

AUV 水下回收光学导引(特邀)

李学龙^{1,2*}, 孙哲^{1,2**}, 吴国俊^{3***}¹西北工业大学光电与智能研究院, 陕西 西安 710072;²中国电信人工智能研究院(TeleAI), 上海 200232;³中国科学院西安光学精密机械研究所, 陕西 西安 710119

摘要 自主水下航行器(AUV)是“进入海洋、探测海洋、利用海洋”的重要工具,AUV 水下回收光学导引技术一直以来都是国内外的研究热点。在总结分析国内外 AUV 水下回收单目视觉、双目视觉、位置探测器导引技术的基础上,重点讨论了各类 AUV 水下回收光学导引技术的实现原理、方法、发展现状与趋势,详细介绍了基于图像传感器和位置探测器的光学导引技术。未来 AUV 水下回收光学导引技术将从单一功能模块向智能系统生态演进,为构建涉水智能光电技术与装备体系提供核心技术基础,以满足海洋科考、资源勘探、水下安防等领域的重大应用需求。

关键词 光学导引; 自主水下航行器; 图像信号处理; 位置信号处理; 涉水视觉; 涉水光学

中图分类号 TP242.3

文献标志码 A

DOI: 10.3788/AOS241803

1 引言

自主水下航行器(AUV)水下回收光学导引技术是实现 AUV 与水下回收站的物理连接和无线通信,完成能源补充、数据传输、指令下达的关键技术,支撑 AUV 实现更高的作业效率、更长的作业时间、更大的作业范围,保证 AUV 安全、连续、有效地完成回收作业,是涉水智能光电技术与装备的核心技术,属于涉水光学与涉水视觉技术体系,在海洋科考、资源勘探、水下安防等领域具有重要意义^[1-4]。

AUV 水下回收的主要方式包括水面母船回收和水下对接回收,其中 AUV 水下回收具有更自主、更隐蔽、更高效的特点,因此 AUV 水下回收光学导引成为了国内外水下无人装备技术的研究热点之一。目前,按照对接装置形式,AUV 水下回收可以分为平台式和坞站式^[5]。平台式 AUV 水下回收能够实现全方位对接,但是机械设计复杂^[6]。坞站式 AUV 水下回收使用喇叭口形回收站,允许 AUV 以一定姿态误差驶入回收站,尽管只能实现单方向对接,但是具有机械结构简单、便于布放的特点,有利于实现自主式回收,是目前 AUV 水下回收最主要的形式之一^[7]。

坞站式 AUV 水下回收光学导引过程通常分为两个阶段,即中远距离导引和近距离末端导引,利用声学导引、电磁导引和光学导引单独或组合完成。中远距离导引通常采用声学导引,或者融合惯性导航、多

普勒计程仪等设备,以组合导引方式使 AUV 接近回收站。其中,声学导引感知距离远,可达 10 km 量级,但分辨率低、数据更新慢,容易受到目标结构或海底等声反射源的影响,且难以在近距离实现高精度定位。近距离末端导引是 AUV 驶入回收站附近,对接进入回收站的最终阶段,直接决定了 AUV 能否成功回收。因此,近距离末端导引定位精度要求更高,通常采用电磁导引或者光学导引方法。其中,电磁信号在水下衰减的速度很快,通用性较差,作用范围受限。光学导引利用光电探测器对水下回收站的导引标识物进行检测和识别,实现相对位姿解算,具有精度高、隐蔽性好的特点。AUV 水下回收光学导引的作业过程示意图如图 1 所示。为了提高 AUV 末端水下回收光学导引成功率,需要针对性地设计 AUV 水下回收光学导引方法^[8]。

AUV 水下回收光学导引根据导引识别方式可以分为基于图像传感器的光学导引和基于位置探测器的光学导引,如图 2 所示。基于图像传感器的视觉导引方法包括单目视觉导引和双目视觉导引,通过检测回收基站上安装的导引灯信号或者特征图像,与已知特征进行比对计算,确定回收站的相对位置。基于位置传感器的光学导引装置包括四象限光电探测器和多象限光电探测器,通过检测光斑位置对回收站导引灯进行检测和定位。目前国内外开展了大量关于 AUV 水下回收光学导引方面的研究,本文总结了近年来报道

收稿日期: 2024-11-26; 修回日期: 2025-02-28; 录用日期: 2025-04-22; 网络首发日期: 2025-04-25

基金项目: 国家重点研发计划(2022YFC2808003)、陕西省自然科学基金面上项目(2024JC-YBMS-468)

通信作者: *xuelong_li@ieee.org; **sunzhe@nwpu.edu.cn; ***wuguojun@opt.ac.cn

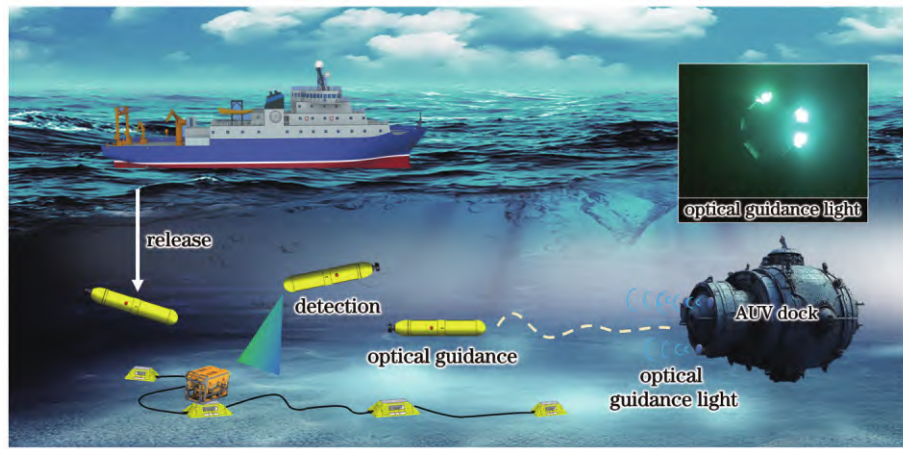


图1 AUV水下回收光学导引示意图

Fig. 1 Schematic diagram of optical guidance for AUV underwater dock

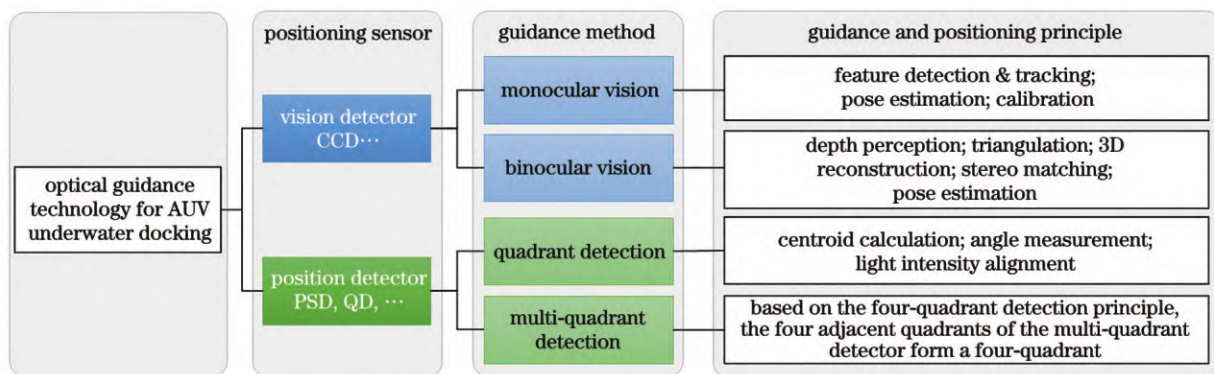


图2 AUV水下回收光学导引技术

Fig. 2 Optical guidance technologies for AUV underwater dock

的相关理论、方法与研究现状,展望了AUV水下回收光学导引的发展方向。

2 基于图像传感器的水下光学导引

2.1 原理与挑战

基于图像传感器的光学导引的原理是通过捕捉回收站笼口的主动或者被动标识物图像,利用图像的特征识别和透视几何计算AUV相对回收站的位置和姿态信息,该方法具有导引标识形式不受限制、任务可扩展性强等优点。其中:被动标识通常使用二维码、几何形状图案等,具备唯一性,但可视距离近;主动标识则采用光信标发射光信号,具有更高的可见性,可视距离远。然而,基于图像传感器的水下光学导引在实际应用中面临水下光场退化畸变、信标特征表征退化、跨域解算泛化受限等挑战。水下光场在传输过程中呈现非均匀衰减特性,水体介质对光的吸收与散射导致AUV图像传感器捕获的图像出现几何失真、对比度降低,有效探测范围减小等现象;水下信标纹理单一,受光场能量空间再分布的影响,水下信标特征点提取与匹配困难,这进一步增加了基于图像传感器的水下光学导引难度;水下光场的折射率与盐度、浊度参数存在强耦合

效应,导致成像系统在不同区域的位姿解算误差呈非线性变化,使水下光学导引算法的泛化性受限。

针对此,在基于图像传感器的水下光学导引研究方面,研究人员开发了水下图像复原与增强技术,以提高图像信噪比,提升特征点提取和匹配准确率。结合深度学习技术,构建神经网络模型,利用非线性激活函数和优化算法,在数据集和算力支持下,学习数据的复杂特征和内在规律,实现更为准确和稳定的特征识别,提升AUV水下回收光学导引性能。

2.2 单目视觉导引国内外研究现状

1) 传统的单目视觉导引

单目视觉导引使用单目相机识别水下回收站的信标,计算信标在画面中的位置、角度等信息,结合成像模型、AUV运动方程或位置状态,获得AUV与回收站的相对位置和姿态,进而控制AUV实现水下对接回收。单目视觉导引按照导引方法可以分为灯信标、图像信标、灯信标与图像信标融合、灯信标与其他类型信标融合的方法。

(1) 灯信标

AUV单目视觉水下回收光学导引使用灯信标的原理是:利用水下单目视觉系统识别并追踪特定光源

信号(即灯信标),从而为AUV水下回收提供精确引导定位的方法。灯信标是主动光信标,水下可视距离远,广泛应用于AUV水下回收单目视觉导引。Lee等^[9-10]利用单目视觉实现AUV导引,并在回收站设计了点状灯信标作为特征识别点。在此基础上,Park等^[11-12]基于单目相机的ISiMI水下平台,设计了末端回收光学导引算法。Palomeras等^[13]基于有源光信标,实现了距离定位。Fan等^[14]使用单灯光信标,基于单目相机的像素位置确定AUV相对回收站中心线的角度,提高了AUV远距离(25 m)的水下回收光学导引成功率。上述研究提供了单目视觉导引的基本范式,但是单灯信标的光学导引定位精度较低。

为了进一步提高AUV水下回收成功率,研究人员通过增加灯信标的数量提高光学(视觉)导引定位精度。多灯信标光学导引定位解算的核心是求解PnP(Perspective- n -Point),即求解3D到2D点对的物体运动定位,具体过程是:已知灯信标在世界坐标系下的坐标,以及灯信标在相机的图像平面的像素坐标,求解相机的姿态和内外参数。PnP算法的几何结构如图3所示,根据成像平面上多个目标特征点的投影信息的标

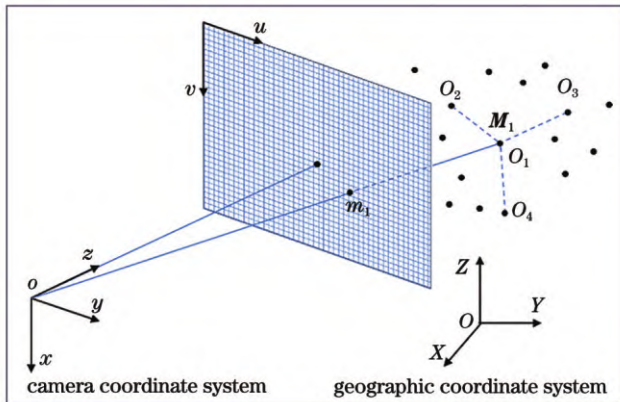


图3 PnP算法几何结构示意图

Fig.3 Schematic diagram of geometric structure of PnP algorithm

定参数,以及成像坐标系、相机坐标系和目标坐标系的投影变换关系,计算目标6个自由度的运动信息,实现回收站目标位姿估计。图3中 $o-xyz$ 为相机坐标系, $O-XYZ$ 为世界坐标系, (u,v) 为3D点经相机成像模型投影到图像平面后的像素坐标, O_1, O_2, O_3, O_4 为世界坐标系中的3D点, M_1 为相机投影矩阵, m_1 为相机传感器平面的成像点。Yan等^[15]设计了L形4灯光学阵列的叉式载杆对接装置,随后Zhang等^[16]利用环境光自适应选择分割阈值方法以识别导引灯,通过设计单目标分割后的跟踪定位方法,提高了水下回收光学导引成功率。Trslic等^[17]设计了不规则标识的4灯排布,获得了导引灯在单目相机中的大小和位置信息,解算AUV与回收站的相对位姿,该方法对于导引灯在相机画面中的位置和大小精度要求较高。Su等^[18]设计了6灯光信标排布结构,进一步提高了导引定位精度。为了在回收站光信标被遮挡时仍能对AUV完成导引,Figueiredo等^[19]建立了高性能且低算力需求的单目视觉导引系统,从算法层面对异常值和错误检测进行排除。通过上述研究可以看出,在AUV水下回收光学导引过程中增加导引灯信标的数量,能够显著提高单目视觉导引定位精度。

本课题组研制了水下光场自适应的单目视觉导引测试系统,如图4所示。该系统采用呈方形排布的光源阵列作为合作目标,使用水下大视场单目相机实时采集光源阵列的图像数据。基于涉水视觉基本理论,提出了双域协同分割算法,构建了水下光场传输物理约束,通过在空域采用暗通道先验引导的显著性检测显著提升光斑识别精度,通过设计空间欧氏距离约束的灯阵匹配算法,实现导引灯像素坐标与空间坐标的精准匹配。利用PnP算法进行单目位姿估计,实现了2~10 m水下距离的精确定位(定位精度80 mm@2 m和500 mm@10 m),输出单目相机与导引灯阵列之间的相对姿态信息,为AUV水下回收光学导引提供技术基础。

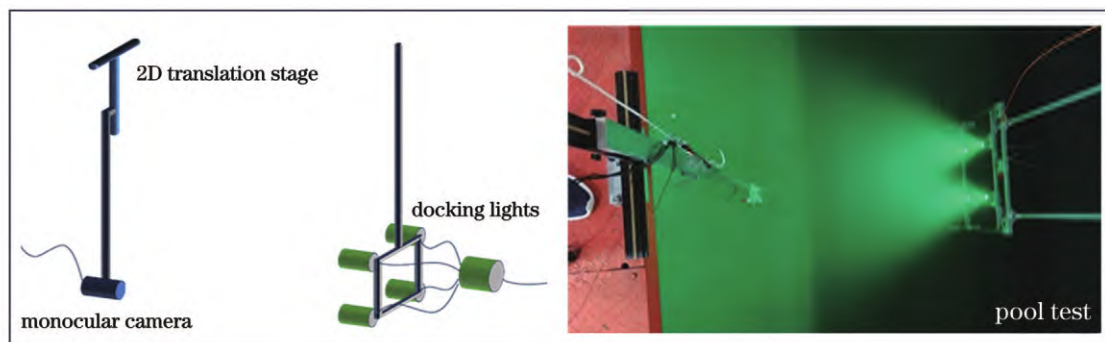


图4 水下单目视觉导引测试系统

Fig.4 Underwater monocular vision guidance test system

(2) 图像信标

尽管多灯信标能够显著提高AUV水下回收单目

视觉定位精度,但是由于单目视觉缺少深度信息,水下回收站灯信标坐标解算过程中容易出现无解或多解情

况。针对此,研究人员开展了水下图像信标的单目视觉导引研究。Rannestad等^[20]将ArUco标记在水下回收站上,极大增加PnP中的对应点数 n ,提升了单目视觉导引系统的鲁棒性与解算精度。但是,该方法在浑浊水环境或者水下远距离条件下,利用单目视觉二值化方法提取水下图像信标特征点时,定位精度较低。针对此,Lv等^[21]提出了一种自适应的图像二值化阈值计算方法,提高了信标特征点提取的成功率,如图5(a)~(c)所示。受水下湍流的影响,AUV水下航行姿态通常不稳,Yan等^[22]提出了关联分析的图像匹配算法,采用图像灰度直方图的灰度绝对关联度作为适应度函数,提高了图像匹配精度和鲁棒性。尽管使用图像信

标的单目视觉导引方法相对于灯信标的鲁棒性较好,但是该方法基于被动成像原理,导引距离相对受限。

(3) 灯信标与图像信标融合

为了增大单目视觉导引信标的导引距离、提高系统鲁棒性,Zhang等^[23]基于ArUco (Augmented Reality University of Cordoba)码,利用标记过的单目相机定位单灯光信标,提高了AUV与回收站的相对位置精度^[24]。Zhang等^[25]设计了光信标和三角形黑白码联合的单目视觉导引方法,降低了差异值测量和惯性导航造成的累计误差,如图5(d)所示。该研究证明通过融合灯信标与图像信标,能够在保证单目视觉导引距离的同时,提高导引定位精度,并提升系统鲁棒性。

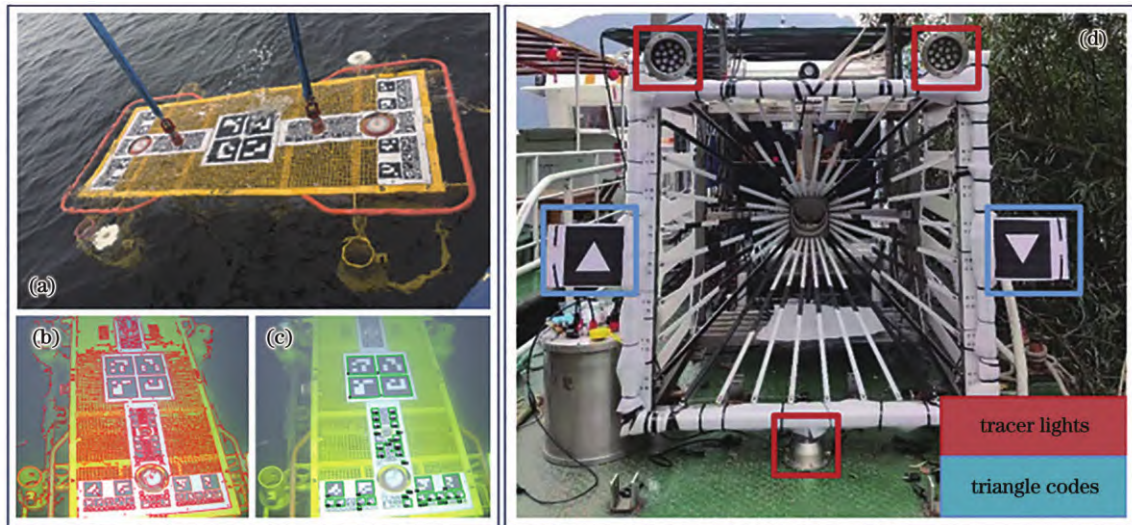


图5 平台式与笼式回收站。(a)平台式回收站;(b)平台式回收站轮廓提取;(c)平台式回收站的标记检测与识别;(d)光信标与三角形黑白码结合的回收站

Fig. 5 Platform and cage docks. (a) Platform dock; (b) contour extraction of platform dock; (c) marker detection and recognition of platform dock; (d) dock with light beacons combined with triangular black and white codes

(4) 灯信标与其他信标融合

为了进一步提高浑浊水环境中AUV回收的导引定位精度,Lin等^[26]建立了电磁信标和光信标信息的融合模型,显著提高了导引系统的鲁棒性,静态和动态的水下回收光学导引实验证明定位精度误差小于20 cm。为了减小AUV水下回收光学导引过程中水下湍流、AUV姿态偏差等因素造成的干扰,Zhang等^[27]提出了基于粒子滤波的融合定位算法,将单目视觉导引与AUV携带的全球定位系统、惯性导航系统、多普勒速度仪相结合,显著提高了AUV单目视觉导引定位精度。通过上述研究可以看出,AUV水下回收单目视觉导引朝着多信标信息与多源定位设备融合的方向发展。

综上所述,单目视觉导引方法主要集中在通过改变信标的分布、颜色、纹理、发光,以及标识物所含信息等方面,提升信标的可识别性与可靠性。表1总结了单目视觉导引定位方法的研究现状,可以看出各研究团队在信标的布局和设计上均有所不同,单目视觉导引方法利用已知特征点之间的真实距离与图像上的几

何距离进行距离测算,或者结合AUV运动状态方程、全球定位系统、惯性导航、多普勒速度仪等传感器,确定AUV与回收站的相对位置。目前,AUV水下回收单目视觉导引的最高定位精度优于10 cm量级。

2) 深度学习的单目视觉导引

深度学习方法具有特征学习和模式识别能力,广泛应用于灯信标和图像信标的AUV单目视觉水下回收光学导引,其基本原理是:基于单目相机拍摄到的画面,利用深度学习方法对灯信标或图像信标进行识别,以增大对信标轮廓的识别距离、提升识别精度,获取高精度的AUV位置和姿态,实现单目视觉导引的AUV水下回收。

(1) 灯信标

基于灯信标的深度学习单目视觉导引方法主要基于技术成熟的YOLO深度神经网络快速识别信标边界,提高分类概率。Singh等^[28]使用YOLO算法对5灯信标进行识别,利用PnP方法对AUV位置和姿态进行解算。Karimanzira等^[29]基于区域卷积神经网络(R-

CNN),通过单目相机识别回收站的侧壁、箱底、后壁,获取 AUV 的朝向和位置。Ren 等^[30]使用 YOLOV3 算法检测蓝绿光信标,使用 P4P 算法对 AUV 进行定位,在近距离时使用 Aruco 码实现精确引导对接。Zhang 等^[31]使用事件相机,基于事件流,使用脉冲神经网络识别 4 灯信标,利用 P4P 算法进行 AUV 位置和姿态估计。Alla 等^[32]采用了迁移学习的目标检测算法,基于 YOLOV4 模型检测光信标,提高了不同水质下光信标的检测和跟踪性能。Zhang 等^[33]提出了基于 YOLOV5 的 4 灯光信标单目定位方法,融合注意力协调机制,提

高了光信标的识别率。同样基于该算法,Jyothi 等^[34]和 Lu 等^[35]完成水下回收站的目标检测,实现了 AUV 与回收站的相对位姿解算。Liu 等^[36-38]搭建了均匀分布的 8 灯信标回收站,利用卷积神经网络对信标的整体轮廓进行识别,使用顺时针分割处理非均匀散射,提高了光信标的识别率,如图 6(a)所示。随后,该课题组设计了两阶段的主动信标跟踪算法,实现了 AUV 对回收站光信标的识别定位^[39]。上述研究表明,基于深度学习的单目视觉导引方法能够显著提高灯信标的识别率。

表 1 单目视觉导引定位方法的研究现状

Table 1 Summary of status of monocular visual guidance and localization methods

Author	Experiment environment	Range	Frequency or time	Evaluating indicator	Computing platform	Ref. No
Park <i>et al.</i>	Pool	≤ 10 m	10–15 Hz	Average position error in camera coordinate: 200 pixel	PC/104+ standard single-board computer	[11–12]
Palomeras <i>et al.</i>	Tank	≤ 10 m	15 Hz	—	—	[13]
Fan <i>et al.</i>	Pool	≤ 2.8 m	—	Lateral positioning error is ≤ 0.2 m; longitudinal and vertical positioning errors are ≤ 0.15 m	—	[14]
Zhang <i>et al.</i>	Pool	—	—	Position error is 0.1 m	—	[16]
Trslic <i>et al.</i>	Sea	—	—	—	PC	[17]
Figueiredo <i>et al.</i>	Pool	≤ 2.5 m	20 Hz	Position error is 0.1 m	Raspberry Pi	[19]
Zhang <i>et al.</i>	Pool	≤ 10 m	30 s	—	—	[23]
Zhang <i>et al.</i>	Pool	≤ 5 m	50 s	—	—	[24]
Rannestad <i>et al.</i>	Sea	≤ 6 m	26–30 Hz	—	—	[20]
Lv <i>et al.</i>	Pool	≤ 10 m	—	Position error (X or Y) is ≤ 80 mm Position error (Z) ≤ 140 mm	—	[21]
Yan <i>et al.</i>	Pool	—	3.07 s	Matching accuracy is 100%	Intel® TM6600 CPU, 2.4 GHz, 1.0 GB EMS memory,	[22]
Zhang <i>et al.</i>	Pool	≤ 1.3 m	13.5709 ms	—	ThinkPad P52s machine	[23]
Lin <i>et al.</i>	pool	≤ 1.5 m	35 Hz	Error range is 3–20 cm	PC	[26]
Zhang <i>et al.</i>	Lake	≤ 6 m	—	Angle error is 1°	—	[27]
Zhou <i>et al.</i>	Tank	—	—	—	—	[40]
Wat <i>et al.</i>	Air	≤ 7 m	30 s	Average angle error is 10°	PC	[41]

(2) 图像信标

在图像信标的单目视觉导引领域,2017 年 Yahya 等^[42]利用深度学习技术识别回收站前后多个不同的图案标识,如图 6(b)所示,以确定 AUV 方向,解决了 AUV 导引过程中的信标丢失问题。Sans-Muntadas 等^[43]利用卷积神经网络将单目相机所拍摄到的图像与 AUV 位姿匹配,结合导航地图,实现 AUV 水下回收光学导引。尽管利用深度学习技术显著提高了图像信标的识别匹配成功率,但是该方法存在水下数据集少、难以获取且缺少对应真实值的问题。针对此,Johnson-

Roberson 等^[44]仿真生成了带有注释的水下数据集。Olson 等^[45]在三维渲染框架中,将真实水下图像投影到随机高度的景观上,合成了立体图像对和相应的场景深度图。di Stefano 等^[46-47]利用对抗生成网络和风格转换水下人工图像数据,基于多类型深度神经网络框架(如卷积神经网络、残差神经网络、Transformer 模型等)对图像信标进行识别。同时,为了降低 AUV 嵌入式设备的算力需求,利用知识蒸馏的方式,更加高效地训练卷积神经网络用于识别回收站^[48]。

综上所述,深度学习单目视觉导引定位方法研究

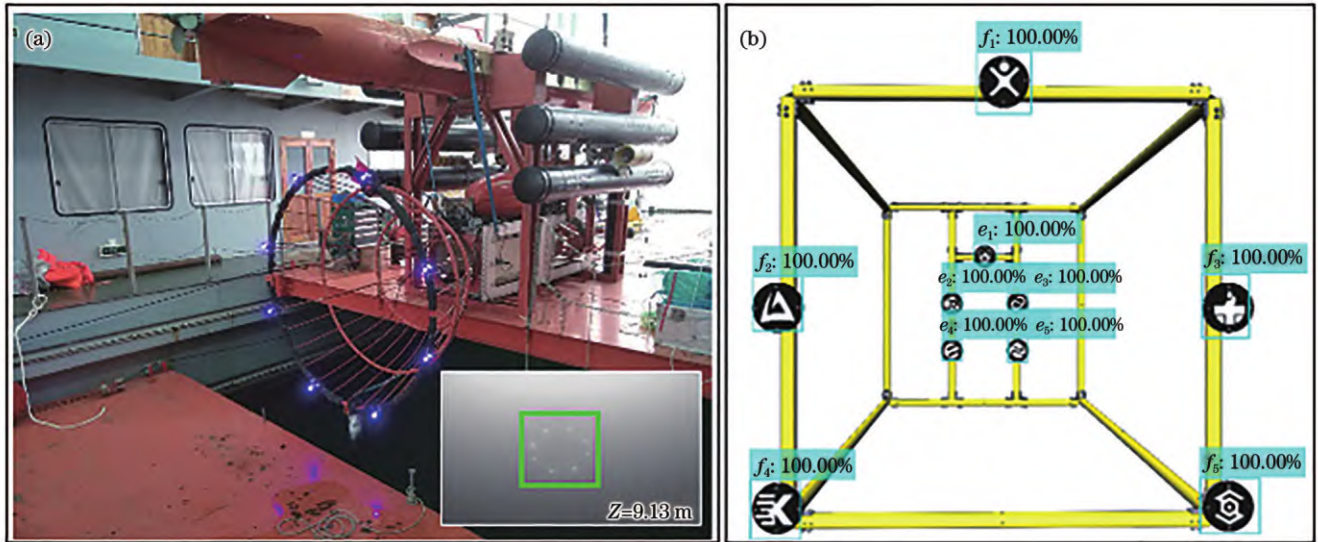


图6 水下导引标识物与光信标识别。(a)标识物分类识别;(b)8灯光信标识别

Fig. 6 Identification of underwater guidance markers and light beacons. (a) Classification and identification of markers; (b) identification of 8 lights

现状如表2所示,主要分为两类:一类是基于目标特征增强的位姿解析方法,这类方法通过构建水下光场补偿机制,提升光学信标在复杂水下光场条件下的特征辨识度,结合PnP算法实现AUV位姿解算;另一类是利用深度学习方法将拍摄到的信标画面与AUV位姿进行匹配,结合全球定位系统、惯性导航系统、多普勒

速度仪等传感器数据,确定AUV的相对位姿。在深度学习赋能下,目前AUV水下回收的单目视觉导引的最高定位精度可达3 cm,可靠性得到进一步提高。但是单目视觉方法获取的深度信息受限,且深度学习方法对算力要求相对较高,AUV与水下回收站的相对位姿解算频率较低。

表2 深度学习单目视觉导引定位方法的研究现状

Table 2 Summary of status of deep learning based monocular visual guidance and localization methods

Author	Method	Environment	Range	Frequency	Evaluating indicator	Computing platform	Ref. No
Singh <i>et al.</i>	YOLO	Pool	≤ 15 m	6 Hz	Average location error is 890 mm	NVIDIA Jetson TX2	[28]
Ren <i>et al.</i>	YOLOV3 & P4P	Pool	≤ 10 m	—	Average location error is 30 mm	—	[30]
Zhang <i>et al.</i>	SNN & P4P	Pool	—	—	—	—	[33]
Jyothi <i>et al.</i>	YOLOV5	Tank	≤ 10 m	—	Success rate is 95%	128 core NVIDIA Maxwell edgeArm cortex-A57 MP	[36]
Lu <i>et al.</i>	YOLOV5	Pool	≤ 15 m	—	—	Intel® Xeon® Silver 4110 CPU @2.10 GHz 16 GB GeForce RTX 3060 Ti 11 GB	[37]
Liu <i>et al.</i>	CNN	Sea	—	—	Success rate is 88%	—	[40]
Yahya <i>et al.</i>	CNN	Tank	≤ 8 m	60 Hz	Success rate is 95%	2.6 GHz Intel Core i7	[42]

2.3 双目视觉导引国内外研究现状

1) 双目视觉定位原理

AUV水下回收的双目视觉定位基于双目立体视觉技术,通过两个摄像头从不同视角采集图像,利用视差信息解算目标物体的三维空间位置,从而实现精准导引与对接。在双目视觉导引过程中,首先需要对两个相机拍摄的回收站导引灯图像进行中心点识别,获得回收站导引灯中心点的相机像素坐标^[49],如图7(a)

所示。双目视觉基于视差原理,利用成像设备从不同位置获取被测目标的两幅图像,计算图像对应点间的位置偏差,获取目标的三维几何信息^[50],进而确定目标回收站相对于AUV的位置,如图7(b)所示。

双目视觉定位的具体原理与过程如下:世界坐标系中被测目标表面上任一点 $P(X_w, Y_w, Z_w)$ 在左相机坐标系中像面上的投影点为 $p(u_l, v_l)$,在右相机坐标系中像面上的投影点为 $p(u_r, v_r)$ 。若仅用单目视觉

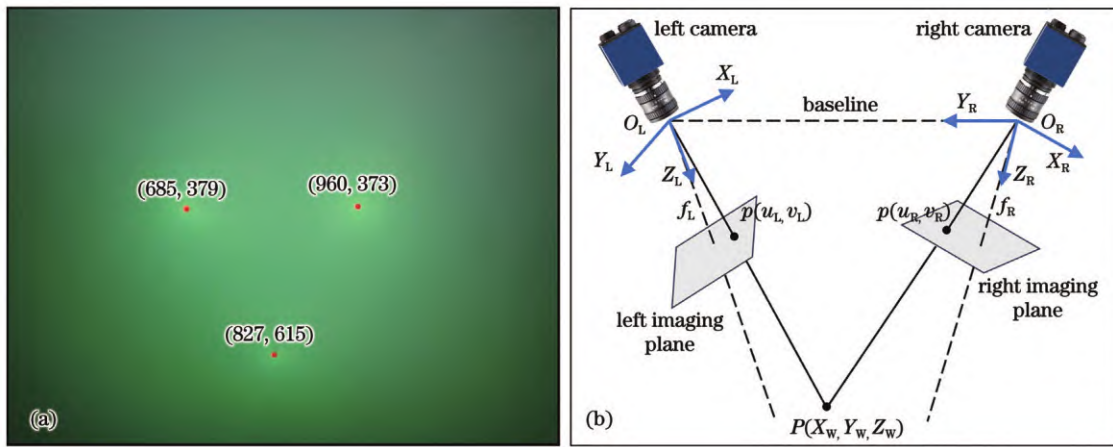


图7 水下双目视觉定位原理。(a)回收站导引灯中心点的相机像素坐标;(b)双目立体视觉原理

Fig. 7 Underwater binocular vision positioning principle. (a) Pixel coordinates of camera at guide light center in dock; (b) principle of binocular stereo vision

系统观察,只能判断点 $P(X_w, Y_w, Z_w)$ 位于左相机光心 O_L 与点 $p(u_L, v_L)$ 所构成的射线上,或位于右相机光心 O_R 与点 $p(u_R, v_R)$ 所构成的射线上。同时使用基于两个相机的双目视觉系统,确定像点 (u_L, v_L) 和 (u_R, v_R) 对应于同一空间特征点 $P(X_w, Y_w, Z_w)$,通过两个像点与相机光心所构成的两直线交点,计算出点 $P(X_w, Y_w, Z_w)$ 的确切位置,即可得到回收站导引灯 P 点的实时空间3维坐标。已知世界坐标系下3个导引灯中心点,利用3个导引灯分别在世界坐标系和相机坐标系下的坐标进行解算,得到相机坐标系和世界坐标系的旋转矩阵和平移矩阵,即AUV导引头相对于回收站导引灯的旋转矩阵和平移矩阵,进而获得AUV导引头段装置的航向角、俯仰角和偏转角,以及世界坐标系下的坐标,实现AUV与回收站的相对位姿解算^[51]。算法总体分为以下4步:1)确定相机内外参数,进行畸变校正,规定3个导引灯的世界坐标系坐标;2)对导引灯中心点进行特征识别,得到像素坐标;3)利用像素坐标进行距离测算,得到相机坐标系坐标;4)利用相机坐标系坐标和世界坐标系坐标进行解算,得到旋转矩阵和平移矩阵。

2) 相机标定

双目立体视觉测量可分为相机标定、特征提取、立体匹配和三维重建4个步骤^[52-54]。双目相机标定是通过双目相机系统的参数进行精确测量和校准,建立两个相机之间的几何关系模型,从而实现对三维场景的深度感知和空间位置计算。传统标定方法基于成像场景设计标定模型,在几何约束基础上选择最优方法解算相机参数,涵盖基于参照物、主动视觉、自标定等的相机标定方法。

(1) 基于参照物的相机标定方法

基于参照物的相机标定方法是已知参照物的形状尺寸等几何信息,通过提取靶标影像角点作为控制点,构建像素与空间坐标对应关系的方程组,最后优

化计算参数,按照参照物的维度分为一维直线标定、二维平面标定和三维立体标定。为了解决大视场下因多图像传感器缺乏重叠区域而难以校准的问题,Ge等^[55]提出了利用线性编码的一维目标对多图像传感器进行外部参数校准的方法。Chen等^[56]提出了含有多个不同周期的相移楔形光栅,基于旋转不变的局部二值编码模式进行全局编码,解决了二维目标标定方法需要相机捕获所有标定点,导引难以区分标定点的问题。在二维平面标定中,基于棋盘格标定板的张正友标定法^[57]是这类方法的经典代表,具有成像约束强、标定过程简单、算法鲁棒性高的特点。在3维立体标定方面,Zhang等^[58]提出了基于三角形基元的实心球体相机标定方法,该方法适用于高精度三维信息重建场景。

(2) 基于主动视觉的相机标定方法

基于主动视觉的相机标定方法通过精确控制相机或标定目标,获取不同视角的多幅图像,再利用可量化的运动约束确定相机内外参数。Hartley^[59]提出了一种相机自校准方法,从同一位置向不同方向获取至少3张图片,通过图像之间的点匹配分析进行标定校准。为了解决相机标定需要相机在3维空间内做任意方向平移运动的问题,杨长江等^[60]提出了仅需控制相机在同一平面做4组运动的相机标定方法。随后,吴福朝等^[61-62]设计了线性和非线性的相机标定算法,仅需一次平移运动和两次任意的缸体运动,就可以线性唯一地确定相机内参。

(3) 相机自标定方法

相机自标定方法利用多针图像对应点的几何一致性约束关系,标定相机的内外参数矩阵^[63-64],不需要设置标定物或精确控制相机位移,不依赖于拍摄场景的结构信息和运动信息^[65],包括直接求解Kruppa方程、基于绝对二次曲线和绝对二次曲面的方法^[66-67]、Pollefeys模约束标定^[68-69]、可变内参数下的分层逐步

标定^[70]等。Vasconcelos 等^[71]利用相机与其他被标定摄像机的成对关系进行内外标定,但镜像畸变对算法的影响仍有待验证。Kim 等^[72]基于多视图理论,利用热红外相机所拍摄的多方向图像实现相机标定。自标定方法本质上都基于绝对二次曲线欧氏不变性约束求解非线性方程组,实现未知场景下的相机标定^[73],但是该方法对噪声、初值选择较为敏感,且鲁棒性低^[74]。

目前,基于参照物、主动视觉和自标定的相机标定方法在双目视觉技术中得到广泛应用,但存在鲁棒性低、稳定性差及高度依赖初值等问题。这些方法通常针对特定场景,而不同类型相机的设计差异较大,需要单独进行建模标定^[75]。参数标定的稳定性与精度并非简单负相关,基于最小二乘法进行优化时易出现最优解偏离问题。此外,双目相机光学成像系统的噪声也会影响标定精度。大多数方法依赖部分参数设计值或线性求解初值,但在双目相机装配偏差大、噪声严重时,优化结果不稳定^[76]。

(4) 智能相机标定方法

深度神经网络能够准确描述输入、输出在欧氏空间的高维非线性映射,因此被应用于双目相机标定。以目标点像素坐标作为网络输入,对应的世界坐标作为理想输出,在信息正反向传播过程中修正隐藏层神经元权重,通过最小化实际与期望的输出误差,获得精确标定模型^[77-78]。Wen 等^[79]通过将前馈人工神经网络应用于传统标定方法中,校正传统模型误差,将精度提升了 2 倍。Jun 等^[80]提出了基于神经网络的简单灵活标定方法,解决了物像平面近似平行时的病态标定问题。吕朝辉等^[81]设计了基于神经网络的直接映射、镜头畸变校正、空间位置补偿方法,用于双目立体视觉相机标定。Jin 等^[82]在校准过程中引入相机与模板之间的角度,提出了 Harris 角点提取与深度神经网络相结合的单目相机标定方法。Ding^[83]将误差反向传播神经网络应用于 Kinect 深度相机标定,将目标角点信息作为训练集,利用双神经网络进行多方向像素标定,依据误差对称性建立误差补偿模型,减小深度测量误差。胡志新等^[84]利用改进遗传算法的误差反向传播神经网络完成双目相机标定,解决了标定收敛速度慢、易陷入局部最优解的问题。

智能相机标定方法作为一种隐式校正技术,展现出实时性和鲁棒性优势,但仍面临一些问题。智能相机标定方法严重依赖数据集,要求图像包含相机信息并具有一定的分辨率。然而,现有数据零散且不同相机之间难以通用,无法用于构建大规模、通用的标定数据集,这限制了该方法的泛化性。而数据集特征直接影响神经网络性能,因此标定结果的鲁棒性低、精度低。目前大多数研究依赖单一训练集,需要拟合多项任务,且相关性强的参数容易导致局部最优解问题。目前主流的相机标定方法如表 3 所示。

表 3 相机标定方法的研究现状

Table 3 Summary of status of camera calibration methods

Author	Complexity	Robust	Cost	Accuracy/ error	Ref. No
Ge <i>et al.</i>	Low	Low	Low	0.19 pixel	[55]
Chen <i>et al.</i>	Middle	High	Middle	0.05 pixel	[56]
Zhang <i>et al.</i>	Low	Middle	Low	1.3 pixel	[57]
Zhang <i>et al.</i>	High	Middle	High	0.03 cm	[58]
Hartley	Low	High	Low	0.5 pixel	[59]
Yang <i>et al.</i>	Low	High	Low	Low	[60]
Wu <i>et al.</i>	Middle	Middle	Middle	Low	[61]
Wu <i>et al.</i>	Low	Middle	Low	Low	[62]

3) 畸变校正

双目视觉成像系统的几何失真会造成图像像素位置或形状发生变化,即发生图像畸变^[85-86],显著影响定位精度。尤其在涉水视觉中,由于光的折射率突变、导流视窗曲面效应,以及前后向散射干扰,双目视觉成像系统易产生复合型几何畸变,主要表现为桶形畸变和枕形畸变^[87-89]。受相机与所拍摄物体之间角度的影响,这两类畸变将改变拍摄到的目标物体几何形状。此外,图像中还存在旋转畸变、波浪畸变、剪切畸变等^[90-95]。为了减小 AUV 航行阻力,AUV 舰首通常设计成流线型,双目相机透过导流视窗拍摄的图像容易产生不规则畸变。为了保证双目视觉定位精度,提高 AUV 水下回收成功率,亟须校正双目视觉的图像畸变。

目前图像畸变校正方法主要包括两类:一类方法是使用多视角图像,利用更加全面的场景信息提高校正精度,但是该类方法的实用性受限^[96-99];另一类方法利用消失点、纹理、轮廓等特征检测校正图像,但是该类方法主要取决于特征检测的精度^[100-105]。随着深度学习技术的发展,研究人员设计了基于模型感知预训练技术的 RDTR^[106]、DaFIR^[107]算法进行特征提取和畸变的映射回归,实现枕形畸变或桶形畸变的校正。DR-GAN^[108]和 SimFIR^[109]算法的出现,标志着畸变校正方法向端到端框架和自监督学习方向转变,这类方法可直接预测校正后的图像,但是仅针对特定畸变。为了解决图像中的多类型畸变,Rong 等^[110]设计了 GeoNet 用于预测畸变图像的畸变类型和位移场,将预测的畸变类型和已知的畸变模型相结合,生成一系列标准位移场,利用霍夫变换分析网络识别的位移场,确定可能缺失的部分,然后使用标准位移场拟合缺失的部分,以更准确地校正畸变。

涉水视觉的光场畸变校正是实现高精度双目视觉导引的核心前提,当前研究通过融合水下光场传输模型与深度表征学习,已突破静态环境亚像素级校正精度瓶颈。然而,湍流引起的瞬态波前畸变仍缺乏有效

补偿机制。针对此,亟须建立动态光场相位补偿模型,形成实时波前校正架构,开发轻量化神经网络架构,以实现嵌入式平台毫秒级实时校正,满足AUV动态回收场景的鲁棒性需求。

4) 国内外研究进展

国外方面,Myint等^[111-115]设计了模型匹配和单步遗传算法,利用水下无人遥控机器人(ROV)进行回收站三维姿态识别,如图8所示。为了进一步解决深海中光照变化对于图像识别的影响,该团队在模型匹配和单步遗传算法的基础上,添加了动态光适应系统,提

高了不同光照条件下的识别有效性^[116]。随后,该团队将单步遗传算法优化为多步遗传算法,进一步提高了导引定位精度,并在海底50 m处布放了矩形框回收站,开展了海试验证^[115-119]。由于海水浊度极大降低了信标识别精度,Nakamura等^[120-123]将之前的被动信标设计为能够发光的主动信标,提高了匹配适应值。基于上述研究,开展了不同浊度水体的水池试验与海试,研究结果证明该算法仅在FTU (Formazin Turbidity Unit) 小于10的水体中才有效^[124-128]。目前,国外双目视觉导引技术已较为成熟,并成功应用于AUV水下回收。

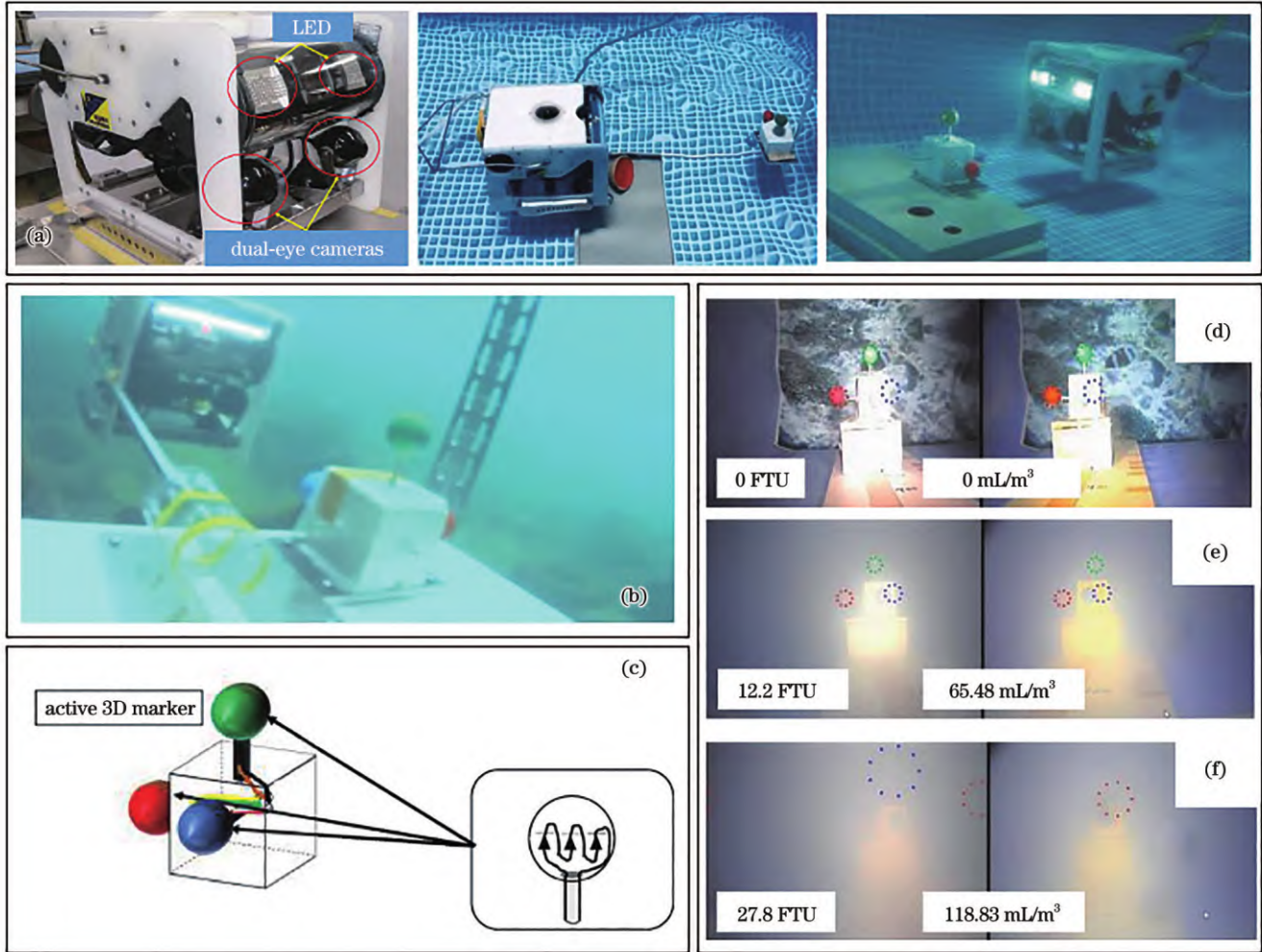


图8 Myint团队的双目视觉导引技术研究^[111-115]。(a) ROV双目视觉导引水池试验;(b) ROV双目视觉导引海试;(c)主动发光3D信标;(d)~(f)不同浊度水体中双目相机拍摄到的3D信标画面和识别结果

Fig. 8 Research on binocular vision guidance technology by Myint's team^[111-115]. (a) ROV binocular vision guidance pool test; (b) ROV binocular vision guidance sea trial; (c) active light-emitting 3D beacon; (d)~(f) 3D beacon images and recognition results captured by binocular cameras in water with different turbidity

国内方面,施小成等^[129]设计了纹理控制的金字塔标准化交叉关联算法,初步解决了AUV水下回收过程中导引标志的距离测量问题。随后,Li等^[130]采用单目视觉对应点匹配余弦定理推算AUV位姿,结合双目视觉定位算法,提高了AUV水下回收光学导引定位精度。Zhong等^[131]提出了基于自适应权重的特征提取方法,该方法在保证提取特征点准确的同时,显著降

低了AUV边缘算力要求。为了解决AUV双目视觉水下回收光学导引过程中单一相机受遮挡或者发生故障造成无法定位的问题,朱志鹏等^[132]提出了单双目切换的视觉定位方法,基于自适应阈值法区分图像中的光源与背景,通过中心加权质心检测法提取光源中心,并以光源中心点数量为判定条件选取单目或双目视觉定位方法。为了进一步提高双目视觉定位精度和鲁棒

性,徐硕等^[133-134]利用图像自适应二值化算法和伪光源剔除算法,实现了相机坐标系下回收站的三维定位。随后,该团队提出了将自适应算法和几何穿线法相结合的光信标检测技术,显著提高了 AUV 双目视觉水下回收光学导引的定位精度。

本课题组基于水下光场调控理论,构建了融合偏振成像与动态曝光的双目视觉光学导引系统,实现了复杂涉水视觉环境中的高精度目标定位。在光场调控层面,提出了非对称导引灯组结构,建立水下光场分布模型,以同步获取目标距离、方位角及三维姿态角等多元导引定位参量,实现导引灯组的精准空间定位^[135]。针对水下光场散射效应造成的成像对比度下降问题,

建立了偏振差分成像的水下目标增强模型。利用水体散射光与目标反射光的偏振态正交性,构建 Stokes 矢量空间的水下光场解算体系,通过多偏振通道信息融合,有效抑制散射噪声,提升目标对比度^[136][图 9(a)]。针对水下光场能量呈指数衰减的特性,设计了基于光场强度分布的自适应曝光控制策略,通过建立距离-曝光时间标定曲线在 1~9 m 作用范围内实现动态曝光补偿[图 9(b)],既避免近场过曝导致的光晕效应,又保证远场目标的信噪比阈值,形成了覆盖全作用距离的光场能量均衡探测方案^[137]。该研究通过光场多参量联合优化,显著提升了水下光学导引系统的环境适应性和定位精度。

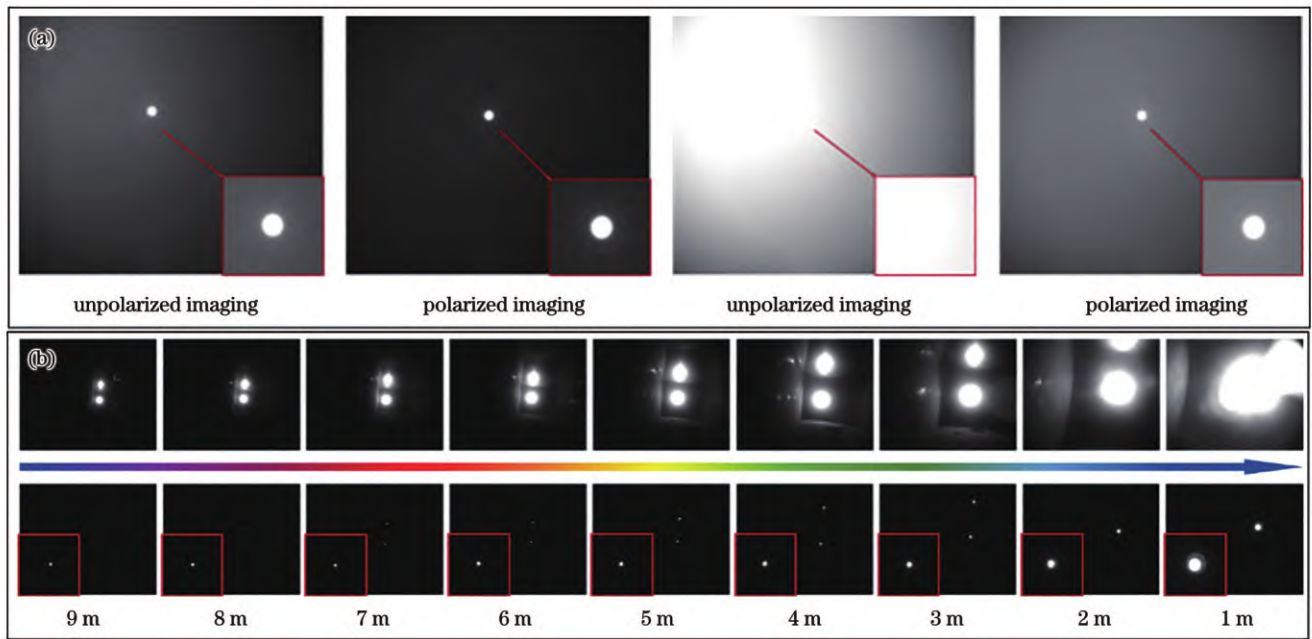


图 9 距离-曝光时间标定图和固定曝光与自动曝光效果对比。(a)水下常规成像与偏振成像对比;(b)水下固定曝光与自动曝光效果对比

Fig. 9 Distance-exposure time calibration chart and effectiveness comparison between fixed exposure and automatic exposure.

(a) Comparison between underwater unpolarized imaging and polarized imaging; (b) comparison of effects between underwater fixed exposure and automatic exposure

为验证系统性能,本课题组搭建了水下光场可控实验平台,将双目导引系统集成于 AUV 导引头,形成闭环控制架构。在静态水池中,通过高精度位移平台建立光场基准坐标系,实施了多距离点的光场强度标定与几何校准,完成了不同距离处的测量误差分析工作。开展了 AUV 水下回收光学导引试验,测得了导引过程中的目标距离、方位角及姿态角等信息,实现 AUV 水下回收光学导引对接,验证了双目视觉导引系统的有效性,如图 10(a)~(d)所示。图 10(e)~(i)显示了水池静态场景下进行导引系统精度指标验证的过程,定位误差小于 3 cm@2 m,方位角误差小于 0.12°,姿态角误差小于 0.3°。该研究通过光场特性与视觉算法的深度融合,为 AUV 水下回收光学导引的精准对接提供了创新性技术解决方案。

基于水下光场传输特性的双目视觉导引方法,显著提高了信标的图像识别率、增加了识别距离,双目视觉导引使 AUV 具备了测距能力,能够获得更多的回收站相对位姿信息,解算结果更加精确,显著提升了复杂涉水视觉环境中的信标识别性能与空间解析能力。如表 4 所示,已报道的定位精度误差为 10 mm 左右,最远水下回收光学导引距离达 30 m。另外,尽管传统双目视觉方法不需要高算力,但每次解算速度需要几毫秒到几百毫秒,解算速度有待进一步提高。目前,亟须进行轻量化光场解析模型与异构计算架构的协同优化,以突破毫秒级实时解算的工程瓶颈。

综上所述,单目视觉导引和双目视觉导引在 AUV 水下回收光学导引过程中的共同之处在于:通过多样化标识分布、颜色、纹理设计,或者使用主动发光信标,

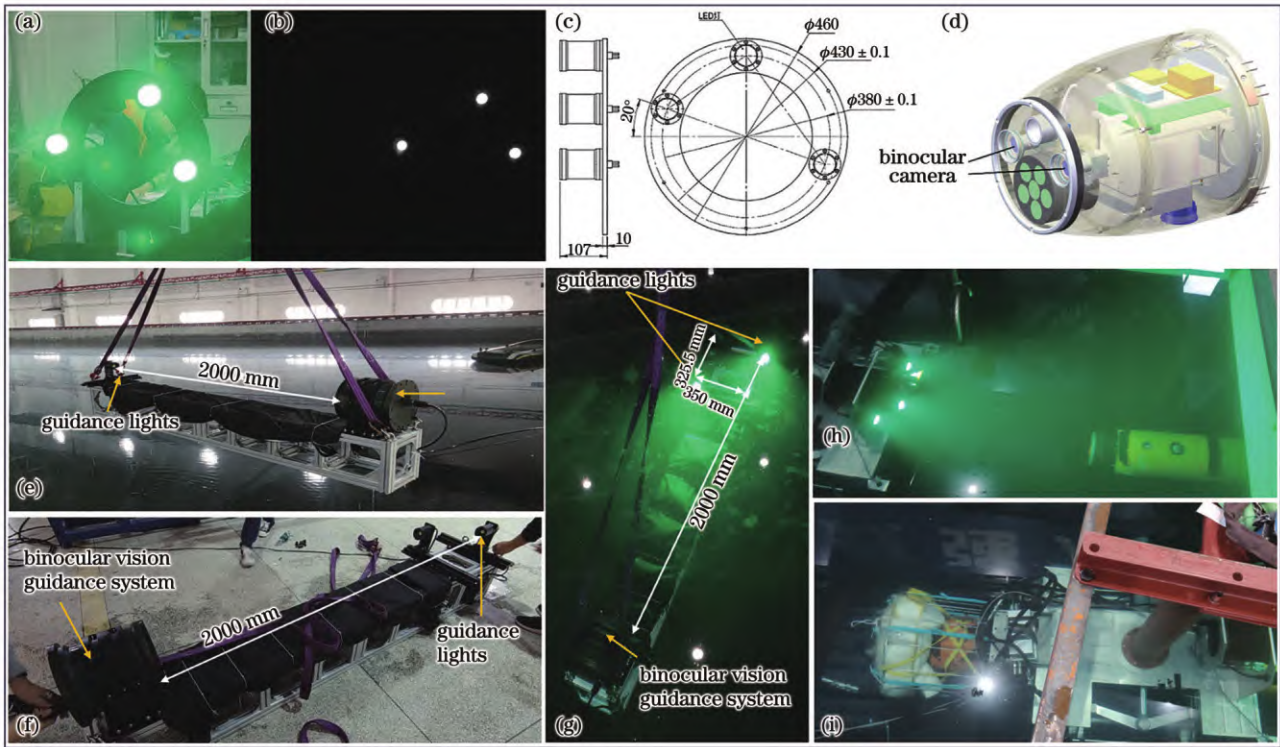


图 10 AUV 水下回收光学定位与导引试验。(a)导引灯陆上测试;(b)双目相机识别导引灯;(c)导引灯结构设计;(d)AUV 导引头设计图;(e)光学导引定位支架;(f)陆上姿态角与方位角测试;(g)水下姿态角与方位角测试;(h)(i)水下 AUV 回收试验

Fig. 10 AUV underwater recovery positioning and optical guidance tests. (a) Onshore test of guidance lights; (b) recognition of guidance lights by binocular cameras; (c) guidance light structure design; (d) AUV guidance head design diagram; (e) optical guidance positioning bracket; (f) onshore attitude and azimuth test; (g) underwater attitude and azimuth test; (h)(i) underwater AUV docking tests

表 4 双目视觉导引定位方法的研究现状

Table 4 Summary of status of binocular vision guidance and localization methods

Author	Environment	Range	Frequency/ time	Evaluating indicator	Computing platform	Ref. No
Myint <i>et al.</i>	Pool	(0.35 m, 2 m)	165 ms	Average location error: 40 mm	ROV	[111]
Myint <i>et al.</i>	Sea trial 50 m depth	(0.35 m, 15 m)	33 ms	Average location error: 10 mm	ROV	[115]
Shi <i>et al.</i>	Lab	—	258 ms	Average location error: 5 m	Pentium 42 GHz	[129]
Li <i>et al.</i>	Pool	≤ 30 m	0.5 s	Average location error: 300 mm	WL-3 mini-micro AUV	[130]
Zhu <i>et al.</i>	Pool	≤ 10 m	7.5 ms	Average location error (@10 m): 140 mm Average location error (@5 m): 92 mm Average location error (@3 m): 53 mm	—	[132]
Xu <i>et al.</i>	Pool & lake	Pool: ≤ 20 m Lake: ≤ 7 m	21 ms	Average location error (X, pool): 14 mm Average location error (Y, pool): 7.8 mm Average location error (X, lake): 2.8 mm Average location error (Y, lake): 3.2 mm	Intel i7-8750 processor and an NVIDIA GTX 1060 graphics	[133– 134]
Guo <i>et al.</i>	Pool	≤ 2.81 m	—	Average location error: 200 mm	Amphibious spherical robot	[138]
Hsu <i>et al.</i>	Tank	≤ 0.95 m	—	—	—	[139]

结合特征学习与识别算法进行图像处理,提高信标的识别率、增加识别距离。二者的区别在于:在AUV位姿解算时,两种方法获取的信息不同。单目视觉导引通过已知标识点之间的距离和图像中标识点之间的几何距离关系,确定AUV与回收站之间的距离,但距离测算的精度低,通常需要结合全球定位系统、惯性导航系统、多普勒速度仪或AUV运动状态方程,辅助确定AUV相对位姿。双目视觉导引通过两个共面光轴相机对同一标识点的视差完成距离测算,获取其在水下坐标系中的三维坐标,有助于提高后续位姿解算精度。针对此,基于图像传感器的光学导引方法,未来亟须:构建水下光场物理模型与神经架构的协同优化策略,实现动态湍流引起的波前畸变补偿;探索多模态传感器融合范式,以将水下成像、雷达点云与双目深度等信息进行贝叶斯融合,突破单一视觉模态的视距与精度矛盾;开发轻量化计算架构,以提升位姿解算频率,满足AUV水下动态回收场景的实时性需求。

3 基于位置探测器的光学导引

3.1 原理与挑战

基于位置探测器的光学导引原理是:利用光学信号处理方法,确定光斑在位置探测器光敏面上的质心位置,进而实现目标精确定位的技术。该技术主要应

用在两个方面:一方面,在激光照射待测目标后,位置探测器通过检测反射光斑在光敏面的位置变化,计算其质心移动,从而精确计算被测目标的线位移或角位移;另一方面,位置探测器解算出的光斑质心位置,为电机或快速反射镜等执行机构提供脱靶量,用于实现实时动态跟踪^[140-141]。这两方面应用的核心都是对照射到位置探测器的光斑质心进行实时测量,因此光电位置探测器的光斑位置检测性能对测量精度十分重要。常用的光电位置探测器包括位置敏感探测器(PSD)、电荷耦合器件(CCD)和四象限光电探测器(QD)^[142-143]。其中,四象限光电探测器是通过将一个光电探测器按照笛卡儿坐标系分成4块(对应探测器的4个象限)并利用光刻技术制备的。当入射光斑落入光敏面的不同位置时,各象限会输出不同强度的电信号,通过处理4个象限输出的信号强度,可以确定光敏面上入射光斑的质心位置^[144],该方法具有检测分辨率高、响应时间短等优势^[145-146],常用于光电测控通信领域^[147-149]。

基于位置探测器的水下回收光学导引方法通过检测回收站导引灯光斑位置信息变化,解算AUV与回收站导引灯的相对位置信息,目前用于AUV水下回收光学导引的位置探测器主要是四象限光电探测器,其原理流程图如图11所示。

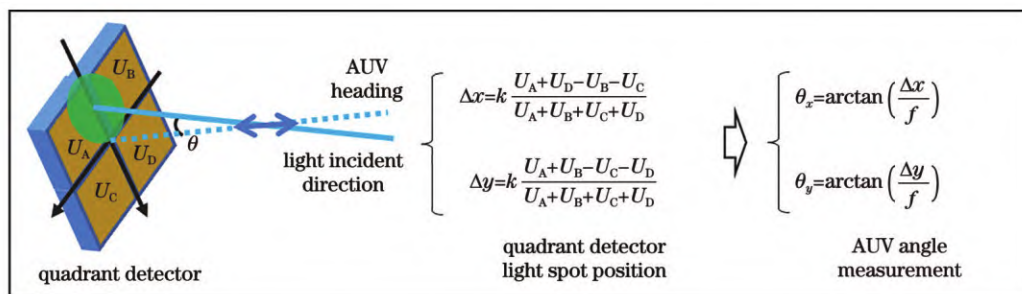


图11 四象限光电探测器工作原理

Fig. 11 Principle of four-quadrant photodetector

3.2 四象限光电探测器的光学导引国内外研究现状

位置探测器应用研究起源于20世纪60年代末^[150],最早应用于激光制导。位置探测器应用于AUV水下回收光学导引的研究最早可以追溯到1997年,Cowen等^[151]使用四象限光电探测器对回收站发出的导引光源脉冲信号进行检测,为AUV控制系统提供回收站位置指引。该光学导引系统在水质浑浊导致视距不足2 m的情况下,对导引光信标的最大可靠捕获距离达15 m;而在水质清澈的条件下,对导引光信标的可靠捕获距离可达28 m。2017年美国海军信息战中心太平洋分部将基于四象限光电探测器的光学导引方法首次应用于Remus 600型水下无人潜航器,利用数学解析方法对AUV惯性位置进行坐标转换和光学测角修正,以为控制系统提供位置信息,并进行

了水下回收光学导引海试,对接成功率为41.7%,如图12所示。海试结果表明,四象限光电探测器能够支持高频信号反馈,有利于AUV水下高速回收对接,但是由于单导引灯无法对AUV位置进行精确解算,且AUV在抵达距离回收站约3 m位置后,已处于探测器信号盲区,因此AUV水下回收光学导引的成功率较低^[152-153]。

为了进一步提高AUV与水下回收站三维空间相对位置的解算精度,2021年本课题组提出了多导引光源的四象限探测光学导引方法,利用四象限光电探测器的能量中心识别原理,通过各象限形成的光电信号强度比例估计光斑中心的偏移量,进而通过简单的几何关系获得光源入射方向角^[154-155]。通过数值迭代求解方式解算AUV与水下回收站的相对位置,该方法

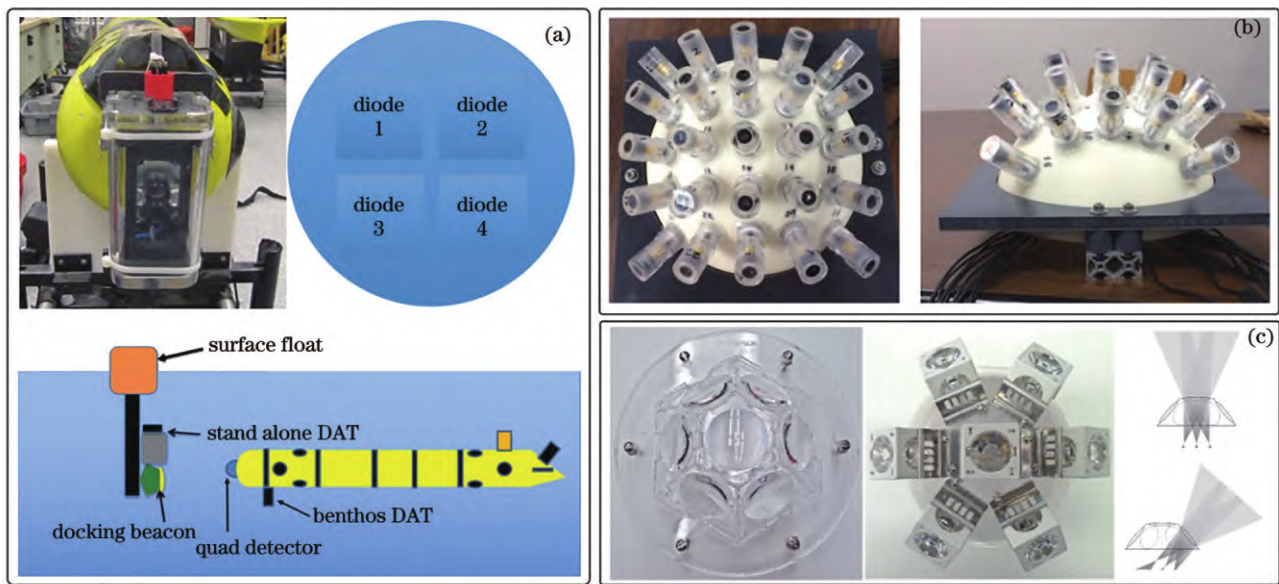


图 12 基于光电探测阵列的光学导引探测器件。(a)安装四象限光电探测器的 REMUS 600;(b)半球形 5×5 光学探测器阵列;
(c) 7 个光电探测器及透镜组成半球形阵列

Fig. 12 Optical guidance detector elements based on photoelectric detection arrays. (a) REMUS 600 with four-quadrant photodetectors;
(b) hemispherical 5×5 optical detector array; (c) 7 photodetectors and lenses form a hemispherical array

具有高频光学编码探测、位置解算准确、算力功耗需求小的优点,但存在视场角小、迭代算法复杂且易失败等问题,难以对所有的几何关系进行准确建模,导致该方法频繁出现无解和误解的问题,仅在部分情况下能够得出正确结果。

3.3 多象限光电探测器的光学导引国内外研究现状

1) 国外研究进展

为了扩大探测范围,研究人员基于四象限光电探测器原理,开发出了多象限光电探测器。2001 年 Evans 等^[156]基于 SWIMMER 水下无人平台,搭载了半球形 5×5 光电探测阵列,光学导引距离为 5 m,对应的定位精度为 0.2 m,探测角度为 10° ,频率为 1 Hz。2017 年 Eren 等^[157]提出了用于 AUV 水下回收光学导引的半球形 5×5 光电探测器阵列传感器系统,使用较少的光学元件有效增大了 AUV 的探测视场。由于位置探测器具有探测频率高的优势,水下通信导引一体化的研究方案随之被提出。2012 年 Simpson 等^[158]设计了由 7 个光电探测器及透镜组成的半球形阵列系统,使用基于数据特征的遍历匹配算法估计角度,论证了光学通信导引一体化的可行性。2022 年 Chen 等^[159]设计了由 7 个光电探测器及透镜组成的平面阵列的通信和定位系统,通过估计入射光的立体角,确定目标的二维位置。该研究证明通过提高导引光源维度提高位置解算精度的可行性。

2) 国内研究进展

针对水下光场动态感知与高精度智能导引定位需求,本课题组创新性地提出了基于智能光电探测导引硬件与算法一体化设计的多象限测角光学导引定位方

法^[160],采用硬件-算法协同优化架构,通过多象限光电传感器阵列与深度学习技术的深度融合,实现了 AUV 水下回收的大视场高精度光学导引定位。

在硬件架构层面,为了实现水下回收光学导引大视角的高精度定位,该体系以 8×8 阵列多象限光电探测器为核心感知单元,使用 64 个性能完全相同的硅光电二极管或 PIN/APD 探测器,按直角坐标排列成 8×8 排布的多象限。当光斑经过光学系统照射到探测器表面时,光斑会落在任意相邻的 4 个阵元,这 4 个阵元构成一个四象限光电探测器,各象限接收的光强与其覆盖的光斑面积成正比,产生对应的光电流信号。若光斑中心与探测器中心重合,4 个象限的输出电信号相等,表明无偏移;若光斑偏离中心,各象限信号强度产生差异,通过和差算法将 4 个象限的电流差值转换为归一化偏移量,从而精确计算光斑质心的位置坐标。相较于传统的四象限光电探测器,该体系从硬件层面显著增大了 AUV 的光学导引在导引灯的几何排布层面,使用正三角形的导引灯排布方式,为算法解算提供了无歧义的几何约束,确保导引定位偏角和位置的一一对应关系,这是算法层面产生唯一位置真值解的重要前提条件。

在算法层面,本课题组突破性地融合深度神经网络,构建了 AUV 与回收站导引灯物理几何模型的多分支回归网络,基于 AUV 水下回收光学导引过程建模,并采用物理仿真策略生成了偏角和位置一一对应的数据集,设计了输入对应单一输出的多分支结构位置解算回归网络,建立了多维度空间定位精度约束目标函数。经过仿真实验验证,发现本课题组提出的多

分支回归网络 AUV 光学导引定位方法显著提升了位置解算精度和速度,克服了现有算法无解或频繁误解的难题,在导引灯正三角形排布的情况下,使用多分支网络定位方法,获得了更接近于真值的唯一近似解,有效解决了基于图像探测器光学导引的计算复杂度高,基于位置探测器光学导引的视场角小、迭代算法易出错、解算速度慢的问题。

在系统工程海试验证阶段,本课题组在动态水场环境中构建了多参数耦合的光场基准坐标系,通过全球定位系统提供真值参考。将多分支回归网络的光学导引定位方法部署到了 ROV 上,通过海试验证了该方法的光学导引定位位置解算精度,证明该方法能够满足 AUV 导引定位要求的视场角^[6],如图 13 所示。本课题组在自然资源部北海海洋技术中心海港港池开展了海试验证,在港池内布置正三角形排布的导引灯组,将多象限光电探测器部署到 ROV 上,将全球定位系统终端通过刚性连接固定在 ROV 上且始终露出水面,以获取已知误差范围的位置真实值。在海试过程中,ROV 在港池内按照一定轨迹向导引灯组运动,多象限光电

探测器实时接收光信号并将其解算为导引灯偏角数据,与全球定位系统获取的 ROV 实际位置坐标一一对应,作为位置解算精度的参考。同时,在整个过程中,使用测试软件进行可视化记录。为了验证多分支网络光学导引定位方法的位置解算精度,首先对所有偏角数据和全球定位系统的坐标数据进行预处理,包括去噪、坐标数据与姿态筛选,以解算出预测位置坐标。对绝对坐标误差进行计算和统计,得到 AUV 的位置坐标解算速度为 5.650 ms/次,坐标解算误差的均值为 58.292 mm@0.8~20 m,标准差为 43.347 mm@0.8~20 m,最佳平均坐标解算精度达 7.107 mm@2~3 m。海试结果表明,利用基于多象限测角的多分支网络光学导引定位方法得到的位置解算速度和精度能够满足 AUV 末端水下回收光学导引需求。与现有光学导引方法相比,利用基于多象限测角的多分支网络光学导引定位方法的解算速度更快,算力功耗需求更低,能量消耗更少,为 AUV 完成能源补充、数据传输、指令下达提供了可靠的水下回收光学导引方法,为复杂涉水视觉环境中 AUV 精准回收提供了新的技术范式。

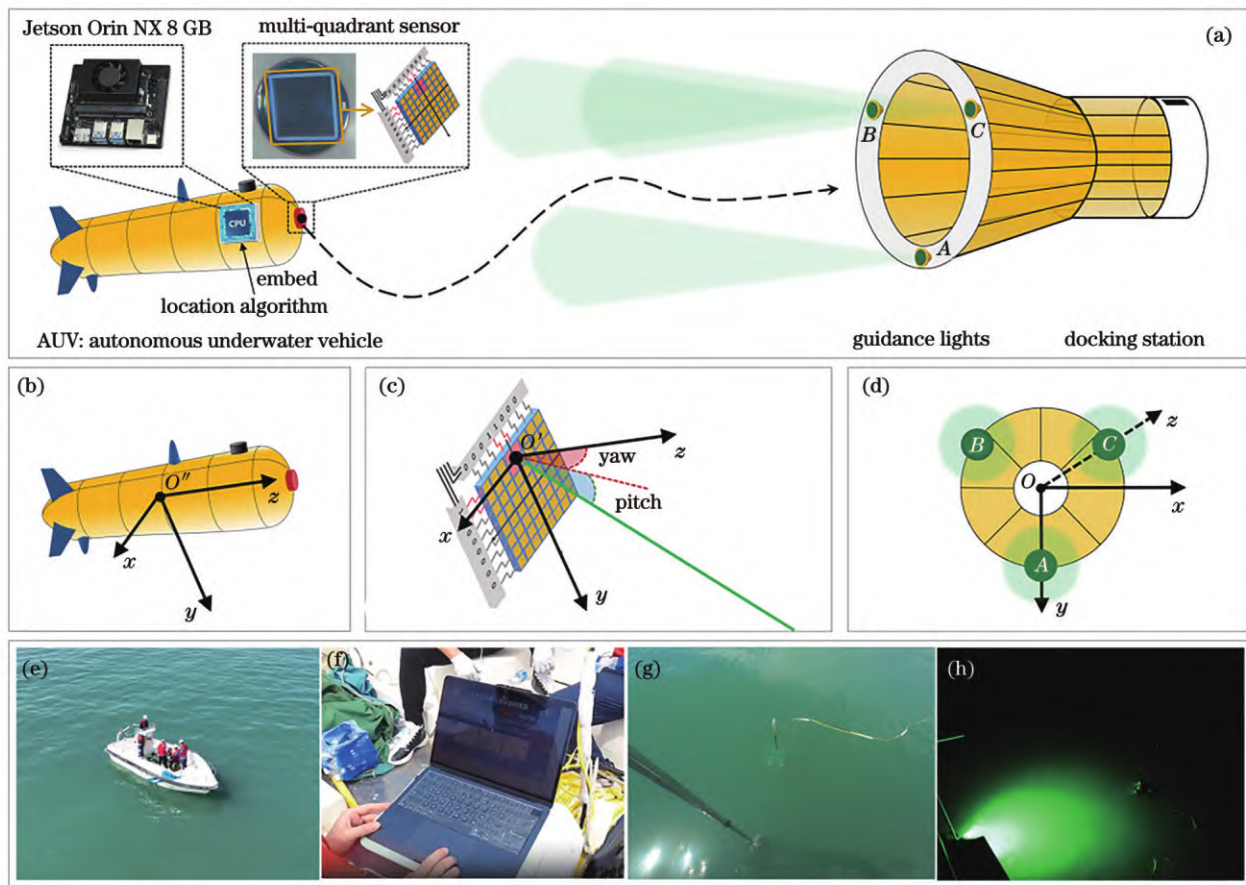


图 13 AUV 导引定位系统示意图。(a) AUV 水下回收光学导引过程;(b) AUV 坐标系图;(c)导引灯偏角几何定义;(d)导引灯坐标系;(e) ROV 远端布防;(f)实时解算定位;(g)白天海试导引定位;(h)夜间海试导引定位

Fig. 13 Diagrams of AUV guidance location system. (a) AUV underwater recovery optical guidance process; (b) AUV coordinate system diagram; (c) guidance light deflection angle geometric definition; (d) guidance light coordinate system diagram; (e) ROV remote arming; (f) real-time solution and location; (g) daytime sea trial guidance location; (h) night sea trial guidance location

基于水下光场传输的物理几何模型构建的多象限探测光学导引方法,通过融合光学物理几何模型与深度表征学习机制,实现了 AUV 水下回收毫米级定位精度与毫秒级实时解算的突破性进展,具有高频响应和低算力需求的特点,已成为 AUV 水下回收光学导引的核心方案之一。如表 5 所示,本课题组已实现定位精度误差小于 8 mm、解算速度达 5.65 ms/次的技术指标,相比于已报道的技术指标具备显著优越性,但仍然需要在环境适应性、多源融合及通信导引协同方向持续突破。针对此,基于

位置探测器的光学导引方法,未来亟须构建水下成像、雷达点云与光学深度等信息的多模态融合框架,通过联邦卡尔曼滤波器实现多源信息最优估计,显著减小远距离定位误差。研发轻量化的边缘计算框架,通过将高效网络算法嵌入光场传输方程实现实时波前补偿,同时开发量子点光电探测器,形成智能光电协同架构,以提升系统在浑浊水体下的目标特征提取能力。智能算法与智能光电器件的进一步结合,将推动该技术向更高精度、更高鲁棒性和更低成本方向发展。

表 5 基于位置探测器的光学导引定位方法研究现状

Table 5 Summary of status of optically guided localization methods based on position detectors

Author	Environment	Range	Frequency/ time	Evaluating indicator	Computing platform	Ref. No
Martin <i>et al.</i>	Sea trial	≤ 10 m	—	Recovery success rate: 48.40%	—	[152]
Fletcher <i>et al.</i>	Sea trial	≤ 10 m	—	Recovery success rate: 41.70%	i7-3770 CPU	[153]
Evans <i>et al.</i>	Pool	≤ 5 m	1 Hz	Positioning accuracy: 0.2 m; detection angle: 10°	—	[156]
Eren <i>et al.</i>	Pool	≤ 8.5 m	5 Hz	Positioning accuracy of x , y and z axes: 0.13 m, 0.13 m and 0.18 m; pitch and yaw errors: $<5^\circ$	—	[157]
Ours	Simulation & Sea trial	0.8— 20 m	5.560 ms	0.8–20 m absolute coordinate error: 58.292 mm; ≤ 3 m absolute coordinate error: 6.647 mm	ARM Cortex-A78 AE v8.2 64 CPU	[160]

4 总结与展望

AUV 水下回收光学导引作为 AUV 安全、连续、有效作业的关键技术,近年来受到了国内外学者的关注。本文分析了基于图像传感器的光学导引技术和基于位置探测器的光学导引技术。基于图像传感器的光学导引方法的信息获取直观、定位精度高,是目前 AUV 水下回收光学导引的主流方法;基于位置探测器的光学导引方法的探测速率高、定位精度高,具备通信探测一体化的潜力。本文对这两类 AUV 水下回收光学导引的理论技术与应用领域进行了综述,分析了当前 AUV 水下回收光学导引领域的难点与问题,最后以任务需求、技术发展为导向,对该领域的发展方向与趋势进行了总结分析。

针对复杂涉水视觉环境中 AUV 水下回收光学导引的实际任务需求,尽管目前设计了光学导引定位算法,构建了相关数据集,但是基准真值数据的采集较难。另外,在 AUV 水下回收光学导引过程中,AUV 控制系统存在延迟、噪声干扰,以及 AUV 自身存在姿态抖动,而基于图像传感器的光学导引方法的探测频率较低(24~120 frame/s),且数据量大,处理速度慢。因此,迫切需提高 AUV 水下回收光学导引的精度、速度、鲁棒性与智能化水平。面向未来海洋智能光电装

备体系化发展需求,AUV 水下回收光学导引技术的革新亟待三个核心方向的突破:智能算法架构升级、通信感知范式重构,以及多模态协同感知体系构建。在海洋作业智能化进程加速的背景下,亟须构建具备自主决策能力、高频实时响应特性和多源融合优势的新一代水下智能导引系统,为深远海无人装备集群化作业提供关键技术支撑。

首要突破方向在于构建基于人工智能范式的水下光学导引技术。针对传统视觉算法在动态水下环境中存在的计算延迟与精度瓶颈,需要革新传统算法架构,深度融合深度学习技术特别是大模型处理方法,突破高速度、高稳定、高准确率的水下光学智能导引技术。开发水下目标信号探测与处理的轻量化模型,探索适配的模型量化技术,以动态平衡模型计算复杂度与特征表征能力,突破边缘计算设备的模型压缩极限,实现大模型在边缘设备的离线部署,满足复杂指令规划与高级任务执行需求,加强 AUV 自主决策与自适应能力,在动态变化的海洋环境中,高速度、高稳定、高精度地完成水下回收光学导引。

关键发展路径在于建立通信导引一体化的水下光学导引系统。水体对电磁波的强吸收,导致传统无线通信技术难以应用于水下;水声通信技术具备水下远距离传输能力,但存在带宽较小、通信延迟较大等问

题,而光通信技术在近距离具备大带宽和小延迟的特点。面向未来AUV水下通信导引一体化的发展趋势,传统图像传感器无法提供高频编码能力,传统位置探测器的载波能力有限和灵敏度较低,亟须研发高频、高灵敏度的多象限光电探测器,以提升微弱光信号的捕获灵敏度。创新性地采用脉冲位置调制与正交频分复用结合的混合调制方案,利用导引光脉冲的时隙分配实现通信信道复用,使定位精度与通信容量同步跃升,形成水下高速通信与精确导引的协同机制。构建导引灯组与探测阵列的探测链路,在回收站部署蓝绿光发射阵列,通过在AUV上搭载多象限光电探测器,可在实现水下精准定位的同时,完成高速率通信。

最终目标在于打造多源异构数据深度融合的水下光学导引体系。通过融合AUV多类型传感器数据,如全球定位系统、惯性导航系统、多普勒速度仪等,实时获取AUV水下位姿信息,以显著提高光学导引定位精度。设计分层联邦滤波系统融合惯性导航系统与多普勒速度仪数据,集成光学导引与超短基线声学定位,采用卡尔曼滤波实现高频位姿预测,并补偿累计误差。通过设计时空对齐的联邦学习框架,将光学图像、惯性导航、多普勒测速、声学定位等多源信息映射至统一特征空间,利用图神经网络建立跨模态关联模型,提升水下视觉特征的提取准确性与目标感知精度,实时精准获取回收站的信标信息。在硬件层面,通过构建嵌入式异构计算平台,实现传感器数据并行预处理与GPU加速深度学习推理,从采集到决策输出的端到端延迟控制在毫秒量级,满足AUV水下回收光学导引全流程实时性要求。

上述技术、系统与体系的突破,将推动AUV水下回收光学导引技术从单一功能模块向智能系统生态演进,形成AUV水下回收智能光学导引新范式,为构建涉水智能光电装备体系奠定核心技术基础,支撑未来海洋科考、资源勘探、水下安防等领域的重大应用需求。

参 考 文 献

- [1] 李学龙. 涉水光学[J]. 中国科学: 信息科学, 2024, 54(2): 227-280.
Li X L. Water-related optics[J]. Scientia Sinica (Informationis), 2024, 54(2): 227-280.
- [2] 李学龙. 涉水视觉[J]. 电子学报, 2024, 52(4): 1041-1082.
Li X L. Water-related vision[J]. Acta Electronica Sinica, 2024, 52(4): 1041-1082.
- [3] Sun Z, Li X L. Water-related optical imaging: from algorithm to hardware[J]. Science China Technological Sciences, 2024, 68(1): 1100401.
- [4] 宋保维, 潘光, 张立川, 等. 自主水下航行器发展趋势及关键技术[J]. 中国舰船研究, 2022, 17(5): 27-44.
Song B W, Pan G, Zhang L C, et al. Development trend and key technologies of autonomous underwater vehicles[J]. Chinese Journal of Ship Research, 2022, 17(5): 27-44.
- [5] 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.
- [6] 燕奎臣, 吴利红. AUV水下对接关键技术研究[J]. 机器人, 2007, 29(3): 267-273.
Yan K C, Wu L H. A survey on the key technologies for underwater AUV docking[J]. Robot, 2007, 29(3): 267-273.
- [7] 朱若澜, 李晨, 杨敏, 等. 自主水下航行器多分支网络光学导引定位方法[J]. 中国科学: 信息科学, 2025, 55(1): 110-128.
Zhu R L, Li C, Yang M, et al. Multi-branch network-based optical guidance and positioning method for autonomous underwater vehicles[J]. Scientia Sinica (Informationis), 2025, 55(1): 110-128.
- [8] 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]. Torpedo Technology, 2008, 16(6): 10-14.
- [9] Lee P M, Jeon B H, Lee C M. A docking and control system for an autonomous underwater vehicle[C]//OCEANS '02 MTS/IEEE, October 29-31, 2002, Biloxi, MI, USA. New York: IEEE Press, 2002: 1609-1614.
- [10] Lee P M, Jeon B H, Kim S M. Visual servoing for underwater docking of an autonomous underwater vehicle with one camera[C]//Oceans 2003, September 22-26, 2003, San Diego, CA, USA. New York: IEEE Press, 2003: 677-682.
- [11] Park J Y, Jun B H, Lee P M, et al. Experiment on underwater docking of an autonomous underwater vehicle 'ISiMI' using optical terminal guidance[C]//OCEANS 2007-Europe, June 18-21, 2007, Aberdeen, UK. New York: IEEE Press, 2007: 1-6.
- [12] Park J Y, Jun B H, Lee P M, et al. Experiments on vision guided docking of an autonomous underwater vehicle using one camera[J]. Ocean Engineering, 2009, 36(1): 48-61.
- [13] Palomeras N, Vallicrosa G, Mallios A, et al. AUV homing and docking for remote operations[J]. Ocean Engineering, 2018, 154: 106-120.
- [14] 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.
- [15] Yan Z P, Gong P, Zhang W, et al. Autonomous underwater vehicle vision guided docking experiments based on L-shaped light array[J]. IEEE Access, 2019, 7: 72567-72576.
- [16] Zhang W, Li Z X, Gong P, et al. Visual location method based on asymmetric guiding light array in UUV recovery progress[C]//2019 IEEE International Conference on Robotics and Biomimetics (ROBIO), December 6-8, 2019, Dali, China. New York: IEEE Press, 2019: 2671-2675.
- [17] Trslie P, Rossi M, Robinson L, et al. Vision based autonomous docking for work class ROVs[J]. Ocean Engineering, 2020, 196: 106840.
- [18] Su X, Xiang X B, Dong D L, et al. Visual LOS guided docking of over-actuated underwater vehicle[C]//Global Oceans 2020: Singapore-U.S. Gulf Coast, October 5-30, 2020, Biloxi, MS, USA. New York: IEEE Press, 2020: 1-5.
- [19] Figueiredo A B, Matos A C. MViDO: a high performance monocular vision-based system for docking a hovering AUV[J]. Applied Sciences, 2020, 10(9): 2991.
- [20] Rannestad T L, Waldum A, Ludvigsen M. Visual close-range navigation and docking of underwater vehicles[C]//OCEANS 2023-Limerick, June 5-8, 2023, Limerick, Ireland. New York: IEEE Press, 2023: 1-10.
- [21] Lv F T, Xu H X, Shi K, et al. Estimation of positions and poses of autonomous underwater vehicle relative to docking station based on adaptive extraction of visual guidance features[J]. Machines, 2022, 10(7): 571.
- [22] Yan Z P, Jiang L, Zhao Y F, et al. A novel image matching

- algorithm application in vision guided AUV docking[J]. *Energy Procedia*, 2012, 17: 991-1000.
- [23] Zhang L C, Wang R Z, Zhang F H, et al. Design of vision-based self assembling method for dual AUVs[C]//2018 OCEANS-MTS/IEEE Kobe Techno-Oceans (OTO), May 28-31, 2018, Kobe, Japan. New York: IEEE Press, 2018: 1-6.
- [24] Zhang L C, Li Y C, Pan G, et al. Terminal stage guidance method for underwater moving rendezvous and docking based on monocular vision[C]//OCEANS 2019-Marseille, June 17-20, 2019, Marseille, France. New York: IEEE Press, 2019: 1-6.
- [25] Zhang Z Y, Zhong L J, Lin M W, et al. Triangle codes and tracer lights based absolute positioning method for terminal visual docking of autonomous underwater vehicles[J]. *Industrial Robot: the International Journal of Robotics Research and Application*, 2024, 51(2): 269-286.
- [26] Lin M W, Lin R, Li D J, et al. Light beacon-aided AUV electromagnetic localization for landing on a planar docking station[J]. *IEEE Journal of Oceanic Engineering*, 2023, 48(3): 677-688.
- [27] Zhang Z Y, Ding W J, Wu R D, et al. Autonomous underwater vehicle cruise positioning and docking guidance scheme[J]. *Journal of Marine Science and Engineering*, 2024, 12(6): 1023.
- [28] Singh P, Gregson E, Ross J, et al. Vision-based AUV docking to an underway dock using convolutional neural networks[C]//2020 IEEE/OES Autonomous Underwater Vehicles Symposium (AUV), September 30-October 2, 2020, St Johns, NL, Canada. New York: IEEE Press, 2020: 1-6.
- [29] Karimanzira D, Pfitzenreuter T, Renkewitz H. Deep learning for long and short range object detection in underwater environment[EB/OL]. [2024-12-12]. <https://symbiosisonlinepublishing.com/robotics-automation/robotics-automation41.pdf>.
- [30] Ren R Z, Zhang L C, 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.
- [31] Zhang F H, Zhong Y H, Chen L Y, et al. Event-based circular detection for AUV docking based on spiking neural network[J]. *Frontiers in Neurorobotics*, 2022, 15: 815144.
- [32] Alla D N V, Jyothi V B N, Venkataraman H, et al. Vision-based deep learning algorithm for underwater object detection and tracking[C]//OCEANS 2022-Chennai, February 21-24, 2022, Chennai, India. New York: IEEE Press, 2022: 1-6.
- [33] Zhang B, Zhong P, Yang F, et al. Fast underwater optical beacon finding and high accuracy visual ranging method based on deep learning[J]. *Sensors*, 2022, 22(20): 7940.
- [34] Jyothi V B N, Akash S J, Ramadass G A, et al. Design and development of deep learning-aided vision guidance system for AUV homing applications[J]. *IEEE Embedded Systems Letters*, 2024, 16(2): 198-201.
- [35] 卢有旺, 夏英凯, 徐国华, 等. 面向 UUV 对接的视觉引导三维轨迹跟踪控制研究[J]. *中国舰船研究*, 2024, 19(1): 290-304.
- Lu Y W, Xia Y K, Xu G H, et al. Study on vision-guided 3D tracking control for UUV docking[J]. *Chinese Journal of Ship Research*, 2024, 19(1): 290-304.
- [36] Liu S, Xu H L, Lin Y, et al. Visual navigation for recovering an AUV by another AUV in shallow water[J]. *Sensors*, 2019, 19(8): 1889.
- [37] Liu S, Ozay M, Okatani T, et al. Detection and pose estimation for short-range vision-based underwater docking[J]. *IEEE Access*, 2018, 7: 2720-2749.
- [38] Liu S, Ozay M, Okatani T, et al. Learning deep representations and detection of docking stations using underwater imaging[C]//2018 OCEANS-MTS/IEEE Kobe Techno-Oceans (OTO), May 28-31, 2018, Kobe, Japan. New York: IEEE Press, 2018: 1-5.
- [39] Sun K, Han Z K. Autonomous underwater vehicle docking system for energy and data transmission in cabled ocean observatory networks[J]. *Frontiers in Energy Research*, 2022, 10: 960278.
- [40] Zhou J J, Zhang W, Wu D, et al. Underwater recovery realization for an AUV using positioning-to-line strategy[C]//2014 IEEE Conference and Expo Transportation Electrification Asia-Pacific (ITEC Asia-Pacific), August 31-September 3, 2014, Beijing. New York: IEEE Press, 2014: 1-5.
- [41] Watt G D, Roy A R, Currie J, et al. A concept for docking a UUV with a slowly moving submarine under waves[J]. *IEEE Journal of Oceanic Engineering*, 2016, 41(2): 471-498.
- [42] Yahya M F, Arshad M R. Detection of markers using deep learning for docking of autonomous underwater vehicle[C]//2017 IEEE 2nd International Conference on Automatic Control and Intelligent Systems (I2CACIS), October 21-21, 2017, Kota Kinabalu, Malaysia. New York: IEEE Press, 2017: 179-184.
- [43] Sans-Muntadas A, Kelasidi E, Pettersen K Y, et al. Learning an AUV docking maneuver with a convolutional neural network [J]. *IFAC Journal of Systems and Control*, 2019, 8: 100049.
- [44] Johnson-Roberson M, Barto C, Mehta R, et al. Driving in the matrix: can virtual worlds replace human-generated annotations for real world tasks? [EB/OL]. (2016-10-06) [2024-12-12]. <https://arxiv.org/abs/1610.01983v2>.
- [45] Olson E A, Barbalata C, Zhang J M, et al. Synthetic data generation for deep learning of underwater disparity estimation [C]//OCEANS 2018 MTS/IEEE Charleston, October 22-25, 2018, Charleston, SC, USA. New York: IEEE Press, 2018: 1-6.
- [46] Chavez-Galaviz J, Mahmoudian N. Underwater dock detection through convolutional neural networks trained with artificial image generation[C]//2022 International Conference on Robotics and Automation (ICRA), May 23-27, 2022, Philadelphia, PA, USA. New York: IEEE Press, 2022: 4621-4627.
- [47] Chavez-Galaviz J, Li J W, Bergman M, et al. An efficient detection and control system for underwater docking using machine learning and realistic simulation: a comprehensive approach[EB/OL]. (2023-11-02)[2024-12-12]. <https://arxiv.org/abs/2311.01522v2>.
- [48] Chavez-Galaviz J, Mahmoudian N. Efficient underwater docking detection using knowledge distillation and artificial image generation[C]//2022 IEEE/OES Autonomous Underwater Vehicles Symposium (AUV), September 19-21, 2022, Singapore. New York: IEEE Press, 2022: 1-7.
- [49] Omachi M, Omachi S. Traffic light detection with color and edge information[C]//2009 2nd IEEE International Conference on Computer Science and Information Technology, August 8-11, 2009, Beijing, China. New York: IEEE Press, 2009: 284-287.
- [50] Foley J M. Binocular distance perception[J]. *Psychological Review*, 1980, 87(5): 411-434.
- [51] Lee H M, Choi W C. Algorithm of 3D spatial coordinates measurement using a camera image[J]. *Journal of Image and Graphics*, 2015, 3(1): 30-33.
- [52] Scharstein D, Szeliski R, Zabih R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms[C]//Proceedings IEEE Workshop on Stereo and Multi-Baseline Vision (SMBV 2001), December 9-10, 2001, Kauai, HI, USA. New York: IEEE Press, 2001: 131-140.
- [53] di Stefano L, Marchionni M, Mattoccia S. A fast area-based stereo matching algorithm[J]. *Image and Vision Computing*, 2004, 22(12): 983-1005.
- [54] Zhang Z Y, Deriche R, Faugeras O, et al. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry[J]. *Artificial Intelligence*, 1995, 78(1/2): 87-119.
- [55] Ge P X, Wang Y H, Wang H Q, et al. Multivision sensor extrinsic calibration method with non-overlapping fields of view

- using encoded 1D target[J]. *IEEE Sensors Journal*, 2022, 22(13): 13519-13528.
- [56] Chen X C, Song X K, Wu J, et al. Camera calibration with global LBP-coded phase-shifting wedge grating arrays[J]. *Optics and Lasers in Engineering*, 2021, 136: 106314.
- [57] Zhang Z. A flexible new technique for camera calibration[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, 22(11): 1330-1334.
- [58] Zhang X W, Ren Y F, Zhen G Y, et al. Camera calibration method for solid spheres based on triangular primitives[J]. *Precision Engineering*, 2020, 65: 91-102.
- [59] Hartley R I. Self-calibration of stationary cameras[J]. *International Journal of Computer Vision*, 1997, 22(1): 5-23.
- [60] 杨长江, 汪威, 胡占义. 一种基于主动视觉的摄像机内参数自标定方法[J]. *计算机学报*, 1998, 21(5): 428-435.
- Yang C J, Wang W, Hu Z Y. An active vision based camera intrinsic parameters self-calibration technique[J]. *Chinese Journal of Computers*, 1998, 21(5): 428-435.
- [61] 吴福朝, 阮宗才, 胡占义. 非线性模型下的摄像机自标定[J]. *计算机学报*, 2002, 25(3): 276-283.
- Wu F C, Ruan Z C, Hu Z Y. Self-calibration of a camera with a non-linear model[J]. *Chinese Journal of Computers*, 2002, 25(3): 276-283.
- [62] 吴福朝, 胡占义. 线性确定无穷远平面的单应矩阵和摄像机自标定[J]. *自动化学报*, 2002, 28(4): 488-496.
- Wu F C, Hu Z Y. Linear determination of the infinite homography and camera self-calibration[J]. *Acta Automatica Sinica*, 2002, 28(4): 488-496.
- [63] Hartley R I. Projective reconstruction and invariants from multiple images[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1994, 16(10): 1036-1041.
- [64] Pollefeys M, van Gool L, Oosterlinck A. The modulus constraint: a new constraint self-calibration[C]//*Proceedings of 13th International Conference on Pattern Recognition*, August 25-29, 1996, Vienna, Austria. New York: IEEE Press, 1996: 349-353.
- [65] 孟晓桥, 胡占义. 摄像机自标定方法的研究与进展[J]. *自动化学报*, 2003, 29(1): 110-124.
- Meng X Q, Hu Z Y. Recent progress in camera self-calibration[J]. *Acta Automatica Sinica*, 2003, 29(1): 110-124.
- [66] Triggs B. Autocalibration and the absolute quadric[C]//*Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, June 17-19, 1997, San Juan, PR, USA. New York: IEEE Press, 1997: 609-614.
- [67] Heyden A, Astrom K. Flexible calibration: minimal cases for auto-calibration[C]//*Proceedings of the Seventh IEEE International Conference on Computer Vision*, September 20-27, 1999, Kerkira, Greece. New York: IEEE Press, 1999: 350-355.
- [68] Pollefeys M, Koch R, van Gool L. Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters[C]//*Sixth International Conference on Computer Vision*, January 7, 1998, Bombay, India. New York: IEEE Press, 1998: 90-95.
- [69] Pollefeys M. Self-calibration and metric 3D reconstruction from uncalibrated image sequences[D]. Louvain: Catholic University of Louvain, 1999.
- [70] Hartley R I, Hayman E, de Agapito L, et al. Camera calibration and the search for infinity[C]//*Proceedings of the Seventh IEEE International Conference on Computer Vision*, September 20-27, 1999, Kerkira, Greece. New York: IEEE Press, 1999: 510-517.
- [71] Vasconcelos F, Barreto J P, Boyer E. Automatic camera calibration using multiple sets of pairwise correspondences[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(4): 791-803.
- [72] Kim J O, Lee D. Mathematical model for a calibration of multiple-view thermal camera[C]//*2021 International Conference on Electronics, Information, and Communication (ICEIC)*, January 31-February 3, 2021, Jeju, Republic of Korea. New York: IEEE Press, 2021: 1-3.
- [73] Faugeras O D, Luong Q T, Maybank S J. Camera self-calibration: theory and experiments[M]//Sandini G. *Computer vision—ECCV'92. Lecture notes in computer science*. Berlin: Springer Berlin Heidelberg, 1992, 588: 321-334.
- [74] Rög M, Rzonca A. The impact of photo overlap, the number of control points and the method of camera calibration on the accuracy of 3D model reconstruction[J]. *Geomatics and Environmental Engineering*, 2021, 15(2): 67-87.
- [75] 王子昂, 刘秉琦, 黄富瑜. 基于一种通用相机模型的双目超大视场红外相机标定[J]. *光电与控制*, 2020, 27(5): 108-111.
- Wang Z A, Liu B Q, Huang F Y. Binocular ultra-wide-angle infrared camera calibration based on generic camera model[J]. *Electronics Optics & Control*, 2020, 27(5): 108-111.
- [76] Bhardwaj R, Tummala G K, Ramalingam G, et al. AutoCalib: automatic traffic camera calibration at scale[J]. *ACM Transactions on Sensor Networks*, 2018, 14(3/4): 1-27.
- [77] 赵清杰, 孙增圻, 兰丽. 摄像机神经网络标定技术[J]. *控制与决策*, 2002, 17(3): 336-338, 342.
- Zhao Q J, Sun Z Q, Lan L. Neural network technique in camera calibration[J]. *Control and Decision*, 2002, 17(3): 336-338, 342.
- [78] Raza S N, ur Rehman H R, Lee S G, et al. Artificial intelligence based camera calibration[C]//*2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC)*, June 24-28, 2019, Tangier, Morocco. New York: IEEE Press, 2019: 1564-1569.
- [79] Wen J, Schweitzer G. Hybrid calibration of CCD cameras using artificial neural nets[C]//*Proceedings 1991 IEEE International Joint Conference on Neural Networks*, November 18-21, 1991, Singapore. New York: IEEE Press, 1991: 337-342.
- [80] Jun J, Kim C. Robust camera calibration using neural network[C]//*Proceedings of IEEE. IEEE Region 10 Conference. TENCON 99. 'Multimedia Technology for Asia-Pacific Information Infrastructure'*, September 15-17, 1999, Cheju, Republic of Korea. New York: IEEE Press, 1999: 694-697.
- [81] 吕朝辉, 张兆杨, 安平. 基于神经网络的立体视觉摄像机标定[J]. *机械工程学报*, 2003, 39(9): 93-96.
- Lu C H, Zhang Z Y, An P. Camera calibration for stereo vision based on neural network[J]. *Chinese Journal of Mechanical Engineering*, 2003, 39(9): 93-96.
- [82] Jin L G, Rui L G. Camera calibration for monocular vision system based on Harris corner extraction and neural network[C]//*2011 International Conference on Consumer Electronics, Communications and Networks (CECNet)*, April 16-18, 2011, Xianning, China. New York: IEEE Press, 2011: 1-4.
- [83] Ding X Z. Research on kinect calibration and depth error compensation based on BP neural network[C]//*2020 International Conference on Computer Vision, Image and Deep Learning (CVIDL)*, July 10-12, 2020, Chongqing, China. New York: IEEE Press, 2020: 596-600.
- [84] 胡志新, 王涛. 改进遗传算法优化BP神经网络的双目相机标定[J]. *光电与控制*, 2022, 29(1): 75-79.
- Hu Z X, Wang T. Binocular camera calibration based on BP neural network optimized by improved genetic algorithm[J]. *Electronics Optics & Control*, 2022, 29(1): 75-79.
- [85] Muhammad K, Ahmad J, Lv Z H, et al. Efficient deep CNN-based fire detection and localization in video surveillance applications[J]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2019, 49(7): 1419-1434.
- [86] Guo H B, Sheng B, Li P, et al. Multiview high dynamic range image synthesis using fuzzy broad learning system[J]. *IEEE Transactions on Cybernetics*, 2021, 51(5): 2735-2747.
- [87] Huang K, Ziauddin S, Zand M, et al. One shot radial distortion correction by direct linear transformation[C]//*2020 IEEE*

- International Conference on Image Processing (ICIP), October 25–28, 2020, Abu Dhabi, United Arab Emirates. New York: IEEE Press, 2020: 473-477.
- [88] Zhao H, Ying X H, Shi Y J, et al. RDCFace: radial distortion correction for face recognition[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13–19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 7718-7727.
- [89] Sheng B, Li P, Jin Y X, et al. Intrinsic image decomposition with step and drift shading separation[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2020, 26(2): 1332-1346.
- [90] Singh S, Singh M. Explosives detection systems (EDS) for aviation security[J]. *Signal Processing*, 2003, 83(1): 31-55.
- [91] Hill D L, Batchelor P G, Holden M, et al. Medical image registration[J]. *Physics in Medicine & Biology*, 2001, 46(3): R1-R45.
- [92] Qian B, Chen H, Wang X N, et al. DRAC 2022: a public benchmark for diabetic retinopathy analysis on ultra-wide optical coherence tomography angiography images[J]. *Patterns*, 2024, 5(3): 100929.
- [93] Dai L, Wu L, Li H T, et al. A deep learning system for detecting diabetic retinopathy across the disease spectrum[J]. *Nature Communications*, 2021, 12: 3242.
- [94] Xie Z F, Zhang W L, Sheng B, et al. BaGFN: broad attentive graph fusion network for high-order feature interactions[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2023, 34(8): 4499-4513.
- [95] Cheng Z Z, Yang Q X, Sheng B. Deep colorization[C]//2015 IEEE International Conference on Computer Vision (ICCV), December 7–13, 2015, Santiago, Chile. New York: IEEE Press, 2015: 415-423.
- [96] Ma Z L, Wang J L, Cui C C, et al. Geometric distortion correction for the underwater images[C]//OCEANS 2018 MTS/IEEE Charleston, October 22–25, 2018, Charleston, SC, USA. New York: IEEE Press, 2018: 1-5.
- [97] Tsai R. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses[J]. *IEEE Journal on Robotics and Automation*, 1987, 3(4): 323-344.
- [98] Barreto J P, Daniilidis K. Fundamental matrix for cameras with radial distortion[C]//Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1, October 17–21, 2005, Beijing, China. New York: IEEE Press, 2005: 625-632.
- [99] Fan J L, Zhang J, Maybank S J, et al. Wide-angle image rectification: a survey[J]. *International Journal of Computer Vision*, 2022, 130(3): 747-776.
- [100] Hartley R, Kang S B. Parameter-free radial distortion correction with center of distortion estimation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, 29(8): 1309-1321.
- [101] Kannala J, Brandt S S. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, 28(8): 1335-1340.
- [102] Sun J, Wang H P, Zhu X L. A fast underwater calibration method based on vanishing point optimization of two orthogonal parallel lines[J]. *Measurement*, 2021, 178: 109305.
- [103] 赵亚凤, 胡峻峰. 一种双正交消隐点的双目相机标定方法[J]. *液晶与显示*, 2016, 31(10): 958-966.
- Zhao Y F, Hu J F. Binocular self calibration using two pairs of orthogonal vanishing points[J]. *Chinese Journal of Liquid Crystals and Displays*, 2016, 31(10): 958-966.
- [104] Barreto J P, Roquette J, Sturm P, et al. Automatic camera calibration applied to medical endoscopy[C]//Proceedings of the British Machine Vision Conference 2009, September 7–10, 2009, London. London: British Machine Vision Association, 2009: 52.1-52.10.
- [105] Bukhari F, Dailey M N. Automatic radial distortion estimation from a single image[J]. *Journal of Mathematical Imaging and Vision*, 2013, 45(1): 31-45.
- [106] Devernay F, Faugeras O. Straight lines have to be straight[J]. *Machine Vision and Applications*, 2001, 13(1): 14-24.
- [107] Zhang B X, Sheng B, Li P, et al. Depth of field rendering using multilayer-neighborhood optimization[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2020, 26(8): 2546-2559.
- [108] Nazir A, Cheema M N, Sheng B, et al. OFF-eNET: an optimally fused fully end-to-end network for automatic dense volumetric 3D intracranial blood vessels segmentation[J]. *IEEE Transactions on Image Processing*, 2020, 29: 7192-7202.
- [109] Chen Z H, Gao T, Sheng B, et al. Outdoor shadow estimating using multiclass geometric decomposition based on BLS[J]. *IEEE Transactions on Cybernetics*, 2020, 50(5): 2152-2165.
- [110] Rong J P, Huang S Y, Shang Z Y, et al. Radial lens distortion correction using convolutional neural networks trained with synthesized images[M]//Lai S H, Lepetit V, Nishino K, et al. *Computer vision-ACCV 2016. Lecture notes in computer science*. Cham: Springer International Publishing, 2017, 10113: 35-49.
- [111] Myint M, Yonemori K, Yanou A, et al. Visual-servo-based autonomous docking system for underwater vehicle using dual-eyes camera 3D-pose tracking[C]//2015 IEEE/SICE International Symposium on System Integration (SII), December 11–13, 2015, Nagoya, Japan. New York: IEEE Press, 2015: 989-994.
- [112] Myint M, Yonemori K, Yanou A, et al. Visual servoing for underwater vehicle using dual-eyes evolutionary real-time pose tracking[J]. *Journal of Robotics and Mechatronics*, 2016, 28(4): 543-558.
- [113] Myint M, Yonemori K, Yanou A, et al. Robustness of visual-servo against air bubble disturbance of underwater vehicle system using three-dimensional marker and dual-eye cameras [C]//OCEANS 2015-MTS/IEEE Washington, October 19–22, 2015, Washington, DC, USA. New York: IEEE Press, 2015: 1-8.
- [114] Li X, Nishida Y, Myint M, et al. Dual-eyes vision-based docking experiment of AUV for sea bottom battery recharging [C]//OCEANS 2017-Aberdeen, June 19–22, 2017, Aberdeen, UK. New York: IEEE Press, 2017: 1-5.
- [115] Myint M, Yonemori K, Yanou A, et al. Dual-eyes visual-based sea docking for sea bottom battery recharging[C]//OCEANS 2016 MTS/IEEE Monterey, September 19–23, 2016, Monterey, CA, USA. New York: IEEE Press, 2016: 1-7.
- [116] Myint M, Yonemori K, Yanou A, et al. Visual-based deep sea docking simulation of underwater vehicle using dual-eyes cameras with lighting adaptation[C]//OCEANS 2016-Shanghai, April 10–13, 2016, Shanghai, China. New York: IEEE Press, 2016: 1-8.
- [117] Lwin K N, Myint M, Yonemori K, et al. Dual-eye vision-based docking experiment in the sea for battery recharging application [J]. *SICE Journal of Control, Measurement, and System Integration*, 2019, 12(2): 47-55.
- [118] Lwin K N, Yonemori K, Myint M, et al. Autonomous docking experiment in the sea for visual-servo type underwater vehicle using three-dimensional marker and dual-eyes cameras[C]//2016 55th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE), September 20–23, 2016, Tsukuba, Japan. New York: IEEE Press, 2016: 1359-1365.
- [119] Lwin K N, Myint M, Mukada N, et al. Sea docking by dual-eye pose estimation with optimized genetic algorithm parameters[J]. *Journal of Intelligent & Robotic Systems*, 2019, 96(2): 245-266.
- [120] Nakamura S, Yamada D, Mukada N, et al. Development of dual-eyes docking system for AUV with lighting 3D marker[C]//

- OCEANS 2018 MTS/IEEE Charleston, October 22–25, 2018, Charleston, SC, USA. New York: IEEE Press, 2018: 1–9.
- [121] Lwin K N, Mukada N, Myint M, et al. Docking at pool and sea by using active marker in turbid and day/night environment[J]. *Artificial Life and Robotics*, 2018, 23(3): 409–419.
- [122] Yamada D, Mukada N, Myint M, et al. Docking experiment in dark environments using active/lighting marker and HSV correlation[C]//2018 OCEANS-MTS/IEEE Kobe Techno-Oceans (OTO), May 28–31, 2018, Kobe, Japan. New York: IEEE Press, 2018: 1–9.
- [123] Kanda Y, Myint M, Mukada N, et al. Applying lighting marker and stereo-vision to V-shaped-thruster vehicle for AUV deep sea docking[C]//OCEANS 2018 MTS/IEEE Charleston, October 22–25, 2018, Charleston, SC, USA. New York: IEEE Press, 2018: 1–8.
- [124] Myint M, Lwin K N, Mukada N, et al. Stereo vision-based 3D pose estimation under turbid water for underwater vehicles [EB/OL]. [2024-12-12]. <https://visual-servo.com/pdf/thesis/pdf2017/RSJ2017/Myo.pdf>.
- [125] Myint M, Yonemori K, Lwin K N, et al. Dual-eyes vision-based docking system for autonomous underwater vehicle: an approach and experiments[J]. *Journal of Intelligent & Robotic Systems*, 2018, 92(1): 159–186.
- [126] Myint M, Lwin K N, Mukada N, et al. Experimental verification of turbidity tolerance of stereo-vision-based 3D pose estimation system[J]. *Journal of Marine Science and Technology*, 2019, 24(3): 756–779.
- [127] Lwin K N, Mukada N, Myint M, et al. Visual docking against bubble noise with 3-D perception using dual-eye cameras[J]. *IEEE Journal of Oceanic Engineering*, 2020, 45(1): 247–270.
- [128] Hsu H Y, Mukada N, Yamada D, et al. Improving pose estimation accuracy and expanding of visible space of lighting 3D marker in turbid water[C]//2019 IEEE Underwater Technology (UT), April 16–19, 2019, Kaohsiung, Taiwan, China. New York: IEEE Press, 2019: 1–8.
- [129] 施小成, 王晓娟. 一种面向AUV水下对接的双目视觉测距方法[J]. *计算机测量与控制*, 2008, 16(10): 1460–1462, 1488.
- Shi X C, Wang X J. A ranging method in AUV underwater docking by binocular vision[J]. *Computer Measurement & Control*, 2008, 16(10): 1460–1462, 1488.
- [130] Li Y, Jiang Y Q, Cao J, et al. AUV docking experiments based on vision positioning using two cameras[J]. *Ocean Engineering*, 2015, 110: 163–173.
- [131] Zhong L J, Li D J, Lin M W, et al. A fast binocular localisation method for AUV docking[J]. *Sensors*, 2019, 19(7): 1735.
- [132] 朱志鹏, 张晓文, 韩祥. 基于单双目切换的自主式水下机器人视觉定位方法[J]. *船舶工程*, 2023, 45(2): 131–139, 144.
- Zhu Z P, Zhang X W, Han X. AUV visual localization method based on monocular and binocular switching[J]. *Ship Engineering*, 2023, 45(2): 131–139, 144.
- [133] 徐硕, 姜言清, 李晔, 等. 智能水下机器人自主回收的双目视觉定位[J]. *哈尔滨工程大学学报*, 2022, 43(8): 1084–1090.
- Xu S, Jiang Y Q, Li Y, et al. A stereo vision localization method for autonomous recovery of autonomous underwater vehicle[J]. *Journal of Harbin Engineering University*, 2022, 43(8): 1084–1090.
- [134] Xu S, Jiang Y Q, Li Y, et al. A stereo visual navigation method for docking autonomous underwater vehicles[J]. *Journal of Field Robotics*, 2024, 41(2): 374–395.
- [135] 陈青艳, 吴国俊, 郝歌扬. 一种多相机成像系统的图像配准方法: CN118762059A[P]. 2024-10-11.
- Chen Q Y, Wu G J, Hao G Y. Image registration method for multi-camera imaging system: CN118762059A[P]. 2024-10-11.
- [136] 封斐, 吴国俊, 吴亚风, 等. 基于全局估计的水下偏振成像算法[J]. *光学学报*, 2020, 40(21): 2111002.
- Feng F, Wu G J, Wu Y F, et al. Algorithm for underwater polarization imaging based on global estimation[J]. *Acta Optica Sinica*, 2020, 40(21): 2111002.
- [137] 徐菁焱, 吴国俊, 董晶, 等. 基于Stokes矢量差分法的背景光偏振特性研究[J]. *物理学报*, 2023, 72(24): 244201.
- Xu J H, Wu G J, Dong J, et al. Research on polarization characteristics of background light by modified polarization difference imaging method[J]. *Acta Physica Sinica*, 2023, 72(24): 244201.
- [138] Guo S X, Liu Y, Shi L W, et al. Binocular camera-based a docking system for an amphibious spherical robot[C]//2018 IEEE International Conference on Mechatronics and Automation (ICMA), August 5–8, 2018, Changchun, China. New York: IEEE Press, 2018: 1621–1626.
- [139] Hsu H Y, Toda Y, Yamashita K, et al. Stereo-vision-based AUV navigation system for resetting the inertial navigation system error[J]. *Artificial Life and Robotics*, 2022, 27(1): 165–178.
- [140] 徐飞飞, 纪明, 赵创社. 快速偏转反射镜研究现状及关键技术[J]. *应用光学*, 2010, 31(5): 847–850.
- Xu F F, Ji M, Zhao C S. Status of fast steering mirror[J]. *Journal of Applied Optics*, 2010, 31(5): 847–850.
- [141] 周培松, 卢荣胜, 周维虎, 等. 四象限探测器用于激光跟踪仪目标脱靶量测量[J]. *光电技术应用*, 2014, 29(1): 75–79.
- Zhou P S, Lu R S, Zhou W H, et al. Measurement system for laser tracker's target-missing quantity based on four-quadrant detector[J]. *Electro-Optic Technology Application*, 2014, 29(1): 75–79.
- [142] Wang X Q, Xu D, Tan M, et al. Analysis for the characteristic of a four-quadrant photo-detector applied to a novel laser aligning system[J]. *Proceedings of SPIE*, 2006, 6150: 61504V.
- [143] Qian J Q, Cui Y M, Xu P. The study for measuring rotor speed and direction with quadrant photoelectric detector[J]. *Measurement*, 2008, 41(6): 626–630.
- [144] Hu Y B, Wang M. Study on the technology of mutual alignment based on the four-quadrant photo electric detector[J]. *Proceedings of SPIE*, 2015, 9795: 979501.
- [145] 姜清华, 李立京, 刘腾飞. 基于四象限探测器的导引头信号处理系统设计[J]. *电子测量技术*, 2018, 41(2): 88–92.
- Jiang Q H, Li L J, Liu T F. Design of seeker signal processing system based on four-quadrant detector[J]. *Electronic Measurement Technology*, 2018, 41(2): 88–92.
- [146] 刘鹏飞, 宋翠莲, 马昱超. 一种用于微光探测的四象限探测器设计[J]. *国外电子测量技术*, 2020, 39(5): 99–104.
- Liu P F, Song C L, Ma Y C. Design of a four quadrant detector for shimmer detection[J]. *Foreign Electronic Measurement Technology*, 2020, 39(5): 99–104.
- [147] Ma S W, Chen W J, Gao W, et al. Research on the laser angle deception jamming technology of laser countermeasure[J]. *Proceedings of SPIE*, 2015, 9671: 96710B.
- [148] He C, Chang J T, Wang Y, et al. Linear polarization optimized Stokes polarimeter based on four-quadrant detector[J]. *Applied Optics*, 2015, 54(14): 4458–4463.
- [149] Wang S Z, Chang S P, Xie T B, et al. Applied research on four-quadrant optoelectronic detector for surface topography measurement[J]. *Proceedings of SPIE*, 2007, 6723: 672329.
- [150] Marquina A V, González L A, Noriega J R. Design and characterization of a four-quadrant detector and its application to a shape recovering system[J]. *Proceedings of SPIE*, 2006, 6294: 629410.
- [151] Cowen S, Briest S, Dombrowski J. Underwater docking of autonomous undersea vehicles using optical terminal guidance [C]//Oceans '97. MTS/IEEE Conference Proceedings, October 6–9, 1997, Halifax, NS, Canada. New York: IEEE Press, 1997: 1143–1147.
- [152] Martin S, Fletcher B, Flores G, et al. Characterizing the critical parameters for docking unmanned underwater vehicles[C]//OCEANS 2016 MTS/IEEE Monterey, September 19–23,

- 2016, Monterey, CA, USA. New York: IEEE Press, 2016: 1-7.
- [153] Fletcher B, Martin S, Flores G, et al. From the lab to the ocean: characterizing the critical docking parameters for a free floating dock with a REMUS 600[C]//OCEANS 2017-Anchorage, September 18–21, 2017, Anchorage, AK, USA. New York: IEEE Press, 2017: 1-7.
- [154] 吴国俊, 邵建峰, 吕小鹏, 等. 水下无人航行器回收系统及回收方法: CN112013773A[P]. 2020-12-01.
Wu G J, Shao J F, Lü X P, et al. Underwater unmanned vehicle recovery system and recovery method: CN112013773A [P]. 2020-12-01.
- [155] 吕小鹏, 吴国俊, 吴亚风, 等. 一种基于多元探测器的水下无人潜航器回收方法: CN118936509A[P]. 2024-11-12.
Lü X P, Wu G J, Wu Y F. A method for recovering underwater unmanned submersible based on multi-element detector: CN118936509A[P]. 2024-11-12.
- [156] Evans J C, Keller K M, Smith J S, et al. Docking techniques and evaluation trials of the SWIMMER AUV: an autonomous deployment AUV for work-class ROVs[C]//MTS/IEEE Oceans 2001. An Ocean Odyssey. Conference Proceedings, November 5–8, 2001, Honolulu, HI, USA. New York: IEEE Press, 2001: 520-528.
- [157] Eren F, Pe'eri S, Thein M W, et al. Position, orientation and velocity detection of unmanned underwater vehicles (UUVs) using an optical detector array[J]. Sensors, 2017, 17(8): 1741.
- [158] Simpson J A, Hughes B L, Muth J F. Smart transmitters and receivers for underwater free-space optical communication[J]. IEEE Journal on Selected Areas in Communications, 2012, 30 (5): 964-974.
- [159] Chen Y H, Duan Z Y, Zheng F Z, et al. Underwater optical guiding and communication solution for the AUV and seafloor node[J]. Applied Optics, 2022, 61(24): 7059-7070.
- [160] 孙哲, 李学龙. 基于计算成像的无人航行器加密导引系统及方法: CN116320198A[P]. 2023-06-23.
Sun Z, Li X L. Encrypted guidance system and method for AUV based on computational imaging: CN116320198A[P]. 2023-06-23.

Optical Guidance for AUV Underwater Docking (Invited)

Li Xuelong^{1,2*}, Sun Zhe^{1,2**}, Wu Guojun^{3***}

¹*School of Artificial Intelligence, Optics and Electronics (iOPEN), Northwestern Polytechnical University, Xi'an 710072, Shaanxi, China;*

²*Institute of Artificial Intelligence (TeleAI), China Telecom, Shanghai 200232, China;*

³*Xi'an Institute of Optics and Precision Mechanics of Chinese Academy of Sciences, Xi'an 710119, Shaanxi, China*

Abstract

Significance Autonomous Underwater Vehicles (AUVs) are pivotal tools for ocean exploration, resource utilization, and environmental monitoring. The underwater docking process, which enables AUVs to physically connect with recovery stations for energy replenishment and data transmission, is critical for enhancing operational efficiency and mission continuity. Traditional guidance methods, such as acoustic and electromagnetic systems, face limitations in precision, robustness, and adaptability to complex underwater environments. Acoustic guidance suffers from low resolution and susceptibility to multipath interference, while electromagnetic signals degrade rapidly in water. Optical guidance, leveraging high-resolution visual or photoelectric detection, has emerged as a promising solution for close-range docking due to its superior accuracy, real-time performance, and stealth advantages. This review highlights advancements in optical guidance technologies, focusing on monocular vision, binocular vision, and position detectors, and outlines their transformative potential in enabling reliable AUV underwater recovery.

Progress 1) Monocular vision guidance. Monocular vision systems utilize a single camera to detect active or passive optical markers on docking stations. Active markers, such as LED arrays, offer long visibility ranges but require precise geometric configurations to avoid ambiguity in pose estimation. Passive markers (e.g., ArUco codes or geometric patterns) provide unique identification but are limited by shorter detection distances. Recent studies have improved robustness through multi-marker fusion and deep learning. For instance, irregularly arranged four-light beacons [Figs. 5(a)–(c)] and hybrid markers combining LEDs with black-and-white codes [Fig. 5(d)] enhance feature matching accuracy. Deep learning frameworks like YOLOV5 and CNN-based models [Fig. 6(a)] further optimize marker recognition in turbid water. Currently, deep learning-enhanced monocular visual guidance achieves sub-3 cm localization accuracy by combining beacon recognition with PnP algorithms or fusing visual data with multi-sensor inputs, but faces challenges in underwater optical attenuation, high computational demands, and limited real-time pose estimation frequency.

2) Binocular vision guidance. Binocular vision systems leverage stereo cameras to resolve depth through disparity analysis [Fig. 7(b)]. By correlating pixel coordinates of guide lights in dual images, 3D coordinates are derived using

triangulation. Key advancements include camera calibration and distortion correction. Traditional calibration methods (e.g., Zhang's checkerboard approach) ensure sub-pixel accuracy, while neural networks address nonlinear distortions caused by underwater optical windows [Fig. 7(a)]. Binocular vision guidance enhances beacon recognition range and accuracy, enabling AUVs to achieve sub-centimeter localization precision (~ 10 mm error) and 30-meter docking range, though its computational speed (milliseconds to hundreds of milliseconds per cycle) requires further optimization despite low hardware demands.

3) Position detector-based guidance. Position detectors, such as quadrant photodetectors (QPDs), track light spots from docking station beacons. These systems excel in high-speed tracking and angular resolution but require precise optical alignment. Experimental validations demonstrate their robustness in turbulent flows, achieving angular accuracies within 0.1° . The sea trial demonstrated that the multi-branch network optical guidance method, based on multi-quadrant photoelectric detection and real-time angle data processing, achieved an AUV position resolution speed of 5.650 ms/cycle and a mean coordinate error of 58.292 mm (best 7.107 mm at 2–3 m), fulfilling precision and efficiency requirements for terminal docking with lower computational power and energy consumption compared to existing methods.

Conclusions and Prospects Optical guidance for AUV underwater docking, a cornerstone technology enabling safe, continuous, and efficient marine operations, has garnered significant attention from researchers globally. This study systematically reviews two primary optical guidance paradigms: image sensor-based methods and position detector-based methods. Image sensor-based approaches, characterized by intuitive data acquisition and high positioning accuracy, dominate current practices by leveraging visual or photoelectric sensing to extract beacon features and resolve relative pose. Meanwhile, position detector-based methods, exemplified by multi-quadrant photoelectric detectors, highlight advantages in detection speed and communication-integration potential.

Despite progress, critical challenges persist: benchmark dataset limitations. While datasets have been developed, acquiring high-fidelity ground truth data remains arduous due to dynamic underwater environments and system-induced noise. Image sensor-based methods suffer from low frame rates, exacerbating latency and computational burdens during real-time processing. Position detectors, though faster, lack sufficient modulation bandwidth for high-speed communication.

To address these gaps, future advancements should focus on three synergistic directions:

1) High-speed, stable, and intelligent guidance systems. The integration of deep learning architectures, particularly large-scale models, with edge-computing frameworks will enhance real-time decision-making capabilities. Model quantization and lightweight design facilitate deployment on embedded devices, ensuring adaptive navigation in dynamic underwater scenarios.

2) Integrated optical-acoustic communication guidance. The development of multi-quadrant photodetectors with high-frequency modulation capabilities enables unify positioning and communication functions. Utilizing optical communication's short-range high-bandwidth advantages while compensating for acoustic latency bridges the gap between near-field precision and long-range connectivity.

3) Multi-sensor fusion perception. The fusion of heterogeneous sensor data (e.g., GPS, INS, DVL) with optical guidance through advanced communication protocols and collaborative control algorithms enhances system performance. The incorporation of deep learning enables robust feature extraction and target perception, achieving centimeter-level accuracy and cross-domain sensor synergy. By synergizing these innovations, AUV underwater docking systems will evolve toward autonomous, resilient, and intelligent operation, unlocking new frontiers in marine exploration, infrastructure maintenance, and underwater robotics.

Key words optical guidance; autonomous underwater vehicle; image signal processing; position signal processing; water-related vision; water-related optics