /s	
Method in Ref. [13]	24.6	29.7	53.3	28.4	227.5
Our approach	5.03	5.65	5.62	5.54	124.8

5.2 实验结果讨论

图 4 比较了文献[13]方法与本文方法的视觉显著物体的检测结果。通过比较发现文献[13]方法检测出的视觉注意区域只能获得显著物体所在的大致区域[如图 4(a1)~(a4)]。而本文方法能够突出整个显著区域并准确检测出显著物体的显著区域[如图 4(b1)~(b4)]。另外,本文方法获取显著物体轮廓的同时还能确定出物体数目以及物体间的相互关系(表 2),达到准确检测显著物体的目的。

表 3 中对比了两种方法的复杂度。在时间复杂度上,本文方法的 4 个场景检测用时明显优于文献

[13]方法(表 3)。在空间复杂度上,两种方法在微机上运行时,本文方法平均占用存储空间只是文献[13]方法的 54.9%(表 3)。

在场景 2~4 中,显著物体彼此之间存在相互遮挡,通过进一步分析发现:当物体相互遮挡率大于 90%时,本文方法会将两个物体误检测为同一个物体[如图 5(d)]。由于传感器的深度传感器由红外线发射器和红外摄像头组合而成,当物体表面被大面积遮挡时,受遮挡物影响,被遮挡物体表面上无法形成光散斑,其表面深度信息无法获得,即深度值为 0。[如图 6(d)中红色椭圆所示]。

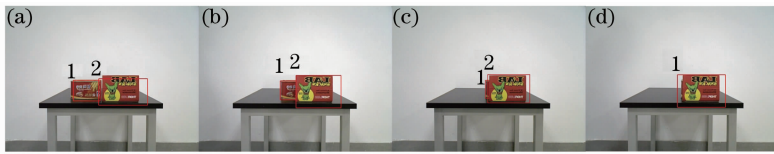


图 5 不同遮挡情况下本文方法的检测结果。(a) 遮挡率为 0%; (b) 遮挡率为 45%; (c) 遮挡率为 90%; (d) 遮挡率为 91%

Fig. 5 Detection results obtained by the proposed approach at different cover rates. (a) Cover rate of 0; (b) cover rate of 45%; (c) cover rate of 90%; (d) cover rate of 91%

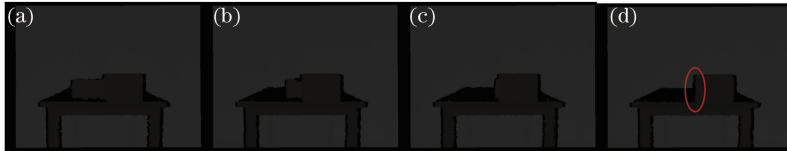


图 6 不同遮挡情况下深度图。(a) 遮挡率为 0%; (b) 遮挡率为 45%; (c) 遮挡率为 90%; (d) 遮挡率为 91%

Fig. 6 Depth maps at different cover rates. (a) Cover rate of 0 (b) cover rate of 45%; (c) cover rate of 90%; (d) cover rate of 91%

6 结 论

提出了一种融合深度信息的室内 RGB 图像视觉显著物体快速检测方法。与传统检测方法相比,通过多特征视觉注意机制显著性检测模型获得 RGB 图像的显著图,不仅降低了注意机制运算复杂度,而且提高了检测效率。此外,利用显著区域生长策略,解决了传统显著区域分割不准确和漏分割等缺点。通过融合深度信息获得显著区域内物体的数目以及相互之间的位置关系。

将该方法与文献[13]方法进行比较,实验结果表明该方法实现了室内显著物体快速、准确的检测。这也为后续物体重建和定位奠定了良好的基础。未

来将对室内非平面物体和凹凸曲面物体以及多个物体相互遮挡情况下的快速检测方法、二次视点规划方法等做进一步的研究。

参 考 文 献

- 1 H Masuzawa, J Miura. Observation planning for efficient environment information summarization [C]. IEEE RSJ International Conference on Intelligent Robots and Systems, 2009. 5794-5800.
- 2 S Ekvall, D Kragic. Receptive field cooccurrence histograms for object detection [C]. IEEE RSJ International Conference on Intelligent Robots and Systems, 2005. 84-89.
- 3 C Choi, H I Christensen. Cognitive vision for efficient scene processing and object categorization in highly cluttered environments [C]. IEEE RSJ International Conference on Intelligent Robots and Systems, 2009. 4267-4274.

- 4 K A Ehinger, S B Hidalgo, A Torralba, *et al.*. Modelling search for people in 900 scenes; A combined source model of eye guidance[J]. *Visual Cognition*, 2009, 17(6-7): 945-978.
- 5 A Borji, D N Sihite, L Itti. Quantitative analysis of human-model agreement in visual saliency modeling; A comparative study[J]. *IEEE Transactions on Image Processing*, 2013, 22(1): 55-69.
- 6 Zhang Libao, Wang Pengfei. Fast detection of regions of interest in high resolution remote sensing image[J]. *Chinese J Lasers*, 2012, 39(7): 0714001.
张立保, 王鹏飞. 高分辨率遥感影像感兴趣区域快速检测[J]. *中国激光*, 2012, 39(7): 0714001.
- 7 A Borji, L Itti. State-of-the-art in visual attention modeling[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(1): 185-207.
- 8 D Walther. *Color Opponencies for Bottom-up Attention* [D]. Pasadena; California Institute of Technology, 2006. 98-99.
- 9 Wang Weihua, Li Zhijun, He Yan, *et al.*. Real-time target detection algorithm for infrared search and track system based on region of interest extraction [J]. *Chinese J Lasers*, 2012, 39(11): 1109001.
王卫华, 李志军, 何艳, 等. 一种基于兴趣区提取的红外搜索系统目标实时检测算法[J]. *中国激光*, 2013, 39(11): 1109001.
- 10 J Theeuwes. Feature-based attention; It is all bottom-up priming [J]. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 2013, 368(1628): 20130055.
- 11 Zhang Libao. Fast detection of visual saliency regions in remote sensing image based on region growing [J]. *Chinese J Lasers*, 2012, 39(11): 1114001.
张立保. 基于区域增长的遥感影像视觉显著目标快速检测[J]. *中国激光*, 2013, 39(11): 1114001.
- 12 Zhang Jing, Shen Lansun, Gao Jingjing. Regions of interest detection based on visual attention mechanism[J]. *Acta photonica Sinica*, 2009, 38(6): 1561-1565.
张菁, 沈兰荪, 高静静. 基于视觉注意机制的感兴趣区检测[J]. *光子学报*, 2009, 38(6): 1561-1565.
- 13 S Goferman, M L Zelnik, A Tal. Context-aware saliency detection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(10): 1915-1926.
- 14 X Hou, L Zhang. Saliency detection; A spectral residual approach[C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2007. 1-8.
- 15 S J Russell, P Norvig, J F Canny, *et al.*. *Artificial Intelligence: A Modern Approach* [M]. Englewood Cliffs; Prentice Hall, 1995.
- 16 A Rosenfeld, A C Kak. *Digital Picture Processing*[M]. London; Academic Inc. (London) Ltd. Press, 2014.

栏目编辑:张浩佳