

# 基于 Mask RCNN 的目标识别与空间定位

欧攀, 路奎, 张正, 刘泽阳

(北京航空航天大学 仪器科学与光电工程学院, 北京 100191)

**摘要:** 在机械臂的自主抓取系统研究中, 为了自动获取目标物体的空间位置, 采用 Kinect 深度传感器采集 RGB 图像, 利用改进的深度学习算法 Mask RCNN 对 RGB 图像上的目标进行识别与分割, 并通过 Kinect 深度传感器模型, 将二维图像坐标转换成三维空间坐标, 对目标物体进行三维建模, 达到空间定位的目的; 通过大量数据训练的 Mask RCNN 算法, 可以同时识别多种特征差异很大的目标物体, 具有广泛的应用空间; 经过实验表明, 获得的目标物体的三维空间坐标较为准确, 且受环境影响较小, 对机械臂抓取系统的研究具有较为重要的意义。

**关键词:** 空间定位; 目标识别; 深度传感器; 深度学习; 机械臂

## Target Recognition and Spatial Location Based on Mask RCNN

Ou Pan, Lu Kui, Zhang Zheng, Liu Zeyang

(School of Instrumentation Science and Opto-Electronics Engineering, Beihang University, Beijing 100191, China)

**Abstract:** In the research of autonomous grasping system of manipulator, in order to acquire the spatial coordinates of the object automatically, Kinect depth sensor is used to collect RGB image, improved depth learning algorithm Mask RCNN is used to recognize and segment the target on RGB image, and through the Kinect depth sensor model, the two-dimensional image coordinates are transformed into three dimensional space coordinates, and the object is modeled in three-dimensional to achieve the purpose of spatial positioning. Mask RCNN algorithm trained by a large amount of data can recognize many objects with different features at the same time, so it has wide application space. Experiments show that the three-dimensional coordinates of the target object are more accurate and less affected by the environment, it is of great significance to the research of the manipulator grasping system.

**Keywords:** spatial positioning; target recognition; depth sensor; deep learning; mechanical arm

## 0 引言

目前工厂车间对于机械臂的使用越来越广泛, 机械臂的自主抓取大多是基于视觉的控制, 如视觉定位、视觉伺服<sup>[1]</sup>, 但是, 当环境较复杂时, 就会导致识别定位精度大大下降。所以, 精确的确定目标物体的位置是急需解决的问题。微软公司研发的 Kinect 深度传感器能同时获取周围空间的彩色二维图像及深度信息, 通过深度传感器自身的参数可以求出空间三维点云<sup>[2]</sup>, 且价格便宜, 因此广泛应用于机器人系统的研究中。蒋焕煜<sup>[3]</sup>等使用双目视觉传感器对成熟番茄进行定位, 采用自适应阈值法对 RGB 图像进行分割, 并将形心匹配与区域匹配相结合来计算目标番茄的位置信息。该方法精度较高, 对指导机械臂采摘成熟番茄具有重要的作用, 但是双目视觉传感器标定较为繁琐, 计算复杂, 而且只能可见光的环境下进行识别。钟泽宇<sup>[4]</sup>等使用 Kinect 深度传感器同时采集 RGB 图和深度图, 在 RGB 图上通过融合 HSI 色彩空间和 Lab 色彩空间的分割结果, 分割出包含西红柿的区域, 并将该区域转换成三维空间点云, 计算点云的重心达到对成熟番茄定位的目的。该

方法对西红柿的定位效果较好, 但只适用于像西红柿等颜色特征较为明显的简单目标物体, 很难扩展到实际工厂使用场景中。吴帅<sup>[5]</sup>等根据深度传感器模型, 将 Kinect 采集的深度图转换成空间点云, 然后通过限制深度阈值和颜色信息将人体的点云数据分割出来, 达到对人体的重构和定位。该方法使用简单, 计算量小, 不需要对相机进行标定, 但在使用过程中需要限制人体与深度传感器之间的距离, 使用的局限性较大。

直接在三维点云中分割出目标物体, 误差大, 处理速度慢, 不能够满足实时的要求, 本文仍然先从二维图像出发, 先在二维彩色将目标物体识别出来, 再进行空间定位。相较于传统的手工提取特征的算法, 深度学习算法的优势在于特征提取环节不需要使用者预先选定提取何种特征, 而是采用一种通用的学习过程, 使模型从大规模数据中学习进而学得目标具备的特征。王鹏<sup>[6]</sup>等利用深度学习算法进行了人体图像的识别与分割, 识别精度达到 88.77%, 单张图片分割耗时约为 60.7 ms, 基本上满足实时性的要求。伍锡如<sup>[7]</sup>等通过 12 层的深度学习算法, 从众多物体中识别出了中国象棋, 识别精度高达 99.6%。本文先使用改进的 Mask RCNN 深度学习算法从 RGB 图像中识别并分割出目标物体的区域, 然后结合深度传感器的数学模型, 将分割得到的区域像素坐标转换成三维空间坐标, 达到分割点云的目的, 完成对物体空间定位的任务。通过实验结果分析,

收稿日期: 2018-11-12; 修回日期: 2018-12-13。

基金项目: 国家自然科学基金(61675031)。

作者简介: 欧攀(1977-), 男, 博士, 副教授, 主要从事 3D 扫描成像、光纤激光器、光纤传感器、激光雷达等方面的研究。

本文提出的方法可以同时识别出特征差异很大的目标物体, 且满足实时性的需求, 为机械臂的自主识别抓取提供了很好的理论基础。

## 1 测量仪器及原理简介

### 1.1 实验设备简介

本文使用的实验设备是由微软生产的 3D 视觉传感器 Kinect, 该设备同时含有 RGB 彩色传感器和红外传感器, 有效的探测距离大约为 0.8~3.5 m。Kinect 可以同时获取场景内的深度图及彩色图, 并且可以根据彩色信息和深度信息完成具有颜色信息的点云数据的生成。

深度传感器是利用光编码技术对目标空间的深度信息进行计算。该技术通过红外传感器结合红外光源获取深度信息, 在传感器有效距离之内, 每隔一段距离取参考平面, 记录下参考平面上的红外光源产生的衍射散斑图案, 当光源在待测物体上产生衍射图样后, 将该图像与几百张参考图像做运算得到一系列相关度图像。依据相关图像中的峰值从而得到测量图像的深度信息<sup>[8]</sup>。

### 1.2 深度传感器模型

空间定位主要是通过获取的点云数据来进行计算的, 和普通相机一样, 对于获取的深度图上的每个像素坐标都可以通过针孔摄像机模型计算得到其对应的三维坐标值。针孔相机模型<sup>[5]</sup>如式 (1) 所示:

$$z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{f_x}{dx} & 0 & u_0 \\ 0 & \frac{f_y}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} [R \quad t] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (1)$$

其中:  $u, v$  是深度图像坐标,  $u_0, v_0$  是图像坐标系的原点,  $f_x, f_y$  为相机镜头  $x, y$  方向上的焦距,  $dx, dy$  是  $x, y$  方向上单位像素的长度,  $R, t$  为摄像机坐标系到世界坐标系的旋转矩阵与平移矢量, 而深度传感器的摄像机坐标系与世界坐标系重合, 因此  $R$  为单位阵,  $t$  为零矢量; 深度值  $z_c$  与世界坐标系中的  $Z$  坐标相同。因此式 (1) 可化为:

$$z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{f_x}{dx} & 0 & u_0 \\ 0 & \frac{f_y}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ z_c \\ 1 \end{bmatrix} \quad (2)$$

由式 (2) 可计算出物体三维坐标为:

$$\begin{cases} X = \frac{(u - u_0) \cdot z_c \cdot dx}{f_x} \\ Y = \frac{(v - v_0) \cdot z_c \cdot dy}{f_y} \\ Z = z_c \end{cases} \quad (3)$$

因此, 利用摄像机模型就可以将深度图上的二维点坐标转换成三维空间点, 通过对空间点云的处理即可完成空间定位的目的。

### 1.3 目标识别与分割

本文主要是通过 RGB 图中识别并分割出目标物体,

间接达到分割点云的目的。因此, 在 RGB 图像中对目标物体识别分割的精度直接影响着后续点云的分割。传统的图形分割方法大都是基于图像图形学方法的, 如 Chuang<sup>[9]</sup>等利用颜色直方图, 通过双局部阈值变换实现了图像的语义分割, 这种分割方法对颜色特征有较高的要求, 颜色不明显时很难做到精确的分割。后来, 浅层机器学习算法得到了发展, 算法可以学习到一些浅层的特征, 分割效果较图像图形学的分割方法得到了提升。如毛凌<sup>[10]</sup>等提出一种改进的条件随机场模型, 完成了目标检测、与图像分割的任务。浅层的机器学习需要借助人工设计的部分特征才能完成识别的任务, 很难学习到物体深层的特征, 且受人的经验影响。

随着卷积神经网络<sup>[11]</sup>的提出, 深度学习算法得到了迅速的发展, 它能够通过大量的数据学习到物体的深层特征, 而不需要人为的指定特征。本文采用的深度学习分割算法是由 Facebook 提出 Mask RCNN<sup>[12]</sup> 分割算法, 算法的结构和流程图如图 1 所示, 输出包含三个分支, 分别是目标的分类、目标包围框的坐标及目标物体的二值掩码。其中分类和回归部分都是由基于区域的目标检测网络 Faster RCNN<sup>[13]</sup>完成, 目标物体的像素级分割由全卷积神经网络 FCN<sup>[14]</sup>完成。从图 1 可以看出, 对于输入的图片, 首先进行的是将图片输入到卷积神经网络中进行特征提取, 对于得到特征图 (Feature Map) 通过目标估计网络 (RPN) 网络提取可能存在的目标区域 (region of interest, ROI)。这些 ROI 经过 ROI Align 层被映射成固定维数的特征向量, 其中两个分支经过全连接层进行分类和包围框的回归, 另一分支经过全卷积进行上采样得到分割图。训练采用的多任务损失函数, 通过学习不断的减小损失函数的值, 直到获得全局最优解。损失函数的公式如式 (4), 式中的三项分别为分类误差、包围框误差和分割误差。

$$L = L_{cls} + L_{box} + L_{mask} \quad (4)$$

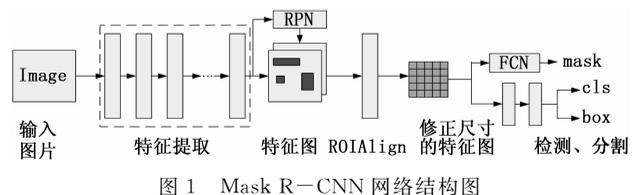


图 1 Mask R-CNN 网络结构图

Mask RCNN 算法在提取 ROI 时采用的是 RPN 网络, RPN 网络可以接受任意大小的输入图片, 输出为可能存在目标的区域。Mask RCNN 为了减少计算量, 采用共享卷积的形式进行 ROI 的提取, 即输入 RPN 网络的图片为经过卷积网络的最后一层输出的特征图, RPN 的运行机制如图 2 所示, 运行  $3 \times 3$  的滑动窗口在特征图上进行遍历, 为了估算原始目标在每个滑动窗口上位置上的大小, 定义了  $k$  种不同形状的包围框, 根据目标物体大小的不同选择不同尺寸的包围框, 本文采用的是  $128 \times 128$ 、 $256 \times 256$ 、 $512 \times 512$  三种包围框大小, 长宽比设置为 1: 1、1: 2、2: 1 三种比例, 共九种包围框。通过非极大值抑制的方法, 挑选

出最有可能存在目标的区域，随后对该区域进行检测与分割。

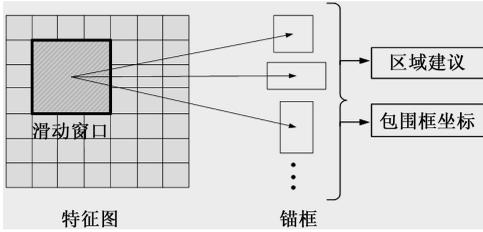


图 2 RPN 运行机制

RPN 网络采用的是固定大小的滑动窗口的方式在特征图上进行遍历，固定的滑动窗口在实际使用中可能存在一定的限制，因为目标可能大也可能小，同时窗口的滑动会到来很多重复的估计目标。基于以上问题提出了一种改进的 RPN 网络，其结构如图 3 所示，将固定滑动窗口操作改成固定尺寸分割策略，分割后的每一个分割窗口同样经过不同尺寸的锚框进行 ROI 提取，之后的操作和 RPN 一样。由于采用非滑动窗口的分割方法，改进的 RPN 产生了更少的估计目标区域，如按图 3 的方式进行，产生的估计目标区域个数为  $k * (2 * 2 + 3 * 3 + 5 * 5)$ ，多种分割方法相当于构建了多种尺度的 RPN，效率更高，更节省时间。

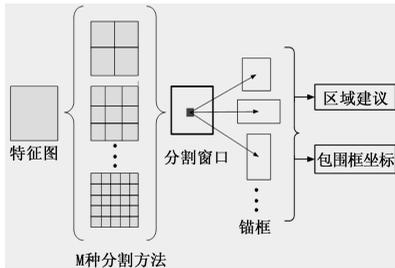


图 3 改进的 RPN 运行机制

## 2 实验结果及分析

### 2.1 目标识别与分割

在使用改进 Mask R-CNN 网络前，需要先对该网络进行训练。由于人工采集的数据集较少，为了防止算法在训练过程过拟合，先是基于 MS COCO 公用数据集进行预训练，样本数量约 8 万多张，并且每张图片中的物体类别都进行了语义标注，很适合用于图像分类与语义分割。使用 Kinect 分别采集了圆柱形物体、球形物体、锥形物体和工厂常见的其他目标物体，并用 LableMe 进行人工标记得到标签图。本文首先采用的数据增广的方法是对原数据进行随机旋转一定的角度形成新的图像，通过对原图进行随机旋转可以部分消除样本多样性不足的问题。采用的另一种数据增广的方法是改变图像 RGB 通道的强度，目的是为了消除光照强度和颜色变化对实验结果带来的影响，具体的做法是对 RGB 图像的像素值做主成分分析，然后对图像中的每一个像素添加一个随机倍数的主成分<sup>[15]</sup>。

$$[p_1, p_2, p_3][\alpha_1 \lambda_1, \alpha_2 \lambda_2, \alpha_3 \lambda_3]^T \quad (4)$$

式中， $p$  和  $\lambda$  分别为图像 RGB 的像素值的  $3 * 3$  协方差矩阵的特征向量和特征值， $\alpha$  是随机变量，且随机性符合平均值为 0，标准偏差为 0.1 的高斯随机分布。

由于采集设备的原因，采集到的设备会存在一定的图像噪声，所以在将图像输入到训练模型之前，进行简单的中值滤波操作。训练过程分为三个阶段，第一阶段采用学习率为 0.001 训练头部输出网络层，第二阶段训练以学习率为 0.001 训练四层以后的网络层，第三阶段将学习率降低 10 倍训练整个网络层，直到训练的误差不再变化。其过程如图 4 所示。

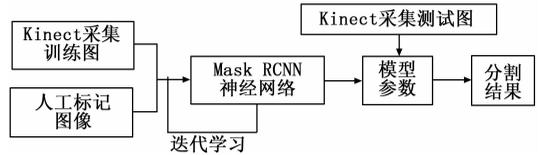


图 4 图像分割实验流程图

为了准确客观地评价所得分类器的性能，用重叠率来度量图像最终识别和分割的精度，重叠率的计算如式 (5)。其中， $A_p$  为 Mask R-CNN 网络预测分割的区域， $A_{GT}$  为目标的实际区域。

$$score = \frac{area(A_p \cap A_{GT})}{area(A_p \cup A_{GT})} \quad (5)$$

为了验证改进 Mask R-CNN 对于目标分割的可靠性、准确性与实时性，本实验分别在正常光照、黑暗光照、强光照、多色彩组成的物体（为了验证不受颜色的干扰）及不同拍摄距离、不同程度的遮挡、不同背景情况下，分别用深度传感器采集多幅图片进行识别与分割，正常光照下得到的实验结果如图 5 所示，以及在 GTX 1080 显卡下的平均准确率及消耗时间如表 1 所示。



(a) 输入图像 (b) 输出图像

图 5 目标识别结果。

表 1 不同环境下算法的准确率与平均消耗时间

实验环境	目标数量	消耗时间/s	准确率/%
正常光照	3	0.432	98.2
暗光	3	0.478	92.4
强光	3	0.531	88.1
多色彩物体	3	0.473	88.4
不同的距离拍摄	3	0.492	93.2
复杂的背景	3	0.538	91.2

通过表 1 数据可以看出，改进的 Mask RCNN 网络在不同条件下，对目标的识别精度相对较高，受环境影响较小，不会因为光线变暗等环境因素导致识别精度变得很低。经

分析可知, Mask RCNN 模型是通过大量数据学习到目标物体的特征, 而不是人为设定的, 这样学来的特征更能适应环境的变化, 确定就是需要准备大量的数据, 在数据量不足的情况下, 准确率可能会下降。

经过改进的 Mask RCNN 算法之后会得到三个输出, 分别是目标的类别、包围框的位置及物体的掩码。为了得到目标物体所对应的深度图上的区域, 需要对分割算法的效果图进行进一步的处理。通过提取被识别到的物体的掩码, 并将背景的像素值设为 0, 得到的效果如图 6 所示。



(a) 目标1分割结果; (b) 目标2分割结果; (c) 目标3分割结果

图 6 目标分割结果

### 2.2 点云分割

Kinect 可以同时采集同一场景下的深度图和 RGB 图, 并且像素坐标是对应的。通过图 6, 可以得到目标物体在深度图对应的区域。利用式 (3) 可以将深度图中的点映射到三维空间中, 将每个二维点转换成基于 Kinect 坐标系的三维空间点, 得到的空间点云数据可视化如图 7 所示。

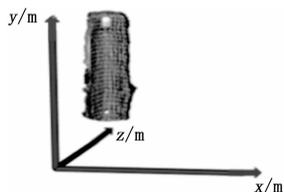


图 7 目标 2 点云分割的结果

### 2.3 实验结果及误差分析

本实验通过计算分割获得的点云的重心对目标物体进行定位, 定位的误差定义为:

$$\Delta = (x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2 \quad (6)$$

手动测量的精度为 1 mm, 深度传感器的测量单位为 mm, 测量结果保留一位小数, 实验测量与手动测量的平均相对误差  $\sigma$  为:

$$\sigma = \frac{\sum_{i=1}^n \Delta}{n(|x| + |y| + |z|)} \quad (7)$$

通过多次实验, 得到的数据如表 2 所示。

表 2 自然光下手动测量与深度传感器测量的结果

手动/cm	深度传感器/cm	误差/cm	相对误差/%
(10.2, 2.1, 83.2)	(10.8, 2.6, 82.8)	1.5	1.57
(18.4, 6.8, 92.8)	(17.9, 6.0, 94.2)	2.7	1.84
(-20.1, 4.2, 100.8)	(-22.0, 4.8, 101.5)	2.2	1.76
(26.3, 3.6, 104.0)	(26.9, 4.3, 103.0)	2.3	1.71
(-10.4, 10.8, 120.4)	(-11.2, 10.1, 121.1)	2.2	1.55

如表 3 所示, 使用了几种传统方法和基于卷积神经网络的 Mask RCNN 算法、改进的 Mask RCNN 算法, 在正常光照下, 对形状不一的工件进行了 50 次对比试验, 统计了平均识别率, 定位的精度和消耗的时间。

表 3 不同的方法性能对比

方法	识别率/%	定位精度/cm	消耗时间/s
Kinect+HSI+Lab	85.2	5.3	0.83
双目视觉+Hausdorff	92.3	3.3	1.45
Mask RCNN	98.0	2.3	0.45
改进的 Mask RCNN	98.5	2.3	0.23

通过表 3 可以发现, 文献 [3] 提出的基于 HSI 和 Lab 色彩空间的定位方法与原文相比识别率有所下降, 主要原因是 HSI 和 Lab 只能对颜色特征比较明显的物体识别率比较好, 不具有广泛的实用性。传统使用的比较多的双目视觉方法比较复杂, 用起来繁琐, 而且基于 Hausdorff 距离的相似度匹配需要提前制作模板库, 通过计算大量的距离匹配到目标, 比较消耗时间。本文所使用的 Mask RCNN 深度学习算法能从大量数据中学习大目标的深层次特征, 相较于人工设计的特征更能表达物体之间的差别, 因此受外界因素影响较小, 而且经改进的 Mask RCNN 算法识别时间在 0.25 s 左右, 基本满足实时性, 具有很好的使用价值。

### 3 结论

本课题使用了深度传感器与深度学习算法相结合, 对目标物体进行识别、分割与定位, 为机械臂实现自主识别抓取提供了较好的技术指导。根据机械臂实际可能抓取物体种类的需求, 提前训练深度学习算法, 使其学习目标物体深层次的特征, 然后结合深度传感器重建出目标物体的点云, 达到对目标物体空间定位的目的, 相比于传统的基于人为指定特征进行分割具有较好的实用性。实验证明定位误差在 2% 以内, 可行性较好, 对机械臂抓取系统的研究具有很重要的作用。

#### 参考文献:

[1] 方勇纯. 机器人视觉伺服研究综述 [J]. 智能系统学报, 2008, 3 (2): 109-114.

[2] Shum H P H, Ho E S L, Jiang Y, et al. 3D Reconstruction using the Kinect sensor and its application to a visualization system [A]. IEEE International Conference on Systems, Man, and Cybernetics [C]. 2012: Paper 6378311.

[3] 蒋焕煜, 彭永石, 申川. 基于双目立体视觉技术的成熟番茄识别与定位 [J]. 农业工程学报, 2008, 24 (8): 279-283.

[4] 钟泽宇, 周海婷, 古炜豪, 等. 基于 Kinect 传感器的机械手自主抓取技术研究 [J]. 仪表技术与传感器, 2016, 17 (7): 123-126.

[5] 欧攀, 吴帅, 周错. 基于深度传感器骨骼追踪的快速人体测量方法 [J]. 激光及光电子进展, 2017, 54 (7): 121206.

[6] 王 鹏, 方志军, 赵晓丽, 等. 基于深度学习的人体图像分割算法 [J]. 武汉大学学报 (理学版), 2017, 63 (5): 466 - 470.

[7] 伍锡如, 黄国明, 孙立宁. 基于深度学习的工业分拣机器人快速视觉识别与定位算法 [J] 机器人, 2016, 38 (6): 711 - 719.

[8] Gonzalez-Jorge H, Riveiro B, Vazquez-Fernandez E, et al. Metrological evaluation of Microsoft Kinect and Asus Xtion sensors [J]. Measurement, 2013, 46 (6): 1800 - 1806.

[9] 林 川, 潘盛辉, 韩峻峰, 等. 基于人类视觉系统的交通标志优化分割方法 [J]. 武汉大学学报 (理学版), 2011, 57 (3): 236 - 240.

[10] 毛 凌, 谢 梅. 基于高阶 CRF 模型的图像语义分割 [J]. 计算机应用研究, 2013, 30 (11): 3514 - 3517.

[11] Lawrence S, Giles C L, Tsoi A C, et al. Face recognition: a convolutional neural network approach [J]. IEEE Trans on

Neural Net-work. 1997, 8 (1): 98 - 113.

[12] He K, Gkioxari G, Dollar P, et al. Mask R-CNN [A]. IEEE Conference on Computer Vision and Pattern Recognition [C]. Washington, DC: IEEE Computer Society, 2018. 1.

[13] Ren Shaoqing, He Kaiming, Ross Girshick, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [A]. IEEE Conference on Computer Vision and Pattern Recognition [C]. Washington, DC: IEEE Computer Society, 2017: 1506 - 1519.

[14] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation [A]. IEEE Conference on Computer Vision and Pattern Recognition [C]. Washington, DC: IEEE Computer Society, 2015: 3431 - 3440.

[15] 周云成, 许童羽, 郑 伟, 等. 基于深度卷积神经网络的番茄主要器官识别分类方法 [J]. 农业工程学报, 2015, 33 (15): 219 - 226.



(上接第 166 页)

$$\Delta \epsilon = \frac{\Delta \delta}{360} \times \Delta \theta \tag{11}$$

将实验结果进行拟合, 得出不确定度, 由图可知, 随着速度的增加产生的滑动可能性越大, 大半径转弯比小半径转弯的滑动偏差要小。通过编码器数值, 对滑动偏差进行分析, 可以帮助机器人确定在定位中的不确定度, 通过多传感器如激光雷达、编码器、惯性测量元件等, 根据卡尔曼滤波理论, 对先锋机器人进行定位, 进而使定位结果更合理。通过 KNN 算法进行路面识别, 离线数据进行在线预测误差范围。

通过实验仿真我们得到了里程计在滑移转向机器人运动中的不确定度误差模型, 并得出当左右两轮相对速度过大时产生的误差比较大。在导航时, 为保证导航的定位精度, 我们应该避免左右两轮相对速度差过大。在实际应用过程中, 通过惯性测量元件及激光雷达与里程计得到的位置数据进行比较, 可以得出当前路面状况与里程计不确定度的离线数据, 为滑移转向机器人定位导航提供依据。

参考文献:

[1] Thrun S, Burgard W, Fox D. 概率机器人 [M]. 曹红玉等译. 北京: 机械工业出版社, 2017.

[2] 杨晶东, 杨敬辉, 洪炳熔. 一种有效的移动机器人里程计误差建模方法 [J]. 自动化学报, 2009, 35 (2): 168 - 173.

[3] Wu Y, Wang T M, Liang J H, et al. Experimental kinematics modeling estimation for wheeled skid-steering mobile robots [A]. IEEE International Conference on Robotics and Biomimetics [C]. 2014.

[4] 吴 耀, 王田苗, 王晓刚, 等. 基于自适应卡尔曼滤波的侧滑移动机器人运动模型估计 [J]. 电子与信息学报, 2015, 37 (12): 3016 - 3024.

[5] 郭晓林, 潘家平, 于诏城. 轮式车辆稳态滑移转向特性研究

[J]. 农业装备与车辆工程, 2012, 50 (12): 27 - 30.

[6] 陈晋市. 滑移装载机行走系统研究 [D]. 吉林: 吉林大学, 2012.

[7] 邵春梅. 路面摩擦系数检测方法及其纵、横向摩擦系数关联性研究 [D]. 西安: 长安大学, 2011.

[8] 张高峰. 轮式滑移装载机转向过程分析及仿真 [D]. 太原: 太原科技大学, 2014.

[9] 崔明月, 孙建华, 李永福, 等. 轮子纵向打滑条件下的移动机器人自适应跟踪控制 [J]. 控制与决策, 2013, 28 (5): 664 - 670.

[10] 安 雷, 张国良, 张维平, 等. 移动机器人扩展卡尔曼滤波定位与传感器误差建模 [J]. 信息与控制, 2012, 41 (4): 406 - 412.

[11] Reinstein M, Kubelka V, Zimmermann K. Terrain adaptive odometry for mobile skid-steer robots [A]. IEEE International Conference on Robotics & Automation [C]. Karlsruhe, Germany, 2013: 4691 - 4696.

[12] Weiss C, Fechner N, Stark M, et al. Comparison of different approaches to vibration-based terrain classification [A]. The European Conference on Mobile Robots [C]. Freiburg, Germany, 2007.

[13] Tick D, Rahman T, Busso C, et al. Indoor robotic terrain classification via angular velocity based hierarchical classifier selection [A]. IEEE International Conference on Robotics & Automation [C]. River Centre, Saint Paul, Minnesota, USA, 2012: 3594 - 3600.

[14] 宋 宇, 陈无畏, 陈黎卿. 基于 ADAMS 与 Matlab 的车辆稳定性控制联合仿真研究 [J]. 机械工程学报, 2011, 47 (16): 86 - 92.

[15] 张高峰. 轮式滑移装载机转向过程分析及仿真 [D]. 太原: 太原科技大学, 2014.

[16] 李 强. 基于振动信号的轮式机器人地面分类方法研究 [D]. 哈尔滨: 哈尔滨工程大学, 2013.