基于异构数据特征融合的非规则回转体位姿估计算法

程爱玲

浙江博城机器人科技有限公司，浙江 杭州 037001

DOI: 10.61369/TACS.2026080043

摘　　要： 非规则回转体是精密制造、工业装配、机器人抓取领域核心零部件，高精度位姿估计是其自动化作业的关键。针对现有算法在低纹理、局部遮挡工况下位姿估计精度差，三维点云运算效率偏低，传统异构融合算法特征利用率不足等工程问题，本文提出 RGB 图像与三维点云异构数据融合的三阶段非规则回转体位姿估计算法。算法分为目标分割、初始位姿估计、位姿求精三大模块：分割阶段嵌入轻量化注意力机制优化实例分割网络，提纯目标数据、优化边缘分割效果；初始位姿估计阶段依托回转体固有几何特征，设计高斯曲率自适应关键点分层采样策略，削减点云冗余数据，融合图像纹理、点云几何两类特征完成初始位姿求解；位姿求精阶段迭代修正二维 – 三维匹配误差，优化位姿参数。工业工况实验表明，该算法适配低纹理、轻微遮挡复杂场景，相较于单模态与常规融合算法，位姿估计精度、鲁棒性与运算效率优势明显，能够满足工业精密自动化作业高精度使用要求。

关 键 词： 非规则回转体；位姿估计；异构数据；特征融合；高斯曲率；注意力机制

Non-Regular Rotary Body Pose Estimation Algorithm Based on Heterogeneous Data Feature Fusion

Cheng Ailing

Zhejiang Bocheng Robotics Technology Co., Ltd., Hangzhou, Zhejiang 037001

Abstract: Irregular rotation bodies are core components in precision manufacturing,industrial assembly,and robotic grasping,with high-precision pose estimation being the key to their automated operations.To address engineering challenges such as poor pose estimation accuracy in low-texture and partial occlusion scenarios,low computational efficiency of 3D point clouds, and insufficient feature utilization in traditional heterogeneous fusion algorithms,this paper proposes a three-stage pose estimation algorithm for irregular rotation bodies through RGB image and 3D point cloud heterogeneous data fusion.The algorithm is divided into three major modules: target segmentation,initial pose estimation, and pose refinement.The segmentation phase incorporates a lightweight attention mechanism to optimize instance segmentation networks,purify target data, and enhance edge segmentation effects. The initial pose estimation phase leverages the inherent geometric features of rotation bodies,designing a Gaussian curvature-adaptive keypoint hierarchical sampling strategy to reduce redundant point cloud data and fuse image texture and point cloud geometric features for initial pose solution.The pose refinement phase iteratively corrects 2D-3D matching errors and optimizes pose parameters.Industrial experiments demonstrate that this algorithm adapts to complex scenes with low texture and minor occlusion,showing significant advantages in pose estimation accuracy,robustness,and computational efficiency compared to single-modal and conventional fusion algorithms,meeting the high-precision requirements of industrial automated operations.

Keywords: irregularly shaped rotation bodies; pose estimation; heterogeneous data; feature fusion; Gaussian curvature; attention mechanism

引言

智能制造发展使视觉位姿估计广泛用于装配、抓取等场景，该技术求解目标空间位姿以支撑自动化作业。非规则回转体纹理稀疏、曲面复杂，叠加现场光照、遮挡干扰，高精度位姿估计存在较大技术难点。现有方法各有短板：RGB 图像法缺少深度信息，低纹理构件特征提取失效；点云法数据冗余，实时性差；多模态融合算法可互补两类数据，但经典算法计算量大，现有改进未利用回转体几何特征，曲面有效特征挖掘不足。面向工业精密作业场景，本文提出三阶段异构融合位姿估计算法，融合注意力分割、曲率分层采样与迭代求精，实现非规则回转体高效、鲁棒、高精度位姿估计。

Copyright © This Work is Licensed under A Commons Attribution-Non Commercial 4.0 International License.

一、相关工作

（一）基于 RGB 图像的位姿估计方法

传统 RGB 位姿估计主流技术路线为“关键点检测 - 特征匹配 -PnP 位姿解算”，通过提取目标角点、边缘等二维特征，建立二维像素与三维模型坐标映射关系，求解位姿参数[1]，在纹理丰富、无遮挡、光照稳定的理想场景下性能良好。为解决低纹理目标检测难题，学界引入模板匹配算法，通过比对实景与模板图像的梯度、法向量相似度匹配目标，一定程度上改善了低纹理目标特征提取失效的问题[2]。

但该类方法局限性显著，模板匹配高度依赖完整目标图像，光照突变、局部遮挡会导致匹配精度骤降，无法适配工业复杂工况。同时，二维图像缺乏三维深度信息，难以捕捉非规则回转体曲面径向差异特征，无法实现高精度位姿解算，难以满足精密工业作业需求。

（二）基于三维数据的位姿估计方法

随着3D 视觉传感器普及，点云、体素等三维数据广泛应用于位姿估计任务，彻底解决了二维图像无深度、低纹理目标特征缺失的痛点[3]。三维位姿估计方法通过提取目标空间结构、曲面形态、轮廓边界等几何特征，直接回归位姿参数，对低纹理、无纹理回转体具备天然适配性。

该类方法存在突出工程短板：原始三维点云包含海量冗余坐标信息，直接输入网络会导致模型参数量大、算力需求高、推理速度慢，难以适配工业实时检测场景。同时，纯三维数据仅包含几何信息，缺失纹理、色彩细节，无法区分几何形态相似、表面特征不同的目标，复杂工况下泛化能力与鲁棒性不足。

（三）基于异构数据融合的位姿估计方法

异构数据融合技术整合二维纹理与三维几何特征，弥补单一模态的特征缺陷，是当前高精度位姿估计的核心研究方向。经典 DenseFusion 算法采用逐像素密集融合策略，同步挖掘图像与点云多维度特征，大幅提升位姿估计精度，成为该领域基准模型[4-5]。

后续研究通过冗余特征剔除、关键数据筛选优化模型运算效率，但现有算法针对性不足，未聚焦非规则回转体的回转结构与曲面差异化特性，对曲面交接、曲率突变等核心位姿特征区域关注度不足，有效特征利用率偏低，模型精度与工况适应性仍有待优化。

二、本文算法设计

本文结合非规则回转体复杂曲面特性与工业检测难点，设计目标分割、初始位姿估计、位姿求精三级递进式异构特征融合算法。通过面结构光传感器同步采集 RGB 图像与三维点云数据，经注意力机制精准分割目标，依托高斯曲率自适应采样提取核心几何特征，融合多模态特征完成初始位姿预测，最后通过迭代偏差修正实现位姿精准优化，兼顾算法精度、鲁棒性与实时性[6-7]。

（一）目标分割阶段

非规则回转体边缘轮廓不规则、曲面结构复杂，传统分割网络易出现边缘模糊、目标误分割、背景噪声残留等问题，直接影响后续位姿估计精度。本文在实例分割网络中嵌入轻量化注意力机制模块[8]，通过通道与空间双重注意力维度优化特征提取效果。

通道注意力自适应筛选有效特征通道，强化纹理、几何核心特征权重，抑制冗余噪声特征；空间注意力聚焦目标边缘、曲面交接等关键区域，强化差异化几何特征表达。该模块可有效优化目标分割边缘精度，剔除背景干扰与无效点云，输出高质量目标掩模与纯净裁剪点云，从源头降低分割误差对位姿解算的影响。

（二）初始位姿估计阶段

为解决传统密集融合算法运算复杂度高、非规则回转体核心特征挖掘不充分的问题，本文结合回转体几何特性，提出基于高斯曲率的自适应关键点分层采样策略。非规则回转体的核心位姿特征集中于曲面交接区域，该区域存在显著的高斯曲率突变，可精准表征曲面不规则特性，是位姿解算的核心依据[9]。

本文设计内外双层采样机制实现精细化特征筛选：外层全局采样依据回转体整体曲率分布规律，均匀筛选关键点，完整保留目标回转结构与全局空间位姿特征；内层局部采样针对曲率突变、径向轮廓差异区域密集采样，重点挖掘不规则曲面的差异化特征，强化核心位姿表达能力。该策略在保留有效特征的前提下，大幅剔除点云冗余信息，降低模型运算复杂度。随后网络融合图像纹理色彩特征与采样点云精细化几何特征，通过多层感知机完成位姿参数回归，输出高精度初始位姿结果。

（三）位姿求精阶段

初始位姿估计结果存在细微偏差，无法满足精密装配、高精度抓取的严苛需求。本文设计迭代式位姿求精模块，以初始位姿参数为基础，通过坐标变换校正三维点云姿态，建立校正点云与二维实景图像的特征映射关系。通过迭代计算二维像素特征与三维点云几何特征的匹配偏差，构建位姿偏差损失函数，借助反向传播持续修正旋转、平移参数，逐步缩小匹配误差。经多轮迭代优化，彻底消除初始位姿细微偏差，输出高稳定性、高精度的最优位姿参数。

三、算法创新点分析

本文针对非规则回转体位姿估计的行业痛点，从框架设计、特征优化、精度迭代三个维度完成创新，核心创新点如下：

（1）构建三级递进式异构数据融合位姿估计框架。打破传统单阶段位姿估计模式，设计目标分割、初始位姿估计、位姿求精全流程闭环框架，通过数据净化、特征融合、偏差迭代优化的递进逻辑，适配非规则回转体复杂几何结构，显著提升复杂工况下位姿估计的稳定性。

（2）注意力机制赋能精准目标分割。将轻量化注意力模块嵌入实例分割网络，自适应强化目标核心特征、抑制背景干扰，有效解决传统网络边缘模糊、误分割问题，优化不规则回转体边缘分割效果，从数据源头保障位姿估计精度。

（3）提出高斯曲率自适应分层采样策略。结合非规则回转体专属几何特性，通过内外双层采样兼顾全局回转结构与局部曲面差异特征，在精简点云数据、降低运算复杂度的同时，最大化挖掘核心位姿特征，解决了传统算法特征利用率低、曲面信息挖掘不足的短板，大幅提升低纹理复杂曲面目标的位姿估计精度。

 Copyright © This Work is Licensed under A Commons Attribution-Non Commercial 4.0 International License.

四、实验与结果分析

（一）实验环境与参数设置

本文算法基于 PyTorch 框架、Python3.8语言搭建，硬件平台为 Intel Core i7-12700H 处理器、32GB 内存、NVIDIA RTX 3060 12GB 显卡，搭配 CUDA11.3加速库及 OpenCV、Open3D 视觉工具库，满足三维点云运算与模型训练推理需求。模型采用 Adam 优化器，初始学习率10^{-4}，衰减系数0.95，每10轮迭代更新学习率；batch_size 设置为16，最大迭代轮数100，开启早停机制防止过拟合。高斯曲率采样设置外层全局采样1024点、内层局部采样512点，兼顾特征完整性与精准性。

（二）数据集与预处理

目前公开数据集缺乏非规则回转体工业样本，本文采用自制数据集开展实验。通过面结构光传感器采集非规则轴类、偏心回转体、异形曲面回转体等工业零部件数据，覆盖不同光照、轻微遮挡、无遮挡工况，共采集4200组 RGB 图像－点云成对有效样本。数据集按8:1:1划分为训练集、验证集、测试集。为提升模型泛化能力，对图像进行随机裁剪、亮度微调、高斯模糊增强，对原始点云完成去噪、下采样、地面冗余点剔除预处理，保证输入数据一致性。

（三）评价指标

选用工业位姿估计通用评价指标：平移误差（mm）、旋转误差（°）、单帧推理时间（ms），分别表征位置估计精度、姿态估计精度与算法实时性，误差数值越低、推理时间越短，算法综合性能越优异。

（四）对比实验与结果分析

选取 PnP 关键点算法、PointNet 算法、DenseFusion 算法作为对比模型，在统一实验环境与测试集下开展对比实验，结果如表1所示。

表1 不同算法位姿估计性能对比

算法模型	平移误差 /mm	旋转误差 /°	单帧推理时间 /ms
PnP 关键点算法	2.86	2.15	42.3
PointNet 算法	1.92	1.36	86.5
DenseFusion 算法	1.25	0.89	68.2
本文算法	0.68	0.42	51.6

由实验结果可知，传统 PnP 算法精度最低，低纹理场景特征提取失效问题突出；PointNet 纯三维算法精度优于二维算法，但推理速度慢、实时性差；DenseFusion 融合算法性能均衡，但针对性不足。本文算法相较 DenseFusion 算法，平移误差降低45.6%，旋转误差降低52.8%，推理速度提升24.3%，实现精度与实时性双重优化，综合性能最优。

（五）消融实验

为验证各创新模块的有效性，设置消融实验，结果如表2所示。单独添加注意力机制可优化分割精度，降低位姿误差；引入高斯曲率分层采样后，核心特征提取能力提升，精度进一步优化；叠加迭代求精模块后，初始位姿细微偏差被完全修正。各模块协同作用，可最大化提升算法性能，证明本文创新设计的合理性与有效性。

表2 模块消融实验结果

实验模型	平移误差 /mm	旋转误差 /°
基础融合网络	1.21	0.85
基础网络 + 注意力机制	0.93	0.64
基础网络 + 注意力机制 + 曲率采样	0.76	0.49
本文完整算法（全模块）	0.68	0.42

（六）复杂工况鲁棒性分析

针对光照波动、局部遮挡两类工业常见干扰工况开展鲁棒性测试。结果表明，传统融合算法在光照变化、轻微遮挡工况下，位姿误差涨幅分别超30%、40%，而本文算法误差涨幅仅为12%、15%左右。得益于注意力机制的抗干扰能力与多模态特征融合的互补优势，本文算法对复杂工况适应性更强，工业落地鲁棒性更优异。

五、结论

针对非规则回转体低纹理、易遮挡、曲面复杂导致的位姿估计精度低、鲁棒性差，以及传统融合算法运算效率低、特征挖掘不充分的问题，本文提出基于异构数据特征融合的三阶段位姿估计算法。算法通过注意力机制优化目标分割质量，从源头减少检测误差；依托高斯曲率自适应分层采样策略，精准挖掘曲面核心几何特征，平衡算法精度与运算效率；通过多模态特征融合与迭代求精机制，实现位姿参数高精度优化。

本文算法充分整合 RGB 图像纹理特征与三维点云几何特征，弥补单一模态缺陷，针对性适配非规则回转体几何特性，在低纹理、光照干扰、轻微遮挡等复杂工况下，相较传统算法精度与鲁棒性显著提升，同时优化了模型推理效率，可有效适配工业精密装配、机器人精准抓取等高精度作业场景。未来将进一步优化关键点动态采样策略、简化网络结构，提升算法实时性与泛化能力，同时针对重度遮挡、强光逆光等极端工况优化模型性能，拓展算法工业应用场景。

参考文献

[1] 魏隆星 . 面向智能工厂应用的工件位姿估计算法研究 [D]. 东北大学 ,2023.

[2] 邓天睿 . 基于三维点云点对特征的工件位姿估计方法研究 [D]. 西南科技大学 ,2024.

[3] 陈海永 , 李龙腾 , 陈鹏 , 等 . 复杂场景点云数据的6D 位姿估计深度学习网络 [J]. 电子与信息学报 ,2022,44(5):1847-1855.

[4] 夏文明 , 曾庞钗 , 陈斌 , 等 . 数字技术赋能数字化工厂实训室建设研究 [J]. 科技与创新 ,2025,(18):188-191.

[5] 疏阳 , 孙轶琳 , 梅振宇 , 等 . 基于多源数据的个体活动序列多步预测 [J]. 浙江大学学报 (工学版),2025,59(11):2317-2325.

[6] 房乾刚 . 基于CMM 的涡轮叶片叶型测量数据补偿和配准方法研究 [D]. 重庆大学 ,2024.

[7] 孙梦 . 三维点云数据配准技术研究 [D]. 长春工业大学 ,2023.

[8] 赵绍轩 , 周晓光 , 侯东阳 , 等 . 形状与位置先验信息引导的遥感影像建筑物域适应提取方法 [J]. 遥感学报 ,2025,29(12):3609-3625.

[9] 任可 . 基于多尺度关键点提取的关系感知型点云补全方法 [D]. 电子科技大学 ,2025.

Copyright © This Work is Licensed under A Commons Attribution-Non Commercial 4.0 International License.