



(12) 发明专利

(10) 授权公告号 CN 112396655 B

(45) 授权公告日 2023. 01. 03

(21) 申请号 202011290504.3

(22) 申请日 2020.11.18

(65) 同一申请的已公布的文献号  
申请公布号 CN 112396655 A

(43) 申请公布日 2021.02.23

(73) 专利权人 哈尔滨工程大学  
地址 150001 黑龙江省哈尔滨市南岗区南  
通大街145号哈尔滨工程大学科技处  
知识产权办公室

(72) 发明人 苏丽 宋浩

(51) Int. Cl.  
G06T 7/73 (2017.01)  
G06V 10/774 (2022.01)  
G06V 10/82 (2022.01)  
G06N 3/08 (2006.01)

(56) 对比文件

- CN 110930454 A, 2020.03.27
- CN 109086683 A, 2018.12.25
- CN 110930452 A, 2020.03.27
- CN 111368733 A, 2020.07.03
- CN 111126269 A, 2020.05.08
- CN 111915677 A, 2020.11.10
- CN 110533721 A, 2019.12.03
- CN 111862201 A, 2020.10.30
- CN 109064514 A, 2018.12.21
- CN 110473284 A, 2019.11.19
- US 2019304134 A1, 2019.10.03
- US 2020311956 A1, 2020.10.01

审查员 李小敏

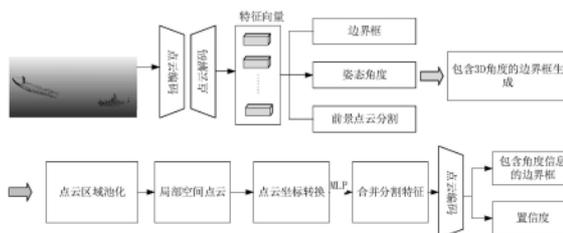
权利要求书1页 说明书6页 附图1页

(54) 发明名称

一种基于点云数据的船舶目标6D位姿估计方法

(57) 摘要

本发明公开了一种基于点云数据的船舶目标6D位姿估计方法,包括:步骤1:获得海上场景的船舶点云数据集,数据集标签包括目标类别、目标的三维坐标、目标的三维尺寸以及目标的三维位姿;步骤2:构建神经网络,采用PointNet++进行逐点的点云特征提取,得到逐点的高维特征;步骤3:以自下而上的方案生成3D边界框提案,基于3D边界框生成真实分割掩模,分割前景点并同时从分割点生成带有角度信息的边界框提案,用于RCNN的输入;步骤4:基于步骤3得到的提案以及前景分割特征和空间特征进行提案的优化以输出最终的分类及3D边框和姿态角度。该方法实现了采用端到端学习的方式进行三维目标的位姿估计效果,提升了位姿估计的实时性。



1. 一种基于点云数据的船舶目标6D位姿估计方法,其特征在于,包括以下步骤:

步骤1:获得海上场景的船舶点云数据集,数据集标签包括目标类别、目标的三维坐标、目标的三维尺寸以及目标的三维位姿;

步骤2:构建神经网络,采用PointNet++进行逐点的点云特征提取,得到逐点的高维特征;

步骤3:以自下而上的方案生成3D边界框提案,基于3D边界框生成真实分割掩模,分割前景点并同时从分割点生成带有角度信息的边界框提案,用于RCNN的输入;所述基于3D边界框生成真实分割掩模,分割前景点并同时从分割点生成带有角度信息的边界框提案具体为:对每个点进行二分类,前景和背景得分,再经过一个sigmoid函数归一化到0~1,认为得分高于阈值的为前景点,分割前景点并同时从分割点生成带有角度信息的边界框提案,以前景点为中心,总数为N个,在每个点上利用回归值以及设置的平均尺寸生成初始提案,大小为(batch\_size\*N,9),9维向量分别为 $[x, y, z, h, w, l, rx, ry, rz]$ ,即目标的中心位置、高宽长以及分别在xyz方向的旋转角度,再根据分类的得分进行排序,利用非极大值找出每个batch前512个边界框,返回边界框和角度信息以及置信度得分,分割阶段采用交叉熵损失函数作为二分类网络的损失函数,对初始提案的预测,其中x、z和偏航角ry基于bin的损失函数回归计算,尺寸信息(h,w,l)和旋转角(rx,rz)采用平滑L1损失进行计算;

步骤4:基于步骤3得到的提案以及前景分割特征和空间特征进行提案的优化以输出最终的分类及3D边框和姿态角度。

2. 根据权利要求1所述的一种基于点云数据的船舶目标6D位姿估计方法,其特征在于:步骤2所述采用PointNet++进行逐点的点云特征提取,得到逐点的高维特征具体为:点集的特征提取由三部分组成,包括Sampling layer、Grouping layer、Pointnet layer,采样层的采样算法使用迭代最远点采样方法,在输入点云中选择一个系列点,由此定义出局部区域的中心,然后构建局部邻域,在给定的距离内寻找点,然后用全连接层进行特征提取,最后做池化操作得到高级特征,将点集上采样到原始点集数量,得到输入点集中逐点的高维特征。

## 一种基于点云数据的船舶目标6D位姿估计方法

### 技术领域

[0001] 本发明涉及一种基于点云数据的船舶目标6D位姿估计方法,本发明涉及点云领域、海上场景的船舶目标位姿估计、深度学习和神经网络领域。

### 背景技术

[0002] 位姿估计在计算机视觉领域扮演着十分重要的角色。在使用视觉传感器估计机器人位姿进行控制、机器人导航、增强现实以及其它方面都有着极大的应用。目标的位姿估计问题是确定物体在3D空间中的空间位置,以及物体绕坐标轴旋转的角度,其中绕Z轴转动为偏航角(Yaw),绕Y轴转动为俯仰角(Pitch),绕X轴转动为横滚角(Roll)。近年来,6D位姿估计的方法可以分为四大类,分别为基于模型对应点的方法、基于模板匹配的方法、基于投票的方法以及基于回归的方法。基于模型对应点的方法主要针对的是纹理丰富的物体,基于模板匹配的方法主要针对弱纹理或无纹理的图像。针对船舶目标的海上场景的特殊性,例如海上水面光照的影响,以及天气的多变影响图像质量。传统的三维姿态估计算法一般只能对单个目标进行检测且耗时较大。

### 发明内容

[0003] 针对上述现有技术,本发明要解决的技术问题是提供一种基于点云数据的船舶目标6D位姿估计方法,通过学习点云信息的特征和6D位姿之间的联系,通过在基于深度学习进行目标检测的基础上进行改进,再进一步回归出目标的6D位姿。

[0004] 为解决上述技术问题,本发明的一种基于点云数据的船舶目标6D位姿估计方法,包括以下步骤:

[0005] 步骤1:获得海上场景的船舶点云数据集,数据集标签包括目标类别、目标的三维坐标、目标的三维尺寸以及目标的三维位姿;

[0006] 步骤2:构建神经网络,采用PointNet++进行逐点的点云特征提取,得到逐点的高维特征;

[0007] 步骤3:以自下而上的方案生成3D边界框提案,基于3D边界框生成真实分割掩模,分割前景点并同时从分割点生成带有角度信息的边界框提案,用于RCNN的输入;

[0008] 步骤4:基于步骤3得到的提案以及前景分割特征和空间特征进行提案的优化以输出最终的分类及3D边框和姿态角度。

[0009] 本发明还包括:

[0010] 1.步骤2中采用PointNet++进行逐点的点云特征提取,得到逐点的高维特征具体为:点集的特征提取由三部分组成,包括Sampling layer、Grouping layer、Pointnet layer,采样层的采样算法使用迭代最远点采样方法,在输入点云中选择一个系列点,由此定义出局部区域的中心,然后构建局部邻域,在给定的距离内寻找点,然后用全连接层进行特征提取,最后做池化操作得到高级特征,将点集上采样到原始点集数量,得到输入点集中逐点的高维特征。

[0011] 2.步骤3中基于3D边界框生成真实分割掩模,分割前景点并同时从分割点生成带有角度信息的边界框提案具体为:对每个点进行二分类,前景和背景得分,再经过一个sigmoid函数归一化到0~1,认为得分高于阈值的为前景点,分割前景点并同时从分割点生成带有角度信息的边界框提案,以前景点为中心,总数为N个,在每个点上利用回归值以及设置的平均尺寸生成初始提案,大小为(batch\_size\*N,9),9维向量分别为[x,y,z,h,w,l,rx,ry,rz],即目标的中心位置、长宽高以及分别在xyz方向的旋转角度,再根据分类的得分进行排序,利用非极大值找出每个batch前512个边界框,返回边界框和角度信息以及置信度得分,分割阶段采用交叉熵损失函数作为二分类网络的损失函数,对初始提案的预测,其中x、z和偏航角ry基于bin的损失函数回归计算,尺寸信息(h,w,l)和旋转角(rx,rz)采用平滑L1损失进行计算。

[0012] 本发明的有益效果:船舶感知环节是智能船舶发展的路径之一,对船舶目标的位置和姿态识别是船舶感知或对船舶跟踪识别不可缺少的环节。在目标检测的基础上加上对船舶目标的姿态估计能够增强感知精度。传统的三维姿态估计算法一般只能对单个目标进行检测且耗时较大,所以本算法采用基于深度学习神经网络,采用学习的方式对先验数据集进行训练,大大提升了算法的实时性,并可以进行复杂场景的位姿估计。该方法实现了采用端到端学习的方式进行三维目标的位姿估计效果,提升了位姿估计的实时性。

## 附图说明

[0013] 图1是网络中生成器的结构示意图;

[0014] 图2为算法位姿估计结果的可视化示意图。

## 具体实施方式

[0015] 下面结合附图对本发明具体实施方式做进一步说明。

[0016] 本发明的目的是这样实现的:

[0017] 1、获取数据集

[0018] 通过人工制作获取点云数据集,数据集要求海上场景的船舶点云数据,可以是包含单船、多船、遮挡等情况下的船舶目标。同时每幅场景的点云数据应包含相应的船舶位姿标签,标签的结构包括类别、目标的三维坐标、目标的三维尺寸以及目标的三维位姿即船舶目标相应的偏航角、横滚角以及俯仰角三个角度。

[0019] 2、构建神经网络

[0020] 该算法基于3D目标检测网络PointRCNN,3D目标检测算法目标仅仅是对3D边框的输出,在此基础上增加网络角度分支以估计目标的6D位姿。网络主要包含两部分:基于PointNet++的点云特征提取和基于3D目标检测网络PointRCNN的改进。其中PointNet++网络作为主干网络,负责从输入点云数据进行特征提取。

[0021] 基于PointRCNN的改进的6D位姿估计网络包括两个阶段,第一阶段旨在以自下而上的方案生成3D边界框提案并在此基础上同时得出目标的三维姿态角度,基于边界框生成真实分割掩模,分割前景点并同时从分割点生成少量边界框提案。这样的策略避免了在整个三维空间中使用大量锚框。第二阶段进行规范的带角度的边框优化。在生成边框及角度提案之后,采用点云区域池化操作进行池化来自第一阶段学到的点表示。与直接估计全局

边框坐标的方法不同,池化后的点被转换为规范坐标并与点特征以及第一阶段的分割掩模组合完成坐标优化。该策略充分利用了第一阶段的分割和提案子网络提供的信息。

[0022] 3、6D位姿估计网络主要步骤

[0023] (1) 3D边框及角度提案的生成

[0024] 主干网络采用PointNet++进行逐点的点云特征提取。输入为N\*3的点云数据,对于每一个N×3的点云输入,点集的特征提取由三部分组成,分别为Sampling layer、Grouping layer、Pointnet layer。采样层的采样算法使用迭代最远点采样方法iterative farthest point sampling (FPS)。在输入点云中选择一个系列点,由此定义出局部区域的中心。之后进行构建局部邻域,在给定的距离内寻找点,然后用全连接层进行特征提取,最后做池化操作得到高级特征,将点集上采样回原始点集数量,得到输入点集中逐点的高维特征。

[0025] 3D场景中的目标是自然分离的,彼此不重叠。所有三维目标的分割掩模都可以通过3D边界框注释直接获得,即3D框内的3D点被视为前景点。经过主干网络处理后的逐点的特征,前景预测分支即对每个点进行预测,点云的真实分割掩模由三维边框确定。

[0026] 在前景预测分支的同时进行3D边框以及三维角度的预测分支,采用全连接层。由于船舶目标的包围边框坐标以及偏航角的尺度变化较大,但横滚角与俯仰角的变化尺度通常较小,对于包围边框的水平方向坐标x,z以及偏航角采用基于Bin的思想回归计算。具体来说将每个前景点的周围区域沿着X轴和Z轴分割成一系列离散的箱子,在这两个方向上采用基于Bin的交叉熵损失分类并进行残差回归而不是直接进行平滑L1损失回归。对于大多数海上场景中,船舶目标的水平方向尺度变化较大,但在垂直方向上的中心坐标尺度变化通常较小,且姿态角度俯仰角与横滚角都在一个非常小的尺度范围内变化,算法采用直接用平滑L1损失进行回归即可以得到船舶的横滚角、俯仰角度和船舶高度位置的准确值。

[0027] (2) 点云区域池化

[0028] 在获得三维边界框方案后,根据之前生成的边框以及角度分支来优化框的位置和方向。根据每个三维边框的位置,对每个点及其特征进行池化。稍稍放大后的边界框内的点及其特征将被保留。然后使用分割掩模区分稍微放大框内的前景、背景点。没有内部点的提案将会被消除。

[0029] (3) 规范3D边界框及角度回归提案优化

[0030] 池化后的点集及其相关特征都被送入第二阶段子网络,以优化边框的位置、角度和前景目标的可信度。将上述特征进行组合,经过3个PointNet++提出的Sampling layer得到高维特征,然后使用分类和回归分支进行预测输出目标的坐标尺寸以及姿态角度信息。

[0031] 损失函数如下:

$$[0032] \quad bin_x^{(p)} = \frac{x^p - x^{(p)} + s}{\delta}, \quad bin_z^{(p)} = \frac{z^p - z^{(p)} + s}{\delta}$$

$$[0033] \quad res_u^{(p)} = \frac{1}{\delta} u^p - u^{(p)} + s - (bin_u^{(p)} \cdot \delta + \frac{\delta}{2}), \quad res_y^{(p)} = y^p - y^{(p)}$$

$$[0034] \quad L = \frac{1}{||B||} \sum_{i \in B} F_{cls}(prob_i, lable_i) + \frac{1}{||B_{pos}||} \sum_{i \in B_{pos}} L_{bin}^{(p)} + L_{res} F_{reg}$$

[0035] 其中基于bin的损失函数 $L_{bin}^{(p)}$ :

$$[0036] \quad L_{bin}^{(p)} = \sum_{u \in \{x, z, \theta\}} (F_{cls}(\widehat{bin}_u^{(p)}, bin_u^{(p)}) + F_{reg}(\widehat{res}_u^{(p)}, res_u^{(p)}))$$

[0037] 残差损失函数 $L_{res}^{(p)}$ 为:

$$[0038] \quad L_{res}^{(p)} = \sum_{v \in \{y, h, w, l, \varphi, \psi\}} F_{reg}(\widehat{res}_v^{(p)}, res_v^{(p)})$$

[0039] 其中

$$[0040] \quad res_u^{(p)} = \frac{1}{\delta} u^p - u^{(p)} + s - (bin_u^{(p)} \cdot \delta + \frac{\delta}{2}), \quad res_y^{(p)} = y^p - y^{(p)}$$

$$[0041] \quad bin_x^{(p)} = \frac{x^p - x^{(p)} + s}{\delta}, \quad bin_z^{(p)} = \frac{z^p - z^{(p)} + s}{\delta}$$

[0042] B为第一阶段的提案集合,  $B_{pos}$ 为回归的正提案,  $prob_i$ 为置信度,  $lable_i$ 为相应的标签,  $F_{cls}$ 为交叉熵损失函数用于计算预测的置信度。

[0043] 位姿估计(Pose estimation)在计算机视觉领域扮演着十分重要的角色。在使用视觉传感器估计机器人位姿进行控制、机器人导航、增强现实以及其它方面都有着极大的应用。目标的位姿估计问题是确定物体在3D空间中的空间位置,以及物体绕坐标轴旋转的角度,其中绕Z轴转动为偏航角(Yaw),绕Y轴转动为俯仰角(Pitch),绕X轴转动为横滚角(Roll)。

[0044] 结合图1和图2,本发明步骤如下:

[0045] 步骤1、准备船舶场景的位姿数据集,数据集标签应包括目标类别、目标的三维坐标、目标的三维尺寸以及目标的三维位姿即船舶目标相应的偏航角、横滚角以及俯仰角三个角度。

[0046] 步骤2、基于PointNet++的点云特征提取,特征提取部分主要由Set Abstraction子网络构成点集的特征提取由三部分组成,对输入的点云进行排列,基于采样算法迭代最远点采样方法iterative farthest point sampling(FPS)寻找最远点,在输入点云中选择这些一系列最远点后,由此定义出局部区域的中心。之后进行构建局部邻域,在给定的距离内寻找点,然后用全连接层进行特征提取,最后做池化操作得到高级特征,将点集上采样到原始点集数量,得到输入点集中逐点的高维特征将输入点云数据经过主干网络得到逐点的高维特征。

[0047] 步骤3、RPN阶段,旨在以自下而上的方案生成3D边界框提案,基于3D边界框生成真实分割掩模,即对每个点进行二分类,前景和背景得分,再经过一个sigmoid函数归一化到0~1,认为得分高于阈值的为前景点,分割前景点并同时从分割点生成少量带有角度信息的边界框提案,以前景点(总数为N个)为中心,在每个点上,利用回归值以及设置的平均尺寸,生成初始提案,大小为(batch\_size\*N,9),9维向量分别为[x,y,z,h,w,l,rx,ry,rz],即目标的中心位置和长宽高以及分别在xyz方向的旋转角度。再根据分类的得分进行排序,利用非极大值找出每个batch前512个边界框,返回边界框和角度信息,以及置信度得分用于RCNN阶段的输入,分割阶段采用交叉熵损失函数作为二分类网络的损失函数,对初始提案的预测,其中x、z和偏航角ry基于bin的损失函数回归计算,尺寸信息(h,w,l)和旋转角(rx,rz)采用平滑L1损失进行计算。

[0048] 步骤4、RCNN阶段,旨在基于RPN阶段得到的少量提案以及前景分割特征和空间特征进行提案的优化以输出最终的分类及3D边框和姿态角度。根据RPN阶段的提案,计算感兴趣区域ROI与真值之间的IOU,IOU大于0.55是将真值分给预测的提案进行微调。将分割的掩码与点云坐标和高维特征进行连接重组,将合并的特征向量由PointNet++的Set Abstraction子网络得到高级特征,由分类层和回归层网络进行预测,其中分类层为交叉熵损失,回归层为基于bin的损失以及平滑L1损失。

[0049] 实施例:

[0050] 1、制作含有目标类别、目标的三维坐标、目标的三维尺寸以及目标的三位姿即船舶目标相应的偏航角、横滚角以及俯仰角三个角度的数据集以供训练。

[0051] 2、构建神经网络。

[0052] 主干网络采用PointNet++进行逐点的点云特征提取。点集的特征提取由三部分组成,分别为Sampling layer、Grouping layer、Pointnet layer。采样层的采样算法使用迭代最远点采样方法iterative farthest point sampling (FPS)。在输入点云中选择一个系列点,由此定义出局部区域的中心。之后进行构建局部邻域,在给定的距离内寻找点,然后用全连接层进行特征提取,最后做池化操作得到高级特征,将点集上采样到原始点集数量,得到输入点集中逐点的高维特征,输入有 $n*3$ 到 $n*128$ 的转换。

[0053] 3、6D位姿估计网络

[0054] 整体网络结构如图1所示。经过主干网络处理后的逐点的特征,高维特征经过三个后续结构即前景预测分支、边框回归以及角度回归分支。前景预测分支即对每个点进行预测,点云的真实分割掩码由三维边框确定。在前景预测分支的同时进行3D边框以及三维角度的预测分支,采用全连接层。由于船舶目标的包围边框坐标以及偏航角的尺度变化较大,但横滚角与俯仰角的变化尺度通常较小,对于包围边框的水平方向坐标 $x, z$ 以及偏航角采用基于Bin的思想回归计算。具体来说将每个前景点的周围区域沿着X轴和Z轴分割成一系列离散的箱子,在这两个方向上采用基于Bin的交叉熵损失分类并进行残差回归而不是直接进行平滑L1损失回归。对于大多数海上场景中,船舶目标的水平方向尺度变化较大,但在垂直方向上的中心坐标尺度变化通常较小,且姿态角度俯仰角与横滚角都在一个非常小的尺度范围内变化,算法采用直接用平滑L1损失进行回归即可以得到准确值。

[0055] 在获得三维边界框方案后,根据之前生成的边框以及角度分支来优化框的位置和方向。根据每个三维边框的位置,对每个点及其特征进行池化。稍稍放大后的边界框内的点及其特征将被保留。然后使用分割掩模区分稍微放大框内的前景、背景点。没有内部点的提案将会被消除。

[0056] 池化后的点集及其相关特征都被送入第二阶段子网络,以优化边框的位置、角度和前景目标的可信度。将上述特征进行组合,经过3个PointNet++提出的Sampling layer得到高级特征,然后使用分类和回归分支进行预测输出目标的坐标尺寸以及姿态角度信息,对输出信息的可视化结果如图2所示。

[0057] 损失函数如下:

$$[0058] \quad bin_x^{(p)} = \frac{x^p - x^{(p)} + s}{\delta}, \quad bin_z^{(p)} = \frac{z^p - z^{(p)} + s}{\delta}$$

$$[0059] \quad res_u^{(p)} = \frac{1}{\delta} u^p - u^{(p)} + s - (bin_u^{(p)} \cdot \delta + \frac{\delta}{2}), \quad res_y^{(p)} = y^p - y^{(p)}$$

$$[0060] \quad L = \frac{1}{\|B\|} \sum_{i \in B} F_{cls}(prob_i, lable_i) + \frac{1}{\|B_{pos}\|} \sum_{i \in B_{pos}} L_{bin}^{(p)} + L_{res}^{(p)} F_{reg}$$

[0061] 其中基于bin的损失函数 $L_{bin}^{(p)}$ :

$$[0062] \quad L_{bin}^{(p)} = \sum_{u \in \{x, z, \theta\}} (F_{cls}(\widehat{bin}_u^{(p)}, bin_u^{(p)}) + F_{reg}(\widehat{res}_u^{(p)}, res_u^{(p)}))$$

[0063] 残差损失函数 $L_{res}^{(p)}$ 为:

$$[0064] \quad L_{res}^{(p)} = \sum_{v \in \{y, h, w, l, \phi, \psi\}} F_{reg}(\widehat{res}_v^{(p)}, res_v^{(p)})$$

[0065] 其中

$$[0066] \quad res_u^{(p)} = \frac{1}{\delta} u^p - u^{(p)} + s - (bin_u^{(p)} \cdot \delta + \frac{\delta}{2}), \quad res_y^{(p)} = y^p - y^{(p)}$$

$$[0067] \quad bin_x^{(p)} = \frac{x^p - x^{(p)} + s}{\delta}, \quad bin_z^{(p)} = \frac{z^p - z^{(p)} + s}{\delta}$$

[0068] B为第一阶段的提案集合,  $B_{pos}$ 为回归的正提案,  $prob_i$ 为置信度,  $lable_i$ 为相应的标签,  $F_{cls}$ 为交叉熵损失函数用于计算预测的置信度。

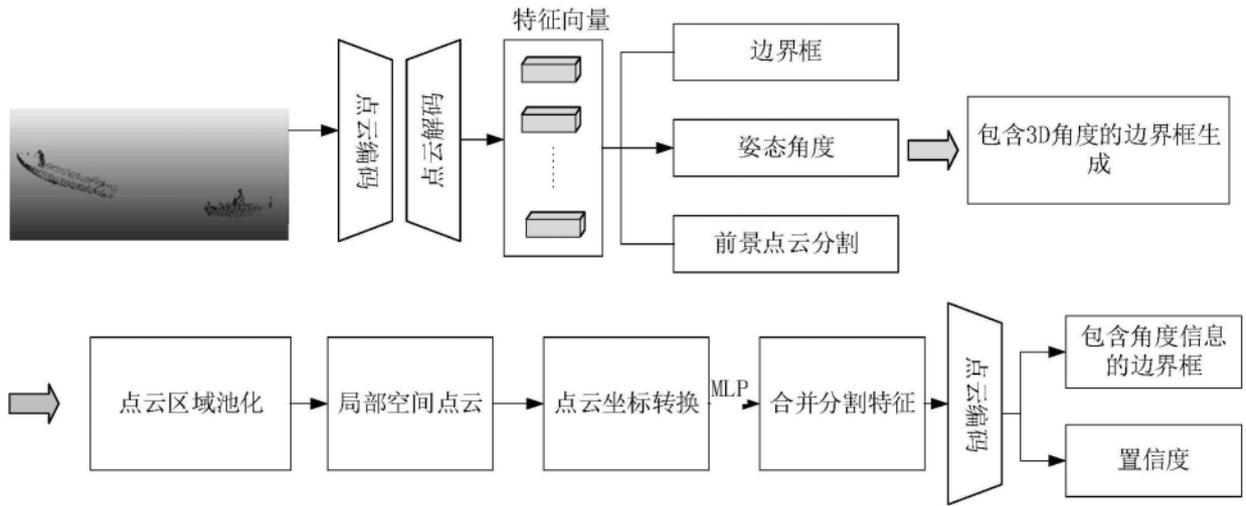


图1

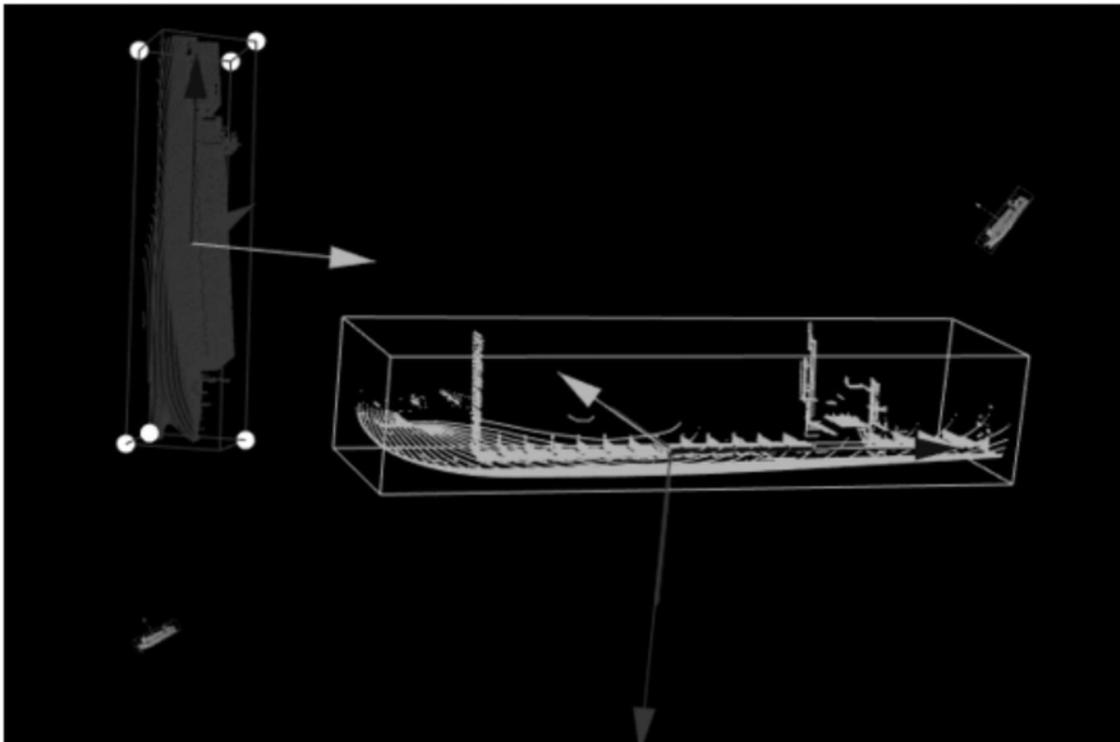


图2