

控制与决策

Control and Decision

基于单双目融合的AUV坐落式回收光视觉引导算法

祝志坤, 卢丙举, 李一辰, 王凯, 于文彬

引用本文:

祝志坤, 卢丙举, 李一辰, 等. 基于单双目融合的AUV坐落式回收光视觉引导算法[J]. *控制与决策*, 2025, 40(1): 28-37.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2024.0346>

您可能感兴趣的其他文章

Articles you may be interested in

[基于动态蚁群劳动分工模型的多AUV任务分配方法](#)

A multi-AUV dynamic task allocation method based on antecolony labor division model

控制与决策. 2021, 36(8): 1911-1919 <https://doi.org/10.13195/j.kzyjc.2019.1312>

[基于多层级特征的机械臂单阶段抓取位姿检测](#)

Single-stage grasp pose detection of manipulator based on multi-level features

控制与决策. 2021, 36(8): 1815-1824 <https://doi.org/10.13195/j.kzyjc.2019.1840>

[USV&AUV水下目标协同搜索与打击航迹规划](#)

Coordinated path planning of USV&AUV for an underwater target

控制与决策. 2021, 36(4): 825-834 <https://doi.org/10.13195/j.kzyjc.2019.1038>

[基于多块信息提取的AUV资源勘查系统故障检测](#)

Fault detection of AUV resource exploration system based on multi-block information extraction

控制与决策. 2021, 36(4): 790-800 <https://doi.org/10.13195/j.kzyjc.2019.0732>

[FMM与改进GBNN模型相结合的多AUV实时围捕算法](#)

Multi-AUV real-time hunting control based on FMM and improved GBNN model

控制与决策. 2020, 35(12): 2845-2854 <https://doi.org/10.13195/j.kzyjc.2019.0393>

基于单双目融合的AUV坐落式回收光视觉引导算法

祝志坤¹, 卢丙举², 李一辰^{1†}, 王凯², 于文彬¹

(1. 上海交通大学 电子信息与电气工程学院, 上海 200240; 2. 河南省水下智能装备重点实验室, 郑州 450000)

摘要: 自主水下航行器(AUV)在任务进行过程中或完成后需通过自主回收实现能源补充与数据下载等操作, 是否能进行高效、精准地回收引导决定了AUV的回收效率, 成为其能否广泛应用的关键. 针对AUV坐落式回收过程中的近距离光学引导定位问题, 提出一种基于深度学习的单双目位姿测量算法. 首先, 面向恶劣的水下光学成像条件, 结合暗通道先验去雾和YOLO v9目标检测网络, 实现一种可适应不同水质、光照强度且鲁棒性强、可靠性高的引导光源提取算法. 同时, 针对回收过程中的特征匹配问题, 设计一种不依赖于AUV速度的全向特征匹配算法, 实现3D-2D特征匹配. 此外, 针对坐落式回收典型的多阶段引导特点, 分别基于PnP原理和SVD分解设计面向不同阶段的单、双目引导定位算法. 最后, 基于多次仿真和实物实验, 验证算法在精确位姿估计方面的可行性和有效性.

关键词: 水下自主航行器; 坐落式回收; 深度学习; 特征点提取; 位姿估计; 目标检测

中图分类号: TP274+.2

文献标志码: A

DOI: 10.13195/j.kzyjc.2024.0346

引用格式: 祝志坤, 卢丙举, 李一辰, 等. 基于单双目融合的AUV坐落式回收光视觉引导算法[J]. 控制与决策, 2025, 40(1): 28-37.

Light visual guidance algorithm for AUV situated recovery based on monocular and binocular fusion

ZHU Zhi-kun¹, LU Bing-ju², LI Yi-chen^{1†}, WANG Kai², YU Wen-bin¹

(1. School of Electronic Information and Electrical Engineering, Shanghai Jiaotong University, Shanghai 200240, China; 2. Key Laboratory of Underwater Intelligent Equipment in Henan Province, Zhengzhou 450000, China)

Abstract: The autonomous underwater vehicle (AUV) needs to perform operations such as energy replenishment and data download through autonomous recovery during or after a mission. The efficiency and accuracy of the recovery guidance determine the recovery efficiency of the AUV, which is crucial for its widespread application. To address short-range optical guidance and positioning in AUV recovery, this paper proposes a deep learning-based monocular and binocular pose measurement algorithm. Firstly, to address harsh underwater imaging conditions, a robust and reliable guided light source extraction algorithm is implemented, combining dark channel prior dehazing and the YOLO v9 target detection network, adaptable to different water qualities and light intensities. At the same time, in response to the feature matching problem in the recovery process, an omnidirectional feature matching algorithm that does not depend on the AUV speed to achieve 3D-2D feature matching is designed. In addition, in view of the typical multi-stage guidance characteristics of situated recovery, single and binocular guidance and positioning algorithm for different stages are designed based on the PnP principle and SVD decomposition. Finally, based on multiple simulations and physical experiments, the feasibility and effectiveness of the proposed algorithm in accurate pose estimation are verified.

Keywords: AUV; situated recovery; deep learning; feature point extraction; pose estimation; object detection

0 引言

自主式水下航行器(AUV)是海洋开发的重要工具, 可携带多种任务模块和传感器, 在作业效率、自主能力和环境适应性等诸多领域上具有优势, 可用于海底勘测、水下救援、水下目标搜索等关键任务^[1-2]. 通

常, AUV需通过自身携带能源完成水下作业. 在任务进行过程中或完成后, 需通过自主回收实现能源补充、数据下载、设备检修等操作^[3]. 因此, AUV能否进行精准、高效回收, 是决定其能否被广泛应用的关键技术.

收稿日期: 2024-03-31; 录用日期: 2024-07-06.

基金项目: 河南省水下智能装备重点实验室基金项目(KL02B2301); 国家自然科学基金项目(62203299, 62373246); 中央引导资金项目(Z20221343002).

†通讯作者. E-mail: liyichensjtu@sjtu.edu.cn.

AUV主要回收方式可分为两类,母船水面起吊回收和 underwater 自主对接回收。母船水面起吊回收通过 AUV 自主靠近母船,利用母船的起吊装置实现升降回收^[4]。这种方式易受风浪影响,隐蔽性差,难以适应 AUV 活动范围较大的应用场景。水下自主对接回收是指 AUV 依据水声、视觉信息等导引,运动至固定或移动的回收平台,实现类似空/天对接的回收方式。相较于母船水面起吊,此方式隐蔽性更好、受风浪等环境因素影响更小。因此,水下自主对接回收成为了 AUV 回收领域的研究热点^[5-6],具体可分为远距离和近距离两个阶段。远距离阶段,AUV 从远处靠近回收平台,对定位精度要求低,常使用声学引导;近距离阶段,AUV 逐渐调整自身位姿直至坐落至回收平台,对定位精度要求高,仅凭声学信息难以实现。光视觉信息具有适应性强、信息丰富、更新频率快等优点,可用于 AUV 回收近距离阶段的自身位姿测量。因此,本文主要研究对象为 AUV 回收近距离阶段基于光视觉的位姿测量。

近年来,基于光视觉信息的位姿测量策略广泛应用于 AUV 回收过程近距离阶段,并多次成功完成回收实验。Xu 等^[7]为提供基于立体视觉的导航策略,利用自适应选择算法进行光源检测,通过滤光片匹配精确地跟随光标以提高定位精度;Fan 等^[8]采用远距离超短基线定位系统(USBL)导航与近距离光视觉导航相结合的方式实现 AUV 的对接回收,并成功进行了对接实验,但其视觉检测算法抗干扰能力较弱,在复杂环境下容易误检或漏检;Ren 等^[9]提出了一种基于视觉信息引导的 AUV 自主重建对接环方法,利用蓝绿光信标和 YOLO v3 实现长距离引导,通过 Aruco 标记和核相关滤波(KCF)算法确保短距离高精度对接,有效兼顾了广泛制导范围和近距离制导精度;Li 等^[10]设计了基于单双目视觉的定位算法,并利用双层控制系统实现了 AUV 对接回收。然而,在 AUV 速度较快时该定位算法可能会变得不稳定甚至失效;Yahya 等^[11]提出了一种基于颜色信息的引导灯光源跟踪算法,但其仅能在 AUV 低速的情况下成功识别并跟踪;Zhong 等^[12]提出了一种自适应加权最大类间方差法(OTSU)特征提取方法,实现了引导灯光源的提取,进而实现了自身位姿的测量。综上,虽然基于光视觉的位姿测量策略可实现精准的位姿测量,但也存在检测算法抗干扰能力差和 AUV 高速情况下失效等问题。

目前,国内外大多数团队在 AUV 回收过程中的近距离阶段均利用水下图像信息,通过特定算法确定引导灯的位置、大小、类别等信息,即基于光学的目标

检测技术。根据使用方法不同,可将目标检测技术分为基于传统算法的水下目标检测算法和基于深度学习的水下目标检测算法。其中,传统算法包括 Harris 角点提取^[13]、Fast 角点提取^[14]等,其需在检测完成后加以筛选合并,得到与三维世界匹配的特征点,鲁棒性较差。深度学习类算法对参考数据集的特征进行学习和拟合,泛化能力较强,可更好地适用于多类场景。因此利用深度学习类方法进行特征提取与匹配成为近年研究的热点之一,被广泛应用于海洋场景下的目标检测、跟踪、定位等任务。因此,本文采用基于深度学习的目标检测算法,实现对水下引导灯光源的精准检测。

本文主要贡献在于面向 AUV 坐落式回收过程近距离阶段的自身位姿测量问题,提出一种基于单双目相机相融合的位姿测量框架。首先,针对引导灯光源检测问题,将暗通道先验去雾思路^[15]与 YOLO v9 目标检测网络^[16]相融合,实现具有高鲁棒性和准确性的引导光源检测算法;其次,针对 AUV 相对回收平台的位姿估计,设计基于匈牙利匹配的全向特征匹配方法,进而实现一种不依赖 AUV 速度的位姿测量算法;同时,基于光视觉提取信息,针对坐落式回收的多阶段特性,设计单双目融合策略与位姿估计方法;最后,为了验证位姿测量算法以及测试极限情况,本文在实物实验基础上构建 AUV 坐落式回收仿真场景,为算法验证与优化提供平台支撑。

1 问题描述

AUV 坐落式回收过程分为两个阶段:远距离声学导航阶段和近距离光学导引阶段。在第 1 阶段中,AUV 使用声学定位,逐渐从远距离工作区域接近回收平台。同时,启动第 2 阶段近距离光学定位导引。此时,AUV 通过自身携带的单双目相机,观察回收平台的引导光源,进而实现自身对于回收平台的相对位姿估计。本文研究的主要目标为光学定位导引阶段的自身位姿测量问题,如图 1 所示。

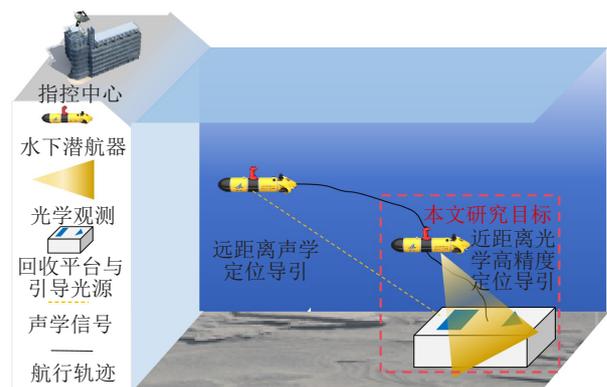


图 1 AUV 回收

由于光学导引阶段对定位精度和实时性要求高且情况更为复杂,本文提出一种基于深度学习的单双目位姿测量算法,实现在检测精度和抗干扰能力上的提升.在回收过程中,首先在相机拍摄到图像后,利用暗通道先验去雾算法^[15]和YOLO v9目标检测^[16]等,确定引导灯坐标.进而,利用特征匹配算法,确定图像特征点与三维世界中特征点的对应关系.在单目

位姿测量中,特征匹配完成后,利用PnP (perspective-n-point)算法得到AUV自身位姿.在双目位姿测量中,通过三角测量原理,确定匹配后特征点在双目相机坐标系下的三维坐标.最后,通过基于奇异值分解(SVD)的位姿求解算法得到AUV自身位姿.本文将核心算法设计为光视觉处理和基于光视觉的位姿估计两部分,图2为整体算法流程.

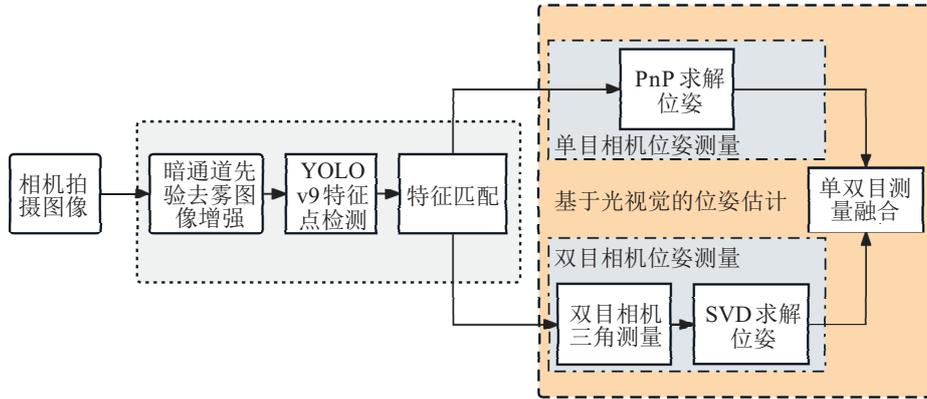


图2 算法流程

2 光视觉处理

光视觉处理的主要目的是获取实拍图像的二维特征点位置及其与现实世界的三维特征点的对应关系.二者为位姿估计模块的输入信息,是进行位姿估

计的必要参数.光视觉处理算法首先利用引导灯光源检测得到二维特征点位置,进而利用特征匹配得到其与三维特征点的对应关系.其整体流程如图3所示.

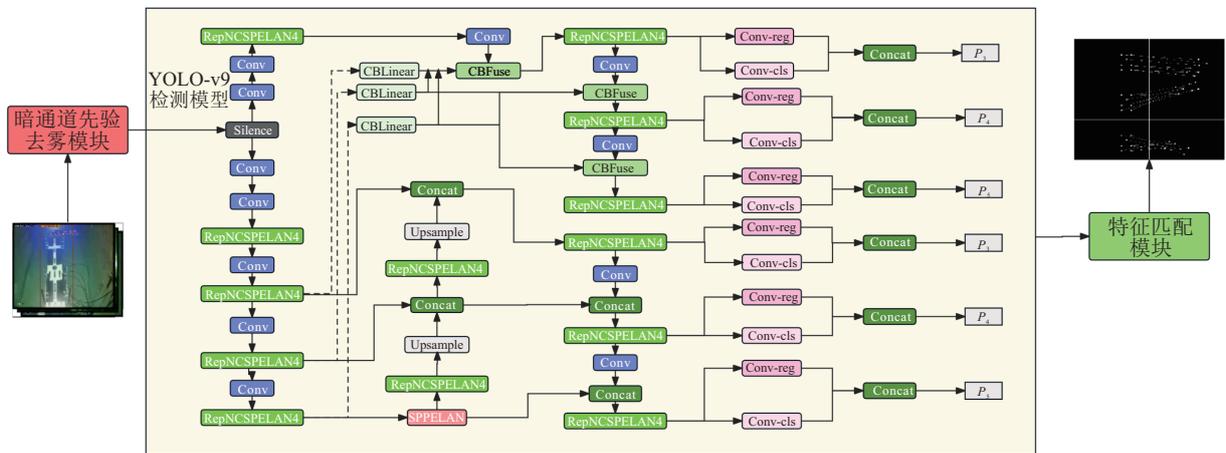


图3 引导灯光源检测结构

2.1 引导灯光源检测

在获取水下图像后,首先利用暗通道先验去雾进行图像增强.基于图像增强,利用YOLO v9目标检测网络以及阈值分割、连通域检测和质心提取等方法,确定引导灯的坐标.

2.1.1 暗通道先验去雾图像增强

由于水介质不均、光照衰减和散射等情况的存在,水下图像存在对比度损失、细节模糊和颜色偏

差^[17]等质量退化问题.暗通道去雾^[15]可以增强图像细节,提升图像质量,进而提高目标检测的准确性和稳定性.

暗通道是指自然图像中,一定大小区域内,像素的最小值所构成的图像.对于有雾图像 $I(x)$,其暗通道表示为

$$I_{\text{dark}}(x) = \min_{y \in \Omega(x)} (\min_{c \in \{r, g, b\}} I^c(y)). \quad (1)$$

其中: x 为图像中的点, c 为 r, g, b 三通道中的某一通

道, $\Omega(x)$ 为一定大小的矩形窗口。

暗通道先验原理^[15]指出, 无雾图像暗通道的值很小, 几乎趋近于0。暗通道去雾算法根据此原理, 利用暗通道计算大气光值和透射率, 进而去除雾气效果, 复原得到无雾图像, 即增强后的图像。

首先, 构建雾气退化模型^[15]为

$$I(x) = J(x)t(x) + A(1 - t(x)). \quad (2)$$

其中: x 为图像中的点, $I(x)$ 为有雾图像, $J(x)$ 为无雾图像, $t(x)$ 为透射率, A 为大气光值。

为获取大气光值, 从暗通道 $I_{\text{dark}}(x)$ 中选取像素值最大的前0.1%的像素点, 并找到其对应于有雾图像 $I(x)$ 的像素位置以及每个通道像素的平均值, 从而计算出每个通道 c 的大气光值 A^c 。

对于图像中的每一点 x 以及其局部窗口 $\Omega(x)$, 计算其透射率 $t(x)$ 为

$$t(x) = 1 - \omega \times \min_{y \in \Omega(x)} \left(\min_{c \in \{r, g, b\}} \frac{I^c(y)}{A^c} \right). \quad (3)$$

其中: ω 控制去雾程度, 0意味着不去雾, 1意味着全部去雾。

最后, 将各通道 c 的大气光值 A^c 和透射率 $t(x)$ 代入式(2), 计算复原得到增强后的图像, 即

$$J^c(x) = \frac{I^c(x) - A^c}{t(x)} + A^c. \quad (4)$$

2.1.2 基于YOLO v9的目标检测

本文目标检测采用该领域最新的前沿模型(SOTA)YOLO v9神经网络^[16]。在面向AUV回收的光视觉处理过程中, 其输入为暗通道先验去雾增强后的图像, 输出为图像中引导灯的边界框。通过对边界框内的图像进行阈值分割、连通域检测和质心提取, 得到引导灯特征点在图像中的坐标。

目前, AUV通常搭载低功耗、小体积的嵌入式处理器, 如NVIDIA Jetson系列、ARM架构的处理器或

者定制的嵌入式计算单元等。受限于功耗、散热等因素, 这些处理器虽然具有一定的推理能力, 但算力较低且难以发挥到最佳性能。针对上述问题, YOLO v9在高效层聚合网络(E-LAN)模型^[18]的基础上, 提出了广义E-LAN模型(GELAN), 它同时将参数数量、计算复杂性、准确性和推理速度等因素考虑在内, 使得其可以通过选择不同的计算块适应不同的推理设备, 适应于水下回收时算力受限的场景。在卷积方面, GELAN模型仅使用传统卷积, 实现了比深度卷积更高的参数使用率。此外, YOLO v9从可逆函数的角度对现有的神经网络架构进行了分析, 引入可编程梯度信息(PGI)和辅助可逆支路, 解决了深度监督只能用于极深神经网络架构的问题^[16]。YOLO v9通过融合GELAN和PGI, 实现对水下引导光源特征信息的高效学习, 从而增强对深层特征和小尺寸特征的提取能力, 进而显著提升水下引导光源检测的鲁棒性和准确性。

2.2 特征匹配

特征匹配算法的目的是确立检测得到的二维特征点与现实世界中的三维特征点的对应关系。常用的特征匹配算法有尺度不变特征转换(SIFT)算法^[19]、加速稳健特征(SURF)算法^[20]等。在AUV的回收过程中, 先前工作中涉及的最快速度约2节^[3], AUV较快的移动速度导致相机前后帧差异显著提升, 导致依赖先前帧信息的算法容易失效。上述算法根据特征点周围像素的分布情况推理匹配关系, 耗时较长且依赖先前帧的信息, 难以适应于AUV速度较快的情况。因此, 本文设计一种基于匈牙利匹配的全向特征匹配算法, 其速度快且不依赖于先前帧信息, 进而可适应于AUV速度较快的情况。算法流程如图4所示, 具体可分为对真实图像处理和对相机实拍图像处理两部分, 处理过程分别用上标 o 和 m 表示。

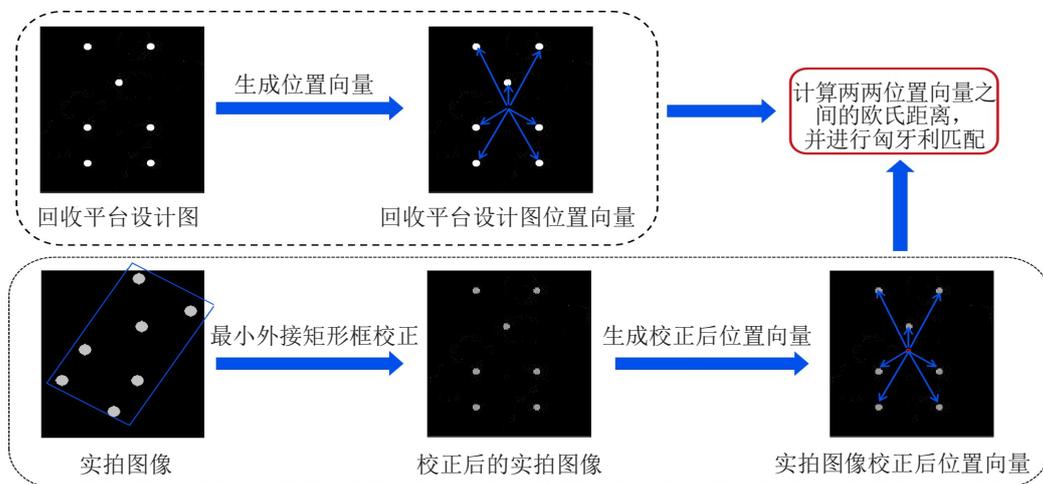


图4 特征匹配流程

在处理回收平台设计图时,假设回收平台设计图上共有 n 个特征点.首先计算所有特征点的质心坐标 q_c^o ,其公式为

$$q_c^o = (q_1^o + q_2^o + \dots + q_n^o)/n, \quad (5)$$

其中 q_i^o 为第 i 个特征点的位置.以质心坐标为向量起点,以各特征点坐标为向量终点,计算得到第 i 个特征点对应的位置向量 p_i^o ,即

$$p_i^o = q_i^o - q_c^o. \quad (6)$$

实拍图像的质心坐标 q_c^m 和各点位置向量 q_i^m 可通过相同方式求得.

为了使得实拍图像的位置向量分布更加接近回收平台设计图的位置向量分布,算法设计对实拍图像中的位置向量进行校正.在获取到实拍图像的位置向量后,根据图像中特征点的分布计算特征点的最小外接矩形框^[21]及其在图像上的偏转角度 θ (以绕 y 轴顺时针方向为正),并根据 θ 对特征向量进行校正(如图4所示).第 i 个特征点对应的位置向量 p_i^m 校正公式为

$$p_i^{m'} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} p_i^m, \quad (7)$$

其中 $p_i^{m'}$ 为校正后的位置向量.

在得到回收平台设计图和实拍图像中两组位置向量后,计算两组向量中两两向量的欧氏距离,得到代价矩阵 D ,即

$$D = \begin{bmatrix} \|p_1^o - p_1^{m'}\| & \|p_1^o - p_2^{m'}\| & \dots & \|p_1^o - p_n^{m'}\| \\ \|p_2^o - p_1^{m'}\| & \|p_2^o - p_2^{m'}\| & \dots & \|p_2^o - p_n^{m'}\| \\ \vdots & \vdots & \ddots & \vdots \\ \|p_n^o - p_1^{m'}\| & \|p_n^o - p_2^{m'}\| & \dots & \|p_n^o - p_n^{m'}\| \end{bmatrix}. \quad (8)$$

以计算得到的欧氏距离为代价,利用匈牙利匹配算法确定两组向量的对应关系 M ,即

$$M = \text{Hm}(D). \quad (9)$$

其中:Hm为匈牙利匹配,其总体流程如图4所示.

需要强调的是,为了能够实现全向匹配,在第1次匹配结束后,将实拍图像中特征点绕 y 轴旋转 180° 后,再重复一次匹配.取两次匹配中平均代价更小的一次匹配作为最终匹配结果.

3 基于光视觉的位姿估计算法

在光视觉处理完成后,基于光视觉的位姿估计算法利用处理得到图像中的二维特征点位置及其与现实世界中的三维特征点的对应关系,计算出AUV相

对于回收平台的位姿.本文针对坐落式回收的多阶段特点,设计适用于各阶段的位姿估计策略.此外,根据单双目相机的不同特点,分别设计相应的位姿测量算法.

3.1 单双目相机融合的位姿估计策略

对于坐落式回收不同阶段的特性,可概括为“仅双目可见”“单双目均可见”“仅单目可见”“单双目均不可见”4个阶段,如图5所示.

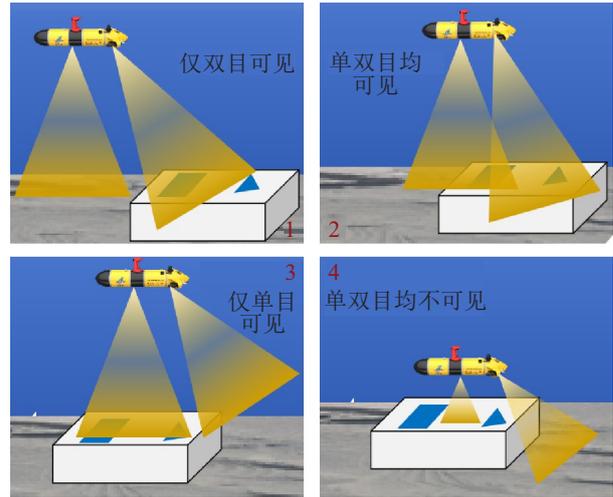


图5 坐落式回收典型分阶段特性

根据相机不同阶段的可视情况,需选择适用的位姿解算方法,以确保全过程位姿估计精度的平滑和稳定.在仅双目可见和仅单目可见阶段,分别采用双目位姿测量算法和单目位姿测量算法计算自身位姿.在单双目均可见阶段,根据双目相机和单目相机拍摄的图像中引导灯的数量,对单双目位姿测量结果进行加权融合.而在单双目均不可见阶段,则暂停位姿推算,保持下降状态,直至回收任务结束.

3.2 单目相机位姿测量算法

对于单目相机,本文利用匹配后的特征点基于PnP原理计算得到自身位姿.PnP是一种用于获取相机姿态的计算机视觉算法,通过已知的三维点和它们在相机图像中的对应投影点,推导出相机的位姿.进而利用单目相机位姿与AUV位姿的转换关系,得到AUV位姿.考虑到引导灯数量情况,本文使用迭代求解法^[22]求解相机位姿,通过迭代优化相机的位姿参数实现相机位姿的估计.该方法结合优化算法和重投影误差最小化,迭代地调整相机的位姿参数,直至找到最小化重投影误差的位姿参数,进而得到相机位姿.

在单目相机特征匹配完成后,可以得到 n 个像素坐标系下的特征点坐标 p_i 和其对应的三维空间点坐

标 q_i , 根据重投影误差函数, 构建优化目标

$$L^* = \arg \min_L \frac{1}{2} \sum_{i=1}^n \left\| p_i - \frac{1}{s_i} \mathbf{K} \mathbf{L} q_i \right\|^2. \quad (10)$$

其中: \mathbf{K} 为相机内参矩阵; $\mathbf{L} = \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ \mathbf{0} & 1 \end{bmatrix}$ 为相机位姿 (\mathbf{R}, \mathbf{T}) 的李群表示, \mathbf{R} 为相机姿态, \mathbf{T} 为相机位置.

通过上述操作可将位姿求解问题转化为优化问题, 利用列文伯格-马夸尔特优化算法, 最小化重投影误差函数, 得到相机的位姿 (\mathbf{R}, \mathbf{T}) .

3.3 双目相机位姿测量算法

对于双目相机, 本文首先通过三角测量确定匹配后特征点在双目相机坐标系下的坐标, 进而采用基于SVD分解的位姿测量算法计算AUV自身位姿.

3.3.1 三角测量算法

在双目相机位姿估计中, 本文首先通过三角测量原理^[22] 计算得到左右目图像中对应二维特征点在双目相机坐标系下的三维坐标, 其是最终AUV自身位姿计算的输入参数之一.

首先, 对单个相机的成像模型进行分析. 令 $[x_{bc}, y_{bc}, z_{bc}]^T$ 为三角测量时的双目相机坐标系下的三维坐标. 令 $[x_c, y_c, z_c]^T$ 为该相机坐标系下的3D点坐标, $[u, v]^T$ 为像素坐标系下坐标. 利用相机内参矩阵 \mathbf{K} , 可得

$$z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \mathbf{K} \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix}. \quad (11)$$

利用相机外参矩阵 \mathbf{P} , 可得

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} & \mathbf{R}_{13} & \mathbf{T}_x \\ \mathbf{R}_{21} & \mathbf{R}_{22} & \mathbf{R}_{23} & \mathbf{T}_y \\ \mathbf{R}_{31} & \mathbf{R}_{32} & \mathbf{R}_{33} & \mathbf{T}_z \end{bmatrix} \begin{bmatrix} x_{bc} \\ y_{bc} \\ z_{bc} \\ 1 \end{bmatrix} = \mathbf{P} \begin{bmatrix} x_{bc} \\ y_{bc} \\ z_{bc} \\ 1 \end{bmatrix}. \quad (12)$$

将式(12)移项可得

$$\begin{cases} x'_c = x_c / z_c = \frac{\mathbf{R}_{11}x_{bc} + \mathbf{R}_{12}y_{bc} + \mathbf{R}_{13}z_{bc} + \mathbf{T}_x}{\mathbf{R}_{31}x_{bc} + \mathbf{R}_{32}y_{bc} + \mathbf{R}_{33}z_{bc} + \mathbf{T}_z}, \\ y'_c = y_c / z_c = \frac{\mathbf{R}_{21}x_{bc} + \mathbf{R}_{22}y_{bc} + \mathbf{R}_{23}z_{bc} + \mathbf{T}_y}{\mathbf{R}_{31}x_{bc} + \mathbf{R}_{32}y_{bc} + \mathbf{R}_{33}z_{bc} + \mathbf{T}_z}. \end{cases} \quad (13)$$

最后, 综合考虑左右目相机成像模型, 构建三维坐标与二维特征点之间的转换关系. 左右目相机的内外参数分别用上标 l 和 r 表示, 左右目相机坐标系下的坐标分别用下标 cl 和 cr 表示. 将左右目相机的内外

参数分别代入式(13), 得到关于特征点在相机坐标系下的三维坐标 $[x_{bc}, y_{bc}, z_{bc}]^T$ 的超定方程

$$\tilde{\mathbf{R}} = \begin{bmatrix} x'_{cl} \mathbf{R}_{31}^l - \mathbf{R}_{11}^l & x'_{cl} \mathbf{R}_{32}^l - \mathbf{R}_{12}^l & x'_{cl} \mathbf{R}_{33}^l - \mathbf{R}_{13}^l \\ y'_{cl} \mathbf{R}_{31}^l - \mathbf{R}_{21}^l & y'_{cl} \mathbf{R}_{32}^l - \mathbf{R}_{22}^l & y'_{cl} \mathbf{R}_{33}^l - \mathbf{R}_{23}^l \\ x'_{cr} \mathbf{R}_{31}^r - \mathbf{R}_{11}^r & x'_{cr} \mathbf{R}_{32}^r - \mathbf{R}_{12}^r & x'_{cr} \mathbf{R}_{33}^r - \mathbf{R}_{13}^r \\ y'_{cr} \mathbf{R}_{31}^r - \mathbf{R}_{21}^r & y'_{cr} \mathbf{R}_{32}^r - \mathbf{R}_{22}^r & y'_{cr} \mathbf{R}_{33}^r - \mathbf{R}_{23}^r \end{bmatrix}, \quad (14)$$

$$\tilde{\mathbf{T}} = \begin{bmatrix} x'_{cl} \mathbf{T}_z^l - \mathbf{T}_x^l \\ y'_{cl} \mathbf{T}_z^l - \mathbf{T}_y^l \\ x'_{cr} \mathbf{T}_z^r - \mathbf{T}_x^r \\ y'_{cr} \mathbf{T}_z^r - \mathbf{T}_y^r \end{bmatrix}, \quad (15)$$

$$[\tilde{\mathbf{R}} \quad \tilde{\mathbf{T}}] \begin{bmatrix} x_{bc} \\ y_{bc} \\ z_{bc} \\ 1 \end{bmatrix} = \mathbf{A} \begin{bmatrix} x_{bc} \\ y_{bc} \\ z_{bc} \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \quad (16)$$

通过求解超定方程(16), 可得双目相机中的二维特征点在双目相机坐标系下的坐标 $[x_{bc}, y_{bc}, z_{bc}]^T$.

3.3.2 基于SVD分解法的位姿求解算法

在双目相机完成三角测量后, 利用所得的双目相机坐标系下的三维空间点坐标 $[x_{bc}, y_{bc}, z_{bc}]^T$ 及其在世界坐标系下对应的三维空间点坐标 $[x_w, y_w, z_w]^T$, 通过基于SVD分解的处理思路^[22], 计算得到AUV相对于回收平台的相对位姿. 首先, 利用相机位姿 (\mathbf{R}, \mathbf{T}) 确定两个坐标系下坐标点的转换关系, 进而构造关于 (\mathbf{R}, \mathbf{T}) 的误差函数并通过SVD分解法最小化误差函数, 从而得到相机位姿. 最后, 利用双目相机位姿与AUV位姿之间的转换关系, 得到AUV位姿.

首先, 构造关于相机位姿 \mathbf{R}, \mathbf{T} 的误差函数

$$E(\mathbf{R}, \mathbf{T}) = \sum_{i=1}^n \|(\mathbf{R} p_i + \mathbf{T}) - q_i\|^2. \quad (17)$$

其中: n 为特征点数目, p_i 和 q_i 分别为第 i 个特征点在双目相机坐标系和世界相机坐标系下的坐标.

设 q_0 和 p_0 分别为两组坐标的质心坐标, 对误差函数进行如下分析和变换:

$$\begin{aligned} E(\mathbf{R}, \mathbf{T}) &= \sum_{i=1}^n \| \mathbf{R}(p_i - p_0) - (q_i - q_0) + \mathbf{T} - q_0 + \mathbf{R}p_0 \|^2 = \\ &= \sum_{i=1}^n (2(\mathbf{R}(p_i - p_0) - (q_i - q_0))^T (\mathbf{T} - q_0 + \mathbf{R}p_0) + \\ & \quad \| \mathbf{R}(p_i - p_0) - (q_i - q_0) \|^2 + \| \mathbf{T} - q_0 + \mathbf{R}p_0 \|^2). \end{aligned} \quad (18)$$

根据质心的定义, 可将式(18)进一步转换为

$$E(\mathbf{R}, \mathbf{T}) = \sum_{i=1}^n \|\mathbf{R}(\mathbf{p}_i - \mathbf{p}_0) - (\mathbf{q}_i - \mathbf{q}_0)\|^2 + n\|\mathbf{T} - \mathbf{q}_0 + \mathbf{R}\mathbf{p}_0\|^2, \quad (19)$$

可将第1项优化至最小,得到 \mathbf{R} 的最优解,进而根据得到的 \mathbf{R} 求解得到 \mathbf{T} .因此有

$$\begin{aligned} \mathbf{R}^* &= \arg \min_{\mathbf{R}} \sum_{i=1}^n \|\mathbf{q}_i - \mathbf{q}_0 - \mathbf{R}(\mathbf{p}_i - \mathbf{p}_0)\|^2 = \\ &= \arg \max_{\mathbf{R}} \sum_{i=1}^n \text{trace}(\mathbf{R}\mathbf{p}_i\mathbf{q}_i^T) = \\ &= \arg \max_{\mathbf{R}} \text{trace}(\mathbf{R}\mathbf{H}), \end{aligned} \quad (20)$$

其中 $\mathbf{H} = \sum_{i=1}^n \mathbf{p}_i\mathbf{q}_i^T$.对 \mathbf{H} 进行SVD分解可得 $\mathbf{H} = \mathbf{U}\mathbf{S}\mathbf{V}^T$,其中 \mathbf{U} 和 \mathbf{V} 均为正交阵.继续对式(20)进行变换,有

$$\begin{aligned} \mathbf{R}^* &= \arg \max_{\mathbf{R}} \text{trace}(\mathbf{R}\mathbf{U}\mathbf{S}\mathbf{V}^T) = \\ &= \mathbf{V}(\arg \max_{\mathbf{R}'} \text{trace}(\mathbf{R}'\mathbf{S}))\mathbf{U}^T = \\ &= \mathbf{V}(\arg \max_{\mathbf{R}'} (r'_{11}\sigma_1 + r'_{22}\sigma_2 + r'_{33}\sigma_3))\mathbf{U}^T. \end{aligned} \quad (21)$$

由于 \mathbf{R} 为正交阵,有 $-1 \leq r'_{11}, r'_{22}, r'_{33} \leq 1$,且有 $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq 0$,所以 \mathbf{R}' 的最优解为单位矩阵,进而可得初步姿态最优解 $\mathbf{R}^* = \mathbf{V}\mathbf{U}^T$.由于 \mathbf{V}, \mathbf{U} 中可能含有反射元素,对初步最优解进行修正,得到最终姿态解

$$\mathbf{R}^* = \mathbf{V} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \det(\mathbf{V}\mathbf{U}^T) \end{bmatrix} \mathbf{U}^T. \quad (22)$$

利用姿态最终解 \mathbf{R}^* ,代入方程(19)得到位置最终解

$$\mathbf{T}^* = \mathbf{q}_0 - \mathbf{R}^*\mathbf{p}_0. \quad (23)$$

4 实验验证与分析

为了验证水下引导灯光源提取算法的有效性,本文在AUV回收实拍数据集上进行训练和测试.此外,为了验证位姿估计算法的准确性,本文基于gazebo平台搭建了仿真验证系统,实现了AUV回收模拟以及误差分析.

4.1 光视觉处理

在去雾增强方面,本文实验使用的操作系统为Ubuntu16.04, GPU为Nvidia RTX 1080Ti,显存为11 G, cuda版本为11.7, Pytorch深度学习框架.设置迭代批量(batch-size)为3,优化算法为带动量的随机梯度下降法^[23](SGD),预设初始学习率为0.01,动量为0.94,衰减系数为0.0005.训练迭代系数设置为2000.在验证集上,当损失函数100轮内均没有下降

时,提前停止迭代求解.

本实验中首先将控制去雾程度的参数 ω 设置为不同值,分别对数据集进行不同程度的先验去雾.进而将不同程度去雾后的图像作为训练YOLOv9的数据集,分别进行训练,并通过对比不同 ω 对应的误差函数曲线验证去雾算法的有效性. YOLOv9采用二元交叉熵损失(BCE loss)估算分类损失,采用分布焦点损失(DFL loss)和CIou loss(complete-IoU loss)^[24]评估边框回归损失.由于AUV回收过程中对于引导灯特征点位置精度要求较高,本实验以评估YOLOv9回归损失的DFL loss和CIou loss作为评估去雾算法有效性的指标.当控制去雾程度参数 ω 设置为不同参数时,实验得到的验证集损失函数变化曲线如图6所示.在图6中DFL loss和BOX loss代表回归损失.

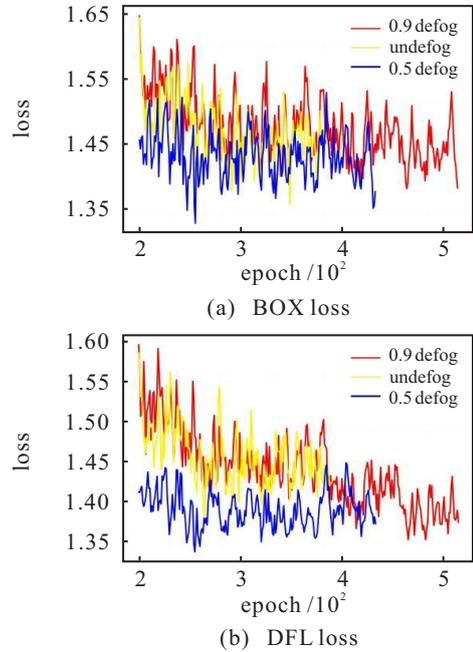


图6 损失函数变化

实验结果显示,当 $\omega = 0.5$ 时,去雾算法可有效去除雾气效果,同时保留图像的主要特征信息,表现出最为理想的收敛效果.随着迭代次数的增加,损失函数保持下降趋势,最终收敛到较小的值.因此,本文将控制去雾程度的参数 ω 设为0.5,以进行暗通道先验去雾图像增强.去雾前后效果如图7所示.



图7 去雾效果对比

在英伟达Jeston AGX Orin边缘计算平台中,基于YOLOv9-e的引导灯光源检测模型平均帧率(FPS)可达3.35帧,基于YOLOv9-c的引导灯光源检测模型平均FPS可达4.11帧.在特征检测方面,本文采用精确率(precision)、召回率(recall)、以及平均精度均值(mAP)等指标对YOLOv9模型的检测能力进行评估.精确率为预测正样本中实际正样本数,可用于评估检测模型的可靠性.召回率为预测得到的正样本占有正类别样本的比例,可用于评估检测模型的全面性和敏感性.平均精度均值综合精确率和召回率两种指标,代表模型的综合检测能力.计算并保存验证集上各指标在每一轮训练中的值,其变化曲线如图8所示.

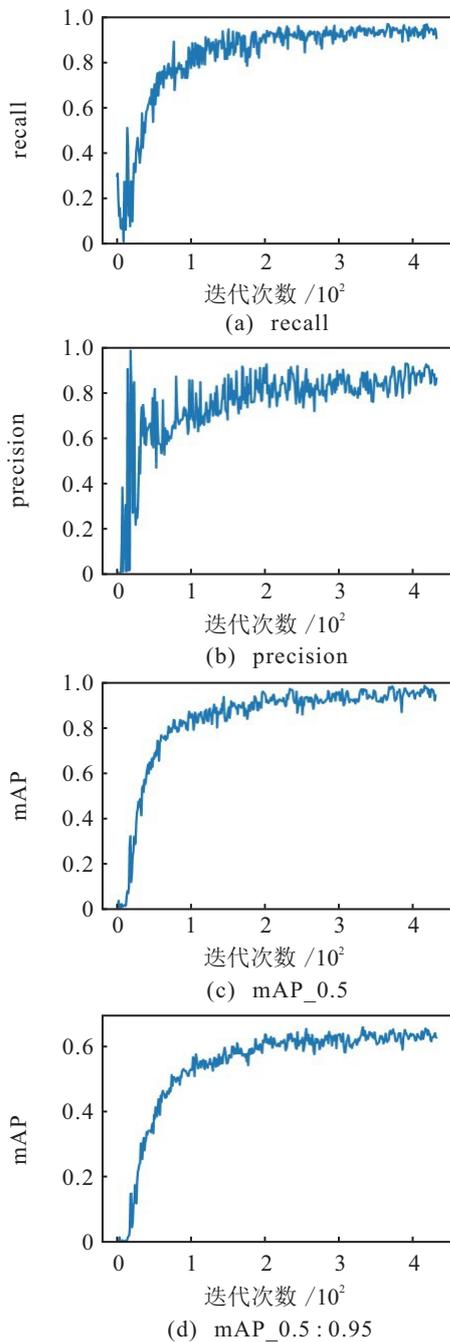


图8 检测模型评估结果

由实验结果可知,模型准确率和召回率最终分别收敛到0.8和0.9左右.当交并比阈值为0.95时,mAP最终收敛到0.9左右;阈值为0.5时,mAP收敛到0.65左右.各评价指标均达到了较好的收敛值,验证了检测算法的准确性和有效性.

图9为检测算法在AUV回收实拍数据集测试集上的检测效果.

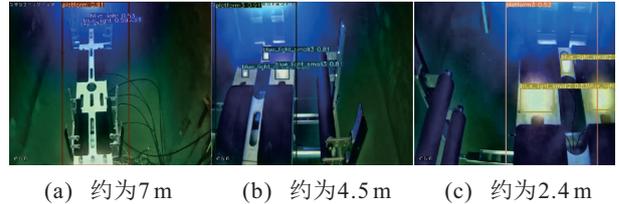


图9 实际检测结果

在图9中:图9(a)~图9(c)呈现了AUV自远处逐渐靠近回收平台过程中的实拍画面;图9(a)距回收平台距离约为7 m,图9(b)距回收平台距离约为4.5 m,图9(c)距回收平台距离约为2.4 m.在AUV距回收平台不同距离时,引导灯尺寸与颜色差异较大.但基于暗通道先验去雾和YOLOv9的水下引导灯光源检测算法可充分学习到图像特征信息,从而在不同距离均达到了较好的引导灯检测效果.

为了验证特征匹配算法的准确性和鲁棒性,本文从不同角度拍摄回收平台照片50张并对其进行特征匹配.在拍摄过程中,需确保所有引导灯均呈现在所拍摄的照片中.同时,每张图片较前一张图片的偏航角变化控制在6 rad~8 rad的范围内,以便尽可能涵盖所有可能的角度变化,并且俯仰角和横滚角分别在-60 rad~60 rad的区间内随机取值,从而获取更全面、多角度的拍摄效果.实验结果表明,本文使用的特征匹配算法在这50张照片上均可以实现正确的特征匹配.图10为部分有代表性的匹配结果.

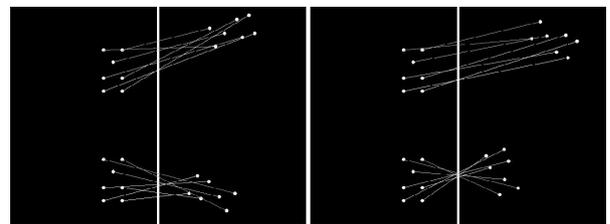


图10 特征匹配结果

4.2 位姿估计算法

位姿估计算法采用仿真实验验证的方式.仿真实验使用的操作系统为Ubuntu16.04,ROS版本为kinetic,gazebo版本为7.0.

首先,本文在gazebo中搭建轻量级、易移植的仿

真机器人. 在控制器方面, 仿真机器人通过插件的方式直接进行速度控制, 实现了几乎零误差、任意速度的角速度和线速度控制. 在传感器方面, 仿真机器人搭配了单目相机、双目相机、惯性测量单元(IMU)等. 本文编写控制程序, 控制仿真机器人在回收平台上方运动, 模拟回收过程. 位姿估计算法以固定频率进行, 记录每次估计的测量值和真实值.

双目相机位姿测量和单目相机位姿测量的误差数据分别如图11和图12所示.

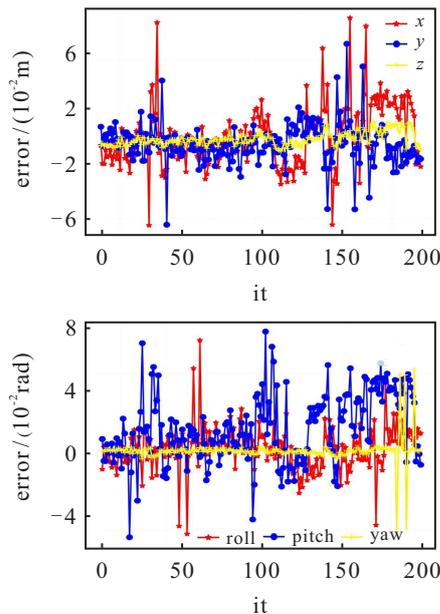


图11 双目相机位置与姿态误差折线

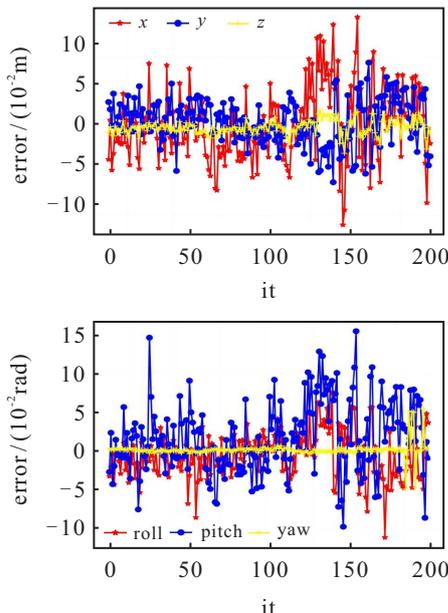


图12 单目相机位置与姿态误差折线

在图11和图12中: (x, y, z) 为AUV位置误差(单位: m), $(roll, pitch, yaw)$ 为AUV姿态误差(单位: rad). 横坐标为估计次数, 纵坐标为测量值与真实值的差值.

对于双目测量, 因为其充分利用了双目图像信息, 单方向最大位置误差在0.08 m左右, 单方向最大姿态误差在0.08 rad左右. 对于单目测量, 由于其只利用单目图像信息, 误差值略有上升. 由于回收平台特征点较少, 对噪声较为敏感, 个别情况下产生了误差较高的现象. 单方向最大位置误差在0.13 m左右, 单方向最大姿态误差在0.15 rad左右. 但除个别情况, 单目测量仍可控制位置误差和姿态误差在0.1 m和0.1 rad以内. 此外, 由于本算法不依赖之前帧信息, 回收过程中位姿误差仅在小范围内波动, 不存在累积误差. 对于位置误差, 在AUV回收过程的前期阶段, AUV主要依赖在X方向和Y方向上的运动来接近回收平台, 因此其位置误差主要表现为X方向和Y方向上的偏差. 对于姿态误差, 由于AUV在俯仰角和横滚角上的运动, 使得实际拍摄图像中各光源的大小产生差异, 这种差异会导致引导灯光源位置的检测出现不准确的情况. 因此, 相较于偏航角方向的误差, 俯仰角和横滚角方向的误差会相对较大.

5 结论

本文针对AUV坐落式回收过程中的自身位姿测量问题, 设计了一种单双目相机融合的位姿估计算法. 针对水下引导灯光源的提取问题, 本文将先验去雾算法与YOLO v9目标检测相结合, 实现了在水下干扰情况下的特征点检测. 针对特征点匹配问题, 本文设计了一种基于匈牙利匹配的全向特征匹配算法, 实现了不依赖于AUV速度的全向特征匹配. 最后, 对于单双目相机, 分别利用PnP算法和基于SVD分解法的位姿求解算法实现最终位姿估计.

实验结果表明: 1) 暗通道先验去雾与YOLO v9相结合可以实现引导灯特征点的准确检测; 2) 本文设计的特征匹配算法可以实现全向特征匹配, 进而实现AUV位姿的精准估计.

参考文献(References)

- [1] Kronen D M. Docking the ocean explorer autonomous underwater vehicle using a low-cost acoustic positioning system and a fuzzy logic guidance algorithm[C]. Florida Atlantic University. Boca Raton, 1997: 1-24.
- [2] 吴宇, 苏析超, 崔佳鹏, 等. USV&AUV水下目标协同搜索与打击航迹规划[J]. 控制与决策, 2021, 36(4): 825-834.
(Wu Y, Su X C, Cui J P, et al. Coordinated path planning of USV & AUV for an underwater target[J]. Control and Decision, 2021, 36(4): 825-834.)
- [3] 孙叶义, 武皓微, 李晔, 等. 智能无人水下航行器水下

- 回收对接技术综述[J]. 哈尔滨工程大学学报, 2019, 40(1): 1-11.
(Sun Y Y, Wu H W, Li Y, et al. Summary of AUV underwater recycle docking technology[J]. Journal of Harbin Engineering University, 2019, 40(1): 1-11.)
- [4] 潘光, 黄明明, 宋保维, 等. AUV回收技术现状及发展趋势[J]. 鱼雷技术, 2008, 16(6): 10-14.
(Pan G, Huang M M, Song B W, et al. Current situation and development trend of AUV recovery technology[J]. Journal of Unmanned Undersea Systems, 2008, 16(6): 10-14.)
- [5] Xia G Q, Zhang Y, Zhang W, et al. Robust adaptive super-twisting sliding mode formation controller for homing of multi-underactuated AUV recovery system with uncertainties[J]. ISA Transactions, 2022, 130: 136-151.
- [6] 岳伟, 季嘉诚, 刘中常, 等. 基于准最大最小模型预测控制的AUV视觉对接[J]. 控制与决策, 2023, 38(7): 1887-1894.
(Yue W, Ji J C, Liu Z C, et al. Quasi-min-max MPC algorithm for visual docking of an autonomous underwater vehicle[J]. Control and Decision, 2023, 38(7): 1887-1894.)
- [7] Xu S, Jiang Y, Li Y, et al. A stereo visual navigation method for docking autonomous underwater vehicles[J]. Journal of Field Robotics, 2024, 41(2): 374-395.
- [8] Fan S S, Liu C Z, Li B, et al. AUV docking based on USBL navigation and vision guidance[J]. Journal of Marine Science and Technology, 2019, 24(3): 673-685.
- [9] Ren R, Zhang L, Liu L, et al. Two AUVs guidance method for self-reconfiguration mission based on monocular vision[J]. IEEE Sensors Journal, 2021, 21(8): 10082-10090.
- [10] Li Y, Jiang Y, Cao J, et al. AUV docking experiments based on vision positioning using two cameras[J]. Ocean Engineering, 2015, 110: 163-173.
- [11] Yahya M F, Arshad M R. Tracking of multiple light sources using computer vision for underwater docking[J]. Procedia Computer Science, 2015, 76: 192-197.
- [12] Zhong L, Li D, Lin M, et al. A fast binocular localisation method for AUV docking[J]. Sensors: Basel, Switzerland, 2019, 19(7): 1735.
- [13] Harris C, Stephens M. A combined corner and edge detector[C]. Alvey vision conference. Manchester, 1988: 147-151.
- [14] Rosten E, Drummond T. Fusing points and lines for high performance tracking[C]. 10th IEEE International Conference on Computer Vision. Beijing, 2005: 1508-1515.
- [15] He K, Sun J, Tang X. Single image haze removal using dark channel prior[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(12): 2341-2353.
- [16] Wang C Y, Yeh I H, Liao H Y M. YOLO v9: Learning what you want to learn using programmable gradient information[J/OL]. 2024, arXiv: 2402.13616.
- [17] Berman D, Levy D, Avidan S, et al. Underwater single image color restoration using haze-lines and a new quantitative dataset[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 43(8): 2822-2837.
- [18] Wang C Y, Liao H Y M, Yeh I H. Designing network design strategies through gradient path analysis[J/OL]. 2022, arXiv: 2211.04800.
- [19] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [20] Bay H, Tuytelaars T, Van Gool L. Surf: Speeded up robust features[J]. Lecture Notes in Computer Science, 2006, 3951: 404-417.
- [21] Shamos M I. Computational geometry[M]. New Haven: Yale University, 1978: 31-142.
- [22] 高翔, 张涛, 刘毅, 等. 视觉SLAM十四讲: 从理论到实践[M]. 北京: 电子工业出版社, 2017: 130-180.
(Gao X, Zhang T, Liu Y, et al. Fourteen lectures on visual SLAM: From theory to practice[M]. Beijing: Publishing House of Electronics Industry, 2017: 130-180.)
- [23] Ruder S. An overview of gradient descent optimization algorithms[J/OL]. 2016, arXiv: 1609.04747.
- [24] Zheng Z, Wang P, Liu W, et al. Distance-IoU loss: Faster and better learning for bounding box regression[C]. Proceedings of the AAAI conference on artificial intelligence. Washington, 2020: 12993-13000.

作者简介

祝志坤(2001—), 男, 博士生, 主要研究方向为水下多智能体光学测量, E-mail: zhuzhikun@sjtu.edu.cn;

卢丙举(1979—), 男, 研究员, 主要研究方向为水下机器人总体设计、智能检测作业, E-mail: bitlcy@qq.com;

李一辰(1993—), 男, 助理研究员, 博士, 主要研究方向为水下多潜器系统定位导航、复合测量感知、多源信息融合, E-mail: liyichensjtu@sjtu.edu.cn;

王凯(1992—), 男, 工程师, 硕士, 主要研究方向为计算机视觉、水下目标识别检测, E-mail: wangkai24@buaa.edu.cn;

于文彬(1983—), 男, 副研究员, 博士, 主要研究方向为海洋传感网络多源融合感知、光学目标跟踪、复合测量, E-mail: yuwenbin@sjtu.edu.cn.