

LIDAR-based 3D Object Perception中文翻译

translation

摘要

本文描述了一种基于激光雷达的地面机器人移动感知系统，包括三维物体检测，分类和跟踪。所提出的系统在我们的自动地面车辆上进行了演示MuCAR-3，使其能够在类似城市交通的场景以及越野车队场景中安全导航。我们的方法的效率源于2D和3D数据处理技术的独特组合。尽管在2.5D占用网格中将点云快速分割成对象，但是在原始3D点云上对对象进行分类。对于域的快速切换，占用网格被增强以充当用于检索3D点的哈希表。与大多数现有的3D点云分类工作相比，实时操作通常是不可能的，这种组合使我们的系统能够以0.1s的帧速率实时执行。

1. 介绍

在本文中，我们解决了将3D扫描数据分割为已知类对象的问题。给定由范围扫描仪获取的3D点集，分割的目标是将点归因于一组候选对象类。在地面机器人移动性的背景下，这种分割能力不仅对于场景理解和规划等高级任务是必不可少的，而且还可以用于扫描配准和机器人定位，例如，在SLAM框架中[1]。此外，了解对象的类在动态环境中特别有用，无论是用于规划还是估计：通过使用适当的动态模型可以改进估计，并且规划可以包含关于特定对象类的典型行为或意图的知识。

我们的感知方法被分解为三个主要步骤：分割，分类和跟踪。分段步骤在占用网格上执行，产生不属于地面的网格单元的连接组件。在有效的操作中，我们确定与分割的对象相对应的所有3D LIDAR点测量。在分类步骤中，我们从对象的点云中提取特征，捕获在每个点周围的固定大小支撑体积上提取的局部空间和反射率属性的分布。在监督学习框架中，训练支持向量机（SVM）分类器以区分感兴趣的类别，例如，我们案例中的其他流量参与者，给出了手工标记的点云示例。

该方法不限于特定的机器人或传感器，但是我们使用我们的车辆MuCAR-3（慕尼黑认知自主机器人车，第3代），配备Velodyne HDL-64 LIDAR的VW Touareg来描述和演示它（见图

1)。

A. 相关工作

随着距离扫描设备成为移动机器人的标准设备，3D扫描分割和分类的任务是增加实际相关性的任务之一。

有趣的是，尽管范围扫描仪是DARPA城市挑战赛2007的主要传感器，但分割主要是在2D占用网格上完成的。如果有的话，分割对象的分类是在2.5D域中完成的，通过拟合L形或边界框并根据简单规则验证它们[2]，[3]。由于严格的竞赛规则，分类可能被省略，这确保了在道路边界内检测到的每个物体只能对应于另一辆车。

相比之下，Anguelov等。人。[4]和Lalonde等。人。[5]描述了为扫描的每个点分配类标签的方法。给定标记的点云，然后对扫描进行分段是直截了当的。虽然为每个点提取的特征没有显着差异 - 两种方法都使用局部点云统计进行特征提取，稍后将详细说明 - 遵循不同的分类范例。安古洛夫等。人。[4]通过以局部特征和点邻域中的标签为条件的概率分布来模拟点的类标签。因此，它们利用扫描中的相邻点应具有相似标签的事实来强制空间连续性。该分布由马尔可夫随机场（MRF）建模，其参数在监督学习阶段中确定，使得得到的分类器最大化所学习的类之间的余量，如SVM那样。虽然[4]中没有给出时序结果，但从[6]可以得出结论，该方法不允许实时使用。

Lalonde等。[5]通过在手标记训练数据集上使用期望最大化（EM）算法拟合高斯混合模型（GMM）来学习每个类的特征分布的参数模型。通过在分类之后运行简单的基于规则的过滤器来解决空间容差，例如，通过将点的标签更改为其邻居中最常见的类。但是，为了使它们的方法实时执行，需要进行一些修改。特别是，它们不再对单个点进行分类，而是对3D体素网格单元中包含的所有点的人工原型点进行分类，使得7000个体素/秒。可以分类。

我们对3D点云中的对象分类采用了一种截然不同的独特方法，即分割基于包含在2.5D占用2网格中的压缩数据。然后，我们再次使用Velodyne的3D点云中包含的丰富信息，再次将域切换为3D，现在仅对扫描的总点云的子集进行分类，并证明每个子集代表单个对象。由于2D和3D数据处理技术的有效组合，可以在自动车辆上实时地对由其3D点云表示的对象进行分类。

2. 对象检测

A. 占用网格

我们使用尺寸为 $100\text{m} \times 100\text{m}$ 的2.5D自我中心占用网格，每个单元覆盖 $0.15\text{m} \times 0.15\text{m}$ 的小地面块。每个单元存储单个值，表示该单元被障碍物占用的程度。在我们的实现中，该值是具有物理单位[m]的度量长度。在我们详细说明其含义和计算之前，请注意在我们的方法中，我们在每次新的LIDAR革新中创建一个新的占用网格，即每0.1秒。因此，我们不会长时间累积数据。这个决定的原因是双重的。首先，Velodyne的一次革新提供了大约100000个3D点，这证明是足够的。其次，如果不以非常高的精度估计传感器的物理运动，则累积占用网格的质量可能容易恶化。传感器姿态估计中的小角度偏差可能导致大的误差。互相注册扫描，例如使用ICP算法[7]或它的一些衍生物，可以解决这个问题，但需要大量额外的计算负荷。

为了计算占用率值，我们首先惯性地校正LIDAR扫描，考虑车辆的运动（利用IMU和测距信息）。这是通过同时移动车辆的坐标系同时将本地LIDAR测量值转换为全局3D空间来完成的。在帧完成之后，所有点都被转换回车辆的最后一个局部坐标系，模拟扫描，好像所有测量都是在单个时间点而不是一个LIDAR旋转的0.1s时间段内进行的。

与Thrun等人类似[8]，将每个单元的值计算为落入相应网格单元的所有点的z坐标的最大绝对差。当激光束击中网格单元并更新其占用值时，我们将激光读取存储在单元中，以便可以查询以便稍后处理，详见第2节。III。图2显示了具有叠加点云的占用网格。

B. 来自分段的对象假设

为了获得初始对象假设，我们接下来通过找到网格单元的连通分量来执行对占用的网格单元的分割。为了应用连通分量算法，首先需要对网格进行二值化。这是通过简单地将所有细胞的占有值与适当的值（MuCAR-3的 0.15m ，从其轮胎的直径得出）进行阈值处理，将所有细胞低于阈值并将所有其他细胞设置为1来实现的。然后，可以应用从机器视图[9]中已知的标准连通分量算法，为每个网格单元 c_i 分配它所属的连通分量的标签 $label_i$ 。

对于每个连接的组件 cc_k ，我们制定一个未知类的对象假设，表示为3D边界框。可以根据属于相应连接分量的所有单元 c_i 的离散网格坐标 $gc_i = (u, v)^T$ 来计算对象的边界框的x轴和y轴， $cc_k = \{gc_i^{ego} | label_i = k\}$ 。

这里，“ego”上标用于表示现在所有网格坐标都在自我坐标系中表示，通过自我和网格单元之间的静态几何关系极大地简化了转换。然后，轴对应于坐标协变矩阵 Σ_{cc_k} 的标准正交特征向量 e_1, e_2 ，按照w.r.t的降序排序。相应的特征值，即 $d_1 \geq d_2$ 。

平面中的盒子尺寸是通过将所有 $gc_i^{ego} \in cc_k$ 转换为由特征向量（即所谓的本征空间）定义的坐标系统， $gc_i^{ego*} = (e1|e2) gc_i^{ego}$ ，并将极值超过结果坐标。对象假设的位置 pos_k 仅仅是连通分量的重心，即 $pos_k = |cc_k|^{-1} \sum (gc_i^{ego} \in cc_k)$ 。确保该2D盒子包围所连接组件的所有网格单元，但缺少一些期望的属性，例如具有所有可能的封闭盒的最小面积。虽然目前的盒子运作良好，但这可能会受到未来的改善。

假设对象的z轴与xy平面正交并将其z维度设置为连通分量的所有单元格部分的最大z坐标，我们获得最终的3D边界框对象假设。

图3示出了将轮廓对象检测算法应用于图2所示的占用网格的结果。注意，以这种方式检测对象不涉及关于对象形状的任何假设，而是执行自由形式对象检测。这与针对自动驾驶车辆的物体检测的大多数工作形成对比，其中通常明确地假设所有感兴趣的物体呈现“L形”，从而限制了可识别的不同类型的物体的数量。

3. 分类

(略。。。)