

DOI:10.16356/j.2097-6771.2026.03.007

基于感兴趣区域线特征相似度的 移动机器人视觉检测对准

武星¹, 李松泽¹, 郭亮^{1,2}, 李兴达¹, 王晨涛¹

(1. 南京航空航天大学机电学院, 南京 210016; 2. 中国航发航空科技股份有限公司, 成都 610503)

摘要: 为了给进气道巡检机器人提供准确的作业位姿, 提出一种基于感兴趣区域线特征相似度的移动机器人视觉检测对准方法。通过移动机械臂末端安装的深度相机采集场景图片, 设计了融合深度学习目标识别与目标轮廓视觉测量的技术架构。先采用 YOLOv8 深度网络从场景图片中识别包含进气口的感兴趣区域 (Region of interest, ROI), 再设计改进随机直线检测 (Random line detection, RLD) 算法在 ROI 区域中检测进气口的轮廓直线。进一步, 提出了进气口四边形轮廓提取方法, 基于线特征相似度进行直线的精简、拼接, 并通过线端连接判断、轮廓闭合筛选、轮廓面积筛选, 提取进气口四边形轮廓。针对该轮廓建立移动机器人相对位姿估计模型, 并设计了位姿反馈控制器完成工位对准。搭建了进气口视觉对准模拟实验系统, 通过 ROI 识别、进气口直线检测、轮廓提取与对准控制实验, 验证了本文方法的可行性与有效性。

关键词: 移动机器人; 视觉对准; 感兴趣区域识别; 直线检测; 轮廓提取

中图分类号: TP242.2

文献标志码: A

文章编号: 1005-2615(2026)03-0537-12

Vision Detection and Alignment for Mobile Robots Based on Similarity of Feature Lines in Region of Interest

WU Xing¹, LI Songze¹, GUO Liang^{1,2}, LI Xingda¹, WANG Chentao¹

(1. College of Mechanical and Electrical Engineering, Nanjing University of Aeronautics & Astronautics, Nanjing 210016, China; 2. AECC Aero Science and Technology Co., Ltd., Chengdu 610503, China)

Abstract: To provide an accurate operating pose for an air-inlet inspection robot, a vision-based detection and alignment method based on line-feature similarity in the region of interest (ROI) is proposed. Scene images are acquired by a depth camera mounted on the end effector of a mobile manipulator. A technical framework combining deep-learning-based object recognition and machine-vision-based contour measurement is innovatively designed. First, a YOLOv8 network is used to identify the ROI containing the inlet opening from scene images. Second, an improved random line detection (RLD) algorithm is designed to detect the contour lines of the inlet opening within the ROI. Furthermore, a quadrilateral contour extraction method for the inlet opening is proposed, in which the detected lines are refined and spliced based on line-feature similarity, and the quadrilateral contour of the inlet opening is extracted through line-end connection checking, contour-closure screening, and contour-area screening. Based on the extracted contour, a relative pose estimation model of the mobile robot is established, and a pose-feedback controller is designed to accomplish workstation alignment. A simulated experimental system for air-inlet visual alignment is built, and experiments on ROI

基金项目: 国家自然科学基金面上基金(52475521); 国防基础科研项目重点项目(JCKY2022209B001)。

收稿日期: 2025-12-30; **修订日期:** 2026-04-23

通信作者: 武星, 男, 副教授, 硕士生导师, E-mail: wustar5353@nuaa.edu.cn。

引用格式: 武星, 李松泽, 郭亮, 等. 基于感兴趣区域线特征相似度的移动机器人视觉检测对准[J]. 南京航空航天大学学报(自然科学版), 2026, 58(3): 537-548. WU Xing, LI Songze, GUO Liang, et al. Vision detection and alignment for mobile robots based on similarity of feature lines in region of interest[J]. Journal of Nanjing University of Aeronautics & Astronautics(Natural Science Edition), 2026, 58(3): 537-548.

recognition, line detection, contour extraction, and alignment control verify the feasibility and effectiveness of the proposed method.

Key words: mobile robot; vision alignment; region-of-interest recognition; line detection; contour extraction

作为飞机推进系统的起始部件,进气道是支撑保护航空发动机的重要结构,具有捕获来流并按流量、流速等向发动机供气的作用^[1]。在飞机服役过程中,进气道可能因异物侵入、叶片疲劳损伤、管壁形变、铆钉脱落等问题存在较严重的安全隐患^[2]。而人工巡检仍是航空发动机进气道巡检的主要作业方式,作业效率低,劳动强度大,误检漏检时有发生。

不同于已现场应用的变电站/配电站/输电线路巡检机器人^[3-4],面向飞机运行维护的巡检机器人在国内还较为少见。文献[5]研究了基于数字孪生的飞机装配状态巡检机器人技术,通过模板匹配等图像处理算法判断线缆支架安装位置、角度是否正确。法国图卢兹大学^[6]针对飞机起飞前外部状态的巡检需求,研究了移动机器人全方位旋转变焦相机的图像处理与视觉识别方法,并用于空客 A320 飞机的氧气舱把手、进气口、静压口和风扇叶片等项目的自动巡检。

在智能制造与运维过程中,移动机器人的作业寻位过程可分为两个阶段:(1)移动机器人导航定位到作业工位,可采用常规的移动机器人全局定位方法;(2)移动机器人对准作业对象,需针对作业对象研究目标检测与对准控制方法。

移动机器人到达作业工位后,目标检测与定位对准技术是实现其与作业对象对准的关键。现有移动机器人对准方法大致可分为 3 类^[7]:(1)基于人工标识的对准方法,如 Wu 等^[8]利用射频识别(Radio frequency identification, RFID)标签及粒子滤波实现机器人与静态目标的相对位姿估计;(2)基于激光雷达点云配准的对准方法,如 Wang 等^[9]结合条纹投影扫描、激光定位与改进迭代最近点(Iterative closest point, ICP)算法完成大型构件点云配准;(3)基于视觉的自然特征对准方法,如 Luo 等^[10]通过实时提取目标图像中的自然特征点实现视觉引导与位姿控制。

虽然激光雷达和深度相机在目标对准中皆有应用,然而本文主要关注视觉检测对准方法,其原因在于该进气道巡检机器人采用机械臂末端安装的深度相机进行异物检测,而采用视觉检测对准方法可使作业对象的对准任务和巡检任务共用同一个视觉传感器,有利于降低系统硬件成本和技术复杂度。

在视觉目标检测领域,基于传统图像处理的目标识别方法通常依赖边缘提取、线段检测与几何约束的逐级处理流程。Von Gioi 等^[11]提出线段检测器(Line segment detector, LSD),可在无需复杂参数调节的条件下实现亚像素精度的线段提取,为规则目标轮廓分析提供了高效的底层结构表示。Song 等^[12]在 Canny 边缘检测框架中引入梯度方向与 Hough 变换进行边缘连接,增强了弱边缘的连续性与完整性。

近年来,基于深度学习的目标检测与区域定位方法发展迅速。Li 等^[13]提出了目标检测与分割的统一框架 Mask DINO 模型,实现了检测框与像素掩膜的联合预测。Cheng 等^[14]提出的 YOLO-World 模型,将 YOLO 框架扩展到开放词汇检测任务,展现了 YOLO 系列在高效检测与工程部署方面的优势。Lukács 等^[15]将 Attention U-Net 用于焊缝线的语义分割与几何尺寸评估,表明 U-Net 类网络在细粒度区域提取方面具有较强潜力。

由于本文视觉检测对准任务不仅需要识别进气口目标并确定其所在图像区域,还需精确测量其四边形轮廓并计算其对准中心点。因此,上述传统图像处理或深度学习类方法直接用于本文任务皆存在一定挑战:(1)传统图像处理方法在复杂场景下提取进气口轮廓的准确率低、处理时间长;(2)基于深度学习的目标检测方法提取的目标包围框与其实际几何轮廓存在较大轮廓误差,难以有效支撑后续四边形轮廓的精确恢复与位姿测量;(3)基于深度学习的目标边缘提取方法对高质量标注样本尤其是像素级掩膜标注依赖较强。在小样本情况下,分割模型容易受标注掩膜稀缺限制而影响训练稳定性与泛化能力^[16-17]。

针对该问题,本文提出了一种融合感兴趣区域(Region of interest, ROI)识别、直线检测、轮廓提取与位姿反馈控制的视觉检测对准方法。研究工作和创新如下:(1)创新设计了融合深度学习目标识别与目标轮廓视觉测量的技术架构,在复杂背景下实现了进气口区域的快速收缩搜索;(2)改进了随机直线检测(Random line detection, RLD)算法,提高了直线检测的鲁棒性;(3)提出了进气口四边形轮廓提取方法,实现了面向位姿估计任务的精确轮廓恢复。

1 视觉对准总体方案

移动机器人巡检作业寻位过程分为两个阶段。(1)粗定位阶段:搭载机械臂对准巡检装置的移动机器人由起始位置运行定位到巡检工位。(2)精定位阶段:调整移动机器人与待检测进气口的相对位姿,保证机械臂巡检装置对准进气口中心。由于阶段(1)可采用常规的移动机器人全局定位方法,本文在此不作深入讨论,主要研究阶段(2)巡检工位的视觉对准方法。

巡检机器人配备全局导航定位装置和机械臂对准巡检装置(深度相机)。先通过全局导航引导移动机器人到达进气口巡检工位附近,并保证移动机器人粗定位误差小于深度相机的视野范围。再由机械臂末端安装的深度相机获取进气口深度图像,通过视觉识别测量与伺服控制完成巡检机器人和进气口的对准,为后续的进气口巡检作业做好准备。

针对阶段(2)的巡检工位精定位,本文提出一种基于ROI线特征相似度的进气口视觉对准总体方案,包括ROI识别、进气口直线检测、进气口轮廓提取和进气口位姿对准等4部分,如图1所示。

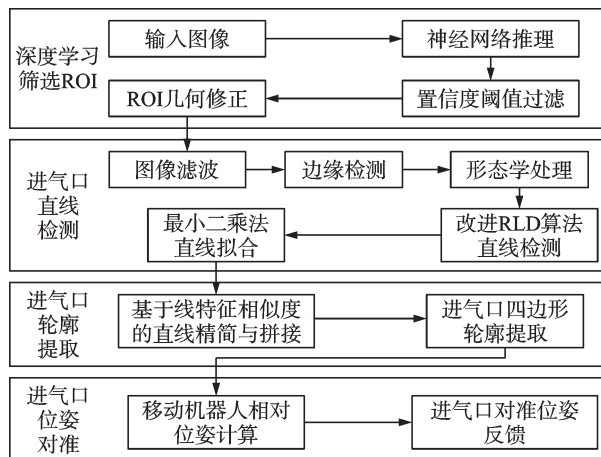


图1 视觉对准总体方案

Fig.1 Overall scheme of vision alignment

在ROI识别阶段,使用深度相机采集含有深度信息的进气口图像,基于深度学习方法识别出包含进气口的ROI区域,对ROI区域完成图像滤波、边缘检测及形态学处理等图像预处理操作。

在进气口直线检测阶段,采用改进RLD算法进行图像直线检测,获取进气口轮廓上的可能直线,并利用最小二乘法拟合直线方程,便于后续进气口的识别与提取。

在进气口轮廓提取阶段,基于线特征相似度进行可能直线的精简与拼接,在消除ROI中剩余冗余直线的同时,减少进气口轮廓的直线断裂问题。

进一步,通过线端连接判断、轮廓闭合筛选、轮廓面积筛选,提取进气口四边形轮廓。

在进气口位姿对准阶段,根据进气口四边形轮廓建立移动机器人相对位姿估计模型,获取移动机器人相对于进气口的侧向位置偏差和姿态角偏差。进一步,针对相对位姿偏差设计位姿反馈控制器,计算移动机器人的平移速度和转向角速度。并基于移动机器人逆运动学模型计算驱动轮角速度。

2 进气口ROI识别

传统直线检测方法通常在全图范围内进行边缘提取与直线搜索,易将背景中的结构边缘一并检出,从而引入大量冗余直线,影响后续轮廓提取与位姿估计的稳定性。为降低复杂背景对直线检测阶段的干扰,本文在进气口直线检测之前增加深度学习ROI识别阶段,先利用神经网络在图像中粗定位进气口的大致区域,然后在ROI内执行图像预处理与改进RLD直线检测及最小二乘拟合等操作,从而缩小搜索空间并提升算法的鲁棒性^[18]。

2.1 网络模型与训练策略

为满足移动机器人巡检工位对准场景下的实时性需求,本文采用轻量化单阶段目标检测网络YOLOv8实现ROI粗定位。单阶段检测网络可直接回归边界框并输出置信度,具有结构简洁、推理速度快等优点,适用于后续仍需结合传统图像处理方法进行精细几何提取的场景。

本文针对进气道模拟件采集所得的进气口图像构建ROI检测数据集。为满足深度学习训练需求,将数据集划分为训练集与验证集,并将标注转换为统一格式,便于网络训练与部署。鉴于样本数量有限,本文采用数据增强策略提升模型泛化能力,包括随机亮度、对比度扰动、随机尺度缩放、噪声扰动与模糊扰动等。数据增强仅作用于训练集,验证集保持原始分布用于客观评估。

考虑到训练样本规模较小,本文采用迁移学习策略:以在通用数据集上预训练的模型参数作为初始化权重,在进气口ROI数据集上进行微调训练。同时结合早停机制和学习率衰减策略,以提高训练效率并降低过拟合风险。YOLOv8训练阶段的优化目标可表示为边界框回归损失、分类损失与目标置信度损失的加权和,即

$$\mathcal{L} = \lambda_{\text{box}} \mathcal{L}_{\text{box}} + \lambda_{\text{cls}} \mathcal{L}_{\text{cls}} + \lambda_{\text{obj}} \mathcal{L}_{\text{obj}} \quad (1)$$

式中: \mathcal{L}_{box} 、 \mathcal{L}_{cls} 、 \mathcal{L}_{obj} 分别对应边界框回归、类别判别和目标存在性预测的误差, λ_{box} 、 λ_{cls} 、 λ_{obj} 分别为相应权重系数。本文ROI检测为单类别任务,分类分支可采用二元交叉熵形式,边界框回归项可采用交

并比(Intersection over union, IoU)类损失进行度量。以任意一对预测框 \hat{b} 与真实框 b 为例,其回归损失可简化表示为

$$\mathcal{L}_{\text{box}} = 1 - \text{IoU}(\hat{b}, b) \quad (2)$$

分类与目标置信度损失可分别写为

$$\mathcal{L}_{\text{cls}} = -[y \ln \hat{c} + (1 - y) \ln(1 - \hat{c})] \quad (3)$$

$$\mathcal{L}_{\text{obj}} = -[y \ln \hat{p} + (1 - y) \ln(1 - \hat{p})] \quad (4)$$

式中: $y \in \{0, 1\}$ 表示样本标签, \hat{c} 为进气口类别置信度输出, \hat{p} 为目标存在置信度输出。通过最小化式(1~4),网络可获得对进气口 ROI 的稳定检测能力,为后续的 ROI 区域识别与裁剪提供可靠输入。

2.2 ROI 目标识别

基于深度学习的 ROI 目标识别流程主要包括图像预处理、网络推理、候选框后处理、ROI 扩张等步骤,如图 2 所示。ROI 目标识别模块输出的裁剪图像将直接作为后续“进气口直线检测”模块的输入。相比在全图执行图像滤波、边缘检测、形态学处理与改进 RLD 直线检测等步骤,在 ROI 内处理可显著减少背景干扰边缘点数量,从而降低冗余直线数量并提升直线拟合与轮廓提取的成功率。具体步骤如下:

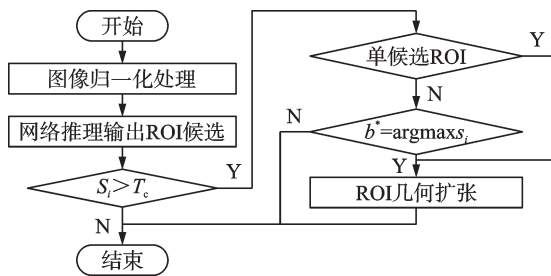


图2 ROI 识别流程图

Fig. 2 Flow chart of ROI recognition

(1)对深度相机采集的进气口图像进行尺寸归一化处理。为兼顾检测精度与实时性,本文将输入图像按固定尺寸进行缩放,并保持原始宽高比,避免几何形变对检测结果造成影响。随后对像素值进行归一化处理,为网络推理提供统一输入分布。

(2)将预处理后的图像输入目标检测网络,输出进气口 ROI 候选集合。每个候选由边界框与置信度组成,记候选框为

$$b_i = (x_{1i}, y_{1i}, x_{2i}, y_{2i}, s_i) \quad (5)$$

式中: (x_{1i}, y_{1i}) 与 (x_{2i}, y_{2i}) 分别为候选框左上角与右下角像素坐标, s_i 为置信度。

(3)为去除重复检测与低质量候选,先对候选框进行置信度阈值筛选($s_i > T_c$),再利用非极大值

抑制对重叠候选进行合并与去重,得到筛选后的候选集合。

(4)当存在多个候选时,本文采用“最高置信度优先”的策略选取最终 ROI 区域。

(5)考虑到检测框可能偏紧导致后续边缘/直线提取缺失,本文对最终 ROI 区域进行边界扩张,以检测框的宽度和高度作为基本尺度,分别沿水平方向和垂直方向向外扩张 8%。

3 进气口直线检测

3.1 图像预处理

本文所用图像预处理方法包括图像滤波、图像边缘检测与图像形态学处理,具体如下:(1)采用中值滤波对图像进行处理,从而抑制图像中的噪声强度,减弱噪声对边缘提取与特征识别的干扰;(2)利用 Canny 边缘检测算子^[19]提取图像中的进气口外轮廓边缘信息;(3)使用膨胀运算对图像边缘检测结果中断开的边缘线进行补全,同时去除错误提取的背景点干扰。

3.2 直线检测

RLD 算法可视为基于 Hough 变换算法的改进,基本原理是在边缘点集合中随机采样,以 3 个点作为一组种子点,根据距离阈值判断这 3 点是否能够构成一条直线,并通过其余点验证直线合理性,能显著提高算法的运行效率^[20]。

然而,RLD 算法使用固定的距离阈值判断 3 个点能否构成直线。当相机与目标物体之间的距离发生变化时,RLD 算法可能会出现误检漏检等问题。因此本文将 RLD 算法的直线判定依据由 3 点间距离阈值更换为 3 点间角度阈值,减少了相机与目标物体间距离变化对算法判定的影响,提升了算法的鲁棒性。基于改进 RLD 算法的直线检测流程如图 3 所示,具体步骤如下:

(1)假定图像中全部边缘点的数据集合为 $O = \{p_1, p_2, p_3, \dots, p_n\}$,其中, $p_i = \{x_i, y_i, z_i\}$ 。设置失败计数器为 f_{time} ,失败次数阈值为 T_{time} ,若构建直线失败次数达到阈值,则认为图像中所有的直线均已提取完成。设置角度阈值为 θ_s ,直线点数阈值为 T_s 。

(2)在数据集合 O 中进行随机采样,选取 3 个点 p_i, p_j, p_k 作为种子点,同时将这 3 个点从数据集合 O 当中取出。

(3)计算选取出的 3 点 p_i, p_j, p_k 中每两点之间的距离 L_i, L_j, L_k ,并对使用 3 点之中任意两点构建

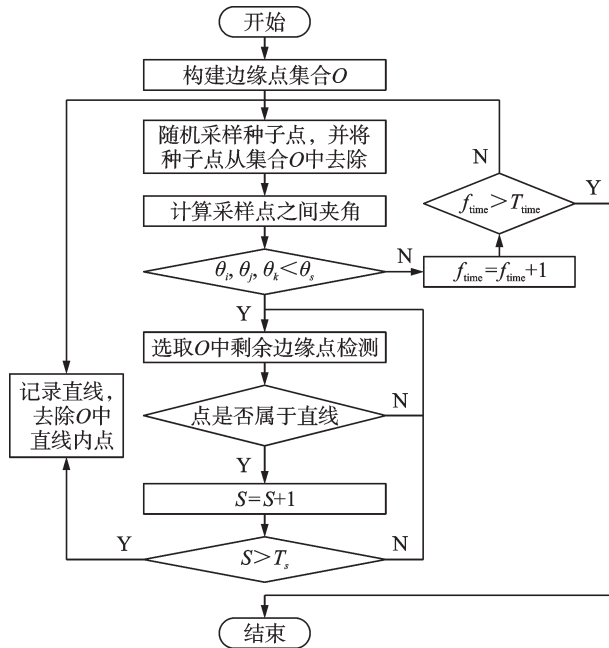


图 3 直线检测流程图

Fig.3 Flow chart of line detection

出的 3 条直线之间的夹角进行计算,计算公式为

$$\theta_i = \arccos \frac{L_j^2 + L_k^2 - L_i^2}{2L_jL_k} \quad (6)$$

$$\theta_j = \arccos \frac{L_i^2 + L_k^2 - L_j^2}{2L_iL_k} \quad (7)$$

$$\theta_k = \arccos \frac{L_i^2 + L_j^2 - L_k^2}{2L_iL_j} \quad (8)$$

(4)此时若 p_i, p_j, p_k 这 3 点共线,则 3 点所形成的三角形为面积接近于零的退化三角形。因此当上述 3 个夹角 $\theta_i, \theta_j, \theta_k$ 之中有一个角度小于阈值 θ_s , 说明这 3 点在几何上近似落在同一支撑直线上,则判断 p_i, p_j, p_k 属于同一条直线,并构建该直线,并进入步骤(5)。否则,将失败计数器 f_{time} 累加。判断 f_{time} 是否大于 T_{time} ,若成立,则结束算法运行,否则返回步骤(2)继续运行。

(5)设置直线点数计数器为 S ,对数据集合 O 中剩余的边缘点进行遍历,并判断是否处于上一步骤中构建出的直线值上,若成立,则将此边缘点视为该直线的内点,并从 O 中取出,对 S 进行累加。

(6)判断 S 是否大于 T_s ,若成立,则记录构建好的直线,并将直线上所有点从 O 中去除,并返回步骤(2)继续运行。若不成立,则返回步骤(5)继续运行。

3.3 直线拟合

当改进 RLD 算法从导航图像中检测出直线内点后,再利用最小二乘法对各组直线内点进行直线拟合,获得各条直线的数学模型

$$\left\{ \begin{aligned} k &= \frac{\sum_{i=1}^n y_i \sum_{i=1}^n x_i - \sum_{i=1}^n x_i y_i}{\left(\sum_{i=1}^n x_i\right)^2 - \sum_{i=1}^n x_i^2} \\ b &= \frac{\sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i - \sum_{i=1}^n \sum_{i=1}^n x_i^2}{\left(\sum_{i=1}^n x_i\right)^2 - \sum_{i=1}^n x_i^2} \end{aligned} \right. \quad (9)$$

式中: (x_i, y_i) 为属于该直线的内点; k, b 为二维直线的斜率与截距。

4 进气口轮廓提取

第 3 节从导航图像中检测了多条可能属于进气口轮廓的直线。为了实现进气口识别对准,本节从多条相互独立的候选直线中提取进气口四边形轮廓,如图 4 所示。实际上,航空发动机进气口具有不同的轮廓形状,本文以四边形轮廓为例研究进气口识别对准问题,并制作进气道模拟件用于移动机器人工位对准的实验验证。



(a) Quadrilateral-profile air inlet (b) Air-inlet analogue

图 4 进气道

Fig.4 Air-inlet

4.1 直线精简与拼接

针对 RLD 算法从导航图像中提取的直线边缘较粗、存在大量冗余及部分断裂问题,本节提出一种基于线特征相似度的直线精简与拼接方法,充分利用二维图像中的直线斜率、三维空间中的直线方向向量及三维空间中的直线距离等特征相似度,去除冗余直线,拼接断裂线段,避免直线检测误差对后续进气口轮廓提取的不利影响。基于线特征相似度的直线精简与拼接流程如图 5 所示。

(1)直线斜率计算:根据每条直线在图像像素坐标系下的起点与终点坐标,计算二维图像中的直线斜率 K_L 。

(2)斜率相似度判断:根据每条直线的斜率,计算二维图像中任意两条直线的斜率相似度 $S_{ij} = K_{L_i}/K_{L_j}$,并将斜率相似度较高的直线分为一类,计算为

$$|1 - S_{ij}| < T_s \quad (10)$$

式中 T_s 为斜率相似度阈值。

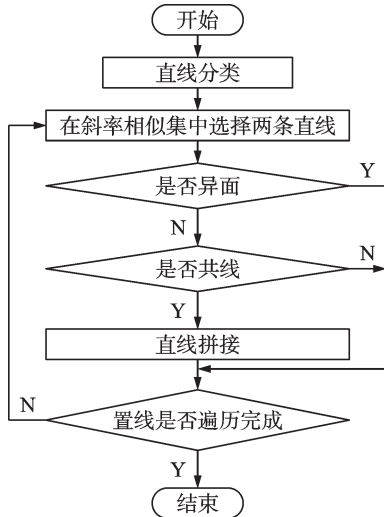


图5 直线拼接流程图

Fig.5 Flow chart of line splicing

(3)斜率相似度分类:将所有直线按照二维图像斜率分类后,处于相同斜率相似集中的两条直线可能具有3种空间关系:①两直线共线;②两直线平行;③两直线异面。因此,需要对斜率相似的两条直线做进一步判断。

(4)三维直线变换:在斜率相似集中按顺序取出两条直线 L_i 与 L_j 。根据深度相机标定时所获得的内参矩阵以及检测的像素点深度值,将直线起点与终点的像素坐标从图像坐标系变换至相机坐标系,从而将平面图像中的二维直线转换为立体空间中的三维直线,其计算公式为

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = Z_c \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} M_{col} \quad (11)$$

式中: Z_c 为图像像素点坐标对应的深度值, M_{col} 为内参矩阵。

(5)直线异面判断:根据三维直线的方向向量判断两条直线是否异面,其计算为

$$\left| \frac{X_i}{X_j} - \frac{Y_i}{Y_j} \right| < T_v \quad (12)$$

$$\left| \frac{X_i}{X_j} - \frac{Z_i}{Z_j} \right| < T_v \quad (13)$$

$$\left| \frac{Y_i}{Y_j} - \frac{Z_i}{Z_j} \right| < T_v \quad (14)$$

式中: T_v 为方向向量的相似度阈值; X 、 Y 、 Z 分别为 L_i 和 L_j 起点坐标与终点坐标的差值。若式(12~14)均成立,则这两条直线不为异面直线,进入步骤(6);否则这两条直线为异面直线,返回步骤(4)。

(6)直线共线判断:根据直线间距离判断两条直线是否共线。在直线 L_i 与 L_j 上以一定的距离间隔选取一系列的关键点,分别计算 L_i 的起点、终点

及关键点与 L_j 的起点、终点及关键点之间的距离。若其中存在一个距离满足距离阈值 T_d ,则两条直线共线,进入步骤(7);否则,两条直线不共线,返回步骤(4)。

(7)直线拼接:将判断为共线的两条直线 L_i 与 L_j 进行拼接。分别计算 L_i 的起点、终点和 L_j 的起点、终点中任意两点之间的距离,将其中相距最远的两点作为拼接后直线 L 的起点与终点,将拼接结果直线 L 保留,去除原直线 L_i 与 L_j ,如图6所示。

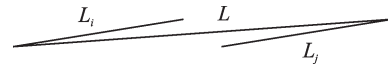


图6 直线拼接示意图

Fig.6 Schematic diagram of straight line splicing

(8)遍历终止条件判断:所有斜率相似集中的直线是否遍历完成。若未完成则返回步骤(4),否则结束直线拼接流程。

4.2 轮廓提取

直线精简与拼接方法可去除导航图像中的大部分冗余直线,并基本解决直线断裂问题,为进气口轮廓提取打下了基础。轮廓提取流程如图7所示,包括线端连接判断、轮廓闭合筛选、轮廓面积筛选等方法,具体步骤如下。

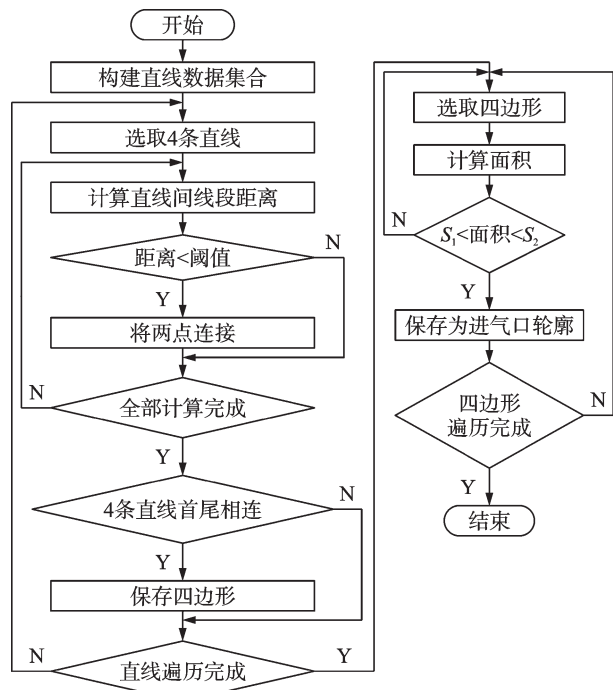


图7 轮廓提取流程图

Fig.7 Flow chart of contour extraction

(1)直线集合构建:假定导航图像中完成拼接后直线集合为 $S_l = \{L_1, L_2, L_3, \dots, L_n\}$,其中 $L_i = \{x_{1i}, y_{1i}, x_{2i}, y_{2i}\}$,而 (x_{1i}, y_{1i}) 和 (x_{2i}, y_{2i}) 为直线的起点与终点。

L515 深度相机,用于获取进气口近距图像与深度信息。导航控制工控机配置 Intel i7 处理器,通过 USB 与深度相机通信,并通过 CAN 总线与底盘驱动器通信,在机载端完成图像处理、相对位姿计算及运动控制。



图 10 模拟实验系统

Fig.10 Simulated experimental system

6.2 实验参数设定

实验测试验证分别从 ROI 识别、直线检测、轮廓提取和目标对准 4 方面展开。表 1 给出了本文方法在图像预处理、直线检测、轮廓拼接及位姿反馈控制中所涉及的关键参数。

表 1 本文方法的参数设置

Table 1 Parameter setting of proposed method

参数名称	取值	说明
T_{low}	50	Canny 低阈值
T_{high}	150	Canny 高阈值
$\theta_s/(^\circ)$	5	三点近似共线角度阈值
T_s	30	最小直线点数阈值
T_{time}	100	随机采样失败终止阈值
T_S	0.9	斜率相似度阈值
T_V	0.98	方向向量相似度阈值
T_d/mm	20	空间距离阈值
K_{Pd}	0.8	侧向偏差比例系数
K_{Dd}	0.12	侧向偏差微分系数
$K_{P\theta}$	1.5	姿态偏差比例系数
$K_{D\theta}$	0.2	姿态偏差微分系数

其中,Canny 边缘检测采用固定双阈值策略。这是因为本文并非直接在全图范围内进行边缘提取,而是先通过 YOLOv8 完成进气口区域的 ROI 定位,然后仅在 ROI 内进行中值滤波、边缘提取和后续直线检测。由于 ROI 内目标相对集中、背景干扰明显减弱,且实验平台中的目标对象、相机安装方式和观测工位条件相对稳定,经预实验比较后,取 $T_{low}=50$ 、 $T_{high}=150$ 可较好兼顾边缘完整性与噪声抑制效果。

改进 RLD 直线检测中的 θ_s 、 T_s 与 T_{time} 主要用于约束三点近似共线判定、短伪直线过滤以及随机采样终止条件。预实验结果表明,当 θ_s 过小时,真实共线点易因边缘抖动而无法通过判定;当其过大时,又会使拐角附近的点误判为共线。类似地, T_s 过小会保留过多噪声线段, T_s 过大则可能误删远距

离条件下的真实轮廓边。 T_{time} 过小会导致算法提前终止,过大又会增加无效计算时间。综合检测精度与运行效率后,最终选取 $\theta_s=5^\circ$ 、 $T_s=30$ 、 $T_{time}=100$ 。

在直线拼接阶段, T_S 、 T_V 和 T_d 分别从斜率一致性、方向一致性及空间邻近性 3 个方面约束候选线段的合并关系。若阈值设置过宽,容易发生不同物理边界之间的误拼接。若设置过严,则会使原本属于同一轮廓边的断裂线段无法连接。综合轮廓完整性与误匹配抑制效果,最终取 $T_S=0.90$ 、 $T_V=0.98$ 、 $T_d=20$ mm。

位姿反馈控制中的 PD 参数采用分步经验整定方法获取。经预实验整定,最终设置 $K_{Pd}=0.8$ 、 $K_{Dd}=0.12$ 、 $K_{P\theta}=1.5$ 、 $K_{D\theta}=0.2$,该组参数能较好满足移动机器人在进气口对准过程中的侧向偏差和姿态角偏差的纠偏控制要求。

6.3 ROI 识别实验

为了验证深度学习 ROI 识别模块在不同工位距离及背景干扰条件下的识别稳定性,本文利用图 10 所示的模拟实验系统开展 ROI 识别实验。深度相机在不同纵向距离下对进气道模拟件采集图像,同时改变环境光照条件,获得包含不同尺度、视角与噪声扰动的测试数据。图 11(a)为原始图像,图 11(b)为该图像的 ROI 识别结果。

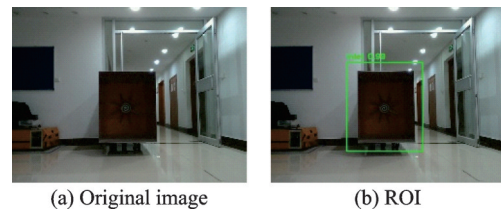


图 11 ROI 识别结果

Fig.11 ROI recognition results

实验结果表明,ROI 识别准确率在 1 m 距离下可达 99.3%,随着检测距离增大,目标尺度减小导致平均 IoU 略有下降,但在 5 m 距离下准确率仍保持在 98.6%,说明所用 ROI 识别模型具有较好的尺度鲁棒性与抗干扰能力。此外,ROI 识别过程仅包含单次前向推理与非极大值抑制等后处理步骤,单帧推理时间约 0.05 s,相比后续直线检测与轮廓提取阶段的计算开销较小,不会对整体视觉对准流程的实时性造成明显影响。

6.4 直线检测实验

为了验证本文改进 RLD 直线提取算法的有效性,将其与 Hough 变换算法、RLD 算法进行实验对比。先针对一张室外半结构化图像进行定性对比实验,检测结果如图 12 所示。其中图 12(a)为原始图像,图 12(b)为 Hough 变换算法检测结果,图 12

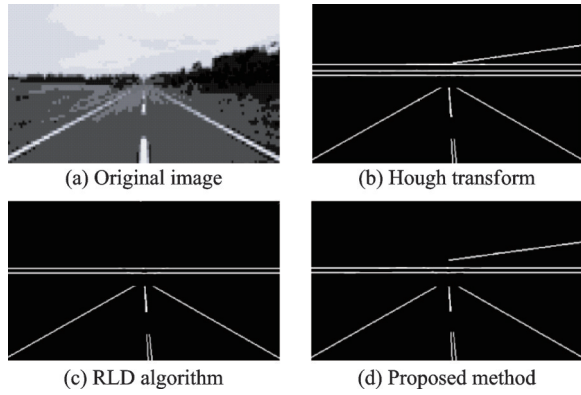


图 12 直线检测结果

Fig.12 Line detection results

(c)为 RLD 算法检测结果,图 12(d)为本文算法检测结果。

再针对存在图像干扰噪声的 100 张测试图片,分别使用 Hough 变换算法、RLD 算法与本文算法对测试图片进行直线检测,检测结果如表 2 所示。由表 2 中数据可知,本文改进 RLD 算法的准确率略优于 Hough 变换算法,优于 RLD 算法;其运行效率与 RLD 算法基本相当,优于 Hough 变换算法,具有更好的综合性能。

表 2 直线检测算法性能对比

Table 2 Performance comparison of line detection algorithms

使用方法	准确率/%	运行时间/s
Hough 变换算法	95	2.563
RLD 算法	90	1.778
本文算法	96	1.780

6.5 轮廓提取实验

为验证深度学习 ROI 识别方法对后续轮廓提取性能的提升效果,以及其对移动机器人纵向观测位置的鲁棒性,本文在不同纵向距离条件下分别开展全图识别与 ROI 内识别的轮廓提取对比实验。深度相机依次在 1、3 和 5 m 这三种距离下对进气道模拟件采集图像,轮廓提取流程均包含直线检测、直线精简与拼接、线端连接判断、轮廓闭合筛选与轮廓面积筛选等步骤。对比实验的区别在于使用 ROI 识别方法时,仅在网络输出的 ROI 区域内执行图像预处理、改进 RLD 直线检测与轮廓提取,有望降低背景干扰直线的影响。本文将两种方法的直线检测与轮廓提取测试结果进行对比,如图 13 所示为两种方法轮廓提取结果,左列为无 ROI 方法的提取结果,右侧为有 ROI 方法的提取结果。

在图 13 中,子图编号(a~c)分别为两种方法在 1、3 和 5 m 距离下的测试结果。在每种纵向位置,分别采集 300 帧进气口图像进行轮廓提取对比实验,测试结果如表 3 所示。由表 3 中数据可知,无 ROI 识别的传统进气口轮廓提取算法受纵向检测距离影响较大:在 1 m 距离下,由于目标在图像中占比较高、背景干扰较弱,传统方法仍可获得较好的轮廓提取效果。而在 3 与 5 m 距离下,随着目标尺度减小和背景结构边缘增多,全图范围内的边缘提取与直线搜索更易引入冗余直线,导致轮廓闭合筛选与面积筛选的稳定性明显下降。

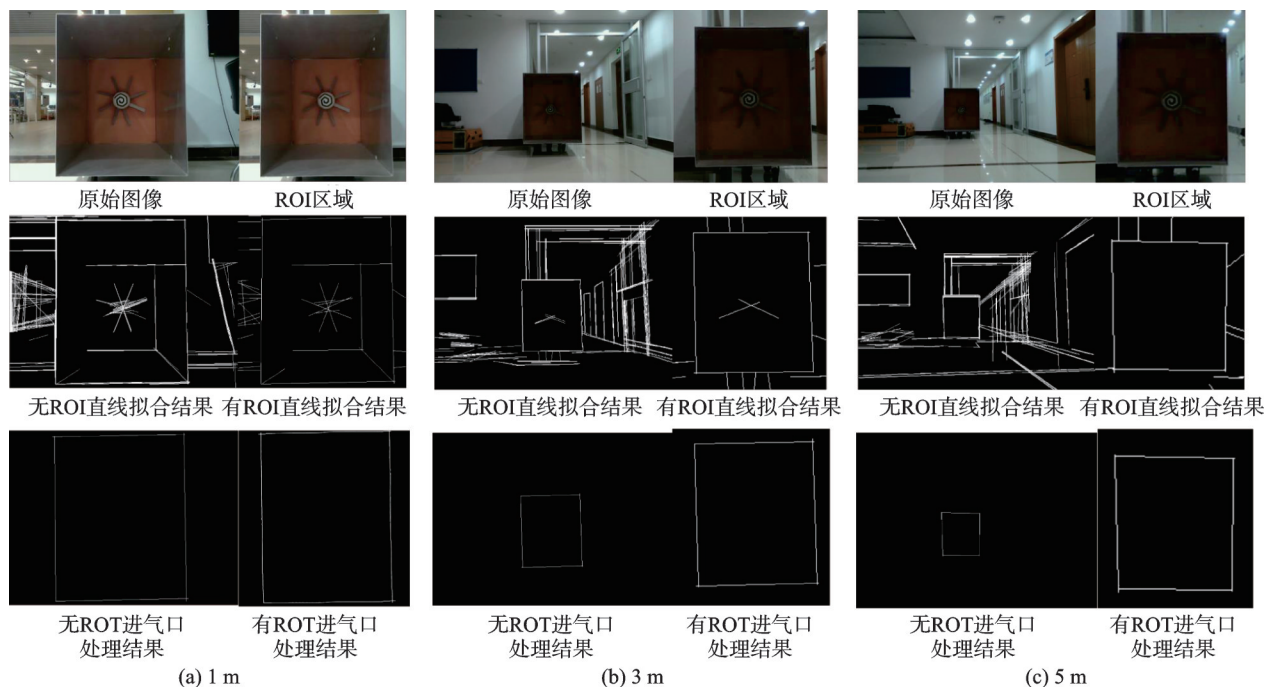


图 13 传统方法与 ROI 方法的轮廓提取结果对比

Fig.13 Comparison of contour extraction results between traditional methods and ROI method

表 3 进气口轮廓提取性能

Table 3 Performances of air-inlet port contour extraction

检测距离/m	无ROI 准确率/%	无ROI 处理时间/s	有ROI 准确率/%	有ROI 处理时间/s
1	97.6	0.825	97.8	0.628
3	90.1	3.471	97.2	0.644
5	76.5	16.384	96.3	0.649

相比之下,在引入ROI识别后,本文方法先利用深度学习模型完成进气口区域的语义粗定位,再在局部区域内执行直线检测与轮廓提取,从而有效抑制了复杂背景对后续几何结构图像处理的干扰,使本文方法在不同距离下的轮廓提取准确率均保持在较高水平,尤其在5 m距离下由76.5%提升至96.3%。

在算法处理效率方面,由于ROI识别方法减少了参与边缘检测与直线提取的像素范围与边缘点数量,改进RLD算法的随机采样与验证过程明显加速。同时,直线精简与拼接阶段参与相似度计算的候选直线数量减少,整体计算量明显降低。在5 m距离下处理时间由16.384 s降至0.649 s,与在1和3 m观测距离下的处理时间基本相当,这表明ROI识别方法相对于传统方法可显著提升远距离、强背景干扰场景的轮廓提取效率。

进一步对比了基于深度学习的目标检测方法(YOLO模型)的检测框中心点与本文深度学习目标识别+目标轮廓视觉测量方法的几何轮廓中心点相对于进气口中心点的定位误差。在3种不同纵向距离下,分别对成功检测样本中的进气口四角点进行人工标注,并以四角点构成四边形的几何中心作为参考真值,再计算YOLO检测框中心点与本文方法几何轮廓中心点的定位误差,如表4所示。

表 4 进气口中心定位误差

Table 4 Positioning errors of center point of inlet opening

检测距离/m	纯学习类 YOLO模型/mm	本文方法/ mm
1	7.4	2.1
3	18.2	2.6
5	33.7	3.5

由表4可见,纯学习类YOLO模型的定位误差在3种纵向距离下均明显大于本文学习+测量融合类方法的测量结果;并且随着纵向距离增加,两种方法定位精度的差距也愈发增大。这是因为

YOLO模型输出的是面向目标检测任务的矩形包围框,其中心位置主要反映目标所在区域的语义位置,而不严格对应目标轮廓的几何中心。若观察目标时存在视角变化、边界留白不对称或背景干扰等情况,检测框中心容易产生系统性偏移。

6.6 目标对准实验

移动机器人先通过全局导航定位装置移动至巡检工位附近,即视觉对准的起始点。移动机器人全局定位精度需保证进气口位于深度相机的视野范围内。移动机器人与进气道模拟件之间的相对位姿如图14(a)所示,以进气口的中点位置为原点建立坐标系,移动机器人的初始位置与进气口中点位置之间的距离在X轴方向上1 m,在Y轴方向上2 m,移动机器人姿态与Y轴的夹角为 30° 。因此,移动机器人的初始侧向位置偏差 $e_d=1$ m,初始姿态角偏差 $e_\theta=30^\circ$ 。移动机器人与进气口完成视觉对准后的相对位姿如图14(b)所示。

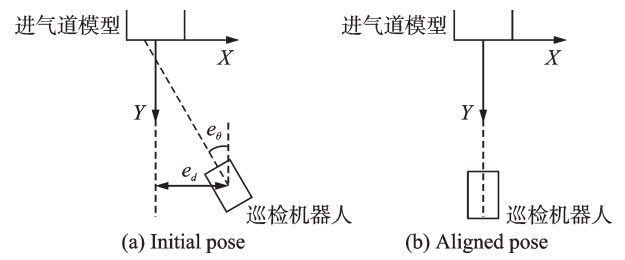


图 14 移动机器人视觉对准示意图

Fig.14 Diagram of vision alignment of mobile robots

在起始点启动移动机器人进行进气口视觉对准实验。通过进气口ROI识别、直线检测、轮廓提取、位姿测量与反馈控制,实现移动机器人的相对位姿偏差消除与工位对准,进气口对准的视频截图如图15所示。

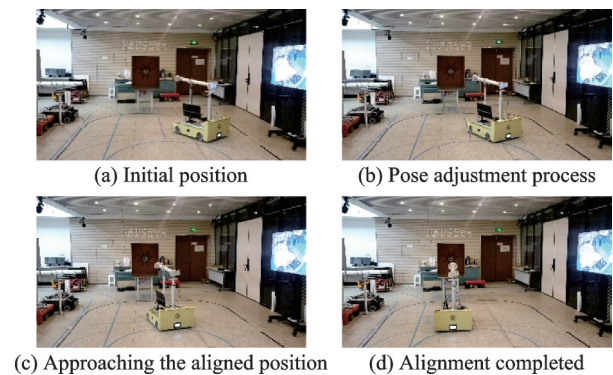


图 15 移动机器人视觉对准照片

Fig.15 Photos of vision alignment of mobile robots

在移动机器人视觉对准过程中,其与进气口之间的相对位姿偏差变化如图16所示。其中图16(a)为侧向位置偏差,图16(b)为姿态角偏差。由

图16中数据可知,视觉对准过程中移动机器人的相对位姿偏差逐渐减小。由于深度相机的测量误差以及移动机器人的控制误差,相对位姿偏差存在一些波动。当移动机器人到达巡检工位目标位姿附近时,侧向距离偏差减小到 ± 10 mm范围内,姿态角偏差减小到 $\pm 3^\circ$ 范围内,满足移动机器人进气道巡检的作业位姿精度要求^[21]。

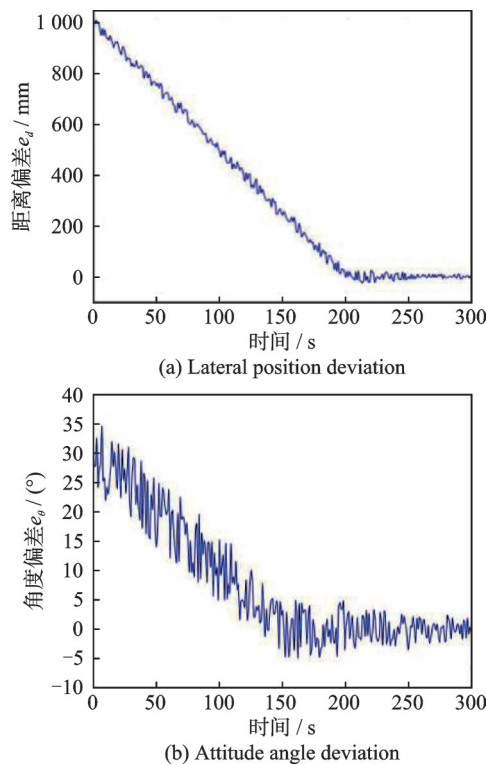


图16 相对位姿偏差数据图

Fig.16 Graphs of relative pose deviation

7 结 论

为了给进气道巡检机器人提供准确的作业位姿,提出一种基于ROI线特征相似度的移动机器人视觉检测对准方法。创新设计了融合深度学习目标识别与目标轮廓视觉测量的技术架构,先采用YOLOv8深度网络从场景图片中识别包含进气口的ROI区域,再改进用于进气口图像直线检测的RLD算法,在ROI区域内采用最小二乘法拟合进气口轮廓上的可能直线。进一步,基于线特征相似度进行直线的精简、拼接,并通过线段连接判定、轮廓闭合筛选及轮廓面积筛选,提取进气口四边形轮廓。根据进气口四边形轮廓计算移动机器人相对位姿,通过位姿反馈控制实现进气口对准。搭建了进气口视觉对准模拟实验系统,通过进气口ROI识别、直线检测、轮廓提取与对准控制实验,验证了本文方法的可行性与有效性。

参考文献:

- [1] HUANG Hexia, SUN Shu, YU Hang, et al. Recent progress in subsonic S-shaped inlets[J]. Journal of Propulsion Technology, 2020, 41(12): 2641-2658.
- [2] 王利敏,张彦军,米百刚,等. 涡桨飞机发动机进气道排除异物特性数值研究[J]. 航空工程进展, 2020, 11(2): 264-271, 278.
- [3] WANG Limin, ZHANG Yanjun, MI Baigang, et al. Numerical simulation on excluding foreign objects from engine inlet of turboprop aircraft[J]. Advances in Aeronautical Science and Engineering, 2020, 11(2): 264-271, 278.
- [4] CHEN Leisheng, WU Xing, SUN Chao, et al. An intelligent vision recognition method based on deep learning for pointer meters[J]. Measurement Science and Technology, 2023, 34: 055410.
- [5] CHEN Minghao, TIAN Yunong, XING Shiyu, et al. Environment perception technologies for power transmission line inspection robots[J]. Journal of Sensors, 2021, 2021: 5559231.
- [6] 侯正航,何卫平. 基于数字孪生的飞机装配状态巡检机器人的建模与控制[J]. 计算机集成制造系统, 2021, 27(4): 981-989.
- [7] HOU Zhenghang, HE Weiping. Modeling and control of digital twin-based aircraft assembly state inspection robot[J]. Computer Integrated Manufacturing Systems, 2021, 27(4): 981-989.
- [8] JOVANČEVIĆ I, LARNIER S, ORTEU J, et al. Automated exterior inspection of an aircraft with a pan-tilt-zoom camera mounted on a mobile robot[J]. Journal of Electronic Imaging, 2015, 24(6): 061110.
- [9] LIU Y, WANG S, XIE Y, et al. A review of sensing technologies for indoor autonomous mobile robots[J]. Sensors, 2024, 24(4): 1222.
- [10] WU Haibing, TAO Bo, GONG Zeyu, et al. A stand-alone RFID-based mobile robot navigation method using single passive tag[J]. IEEE Transactions on Automation Science and Engineering, 2020, 18(4): 1529-1537.
- [11] WANG Jinshan, TAO Bo, GONG Zeyu, et al. A mobile robotic 3-D measurement method based on point clouds alignment for large-scale complex surfaces[J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 7503011.
- [12] LUO B, CHEN H, QUAN F, et al. Natural feature-based visual servoing for grasping target with an aerial manipulator[J]. Journal of Bionic Engineering, 2020, 17(2): 215-228.
- [13] VON GIOI R G, JAKUBOWICZ J, MOREL J M,

- et al. LSD: A line segment detector[J]. *Image Processing On Line*, 2012, 2: 35-55.
- [12] SONG R, ZHANG Z, LIU H. Edge connection based canny edge detection algorithm[J]. *Pattern Recognition and Image Analysis*, 2017, 27(4): 740-747.
- [13] LI F, ZHANG H, XU H, et al. Mask dino: Towards a unified transformer-based framework for object detection and segmentation[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. [S.l.]: IEEE, 2023: 3041-3050.
- [14] CHENG T, SONG L, GE Y, et al. Yolo-world: Real-time open-vocabulary object detection[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. [S.l.]: IEEE, 2024: 16901-16911.
- [15] LUKÁCS H I, BEREGI B Z, PORTELEKI B, et al. Attention U-Net-based semantic segmentation for welding line detection[J]. *Scientific Reports*, 2025, 15: 15276.
- [16] SHEN W, PENG Z, WANG X, et al. A survey on label-efficient deep image segmentation: Bridging the gap between weak supervision and dense prediction [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, 45(8): 9284-9305.
- [17] ZHANG L, JINDAL B, ALAA A, et al. Generative AI enables medical image segmentation in ultra low-data regimes[J]. *Nature Communications*, 2025, 16(1): 6486.
- [18] 张睿, 王尧尧, 段雅琦, 等. 面向飞行机械臂的实时目标检测与定位算法[J]. *南京航空航天大学学报*, 2022, 54(1): 27-33.
ZHANG Rui, WANG Yaoyao, DUAN Yaqi, et al. Real-time object detection and location algorithm for aerial manipulator[J]. *Journal of Nanjing University of Aeronautics & Astronautics*, 2022, 54(1): 27-33.
- [19] 王福斌, 孙海洋, TU P. 边缘扩展的皮带撕裂支持向量机视觉检测[J]. *中国机械工程*, 2019, 30(4): 455-460.
WANG Fubin, SUN Haiyang, TU P. Visual inspection for extended edge belt tearing based on SVM[J]. *China Mechanical Engineering*, 2019, 30(4): 455-460.
- [20] 徐刚锋, 李彪, 沈振康. 一种提取直线的随机方法[J]. *中国图象图形学报*, 2003, 8(12): 1418-1421.
XU Gangfeng, LI Biao, SHEN Zhenkang. An efficient random algorithm for lines detection[J]. *Journal of Image and Graphics*, 2003, 8(12): 1418-1421.
- [21] 石世锋, 叶南, 吴哲, 等. 具有远近视距引导的机械臂多工位精确对准技术[J]. *南京航空航天大学学报*, 2022, 54(3): 419-427.
SHI Shifeng, YE Nan, WU Zhe, et al. A multi-station precise alignment technique of robotic arms guided by far-near-sight vision system[J]. *Journal of Nanjing University of Aeronautics & Astronautics*, 2022, 54(3): 419-427.

(编辑:陈琚)