

基于多关键点检测加权融合的无人机相对位姿估计算法

葛泉波^{1,2,3} 李凯¹ 张兴国⁴

摘要 针对无人机降落阶段中无人船受水面波浪影响导致图像产生运动模糊以及获取无人机相对位姿精度低且鲁棒性差的问题,提出一种基于多模型关键点加权融合的 6D 目标位姿估计算法,以提高位姿估计的精度和鲁棒性。首先,基于无人船陀螺仪得到的运动信息设计帧间抖动模型,通过还原图像信息达到降低图像噪声的目的;然后,设计一种多模型的级联回归特征提取算法,通过多模型检测舰载视觉系统获取的图像,以增强特征空间的多样性;同时,将检测过程中关键点定位形状增量集作为融合权重对模型进行加权融合,以提高特征空间的鲁棒性;紧接着,利用 EPnP (Efficient perspective-n-point) 计算关键点相机坐标系坐标,将 PnP (Perspective-n-point) 问题转化为 ICP (Iterative closest point) 问题;最终,基于关键点解集的离散度为关键点赋权,使用 ICP 算法求解位姿以削弱深度信息对位姿的影响。仿真结果表明,该算法能够建立一个精度更高的特征空间,使得位姿解算时特征映射的损失降低,最终提高位姿解算的精度。

关键词 辅助无人机降落,舰载视觉系统,6D 位姿估计,加权融合,关键点检测,级联特征提取

引用格式 葛泉波,李凯,张兴国. 基于多关键点检测加权融合的无人机相对位姿估计算法. 自动化学报, 2024, 50(7): 1402-1416

DOI 10.16383/j.aas.c230297

Relative Pose Estimation Algorithm for Unmanned Aerial Vehicles Based on Weighted Fusion of Multiple Keypoint Detection

GE Quan-Bo^{1,2,3} LI Kai¹ ZHANG Xing-Guo⁴

Abstract A 6D target pose estimation algorithm based on multiple models keypoints weighted fusion is proposed to address the issue of low accuracy and poor robustness in obtaining the relative pose of unmanned aerial vehicles due to motion blur caused by the influence of water surface waves on the image during the landing phase of unmanned aerial vehicles. This algorithm aims to improve the accuracy and robustness of pose estimation. Firstly, based on the motion information obtained from the unmanned ship gyroscope, an inter frame jitter model is designed to reduce image noise by restoring image information. Then, a cascaded regression feature extraction algorithm with multiple models is designed to detect images obtained by the shipborne visual system through multiple models, in order to enhance the diversity of the feature space; at the same time, the incremental set of keypoint localization shapes during the detection process is used as the fusion weight to weight and fuse the model, in order to improve the robustness of the feature space. This paper uses efficient perspective-n-point (EPnP) to calculate the coordinates of the camera coordinate system for keypoints, and transforms the perspective-n-point (PnP) problem into an iterative closest point (ICP) problem. Finally, based on the dispersion of the keypoints solution set, weights are assigned to keypoints, and the ICP algorithm is used to mitigate the influence of depth information on the pose estimation. The simulation results show that this algorithm can establish a more accurate feature space, reduce the loss of feature mapping during pose estimation, and ultimately improve the accuracy of pose estimation.

Key words Assisting unmanned aerial vehicles landing, shipborne visual system, 6D pose estimation, weighted fusion, keypoint detection, cascading feature extraction

Citation Ge Quan-Bo, Li Kai, Zhang Xing-Guo. Relative pose estimation algorithm for unmanned aerial vehicles based on weighted fusion of multiple keypoint detection. *Acta Automatica Sinica*, 2024, 50(7): 1402-1416

收稿日期 2023-05-23 录用日期 2023-11-20

Manuscript received May 23, 2023; accepted November 20, 2023

国家自然科学基金(62033010),江苏高校“青蓝工程”(R2023Q07)资助

Supported by National Natural Science Foundation of China (62033010) and Qing Lan Project of Jiangsu Province (R2023Q07)

本文责任编辑 陈谋

Recommended by Associate Editor CHEN Mou

1. 南京信息工程大学自动化学院 南京 210044 2. 江苏大数据分析
与智能系统省高校重点实验室 南京 210044 3. 大气环境与装备技术

6D 目标位姿估计是计算机视觉领域的关键技术,其在机器人抓取、增强现实和舰载机降落^[1-2]等领域有着广泛应用. 6D 目标位姿估计的目标是利

协同创新中心 南京 210044 4. 中国飞行试验研究院 西安 710089

1. School of Automation, Nanjing University of Information Science and Technology, Nanjing 210044 2. Jiangsu Provincial University Key Laboratory of Big Data Analysis and Intelligent Systems, Nanjing 210044 3. Jiangsu Collaborative Innovation Center on Atmospheric Environment and Equipment Technology (CICAEET), Nanjing 210044. China Flight Test Research Institute, Xi'an 710089

用相机拍摄到的图像或视频获得目标到相机坐标系的单应性变换. 现阶段 6D 位姿估计方法主要分为传统方法和基于深度学习的方法. 传统 6D 位姿估计方法一般是基于特征匹配来完成的, 首先提取图像中的特征, 如 SIFT^[3]、SURF^[4]、ORB^[5] 等特征点, 然后与先验模板库中的图像进行特征匹配, 再将匹配成功的特征映射关系转化为 PnP (Perspective-n-point)^[6] 问题进行求解, 最终得到目标的位姿信息. 此类方法的精度受制于目标纹理丰富度和模板库数量, 当目标纹理不清晰或模板库准备不够充分时, 其位姿估计精度会大大降低. 随着人工智能技术的发展, 文献 [7-9] 研究了基于深度学习的 6D 位姿估计, 该类方法主要分为直接回归法和间接回归法. 直接回归法实质是训练一个神经网络, 使其能够端到端地通过一幅图像输出目标的旋转和平移向量. 整个网络架构一般由目标检测与目标跟踪、空间映射、位姿恢复和位姿优化 4 个模块组成. 文献 [10-11] 利用多网络结构在语义分割、平移、旋转三个网络子分支共享提取到的特征, 增强了对遮挡的对抗性, 但无法保证该方法的实时性. 文献 [12] 在估计时引入目标深度信息, 去除了对位姿优化的依赖, 但是在面对光照影响时表现不佳. 文献 [13-14] 则是针对未知新物体的类别级位姿估计问题, 通过充分考虑类别的共有特征来提高模型的泛化性, 但在面对多目标或者目标存在噪声时, 难以保证该方法的精度. 间接回归法的实质是寻找一组相同点在两个不同坐标系下的转换关系. 按照先验信息的不同可以分为 2D-3D 对应和 3D-3D 对应, 其中在已知一组点的 2D 相机投影坐标和与之对应的 3D 世界坐标 (即地图是先验的) 时, 可以构建 PnP 问题, 利用 2D 与 3D 间的对应关系进行位姿求解, 得到相机相对于世界坐标系的位姿, 然后利用已知的目标世界坐标系位姿和相机世界坐标系位姿求解出目标相对于相机的位姿. 在关键点定位优化方面, 如 PVnet (Perspective-n-Point network)^[15] 利用每个关键点附近的向量场能够更加精确地定位出关键点位置, 以获取精度更高的特征空间. 不利用目标自身关键点的方法, 例如 BB8^[16]、YOLO-6D^[17]、SSD-6D^[18] 等方法利用基于卷积神经网络的网络框架对输入的 RGB 图像进行目标 3D 包围盒顶点的预测, 然后将其投影到二维平面, 并使用 PnP 方法进行位姿求解. 但是, 以上方法在面对包围盒超出图像边框时会造成错误的位姿估计. 与检测目标特征点的方法相比, 文献 [19-22] 将地标作为先验信息来提取特征, 提高了特征间的联系, 该方法虽可增加整体特征精度, 但缺乏思考噪声对地标纹理的破坏问题. 文献 [23-26] 则考虑到图像存在噪声的影响, 对输入的特征分别使用标志点加权、特征简化以及特征

融合等方法进行降噪, 不过仍缺乏讨论复杂环境对成像设备的干扰问题. 而在已知一组点的 3D 相机坐标系坐标和与之对应的 3D 坐标 (目标的 3D 模型) 时, 可以构建 ICP (Iterative closest point)^[27] 问题, 利用 3D 点与 3D 点间的对应关系对两片点云进行配准, 以求解相机相对于目标坐标系的位姿. 然而在面对大规模点云配准时, 该类方法易产生错误的 3D-3D 映射关系, 在进行位姿解算时会带来很大的误差, 对最终的位姿结果的精度影响较大.

面对实际工程应用, 上述方法均缺少对成像设备受环境干扰的考量且在计算位姿时未考虑特征空间出现错误的情况, 这可能造成实际位姿计算出现偏差, 进而增大应用的危险性. 考虑以上问题, 本文针对无人船辅助无人机完成水面降落的姿态估计问题, 提出一种基于关键点的 6D 位姿估计算法, 利用多模型的关键点检测算法和先验的 3D 点信息构造 PnP 问题, 然后利用 EPnP (Efficient perspective-n-point)^[6] 算法得到每个关键点的相机坐标系坐标, 最终转化为 ICP 问题并求解位姿. 本文主要工作如下: 1) 针对成像设备因受海面随机抖动而产生运动模糊问题, 在预处理阶段利用维纳滤波进行非盲图像复原, 并引入受无人船动平台影响的图像数据集. 2) 考虑单一关键点类型无法处理所有程度的运动模糊问题, 提出关键点类型选择算法. 3) 鉴于稀疏匹配方法存在不稳定的问题, 提出基于多特征提取模型加权融合的级联回归方法, 以增加关键点数量, 提高特征空间的多样性; 同时引入特征空间建立时的筛选机制, 以增加特征空间的鲁棒性和精度. 4) 提出利用关键点的离散度作为权值的远距离 ICP 优化方法, 以减少深度信息对位姿的影响.

本文中动平台指的是无人机的降落平台, 即无人船本身, 无人船在水上场景行驶时与水面波浪存在相对运动, 因此在文中称之为动平台. 其一般受水面风浪流影响, 存在随机的六自由度抖动现象, 容易对成像设备造成一定程度的运动模糊. 根据国际海事组织 (International Maritime Organization, IMO) 要求, 船舶在低载荷的状态下倾角应处于 30° 以下, 同时为验证在大倾角下算法依旧有效, 文中实验结论涉及到的运动倾角被限制在 45° 以下.

1 问题分析

1.1 系统描述

本文建立的基于关键点的无人机位姿估计算法考虑到由于水面环境光的反射会导致无人机自身搭

载的 RGB 视觉传感器成像质量下降, 单单依靠无人机自身解算出的位姿已无法保证其降落精度. 因此通过在无人船上搭载一个 RGB 视觉传感器来辅助无人船获取降落过程中的无人机相对位姿信息, 以保证无人机降落的安全性.

本文研究内容为使用 RGB 相机在水面未知场景下估计出无人机相对于无人船 (即相机) 的位姿, 即在一个未知世界坐标系和无深度信息的前提下计算出无人机的位姿. 在以上环境下单一地使用 PnP 或 ICP 方法无法得到精度较高的结果. 因此本文的思路为首先利用 EPnP 方法将 2D 相机投影点变换到相机坐标系下的 3D 点, 然后用 ICP 方法求解 3D-3D 的变换, 最终得到无人机相对于相机的位姿. 在本文的场景中, 将世界坐标系设置在无人机身上, 以无人机底盘标定板作为 ZOX 平面, 以位于 ZOX 平面的无人机最右后方点作为坐标原点, Z 轴朝向无人机正面, 利用右手准则制定整个世界坐标系. 然后通过 3D 建模得到事先标定的待检测关键点的 3D 世界坐标系坐标 (即无人机自身 3D 坐标), 并且对每个关键点设置不同身份以保证 2D 相机投影点与无人机 3D 世界坐标系坐标间的匹配正确性. 但是本文使用的是 RGB 相机, 无法得到另一组 3D 点 (3D 相机坐标系坐标), 所以一开始无法构建 3D-3D 的对应关系进行 ICP 求解. 于是可以将核心问题转变为获取与无人机 2D 相机投影点和无人机 3D 世界坐标系坐标对应的无人机 3D 相机坐标系坐标. 针对此问题, 通过 EPnP 方法结合刚体结构不变性, 得到控制点的相机坐标系坐标. 然后得到一组相对应的无人机世界坐标系 3D 点及其相机坐标系 3D 点. 因此可以将 2D-3D 问题转化为 3D-3D 问题, 最后利用 ICP 方法求解最终的无人机相对相机的位姿. 在整个流程中误差的主要来源为 2D 关键点的定位, 而产生误差的原因是由于海面波浪导致成像设备晃动, 产生的运动模糊使得关键点定位无法做到百分百的准确, 也就是无法建立完美的特征空间. 在本文中, 针对海面场景会对 2D 关键点定位产生的负面影响进行分析和处理, 其中主要包括海面波浪^[28-29]造成的运动模糊、单一特征采集无法构造优质的特征空间、无人机距离降落平台较远导致物方残差增大. 本文通过图像复原、关键点类型选择、级联特征提取^[30-34]、特征筛选机制^[35]、设计新的特征融合策略和增加关键点权重方法对以上问题进行解决, 以增强最终解算位姿的精确度和鲁棒性.

1.2 研究动机

近年来, 利用 RGB 视觉传感器的无人机降落

系统虽得到了很大的发展, 但在无人机的水面场景应用中仍存在很多不足, 如图 1 所示.

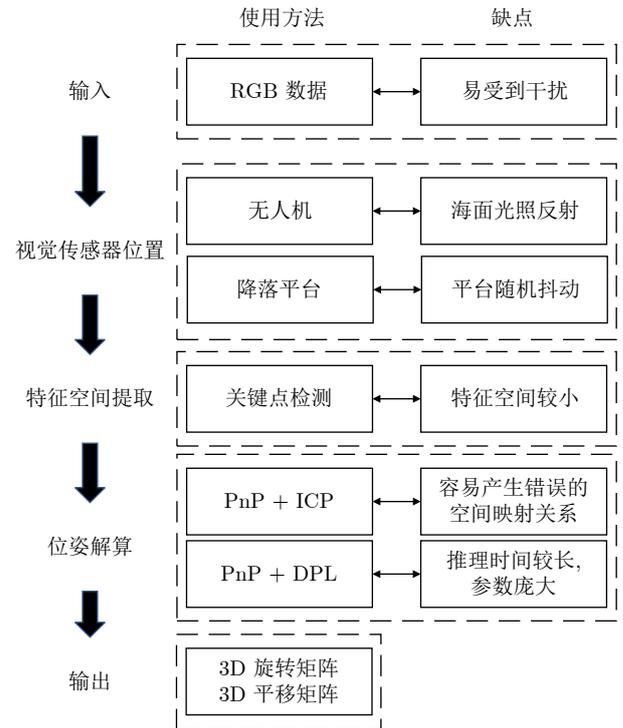


图 1 基于关键点检测的位姿估计框架及其缺点
Fig.1 The pose estimation framework based on keypoint detection and its drawbacks

1) 传统搭载在降落平台的成像系统^[36-41]缺少对环境随机干扰的考虑及应对复杂随机环境的有效处理. 传统陆地成像系统仅需要考虑四自由度且成像设备不易受到抖动等外界因素的影响, 而搭载在无人船上的成像设备时刻会受到来自风浪流的干扰, 从而产生不规则的六自由度抖动现象. 这些不规则的抖动会造成成像时出现一定程度的运动模糊, 导致图像的纹理信息被破坏. 如图 2 所示, 图 a 为未受到船体抖动影响时获取的图像, 其中大部分纹理特征能够清晰辨别; 图 b 为受到轻微船体抖动 ($0^\circ \sim 15^\circ$) 影响时拍摄到的图像, 其纹理特征存在模糊现象, 但依旧能够分辨; 图 c 为受到重度船体抖动 ($15^\circ \sim 45^\circ$) 影响时拍摄到的图像, 可以发现出现了很大的模糊现象, 大量纹理信息被破坏. 由此可见不同程度的平台帧间倾斜抖动均会造成运动模糊, 而这些运动模糊会对图像自身的纹理信息产生干扰和破坏, 进而导致关键点检测信息和空间映射的精度和鲁棒性大大下降, 最终解算出的较低精度位姿信息将增大无人机自主降落的危险性.

2) 在关键点类型及其运动模糊程度不同的情况下, 现有检测算法 (如文献 [20-21, 42-47] 中算



图 2 运动模糊对纹理的影响

Fig.2 Effect of motion blur on texture

法) 缺少统一的计算标准, 泛化性较差. 由于海面波浪随机扰动对图像造成的影响只能进行抑制而无法完全消除, 因此图像无法避免不同程度的纹理损失. 同时各个关键点类型在不同图像纹理损失程度下的适用性也不尽相同. 对于角点来说, 是通过寻找一个区域的梯度极值点来获取的, 在面对运动模糊时, 图像的梯度幅值会降低、梯度方向会被模糊, 但是一个区域的极值点不会产生很大的偏差. 而对于纹理中心点来说, 是需要利用图像自身的特性 (如对称性、高对比度、突出的形状等特征) 来进行更精确的定位, 虽然能够得到更好的定位精度, 但在面对较大的运动模糊时, 图像自身的特性会受到很大程度破坏, 会对纹理中心点的定位带来干扰. 由此可见, 不同的关键点类型存在不同的优势.

3) 面对图像纹理被破坏的情况, 特征空间映射的数量和鲁棒性难以保证, 位姿解算的精度也无法得到有效保证. 根据文献 [6] 可知位姿解算的精度与特征空间的映射数量成正比, 至少需要 10 对以上的特征才能保证较好的位姿解算结果. 而现有间接式 6D 位姿估计算法为了保证整体算法的实时性, 事先设计的特征映射数量往往是能够达到最优位姿解算的最低标准, 既可得到最优的位姿结果又不会大幅增加解算时间. 但当成像设备时刻处于随机运动模糊的情况下, 特征空间映射易产生偏差和丢失, 进而降低解算出的位姿精度和鲁棒性.

4) 传统 ICP 位姿优化算法 (如文献 [19, 48–50] 中算法) 未考虑到距离过远时深度信息会造成关键点定位精度差异增大, 最终导致优化程度下降的问题. 传统 ICP 是通过构建最小二乘问题来最小化重投影误差进而优化位姿, 构建过程中并未对关键点的权值加以区分. 这是因为在近距离时各关键点的精度差异不大, 而在面对深度大幅增加和无人机尺度增大时, 一个像素位置代表的物方实际距离也随之增大. 这便会导致相机坐标系关键点定位的实际误差增大, 最终导致各关键点精度的差异增大. 若是对每个关键点赋予相同的权值, 则会放大精度较低的关键点对位姿的影响, 并稀释精度较高的关键点对位姿计算的贡献, 最终导致位姿解算结果的精度下降.

1.3 解决方案

针对第 1.2 节描述的系统存在的问题, 提出如下解决方法, 如图 3 所示.

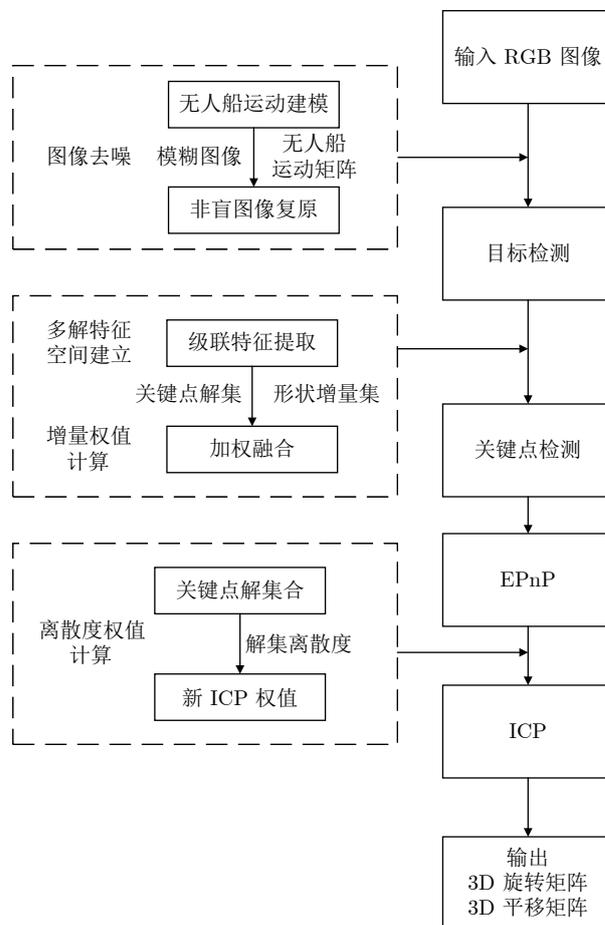


图 3 解决方案

Fig.3 Solution

1) 针对海面波浪随机抖动对图像影响的问题, 根据维纳滤波算法提出利用运动建模进行非盲图像复原的方法, 以削弱动平台抖动对图像的影响. 首先通过先验信息可知, 动平台图像噪声的主要来源为运动模糊. 所以, 通过对无人船进行运动建模, 得到帧间的无人船运动矩阵, 从而获得运动模糊的方向和距离. 然后基于无人船运动矩阵, 利用无人船拍摄得到的模糊图像与已知模糊方向和距离信息反解出清晰图像. 以此方法复原的清晰图像能够有效削弱运动模糊对图像的影响, 为后续构建特征空间提供一个较优质的输入, 最终提高特征的精度.

2) 针对单一关键点类型无法在所有程度的运动模糊下达到最优精度的问题, 提出一种合适关键点类型选择算法, 通过选取应对当前模糊的最优关键点类型, 以增强关键点的鲁棒性和精度. 在进行

关键点检测前需要得到两个先验信息,一个是拍摄图像时无人船的运动信息,另一个是关键点类型与图像运动模糊程度的映射关系.考虑到帧间船体倾角越大,其造成的运动模糊程度越大,对应的图像将遭受更大的纹理损失.因此可以利用无人船的帧间运动信息得到无人船的帧间倾角,然后通过运动模糊程度与帧间倾角呈正相关的先验信息和关键点类型与运动模糊程度的映射关系选择一个合适的的关键点类型进行检测.这样可以保证在海面随机运动模糊下找出适用性更高的关键点,满足在面对海面高度随机扰动时特征所需的高泛化性.

3) 针对特征空间映射因受到噪声的破坏而导致稀疏匹配精度和鲁棒性下降的问题,根据级联回归算法提出一种基于级联回归的多特征提取模型加权融合算法,以增强特征空间的鲁棒性.利用多特征提取模型和级联回归方式实现特征检测和特征空间构建,可以保证关键点数量,有效解决稀疏匹配因特征数量不够而产生的精度较低的问题.同时在回归时利用模型间较为平滑的迭代可以降低多次提取特征带来的累计负担和时间开销.另外,在模型评估中借鉴学习率设置思想:太小收敛太慢,太大易产生波动.在级联回归中增加一个评估指标:形状增量集,利用该指标对模型中间定位过程进行映射,并将其类比于学习率的设置.将每次级联回归产生的关键点形状增量总和的归一化数值作为每种特征提取模型的融合权值,并通过对多种提取模型结果的加权融合来增强整体结果的鲁棒性.使用此算法可以为后续位姿解算提供足够多且鲁棒性高的特征对输入.

4) 针对 ICP 算法在面对远距离目标时物方残差增大而产生的位姿解算精度下降的问题,提出基于关键点解集离散度的加权 ICP 位姿优化算法,以削弱深度信息对位姿优化的影响.首先通过多模型的级联回归方法得到关于同一关键点的多个定位结果,构成关键点定位解集,每个集合包含对该点进行多次检测的结果.将属于同一关键点不同定位结果间的离散度进行求和并做归一化处理,将处理后的系数作为加权 ICP 优化算法的权重.根据定位距离越相近,定位精度和鲁棒性越高的先验知识,离散度低的关键点将获得更高的权值,在后续的位姿优化中起到更大的影响.使用此方法可以减少深度信息对最终位姿的影响,进而提升位姿优化的鲁棒性.

2 基于关键点的无人机相对位姿估计

传统位姿求解的方法受制于纹理的丰富程度,若图像受到噪声的干扰,则位姿求解的精度会受到

很大影响.而基于深度学习去预测 3D 包围盒的 8 个顶点也存在包围盒超出图像范围和稀疏匹配鲁棒性低的问题.考虑到本文所面对的海上环境噪声较多和船体的晃动导致目标不容易出现在图像正中心等,设计一种能够应对水面随机扰动的位姿估计算法.首先利用无人船的帧间运动矩阵通过非盲图像复原算法削弱无人船动平台抖动的影响,同时完成关键点类型的选择.然后利用目标检测算法得到无人机在图像上的具体位置信息.接着利用多模型级联回归特征提取加权融合算法对无人机的关键点进行检测,并利用关键点剔除原则加强特征空间的鲁棒性.最终利用关键点解集的离散度信息完成远距离优化的加权 ICP 位姿估计算法,以得到无人机的相对位姿,流程如图 4 所示.

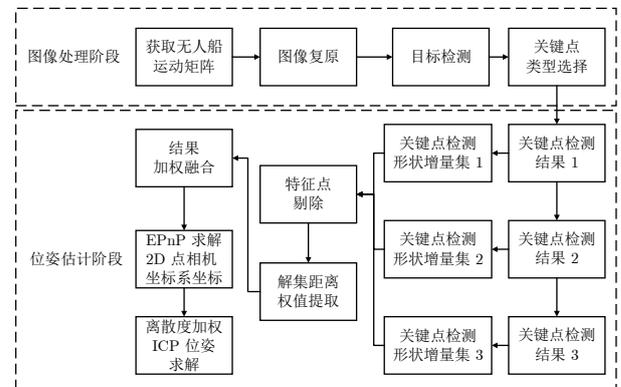


图 4 算法流程图

Fig. 4 The flow chart of algorithm

2.1 非盲图像复原

由于现有的无人机位姿估计系统缺少对成像设备所处环境的考虑,难以处理随机环境干扰下的成像问题,本节针对将成像设备搭载在无人船上时所解决的环境干扰问题提出非盲图像复原的方法,以增强 2D 相机投影点定位的精度.在无人船上通过 RGB 相机去获得无人机图像的过程中,虽然无人船的速度保持为零,但由于水面的波动会导致无人船产生相对运动,从而导致相机产生晃动.这会使得拍摄的图像产生拖尾和扭曲效果,破坏图像细节,最终导致 2D 相机投影点定位的误差增大.因此本文就抑制运动模糊这一环节,首先对无人船进行运动建模来获取帧间运动矩阵,并利用帧间运动矩阵获得此帧的模糊方向和距离,最终通过维纳滤波^[5]得到清晰的图像.

本节考虑的是无人船静止状态(即速度为 0 m/s)且仅受到海面波浪影响的情况下,通过无人船帧间

运动矩阵 $[\mathbf{R}_s|\mathbf{T}_s]$ 对拍摄得到的图像进行复原, 最终得到清晰图像. 以下是无人船帧间运动建模及图像复原过程^[51]

$$\begin{bmatrix} u_0 \\ v_0 \end{bmatrix} = \mathbf{K}\mathbf{R}_s\mathbf{K}^{-1} \begin{bmatrix} u_1 \\ v_1 \end{bmatrix} + \mathbf{T}_s \quad (1)$$

其中, (u_0, v_0) 为模糊图像的像素位置, (u_1, v_1) 为清晰图像的像素位置, $\mathbf{R}_s, \mathbf{T}_s$ 分别为无人船帧间运动的旋转分量和平移分量, \mathbf{K} 为相机的内参矩阵.

由于本系统的图像噪声主要由运动模糊造成, 因此可以利用船体的运动矩阵 $[\mathbf{R}_s|\mathbf{T}_s]$ 得到模糊的方向和距离, 更好地完成图像复原. 下式为图像复原的求解方程

$$\mathbf{Y} = [\mathbf{R}_s|\mathbf{T}_s] * \mathbf{X} + \mathbf{n} \quad (2)$$

其中, \mathbf{X} 为需要复原的清晰图像, \mathbf{Y} 为船载视觉系统拍摄到的图像, $*$ 为卷积操作, \mathbf{n} 为海面波动引起的噪声.

本节通过使用无人船的运动矩阵得到无人机图像的模糊方向和距离, 有效帮助图像进行复原. 在已知运动矩阵的前提下, 利用维纳滤波可以得到复原后的图像 \mathbf{X} , 并且将 \mathbf{X} 与无人船运动矩阵 $[\mathbf{R}_s|\mathbf{T}_s]$ 配对. 这可有效减少动平台对图像产生的噪声影响, 同时为后续关键点检测提供清晰图像 \mathbf{X} 和帧间运动信息 $[\mathbf{R}_s|\mathbf{T}_s]$, 提高关键点检测精度.

2.2 目标检测

考虑到对无人机进行位姿估计有实时性的要求, 因此需要使用目标检测或者目标分割对待处理区域进行缩减, 以减少特征提取所需的时间. 虽然目标分割能够剔除背景因素的干扰, 但其最终得到的目标区域一般为不规则形状, 这对于关键点检测是十分不便的. 同时目标分割耗时相较于目标检测过大^[9], 使得位姿估计要求实时性的初衷难以实现. 因此在本文的研究中, 选用目标检测作为缩减待处理区域的手段, 并针对速度和泛化能力, 最终选择 YOLOV5^[52] 作为目标检测阶段的算法. YOLOV5 目标检测主要有以下几点优势: 1) 能够抽取整张图像信息, 泛化能力强大; 2) 模型推理速度快, 能够在实时场景下进行目标检测; 3) 模型具有高的鲁棒性和精度. 以上几点均符合海面估计无人机位姿的场景.

本节的工作是对第 2.1 节复原后的图像 \mathbf{X} 进行目标检测, 得到一个包含无人机的最小矩形框 (x, y, w, h) , 其中 (x, y) 为无人机所在矩形的中心点坐标, w 为矩形框的宽度, h 为矩形框的高度. 这样做不仅能够快速完成识别的基本任务, 同时可以提供一个最小的目标矩形框, 相当于对图像无用部

分做了一次粗剔除, 因此可以大大减少级联回归特征提取和特征匹配的时间开销.

2.3 级联回归关键点检测

现有的特征提取方法只关注特征定位的结果, 缺少对特征定位过程中信息的充分利用, 这会降低特征的稳定性. 本节对特征提取过程做出了改进, 提出一种加权的多模型级联特征提取算法. 在提取特征的环节, 使用多个并行的级联特征提取模型对目标进行特征提取, 通过提取过程的形状增量集来对每个模型的稳定性进行考量, 最终利用形状增量集的距离和作为权值对多个模型进行加权融合.

级联回归算法是一种通过训练后能够快速定位出目标关键点的算法, 回归的目标并非只有关键点本身, 同时包括回归过程中关键点的形状增量集合^[31-33]. 其原理是通过训练的方法得到一组形状回归器 \mathbf{W} , 形状回归器从初始形状 \mathbf{S}^0 开始, 经过 T 次级联逐步回归到最终的预测形状 \mathbf{S}^T . 其中每一次迭代由不同的回归器完成, 并且每一个回归器的输入都依赖前一个回归器的输出, 所有的级联回归器是通过最小化事先标记好的形状 $\hat{\mathbf{S}}$ 与当前回归形状的对齐误差学习到的. 近几年来深度学习在关键点定位上取得了重大突破, 主流的方法如文献 [33], 其构建局部二元特征 (Local binary features, LBF) 并利用所有局部特征构建一个全局特征, 最后利用全局特征来训练回归器. 文献 [53] 对下降方法做出了改进, 将以往求解海森矩阵和雅可比矩阵的步骤替换为求解两者的乘积, 这样就解决了求解海森矩阵时要用到二阶导的问题, 大大简化了计算复杂度. 使用级联回归算法提取特征能够直接将关注点全部放在目标上且无需对整幅图像做特征提取, 其训练流程如下:

1) 首先输入图像数据 \mathbf{I} 和事先标记好的关键点位置 $\hat{\mathbf{S}}$. 其中, $\hat{\mathbf{S}}$ 的表现形式为 $\{x_0, \dots, x_l, \dots, x_L, y_0, \dots, y_l, \dots, y_L\}$, L 为事先规定的关键点的个数.

2) 优化目标是最小化图像的对齐误差 ε

$$\varepsilon = \min \|\mathbf{S}_i - \hat{\mathbf{S}}_i\| \quad (3)$$

其中, \mathbf{S}_i 是第 i 张图像的回归结果, $\hat{\mathbf{S}}_i$ 是第 i 张图像的真实关键点分布, $\hat{\mathbf{S}}_i$ 是事先给出的.

3) 关键点的迭代依赖于前一次迭代的结果, 利用前一次估计的结果来更新本次迭代, 即^[31-33]

$$\mathbf{S}_i^t = \mathbf{S}_i^{t-1} + \Delta \mathbf{S}_i^{t-1} \quad (4)$$

其中, \mathbf{S}_i^{t-1} 为第 i 张图像的第 $t-1$ 次迭代结果, t 表示第 t 次迭代, $t = 1, 2, \dots, T$, $\Delta \mathbf{S}_i^{t-1}$ 为第 $t-1$

次迭代的形状增量.

4) 将形状增量与前一次的估计结果相加得到下一次的迭代结果

$$\Delta \mathbf{S}^t = \mathbf{W}^t(\mathbf{I}, \mathbf{S}^{t-1}) \quad (5)$$

其中, $\Delta \mathbf{S}^t$ 为第 t 次迭代的形状增量, \mathbf{W}^t 为第 t 次迭代所用的回归器.

5) 加入正则项防止出现过拟合

$$\mathbf{W}^t = \arg \min_{\mathbf{W}^t} \sum_{i=1}^M (\|\Delta \hat{\mathbf{S}}_i^t - \mathbf{W}^t(\mathbf{I}_i, \mathbf{S}_i^{t-1})\|_2^2 + \lambda \|\mathbf{W}^t\|_2^2) \quad (6)$$

其中, $\Delta \hat{\mathbf{S}}_i^t$ 为实际的形状增量, 是已知的. $\mathbf{W} = \{\mathbf{W}^1, \dots, \mathbf{W}^t, \dots, \mathbf{W}^T\}$ 是训练出来的一组回归器, M 表示图像的数量, λ 为正则化系数.

关键点定位回归流程如下^[30-34]:

1) 常用的初始形状 \mathbf{S}^0 有两种选择方式: 平均形状和随机形状. 本文使用的是平均初始形状

$$\mathbf{S}^0 = \frac{\sum_{i=1}^M \hat{\mathbf{S}}_i}{M} \quad (7)$$

2) 根据回归器组进行回归

$$\mathbf{S}^t = \mathbf{S}^{t-1} + \mathbf{W}^t(\mathbf{I}, \mathbf{S}^{t-1}), t = 1, 2, \dots, T \quad (8)$$

本节经过级联特征提取最终可以得到多组关键点的定位解集 $\alpha_f = \{x_f^1, \dots, x_f^L, y_f^1, \dots, y_f^L\}$ 和各方法定位过程中产生的形状增量集 $\Delta \mathbf{S}_f = \{\Delta \mathbf{S}_f^1, \dots, \Delta \mathbf{S}_f^t, \dots, \Delta \mathbf{S}_f^T\}$, $\Delta \mathbf{S}_f^t$ 表示第 f 种特征提取模型的第 t 次迭代的形状增量.

2.4 关键点剔除原则

现有的位姿估计算法缺少对特征空间的质量衡量, 难以处理特征空间被噪声污染的情况. 本节对特征空间的建立和筛选做出改进, 提出一种关键点剔除原则. 在获取第 2.3 节的关键点解集 α_f 后, 利用不同结果间的距离之和作为衡量特征的质量标准, 以此来剔除定位不稳定的关键点, 最终得到一个更加准确的关键点解集 $\beta_f = \{\beta_f^1, \dots, \beta_f^L, \dots, \beta_f^L\}$, 其中 β_f^l 表示第 f 种特征提取模型得到的第 l 个关键点的定位信息, L 表示关键点的总数. 以下是剔除原则:

1) 同一个关键点不同模型定位结果所围成多边形的曼哈顿距离不能过大. 当 $\sum_{f=1}^{F-1} (\|x_f^l - x_{f+1}^l\| + \|y_f^l - y_{f+1}^l\|)$ 过大时, 则说明所融合的特征提取方法得到的结果差异较大, 但是系统无法辨别出哪个关键点更加准确, 因此直接舍弃该关键点, 其中 F

表示特征提取模型的数量.

多边形的曼哈顿距离 $d(x, y)$ 定义如下:

$$d(x, y) = \sum_{f=1}^{F-1} (|x_f - x_{f+1}| + |y_f - y_{f+1}|) \quad (9)$$

其中, (x_f, y_f) 表示第 f 种特征提取模型得到的关键点坐标.

2) 关键点的像素累计差值不能过大, 如果 $\sum_{f=1}^{F-1} |\mathbf{I}_f^l - \mathbf{I}_{f+1}^l|$ 过大, 则说明特征提取到的像素本身的差别很大, 应当舍弃. 其中 \mathbf{I}_f^l 表示第 f 种特征提取模型的第 l 个关键点的像素值.

由于本文的应用场景存在海面或水面波浪随机扰动影响, 无法通过外部措施完全消除这些扰动对图像的影响. 为此设计以上的关键点剔除原则, 以进一步剔除质量较差的特征空间映射对. 首先通过比较同一个关键点定位解集 (多个模型对同一个关键点的不同定位结果) 的曼哈顿距离来剔除定位过于离散的关键点, 选用曼哈顿距离可以避免浮点运算从而减少时间开销和误差累积. 然后通过比较同一关键点定位解集的像素差值来剔除像素差异较大的关键点. 使用以上剔除原则可以在无法完全消除图像噪声的前提下进一步保证关键点信息的鲁棒性和精度, 得到一个精度更高的关键点解集 β_f .

2.5 多模型级联回归特征提取加权融合

本节利用第 2.4 节得到的关键点解集 β_f 对第 2.3 节得到的多模型结果 \mathbf{S}_f^T 进行加权融合.

2.5.1 关键点定位加权融合权值确定法

由于系统通过级联的多特征提取模型提取关键点位置, 所以可以得到每种特征提取模型在级联回归过程中的形状增量轨迹, 可利用这些增量轨迹的距离变化和 (即形状增量集合) 作为每种提取模型的权值^[30-34]. 与学习率一样, 考虑形状增量集合过大和过小的情况. 形状增量集过小可能是由两种情况产生的: 1) 初始定位和最终结果相差不大; 2) 关键点和周围环境难以做出有效的区分 (海面场景与无人机相差较大, 实验中并没有出现该问题). 考虑到以上两种情况在海面实际场景的出现概率, 将形状增量集小的模型赋予大的权重. 而形状增量集过大则说明在定位过程中出现了大幅度的抖动, 此次定位过程在某些环节中出现了来回的波动, 模型无法对波动点的特征进行有效区分, 因此定位的结果是有歧义的, 认为该模型是不鲁棒的, 需要赋予较小的权重. 通过以上方式对模型结果进行质量评估并加权融合, 可以有效减小少数定位误差点带来的特征空间误差.

步骤 1. 首先以级联回归的方式统计每种特征提取模型的每一步定位结果, 得到各方式的形状增量集合

$$\Delta S_f = \{\Delta S_f^1, \dots, \Delta S_f^t, \dots, \Delta S_f^T\} \quad (10)$$

步骤 2. 计算每种特征提取模型所得到的形状增量集的距离绝对值并求和, 得到关键点定位过程中各模型的移动步长

$$\varphi_f = \sum_{t=1}^T |\Delta S_f^t| \quad (11)$$

其中, φ_f 为第 f 种模型所有关键点回归的累积位移和。

步骤 3. 确定每种特征提取模型的归一化权值

$$q_f = \frac{\frac{1}{\varphi_f}}{\sum_{f=1}^F \frac{1}{\varphi_f}}, \quad f = 1, 2, \dots, F \quad (12)$$

其中, q_f 为第 f 种模型的融合权重。

2.5.2 加权融合策略

由于使用多特征模型提取会得到多组解, 为保证每组解都为特征空间映射做出一定贡献, 考虑使用加权的方式进行融合处理。

步骤 1. 根据上文中权值确定法得到权值和各方式优劣排名, 按照形状增量集合距离的绝对值总和从大到小的顺序得到最优、次优、最差三种结果。

步骤 2. 关键点的移动原则是性能差的往性能好的方向移动, 也就是次优方法往最优方法移动, 最差方法往次优和最优方法移动。方向的判定标准如下 (假设特征提取的性能从高到低为 $i > j > k$):

情况 1. 次优往最优移动

如果 $x_i - x_j > 0$, 则 $x'_j = x_j + d$; 如果 $x'_j > x_i$, 则 $x'_j = x_i$; y 方向同理。

情况 2. 最差往次优和最优移动

先往次优移动: 如果 $x_j - x_k > 0$, 则 $x'_k = x_k + d$; 如果 $x'_k > x_j$, 则 $x'_k = x_j$; y 方向同理。然后更新后的点向最优方向移动: 如果 $x_i - x'_k > 0$, 则 $x''_k = x'_k + d$; 如果 $x''_k > x_i$, 则 $x''_k = x_i$; y 方向同理。

移动距离的标准定义为:

$$d_{i,j}^x = q_f |x_i - x_j| \quad (13)$$

$$d_{i,j}^y = q_f |y_i - y_j| \quad (14)$$

其中, $d_{i,j}^x$ 为 x 方向的移动距离, $d_{i,j}^y$ 为 y 方向的移动距离。

步骤 3. 将更新后的特征点取均值

$$\begin{cases} x = \frac{x_i + x_j + x_k}{3} \\ y = \frac{y_i + y_j + y_k}{3} \end{cases} \quad (15)$$

与现有算法使用单一特征提取模型不同, 本文训练多个关键点特征提取模型, 这样在检测过程中可以利用多模型同时检测得到多个结果, 最终进行加权融合来增加整体结果的鲁棒性, 同时在特征提取环节本系统更加关注特征获取过程的稳定性。级联形式的特征提取不仅能够获得最终的检测结果, 还能获得检测过程中关键点信息定位的形状增量集, 利用这些形状增量集的变化信息可以获得对应模型在定位时的稳定性。当回归器足够鲁棒且能够平滑地收敛时, 其将初始点 S^0 移动到 S^T 所花费增量之和也相对较小。因此利用这些关键点增量集的绝对值之和来获得对应特征提取模型的稳定性和鲁棒性, 可以为后续的加权融合过程提供一个权衡标准, 并且使用以上方法可以为后续的位姿解算提供一个优质的特征映射关系。

2.6 远距离优化的 ICP 位姿解算

为保证质量较高的关键点 (指的是融合时位移较少即离散度较低的关键点) 在位姿优化过程中发挥较大的作用, 同时减少深度信息带来的影响, 本节利用第 2.4 节得到的关键点解集 β 计算每个关键点的离散度信息, 为每个关键点赋权值, 以此来增大精度较高关键点的影响。

步骤 1. 通过多特征提取模型得到每个关键点的定位解集合

$$\beta = \{\beta_1, \dots, \beta_f, \dots, \beta_F\} \quad (16)$$

其中, β_f 表示第 f 种特征提取模型得到的关键点定位解集。

步骤 2. 为衡量每个关键点解集的质量, 解集间元素距离求和为

$$D_l = \sum_{f=1}^{F-1} |\beta_f^l - \beta_{f+1}^l|, \quad l = 1, 2, \dots, L \quad (17)$$

其中, β_f^l 表示第 f 种特征提取模型的第 l 个关键点像素坐标, $|\beta_f^l - \beta_{f+1}^l|$ 表示同一个关键点在不同模型下的距离, 其和 D_l 称为关键点离散度。

步骤 3. 归一化处理得到权值

$$D_l = \frac{\sum_{f=1}^{F-1} |\beta_f^l - \beta_{f+1}^l|}{\sum_{f=1}^{F-1} \sum_{l=1}^L |\beta_f^l - \beta_{f+1}^l|}, \quad l = 1, 2, \dots, L \quad (18)$$

步骤 4. EPnP 求解关键点的相机坐标系坐标

利用 EPnP 算法得到目标每个 3D 关键点对应的齐次重心坐标, 然后利用齐次重心坐标在世界坐标系和相机坐标系下是相同的这一特性, 求解出控制点相机坐标系坐标, 最终利用控制点相机坐标系坐标和齐次重心坐标求解出每个 3D 关键点对应的相机坐标系坐标 P_l^c .

步骤 5. 加权 ICP 求解位姿

$$[\mathbf{R}|\mathbf{T}] = \arg \min_{[\mathbf{R}|\mathbf{T}]} \frac{1}{2D_l} \sum_{l=1}^L \|P_l^w - \frac{1}{s} \mathbf{K}[\mathbf{R}|\mathbf{T}]P_l^c\|_2^2 \quad (19)$$

其中, P_l^w 表示第 l 个关键点已知的无人机模型世界坐标系坐标, $[\mathbf{R}|\mathbf{T}]$ 为求解的位姿矩阵, s 为尺度缩放因子, P_l^c 为由 EPnP 算法求解出来的第 l 个关键点的相机坐标系坐标.

不同于传统近距离位姿估计, 由于无人机与无人船间的距离较大, 因此各关键点的深度信息较大, 这会放大物方残差, 明显降低 ICP 位姿解算精度. 本文为基于级联回归得到的每个关键点解集设计了新的权值衡量标准, 利用解集中关键点的离散度作为加权 ICP 位姿解算的权值, 放大了精度较高的解集所对应的关键点对整体位姿的影响.

3 实验仿真

为验证本文所提算法的精确性和鲁棒性, 主要从以下 5 个方面进行仿真实验: 1) 实验一验证本文使用的非盲图像复原方法在辅助关键点信息定位上更加适合应对海面动平台随机影响下的环境; 2) 实验二验证本文建立的关键点类型选择算法能够在不同图像纹理损失下增加关键点信息定位的精度; 3) 实验三验证本文建立的多模型级联回归特征提取加权融合算法在面对动平台影响时能够表现出更高的鲁棒性; 4) 实验四验证本文建立的基于关键点离散度加权的 ICP 位姿优化算法在面对大深度目标时能够表现出较好的准确性; 5) 实验五为对比实验, 主要验证在海面干扰的环境下, 本文的算法能够提供精度高于现有算法的位姿结果.

现在常用的 6D 目标位姿估计真值的获取方式主要分为两种: 一是外部跟踪器获取, 二是安装标定板获取^[54]. 本文实验中无人机的位姿真值是在无人机底部安装一张特殊图案的位姿标定板, 通过标定板可以在图像中检测到标定板的特征点, 进而计算出标定板在相机坐标系下的位姿, 而在本文中无人机 3D 模型所在的坐标系与标定板所在的坐标系一致, 因此可以得到无人机相对于相机的位姿关系.

具体来说, 使用相机拍摄标定板的图像, 通过检测标定板的特征点并将得到的特征点与标定板三维坐标匹配, 最终可以计算出相机相对于标定板的位姿, 即无人机相对于相机的真实位姿. 其中为保证位姿真值的可靠性, 对可能存在误差的特征点定位环节加入人工二次定位, 以确保特征点精度. 使用的无人机和无人船如图 5 所示, 实验中所使用的视觉传感器为单目 RGB 视觉传感器. 实验场景 (如图 6 所示) 设置: 无人船作为拍摄平台以 0.05 s 的时间间隔来获取实验用的图像数据, 无人机作为被拍摄目标完成降落过程. 使用的位姿解算算法为 EPnP+ICP. 具体实验仿真场景和参数如下.

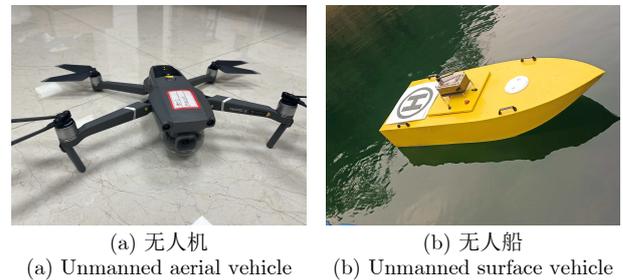


图 5 实验设备

Fig. 5 Experimental equipment



图 6 实验场景

Fig. 6 Experimental scene

3.1 实验平台及实验参数设置

实验设置的数据有无人船前进速度 v , 无人船倾角 ψ ; 无人机三维空间速度 v_x 、 v_y 、 v_z , 无人机世界坐标集合 P_l^w , 无人机世界重心矩阵 \mathbf{A} , 无人机关键点像素坐标集合 $\{(u, v)\}$, 无人机特征空间映射对 \mathbf{F} , 相机内参 \mathbf{K} , 相机畸变矩阵 \mathbf{J} , 控制点坐标集 C_l^w . 本部分实验中所使用的视觉传感器为 RGB 单目视觉传感器, 型号为 MSK-OV1829, 分辨率为 1080 p, 采样频率为 60 fps, 搭设位置为无人船动平台上.

无人船、无人机运动参数分别设置为:

$$\begin{cases} v = 2 \text{ m/s} \\ \psi = 0^\circ \sim 45^\circ \\ v_x = 0.5 \text{ m/s} \\ v_y = 1.5 \text{ m/s} \\ v_z = 0.5 \text{ m/s} \end{cases} \quad (20)$$

单目 RGB 视觉传感器内参设置:

$$\begin{cases} \mathbf{K} = \begin{bmatrix} 712 & 0 & 322 \\ 0 & 709 & 241 \\ 0 & 0 & 1 \end{bmatrix} \\ \mathbf{J} = \begin{bmatrix} -0.18500, & 1.25000, & -0.00298, \\ & & & 0.00549, & -3.87000 \end{bmatrix} \end{cases} \quad (21)$$

EPnP 算法控制点参数设置:

$$\begin{cases} \mathbf{C}_1^w = \frac{1}{L} \sum_{l=1}^L \mathbf{P}_l^w \\ \mathbf{C}_2^w = \mathbf{C}_1^w + \sqrt{\lambda_1} \boldsymbol{\nu}_1 \\ \mathbf{C}_3^w = \mathbf{C}_1^w + \sqrt{\lambda_2} \boldsymbol{\nu}_2 \\ \mathbf{C}_4^w = \mathbf{C}_1^w + \sqrt{\lambda_3} \boldsymbol{\nu}_3 \\ \mathbf{A} = \begin{bmatrix} (\mathbf{P}_1^w)^T - (\mathbf{C}_1^w)^T \\ \vdots \\ (\mathbf{P}_L^w)^T - (\mathbf{C}_1^w)^T \end{bmatrix} \end{cases} \quad (22)$$

其他参数设置:

$$\begin{cases} \mathbf{P}_l^w = \{\mathbf{P}_1^w, \mathbf{P}_2^w, \dots, \mathbf{P}_L^w\} \\ \{(u, v)\} = \{(u_1, v_1), (u_2, v_2), \dots, (u_L, v_L)\} \\ \mathbf{F} = [\mathbf{P}_l^w \approx (u_l, v_l)] \end{cases} \quad (23)$$

其中, λ_i 为 $\mathbf{A}\mathbf{A}^T$ 的特征值, $\boldsymbol{\nu}_i$ 为 $\mathbf{A}\mathbf{A}^T$ 的特征向量, \approx 表示 2D 关键点对应 3D 关键点, $l = 1, 2, \dots, L$.

3.2 实验一

实验场景介绍: 本部分实验主要验证加入非盲图像复原方法后能够降低运动模糊对目标检测精度和关键点检测精度的影响. 使用的数据为受到不同程度动平台抖动影响的图像, 其中因抖动而产生的运动模糊的模糊方向和距离在本文中默认已知. 图 7 为未经过非盲复原的图像和经过非盲复原后的图像分别进行关键点检测的结果. 图 8 为加入复原方法和未加入复原方法得到的目标检测结果. 使用的精度指标为 mAP X , 表示将重叠度阈值设为 X 时判定目标检测成功.

实验结果分析: 由实验一的图 7 和图 8 可知,

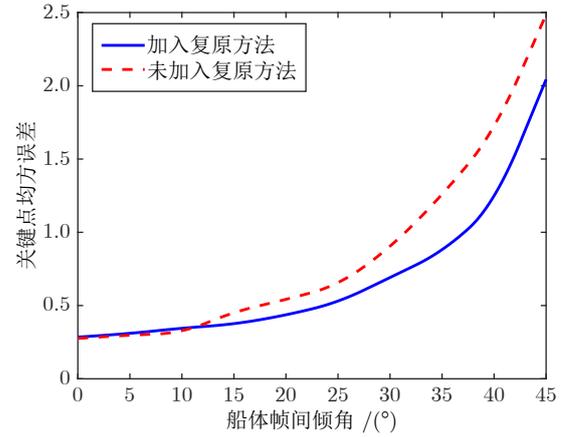


图 7 关键点检测实验

Fig.7 Keypoint detection experiment

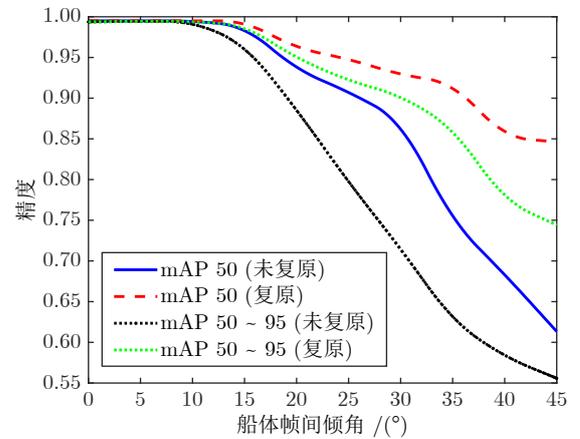


图 8 目标检测实验

Fig.8 Object detection experiment

在海面波浪较小 (即船体帧间倾角较小) 时, 传统方法和加入非盲图像复原方法的目标检测精度和关键点检测精度相差不大且都能表现出较为优秀的效果. 但当船体帧间倾角不断增大时, 产生的运动模糊程度逐渐增大, 目标检测的精度逐步开始下降, 在 $15^\circ \sim 30^\circ$ 的区间内下降逐步扩大, 已经无法保证为后续关键点检测提供一个优质的输入, 进而导致关键点定位精度大幅下降. 在加入图像复原方法后, 目标检测精度和关键点检测精度依旧会随运动模糊的增大而下降, 但下降程度有了明显降低. 因此可以得到以下结论: 非盲复原图像可以在船体倾角较大时提升目标检测精度和关键点检测精度.

3.3 实验二

实验场景介绍: 本部分实验主要验证加入关键点类型选择算法后能够应对因动平台抖动导致的不同程度的图像纹理损失, 达到自适应选择合适关键

点^[26]来应对当前场景的效果. 使用到的数据分为两部分, 一部分为使用角点检测的数据, 另一部分为使用纹理中心点检测的数据, 其中所有图像数据均受到不同程度的船体倾斜影响. 图 9 为分别使用角点和纹理中心点进行检测得到的关键点误差, 竖线为关键点选择阈值线 (16.4°), 竖线左侧使用纹理中心点检测, 右侧使用角点检测. 表 1 为各关键点类型展现最低检测损失时在 30° 内所占的比例.

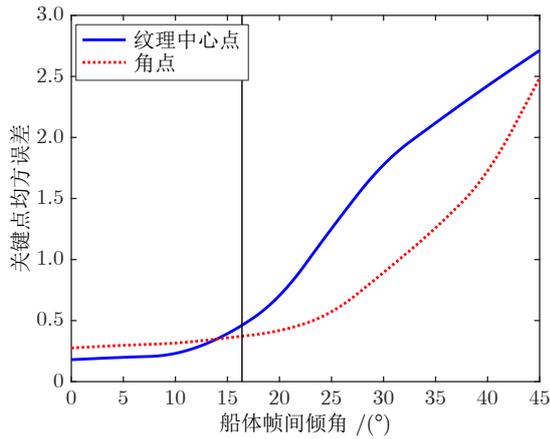


图 9 关键点类型选择实验

Fig.9 Keypoint type selection experiment

表 1 各关键点类型最低检测损失占比

Table 1 Proportion of minimum detection loss for each keypoint type

检测关键点类型	角点检测	纹理中心点检测	选择算法检测
最优占比	54.7%	45.3%	90.6%

实验结果分析: 由图 9 可以看出, 在图像纹理信息丰富时, 将待检测关键点设为纹理中心点得到的精度较高, 而当船体帧间倾角增大、图像受动平台影响导致纹理损失严重时, 将待检测关键点设为角点得到的精度较高, 无论是纹理中心点还是角点其都有自己的适用区间. 相比于传统的单一特征点检测, 本文使用的关键点选择算法充分发挥了不同关键点类型在不同船体倾角下的适用性, 在帧间倾角较小时使用纹理中心点作为待检测点, 相比于使用角点拥有更高的精度; 在帧间倾角较大时使用角点作为待检测点, 相比于使用纹理中心点拥有更高的精度. 同时, 根据表 1, 在 IMO 的 30° 安全倾角范围内, 使用选择算法无论在与任何其他一种关键点类型进行比较时都能有较大的最优占比^[55]. 因此可以得到以下结论: 在经过关键点类型选择后, 整个关键点特征的提取变得更加精确和稳定.

3.4 实验三

实验场景介绍: 本部分实验主要验证单一关键点检测模型易受到一些海面环境因素的影响, 将多种特征提取模型的结果进行融合可以有效减少这些海面因素造成的影响. 使用到的特征提取模型有三个, 使用的数据为经图像预处理后的无人机图像数据. 图 10 为使用三个不同的特征提取模型得到的结果损失, 图 11 为使用加权融合模型得到的结果损失, 表 2 给出了各模型的均值损失和方差比较结果.

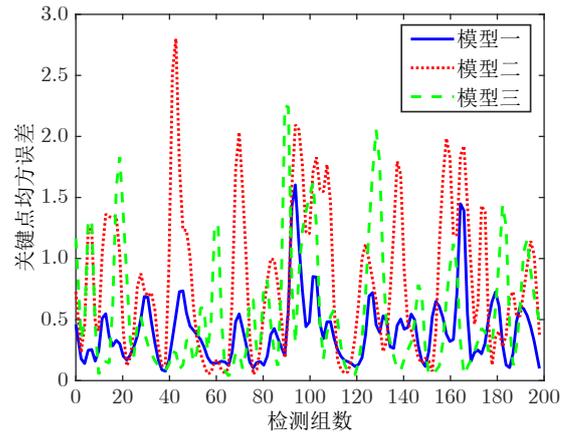


图 10 不同单一模型损失

Fig.10 Different single model loss

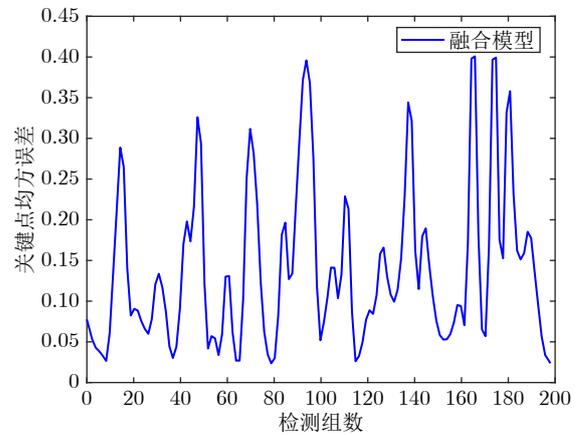


图 11 加权融合模型损失

Fig.11 Weighted fusion model loss

实验结果分析: 由图 10 可知, 传统算法中使用单一模型进行特征提取时会产生一定程度上的不稳定, 通常由模型损失拥有较大的方差所体现. 同时根据表 2 和图 11 可知, 在加入级联的多模型加权融合算法后, 相比于传统单一模型, 融合后的模型依旧能够保持与传统算法同一级别的精度, 并拥有更低的模型损失方差, 表现出了更加强大的稳定性. 因此可以得到以下结论: 在经过加权融合算法对多

表 2 各模型均值损失和方差
Table 2 Mean loss and variance of different models

模型名称	均值损失	方差
模型一	0.4001	0.1184
模型二	0.8213	0.5607
模型三	0.5640	0.4226
加权融合模型	0.1401	0.0157

种特征提取模型结果的融合后,可以在保证精度不丢失的基础上大大提升整体算法的稳定性。

3.5 实验四

实验场景介绍: 本部分实验主要验证利用关键点解集离散度作为权值的 ICP 算法能够在面对大深度环境时表现出较强的优化性能. 使用的数据包括已知的无人机 3D 点分布情况、无人船拍摄到的无人机降落阶段的图像、由融合算法提取到的 2D 关键点坐标和无人机自身的真实位姿数据. 图 12 和图 13 为在不同距离利用加权 ICP 位姿优化的位姿精度, 表 3 给出了不同距离中两种 ICP 算法的均值精度和方差.

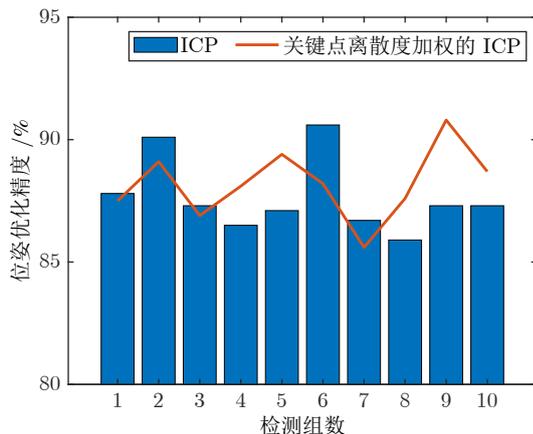


图 12 ICP 优化 (近距离)
Fig.12 ICP optimization (close range)

实验结果分析: 由图 12 和图 13 可知, 在面对较小的深度环境时, 传统的 ICP 算法和利用关键点离散度作为权值的 ICP 算法在位姿估计精度上展现出了相同的水准. 但是在面对较大的深度环境时, 传统 ICP 算法的位姿估计精度和稳定性略低于利用关键点离散度作为权值的 ICP 算法. 由表 3 可知, 在面对大深度环境时, 两种位姿解算精度都会有所下降, 但是本文的 ICP 算法依旧能够保证较低的方差, 能够保证其稳定性. 因此可以得出以下结论: 对不同关键点赋予离散度权值可以在面对较远目标时产生更稳定的位姿解算结果.

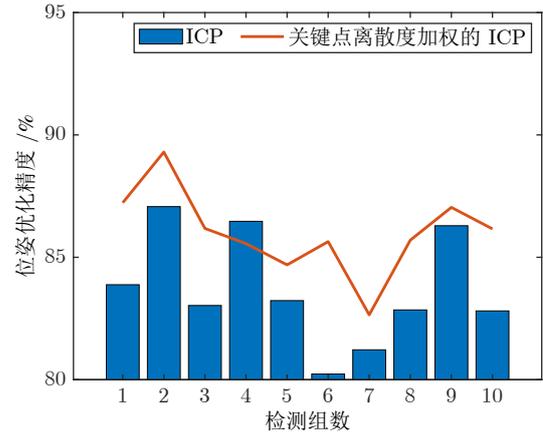


图 13 ICP 优化 (远距离)
Fig.13 ICP optimization (long range)

表 3 不同距离中两种 ICP 算法的均值精度和方差
Table 3 Mean accuracy and variance of two ICP algorithms in different distances

距离	算法	均值精度	方差
近距离 (5 m)	ICP	87.34%	2.30
近距离 (5 m)	关键点离散度加权的 ICP	88.08%	2.06
远距离 (15 m)	ICP	83.71%	5.13
远距离 (15 m)	关键点离散度加权的 ICP	86.05%	2.98

3.6 实验五

实验场景介绍: 本部分实验主要验证在受随机扰动的无人船上对无人机进行位姿估计时, 本文算法相较于现有位姿估计算法和传统位姿估计算法能够提供精度更高的位姿解算结果. 使用到的数据有已知的无人机 3D 点分布情况、无人船拍摄到的无人机降落阶段的图像、由融合算法提取到的 2D 关键点坐标和无人机自身的真实位姿数据. 图 14 为本文算法与其他算法的对比实验图, 使用到的位姿估计评价指标为关键点位姿重构误差, 表 4 给出了不同船体帧间倾角下各算法的位姿估计精度.

实验结果分析: 由图 14 和表 4 可知, 在面对较小幅度 ($0^\circ \sim 15^\circ$) 的海面随机扰动时, 本文算法展现出了与现有位姿估计算法同等水平的估计误差; 在 $15^\circ \sim 30^\circ$ 的船体帧间倾角区间内, 本文算法的误差增长幅度远小于现有的其他算法, 并能够保持在一个较低的水平; 同时在 $30^\circ \sim 45^\circ$ 的船体帧间倾角区间内, 现有算法因特征空间中关键点数量逐步减少出现了位姿精度大幅下降的现象, 而本文利用多模型融合的方式增强了特征空间的鲁棒性, 使得精度下降的趋势得到缓解. 因此可以得到以下结论: 本文的位姿估计算法在面对不同程度的海面随机扰动时能够提供一个更加精确的位姿估计结果.

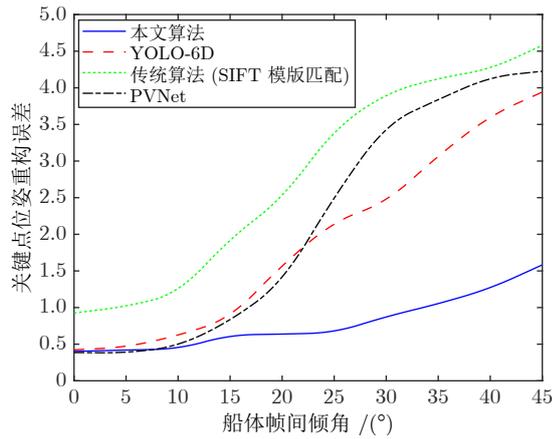


图 14 各算法位姿重构误差

Fig. 14 Pose reconstruction errors of various algorithms

3.7 实验总结

经过以上所有实验发现: 在动平台上应用原有基于关键点的位姿估计技术还是存在一定缺陷的. 本文在原有的框架中加入了图像复原方法、关键点选择算法、多模型级联回归特征提取加权融合算法、特征空间筛选机制和基于关键点离散度加权 ICP 位姿优化算法, 并对各阶段算法的有效性进行了单独实验. 在关键点类型选择环节, 加入关键点选择算法可以使得最优关键点类型在无人船全角度抖动区间占比达到 90.6%. 在特征空间建立阶段, 多模型级联回归特征提取加权融合算法使得关键点均方误差均值损失平均减少了 0.46; 在位姿解算阶段, 基于关键点离散度加权 ICP 位姿优化算法使得位姿解算的稳定性提高了 26%. 最终将以上算法组合在一起进行对比实验, 得到的无人机位姿精度在轻度运动模糊区间 ($0^\circ \sim 15^\circ$) 能够保持在 89% 以上; 在中度运动模糊区间 ($15^\circ \sim 30^\circ$) 能够保持在 81% 以上, 相较于其他算法平均提高了 34.66%; 同时在重度运动模糊区间 ($30^\circ \sim 45^\circ$) 能够保持在 72% 以上. 最终能够在 IMO 要求的船体帧间倾角为 $0^\circ \sim 30^\circ$ 的区间中平均提高了 18.99%.

在长时稳定性方面, 本文系统运行时, 每个环节都是对独立的一张图像做处理, 不同图像处理的

结果并不会相互影响; 每一张图像建立一个独立的特征空间, 不同空间之间互不干扰, 因此上一帧的误差并不会累积到下一帧; 在解算位姿时使用的也是属于该帧图像的特征空间, 因此也不会产生累积误差. 在抗干扰方面, 系统通过非盲图像复原提高图像的细节, 去除图像中的噪声, 增强输入图像的质量; 通过加权的级联特征提取增加特征空间中特征映射对的数量, 又利用定位过程中的形状增量抑制定位偏差模型的融合权重, 以增强 2D 相机投影关键点定位的精度和鲁棒性; 最终利用关键点解集离散度作为权值, 进一步抑制定位存在争议的关键点的位姿解算权重. 综上, 本文系统通过各环节增强特征空间的精度, 并抑制特征空间中质量较差关键点 (多模型定位结果存在较大偏差) 的权重, 在各环节中产生的误差仅针对该帧图像, 并不会产生对后续帧影响的累积误差, 因此本文系统具有长时稳定性和抗干扰性.

4 结束语

针对无人机降落阶段, 因无人船受海面波浪影响对图像产生运动模糊, 导致获取无人机相对位姿精度低且鲁棒性差的问题. 本文首先考虑到海面情况与陆地上的不同, 选用了针对海面随机波浪影响的图像复原方法来增强输入数据的质量. 随后考虑到海面波浪会对图像产生随机的不同程度运动模糊, 设计了关键点类型选择算法, 增强了每次提取关键点的适用性. 然后考虑到单一特征提取模型在面对动平台抖动时, 其特征空间可能被破坏, 运用级联的方式进行特征提取, 利用每次回归的形状增量作为加权融合的权值参考, 以此来提升特征空间的鲁棒性. 最后考虑到场景深度信息较大会放大位姿解算的误差, 设计了一种利用关键点解集离散度作为权值的 ICP 优化算法来增大优化程度, 使得位姿的精度得到有效提高. 通过理论分析和仿真比较表明, 所提出的方法在海面波浪的仿真环境下取得了先进的性能. 特别地, 基于本文研究的无人机降落场景, 其他无人船与无人机协同完成任务的场景将是未来的研究方向.

表 4 不同船体帧间倾角下的各算法位姿估计精度 (%)

Table 4 Pose estimation accuracy of various algorithms under different ship hull inter-frame angles (%)

帧间倾角	0°	5°	10°	15°	20°	25°	30°	35°	40°	45°
本文算法	94.2	93.0	90.4	89.1	88.4	85.3	81.2	82.3	78.3	72.4
PVNet	93.3	94.7	91.3	82.6	71.3	56.7	45.3	44.0	41.3	36.0
YOLO-6D	92.0	92.7	88.7	81.3	72.7	63.3	55.3	50.7	46.7	39.3
传统算法 (SIFT 模版匹配)	83.4	83.7	81.0	73.8	64.8	54.0	45.3	42.6	39.6	33.9

References

- 1 Shen Lin-Cheng, Kong Wei-Wei, Niu Yi-Feng. Ground-and ship-based guidance approaches for autonomous landing of UAV. *Journal of Beijing University of Aeronautics and Astronautics*, 2021, **47**(2): 187–196
(沈林成, 孔维玮, 牛轶峰. 无人机自主降落地基/舰基引导方法综述. *北京航空航天大学学报*, 2021, **47**(2): 187–196)
- 2 Zhen Zi-Yang. Research development in autonomous carrier-landing/ship-recovery guidance and control of unmanned aerial vehicles. *Acta Automatica Sinica*, 2019, **45**(4): 669–681
(甄子洋. 舰载无人机自主着舰回收制导与控制研究进展. *自动化学报*, 2019, **45**(4): 669–681)
- 3 Lowe D G. Object recognition from local scale-invariant features. In: Proceedings of the 7th IEEE International Conference on Computer Vision. Kerkyra, Greece: IEEE, 1999. 1150–1157
- 4 Bay H, Tuytelaars T, Gool L V. SURF: Speeded up robust features. In: Proceedings of the 9th European Conference on Computer Vision. Graz, Austria: Springer, 2006. 404–417
- 5 Rublee E, Rabaud V, Konolige K, Bradski G. ORB: An efficient alternative to SIFT or SURF. In: Proceedings of the International Conference on Computer Vision. Barcelona, Spain: IEEE, 2011. 2564–2571
- 6 Lepetit V, Moreno-Noguer F, Fua P. EPnP: An accurate $O(n)$ solution to the PnP problem. *International Journal of Computer Vision*, 2009, **81**(2): 155–166
- 7 Li Hang-Yu, Zhang Zhi-Long, Li Chu-Wei, Wu Han. Overview of vision-based target pose estimation. In: Proceedings of the 7th High Resolution Earth Observation Academic Annual Conference. Changsha, China: China High Resolution Earth Observation Conference, 2020. 635–648
(李航宇, 张志龙, 李楚为, 吴晗. 基于视觉的目标位姿估计综述. 第七届高分辨率对地观测学术年会论文集. 长沙, 中国: 高分辨率对地观测学术联盟, 2020. 635–648)
- 8 Wang Jing, Jin Yu-Chu, Guo Ping, Hu Shao-Yi. Survey of camera pose estimation methods based on deep learning. *Computer Engineering and Application*, 2023, **59**(7): 1–14
(王静, 金玉楚, 锅苹, 胡少毅. 基于深度学习的相机位姿估计方法综述. *计算机工程与应用*, 2023, **59**(7): 1–14)
- 9 Guo Nan, Li Jing-Yuan, Ren Xi. Survey of rigid object pose estimation algorithms based on deep learning. *Computer Science*, 2023, **50**(2): 178–189
(郭楠, 李婧源, 任曦. 基于深度学习的刚体位姿估计方法综述. *计算机科学*, 2023, **50**(2): 178–189)
- 10 Xiang Y, Schmidt T, Narayanan V, Fox D. PoseCNN: A convolutional neural network for 6D object pose estimation in cluttered scenes. arXiv preprint arXiv: 1711.00199, 2018.
- 11 Do T T, Cai M, Pham T, Reid I. Deep-6DPose: Recovering 6D object pose from a single RGB image. arXiv preprint arXiv: 1802.10367, 2018.
- 12 Liu J, He S. 6D object pose estimation without PnP. arXiv preprint arXiv: 1902.01728, 2019.
- 13 Wang Xiao-Qi, Sang Han-Bo. 6D pose estimation based on category-level prior reconstruction. *Changjiang Information and Communications*, 2023, **36**(1): 15–17
(王晓琦, 桑晗博. 基于类别级先验重建的六维位姿估计方法. *长江信息通信*, 2023, **36**(1): 15–17)
- 14 Sang Han-Bo, Lin Wei-Rao, Ye Long. 3D deep implicit function for category-level object 6D pose estimation. *Journal of Communication University of China (Natural Science Edition)*, 2022, **29**(4): 50–56
(桑晗博, 林巍颢, 叶龙. 基于深度三维模型表征的类别级六维位姿估计. *中国传媒大学学报(自然科学版)*, 2022, **29**(4): 50–56)
- 15 Peng S D, Liu Y, Huang Q X, Bao H J, Zhou X W. PVNet: Pixel-wise voting network for 6DoF pose estimation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA: IEEE, 2019. 4556–4565
- 16 Rad M, Lepetit V. BBS: A scalable, accurate, robust to partial occlusion method for predicting the 3D poses of challenging objects without using depth. In: Proceedings of the IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE, 2017. 3848–3856
- 17 Tekin B, Sinha S N, Fua P. Real-time seamless single shot 6D object pose prediction. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018. 292–301
- 18 Kehl W, Manhardt F, Tombari F, Ilic S, Navab N. Ssd-6D: Making RGB-based 3D detection and 6D pose estimation great again. In: Proceedings of the IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE, 2017. 1530–1538
- 19 Liu J, Sun W, Liu C P, Zhang X, Fan S M, Zhang L X. A novel 6D pose estimation method for indoor objects based on monocular regression depth. In: Proceedings of the China Automation Congress (CAC). Beijing, China: IEEE, 2021. 4168–4172
- 20 Xu S W, Lin F, Lu Y J. Pose estimation method for autonomous landing of quadrotor unmanned aerial vehicle. In: Proceedings of the IEEE 6th Information Technology and Mechatronics Engineering Conference (ITOEC). Chongqing, China: IEEE, 2022. 1246–1250
- 21 Na S, Chen W C, Stork W, Tang S. UAV flight control algorithm based on detection and pose estimation of the mounting position for weather station on transmission tower using depth camera. In: Proceedings of the IEEE 17th International Conference on Control and Automation (ICCA). Naples, Italy: IEEE, 2022. 522–528
- 22 Wang Shuo, Zhu Hai-Jiang, Li He-Ping, Wu Yi-Hong. Relative pose calibration between a range sensor and a camera using two coplanar circles. *Acta Automatica Sinica*, 2020, **46**(6): 1154–1165
(王硕, 祝海江, 李和平, 吴毅红. 基于共面圆的距离传感器与相机的相对位姿标定. *自动化学报*, 2020, **46**(6): 1154–1165)
- 23 Wang Jia-Bao, Zhang Shi-Rong, Zhou Qing-Ya. Vision based real-time 3D displacement measurement using weighted iterative EPnP algorithm. *Chinese Journal of Scientific Instrument*, 2020, **41**(2): 166–175
(汪佳宝, 张世荣, 周清雅. 基于视觉 EPnP 加权迭代算法的三维位移实时测量. *仪器仪表学报*, 2020, **41**(2): 166–175)
- 24 Qiu Xiang, Wang Guo-Shun, Zhao Yang-Yang, Teng You, Yu Li. Multi-pillbox attitude estimation based on YOLOv3 and EPnP algorithm. *Computer Measurement and Control*, 2021, **29**(2): 125–131
(仇翔, 王国顺, 赵扬扬, 滕游, 俞立. 基于 YOLOv3 和 EPnP 算法的多药盒姿态估计. *计算机测量与控制*, 2021, **29**(2): 125–131)
- 25 Ma Tian, Meng Xin, Mou Qi, Li Zhan-Li, He Zhi-Qiang. 6D object pose estimation algorithm based on feature fusion. *Computer Engineering and Design*, 2023, **44**(2): 563–569
(马天, 蒙鑫, 牟琦, 李占利, 何志强. 基于特征融合的 6D 目标位姿估计算法. *计算机工程与设计*, 2023, **44**(2): 563–569)
- 26 Sun Qin-Xuan, Yuan Jing, Zhang Xue-Bo, Gao Yuan-Xi. PLVO: Plane-line-based RGB-D visual odometry. *Acta Automatica Sinica*, 2023, **49**(10): 2060–2072
(孙沁璇, 苑晶, 张雪波, 高远兮. PLVO: 基于平面和直线融合的 RGB-D 视觉里程计. *自动化学报*, 2023, **49**(10): 2060–2072)
- 27 Besl P J, McKay N D. A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis Machine Intelligence*, 1992, **14**(2): 239–256
- 28 Ding Wen-Dong, Xu De, Liu Xi-Long, Zhang Da-Peng, Chen Tian. Review on visual odometry for mobile robots. *Acta Automatica Sinica*, 2018, **44**(3): 385–400
(丁文东, 徐德, 刘希龙, 张大朋, 陈天. 移动机器人视觉里程计综述. *自动化学报*, 2018, **44**(3): 385–400)
- 29 Zhang Wei-Dong, Liu Xiao-Cheng, Han Peng. Progress and challenges of overwater unmanned systems. *Acta Automatica Sinica*, 2020, **46**(5): 847–857
(张卫东, 刘笑成, 韩鹏. 水上无人系统研究进展及其面临的挑战. *自动化学报*, 2020, **46**(5): 847–857)
- 30 Sun Y, Wang X G, Tang X O. Deep convolutional network cascade for facial point detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Portland, USA: IEEE, 2013. 3476–3483
- 31 Dollar P, Welinder P, Perona P. Cascaded pose regression. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Francisco, USA: IEEE, 2010. 1078–1085
- 32 Cao X D, Wei Y C, Wen F, Sun J. Face alignment by explicit shape regression. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Providence, USA: IEEE, 2012. 2887–2894

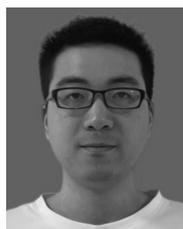
- 33 Ren S Q, Cao X D, Wei Y C, Sun J. Face alignment at 3000 fps via regressing local binary features. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA: IEEE, 2014. 1685–1692
- 34 Kowalski M, Narumiec J, Trzcinski T. Deep alignment network: A convolutional neural network for robust face alignment. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Honolulu, USA: IEEE, 2017. 2034–2043
- 35 Fan Han-Qi, Wu Jin-He. Stereo visual odometry based on Laplace distribution. *Acta Automatica Sinica*, 2022, **48**(3): 865–876
(范涵奇, 吴锦河. 基于拉普拉斯分布的双目视觉里程计. 自动化学报, 2022, **48**(3): 865–876)
- 36 Ferreira T, Bernardino A, Damas B. 6D UAV pose estimation for ship landing guidance. In: Proceedings of the OCEANS 2021: San Diego – Porto. San Diego, USA: IEEE, 2021. 1–10
- 37 Chatzikalymnios E, Moustakas K. Autonomous vision-based landing of UAV's on unstructured terrains. In: Proceedings of the IEEE International Conference on Autonomous Systems (ICAS). Montreal, Canada: IEEE, 2021. 1–5
- 38 Orsulic J, Milijas R, Batinovic A, Markovic L, Ivanovic A, Bogdan S. Flying with cartographer: Adapting the cartographer 3D graph SLAM stack for UAV navigation. In: Proceedings of the Aerial Robotic Systems Physically Interacting With the Environment (AIRPHARO). Biograd na Moru, Croatia: IEEE, 2021. 1–7
- 39 Nie Wei, Wen Huai-Zhi, Xie Liang-Bo, Yang Xiao-Long, Zhou Mu. Indoor localization of UAV using monocular vision. *Journal of Electronics and Information Technology*, 2022, **44**(3): 906–914
(聂伟, 文怀志, 谢良波, 杨小龙, 周牧. 一种基于单目视觉的无人机室内定位方法. 电子与信息学报, 2022, **44**(3): 906–914)
- 40 Lippiello V, Cacace J. Robust visual localization of a UAV over a pipe-rack based on the lie group SE(3). *IEEE Robotics and Automation Letters*, 2022, **7**(1): 295–302
- 41 Gurgu M M, Queralta J P, Westerlund T. Vision-based GNSS-free localization for UAVs in the wild. In: Proceedings of the 7th International Conference on Mechanical Engineering and Robotics Research (ICMERR). Krakow, Poland: IEEE, 2022. 7–12
- 42 Huang Q Z, Feng P Y. Research on vision positioning algorithm of UAV landing based on vanishing point. In: Proceedings of the 3rd International Conference on Intelligent Control, Measurement and Signal Processing and Intelligent Oil Field (ICMSP). Xi'an, China: IEEE, 2021. 68–71
- 43 Wang Y X, Wang H L, Liu B L, Liu Y H, Wu J F, Lu Z Y. A visual navigation framework for the aerial recovery of UAVs. *IEEE Transactions on Instrumentation and Measurement*, 2021, **70**: Article No. 5019713
- 44 Fan Y S, Huang J S, Jia T, Bai C C. A visual marker detection and position method for autonomous aerial refueling of UAVs. In: Proceedings of the IEEE International Conference on Unmanned Systems (ICUS). Beijing, China: IEEE, 2021. 1006–1011
- 45 Moura A, Antunes J, Dias A, Martins A, Almeida J. Graph-SLAM approach for indoor UAV localization in warehouse logistics applications. In: Proceedings of the IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC). Santa Maria da Feira, Portugal: IEEE, 2021. 4–11
- 46 He Yong, Li Zi-Hao, Gao Zheng-Tao. Autonomous and precise landing of UAVs based on visual navigation. *Electronics Optics and Control*, 2023, **30**(4): 88–93
(贺勇, 李子豪, 高正涛. 基于视觉导航的无人机自主精准降落. 电光与控制, 2023, **30**(4): 88–93)
- 47 Lu Jin-Yan, Xu De, Tan Zheng-Ke, Wang Peng, Ren Chao. An automatic alignment strategy of large diameter components with a multi-sensor system. *Acta Automatica Sinica*, 2015, **41**(10): 1711–1722
(卢金燕, 徐德, 覃政科, 王鹏, 任超. 基于多传感器的大口径器件自动对准策略. 自动化学报, 2015, **41**(10): 1711–1722)
- 48 Ou N, Wang J Z, Liu S F, Li J H. Towards pose estimation for large UAV in close range. In: Proceedings of the 37th Youth Academic Annual Conference of Chinese Association of Automation (YAC). Beijing, China: IEEE, 2022. 56–61
- 49 Deng J Y, Qu W D, Fang S Q. A high accuracy and recall rate 6D pose estimation method using point pair features for bin-picking. In: Proceedings of the 34th Chinese Control and Decision Conference (CCDC). Hefei, China: IEEE, 2022. 6056–6061
- 50 Wang H S, Situ H J, Zhuang C G. 6D pose estimation for bin-picking based on improved mask R-CNN and DenseFusion. In: Proceedings of the 26th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA). Vasteras, Sweden: IEEE, 2021. 1–7
- 51 Zhang Wen-Zheng, Wu Chang-Yue, Zhao Wen, Man Wei-Dong, Liu Ming-Yue. Research on UAV image deblurring method based on confrontation network and Wiener filtering. *Radio Engineering*, 2024, **54**(3): 607–614
(张文政, 吴长悦, 赵文, 满卫东, 刘明月. 融合对抗网络和维纳滤波的无人机图像去模糊方法研究. 无线电工程, 2024, **54**(3): 607–614)
- 52 Project webpage [Online], available: <https://github.com/ultra-lytics/yolov5>, March 5, 2024
- 53 Xiong X H, Torre F D L. Supervised descent method and its applications to face alignment. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Portland, USA: IEEE, 2013. 532–539
- 54 Zhang Z Y. Flexible camera calibration by viewing a plane from unknown orientations. In: Proceedings of the 7th IEEE International Conference on Computer Vision. Kerkyra, Greece: IEEE, 1999. 666–673
- 55 Gao Y, He J F, Zhang T Z, Zhang Z, Zhang Y D. Dynamic key-point detection network for image matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, **45**(12): 14404–14419



葛泉波 南京信息工程大学教授. 主要研究方向为状态估计与信息融合, 目标检测识别与跟踪和自主无人系统与试验测试. 本文通信作者.

E-mail: QuanboGe@163.com

(**GE Quan-Bo** Professor at Nanjing University of Information Science and Technology. His research interest covers state estimation and information fusion, object detection recognition and tracking, and autonomous unmanned systems and experimental testing. Corresponding author of this paper.)



李凯 南京信息工程大学硕士研究生. 主要研究方向为无人机视觉位姿估计, 无人机视觉目标跟踪.

E-mail: 20211249528@nuist.edu.cn

(**LI Kai** Master student at Nanjing University of Information Science and Technology. His research interest covers unmanned aerial vehicle visual pose estimation and unmanned aerial vehicle visual object tracking.)



张兴国 中国飞行试验研究院高级工程师. 主要研究方向为飞机试验技术, 智能试验技术.

E-mail: zhangxg011@avic.com

(**ZHANG Xing-Guo** Senior engineer at China Flight Test Research Institute. His research interest covers aircraft test technology and intelligent test technology.)