(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 117214860 B
(45)授权公告日 2024.04.19

(21)申请号 202311027060.8
(22)申请日 2023.08.14
(65)同一申请的已公布的文献号
申请公布号 CN 117214860 A
(43)申请公布日 2023.12.12
(73)专利权人 北京科技大学顺德创新学院
地址 528000 广东省佛山市顺德区大良致慧路2号
(72)发明人 曾慧 叶一彬 李擎 刘启越 杨清港
(74)专利代理机构 北京市广友专利事务所有限责任公司 11237
专利代理师 张仲波 付忠林
(51)Int.Cl.
G01S 7/48(2006.01)
G06T 7/73(2017.01)
G06N 3/045(2023.01)
G06N 3/084(2023.01)
G01S 17/89(2020.01)
G01S 17/08(2006.01)
G01C 21/00(2006.01)
G01C 22/00(2006.01)

(56)对比文件
CN 114663509 A,2022.06.24
CN 114743105 A,2022.07.12
US 2021150228 A1,2021.05.20
CN 111476822 A,2020.07.31
CN 103247075 A,2013.08.14
CN 113284173 A,2021.08.20
KR 20220081261 A,2022.06.15
CN 114663496 A,2022.06.24

审查员 张静

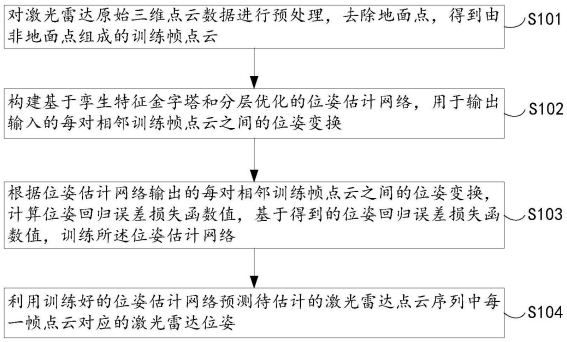
权利要求书3页 说明书10页 附图5页

(54)发明名称

基于孪生特征金字塔和地面分割的激光雷达里程计方法

(57)摘要

本发明提供一种基于孪生特征金字塔和地面分割的激光雷达里程计方法,属于计算机视觉技术领域。所述方法包括:对激光雷达原始三维点云数据进行预处理,去除地面点,得到由非地面点组成的训练帧点云;构建基于孪生特征金字塔和分层优化的位姿估计网络,用于输出输入的每对相邻训练帧点云之间的位姿变换;根据位姿估计网络输出的每对相邻训练帧点云之间的位姿变换,计算位姿回归误差损失函数值,基于得到的位姿回归误差损失函数值,训练所述位姿估计网络;利用训练好的位姿估计网络预测待估计的激光雷达点云序列中每一帧点云对应的激光雷达位姿。采用本发明,能够提高基于激光雷达进行位姿估计任务的精度。



CN 117214860 B

1.一种基于孪生特征金字塔和地面分割的激光雷达里程计方法,其特征在于,包括:

对激光雷达原始三维点云数据进行预处理,去除地面点,得到由非地面点组成的训练帧点云;

构建基于孪生特征金字塔和分层优化的位姿估计网络,用于输出输入的每对相邻训练帧点云之间的位姿变换;

根据位姿估计网络输出的每对相邻训练帧点云之间的位姿变换,计算位姿回归误差损失函数值,基于得到的位姿回归误差损失函数值,训练所述位姿估计网络;

利用训练好的位姿估计网络预测待估计的激光雷达点云序列中每一帧点云对应的激光雷达位姿;

其中,所述对激光雷达原始三维点云数据进行预处理,去除地面点,得到由非地面点组成的训练帧点云包括:

使用地面分割算法对激光雷达原始三维点云数据进行预处理,去除地面点,得到由非地面点组成的训练帧点云;

其中,所述地面分割算法由点云全景分割网络Cylinder3D实现,相邻的两帧点云PC_{t-1}和PC_t分别输入Cylinder3D,Cylinder3D输出逐点的分割标签,分为地面点与非地面点两类,并剔除点云中的非地面点,得到由非地面点组成的训练帧点云PC'_{t-1}和PC'_t,其中,PC'_{t-1}和PC'_t分别表示经过地面分割预处理后的第t-1帧和第t帧点云;

其中,所述位姿估计网络包括:孪生特征金字塔、场景流融合编码模块和位姿分层优化模块;

所述孪生特征金字塔,用于对经过地面分割预处理后的点云PC'_{t-1}和PC'_t进行编码,得到特征向量f_{t-1}和f_t;其中,f_{t-1}和f_t分别表示第t-1帧和第t帧的点云PC'_{t-1}和PC'_t经过孪生特征金字塔输出得到的特征向量;

所述场景流融合编码模块,用于关联孪生特征金字塔编码得到的特征向量f_{t-1}和f_t,并联合几何和语义特征,预测初始相对位姿估计;

所述位姿分层优化模块,用于增量式地优化初始相对位姿估计,通过关注不同尺度下点云特征信息的变化,进行位姿估计的更新。

2.根据权利要求1所述的基于孪生特征金字塔和地面分割的激光雷达里程计方法,其特征在于,所述孪生特征金字塔包括:2个子特征金字塔,2个子特征金字塔的所有网络层共享权值;

每个子孪生特征金字塔由3个不同规模的MBConv3D模块堆叠而成;

每个MBConv3D模块包括:维度扩张单元、卷积单元、压缩激发网络SENet和维度收缩单元;其中,维度扩张单元、卷积单元、压缩激发网络SENet和维度收缩单元依次相连;

维度扩张单元为Shared MLP→BatchNorm→Swish,卷积单元为KPConv→BatchNorm→Swish,维度收缩单元为Shared MLP→BatchNorm;其中,Shared MLP表示共享多层感知器,BatchNorm表示批归一化层,Swish为激活函数,KPConv表示核点卷积层,→表示连接;

在所述维度扩张单元中,Shared MLP通过多层感知机对点云中局部邻域点的特征进行升维;其中,局部邻域点是使用最远点采样方法得到的;经过Shared MLP输出的特征f_{mlp}表示为:

$$f_{mlp} = sharedMLP\left(\left(x_i^k - x_i\right) \oplus f_i^k \oplus f_i\right), k = 1,2,\cdots,K$$

其中，x_i是最远点采样得到的第i个采样点，x_i^k表示在x_i周围的第k个近邻点，f_i和f_i^k分别代表x_i和x_i^k的特征，K表示x_i周围存在的近邻点个数，⊕表示向量的串联操作，sharedMLP()表示多层感知机；

所述KPConv，用于提取局部区域的点特征并将提取到的点特征进行融合，在点x_i处，以g为卷积核的KPConv输出的特征f_{kp}表示为：

$$f_{kp} = \sum_{x_i^n \in \mathcal{N}_x} g(x_i^n - x_i) f_i^n$$

其中，$\mathcal{N}_x = \{x_i^n \mid \|x_i^n - x_i\| \le r\}$表示以点$x_i$为中心、r为半径的卷积区域，$x_i^n$表示在该卷积区域中的第n个点，$f_i^n$表示$x_i^n$的特征，卷积核g在该卷积区域的不同位置具有不同的核函数权重；

所述SENet是针对通道的注意力机制模块，通过压缩和激发特征通道来增强特征；

在所述维度收缩单元中，Shared MLP通过多层感知机降低SENet输出的特征的维度。

3.根据权利要求1所述的基于孪生特征金字塔和地面分割的激光雷达里程计方法，其特征在于，所述场景流融合编码模块包括FlowNet3D、Shared MLP和初始位姿估计子模块；其中，FlowNet3D表示场景流估计网络，Shared MLP表示共享多层感知器；

所述FlowNet3D，用于关联孪生特征金字塔编码得到的特征向量f_{t-1}和f_t，通过学习点云的软对应点关系生成场景流嵌入特征$Flow_0$；

所述Shared MLP使用孪生特征金字塔生成的特征向量f_{t-1}和f_t和场景流嵌入特征$Flow_0$作为输入，输出包含每个点的加权系数的掩码，以降低场景中动态物体点在位姿估计中的影响；

在所述初始位姿估计子模块中，Shared MLP输出的掩码经过Softmax归一化后与场景流嵌入特征$Flow_0$加权求和，求和结果输入全连接层预测初始相对位姿估计，初始的相对位姿向量用平移向量t_0和四元数q_0表示。

4.根据权利要求3所述的基于孪生特征金字塔和地面分割的激光雷达里程计方法，其特征在于，所述位姿分层优化模块包括：第一位姿优化子模块和第二位姿优化子模块，第一位姿优化子模块输出的优化后的相对位姿向量(t_1, q_1)和场景流嵌入特征$Flow_1$为第二位姿优化子模块的输入，第二位姿优化子模块用于输出优化后的相对位姿向量(t_2, q_2)和场景流嵌入特征$Flow_2$；每个位姿优化子模块包括：位姿变换单元、场景流更新与编码单元和位姿更新单元；

所述位姿变换单元，用于通过输入对应的平移向量t_{in}和四元数q_{in}刚性地调整源点云的位置和朝向；其中，位姿变换的过程表示为：

$$[0, xyz'_{t-1}] = q_{in}[0, xyz_{t-1}]q_{in}^{-1} + [0, t_{in}]$$

其中，$xyz_{t-1} \in N \times 3$表示包含了N个点的源点云中所有点的空间坐标集合，$xyz'_{t-1} \in N \times 3$表示变换后的源点云的点空间坐标，$q_{in}$和$t_{in}$为输入位姿；

所述场景流更新与编码单元,用于将经过位姿变换后的源点云与目标点云共同输入FlowNet3D生成场景流嵌入特征,场景流更新与编码单元输入的场景流嵌入特征$Flow_{in}$经过上采样后,与该场景流嵌入特征共同输入SharedMLP进行更新,得到场景流嵌入特征$Flow_{out}$,更新后的场景流嵌入特征$Flow_{out}$再通过SharedMLP生成掩码,并经过Softmax归一化后与场景流嵌入特征$Flow_{out}$加权求和,求和结果输入全连接层,全连接层将加权特征映射为位姿增量Δq和Δt;其中,掩码包含了该尺度下每个点的加权系数;

所述位姿更新单元,用于根据位姿增量对输入的相对位姿向量q_{in}和t_{in}进行更新,其中,位姿更新过程表示为:

$[0,t_{out}]=\Delta q[0,t_{in}]\Delta q^{-1}+[0,\Delta t]$

$q_{out}=\Delta q q_{in}$

其中,Δt和Δq表示场景流更新与编码单元预测的平移向量和四元数的增量,t_{out}和q_{out}表示经过更新后的平移向量和四元数。

5.根据权利要求4所述的基于孪生特征金字塔和地面分割的激光雷达里程计方法,其特征在于,所述根据位姿估计网络输出的每对相邻训练帧点云之间的位姿变换,计算位姿回归误差损失函数值,基于得到的位姿回归误差损失函数值,训练所述位姿估计网络包括:

利用位姿估计网络在3个级别输出的相对位姿向量t_0和q_0、t_1和q_1以及t_2和q_2来监督位姿估计网络进行训练,对于第n级输出的相对位姿向量,计算帧间配准损失函数$\mathbb{L}_{odo}^{n}$:

$$\mathbb{L}_{odo}^{n}=\|\hat{t}-t_n\|\exp(-s_x)+s_x+\left\|\hat{q}-\frac{q_n}{\|q_n\|}\right\|_2\exp(-s_q)+s_q$$

其中,n=0,1,2,$\hat{t}$和$\hat{q}$分别表示由真实的位姿变换矩阵生成的平移向量和四元数,t_n和q_n表示网络在第n级的平移向量和四元数输出,||·||和$||\cdot||_2$分别表示$\mathcal{L}_1$范数和$\mathcal{L}_2$范数,s_x和S_q分别表示平移和旋转的尺度因子;

根据计算得到的帧间配准损失函数$\mathbb{L}_{odo}^{n}$,计算位姿回归误差损失函数$\mathbb{L}_{odo}^{total}$:

$$\mathbb{L}_{odo}^{total}=\sum_{n=0}^{2}\lambda^n\mathbb{L}_{odo}^{n}$$

其中,λ^n表示第n级的位姿损失权重,$\mathbb{L}_{odo}^{n}$表示第n级的位姿损失,$\mathbb{L}_{odo}^{total}$表示激光雷达里程计的总位姿回归损失。

基于孪生特征金字塔和地面分割的激光雷达里程计方法

技术领域

[0001] 本发明涉及计算机视觉技术领域,特别是指一种基于孪生特征金字塔和地面分割的激光雷达里程计方法。

背景技术

[0002] 同时定位与建图(SLAM)是移动机器人研究领域的关键技术之一。经典的SLAM系统通常包括传感器数据读取、前端里程计、后端优化、回环检测、建图等五个部分。里程计作为SLAM系统中的重要步骤之一,其任务是利用传感器采集到的数据估计机器人的运动轨迹。基于视觉和基于激光雷达的里程计方法较为常见。基于视觉的里程计方法容易受到光照、天气等因素的影响较大,导致位姿估计精度较低,而激光雷达能够直接获取机器人周身360°的环境深度信息,这使得基于激光雷达的里程计在许多应用场景中更具有鲁棒性。

发明内容

[0003] 本发明实施例提供了基于孪生特征金字塔和地面分割的激光雷达里程计方法,能够提高基于激光雷达进行位姿估计任务的精度。所述技术方案如下:

[0004] 一方面,提供了一种基于孪生特征金字塔和地面分割的激光雷达里程计方法,该方法应用于电子设备,该方法包括:

[0005] 对激光雷达原始三维点云数据进行预处理,去除地面点,得到由非地面点组成的训练帧点云;

[0006] 构建基于孪生特征金字塔和分层优化的位姿估计网络,用于输出输入的每对相邻训练帧点云之间的位姿变换;

[0007] 根据位姿估计网络输出的每对相邻训练帧点云之间的位姿变换,计算位姿回归误差损失函数值,基于得到的位姿回归误差损失函数值,训练所述位姿估计网络;

[0008] 利用训练好的位姿估计网络预测待估计的激光雷达点云序列中每一帧点云对应的激光雷达位姿。

[0009] 进一步地,所述对激光雷达原始三维点云数据进行预处理,去除地面点,得到由非地面点组成的训练帧点云包括:

[0010] 使用地面分割算法对激光雷达原始三维点云数据进行预处理,去除地面点,得到由非地面点组成的训练帧点云;

[0011] 其中,所述地面分割算法由点云全景分割网络Cylinder3D实现,相邻的两帧点云PC_{t-1}和PC_t分别输入Cylinder3D,Cylinder3D输出逐点的分割标签,分为地面点与非地面点两类,并剔除点云中的非地面点,得到由非地面点组成的训练帧点云PC'_{t-1}和PC'_t,其中,PC'_{t-1}和PC'_t分别表示经过地面分割预处理后的第t-1帧和第t帧点云。

[0012] 进一步地,所述位姿估计网络包括:孪生特征金字塔、场景流融合编码模块和位姿分层优化模块;

[0013] 所述孪生特征金字塔,用于对经过地面分割预处理后的点云PC'_{t-1}和PC'_t进行编

码,得到特征向量f_{t-1}和f_t;其中,f_{t-1}和f_t分别表示第t-1帧和第t帧的点云PC'_{t-1}和PC'_t经过孪生特征金字塔输出得到的特征向量;

[0014] 所述场景流融合编码模块,用于关联孪生特征金字塔编码得到的特征向量f_{t-1}和f_t,并联合几何和语义特征,预测初始相对位姿估计;

[0015] 所述位姿分层优化模块,用于增量式地优化初始相对位姿估计,通过关注不同尺度下点云特征信息的变化,进行位姿估计的更新。

[0016] 进一步地,所述孪生特征金字塔包括:2个子特征金字塔,2个子特征金字塔的所有网络层共享权值;

[0017] 每个子孪生特征金字塔由3个不同规模的MBConv3D模块堆叠而成;

[0018] 每个MBConv3D模块包括:维度扩张单元、卷积单元、压缩激发网络SENet和维度收缩单元;其中,维度扩张单元、卷积单元、压缩激发网络SENet和维度收缩单元依次相连;

[0019] 维度扩张单元为Shared MLP→BatchNorm→Swish,卷积单元为KPConv→BatchNorm→Swish,维度收缩单元为Shared MLP→BatchNorm;其中,Shared MLP表示共享多层感知器,BatchNorm表示批归一化层,Swish为激活函数,KPConv表示核点卷积层,→表示连接;

[0020] 在所述维度扩张单元中,Shared MLP通过多层感知机对点云中局部邻域点的特征进行升维;其中,局部邻域点是使用最远点采样方法得到的;经过Shared MLP输出的特征f_{mlp}表示为:

[0021] $$f_{mlp} = sharedMLP\left(\left(x_i^k - x_i\right) \oplus f_i^k \oplus f_i\right), k = 1,2,\cdots,K$$

[0022] 其中,x_i是最远点采样得到的第i个采样点,x_i^k表示在x_i周围的第k个近邻点,f_i和f_i^k分别代表x_i和x_i^k的特征,K表示x_i周围存在的近邻点个数,$\oplus$表示向量的串联操作,sharedMLP()表示多层感知机;

[0023] 所述KPConv,用于提取局部区域的点特征并将提取到的点特征进行融合,在点x_i处,以g为卷积核的KPConv输出的特征f_{kp}表示为:

[0024] $$f_{kp} = \sum_{x_i^n \in \mathcal{N}_x} g(x_i^n - x_i) f_i^n$$

[0025] 其中,$\mathcal{N}_x = \{x_i^n \mid \|x_i^n - x_i\| \le r\}$表示以点$x_i$为中心、r为半径的卷积区域,$x_i^n$表示在该卷积区域中的第n个点,$f_i^n$表示$x_i^n$的特征,卷积核g在该卷积区域的不同位置具有不同的核函数权重;

[0026] 所述SENet是针对通道的注意力机制模块,通过压缩和激发特征通道来增强特征;

[0027] 在所述维度收缩单元中,Shared MLP通过多层感知机降低SENet输出的特征的维度。

[0028] 进一步地,所述场景流融合编码模块包括FlowNet3D、Shared MLP和初始位姿估计子模块;

[0029] 所述FlowNet3D,用于关联孪生特征金字塔编码得到的特征向量f_{t-1}和f_t,通过学习点云的软对应点关系生成场景流嵌入特征$Flow_0$;

[0030] 所述Shared MLP使用孪生特征金字塔生成的特征向量f_{t-1}和f_t和场景流嵌入特征$Flow_0$作为输入,输出包含每个点的加权系数的掩码,以降低场景中动态物体点在位姿估计中的影响;

[0031] 在所述初始位姿估计子模块中,Shared MLP输出的掩码经过Softmax归一化后与场景流嵌入特征$Flow_0$加权求和,求和结果输入全连接层预测初始相对位姿估计,初始的相对位姿向量用平移向量t_0和四元数q_0表示。

[0032] 进一步地,所述位姿分层优化模块包括:第一位姿优化子模块和第二位姿优化子模块,第一位姿优化子模块输出的优化后的相对位姿向量(t_1,q_1)和场景流嵌入特征$Flow_1$为第二位姿优化子模块的输入,第二位姿优化子模块用于输出优化后的相对位姿向量(t_2,q_2)和场景流嵌入特征$Flow_2$;每个位姿优化子模块包括:位姿变换单元、场景流更新与编码单元和位姿更新单元;

[0033] 所述位姿变换单元,用于通过输入对应的平移向量t_{in}和四元数q_{in}刚性地调整源点云的位置和朝向;其中,位姿变换的过程表示为:

[0034] $$[0, xyz'_{t-1}] = q_{in}[0, xyz_{t-1}]q_{in}^{-1} + [0, t_{in}]$$

[0035] 其中,$xyz_{t-1}\in N\times 3$表示包含了N个点的源点云中所有点的空间坐标集合,$xyz'_{t-1}\in N\times 3$表示变换后的源点云的点空间坐标,q_{in}和t_{in}为输入位姿;

[0036] 所述场景流更新与编码单元,用于将经过位姿变换后的源点云与目标点云共同输入FlowNet3D生成场景流嵌入特征,场景流更新与编码单元输入的场景流嵌入特征$Flow_{in}$经过上采样后,与该场景流嵌入特征共同输入SharedMLP进行更新,生成该尺度下的掩码,掩码经过Softmax归一化后与场景流嵌入特征$Flow_{out}$加权求和,求和结果输入全连接层,全连接层将加权特征映射为位姿增量Δq和Δt;其中,掩码包含了该尺度下每个点的加权系数;

[0037] 所述位姿更新单元,用于根据位姿增量对输入的相对位姿向量q_{in}和t_{in}进行更新,其中,位姿更新过程表示为:

[0038] $[0,t_{out}]=\Delta q[0,t_{in}]\Delta q^{-1}+[0,\Delta t]$

[0039] $q_{out}=\Delta q q_{in}$

[0040] 其中,Δt和Δq表示场景流更新与编码单元预测的平移向量和四元数的增量,t_{out}和q_{out}表示经过更新后的平移向量和四元数。

[0041] 进一步地,所述根据位姿估计网络输出的每对相邻训练帧点云之间的位姿变换,计算位姿回归误差损失函数值,基于得到的位姿回归误差损失函数值,训练所述位姿估计网络包括:

[0042] 利用位姿估计网络在3个级别输出的相对位姿向量t_0和q_0、t_1和q_1以及t_2和q_2来监督位姿估计网络进行训练,对于第n级输出的相对位姿向量,计算帧间配准损失函数$\mathbb{L}_{odo}^{n}$:

[0043] $$\mathbb{L}_{odo}^{n} = \|\hat{t} - t_n\| \exp(-s_x) + s_x + \left\|\hat{q} - \frac{q_n}{\|q_n\|}\right\|_2 \exp(-s_q) + s_q$$

[0044] 其中,n=0,1,2,$\hat{t}$和$\hat{q}$分别表示由真实的位姿变换矩阵生成的平移向量和四元数,tn和q_n表示网络在第n级的平移向量和四元数输出,||·||和$||\cdot||_2$分别表示$\mathcal{L}_1$范数

和$\mathcal{L}_2$范数，s_x和s_q分别表示平移和旋转的尺度因子；

[0045] 根据计算得到的帧间配准损失函数$\mathbb{L}_{odo}^{n}$，计算位姿回归误差损失函数$\mathbb{L}_{odo}^{total}$：

[0046] $$\mathbb{L}_{odo}^{total}=\sum_{n=0}^{2}\lambda^{n}\mathbb{L}_{odo}^{n}$$

[0047] 其中，λ^{n}表示第n级的位姿损失权重，$\mathbb{L}_{odo}^{n}$表示第n级的位姿损失，$\mathbb{L}_{odo}^{total}$表示激光雷达里程计的总位姿回归损失。

[0048] 一方面，提供了一种电子设备，所述电子设备包括处理器和存储器，所述存储器中存储有至少一条指令，所述至少一条指令由所述处理器加载并执行以实现上述基于孪生特征金字塔和地面分割的激光雷达里程计方法。

[0049] 一方面，提供了一种计算机可读存储介质，所述存储介质中存储有至少一条指令，所述至少一条指令由处理器加载并执行以实现上述基于孪生特征金字塔和地面分割的激光雷达里程计方法。

[0050] 本发明实施例提供的技术方案带来的有益效果至少包括：

[0051] 本发明实施例中，对激光雷达原始三维点云数据进行预处理，去除地面点，得到由非地面点组成的训练帧点云；构建基于孪生特征金字塔和分层优化的位姿估计网络，用于输出输入的每对相邻训练帧点云之间的位姿变换；根据位姿估计网络输出的每对相邻训练帧点云之间的位姿变换，计算位姿回归误差损失函数值，基于得到的位姿回归误差损失函数值，训练所述位姿估计网络；利用训练好的位姿估计网络预测待估计的激光雷达点云序列中每一帧点云对应的激光雷达位姿。这样，能够提高基于激光雷达进行位姿估计任务的精度。

附图说明

[0052] 为了更清楚地说明本发明实施例中的技术方案，下面将对实施例描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本发明的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他的附图。

[0053] 图1为本发明实施例提供的基于孪生特征金字塔和地面分割的激光雷达里程计方法的流程示意图；

[0054] 图2为本发明实施例提供的基于孪生特征金字塔和地面分割的激光雷达里程计方法的整体框架示意图；

[0055] 图3为本发明实施例提供的MBConv3D模块的结构示意图；

[0056] 图4为本发明实施例提供的场景流融合编码模块的结构示意图；

[0057] 图5为本发明实施例提供的位姿优化子模块的结构示意图；

[0058] 图6(a)为本发明实施例提供的方法在KITTI里程计数据集中序列09上估计的轨迹示意图；

[0059] 图6(b)为本发明实施例提供的方法在KITTI里程计数据集中序列10上估计的轨迹示意图；

[0060] 图7是本发明实施例提供的一种电子设备的结构示意图。

具体实施方式

[0061] 为使本发明的目的、技术方案和优点更加清楚，下面将结合附图对本发明实施方式作进一步地详细描述。

[0062] 如图1所示，本发明实施例提供了一种基于孪生特征金字塔和地面分割的激光雷达里程计方法，该方法可以由电子设备实现，该电子设备可以是终端或服务器，该方法包括：

[0063] S101，对激光雷达原始三维点云数据进行预处理，去除地面点，得到由非地面点组成的训练帧点云；

[0064] 本实施例中，如图2所示，使用地面分割算法对激光雷达原始三维点云数据进行预处理，去除地面点，得到由非地面点组成的训练帧点云；

[0065] 其中，所述地面分割算法由点云全景分割网络Cylinder3D实现，相邻的两帧点云PC_{t-1}和PC_t分别输入Cylinder3D，Cylinder3D输出逐点的分割标签，分为“地面点”与“非地面点”两类，并剔除点云中的非地面点，得到由非地面点组成的训练帧点云PC'_{t-1}和PC'_t，其中，PC'_{t-1}和PC'_t分别表示经过地面分割预处理后的第t-1帧和第t帧点云。

[0066] S102，构建基于孪生特征金字塔和分层优化的位姿估计网络，用于输出输入的每对相邻训练帧点云之间的位姿变换；

[0067] 本实施例中，所述位姿估计网络包括：孪生特征金字塔、场景流融合编码模块和位姿分层优化模块，位姿估计网络的详细结构如表1所示；其中，

[0068] 所述孪生特征金字塔，用于对经过地面分割预处理后的点云PC'_{t-1}和PC'_t进行编码，得到特征向量f_{t-1}和f_t；其中，f_{t-1}和f_t分别表示第t-1帧和第t帧的点云PC'_{t-1}和PC'_t经过孪生特征金字塔输出得到的特征向量；

[0069] 所述场景流融合编码模块，用于关联孪生特征金字塔编码得到的特征向量f_{t-1}和f_t，并联合几何和语义特征，预测初始相对位姿估计；

[0070] 所述位姿分层优化模块，用于增量式地优化初始相对位姿估计，通过关注不同尺度下点云特征信息的变化，进行位姿估计的更新。

[0071] 表1位姿估计网络结构

[0072]

Module	Operator	In channel	Kernel size	Output Shape
孪生特征金字塔	MBConv3D	3	1	2048×16
	MBConv3D	16	2	1024×32
	MBConv3D	32	2	256×64
场景流融合编码模块	FlowNet3D	131	1	256×64
	SharedMLP	128	1	256×128
	FC	64		1×7
位姿分层优化模块	FlowNet3D	131	1	1024×64
	SetUpconv	32	1	1024×64
	SharedMLP	128	1	1024×64
	FC	64		1×7
	FlowNet3D	131	1	2048×64
	SetUpconv	16	1	2048×64
	SharedMLP	128	1	2048×64
	FC	64		1×7

[0073] 表1中,FC表示全连接层。

[0074] 本实施例中,所述孪生特征金字塔包括:2个子特征金字塔,2个子特征金字塔的所有网络层共享权值;

[0075] 其中一个子孪生特征金字塔对经过预处理的点云PC'_t进行编码,得到特征向量f_t:

[0076] $f_t=Pyramid(PC'_t)$

[0077] 其中,Pyramid()为子孪生特征金字塔;

[0078] 另一个子孪生特征金字塔对经过预处理的点云PC'_{t-1}进行编码,得到特征向量f_{t-1}:

[0079] $f_{t-1}=Pyramid(PC'_{t-1})$

[0080] 本实施例中,每个子孪生特征金字塔由3个不同规模的MBConv3D模块堆叠而成;MBConv3D模块的输入为经过预处理的点云中各点的空间坐标与对应的特征向量;

[0081] 每个MBConv3D模块包括:维度扩张单元、卷积单元、压缩激发网络SENet和维度收缩单元;其中,维度扩张单元、卷积单元、压缩激发网络SENet和维度收缩单元依次相连;

[0082] 如图3所示,维度扩张单元为Shared MLP→BatchNorm→Swish,卷积单元为KPConv→BatchNorm→Swish,维度收缩单元为Shared MLP→BatchNorm;其中,Shared MLP表示共享多层感知器,BatchNorm表示批归一化层,Swish为激活函数,KPConv表示核点卷积层,→表示连接;

[0083] 在所述维度扩张单元中,Shared MLP通过多层感知机对点云中局部邻域点的特征进行升维;其中,局部邻域点是使用最远点采样方法得到的;经过Shared MLP输出的特征f_{mlp}表示为:

[0084] $$f_{mlp}=sharedMLP\left(\left(x_i^k-x_i\right)\oplus f_i^k\oplus f_i\right),k=1,2,\cdots,K$$

[0085] 其中,x_i是最远点采样得到的第i个采样点,x_i^k表示在x_i周围的第k个近邻点,f_i和f_i^k分别代表x_i和x_i^k的特征,K表示x_i周围存在的近邻点个数,$\oplus$表示向量的串联操作,sharedMLP()表示多层感知机;

[0086] 所述KPConv,用于提取局部区域的点特征并将提取到的点特征进行融合,在点x_i处,以g为卷积核的KPConv输出的特征f_{kp}表示为:

[0087] $$f_{kp}=\sum_{x_i^n\in\mathcal{N}_x}g(x_i^n-x_i)f_i^n$$

[0088] 其中,$\mathcal{N}_x=\{x_i^n\mid\|x_i^n-x_i\|\le r\}$表示以点$x_i$为中心、r为半径的卷积区域,$x_i^n$表示在该卷积区域中的第n个点,$f_i^n$表示$x_i^n$的特征,卷积核g在该卷积区域的不同位置具有不同的核函数权重;

[0089] 所述SENet是针对通道的注意力机制模块,通过压缩和激发特征通道来增强特征;

[0090] 在所述维度收缩单元中,Shared MLP通过多层感知机降低SENet输出的特征的维度。

[0091] 由此可知,在孪生特征金字塔中,使用Shared MLP扩展特征向量的维度,KPConv提

取局部区域的点特征并融合,接着SENet通过压缩和激发特征通道来增强特征,最后再使用一层Shared MLP降低特征的维度并输出。

[0092] 如图4所示,场景流融合编码模块的输入为相邻两帧点云(即:源点云和目标点云)的空间坐标(xyz_{t-1}和xyz_t)与高维特征向量(f_{t-1}和f_t),输出为包含7个元素的初始相对位姿向量,前4个元素表示3自由度的旋转四元数q,后3个元素表示3自由度相对位移t,其中,xyz_{t-1}和xyz_t分别为第t-1帧和第t帧的点云PC'_{t-1}和PC'_t的空间坐标。

[0093] 本实施例中,所述场景流融合编码模块包括场景流估计网络FlowNet3D、Shared MLP和初始位姿估计子模块;

[0094] 所述FlowNet3D,用于关联孪生特征金字塔编码得到的特征向量f_{t-1}和f_t,通过学习点云的软对应点关系生成场景流嵌入特征$Flow_0$;其中,FlowNet3D生成的场景流嵌入特征$Flow_0$表示为:

[0095] $Flow_0=FlowNet3D(f_{t-1},f_t)$

[0096] 其中,f_{t-1}和f_t分别表示第t-1帧和第t帧的点云经过特征金字塔输出得到的特征向量,f_{flow}表示场景流嵌入特征,FlowNet3D()为所述FlowNet3D模块;

[0097] 所述Shared MLP使用孪生特征金字塔生成的特征向量f_{t-1}和f_t和场景流嵌入特征$Flow_0$作为输入,输出包含每个点(该点为点云中所有的点,包括:静态物体点和动态物体点)的加权系数的掩码。由于场景中可能同时存在可靠的静态物体和随机运动的动态物体,所以每个点对于全局位姿估计的参考价值并不一致,该Shared MLP为每个点分配一个加权系数,参考价值更高的静态物体点对应更大的权重,同时,动态物体点将会被分配较小的权重;

[0098] 在所述初始位姿估计子模块中,Shared MLP输出的掩码经过Softmax归一化后与场景流嵌入特征$Flow_0$加权求和,求和结果输入全连接层预测初始相对位姿估计,初始的相对位姿向量用平移向量t_0和四元数q_0表示。

[0099] 如图2所示,所述位姿分层优化模块包括:第一位姿优化子模块和第二位姿优化子模块,第一位姿优化子模块输出的优化后的相对位姿向量(t_1,q_1)和场景流嵌入特征$Flow_1$为第二位姿优化子模块的输入,第二位姿优化子模块用于输出优化后的相对位姿向量(t_2,q_2)和场景流嵌入特征$Flow_2$;每个位姿优化子模块的输入包括需要优化的相对位姿向量和场景流嵌入特征、源点云的空间坐标xyz_{t-1}、目标点云的空间坐标xyz_t、源点云的特征向量f_{t-1}和目标点云的特征向量f_t,输出优化后的相对位姿向量和场景流嵌入特征;

[0100] 每个位姿优化子模块包括:位姿变换单元、场景流更新与编码单元和位姿更新单元;

[0101] 如图5所示,所述位姿变换单元,用于通过输入对应的平移向量t_{in}和四元数q_{in}刚性地调整源点云的位置和朝向;其中,位姿变换的过程表示为:

[0102] $$\left[0, xyz'_{t-1}\right] = q_{in}[0, xyz_{t-1}]q_{in}^{-1} + [0, t_{in}]$$

[0103] 其中,$xyz_{t-1}\in N\times 3$表示包含了N个点的源点云中所有点的空间坐标集合,$xyz'_{t-1}\in N\times 3$表示变换后的源点云的点空间坐标,q_{in}和t_{in}为输入的相对位姿向量;

[0104] 所述场景流更新与编码单元,用于将经过位姿变换后的源点云与目标点云共同输入FlowNet3D生成场景流嵌入特征,场景流更新与编码单元输入的场景流嵌入特征$Flow_{in}$经

过上采样(Set Upconv)后,与该场景流嵌入特征共同输入SharedMLP进行更新,得到场景流嵌入特征$Flow_{out}$,更新后的场景流嵌入特征$Flow_{out}$再通过SharedMLP生成掩码,并经过Softmax归一化后与场景流嵌入特征$Flow_{out}$加权求和,求和结果输入全连接层,全连接层将加权特征映射为位姿增量Δq和Δt;其中,掩码包含了该尺度下每个点的加权系数,Set Upconv表示可学习的上采样层;

[0105] 所述位姿更新单元,用于根据位姿增量对输入的相对位姿向量q_{in}和t_{in}进行更新,其中,位姿更新过程表示为:

[0106] $[0,t_{out}]=\Delta q[0,t_{in}]\Delta q^{-1}+[0,\Delta t]$

[0107] $q_{out}=\Delta q q_{in}$

[0108] 其中,Δt和Δq表示场景流更新与编码单元预测的平移向量和四元数的增量,t_{out}和q_{out}表示经过更新后的平移向量和四元数。

[0109] S103,根据位姿估计网络输出的每对相邻训练帧点云之间的位姿变换,计算位姿回归误差损失函数值,基于得到的位姿回归误差损失函数值,训练所述位姿估计网络;

[0110] 在本实施例中,位姿估计网络在3个级别输出的相对位姿向量t_0和q_0、t_1和q_1、t_2和q_2都用来监督网络进行训练,对于第n(n=0,1,2)级输出的相对位姿向量,计算帧间配准损失函数$\mathbb{L}_{odo}^{n}$:

[0111] $$\mathbb{L}_{odo}^{n}=\|\hat{t}-t_n\|\exp(-s_x)+s_x+\left\|\hat{q}-\frac{q_n}{\|q_n\|}\right\|_2\exp(-s_q)+s_q$$

[0112] 其中,$\hat{t}$和$\hat{q}$分别表示由真实的位姿变换矩阵生成的平移向量和四元数,t_n和q_n表示网络在第n级的平移向量和四元数输出,||·||和$||\cdot||_2$分别表示$\mathcal{L}_1$范数和$\mathcal{L}_2$范数,s_x和s_q分别表示平移和旋转的尺度因子;

[0113] 接着,计算最终的位姿回归误差损失函数$\mathbb{L}_{odo}^{total}$:

[0114] $$\mathbb{L}_{odo}^{total}=\sum_{n=0}^{2}\lambda^n\mathbb{L}_{odo}^{n}$$

[0115] 其中,λ^n表示第n级的位姿损失权重,$\mathbb{L}_{odo}^{n}$表示第n级的位姿损失,$\mathbb{L}_{odo}^{total}$表示激光雷达里程计的总位姿回归损失。

[0116] 本实施例中,在训练时,将一个批次的激光雷达序列中的所有三维点云数据输入到点云全景分割网络Cylinder3D中执行数据预处理,得到去除地面点的训练帧点云再输入位姿估计网络中,对位姿估计网络进行训练;其中,根据位姿估计网络预测的每对相邻训练帧之间的相对位姿向量,计算位姿回归误差损失函数值,基于得到的位姿回归损失函数值,采取端到端的训练方式并通过反向传播来训练整个位姿估计网络。

[0117] S104,利用训练好的位姿估计网络预测待估计的激光雷达点云序列中每一帧点云对应的激光雷达位姿。

[0118] 在本实施例中,为了验证本发明实施例提供的激光雷达里程计方法的有效性,使用KITTI里程计数据集评估测试其性能:

[0119] (1)相对位移均方误差(Rel.trans.):一个序列中全部长度为100、200、……、800

米的子序列的平均位移RMSE(Root Mean Square Error),以%度量,即每100米偏差的米数,数值越小越好。

[0120] (2)相对旋转均方误差(Rel.rot.):一个序列中全部长度为100、200、……、800米的子序列的平均旋转RMSE,以deg/m度量,数值越小越好。

[0121] 在本实施例中,应用了KITTI里程计数据集中00-08这9个序列作为训练集与验证集训练位姿估计网络,并用09-10这两个序列测试所述的基于孪生特征金字塔和分层优化的位姿估计网络的性能。

[0122] KITTI里程计数据集是目前国际上主流的自动驾驶数据集之一,包含市区、乡村、高速公路等道路场景,数据集包含双目图像、激光雷达点云以及实际轨迹。

[0123] 在本实施例中,位姿回归误差损失函数的超参数$\lambda^0=0.8$,$\lambda^1=0.4$,$\lambda^2=0.2$,s_x和s_q参数的初始值分别设置为0.0和-2.5,并在训练过程中不断更新。位姿估计网络的训练过程中,初始学习率为10^{-3},并随着训练的进行逐渐减小,每经过10轮迭代,学习率变为上一轮的0.5倍,采用Adam优化器进行90次迭代,每轮迭代的批次大小为16,每批次包含16对相邻帧点云。

[0124] 为了验证本发明所述方法的性能,本实施例中,选择了基于传统方法和基于深度学习的激光雷达里程计方法进行了对比,实验结果如表2所示。本实施例在KITTI序列09、10中生成的轨迹分别如图6(a)和图6(b)所示,其中,虚线轨迹为真实的轨迹,实线轨迹为本实施例中估计出的轨迹。

[0125] 表2 KITTI数据集中本实施例方法与其他方法对比

[0126]

方法	09 Rel. trans.(%)	09 Rel. rot.(deg/m)	10 Rel. trans.(%)	10 Rel. rot.(deg/m)
ICP	3.95	1.71	6.13	2.60
GICP	1.97	0.77	1.31	0.62
LOAM(无后端优化)	8.10	4.30	12.67	8.79
LOAM(带后端优化)	**0.77**	0.48	0.79	0.57
LO-Net	1.37	0.58	1.80	0.93
ENCODE	0.93	**0.36**	0.83	**0.34**
DMLO	1.10	0.61	1.12	0.64
PWCLO-Net	0.79	0.35	1.69	0.62
本实施例	0.90	**0.36**	**0.62**	0.38

[0127] 本实施例中,在表2所比较的方法中,ICP、GICP、LOAM是传统的非学习方法,在这些传统方法中,具有后端优化的LOAM取得了最好的结果;LO-Net、ENCODE、DMLO、PWCLO-Net等都是基于学习的方法。据我们所知,在基于深度学习的方法中,PWCLO-Net是此前所有基于深度学习的方法中精度最高的,本实施例的方法与之相比,特征金字塔和位姿优化的层数更浅,并且由于更好地利用了点云的几何、语义信息,本实施例所述的方法在基于深度学习的方法中取得了最好的性能。

[0128] 为了验证本实施例所述的方法各部分的意义,本实施例中还进行了消融实验。实验结果如表3所示,其中,第二行中的“w/o Ground segmentation”表示去除数据的地面分割预处理,此时位姿估计网络的输入为完整的激光雷达点云帧。第三行的“with Shared MLP”表示将孪生特征金字塔中的MBConv3D模块全部替换为Shared MLP层。第四行的“w/o

Pose refinement”表示去除网络中的位姿分层优化网络,此时场景流融合编码网络输出的6自由度位姿向量直接作为最终的输出。第五行的“w/o Mask”表示去除网络中所有输出为掩码的模块。最后一行表示本文完整的方法的实验结果。

[0129] 表3消融实验结果

[0130]

方法	09 Rel. trans.(%)	09 Rel. rot.(deg/m)	10 Rel. trans.(%)	10 Rel. rot.(deg/m)
w/o Ground segmentation	1.31	0.43	1.11	0.51
with Shared MLP	2.83	0.90	2.06	0.98
w/o Pose refinement	5.77	2.08	8.13	4.37
w/o Mask	2.12	0.68	2.13	1.16
本实施例	**0.90**	**0.36**	**0.62**	**0.38**

[0131] 综上,本发明实施例所述的基于孪生特征金字塔和地面分割的激光雷达里程计方法,至少具有以下优点:

[0132] 1)本发明是一种基于孪生特征金字塔和分层优化的激光雷达里程计方法,通过深度神经网络仅使用激光雷达的点云信息,不使用任何其他信息来估计两帧间的位姿变换;

[0133] 2)针对道路环境场景中地面点对里程计贡献度低、信息冗余的问题,本实施例提出了使用预处理筛除原始点云中的地面点,使得用于位姿估计的点云特征密度更高,提高了位姿估计网络的收敛速度和泛化能力;

[0134] 3)本实施例设计了一个全新的逐点特征提取模块MBConv3D,旨在通过建立局部点的分布特征来捕获物体表面的变形、加强局部点邻域的特征聚合,并基于MBConv模块的结构优化计算效率;

[0135] 4)本实施例在KITTI数据集上进行评估实验和消融实验验证本实施例所提出的方法。实验结果表明,本实施例的方法在大部分序列中与最优的激光雷达里程计效果相当,在测试序列中甚至比基于深度学习的所有方法精度更高,因此,本实施例提供的基于孪生特征金字塔和地面分割的激光雷达里程计方法,能够提高基于激光雷达进行位姿估计任务的精度。

[0136] 图7是本发明实施例提供的一种电子设备600的结构示意图,该电子设备600可因配置或性能不同而产生比较大的差异,可以包括一个或一个以上处理器(central processing units,CPU)601和一个或一个以上的存储器602,其中,所述存储器602中存储有至少一条指令,所述至少一条指令由所述处理器601加载并执行以实现上述基于孪生特征金字塔和地面分割的激光雷达里程计方法。

[0137] 在示例性实施例中,还提供了一种计算机可读存储介质,例如包括指令的存储器,上述指令可由终端中的处理器执行以完成上述基于孪生特征金字塔和地面分割的激光雷达里程计方法。例如,所述计算机可读存储介质可以是ROM、随机存取存储器(RAM)、CD-ROM、磁带、软盘和光数据存储设备等。

[0138] 本领域普通技术人员可以理解实现上述实施例的全部或部分步骤可以通过硬件来完成,也可以通过程序来指令相关的硬件完成,所述的程序可以存储于一种计算机可读存储介质中,上述提到的存储介质可以是只读存储器,磁盘或光盘等。

[0139] 以上所述仅为本发明的较佳实施例,并不用以限制本发明,凡在本发明的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本发明的保护范围之内。

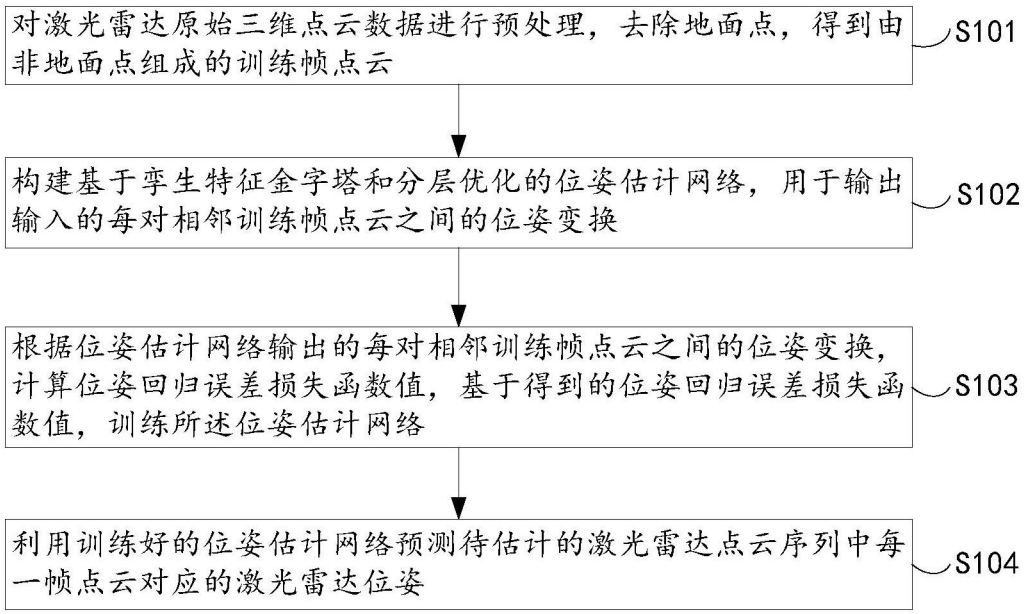

对激光雷达原始三维点云数据进行预处理，去除地面点，得到由非地面点组成的训练帧点云
S101
构建基于孪生特征金字塔和分层优化的位姿估计网络，用于输出输入的每对相邻训练帧点云之间的位姿变换
S102
根据位姿估计网络输出的每对相邻训练帧点云之间的位姿变换，计算位姿回归误差损失函数值，基于得到的位姿回归误差损失函数值，训练所述位姿估计网络
S103
利用训练好的位姿估计网络预测待估计的激光雷达点云序列中每一帧点云对应的激光雷达位姿
S104

图1

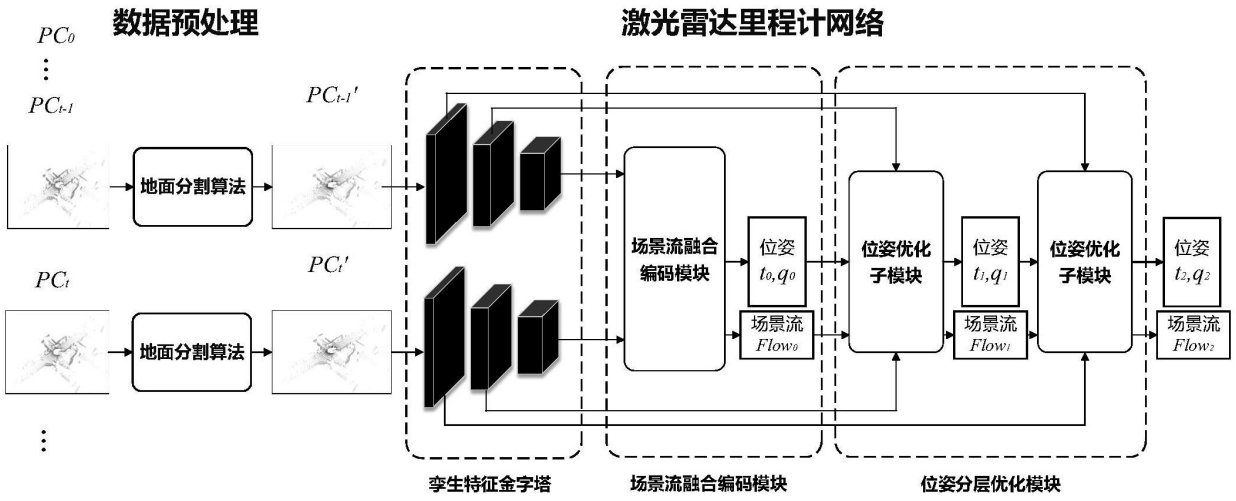

数据预处理
激光雷达里程计网络
PC0
PCt-1
PCt-1'
地面分割算法
PCt
PCt'
地面分割算法
场景流融合编码模块
位姿 t0,q0
场景流 Flow0
位姿优化子模块
位姿 t1,q1
场景流 Flow1
位姿优化子模块
位姿 t2,q2
场景流 Flow2
孪生特征金字塔
场景流融合编码模块
位姿分层优化模块

图2

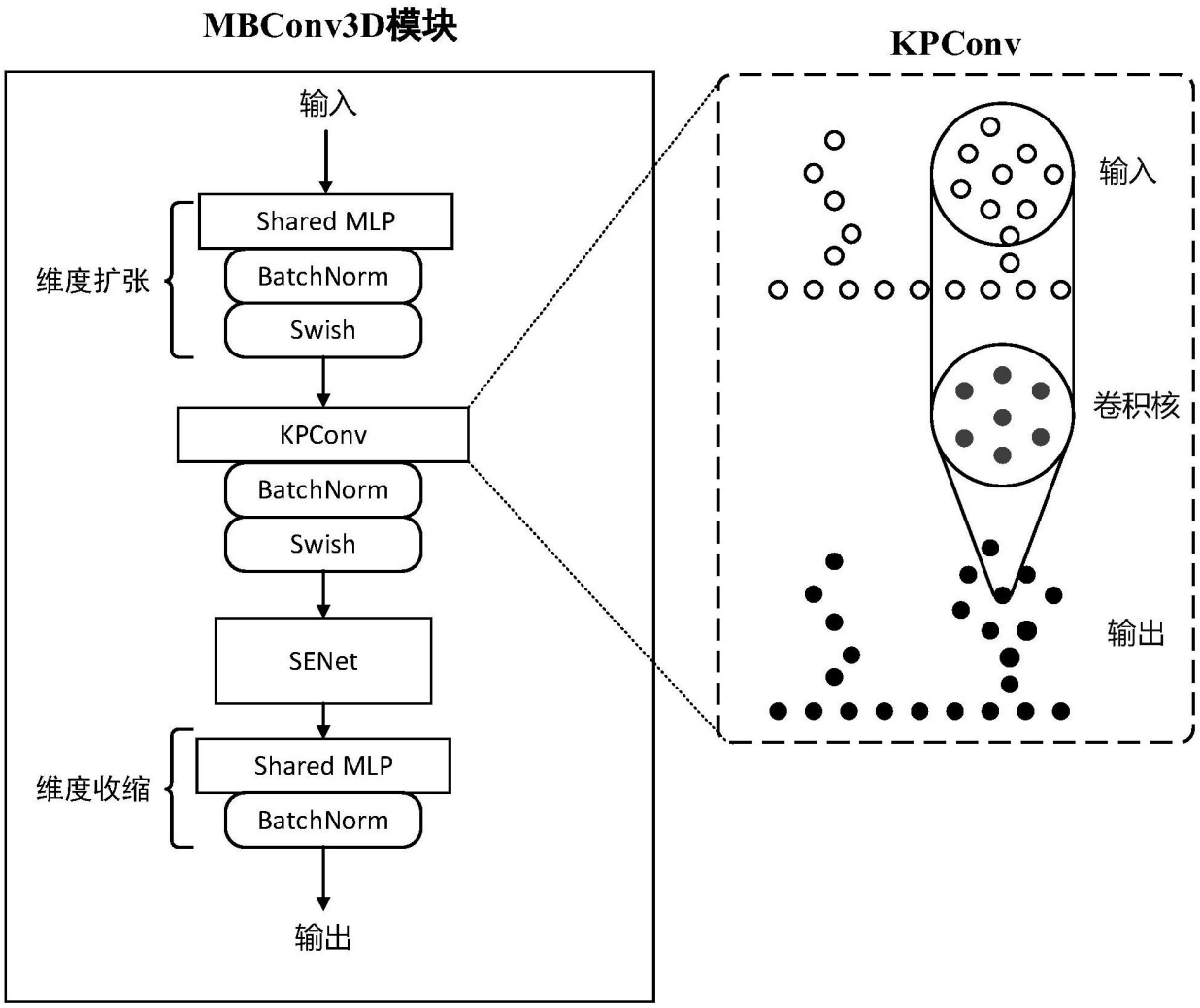

MBConv3D模块
输入
Shared MLP
BatchNorm
Swish
维度扩张
KPConv
BatchNorm
Swish
SENet
Shared MLP
BatchNorm
维度收缩
输出
KPConv
输入
卷积核
输出

图3

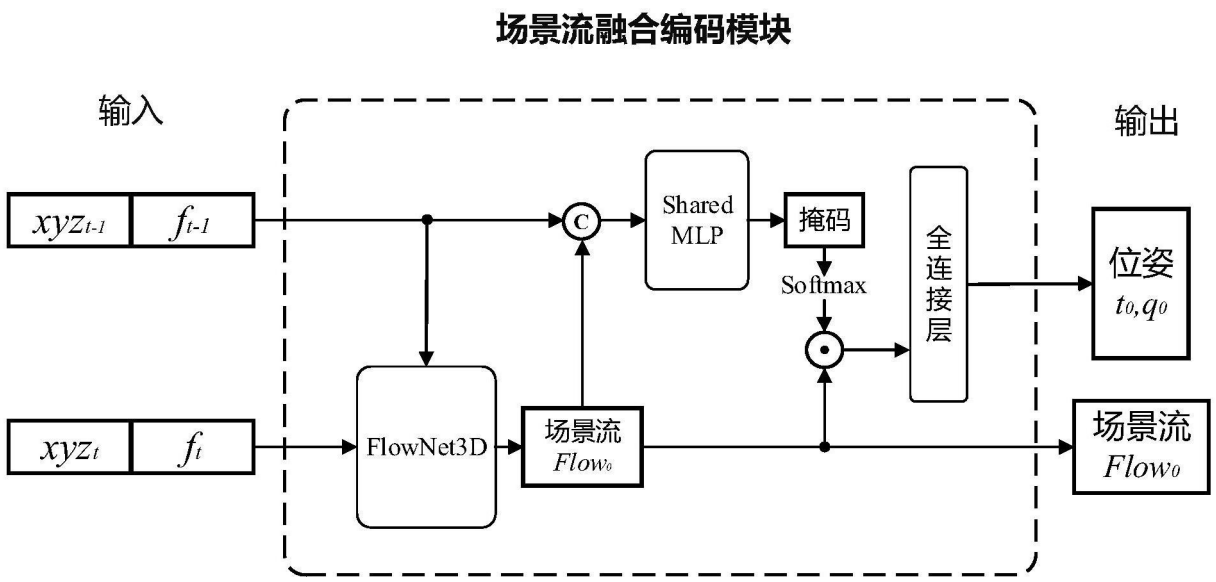

场景流融合编码模块
输入
输出
xyz_{t-1}
f_{t-1}
xyz_t
f_t
C
Shared MLP
掩码
Softmax
全连接层
FlowNet3D
场景流 $Flow_0$
位姿 t_0, q_0
场景流 $Flow_0$

图4

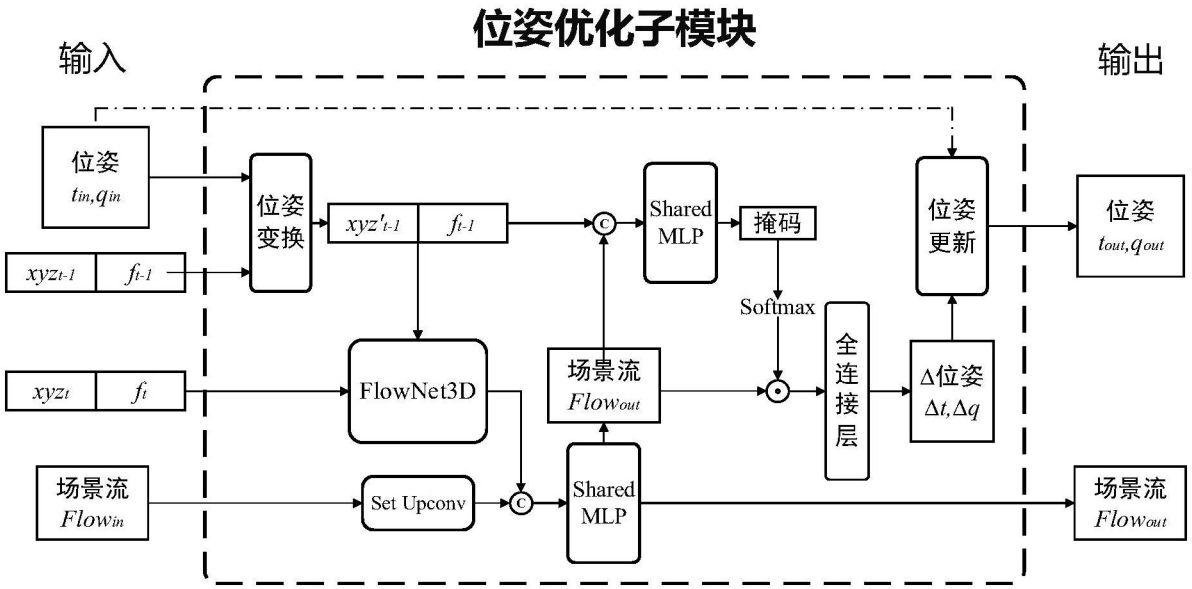

位姿优化子模块
输入
输出
位姿 t_{in}, q_{in}
xyz_{t-1}
f_{t-1}
位姿变换
xyz'_{t-1}
f_{t-1}
Shared MLP
掩码
Softmax
全连接层
位姿更新
位姿 t_{out}, q_{out}
xyz_t
f_t
FlowNet3D
场景流 $Flow_{out}$
Δ位姿 $\Delta t, \Delta q$
场景流 $Flow_{in}$
Set Upconv
Shared MLP
场景流 $Flow_{out}$

图5

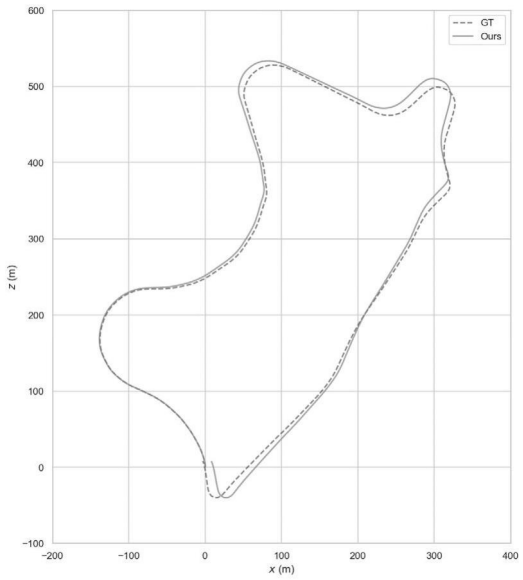

GT
Ours
600
500
400
300
200
100
0
−100
z (m)
−200
−100
0
100
200
300
400
x (m)

图6(a)

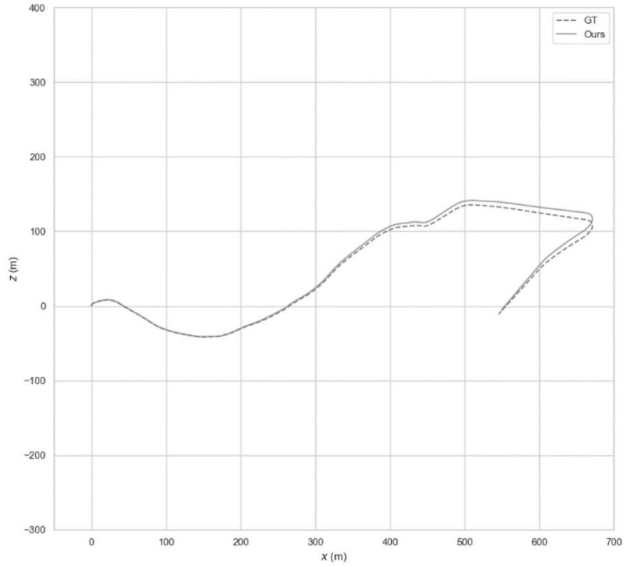

GT
Ours
400
300
200
100
0
−100
−200
−300
z (m)
0
100
200
300
400
500
600
700
x (m)

图6(b)

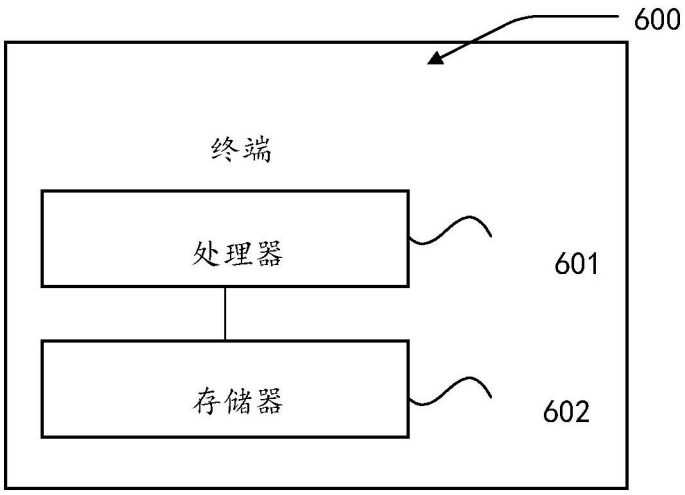


图7