

# 背景自适应的 GrabCut 图像分割算法<sup>①</sup>

杨绍兵<sup>1</sup>, 李磊民<sup>2</sup>, 黄玉清<sup>1</sup>

<sup>1</sup>(西南科技大学 信息工程学院, 绵阳 621010)

<sup>2</sup>(西南科技大学 研究生院, 绵阳 621010)

**摘要:** 图割用于图像分割需用户交互, 基于激光雷达传感器, 提出了阈值法得到目标的外截矩形, 再映射到图像完成交互. 针对 GrabCut 算法耗时、对局部噪声敏感和在复杂背景提取边缘不理想等缺点, 提出了背景自适应的 GrabCut 算法, 即在确定背景像素中选取可能目标像素邻近的一部分像素作为背景像素, 使背景变得简单, 尤其适用于前景像素在整幅图中所占比例较小和在目标像素周围的背景相对简单的情况. 实验结果表明, 所提算法与 GrabCut 算法相比, 减少了图的节点数, 降低了错误率, 有效的提高了运行效率, 提取的目标边缘信息更加完整、平滑.

**关键词:** 图像分割; GrabCut 算法; 高斯混合模型; 激光雷达; 背景自适应

## Adaptive Background Image Segmentation Algorithm Based on GrabCut

YANG Shao-Bing<sup>1</sup>, LI Lei-Min<sup>2</sup>, HUANG Yu-Qing<sup>1</sup>

<sup>1</sup>(School of Information Engineering, Southwest University of Science and Technology, Mianyang 621010, China)

<sup>2</sup>(School of National Postgraduate academy, Southwest University of Science and Technology, Mianyang 621010, China)

**Abstract:** Graph cut needs user interaction on image segmentation, which comes up with that threshold value method, gets the objects by cutting rectangle and then we remaps to image to finish the interaction by laser radar. GrabCut algorithm is sensitive to local noise, and it is time-consuming. In addition, the edge extraction is not ideal under complex background, so an improved GrabCut algorithm is put forward to adapt background automatically in the determined background. The proposed algorithm chooses probable foreground neighboring pixels as background pixels to make background become simple. It is applicable to the case when foreground pixels account for low proportion in the whole image pixels and the background pixels are relatively simple around the foreground. Experimental results show that error rate of the proposed algorithm is reduced and the efficiency is improved in comparison with GrabCut algorithm after reducing nodes number in the graph. In addition, the edge extraction is more complete and smooth.

**Key words:** image segment; GrabCut algorithm; Gaussian mixture model (GMM); laser radar; adaptive background

图像分割是图像处理和计算机视觉领域的基础, 图像分割的算法数量众多, 其中, 图割作为一种结合图像边缘信息和纹理信息鲁棒的能量最小化方法, 得到越来越多的重视, 广泛的应用于图像分割、机器视觉等领域.

2001年 Yuri Y Boykov 和 Maric -PierrJolly<sup>[1]</sup>首次证实了离散能量函数的全局最优化能有效地用于 N-D 图像的精确目标提取, 并提出了一种交互式的基于能量最小化的二值图割算法, 利用最大流/最小割算法得

到全局最优解<sup>[2]</sup>. 许多学者对构建颜色空间、纹理形状以及信息模型和能量函数进行了改进, Han 等<sup>[3]</sup>使用多维非线性结构颜色特征取代 GMM(Gaussian Mixture Model).

2004年, Rother 等<sup>[4]</sup>提出 GrabCut 算法, 是目前目标提取最好的方法之一. Poullot 等<sup>[5]</sup>提出将 GrabCut 用于视频无监督的前景分割, 取得了很好的效果. 针对 GrabCut 耗时的缺陷, 文献[6]提出用分水岭分割把像素聚成超像素, 提高分割效率; 文献[7]提出了降低原

① 收稿时间:2016-06-06;收到修改稿时间:2016-07-25 [doi:10.15888/j.cnki.csa.005637]

图像的分辨率来加快收敛速度; 周良芬等<sup>[8]</sup>人采用分水岭降低了错误率且提高了运行效率. Hua 等<sup>[9]</sup>采用感兴趣区域(ROI)提高算法的准确率. 目前 GrabCut 在工程应用很少, 主要因为图割算法 GMM 模型的迭代求解过程复杂, 运算量大, 而且图割是一种交互式分割算法, 需要借助其他传感器, 为算法提供交互信息. 针对以上两个问题本文提出了一种用激光雷达来实现用户交互的背景自适应的 GrabCut 分割算法.

### 1 GrabCut 算法

Rother 等<sup>[4]</sup>提出了 Grab Cut 算法在 Graph cut 基础上做了一下几个方面改进: 首先, 利用 RGB 三通道 GMM 取代灰度直方图来描述背景像素和前景像素的分布; 其次, 利用迭代求取 GMM 中的各参数取代一次估计完成能量函数最小化; 最后, 通过非完全标记方法, 用户只需框选可能前景区域就可完成交互.

#### 1.1 相关 Graph Cut 分割算法

设  $G=(V,E)$  为一个无向图, 其中  $V$  是一个有限非空的节点集合,  $E$  为一个无序节点对集合的边集, 给定的待分割的图像  $I$ , 要分割出目标和背景, 通过用户交互或者其他传感器的信息确定前景和背景的种子后, 可以对应构建两个特殊的终端节点: 源节点  $S$  和汇节点  $T$ ,  $P$  为像素映射成图的节点集合, 则  $V=(S,T) \cup P$ . 分割后, 源节点  $S$  和目标节点相连, 汇节点  $T$  则和背景节点相连如图 1(c). 要转换成对边加权的图  $G$ , 将图像  $I$  每个像素映射成  $G$  中的一个节点, 像素之间的关系用图  $G$  中边的权重表示. 边分为两种, 终端节点  $S$  和  $T$  分别与像素节点连接、像素节点与像素节点连接, 分别对应的  $t$ -links 和  $n$ -links.

给每个像素  $p_i$  一个二值标号  $l_i \in \{0,1\}$ , 其中 0 代表背景背景像素, 1 代表目标像素, 则标号向量  $L=\{l_1, l_2, \dots, l_N\}$  为二值分割结果. 边的权重(代价)既要考虑两端点所对应像素的位置, 也要考虑像素间的灰度差. 为了获得最优的二值分割结果定义一个  $\lambda$  加权的区域项  $R(L)$  和边界项  $B(L)$  的组合:

$$E(L) = \lambda * R(L) + B(L) \tag{1}$$

其中:

$$R(L) = \sum_{p \in P} R_p(l_p) \tag{2}$$

$$B(L) = \sum_{(p,q)} F_{(p,q)} \delta(l_p, l_q) \tag{3}$$

$$\delta(l_p, l_q) = \begin{cases} 1 & l_p \neq l_q \\ 0 & \text{其他} \end{cases} \tag{4}$$

数据项中,  $R_p(0)$  为像素  $p$  为目标的代价,  $R_p(1)$  为像素  $p$  为背景的代价; 在边界项中, 如果像素  $p$  和  $q$  同属于目标或者背景, 则对应边的代价  $F(p,q)$  比较大; 如果不属于同类, 则  $F(p,q)$  较小. 综上所述, 边集  $E$  中各个边的权重(代价函数)如表 1 所示.

Boykov<sup>[1,2]</sup>的交互式分割过程如图 1 所示, (a) 为一个的二维图像, 将其映射为图  $G$  得到(b)图, 其中  $B$  像素表示背景种子,  $O$  表示前景种子, 由区域项表达式 (2) 得到  $t$ -links, 边界项表达式(3)得到  $n$ -links. 采用最大流/最小割得到最优解.

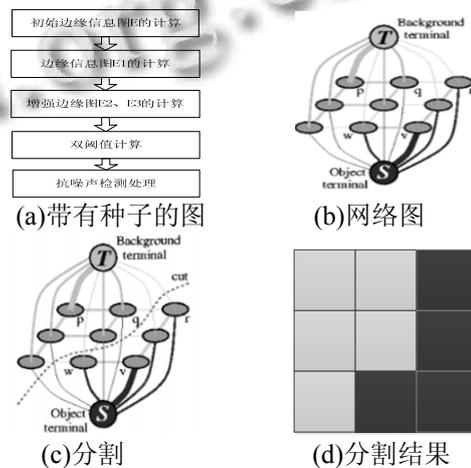


图 1 基于图割理论进行图像分割的流程

表 1 图各个边的权重

边	代价(权重)
$(p,q)$	$C_{(p,q)} \quad (p,q) \in N$
	$K \quad p \in O$
	$0 \quad p \in B$
$(S,p)$	$\lambda R_p(1) \quad p \in I, p \notin (OU B)$
	$K \quad p \in O$ $0 \quad p \in B$
$(p,T)$	$\lambda R_p(1) \quad p \in I, p \notin (OU B)$
	$0 \quad p \in O$ $K \quad p \in B$

#### 1.2 GrabCut 算法原理

GrabCut 采用 RGB 颜色空间模型, 在文献[10]中用  $K$  个高斯分量(一般  $K=5$ )的全协方差 GMM 来分别对目标和背景建模. 存在一个向量  $K=(k_1, \dots, k_n, \dots, k_N)$ ,  $k_n$  表示第  $n$  个像素的高斯分量. GrabCut 采用迭代过程使目标和背景 GMM 的参数更优, 能量函数最小; 此外, GrabCut 的交互更为简单, 只需要框出可能目标像素,

其他的视为背景像素即只需要提供框的两个斜对角坐标就能完成交互能量函数定义为式(5), 其中  $U(L, \theta, z)$  为区域项, 表示一个像素被归类为目标或者背景的惩罚;  $V(L, z)$  为边界项两个像素不连续的惩罚.  $D(l_n, k_n, \theta, z_n)$  为第  $n$  个像素对应的混合高斯建模后归为前景或者背景的惩罚.

$$E(L, \theta, z) = U(L, \theta, z) + V(L, z) \quad (5)$$

$$U(L, K, \theta, z) = \sum_n D(l_n, k_n, \theta, z_n) \quad (6)$$

GrabCut 算法步骤:

初始化:

(1) 用户直接框选可能目标得到初始的 trimap  $T$ . 框外全为背景像素  $T_B$ , 框内为  $T_U$ , 且有  $T_U = \overline{T_B}$ .

(2) 当  $n \in T_B$  则  $l_n = 0$ , 当  $n \in T_U$  则  $l_n = 1$ .

(3) 根据前景和背景的标号, 就可以估计各自的 GMM 参数.

迭代最小化:

(1) 为每个像素分配 GMM 的高斯分量:

$$k_n := \arg \min_{k_n} D_n(l_n, k_n, \theta, z_n);$$

(2) 从给定的图像数据  $z$  中, 学习优化 GMM 参数:

$$\theta := \arg \min_{\theta} U(L, K, \theta, z);$$

(3) 采用能量最小化进行分割估计:

$$\min_{\{l_n, n \in T_U\}} \min_k E(L, K, \theta, Z);$$

(4) 重复上面步骤, 直到收敛.

## 2 激光雷达和GrabCut结合

GrabCut 需要用户交互, 在实际工程运用中, 我们需借助其他传感器的信息. 本文采用 32 线激光雷达完成 GrabCut 所需的交互信息. 激光雷达和 CCD (Charge-coupled Device) 的标定不是本文重点内容, 因此假定激光雷达和 CCD 标定已经完成, 激光雷达和 CCD 图像的像素建立了映射关系.

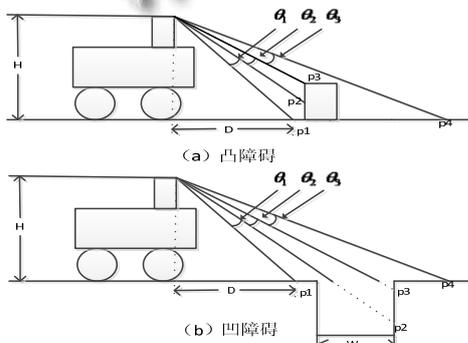


图 2 激光雷达对凹凸障碍物检测原理

如图 2 为激光对凹和凸障碍物检测的原理图,  $H$  为激光雷达相对地面的高度,  $W$  为凹障碍物的宽度,  $p_1, p_2, p_3, p_4$  在一条线上激光雷达扫描与地面相交的四个点,  $\theta_1, \theta_2, \theta_3$  为激光的发射角, 当激光雷达参数一定时  $\theta = \theta_1 = \theta_2 = \theta_3$ ;  $D$  为激光雷达到  $p_1$  的距离,  $\theta$  一定时, 随着  $D$  越大, 激光雷达两线之间的水平距离越远, 也就是说分辨率越低, 则自主机器人(如挖掘机)作业时精度不够, 因此我们把激光雷达的信息和 CCD 图像信息进行融合, 借助图像信息提高精度.

如图 3 所示,  $I$  为一幅 RGB 图, 图中小矩形代表一个像素, 深色部分为检测到的障碍物, 外部的矩形框为所需的交互信息, 由图可知完成算法交互只需求出矩形的斜对角两个坐标  $(x_1, y_1)$  和  $(x_2, y_2)$ . 激光雷达主要的作用是检测障碍物并返回可能目标框的两点坐标.

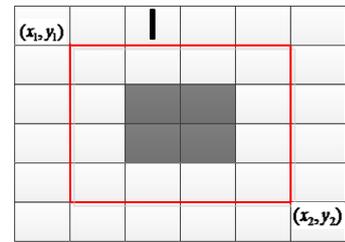


图 3 GrabCut 所需的交互信息

在获取雷达数据后, 本文采用项志宇等<sup>[10]</sup>提出的算法, 首先进行数据滤波, 数据滤波包含两个步骤. 首先, 距离值大于一定阈值的数据点认为是不可靠的点, 直接丢弃; 再采用窗口大小为 3 的中值滤波除去噪声点. 数据滤波后, 把相互之间距离差在一定的阈值范围内的数据点聚成快团. 当得到障碍物的大体轮廓后, 采用外截矩形, 再映射到图像得到图 3 中的两个坐标  $(x_1, y_1)$  和  $(x_2, y_2)$ , 完成交互.

## 3 背景自适应的GrabCut算法

针对 GrabCut 算法耗时、在复杂背景提取边缘不理想等缺点, 提出了背景自适应的 GrabCut 算法. 图割解决图像分割问题时, 需要将图像转化为网络图, 图像较大  $G$  的节点较多, 计算量变大, 因此我们可以根据可能目标像素的个数来自适应背景像素, 这样不仅减少了图  $G$  的节点数, 而且也使背景变得更加简单, 背景的 GMM 更有效, 分割效果更好.

图 4 为背景自适应的 GrabCut 算法原理, 其中  $I$  为 RGB 图像, 深色部分为障碍物. 在为改进之前,  $U$  为

可能目标像素,  $\bar{U}$  为背景像素, 交互时只需得到  $(x_1, y_1)$  和  $(x_2, y_2)$  两个坐标, 再分别进行 GMM, 可求得各像素属于目标或者背景的概率. 改进后, 就可以得到  $U$  为可能目标像素, 图中两个矩形框之间的像素集合  $B$  为背景像素. 在得到  $(x_1, y_1)$  和  $(x_2, y_2)$  两个坐标, 背景我们在此基础上横轴扩展  $m$  个像素, 纵轴扩展  $n$  个像素得到  $B$ , 设  $I$  大小为  $m_0 \times n_0$  的一副图像, 则:

$$\begin{aligned}
 m > 0, n > 0 \\
 x_1 - m \geq 0, y_1 - n \geq 0 \\
 x_2 + m \leq m_0, y_2 + n \leq n_0
 \end{aligned}$$

可得到约束条件:

$$\begin{aligned}
 0 < m \leq \min(x_1, m_0 - x_2) \\
 0 < n \leq \min(y_1, n_0 - y_2)
 \end{aligned}$$

(7)

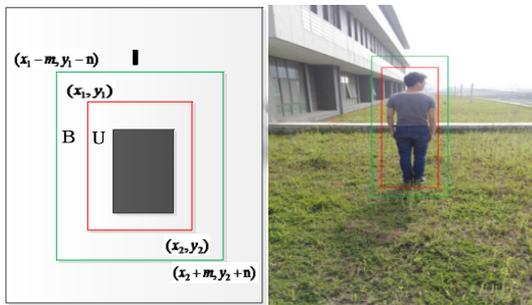


图 4 背景自适应的 GrabCut

改进后的背景根据前景的变化而变化, 从而改变了 GrabCut 的初始化.

初始化:

(1) 通过激光雷达得到  $(x_1, y_1)$  和  $(x_2, y_2)$  两个坐标, 以两个坐标画一个矩形得到  $U$ , 则矩形内为可能目标像素, 得到初始的 trimap  $T$ , 框内为  $T_U$ ,  $(x_1 - m, y_1 - n)$  和  $(x_2 + m, y_2 + n)$  两个坐标所得的矩形框内和  $U$  的差集得到  $T_B$ .

(2) 当  $n \in T_B$  则  $l_n = 0$ , 当  $n \in T_U$  则  $l_n = 1$ .

(3) 根据前景和背景的标号, 就可以估计各自的 GMM 参数.

#### 4 改进算法结果分析

本文选择了不同背景下的 3 幅图, 将改进的算法和文献[4]提出的 GrabCut 算法进行对比分析. 实验 PC 配置为 2.4GHz 的 Intel 双核 CPU4G 内存, 在 windows 平台下, 采用 Visual Studio 2012 配置 opencv2.4.9,  $m$  和  $n$  的值的选取尽量让确定背景像素(两个矩形框之间的像素)单一, 这样 GMM 参数更优, 分割效果更好. 在自动交互的情况下, 通过实验并考虑实时性和分割效

果, 得到式(8)中的  $m$  和  $n$  的值. 当然, 这里仅仅适用一般情况训练所得到的结果, 对于在特殊环境, 还需进一步实验得到参数  $m$  和  $n$  的值.

$$m = n = \min\left(\frac{|x_2 - x_1|}{4}, \frac{|y_2 - y_1|}{4}\right) \tag{8}$$

图 5 中, 图像背景较为复杂且光照较暗, 目标邻近背景相对单一且前景和背景像素区分度不大. (b)为文献[4]提出的算法分割效果较差, 边缘不完整, 人的下半身和头部信息丢失; (c)为本文提出的算法, 分割的目标更加准确, 边缘完整, 目标信息没有丢失.

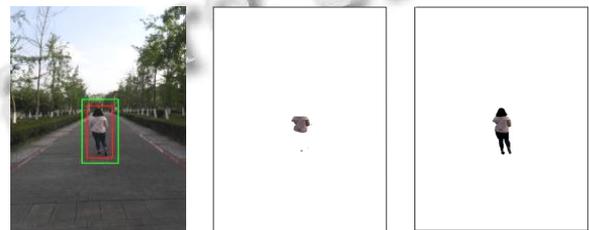


图 5 复杂背景分割结果

图 6 为背景相对单一分割的效果图, (b)和(c)分割的目标信息没有丢失信息都完整, 本文算法边缘信息更光滑.



图 6 背景单一分割结果

图 7 为草地上的分割结果, (b)和(c)总体来讲分割效果都比较好, 但由于头部周围背景相对复杂一些, 所以本文提出的算法分割的边缘更为细致.



图 7 目标像素在整幅图所占比例较大时分割结果

本文采用错误率 *error* 和分割时间两个定量指标对图像分割结果进行客观评价。假设理想分割后目标像素数量为  $N_0$ ，此处用人工手动分割取代理论上的理想分割， $N_1$  为采用其他算法分割后目标像素个数，则可得：

$$error = \frac{|N_1 - N_0|}{N_0} \quad (9)$$

由表 2 对比可知，*error* 显著的减少，由于在图 5 中，前景像素周围的背景像素单一，所以分割错误率减少的最多；由分割时间对比可知，改进的算法提高了运行效率，图 5 采用本文算法运行时间减少的最多，而图 7 减少的时间最少，可以看出本文提出的算法更适用于前景像素在整幅的像素所占比例较少的情况。

表 2 算法分割错误率和时间对比

图像	大小	Error		分割时间/s	
		GrabCut 算法	本文 算法	GrabCut 算法	本文 算法
图 6	450x690	0.4932	0.0350	9.36	1.35
图 7	400x640	0.0464	0.0336	7.74	1.56
图 8	405x679	0.0240	0.0115	8.97	4.68

综上所述，从实时性和错误率来分析，本文的算法提取的目标更加高效、省时，尤其在背景较复杂和前景像素在整幅图的像素所占比例较少时，本文的算法分割的边缘更加光滑、细致，更加节省时间。

## 5 结语

针对在工程实现中，GrabCut 算法分割时需用户交互确定部分背景和可能前景像素，本文提出了对激光雷达信息采用阈值法得到两个坐标信息，再映射到图像完成交互。针对 GrabCut 的局限性，本文提出了背景自适应的 GrabCut 自动分割算法，背景根据可能前景像素的变化而变化，减少了背景像素，从而减少了图的节点数，分割时间显著的减少。此外，通过减少背景像素同时也可以剪除复杂的背景，GMM 建模效果

更有效，错误率降低到 3.5% 以内，分割的目标细节更丰富，提取的目标更完整，同时获得更细致、平滑的边缘，通过算法分割效果分析、错误率和时间开销比较，有效的说明了本文算法的优越性。

## 参考文献

- 1 Boykov YY, Jolly MP. Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images. Proc. 8th IEEE International Conference on Computer Vision, 2001 (ICCV 2001). IEEE. 2001.105-112.
- 2 Boykov Y, Kolmogorov V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. Tissue Engineering, 2005, 11(12): 1631-1639.
- 3 Sezgin M, Sankur B. Survey over image thresholding techniques and quantitative performance evaluation. Journal of Electronic Imaging, 2004.
- 4 Rother C, Kolmogorov V, Blake A. "GrabCut": Interactive foreground extraction using iterated graph cuts. ACM Trans. on Graphics, 2004, 23(3): 307-312.
- 5 Poullot S, Satoh S. VabCut: A video extension of GrabCut for unsupervised video foreground object segmentation. International Conference on Computer Vision Theory and Applications. IEEE. 2014. 362-371.
- 6 徐秋平, 郭敏, 王亚荣. 基于分水岭变换和图割的彩色图像快速分割. 计算机工程, 2009, 35(19): 210-212.
- 7 丁红, 张晓峰. 基于快速收敛 Grabcut 的目标提取算法. 计算机工程与设计, 2012, 33(4): 1477-1481.
- 8 周良芬, 何建农. 基于 GrabCut 改进的图像分割算法. 计算机应用, 2013, 33(1): 49-52.
- 9 Hua S, Shi P. GrabCut color image segmentation based on region of interest. International Congress on Image and Signal Processing. IEEE. 2014.
- 10 项志宇. 针对越野自主导航的障碍物检测系统. 东南大学学报: 自然科学版, 2005, (A02): 71-74.