

# 平行点云: 虚实互动的点云生成与三维模型进化方法

田永林<sup>1,2</sup> 沈宇<sup>2,3</sup> 李强<sup>2,3</sup> 王飞跃<sup>2,4</sup>

**摘要** 三维信息的提取在自动驾驶等智能交通场景中正发挥着越来越重要的作用, 为了解决以激光雷达为主的深度传感器在数据采集方面面临的成本高、样本覆盖不全面等问题, 本文提出了平行点云的框架. 利用人工定义场景获取虚拟点云数据, 通过计算实验训练三维模型, 借助平行执行对模型性能进行测试, 并将结果反馈至数据生成和模型训练过程. 通过不断地迭代, 使三维模型得到充分评估并不断进化. 在平行点云的框架下, 我们以三维目标检测为例, 通过闭环迭代, 构建了虚实结合的点云数据集, 在无需人工标注的情况下, 可达到标注数据训练模型精度的 72%.

**关键词** 平行点云, 虚实互动, 三维视觉模型, 三维目标检测

**引用格式** 田永林, 沈宇, 李强, 王飞跃. 平行点云: 虚实互动的点云生成与三维模型进化方法. 自动化学报, 2020, 46(12): 2572-2582

**DOI** 10.16383/j.aas.c200800

## Parallel Point Clouds: Point Clouds Generation and 3D Model Evolution via Virtual-real Interaction

TIAN Yong-Lin<sup>1,2</sup> SHEN Yu<sup>2,3</sup> LI Qiang<sup>2,3</sup> WANG Fei-Yue<sup>2,4</sup>

**Abstract** The extraction of 3D information is playing an increasingly important role in intelligent traffic scenes such as autonomous driving. In order to solve the problems faced by LiDAR sensor such as the high cost and incomplete coverage of possible scenarios, this paper proposes parallel point clouds and its framework. For parallel point clouds, virtual point clouds are obtained by building artificial scenes. Then 3D models are trained through computational experiments and tested by parallel execution. The evaluation results are fed back to the data generation and the training process of 3D models. Through continuous iteration, 3D models can be fully evaluated and updated. Under the framework of Parallel Point Clouds, we take the 3D object detection as an example and build a point clouds dataset in a closed-loop manner. Without human annotation, it can be used to effectively train the detection model which can achieve the 72% of the performance of model trained with annotated data.

**Key words** Parallel point clouds, virtual-real interaction, 3D vision models, 3D object detection

**Citation** Tian Yong-Lin, Shen Yu, Li Qiang, Wang Fei-Yue. Parallel point clouds: point clouds generation and 3D model evolution via virtual-real interaction. *Acta Automatica Sinica*, 2020, 46(12): 2572-2582

收稿日期 2020-09-26 录用日期 2020-12-11

Manuscript received September 26, 2020; accepted December 11, 2020

国家自然科学基金重点项目 (61533019), 英特尔智能网联汽车大学合作研究中心项目 (ICRI - IACV), 国家自然科学基金项目联合基金 (U1811463), 广州市智能网联汽车重大科技专项 (202007050002) 资助

Supported by the State Key Program of National Natural Science Foundation of China (61533019), the Intel Collaborative Research Institute for Intelligent and Automated Connected Vehicles (ICRI-IACV), the Joint Funds of the National Natural Science Foundation of China (U1811463), and the Key Research and Development Program of Guangzhou (202007050002)

本文责任编辑 袁勇

Recommended by Associate Editor YUAN Yong

1. 中国科学技术大学自动化系 合肥 230027 2. 中国科学院自动化研究所复杂系统管理与控制国家重点实验室 北京 100190 3. 中国科学院大学人工智能学院 北京 100049 4. 青岛智能产业技术研究院 青岛 266000

1. Department of Automation, University of Science and Technology of China, Hefei 230027 2. State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190 3. School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049 4. Qingdao Academy of Intelligent Industries, Qingdao 266000

近年来, 随着计算机视觉技术的快速发展, 诸如图形分类<sup>[1-2]</sup>、目标检测<sup>[3]</sup>、目标跟踪<sup>[4]</sup>、智慧医疗<sup>[5-6]</sup>等二维视觉任务, 正在得到妥善解决和广泛应用. 然而, 在诸多对安全性要求较高的领域如无人驾驶<sup>[7]</sup>、智能机器人<sup>[8]</sup>等, 三维信息的提取和感知则发挥着不可或缺的作用. 以激光雷达和深度相机为代表的深度传感器, 利用其测距功能能够实现对场景深度信息的提取, 其中激光雷达更是凭借其精确的深度测量能力成为室外等较大场景下深度信息获取的首选传感器, 也成为无人驾驶解决方案中的关键组成部分<sup>[9-11]</sup>. 在数据组织方面, 虽然 RGB-D 的数据在数据处理方面存在优势, 但是近期的研究<sup>[12-13]</sup>表明, 相比于 RGB-D 数据类型, 点云数据能够实现更为有效和精确的三维信息感知.

目前, 利用激光雷达获取场景点云信息存在着成本高、标注困难、场景覆盖不全等一系列问题<sup>[14-15]</sup>.

在成本方面, 单枚多线激光雷达的价格往往高达数万人民币. 在标注方面, 由于需要对场景的三维信息进行标注, 但同时点云又存在着稀疏的特点, 导致对感兴趣的目标难以获取其完整的点云信息. 所以, 在标注时需要结合图像等其他信息, 才能对目标的三维信息给出准确标注. 这一过程往往比图像标注更为困难. 此外, 受限于激光雷达设备的昂贵价格, 普通民众无法拥有这样的设备来进行数据采集与分享, 因此点云数据的获取严重依赖于专业团队, 无法做到群策群力, 导致现有的大多数数据集存在数据量小, 样本难以覆盖真实世界的可能情景等问题.

为了解决点云数据生成和标注中存在的困难, 同时对三维模型进行迭代优化, 本文基于 ACP 理论<sup>[16-18]</sup>, 即人工社会 (Artificial societies)、计算实验 (Computational experiments)、平行执行 (Parallel execution), 提出了虚实互动的平行点云计算方法. 首先通过在虚拟环境构建与真实环境平行的人工场景, 对现实世界进行描述、预测和引导, 借助虚拟激光雷达, 获取人工场景的点云表示. 在此基础上, 利用虚实结合的训练方式对三维检测模型进行训练, 完成 3D 检测和跟踪等相关的任务. 之后借助平行执行对模型效果进一步评估, 最后根据反馈结果, 对人工场景的建立、数据采集以及模型训练等过程进行调整优化, 实现整个 ACP 系统的闭环. 在平行点云的框架下, 本文以交通场景下的车辆三维检测任务为例, 以虚实结合的方式构建了无需标注的车辆点云数据集, 缓解了三维检测任务中点云目标标注的困难. 本文的贡献总结如下:

1) 针对点云应用场景, 提出了平行点云计算框架. 通过虚实结合、闭环迭代的方式, 对点云生成以及三维模型优化提供了新的思路.

2) 针对交通场景下的车辆三维检测任务, 构建了平行点云数据集 ShapeKITTI, 显著降低了数据集构建的成本.

3) 基于平行点云数据集, 训练了目标检测模型, 在真实数据上取得了良好的效果.

4) 构建了数据生成与模型评估的闭环, 利用反馈信息进行迭代, 优化了数据生成效果并提升了模型性能.

本文剩余部分安排如下: 第 1 节对相关工作进行总结回顾; 第 2 节阐述了平行点云的理论框架; 第 3 节介绍了三种虚拟点云生成方法并展示了生成效果; 第 4 节以三维检测为例, 基于 ACP 进行了实验; 第 5 节对平行点云进行了思考与展望; 第 6 节对本文进行了总结.

## 1 相关工作综述

基于点云的三维算法近几年受到了来自学术界

和工业界的广泛关注<sup>[19]</sup>, 研究方向包括了点云分类<sup>[20-22]</sup>、点云分割<sup>[23-25]</sup>以及基于点云的目标检测<sup>[7, 26-27]</sup>和跟踪<sup>[28-29]</sup>等. 这些三维算法和模型的发展与进步离不开开源点云数据集<sup>[9-11, 30]</sup>的支撑. 真实数据集的采集和标定是一项耗时耗力的工作<sup>[31]</sup>, 不仅需要昂贵的数据采集设备, 同时数据集的设计和标注同样需要专业的团队来实施. 随着计算机图形学和虚拟现实技术的发展, 软件定义的场景生成和数据采集等技术逐渐获得应用. Prendinger 等<sup>[32]</sup>通过构建虚拟城市实验室来进行驾驶和出行行为的研究. 他们利用开源地图 OpenStreetMap 和大规模城市生成软件 CityEngine, 根据实际地图搭建了仿真城市环境, 之后利用游戏引擎 Unity 添加车辆以及交通设施等, 并模拟了车辆行为以及车辆与道路的交互. 为了提高数据集的多样性同时降低数据采集和标注的成本, Ros 等<sup>[33]</sup>利用 Unity 引擎和 CAD (Computer aided design) 模型构建了逼真的虚拟交通场景, 并自动生成了语义、深度等标注信息. 为了增加数据集的多样性, 他们借助游戏引擎的渲染技术对同一场景设置了不同的天气效果, 为后续研究天气对模型的影响提供了基础. 他们将收集到的虚拟数据同真实数据一起用于训练语义分割模型, 相比于单纯使用真实数据集训练的模型, 混合数据的使用取得了更好的效果, 从而证实了虚拟数据集在视觉模型训练中的有效性. Gaidon 等<sup>[34]</sup>则通过模拟真实 KITTI 数据集<sup>[9]</sup>, 构建了虚拟 KITTI 视频序列. 虚拟 KITTI 数据集在场景结构、物体位置以及物体运动和外观等特征上均和真实 KITTI 数据具有类似的视觉效果, 并自动产生了丰富的视觉信息标注, 包括目标检测、语义分割、目标跟踪、深度、光流等. Li 等<sup>[35]</sup>利用游戏引擎对真实世界中的街道进行了重建并产生了大量多样化的数据, 通过虚实结合的训练方式提高了视觉模型的性能<sup>[36]</sup>. 不同于首先构建虚拟场景进而再收集数据的方法, Johnson-Roberson 等<sup>[37]</sup>直接利用游戏 GTA V (Grand theft auto V) 进行数据的生成, 这种方式大大节省了场景构建所需的时间, 同时生成的数据也更为逼真. 与虚拟图像关注二维信息的采集不同, 虚拟点云还需要捕获像素的深度信息. 虚拟点云不仅可以基于虚拟场景来进行生成, 还可以利用虚实结合的方式, 将真实点云场景和虚拟点云目标进行结合产生, 从而减少工作量, 同时提升逼真效果. Yue 等<sup>[38]</sup>利用软件定义的虚拟雷达, 对 GTA V 游戏中的场景进行了雷达扫描, 从而生成了大量虚拟点云数据. 通过在训练过程中加入虚拟点云, 语义分割模型性能得到了明显提升. Li 等<sup>[39]</sup>利用真实激光雷达扫描得

到背景信息,在此基础上,通过虚拟激光雷达对三维模型进行扫描,将得到的虚拟点云同真实背景融合,构成了混合数据。

上述方法主要关注虚拟数据集的构建以及其在模型训练测试等方面的有效性,而未能将数据生成、模型训练以及系统级评估构成闭环。Wang<sup>[16-18]</sup>提出了 ACP 理论来解决复杂环境的建模与调控问题。近年来该理论在计算机视觉<sup>[40]</sup>、图像生成<sup>[41]</sup>、机器学习<sup>[42]</sup>、自动驾驶<sup>[43]</sup>、艺术创作<sup>[44]</sup>及智能交通<sup>[45]</sup>等领域获得了广泛应用。本文将 ACP 思想引入点云生成和三维模型的性能进化中,旨在通过数据、模型交互的闭环,提升基于点云的三维模型解决复杂实际问题的能力。

## 2 平行点云的基本框架

平行点云是 ACP 思想在点云生成和三维模型训练及评估中的应用。图 1 展示了平行点云的基本框架,主要包括三部分:人工场景构建、计算实验和平行执行。其中人工系统旨在建立与真实环境平行的软件定义的人工场景以及平行传感系统,在真实世界和虚拟环境进行点云数据的收集和标注。计算实验旨在以虚实互动的方式,利用人工场景和真实场景数据集对三维模型进行训练和评估,充分发挥虚拟数据集灵活可控的特性。平行执行则通过虚实系统间的联动,利用人工系统对真实系统进行预测和引导,同时根据真实系统的反馈对人工系统进行调整和优化。该框架旨在借助 ACP 理论在复杂系统的建模方面的优势,解决点云数据收集和标注,以及模型评估和进化等方面存在的问题。将数据生成以及模型训练和评估有机结合并形成闭环,最终实现三维模型性能的持续进化。

### 2.1 人工场景构建与平行传感

人工场景的构建旨在克服现实世界的不可控、不可重复等缺点,通过对现实世界构建多个孪生系统,实现对现实世界的模拟、预测和引导。

在本环节,人工场景利用现实世界的“小数据”生成人工世界的“大数据”,为三维模型提供充足的训练样本。此外,软件定义的人工场景能够轻松实现对模型执行结果的反馈进行响应,是系统闭环的重要环节也是区别于简单仿真系统的关键特性。人工场景的构建可以借助计算机图形学以及虚拟现实等领域的方法和手段,以软件定义的形式实现对现实世界的建模。现有的自动驾驶仿真器、游戏引擎以及游戏等都可以用来实现对现实世界的建模。

平行传感<sup>[46]</sup>利用分布在虚拟和现实两个平行世界的传感器,实现对人工场景和真实场景的数据采集。在平行点云框架中,借助人工场景的虚拟激光雷达传感器和现实世界的真实激光雷达传感器,我们便可以生成包含大量虚拟数据和少量真实数据的混合点云数据。真实激光雷达传感器通过发射和接收激光束来对一定范围内的物体进行测距。垂直方向上,通过发射多束激光束实现对垂直视场的扫描;在水平方向上,通过旋转或光学相控阵等技术实现对水平视场的扫描。图 2 展示了激光雷达的工作示意图。根据实际激光雷达的特性,可以在人工场景中通过软件定义虚拟激光雷达。虚拟激光雷达利用发射的一系列射线与场景中物体进行碰撞来模拟实际场景中激光碰到物体发生反射的过程,可以实现同真实激光雷达类似的效果。此外,软件定义虚拟激光雷达具有多个可调节参数,能够灵活地控制数据采集过程。表 1 展示了虚拟激光雷达的主要参数。

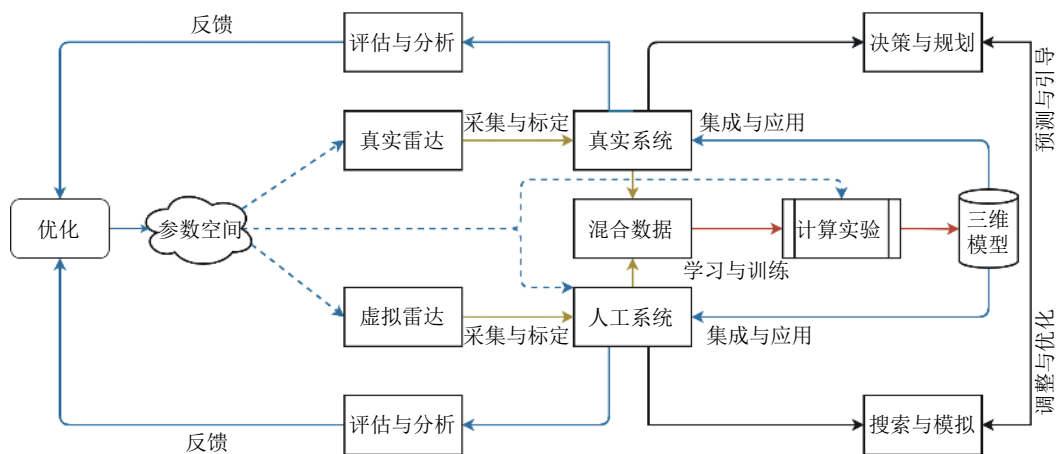
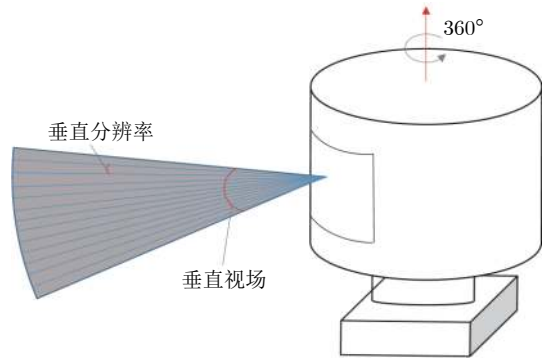


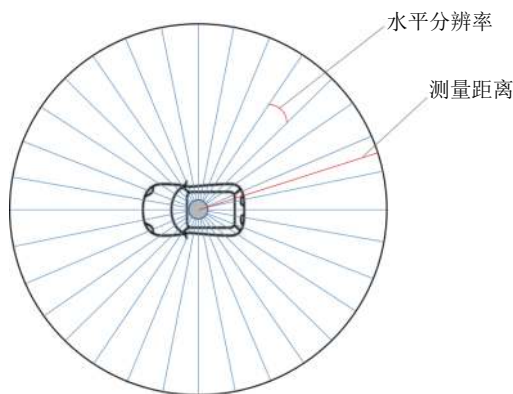
图 1 平行点云框架

Fig.1 The framework of parallel point clouds



(a) 激光雷达垂直平面扫描示意图

(a) The illustration of LiDAR scanning on vertical plane



(b) 激光雷达水平面扫描示意图

(b) The illustration of LiDAR scanning on horizontal plane

图 2 激光雷达扫描示意图

Fig. 2 The illustration of LiDAR scanning

表 1 激光雷达属性描述

Table 1 The description of LiDAR attributes

属性	描述	默认值
Channels	激光器数量	32
Range	测量/射线广播的最大距离(米)	10.0
Points per second	每秒所有激光器产生的点	56000
Rotation frequency	激光雷达旋转频率	10.0
Upper field of view	最高激光的角度(度)	10.0
Lower field of view	最低激光的角度(度)	-30
Atmosphere attenuation rate	测量每米LiDAR强度损失的系数	0.004
Dropoff general rate	随机掉落的总点数比例	0.45
Dropoff intensity limit	对于基于强度的下降, 阈值强度值不超过任何点	0.8

通过软件定义的人工点云场景以及平行传感系统, 我们便能够在真实世界和虚拟世界中进行点云数据的收集和标注. 从真实世界采集的激光雷达数据能反映目标应用场景的特性, 但是受限于真实激光雷达昂贵的成本以及点云标注过程存在的困难, 从真实世界获取点云数据, 其规模难以扩展. 相比

之下, 虚拟场景能够灵活地实现对各类元素的调整, 如天气、场景类型、激光雷达参数等. 借助软件定义的虚拟激光雷达传感器, 便能够灵活地进行点云数据采集, 同时由于点云数据生成由软件定义, 因此其标注几乎无需人工参与, 从而大大降低了数据集构建所需的成本.

## 2.2 基于平行点云的计算实验

在人工场景和真实场景的基础上, 利用平行传感建立起各种三维任务所需要的数据集, 之后便能够开展各类计算实验, 包括三维重建、三维检测、SLAM (Simultaneous localization and mapping)、点云分割等. 本文中, 计算实验的目的在于利用人工场景和真实场景生成的大数据, 来训练各种三维模型, 使其能够在实际场景中得到应用. 模型训练过程可以采用虚实互动的训练方式, 利用大规模带标注的虚拟样本和少量真实样本作为输入, 最小化算法损失, 使模型参数不断更新. 之后利用测试数据对模型进行评估, 调整优化模型结构和训练策略, 使模型达到最佳水平.

基于平行点云的计算实验的关键在于虚实点云数据的有效利用、三维模型的设计以及模型的评估. 三维模型设计超出本文的范围, 这里不予过多讨论. 在数据方面, 相比于图像数据, 激光雷达采集的点云更关注场景的空间结构和物体的形状, 不存在图像中的纹理、风格等方面的问题, 但是来自不同设备和场景的点云, 在点的稀疏程度以及空间布局上仍存在着一定差异. 由于现在的大多数机器学习模型严重依赖于数据的分布状况, 因此一个需要解决的问题是如何克服虚拟数据和真实数据之间的差异. 对此, 目前主要有两种思路, 一种是利用混合数据对模型进行训练, 借助模型自身的学习能力, 同时对虚拟和真实样本进行拟合. 另外一种解决思路是利用领域适应技术<sup>[47]</sup>, 通过实例迁移、特征迁移、参数迁移等方式, 缓解虚拟样本和真实样本的分布差异问题. 近些年随着生成式对抗网络 (Generative adversarial networks, GAN)<sup>[48]</sup> 的发展, 利用生成对抗技术进行领域适应也取得了不错的效果. 该方法主要利用对抗训练的方式调整特征学习和判别过程, 使得真实样本的特征和虚拟样本的特征不可区分, 从而实现不同输入到同一特征空间的转换.

三维模型评估的目的在于利用采集到的虚拟和真实点云数据集对模型进行全面分析. 在这一环节, 人工场景构建的虚拟点云数据集可以充分发挥其灵活可重复的特点, 针对感兴趣的评测指标来构建人工场景并收集数据, 以对模型进行充分全面的评估. 例如为了衡量无人驾驶车三维感知算法在恶劣天气

下的表现能力, 我们可以调节场景生成的效果, 创建阴雨、沙尘等恶劣天气状况, 在此情况下, 虚拟点云数据受环境影响, 将出现更多的噪声点, 同时数据的分布范围也会发生改变, 这将给三维算法的适应能力提出更大挑战. 此外, 还能够对现实世界中出现较少的情况进行模拟, 收集对应的点云数据来训练和评估三维模型. 相比之下, 真实数据往往受限于其规模和多样性, 而难以对三维模型进行全面评价. 在平行点云的框架中, 真实点云数据将主要用于提取其所表征的场景的领域特征, 来指导模型进行跨域的适应训练, 评估模型在虚拟数据和真实数据间的迁移能力. 而虚拟点云数据将更多地用于模型性能的提升和模型泛化能力的评估.

### 2.3 基于平行点云的平行执行

平行执行的目的是主要有两个, 一是通过将模型部署到整个系统中, 根据系统在人工和真实场景中运转情况来进一步检验模型的综合性能; 二是通过虚实系统之间的联动, 利用人工系统对真实系统进行预测和引导, 同时根据真实系统的反馈对人工系统进行调整和优化.

在模型整体性能评估方面, 基于平行点云的平行执行可以缓解三维模型对数据集的依赖, 避免基于固定数据集的评估方法带来的对模型的片面认识. 现有的大多数机器学习模型的训练和评估建立在训练和测试数据同分布的假设之上, 这导致了虽然三维模型在测试集上的评估指标可能很理想, 但是在实际应用中其性能却不尽如人意. 模型部署在实际场景之后, 面对复杂多变的环境, 其适应能力和实际表现往往会大打折扣, 尤其是面临训练和测试集中出现频率较少的场景时, 性能的退化将更为严重. 平行执行则将三维模型的评估和训练全面化、长期化, 将模型在实际和人工场景中进行平行部署, 根据其表现产生反馈, 调整点云场景以及激光雷达参数等环节, 形成数据采集、模型训练和模型评估的闭环, 使三维检测模型不断进化. 针对实际应用三维模型的过程中出现的问题, 人工系统可以将其记录在案, 并利用人工场景可控可重复的特点, 对失败案例进行复现, 通过控制变量的形式, 调整可能导致失败的主要因素, 从而定位问题所在. 之后, 根据反馈信息, 调整点云数据采集、场景生成以及三维模型训练等环节对应的参数空间, 通过重构场景、调整训练数据分布以及修改模型结构或参数等手段, 弥补当前模型存在的不足.

在虚实系统联动方面, 平行执行既可以发挥人工系统在模拟计算方面的优势, 对实际系统的决策规划进行预测和引导, 同时又可以利用真实系统的

反馈对人工系统的决策和规划进行调整与优化, 使人工系统更准确地对真实系统进行建模. 人工场景建立在已有的知识之上, 而实际系统往往是未知和不可控的, 在复杂环境中, 三维模型在实际系统中的运转往往并不按照人们的意愿进行, 因此需要借助人工系统对实际系统的决策进行建模和监督, 通过对比真实系统的实际输出和人工系统的理想输出, 及时发现真实系统潜在的问题, 避免意外事故发生. 同时基于对当前实际系统的模拟, 人工系统可以对未来时刻可能的状况进行预测, 实际系统运行一步, 人工系统可以运行多步, 对可能的状态进行搜索, 从而选择最佳方案, 作为对真实系统决策和规划的引导, 使真实系统尽可能按照理想方案运行.

## 3 虚拟点云的生成途径

虚拟点云可以通过多种方式进行生成, 既可以直接利用三维建模软件模仿真实世界进行生成, 也可以通过虚实结合的方式将虚拟点云同真实点云进行融合, 还可以借助图像深度估计生成伪点云.

### 3.1 基于三维建模的虚拟点云生成方法

基于三维建模的虚拟点云生成借助了仿真软件的强大功能, 通过对真实环境进行高精度的建模, 并使用内置的虚拟传感器对仿真环境进行扫描, 来生成虚拟场景的点云数据. 现如今市场上众多的自动驾驶仿真软件, 例如 CARLA<sup>[49]</sup>、LGSVL<sup>[50]</sup>、VTD<sup>[51]</sup>等, 均支持多种传感模式和环境条件的灵活配置. 这些自动驾驶仿真软件可以完成对真实场景的建模, 以及复杂交通场景的实时生成, 包括道路建模、天气环境模拟、交通信号灯仿真控制等, 并且可以通过自主添加雷达、摄像头等传感器来得到仿真环境的点云数据.

图 3 展示了我们基于 CARLA 仿真环境生成的虚拟点云数据. 我们首先创建客户对象连接 CARLA 模拟器的服务器, 在生成的地图场景的同时创建主体汽车. 在模拟器中主体车辆可以以自动驾驶模式或人工控制模式进行运动. 在运动过程中, 车辆顶部绑定的虚拟激光雷达对周围环境进行扫描得到虚拟点云数据. 通过虚拟场景生成虚拟点云极大地提高了数据生成的灵活性, 能够实现按需定制, 因此十分适合用来对现实场景中较少出现的情况进行个性化生成, 提高无人驾驶算法的性能.

### 3.2 基于伪雷达的点云生成方法

伪雷达是一种利用图像生成点云的方法. 传统激光雷达通过激光发射器以一定频率向外发生激光脉冲, 通过计算到达接收器的反射光得到周围环境

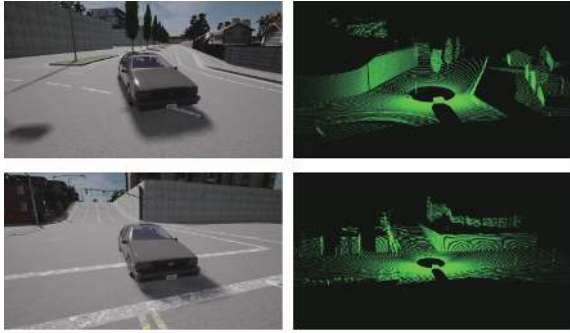


图 3 CARLA 仿真环境及虚拟雷达点云

Fig.3 CARLA simulation environment and virtual LiDAR point clouds

障碍物的方位和距离. 其采集的点云数据在 3D 空间中表现为稀疏的数据组织形式. 由图片中像素深度反投影得到的点云可以弥补传统激光雷达缺失的信息. 伪雷达生成通过立体匹配、深度估计等方法首先计算出图像中像素点  $(u, v)$  的深度信息  $Z(u, v)$ , 再基于图像坐标与世界坐标的变换关系反推出像素点的三维坐标  $(x, y, z)$ . 其中  $(C_U, C_V)$  为相机中心,  $f_U$  和  $f_V$  分别为水平和垂直方向上的焦距. 计算方法为

$$\begin{aligned} z &= Z(u, v) \\ x &= \frac{(u - C_U) \times z}{f_U} \\ y &= \frac{(v - C_V) \times z}{f_V} \end{aligned} \quad (1)$$

图 4 展示了从 KITTI 双目图像数据生成伪雷达的效果图. 我们利用文献 [52] 提出的方法, 首先利用双目深度估计网络, 对图像数据的深度进行估计, 得到深度图像, 之后利用投影矩阵和矫正矩阵, 将深度映射到三维空间形成点云. 可以看到相比于真实雷达数据, 伪雷达数据提供了更为稠密的三维信息, 而且由于不需要使用昂贵的激光雷达, 点云数据生成的成本得到了极大的降低.

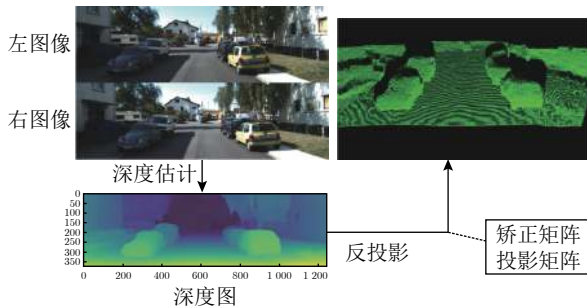


图 4 KITTI 伪雷达点云生成方法

Fig.4 Pseudo LiDAR point clouds generation with KITTI dataset

### 3.3 虚实结合的点云生成方法

虚实结合的生成方式主要考虑到了数据生成成本和生成质量的问题. 通过将真实数据与虚拟数据进行合成并调整数据分布, 一方面缩小合成数据同真实数据的差异, 另一方面提升数据的多样性. 一种有效的手段是通过将前景背景进行分离, 将真实点云背景同虚拟目标进行合成. 该方法利用真实激光雷达的数据来提取背景点云, 之后借助现有的三维模型以及虚拟点云技术, 针对感兴趣的目标生成点云并融合到真实点云背景中. 由于目标位置和类型可以人为定义, 因此, 这种方式可以在很大程度上减少数据标注的工作量. 同时, 考虑到大多数基于点云的模型对前景的类别、位置等信息较为关注, 通过虚拟目标和虚拟激光雷达合成目标的丰富点云将有助于提高三维模型对感兴趣目标的检测能力. 此外, 为了进一步体现虚实互动的灵活性, 除了添加感兴趣的正样本目标之外, 我们还可以有针对性地加入现实世界中缺乏的负样本目标 (如出现位置不符合常理的交通设施、建筑物、树木甚至是动物等), 来提高模型的辨别能力.

图 5 展示了我们虚实结合的生成方式的示意图. 我们将 KITTI 数据集作为真实点云数据来源, 将 ShapeNet 数据集<sup>[53]</sup>作为虚拟点云来源, 首先对 CAD 模型进行点云化处理, 通过提取各个部件的表面点云, 将 CAD 模型转换为点云表示. 之后根据真实点云的雷达位置对虚拟点云遮挡进行模拟, 实现同真实数据相似的数据分布. 最终将二者进行合

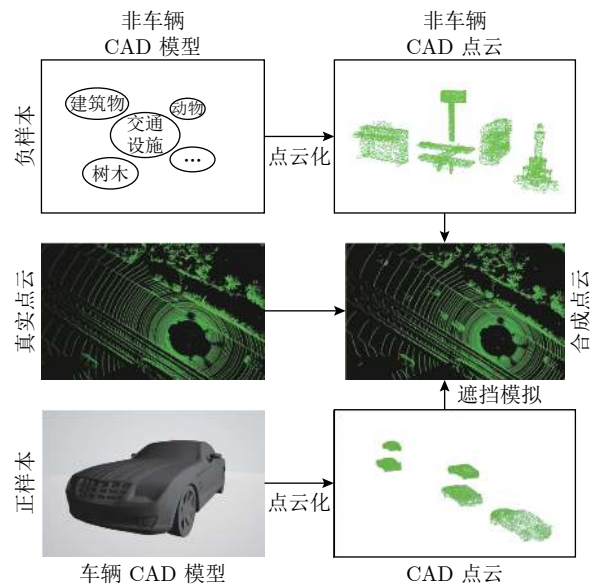


图 5 虚实结合的点云生成方法

Fig.5 LiDAR point clouds generation via virtual-real interaction

成。可以看到,虚实结合的方式能够获得较为逼真的点云数据,同时实现了对前景目标进行灵活操控,因此能够作为一种有效的数据增强手段对三维模型进行训练。

## 4 虚实结合的三维检测模型训练

### 4.1 平行点云数据集 ShapeKITTI

基于第 3.3 节中的数据生成方法,我们以虚实结合的方式构建了新的点云数据集 ShapeKITTI。该数据集从 KITTI 3712 帧数据中获取了背景点云信息,同时基于虚拟雷达对 ShapeNet 中的 200 个车辆数据进行处理,得到目标的虚拟点云。为了减少无关信息干扰,我们对真实 KITTI 训练集中的前景点云进行了去除,使得 ShapeKITTI 数据集中感兴趣的前景目标均来自虚拟雷达。为了模拟真实点云中的点的消失情况,我们在数据生成过程中,以 0.6 的概率随机对目标点云进行了丢弃。同时为了模拟噪声的干扰,对点的位置进行了扰动处理。

图 6 展示了 ShapeKITTI 与 KITTI 数据的可视化结果对比,其中,左侧图像来自真实 KITTI 数据集,右侧图像来自 ShapeKITTI 数据集。可以看出,利用我们提出的虚实结合的点云数据生成方式,场景的真实性得到很好地保障,同时前景目标的生成可以由人工定义,在避免数据标注的前提下,大大提升了数据的可控性和灵活性。

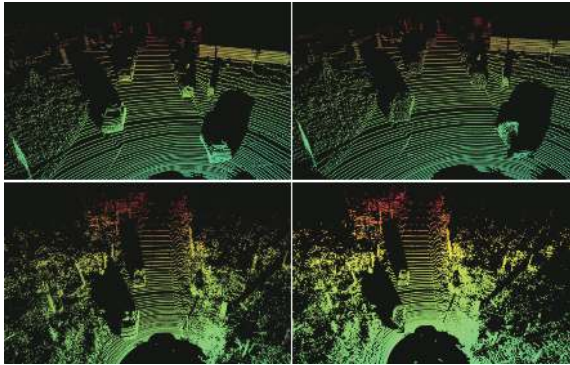


图 6 KITTI (左) 与 ShapeKITTI (右) 点云数据可视化对比

Fig.6 The visualization of point clouds from KITTI (left) and ShapeKITTI (right)

### 4.2 三维目标检测训练

本小节以针对车辆的三维目标检测模型为例,对基于平行点云的数据生成方法的有效性进行了验证。我们选择了 PointPillars<sup>[26]</sup> 和 SECOND<sup>[54]</sup> 模型作为实验对象,这两个模型具有体积小、速度快等优点,能够对目标进行实时检测。实验中,我们在

ShapeKITTI 数据上对三维检测模型进行了训练,在真实 KITTI 验证集上进行模型评估,评估指标包括鸟瞰图检测精度 (BEV AP)、三维检测精度 (3D AP) 以及角度识别精度 (AOS AP)。所有指标均为交并比 (Intersection over union, IoU) 为 0.7 下的结果。表 2 记录了基于 ShapeKITTI 训练的 PointPillars 模型在真实 KITTI 验证集上的效果。模型训练过程中,使用了 Adam 优化器,在 ShapeKITTI 数据集上训练了 20 轮,初始学习率为 0.003,动量为 0.9,批数据大小为 16。PointPillars 模型栅格化尺寸为 [0.16, 0.16, 4]。表 3 记录了基于 ShapeKITTI 数据集训练的 SECOND 模型在真实 KITTI 验证集上的效果。模型训练过程中,使用了 Adam 优化器,在 ShapeKITTI 数据集上训练了 20 轮,初始学习率为 0.0005,动量为 0.9,批数据大小为 16。SECOND 模型的栅格化尺寸为 [0.05, 0.05, 0.05]。在表 2 和表 3 中,我们同时给出了在真实 KITTI 数据集上进行训练并在真实 KITTI 验证集上的评估结果。通过对比可以发现,我们提出的数据生成方法,在无需数据标注的情况下取得了不错的效果,其中基于 PointPillars 的模型结果 (3D AP) 在中等模式下达到了真实数据训练结果的 72%,证明了我们数据生成方式的有效性。

表 2 基于 ShapeKITTI 的三维目标检测模型 (PointPillars) 平均精度 (%)

Table 2 The average precision of 3D object detector (PointPillars) based on ShapeKITTI dataset (%)

评测项目	简单模式	中等模式	困难模式
3D AP (Real)	85.21	75.86	69.21
3D AP (Ours)	71.44	54.39	52.30
BEV AP (Real)	89.79	87.44	84.77
BEV AP (Ours)	82.81	70.44	65.68
AOS AP (Real)	90.63	89.34	88.36
AOS AP (Ours)	84.02	69.65	68.08

表 3 基于 ShapeKITTI 的三维目标检测模型 (SECOND) 平均精度 (%)

Table 3 The average precision of 3D object detector (SECOND) based on ShapeKITTI dataset (%)

评测项目	简单模式	中等模式	困难模式
3D AP (Real)	87.43	76.48	69.10
3D AP (Ours)	59.67	41.18	37.67
BEV AP (Real)	89.52	87.28	83.89
BEV AP (Ours)	73.55	53.37	52.34
AOS AP (Real)	90.49	89.45	88.41
AOS AP (Ours)	75.01	54.72	54.13

### 4.3 基于闭环反馈的三维模型进化实验

本环节借助平行执行的思想, 将模型评估的反馈与数据生成环节构成闭环, 来优化数据生成过程进而实现三维检测模型性能的进化. 实验中, 我们使用了 ShapeKITTI 和真实 KITTI 数据集. 其中, 真实 KITTI 数据集被用作验证集, 以获得模型在实际场景中的性能. ShapeKITTI 数据被用作训练集, 借助其生成灵活的特点, 我们将模型评估结果用作数据生成的指导, 研究了闭环驱动的三维模型进化方法.

表 4 展示了基于闭环反馈的三维检测模型性能进化过程, 图 7 展示了模型性能变化的直方图. 在初始的数据生成阶段 G0 中, 我们只利用了 10 个汽车的 CAD 模型进行 ShapeKITTI 数据集的构建, 根据在实际 KITTI 数据集上进行测试的结果, 我们发现基于 G0 训练的模型只能完成对部分类型车辆的检测, 因此在数据生成阶段 G1 中, 将 CAD 模型的数量增加至 200 个, 由此给三维目标检测器的性能带来了 5.94% 的提升. 而基于 G1 的训练结果表明, 模型对物体的尺寸估计效果欠佳, 为此, 我们依据经验, 将常见车辆目标的平均尺寸作为 ShapeKITTI 数据生成指导, 改进后的数据生成方式 G2 使模型性能提高了 5.94%. 但同时数据生成方式 G2, 缺乏

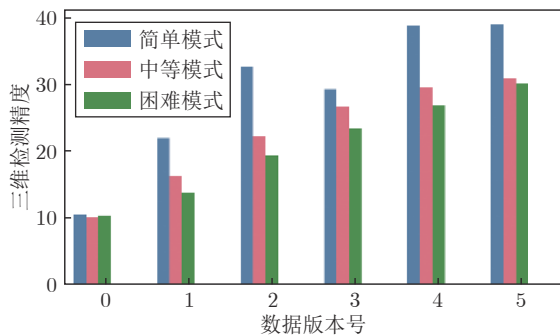


图 7 点云数据生成对三维检测模型性能的影响

Fig. 7 The influence of point clouds generation on the 3D detector performance

对真实数据密度的有效模拟, 导致训练出的模型无法对稀疏目标进行较好的检测, 所以我们提出降低 ShapeKITTI 数据集中模型点云的密度, 新的数据生成方式 G3 将模型性能提高了 4.45%. 之后, 通过对模型分析, 又进一步调整了 ShapeKITTI 数据集中目标的高度和目标密度, 使训练出的模型性能提升了 2.89% 和 2.3%.

## 5 思考与展望

### 5.1 利用平行点云构建多样化数据集

训练数据的多样性直接影响着模型的性能, 构建能够高度覆盖模型运行环境的数据集有助于提升模型的鲁棒性和泛化能力. 在现实世界中, 数据的采集存在着诸多不可控因素, 导致建立的数据集往往不能实现对实际环境的充分覆盖, 这也给算法的落地带来了一定的安全隐患. 相比较真实数据而言, 虚拟数据在多样性上存在着天然的优势. 由于其生成过程完全由算法和参数等可控因素决定, 因此可以很方便地模拟出各式各样的数据样本, 能够对真实数据起到很好的补充作用.

图 8 和图 9 展示了我们利用虚拟雷达模拟的物体的距离和角度的变化所带来的生成效果上的差

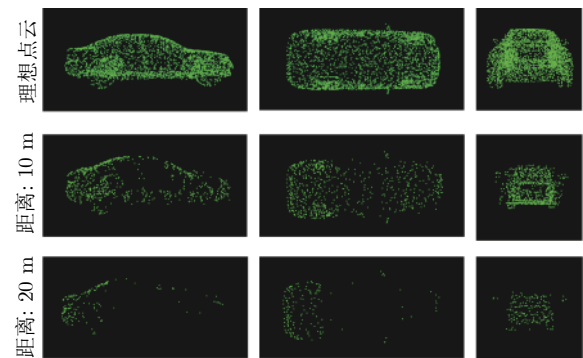


图 8 不同距离下的虚拟点云生成效果

Fig. 8 The visualization of virtual point clouds under different distance

表 4 基于闭环反馈的三维检测模型性能进化过程 (%)

Table 4 The evolutionary process of 3D detection model based on the closed-loop feedback (%)

序号	AP (简单)	AP (中等)	AP (困难)	反馈	建议	AP提升
G0	10.64	10.32	10.53	仅对少数目标能有效检测	增加CAD模型种类	—
G1	22.14	16.46	13.98	目标尺寸估计欠佳	改进CAD模型尺寸	6.14 (G0 to G1)
G2	32.85	22.40	19.52	对稀疏目标检测效果欠佳	降低前景目标密度	5.94 (G1 to G2)
G3	29.43	26.85	23.51	目标高度估计效果欠佳	调整前景目标高度	4.45 (G2 to G3)
G4	38.99	29.74	27.10	对稀疏目标检测效果仍然欠佳	进一步降低前景目标密度	2.89 (G3 to G4)
G5	39.22	31.04	30.30	—	—	2.30 (G4 to G5)



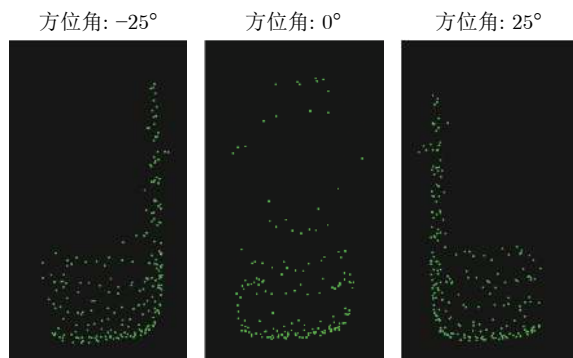


图 9 不同方位角的虚拟点云生成效果

Fig.9 The visualization of virtual point clouds under different azimuths

异. 可以看到随着距离的增加, 物体点云的密度逐渐降低. 随着方位角的不同, 物体所呈现的外观特征也存在明显差异. 由虚拟场景生成数据集不仅在模型训练上存在着巨大潜力, 同时还有助于通过精细化的评估获得对已有模型性能的充分认识.

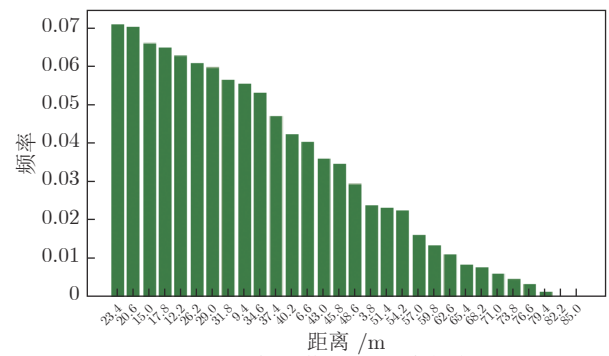
## 5.2 利用平行点云解决三维模型面临的长尾问题

长尾分布广泛存在于自动驾驶、智能交通等领域. 这给三维模型在实际场景中应用时的安全性和可靠性提出了严峻挑战. 如图 10 所示, 三维模型训练所依赖数据集的很多属性 (以距离和角度为例, 按照频率高低进行排序) 也基本服从长尾分布, 因此对稀有样本类别的学习不够充分, 将使得三维模型在实际应用中对可能出现的罕见状况的适应能力较弱, 从而给实际系统的运转带来了不安全因素.

平行点云将数据生成、模型训练、系统评估构成闭环, 以迭代的方式不断更新数据以及三维模型, 使模型性能不断得到进化. 对长尾分布的尾端类别, 由于其现实世界的出现频率较低, 传统的数据收集方法难以实现对样本多样性的有效覆盖. 相比之下, 基于人工场景的点云数据生成可以充分发挥人的主观能动性, 对现实世界稀有的样本进行模拟, 用于模型训练阶段, 提升模型对长尾分布数据的有效解决能力.

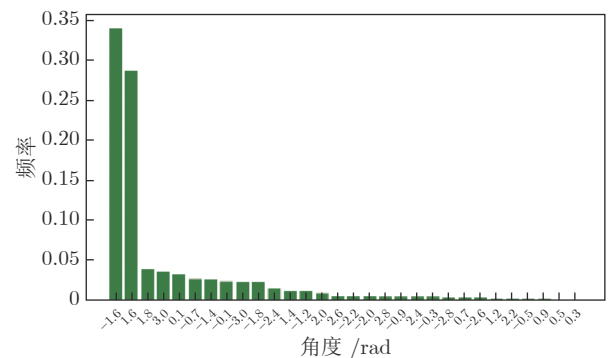
## 6 结论

本文以 ACP 理论为基础, 提出了平行点云的框架. 平行点云将点云数据生成、模型训练以及系统级评估构成闭环, 以一种迭代的方式进行点云数据的生成以及三维模型的进化, 从而为提升三维模型解决实际问题的能力提供了新思路. 基于平行点云的思想, 本文建立了平行点云数据集 ShapeKITTI, 该数据集无需人工标注, 从而大幅降低了数据构建的成本, 我们利用 ShapeKITTI 数据集进行了三维



(a) 目标到激光雷达距离的分布

(a) The distribution of distance between the object and LiDAR sensor



(b) 目标角度的分布

(b) The distribution of orientation of objects

图 10 KITTI 数据集中的长尾分布

Fig.10 The long-tailed distribution in KITTI dataset

检测模型的训练, 证实了虚实结合的点云生成方法的有效性. 最后, 借助平行执行构成了数据生成与模型训练评估的闭环, 通过优化数据生成, 使模型性能得到了持续进化.

## References

- 1 He K M, Zhang X Y, Ren S Q, Sun J. Residual learning for image recognition. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA: IEEE, 2016. 770-778
- 2 Huang G, Liu Z, Van Der Maaten L, Weinberger K Q. Densely connected convolutional networks. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE, 2017. 2261-2269
- 3 Ren S Q, He K M, Girshick R, Sun J. Faster R-CNN: Towards real-time object detection with region proposal networks. In: Proceedings of the 2015 Annual Conference on Neural Information Processing Systems. Montreal, Canada: 2015. 91-99
- 4 Zhu Z, Wu W, Zou W, Yan J J. End-to-end flow correlation tracking with spatial-temporal attention. In: Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018. 548-557
- 5 Shen T Y, Wang J G, Gou C, Wang F-Y. Hierarchical fused model with deep learning and type-2 fuzzy learning for breast cancer diagnosis. *IEEE Transactions on Fuzzy Systems*, 2020, DOI: 10.1109/TFUZZ.2020.3013681
- 6 Shen T Y, Gou C, Wang F-Y, He Z L, Chen W G. Learning from adversarial medical images for X-ray breast mass segmentation. *Computer Methods and Programs in Biomedicine*, 2019, **180**: 105012

- 7 Chen X Z, Ma H M, Wan J, Li B, Xia T. Multi-view 3D object detection network for autonomous driving. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE, 2017. 6526–6534
- 8 Wang C, Xu D F, Zhu Y K, Martín-Martín R, Lu C W, Li F F, Savarese S. Densfusion: 6D object pose estimation by iterative dense fusion. In: Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA: IEEE, 2019. 3343–3352
- 9 Geiger A, Lenz P, Urtasun R. Are we ready for autonomous driving? The KITTI vision benchmark suite. In: Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence, USA: IEEE, 2012. 3354–3361
- 10 Sun P, Kretschmar H, Dotiwalla X, Chouard A, Patnaik V, Tsui P, et al. Scalability in perception for autonomous driving: Waymo open dataset. In: Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, USA: IEEE, 2020. 2443–2451
- 11 Huang X Y, Wang P, Cheng X J, Zhou D F, Geng Q C, Yang R G. The apollo open dataset for autonomous driving and its application. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, **42**(10): 2702–2719
- 12 Wang Y, Chao W L, Garg D, Hariharan B, Campbell M, Weinberger K Q. Pseudo-LiDAR from visual depth estimation: Bridging the gap in 3D object detection for autonomous driving. In: Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA: IEEE, 2019. 8437–8445
- 13 Qian R, Garg D, Wang Y, You Y R, Belongie S, Hariharan B, et al. End-to-end pseudo-LiDAR for image-based 3D object detection. In: Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, USA: IEEE, 2020. 5880–5889
- 14 Elmadawi K, Abdelrazek M, Elsobky M, Eraqi H M, Zahran M. End-to-end sensor modeling for LiDAR Point Cloud. In: Proceedings of the 2019 IEEE Intelligent Transportation Systems Conference (ITSC). IEEE, 2019. 1619–1624
- 15 Fang J, Zhou D F, Yan F L, Zhao T T, Zhang F H, Ma Y, et al. Augmented LiDAR simulator for autonomous driving. *IEEE Robotics and Automation Letters*, 2020, **5**(2): 1931–1938
- 16 Wang Fei-Yue. Parallel system methods for management and control of complex systems. *Control and Decision*, 2004, **19**(5): 485–489, 514  
(王飞跃. 平行系统方法与复杂系统的管理和控制. *控制与决策*, 2004, **19**(5): 485–489, 514)
- 17 Wang F-Y. Parallel control and management for intelligent transportation systems: Concepts, architectures, and applications. *IEEE Transactions on Intelligent Transportation Systems*, 2010, **11**(3): 630–638
- 18 Wang Fei-Yue. Parallel control: A method for data-driven and computational control. *Acta Automatica Sinica*, 2013, **39**(4): 293–302  
(王飞跃. 平行控制: 数据驱动的计算控制方法. *自动化学报*, 2013, **39**(4): 293–302)
- 19 Guo Y L, Wang H Y, Hu Q Y, Liu H, Liu L, Bennamoun M. Deep learning for 3D point clouds: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020
- 20 Su H, Maji S, Kalogerakis E, Learned-Miller E. Multi-view convolutional neural networks for 3D shape recognition. In: Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV). Santiago, Chile: IEEE, 2015. 945–953
- 21 Riegler G, Osman Ulusoy A, Geiger A. OctNet: Learning deep 3D representations at high resolutions. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE, 2017. 6620–6629
- 22 Qi Charles R, Su H, Kaichun M, Guibas L J. Pointnet: Deep learning on point sets for 3D classification and segmentation. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE, 2017. 77–85
- 23 Wu B C, Wan A, Yue X Y, Keutzer K. SqueezeSeg: Convolutional neural nets with recurrent CRF for real-time road-object segmentation from 3D LiDAR point cloud. In: Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA). Brisbane, Australia: IEEE, 2018. 1887–1893
- 24 Hou J, Dai A, Nießner M. 3D-SIS: 3D semantic instance segmentation of RGB-D scans. In: Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA: IEEE, 2019. 4416–4425
- 25 Wang Z J, Lu F. VoxSegNet: Volumetric CNNs for semantic part segmentation of 3D shapes. *IEEE Transactions on Visualization and Computer Graphics*, 2020, **26**(9): 2919–2930
- 26 Lang A H, Vora S, Caesar H, Zhou L B, Yang J, Beijbom O. Pointpillars: Fast encoders for object detection from point clouds. In: Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA: IEEE, 2019. 12689–12697
- 27 Tian Y L, Wang K F, Wang Y, Tian Y L, Wang Z L, Wang F-Y. Adaptive and azimuth-aware fusion network of multimodal local features for 3D object detection. *Neurocomputing*, 2020, **411**: 32–44
- 28 Giancola S, Zarzar J, Ghanem B. Leveraging shape completion for 3D Siamese tracking. In: Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA: IEEE, 2019. 1359–1368
- 29 Mueller M, Smith N, Ghanem B. Context-aware correlation filter tracking. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE, 2017. 1387–1395
- 30 Caesar H, Bankiti V, Lang A H, Vora S, Liong V E, Xu Q, et al. nuScenes: A multimodal dataset for autonomous driving. In: Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, USA: IEEE, 2020. 11618–11628
- 31 Manivasagam S, Wang S, Wong K, Zeng W, Sazanovvich M, Tan S, Yang B, Ma W-C, Urtasun R. LiDARsim: Realistic LiDAR simulation by leveraging the real world. In: Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020. 11167–11176
- 32 Prendinger H, Gajananan K, Zaki A B, Fares A, Molenaar R, Urbano D, et al. Tokyo virtual living lab: Designing smart cities based on the 3D Internet. *IEEE Internet Computing*, 2013, **17**(6): 30–38
- 33 Ros G, Sellart L, Materzynska J, Vazquez D, López A M. The SYNTHIA dataset: A large collection of synthetic images for semantic segmentation of urban scenes. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA: IEEE, 2016. 3234–3243
- 34 Gaidon A, Wang Q, Cabon Y, Vig E. VirtualWorlds as proxy for multi-object tracking analysis. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA: IEEE, 2016. 4340–4349
- 35 Li X, Wang K F, Tian Y L, Yan L, Deng F, Wang F Y. The ParallelEye dataset: A large collection of virtual images for traffic vision research. *IEEE Transactions on Intelligent Transportation Systems*, 2019, **20**(6): 2072–2084
- 36 Tian Y L, Li X, Wang K F, Wang F-Y. Training and testing object detectors with virtual images. *IEEE/CAA Journal of Automatica Sinica*, 2018, **5**(2): 539–546
- 37 Johnson-Roberson M, Barto C, Mehta R, Sridhar S N, Rosaen K, Vasudevan R. Driving in the matrix: Can virtual worlds replace human-generated annotations for real world tasks? In: Proceedings of the 2017 IEEE International Conference on Robotics and Automation (ICRA). Singapore: IEEE, 2017. 746–753
- 38 Yue X Y, Wu B C, Seshia S A, Keutzer K, Sangiovanni-Vincentelli A L. A LiDAR point cloud generator: From a virtual world to autonomous driving. In: Proceedings of the 2018 ACM International Conference on Multimedia Retrieval. Yokohama, Ja-

pan: ACM, 2018. 458–464

- 39 Li W, Pan C W, Zhang R, Ren J P, Ma Y X, Fang J, et al. AADS: Augmented autonomous driving simulation using data-driven algorithms. *Science Robotics*, 2019, 4(28): arXiv: 1901.07849v3
- 40 Wang Kun-Feng, Gou Chao, Wang Fei-Yue. Parallel vision: An ACP-based approach to intelligent vision computing. *Acta Automatica Sinica*, 2016, 42(10): 1490–1500  
(王坤峰, 苟超, 王飞跃. 平行视觉: 基于 ACP 的智能视觉计算方法. *自动化学报*, 2016, 42(10): 1490–1500)
- 41 Wang Kun-Feng, Lu Yue, Wang Yu-Tong, Xiong Zi-Wei, Wang Fei-Yue. Parallel imaging: A new theoretical framework for image generation. *Pattern Recognition and Artificial Intelligence*, 2017, 30(7): 577–587  
(王坤峰, 鲁越, 王雨桐, 熊子威, 王飞跃. 平行图像: 图像生成的一个新型理论框架. *模式识别与人工智能*, 2017, 30(7): 577–587)
- 42 Li Li, Lin Yi-Lun, Cao Dong-Pu, Zheng Nan-Ning, Wang Fei-Yue. Parallel learning — A new framework for machine learning. *Acta Automatica Sinica*, 2017, 43(1): 1–8  
(李力, 林懿伦, 曹东璞, 郑南宁, 王飞跃. 平行学习 — 机器学习的一个新型理论框架. *自动化学报*, 2017, 43(1): 1–8)
- 43 Wang F-Y, Zheng N N, Cao D P, Martinez C M, Li L, Liu T. Parallel driving in CPSS: A unified approach for transport automation and vehicle intelligence. *IEEE/CAA Journal of Automatica Sinica*, 2017, 4(4): 577–587
- 44 Guo Chao, Lu Yue, Lin Yi-Lun, Zhuo Fan, Wang Fei-Yue. Parallel art: Artistic creation under human-machine collaboration. *Chinese Journal of Intelligent Science and Technology*, 2019, 1(4): 335–341  
(郭超, 鲁越, 林懿伦, 卓凡, 王飞跃. 平行艺术: 人机协作的艺术创作. *智能科学与技术学报*, 2019, 1(4): 335–341)
- 45 Lv Yi-Sheng, Chen Yuan-Yuan, Jin Jun-Chen, Li Zhen-Jiang, Ye Pei-Jun, Zhu Feng-Hua. Parallel transportation: Virtual-real interaction for intelligent traffic management and control. *Chinese Journal of Intelligent Science and Technology*, 2019, 1(1): 21–33  
(吕宜生, 陈圆圆, 金峻臣, 李镇江, 叶佩军, 朱风华. 平行交通: 虚实互动的智能交通管理与控制. *智能科学与技术学报*, 2019, 1(1): 21–33)
- 46 王飞跃. 平行传感与平行感知: 平行传感器的原理、框架与设计. 青岛智能产业技术研究院报告, 青岛, 中国, 2015.
- 47 Zhang H, Luo G Y, Tian Y L, Wang K F, He H B, Wang F-Y, et al. A virtual-real interaction approach to object instance segmentation in traffic scenes. *IEEE Transactions on Intelligent Transportation Systems*, 2020, DOI: 10.1109/TITS.2019.2961145.
- 48 Goodfellow I. NIPS 2016 tutorial: Generative adversarial networks. arXiv: 1701.00160, 2016.
- 49 Dosovitskiy A, Ros G, Codevilla F, López A M, Koltun V. CARLA: An open urban driving simulator. In: Proceedings of the 1st Annual Conference on Robot Learning. Mountain View, USA: PMLR, 2017. 1–16
- 50 Rong G D, Shin B H, Tabatabae H, Lu Q, Lemke S, Možeiko M, et al. LGSVL simulator: A high fidelity simulator for autonomous driving. arXiv: 2005.03778, 2020.
- 51 Von Neumann-Cosel K, Dupuis M, Weiss C. Virtual test drive provision of a consistent tool-set for [D, H, S, V]-in-the-loop. In: Proceedings of the Driving Simulation Conference. Monaco, 2009.
- 52 You Y R, Wang Y, Chao W L, Garg D, Pleiss G, Hariharan B, et al. Pseudo-LiDAR++: Accurate depth for 3D object detection in autonomous driving. In: Proceedings of the 8th International Conference on Learning Representations. Addis Ababa, Ethiopia: OpenReview, 2020.
- 53 Chang A X, Funkhouser T, Guibas L, Hanrahan P, Huang Q X, Li Z M, et al. ShapeNet: An information-rich 3D model repository. arXiv: 1512.03012, 2015.
- 54 Yan Y, Mao Y X, Li B. SECOND: Sparsely embedded convolutional detection. *Sensors*, 2018, 18(10): 3337



**田永林** 中国科学技术大学与中国科学院自动化研究所联合培养博士研究生. 2017 年获得中国科学技术大学自动化系学士学位. 主要研究方向为计算机视觉, 智能交通.

E-mail: tyldyx@mail.ustc.edu.cn

(**TIAN Yong-Lin** Ph.D. candidate

in the Department of Automation, University of Science and Technology of China and the Institute of Automation, Chinese Academy of Sciences. He received his bachelor degree from University of Science and Technology of China in 2017. His research interest covers computer vision and intelligent transportation system.)

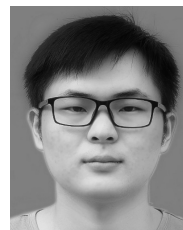


**沈宇** 中国科学院大学人工智能学院博士研究生. 主要研究方向为视频预测, 平行强化学习, 机器人自主导航.

E-mail: shenyu2015@ia.ac.cn

(**SHEN Yu** Ph.D. candidate at the

School of Artificial Intelligence, University of Chinese Academy of Sciences. His research interest covers video prediction, parallel reinforcement learning, and autonomous navigation of robots.)



**李强** 中国科学院大学人工智能学院硕士研究生. 主要研究方向为图像识别, 平行无人车系统.

E-mail: liqiang2018@ia.ac.cn

(**LI Qiang** Master student at the

School of Artificial Intelligence, University of Chinese Academy of Sciences. His research interest covers image identification and parallel unmanned vehicle system.)



**王飞跃** 中国科学院自动化研究所复杂系统管理与控制国家重点实验室研究员. 主要研究方向为智能系统和复杂系统的建模、分析与控制. 本文通信作者.

E-mail: feiyue.wang@ia.ac.cn

(**WANG Fei-Yue** Professor at the

State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences. His research interest covers modeling, analysis, and control of intelligent systems and complex systems. Corresponding author of this paper.)