

# 基于 CNN-RNN 模型的巡检机器人 位姿校正与建图技术研究

张 焰<sup>1,2</sup>

(1. 中曜达数能生态科技(浙江)有限公司, 杭州 310015;

2. 浙江大学 电气工程学院, 杭州 310000)

**摘要:** 针对现有巡检机器人位姿校正精度不足及环境建图效果不佳的问题, 提出一种融合多传感器数据与深度学习的位姿校正与定位建图技术, 以提高巡检机器人的定位精度与作业效率; 采用摄像头、轮编码器与陀螺仪采集位姿数据, 构建 CNN-RNN 复合模型 ( $3 \times 3$  卷积核、20 个卷积层、5 层 RNN 隐含层) 进行异构数据融合; 基于 SIFT 算法提取环境图像特征点, 通过变换矩阵配准图像, 利用 BiGEMAP 生成平面栅格地图; 引入动态概率模型和改进支持向量机优化区域通行判定; 消融实验表明, CNN-RNN 模型在 26 次迭代后位姿校正准确率达 99.5%; 图像拼接重叠区域比例最低为 1.02%, 区域通行误判率降至 1%, 所建地图与实际环境 SSIM 值达 1.0; 通过融合 CNN-RNN 模型与多传感器数据, 显著提升了位姿校正精度与环境建图质量, 为复杂场景下的巡检机器人自主导航提供了可靠技术支持。

**关键词:** 位姿校正; 机器人定位; 平面栅格地图构建; 巡检机器人; 消融实验

## Research on Pose Correction and Mapping Techniques for Inspection Robots Based on CNN-RNN Model

ZHANG Yan<sup>1,2</sup>

(1. Olympians Sunergy Ecological Technology (Zhejiang) Co., Ltd., Hangzhou 310015, China;

2. College of Electrical Engineering, Zhejiang University, Hangzhou 310000, China)

**Abstract:** To address the low accuracy of pose correction and poor performance of environmental mapping in existing inspection robots, a novel technique integrating multi-sensor data and deep learning was proposed to enhance positioning accuracy and operational efficiency. The pose data from cameras, wheel encoders, and gyroscopes were fused to build a composite model with convolution neural network and recurrent neural network (CNN-RNN), which is composed of  $3 \times 3$  convolution kernels, 20 convolutional layers, and 5 RNN hidden layers. Based on a scale-invariant feature transform (SIFT) algorithm, the environmental image feature were extracted, and the planar grid map was generated by transformation matrix matching image and BiGEMAP software. Introducing dynamic probabilistic models and improving support vector machines (SVM) to classify the passable region. Ablation experiments show that the pose correction accuracy of the CNN-RNN model reaches up to 99.5% after 26 iterations, the minimum overlapping ratio of stitched images is 1.02%, with a passability misjudgment rate of 1%, and the structural similarity (SSIM) value between the constructed map and the actual environment reaches up to 1.0. By integrating the CNN-RNN model and multi-sensor data, the precision of pose correction and the quality of environmental mapping are significantly improved, offering a robust solution for autonomous navigation of inspection robots in complex environments.

**Keywords:** pose correction; robot positioning; plane raster map construction; inspection robot; ablation experiment

## 0 引言

巡检机器人作为智能制造的重要组成部分, 近年来

在各行各业的应用越来越广泛, 其能够自主巡视、检测和监测设备、设施和环境, 通过制定好的路线自主移

收稿日期:2025-05-09; 修回日期:2025-06-09。

作者简介:张 焰(1973-),男,硕士,高级工程师。

引用格式:张 焰. 基于 CNN-RNN 模型的巡检机器人位姿校正与建图技术研究[J]. 计算机测量与控制, 2025, 33(7):295-303,320.

动, 并使用各种传感器和摄像头对设备和环境进行实时监测和数据采集<sup>[1]</sup>。但是, 巡检机器人执行任务过程中, 会受到多种因素的影响, 例如振动、光线等, 致使巡检机器人位姿出现一定的误差, 若是不对其进行校正处理, 一定会降低巡检机器人任务操作的质量与效率。此外, 巡检机器人需要在未知环境中进行自主导航和巡检, 因此需要对未知环境进行感知和建模<sup>[2]</sup>。通过定位建图技术, 机器人可以构建未知环境的平面栅格地图, 并根据平面栅格地图进行自主导航, 规划最优巡检路径, 提高巡检效率。与此同时, 机器人还可以根据实时环境信息调整巡检路径, 避免障碍物和潜在危险。由此可见, 如何设计一种精准的、高效的巡检机器人位姿校正与定位建图技术, 对于巡检机器人的未来发展具有至关重要的作用。

通常情况下, 巡检机器人应用环境较为复杂, 加之机器人内部传感器自身的缺陷或者问题, 导致其位姿校正与定位建图要求较高。就现有研究成果来看, 已有技术由于应用模型、算法或者技术的性能较差, 而无法保障巡检机器人位姿校正精度与定位建图效果。其中, 文献 [3] 针对工业生产环境中工件位姿多变、遮挡堆叠及大小各异等问题, 提出了使用 Res2Net 替换原残差结构的 PVNet 改进版, 通过回归像素级单位向量并经投票算法确定关键点, 最终利用 EPnP 算法求解工件位姿。设计技术依赖于深度相机的图像质量, 若图像受光照、反光或污染影响, 会导致特征提取不准确, 从而影响最终位姿估计的精度<sup>[3]</sup>, 此外 Res2Net 残差结构的引入会增加模型的复杂度, 也会提升模型参数优选的难度。; 文献 [4] 针对斜坡、特征退化及 GNSS 信号丢失等复杂环境下的连续精确定位问题, 提出结合地面约束与多传感器融合的 SLAM 方案, 通过优化关键帧选取策略、地面点约束及因子图优化, 提高定位精度与鲁棒性。设计技术在极端不平坦或完全无地面特征的环境中可能失效, 多传感器数据采集的融合还依赖同类传感器的相关历史数据和先验知识, 最终会导致使其适应性较差<sup>[4]</sup>; 文献 [5] 针对地下退化环境下的机器人状态估计不完备问题, 提出激光雷达—惯性 SLAM 框架, 结合惯性里程计与激光雷达—惯性里程计级联优化, 引入强度特征降低匹配误差, 通过退化检测确保位姿估计的鲁棒性。在极端尘埃、水雾等严重影响激光雷达信号质量的场景下, 点云特征匹配会失效, 导致提出技术定位精度下降<sup>[5]</sup>; 文献 [6] 引入 IPC 框架, 使机器人在人机交互中主动面对交互伙伴, 通过 3D 身份识别模块识别并选择人类伙伴, 根据任务状态执行动作。设计技术依赖于准确的 3D 身份识别与环境感知, 若环境中存在多个相似外观的人类或遮挡物, 会导致识别错误, 影响交互的流畅性<sup>[6]</sup>; 文献 [7] 针对变电站巡检机器人的

精确定位需求, 介绍了基于里程表、RFID 及机器视觉的定位算法, 并通过 RFID 对机器视觉算法进行改进。比较结果显示, 基于机器视觉的 RFID 定位算法精度最高, 定位误差稳定。RFID 标签的部署与维护成本较高, 且在大范围或复杂地形中, 标签的覆盖密度与位置选择成为挑战。此外, 机器视觉算法易受光照、天气等因素影响, 导致定位不稳定<sup>[7]</sup>。机器人位姿校正与定位建图对于高精度的巡检机器人控制更为重要, 如在电力领域需要利用巡检机器人采集高精度的图像, 以便于准确观测和监控电力系统的运行状态。为了解决上述问题或者缺陷, 提出面向巡检机器人的位姿校正与定位建图技术研究, 希望能够大幅度提升巡检机器人位姿校正的精度以及定位建图的质量, 为巡检机器人的快速发展提供动力支撑。本文的技术路线为: 机器人位姿数据采集→CNN-RNN 复合模型构建→数据特征提取→环境图像拼接→机器人定位建图的实现。本文创新性表现如下:

- 1) 利用 CNN-RNN 实现异构模型的融合, 提升了模型训练能量和记忆性能, 有助于提升机器人位姿校正的精度和效率;
- 2) 利用改进支持向量机实现动态概率建图, 并且提升了 CNN-RNN 模型特征分类的性能;
- 3) 基于 BiGEMAP 软件及通过引入动态概率权重降低图像拼接重叠率。

## 1 基于 CNN-RNN 模型的巡检机器人位姿校正

### 1.1 机器人位姿数据确定与采集

为了提升对巡检机器人位姿校正的精度, 采用摄像头、轮子编码器与陀螺仪采集机器人位姿数据——图像数据、里程计数据与角速度数据, 采用同一时刻的数据并用时间戳标定, 同步对采集到的全部数据进行预处理, 为巡检机器人位姿估计提供精准数据支撑。通过摄像头采集巡检机器人图像数据, 记为  $I(x, y)$ <sup>[8]</sup>。由于光线、温度、湿度、图像采集设备振动、电磁干扰等因素的影响, 图像数据存在着信噪比较低的缺陷, 故采用双边滤波算法对图像数据进行去噪处理, 表达式为:

$$\hat{I}(x, y) = \frac{\sum \omega(x, y) \otimes I(x, y)}{\sum \omega(x, y)} \quad (1)$$

式中,  $\otimes$  为卷积符号,  $\hat{I}(x, y)$  代表经过去噪处理后的图像像素分布。其中,  $\omega(x, y)$  代表像素点  $(x, y)$  的权重系数, 取值在 0—1 之间,  $\omega(x, y)$  系数的取值与当前像素点邻域像素点的密度值、灰度值相关, 且系数数值的动态变化与邻域像素点的密度值、灰度值成正比。巡检机器人里程计数据主要通过轮子编码器获取, 用于估计机器人在环境中的相对位移<sup>[9]</sup>。里程计数据计算公式为:

$$\begin{cases} d = r \times \alpha_c \\ \theta = \frac{d_L - d_R}{D_w} \end{cases} \quad (2)$$

式中， $d$  代表巡检机器人轮子的位移； $r$  代表巡检机器人轮子的直径； $\alpha$  代表编码器脉冲数； $\theta$  代表巡检机器人轮子的转向角度； $d_L$  与  $d_R$  代表巡检机器人左轮位移与右轮位移； $D_w$  代表巡检机器人轮子之间的距离。

巡检机器人角速度数据主要通过陀螺仪获取，用于测量机器人绕某轴的旋转速率。角速度数据计算公式为：

$$\varphi = v \times \Delta t \quad (3)$$

式中， $\varphi$  代表巡检机器人旋转角度； $v$  代表角速度； $\Delta t$  代表数据采集时间间隔。上述过程完成了巡检机器人位姿数据的确定及其采集，为巡检机器人位姿的综合估计提供依据。分别对位移  $d$  和旋转角度  $\varphi$  进行归一化处理，提升数据的一致性：

$$\hat{d} = \frac{d - d_{\min}}{d_{\max} - d_{\min}} \quad (4)$$

式中， $\hat{d}$  为归一化处理后的位移值， $d_{\max}$  和  $d_{\min}$  分别为位移数据的最大值和最小值。

$$\hat{\varphi} = \frac{\varphi - \varphi_{\min}}{\varphi_{\max} - \varphi_{\min}} \quad (5)$$

式中， $\hat{\varphi}$  为归一化处理后的角度值， $\varphi_{\max}$  和  $\varphi_{\min}$  分别为角度数据的最大值和最小值。

### 1.2 CNN-RNN 模型设计

机器人位姿数据的采集为位姿检测和校正提供前提和基础，针对巡检机器人的位姿校正问题，设计技术有效地融合卷积神经网络 (CNN) 与循环神经网络 (RNN)<sup>[10]</sup>。CNN 在图像数据训练和特征提取方面具有显著优势，而 RNN 模型具有良好的记忆功能，在处于较长的非线性数据方面能力较强，CNN-RNN 复合模型能够充分利用 CNN 在图像特征提取方面的优势，以及 RNN 在处理时间序列数据上的能力，并抑制 CNN 与 RNN 的缺陷，从而实现巡检机器人位姿的精确校正。CNN-RNN 复合模型如图 1 所示。

如图 1 所示，CNN-RNN 复合模型主要由五个部分构成，分别为输入部分 (1 个)、CNN 部分 (1 个)、RNN 部分 (2 个) 与输出部分 (1 个)。CNN 部分主要是通过卷积层、池化层与全连接层的协同操作，再基于图像数据获取巡检机器人精确的位姿信息；RNN 部分 (2 个) 主要是通过循环层对里程计数据与角速度数据进行操作与处理，获得相应的巡检机器人位姿信息；输出部分以 CNN 部分与 RNN 部分 (2 个) 输出结果为基础，RNN 的单元为 LSTM (长短时记忆网络)，综合估计当前时刻巡检机器人位姿，并计算其与参考位姿之间的误差，以此为基础，对位姿进行针对性校正。

为避免 CNN 模型由于复杂度过高而出现性能衰退和稳定性降低的情况，本文在 CNN 模型中采用了动态学习率，即模型的学习率取值在 0-1 之间，在不同的

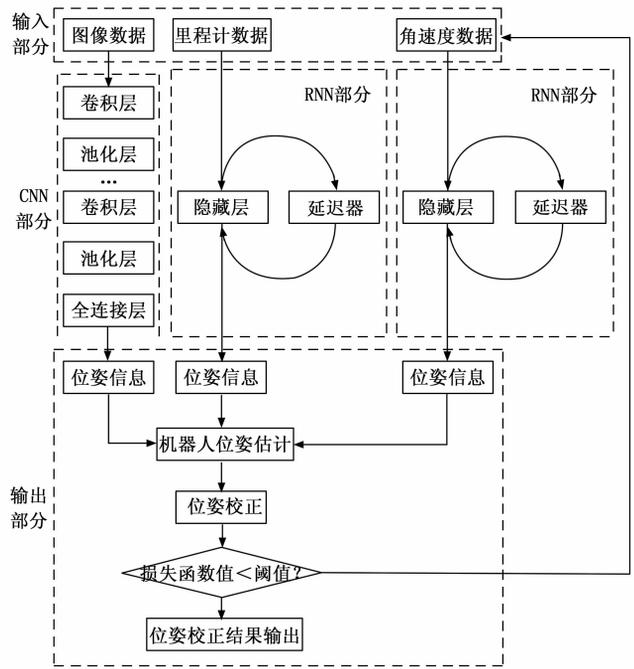


图 1 CNN-RNN 复合模型

训练阶段匹配不同的学习率值，如图 2 所示。

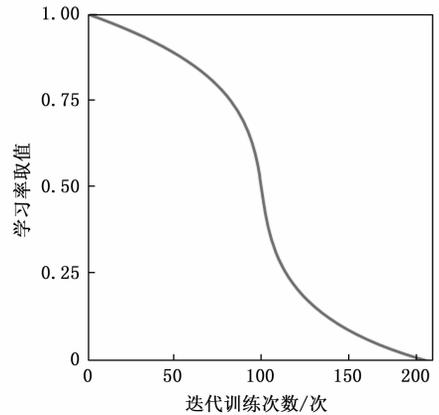


图 2 CNN 模型学习率动态变化示意图

CNN 部分的卷积层数量、池化层数量与 RNN 部分保持相同，以确保模型的复杂度一致。在隐含层的关键参数，如各层权重值、偏置值等，CNN 部分和 RNN 部分保持联动，即在关键参数上保持一致，以便于基于 RNN 模型提升 CNN 模型的数据训练记忆能力，进而提升整体模型的数据特征提取能力。与此同时，以均方误差 (MSE) 为损失函数，衡量复合模型位姿校正结果与真实位姿之间的差异<sup>[11]</sup>。当损失函数值小于阈值时，说明 CNN-RNN 复合模型性能较优，模型输出结果即为最终的位姿校正结果；当损失函数值大于或者等于阈值时，说明 CNN-RNN 复合模型性能较差，需对其进行继续迭代训练。

CNN-RNN 复合模型损失函数表达式为:

$$MSE = \frac{1}{N} \sum_{i=1}^N (P_i - \hat{P}_i)^2 \quad (6)$$

式中,  $MSE$  代表巡检机器人位姿校正结果的均方误差;  $N$  代表样本数量;  $P_i$  代表第  $i$  个样本的位姿校正结果, 记为  $P_i = (X_i, Y_i, \phi_i)$ 。其中,  $(X_i, Y_i)$  代表巡检机器人的位置坐标信息;  $\phi_i$  代表巡检机器人的朝向信息;  $\hat{P}_i$  代表第  $i$  个样本的真实位姿, 记为  $\hat{P}_i = (\hat{X}_i, \hat{Y}_i, \hat{\phi}_i)$ 。

卷积神经网络 (CNN) 与循环神经网络 (RNN) 的融合应用, 可以发挥 CNN 与 RNN 的优势, 抑制 CNN 与 RNN 的缺陷, 并行、精准处理图像数据、里程计数据与角速度数据, 实现对巡检机器人位姿的综合估计与精确校正。

### 1.3 机器人位姿校正方法

接下来将 1.1 节采集的巡检机器人位姿数据作为 CNN-RNN 复合模型的输入数据, 此模型输出结果即为巡检机器人位姿校正结果, 复合模型具体应用流程如下:

1) 输入部分对巡检机器人位姿数据——图像数据  $\hat{I}(x, y)$ 、里程计数据  $(d, \theta)$  与角速度数据  $\varphi$  进行接收, 并将其传输给 CNN 部分与 RNN 部分进行进一步的处理。

2) CNN 部分接收图像数据  $\hat{I}(x, y)$  后, 使用卷积层对其进行卷积操作, 提取图像数据中的局部特征<sup>[13]</sup>, 表达式为:

$$O(x, y) = \sum_m \sum_n \hat{I}(x+m, y+n) \cdot K(m, n) \quad (7)$$

式中,  $O(x, y)$  代表图像数据  $\hat{I}(x, y)$  的局部特征图;  $K(m, n)$  代表卷积核函数;  $(m, n)$  代表卷积核上的位置索引。池化层采用下采样方式对局部特征图  $O(x, y)$  进行继续处理, 在保留局部特征图中重要特征的基础上降低特征数据的整体维度。池化层操作行为表达式为:

$$\hat{O}(x, y) = \max_{(x, y) \in R(x, y)} O(x, y) \quad (8)$$

式中,  $\hat{O}(x, y)$  代表局部特征图  $O(x, y)$  的下采样处理结果;  $R(x, y)$  代表局部特征图上位置  $(x, y)$  对应的池化窗口。全连接层将最后一个池化层输出的局部特征图  $\hat{O}(x, y)$  展平, 并对其进行特征的进一步组合和转换, 从而获取巡检机器人的位姿信息, 表达式为:

$$(x_1, y_1) = \varphi[W \cdot \hat{O}(x, y) + B] \quad (9)$$

式中,  $(x_1, y_1)$  代表基于图像数据  $\hat{I}(x, y)$  获得的巡检机器人位置信息;  $\varphi[\cdot]$  代表激活函数;  $W$  代表权重矩阵;  $B$  代表偏置向量。

3) RNN 部分 (第 1 个) 接收里程计数据  $(d, \theta)$  后, 通过隐藏层的状态更新来捕捉时间序列数据中的依赖关系。循环神经网络 (RNN) 运行过程中, 当前时间步的隐藏状态主要是根据上一个时间步的隐藏状态与当前输入向量进行计算, 计算公式为:

$$h_i = F[W_h \cdot h_{i-1} + v \cdot (d, \theta) + b] \quad (10)$$

式中,  $h_i$  与  $h_{i-1}$  代表当前时间步、上一个时间步的隐藏状态;  $W_h$  与  $v$  代表权重矩阵;  $F[\cdot]$  代表激活函数;  $(d, \theta)$  代表当前时间步的输入向量, 即里程计数据;  $b$  代表偏置向量。

在 RNN 处理完所有时间步的里程计数据后, 从最后一个时间步的隐藏状态中提取巡检机器人位姿信息, 表达式为:

$$\begin{cases} (x_d, y_d) = g[h_i] \\ \phi_d = \delta[h_i] \end{cases} \quad (11)$$

式中,  $(x_d, y_d)$  代表基于里程计数据  $(d, \theta)$  获得的巡检机器人位置信息;  $g[\cdot]$  代表位置信息提取函数;  $\phi_d$  代表基于里程计数据  $(d, \theta)$  获得的巡检机器人朝向信息;  $\delta[\cdot]$  代表朝向信息提取函数。

4) RNN 部分 (第 2 个) 接收角速度数据  $\varphi$  后, 按照上一步骤所示原理提取巡检机器人位姿信息, 记为  $\phi_\varphi$ 。

5) 输出部分以 CNN 部分与 RNN 部分 (2 个) 输出结果为基础, 综合估计巡检机器人的位姿信息, 表达式为:

$$\begin{cases} (\bar{X}, \bar{Y}) = (\beta_1 x_1 + \beta_2 x_d, \beta_1 y_1 + \beta_2 y_d) \\ \bar{\phi} = \chi_1 \phi_d + \chi_2 \phi_\varphi \end{cases} \quad (12)$$

式中,  $(\bar{X}, \bar{Y})$  与  $\bar{\phi}$  代表巡检机器人位置信息与朝向信息的综合估计结果;  $\beta_1$  与  $\beta_2$ 、 $\chi_1$  与  $\chi_2$  代表权重系数。以巡检机器人位姿信息估计结果与参考位姿之间的误差为基础, 对位姿信息进行校正处理, 表达式为:

$$\begin{cases} (X, Y) = (\bar{X} + \Delta X, \bar{Y} + \Delta Y) \\ \phi = \bar{\phi} + \Delta \phi \end{cases} \quad (13)$$

式中,  $(X, Y)$  与  $\phi$  代表位姿信息校正结果采用机器人自带的激光器校准位姿 (每 10 s 校正一次);  $(\Delta X, \Delta Y)$  与  $\Delta \phi$  代表位姿信息估计结果与参考位姿之间的误差。上述过程完成了巡检机器人位姿的精准校正, 获取高精度的巡检机器人位姿信息  $P = (X, Y, \phi)$  (即定位信息), 为后续巡检机器人定位建图奠定坚实的基础。

## 2 巡检机器人建图

### 2.1 获取周围环境图像

对机器人位姿数据进行采集和校正后, 进一步获取机器人的定位信息, 以巡检机器人的定位信息  $P = (X, Y, \phi)$  为基础, 通过调整机器人朝向获取环境图像  $\{f_i(x, y), i = 1, 2, \dots, M\}$ ,  $M$  代表环境图像的总数量。提取并匹配环境图像  $f_i(x, y)$  特征点, 计算相邻环境图像之间的变换矩阵, 以此为基础, 对环境图像进行配准与拼接, 从而获取巡检机器人作业的完整环境图像, 为后续可通行区域与不可通行区域判定提供支撑。环境图像特征点指的是图像中稳定性较高、显著性较强的像素

点, 对其进行提取是环境图像配准与拼接的关键所在。需要注意的是, 特征点必须是显著元素, 在不同的视角、光照和尺度变化下仍能保持稳定<sup>[14]</sup>。设计技术基于 SIFT 算法提取环境图像的特征点, 提取结果为:

$$q_{i,k} = \{(x_{i,k}, y_{i,k}, \sigma_{i,k}, \delta_{i,k}), k = 1, 2, \dots, Q\} \quad (14)$$

式中,  $q_{i,k}$  代表第  $i$  个环境图像的第  $k$  个特征点;  $Q$  代表特征点的总数量;  $\sigma_{i,k}$  代表特征点的尺度信息;  $(x_{i,k}, y_{i,k})$  代表特征点的坐标信息;  $\delta_{i,k}$  代表特征点描述符。

与此同时, 提取相邻环境图像的特征点, 记为  $q_{i+1,l} = (x_{i+1,l}, y_{i+1,l}, \sigma_{i+1,l}, \delta_{i+1,l})$ 。设计技术以欧式距离来衡量特征点之间的相似度, 以此为基础, 对相邻环境图像中的特征点进行匹配<sup>[15]</sup>, 表达式为:

$$\hat{d}(q_{i,k}, q_{i+1,l}) = \sqrt{(\delta_{i,k} - \delta_{i+1,l})^2} \quad (15)$$

式中,  $\hat{d}(q_{i,k}, q_{i+1,l})$  代表特征点  $q_{i,k}$  与  $q_{i+1,l}$  之间的欧式距离。当  $\hat{d}(q_{i,k}, q_{i+1,l})$  大于阈值  $\bar{d}$  时, 认定特征点  $q_{i,k}$  与  $q_{i+1,l}$  匹配; 当  $\hat{d}(q_{i,k}, q_{i+1,l})$  小于或者等于阈值  $\bar{d}$  时, 认定特征点  $q_{i,k}$  与  $q_{i+1,l}$  不匹配。

基于匹配特征点  $(q_{i,k}, q_{i+1,l})$  计算相邻环境图像之间的变换矩阵, 以便将环境图像  $f_{i+1}(x, y)$  变换到与环境图像  $f_i(x, y)$  相同的坐标系下, 为后续环境图像拼接提供便利。相邻环境图像变换矩阵计算公式为:

$$\mathbf{T}_i = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix} \quad (16)$$

式中,  $\mathbf{T}_i$  代表环境图像  $f_i(x, y)$  与  $f_{i+1}(x, y)$  之间的变换矩阵;  $t_x$  与  $t_y$  代表  $f_{i+1}(x, y)$  相对于  $f_i(x, y)$  的平移因子;  $a_{11}$  与  $a_{21}$  代表  $f_{i+1}(x, y)$  相对于  $f_i(x, y)$  的旋转因子;  $a_{12}$  与  $a_{22}$  代表  $f_{i+1}(x, y)$  相对于  $f_i(x, y)$  的缩放因子。需要注意的是, 旋转参数、缩放参数与平移参数均需要根据匹配特征点  $(q_{i,k}, q_{i+1,l})$  之间的位置、方向等信息进行确定。

依据变换矩阵对环境图像  $f_i(x, y)$  进行变换处理, 即  $f'_{i+1}(x, y) = f_{i+1}(x, y) * \mathbf{T}_i$ , 变换处理后的环境图像  $f'_{i+1}(x, y)$  与环境图像  $f_i(x, y)$  位置保持对齐状态<sup>[16]</sup>。设计技术利用自适应权重拼接算法对环境图像进行拼接及其融合, 通过计算重叠区域内像素的加权平均值来实现重叠区域及其拼接痕迹的消除, 从而获取完整的、平滑的巡检机器人环境图像, 表达式为:

$$g(x, y) = \sum_{i=1}^M \kappa_i * f'_i(x, y) \quad (17)$$

式中,  $g(x, y)$  代表环境图像拼接、融合结果;  $f'_i(x, y)$  代表变换处理后的环境图像;  $\kappa_i$  代表环境图像  $f_i(x, y)$  拼接过程中重叠区域内像素的权重系数, 基于当前时刻特征点匹配置信度对权重系数值进行适时调整, 当发现特征点匹配度降低时, 需要重新匹配特征点, 直到匹配度满足需求。上述过程完成了巡检机器人作业环境图像

的采集与拼接, 输出结果为完整的、平滑的环境图像, 为后续区域通行状态判定提供清晰的图像数据依据。

## 2.2 判定通行区域

以上一节输出的巡检机器人作业完整环境图像  $g(x, y)$  为基础, 将其划分为均匀的二维栅格区域, 对区域内部特征进行提取, 以此为基础, 判定二维栅格区域是否能够通行, 为环境平面栅格地图构建提供助力。

设定环境图像  $g(x, y)$  宽度为  $W_g$ , 高度为  $H_g$ , 二维栅格区域尺寸为  $s \times s$ , 则环境图像划分区域总数量的计算公式为:

$$N_\Omega = \frac{W_g \times H_g}{s \times s} \quad (18)$$

式中,  $N_\Omega$  代表环境图像划分区域总数量。

将环境图像区域划分结果记为  $\{\Omega_i(x, y), i = 1, 2, \dots, N_\Omega\}$ , 提取其内部特征——亮度特征、边缘密度特征、纹理特征 (对比度与熵)<sup>[17]</sup>, 表达式为:

$$\begin{cases} L_\Omega = \frac{\sum \Omega_i(x, y)}{s \times s} \\ \rho_\Omega = \frac{\sum \eta[E_\Omega(x, y)]}{s \times s} \\ \bar{T}_\Omega = \sum (x - y)^2 \cdot P_\Omega(x, y) \\ e_\Omega = - \sum P_\Omega(x, y) \cdot \log_2 P_\Omega(x, y) \end{cases} \quad (19)$$

式中,  $L_\Omega$  代表区域  $\Omega_i(x, y)$  的亮度特征。障碍物通常比背景更暗或更亮, 因此亮度特征可以作为一个初步的判定依据;  $\rho_\Omega$  代表区域  $\Omega_i(x, y)$  的边缘密度特征。边缘元素通常表示物体的轮廓, 因此边缘密度高的区域可能包含障碍物, 故将边缘密度特征看作区域是否能够通行的主要判定依据之一;  $\eta[\cdot]$  代表指示函数, 取值为 0 或者 1;  $E_\Omega(x, y)$  代表边缘检测算子在  $(x, y)$  处的输出;  $\bar{T}_\Omega$  与  $e_\Omega$  代表区域  $\Omega_i(x, y)$  的纹理特征 (对比度与熵), 其能够反映区域内像素的空间分布和排列规律, 有助于区分不同地物, 为区域是否能够通行判定提供一定依据;  $P_\Omega(x, y)$  代表灰度共生矩阵中位置  $(x, y)$  的元素值。

将上述提取的环境图像区域特征进行整合, 记为  $A_\Omega = \{L_\Omega, \rho_\Omega, \bar{T}_\Omega, e_\Omega\}$ , 以支持向量机为工具, 确定区域能否通行结果。原始支持向量机训练过程可以作为一个优化问题, 致使其训练过程过于复杂、冗余, 势必会增加设计技术的运算体量, 降低区域能否通行判定的效率。因此, 设计技术引入拉格朗日乘子将支持向量机优化问题转换为对偶问题, 并利用核技巧将其扩展到非线性分类, 最大限度地改进与优化支持向量机性能<sup>[18]</sup>。

基于支持向量机的区域能否通行判定函数公式为:

$$\lambda(A_\Omega) = \text{sgn} \left[ \sum_{j=1}^{N_{\text{sup}}} \vartheta_j \Phi_j \cdot C(A_\Omega, A_\Omega) + b_0 \right] \quad (20)$$

式中,  $\lambda(A_\Omega)$  代表区域  $\Omega_i(x, y)$  能否通行判定结果, 其

取值为 0 或者 1。当  $\lambda(A_{\Omega})$  取值为 0 时, 判定区域  $\Omega_i(x, y)$  不可通行; 当  $\lambda(A_{\Omega})$  取值为 1 时, 判定区域  $\Omega_i(x, y)$  可通行;  $sign[\cdot]$  代表符号函数, 核函数选择 RBF 径向基核函数, 定义样本之间的特征相似度并在特征空间内映射前后样本之间的距离比例;  $N_{total}$  代表训练样本的总数量;  $\vartheta_j$  代表拉格朗日乘子的解, 将拉格朗日约束条件加入到 RBF 径向基核函数之中, 并对函数求导得到优化的解。  $\Phi_j$  代表第  $j$  个训练样本对应的标签, 即可通行区域或不可通行区域;  $C(A_{\Omega_i}, A_{\Omega_j})$  代表核函数, 其主要用于计算训练样本特征  $A_{\Omega_i}$  和测试样本特征  $A_{\Omega_j}$  之间的相似度;  $b_0$  代表偏置项。

### 2.3 环境平面栅格地图构建

依据 2.2 节所示流程对巡检机器人作业环境图像区域能否通行进行判定, 将可通行区域记为  $\Omega_i^Y(x, y)$ , 不可通行区域记为  $\Omega_i^N(x, y)$ , 基于 BiGEMAP 生成巡检机器人作业环境的平面栅格地图, 为巡检机器人自主导航与避障提供重要支持。

一般情况下, 改进支持向量机输出数据  $\lambda(A_{\Omega})$  存在着较大的不确定性<sup>[19]</sup>。因此, 设计技术通过概率模型来实时更新栅格 (区域) 状态。将栅格  $\Omega_i(x, y)$  不可通行概率设置为  $p(o)$ , 可通行概率设置为  $p(f)$ , 更新公式为:

$$\begin{cases} p(o|Z) = \frac{p(Z|o) \cdot p(o)}{p(Z)} \\ p(f|Z) = 1 - p(o|Z) \end{cases} \quad (21)$$

式中,  $p(o|Z)$  与  $p(f|Z)$  代表栅格不可通行概率与可通行概率的更新结果;  $Z$  代表改进支持向量机的观测值;  $p(Z|o)$  代表在栅格不可通行时返回观测值  $Z$  的概率;  $p(Z)$  代表观测值  $Z$  出现的概率。对公式 (19) 进行贝叶斯全概率展开:

$$p(o|Z) = \frac{p(Z|o) \cdot p(o)}{p(Z|o) \cdot p(o) + p(Z|f) \cdot p(f)} \quad (22)$$

式中,  $p(\bar{o})$  表示不可通信的均值概率, 当栅格不可通行概率  $p(o|Z)$  大于可通行概率  $p(f|Z)$  时, 更新栅格状态为不可通行区域; 当栅格不可通行概率  $p(o|Z)$  小于可通行概率  $p(f|Z)$  时, 更新栅格状态为可通行区域。当移动障碍物的位置发生变化时, 贝叶斯不可通信概率  $p$  的值同步调整, 并依据当前时刻栅格状态更新、标记的结果, 建立环境平面栅格地图, 如图 3 所示。

综上所述在 CNN-RNN 复合模型、改进支持向量机、BiGEMAP 软件的应用背景下, 实现了巡检机器人的位姿校正与定位建图, 为巡检机器人的后续应用提供一定的帮助<sup>[20]</sup>。

### 2.4 回环检测与机器人行进误差控制

对机器人的运动路径进行回环检测, 能够提供较长时间周期内的位姿约束并在同步定位与地图构建中消除

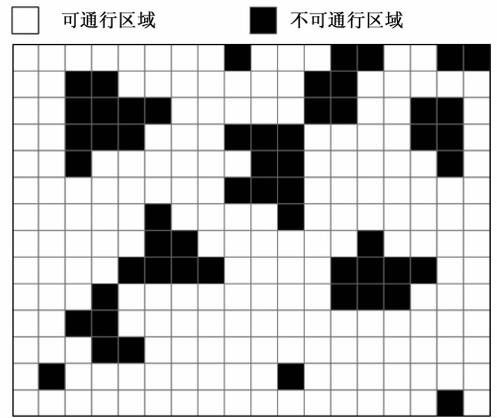


图 3 平面栅格地图示例图

随运动路径增加而产生的累计误差, 回环检测的具体步骤如下:

- 1) 设定机器人移动的起止时刻、结束时刻和时间间隔。
- 2) 通过线性插值的方式, 实现对机器人移动轨迹的动态拟合和动态补偿。
- 3) 将机器人不同时点的移动点动态拟合成光滑的轨迹曲线。
- 4) 基于拟合后的最优曲线修正机器人真实的移动轨迹。

将机器人开始移动的时刻设定为  $t_0$ , 结束的时刻设定为  $t_M$ , 获取到从  $t_0$  到  $t_M$  的时间间隔, 并获取到全部图像统一位置的点坐标。为避免不同时刻时间戳的不同, 基于线性插值的方式对不同时刻的时间间隔实施动态补偿, 插值的过程描述如下:

$$t_{q+1} = \frac{t_q - t_0}{t_{q+1} - t_q} \times t_q \quad (23)$$

式中,  $t_{q+1}$  为插入  $t_q$  的下一个时刻, 通过连续插入采样点能够将机器人的运动轨迹连成曲线, 对  $t_{q+1}$  时刻的采样点  $P_{t+1}$  求一阶导得到  $\bar{P}_{t+1}$ , 在  $\bar{P}_{t+1}$  不为零的前提下对采样点  $P_{t+1}$  的位置坐标进行修正:

$$P'_{t+1} = t_{q+1} \times \bar{P}_{t+1} \quad (24)$$

式中,  $P'_{t+1}$  表示修正后的图像采样点, 当  $P'_{t+1}$  与  $P_{t+1}$  三个轴向的坐标值差距过大时, 可以通过增加插值密度缩小采样点之间的偏差。基于回环检测能够实时判定误差情况, 并及时修正机器人的运行轨迹误差。

## 3 实例测试与结果分析

### 3.1 测试环境搭建

选取某型号巡检机器人作为测试对象, 基于摄像头采集巡检机器人图像数据, 并同步利用激光器标定机器人的运动偏差。采用全景相机获取其作业环境图像, 通过 BiGEMAP 软件生成测试环境的平面栅格地图, 通过

引入动态概率权重降低图像拼接重叠率，具体如图 4 所示（测试场地的长宽为 100 m \* 100 m）。

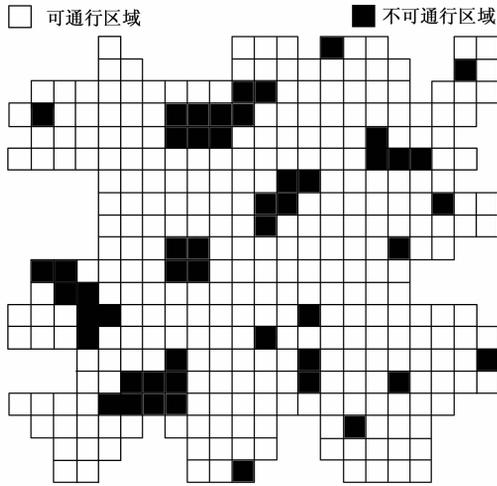


图 4 测试环境的平面栅格地图示意图

如图 4 所示，巡检机器人作业环境呈现出高度的复杂性，不仅体现在作业环境空间的不规则上，还体现在障碍物（不可通行区域）的不规则分布上，能够最大限度地提升设计技术应用性能测试的可信度与精准度。

### 3.2 消融实验

本文设计技术在巡检机器人位姿校正中引入并应用了 CNN-RNN 复合模型，通过消融实验验证其应用性能，各模型的训练数据的划分、训练次数等调整设置相同。实验模型如下所示。

- 1) 实验模型 1：即 CNN-RNN 复合模型（设计技术应用的），结构为输入部分（1 个）、CNN 部分（1 个）、RNN 部分（2 个）与输出部分（1 个）；
- 2) 实验模型 2：即 CNN 结构模型，结构为输入部分（1 个）、CNN 部分（1 个）与输出部分（1 个）；
- 3) 实验模型 3：即 RNN 复合模型，结构为输入部分（1 个）、RNN 部分（1 个）与输出部分（1 个）；
- 4) 实验模型 4：即简化 CNN-RNN 复合模型，结构为输入部分（1 个）、CNN 部分（1 个）、RNN 部分（1 个）与输出部分（1 个）。

应用上述 4 个实验模型对巡检机器人位姿进行校正处理，通过位姿校正准确率显示实验模型的应用性能，如图 5 所示。

实验模型 1 经过 26 次迭代达到巡检机器人位姿校正准确率最大值 99.5%；实验模型 2 经过 38 次迭代达到巡检机器人位姿校正准确率最大值 97%；实验模型 3 经过 37 次迭代达到巡检机器人位姿校正准确率最大值 98%；实验模型 4 经过 34 次迭代达到巡检机器人位姿校正准确率最大值 96%。各不同模型在各自完成迭代次数后准确率的增长情况，如表 1 所示。

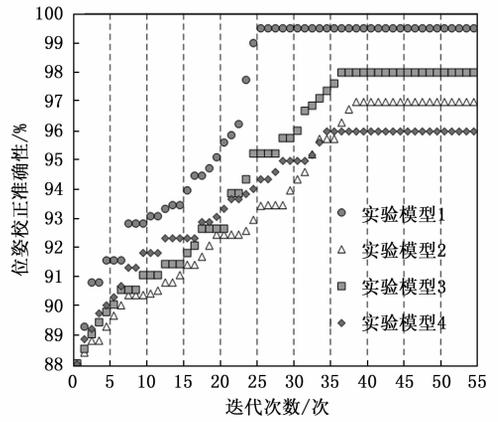


图 5 位姿校正准确率示意图

表 1 位姿校正准确率的增长情况

模型	1	2	3	4
迭代次数/次	26	39	37	34
准确率增加 %	10.2	8.3	9.3	7.5

通过数据比对发现：实验模型 1（即设计技术应用的 CNN-RNN 复合模型）巡检机器人位姿校正准确率最高，并且模型收敛速度最快，说明设计技术改进与应用了的 CNN-RNN 复合模型具备更好的巡检机器人位姿校正性能。这主要是因为 CNN-RNN 复合模型充分利用了 CNN 与 RNN 的优势，并通过融合方式避免了两者的缺陷，从而大幅提升了巡检机器人位姿校正的准确率，准确率提升的净值最大。

### 3.3 巡检机器人环境图像拼接性能分析

应用设计技术、对比技术 1（基于级联优化和强度特征的机器人定位建图技术<sup>[21]</sup>）与对比技术 2（基于 Transformer 的机器人定位建图技术<sup>[22]</sup>）对巡检机器人环境图像进行拼接处理，以拼接环境图像重叠区域比例来显示应用技术的应用性能。对比技术 1 的子系统数量为 5，系统约束值的取值范围为 [0.8 - 1]，对比技术 2 神经网络参数值设定与本文算法一致。

常规情况下，拼接环境图像重叠区域比例越大，表明巡检机器人环境图像拼接性能越差；反之，拼接环境图像重叠区域比例越小，表明巡检机器人环境图像拼接性能越好。设计技术、对比技术 1 与对比技术 2 应用后拼接环境图像重叠区域比例如表 2 所示。

相较于对比技术 1 与对比技术 2 来看，设计技术应用后拼接环境图像重叠区域比例更小，其最小值达到了 1.02%，说明设计技术巡检机器人环境图像拼接性能更优，能够为后续巡检机器人作业环境平面栅格地图建立提供更加精准的环境图像支撑。引入 SSIM（Structural Similarity）结构相似性指标，评价拼接前后图像的相似度：

表 2 拼接环境图像重叠区域比例数据表 %

测试次数	设计技术	对比技术 1	对比技术 2
1	2.31	10.23	8.45
2	1.02	8.56	9.26
3	1.89	9.45	9.77
4	3.45	12.03	9.91
5	2.01	11.02	12.03
6	2.00	8.45	10.02
7	2.89	7.15	13.56
8	2.78	6.02	15.78
9	2.45	5.46	19.56
10	1.23	9.21	9.15
11	1.56	13.02	9.36
12	1.87	15.46	7.45
13	1.44	9.23	6.26
14	1.62	5.80	15.24
15	2.13	10.28	11.14

$$SSIM(z, z') = \frac{(2\mu_z \mu_{z'} + c_1 \xi^2)(2\sigma_{z, z'} + c_2 \xi^2)}{(\mu_z^2 + \mu_{z'}^2 + c_1 \xi^2)(\sigma_z^2 + \sigma_{z'}^2 + c_2 \xi^2)} \quad (21)$$

式中,  $z$  和  $z'$  表示拼接前后的图像,  $\mu_z$  和  $\mu_{z'}$ ,  $\sigma_z$  和  $\sigma_{z'}$  分别表示拼接前后图像的均值和方差,  $c_1$  和  $c_2$  分别为图像动态调节的参数, 取值范围在 0.01~0.10 之间,  $\xi$  表示图像的动态范围。SSIM 的值越趋近于 1, 表明图像的拼接效果越好, 各算法图像拼接的 SSIM 的值表现如表 3 所示。

表 3 图像拼接 SSIM 值统计

算法	设计技术	对比技术 1	对比技术 2
SSIM 值	0.96	0.92	0.91

### 3.4 巡检机器人作业区域能否通行判定性能分析

应用设计技术、对比技术 1、对比技术 2 和对比技术 3 (ORB-SLAM3, 引入该技术进一步验证低照度和高动态场景下各算法的性能) 对巡检机器人作业区域能否通行进行判定, 通过区域通行误判率来显示应用技术的性能优劣, 如图 6 所示。

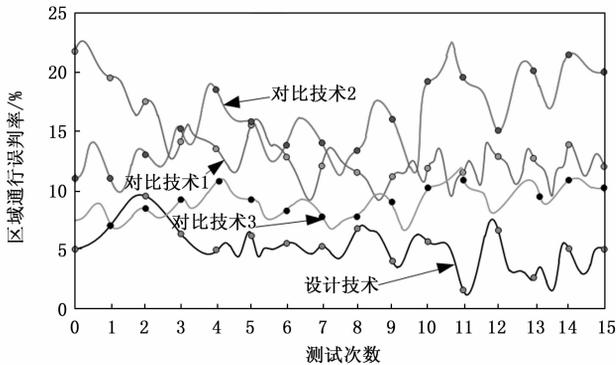


图 6 区域通行误判率示意图

设计技术应用后区域通行误判率远低于对比技术 1 与对比技术 2, 其最小值达到了 1%。这主要是因为设计技术应用了改进支持向量机, 通过引入核函数技巧, 有效提升了模型在处理非线性、高维数据时的分类精度与泛化能力。对比技术 3 在误判率方面相对于对比技术 2 和对比技术 1 误判率得到一定改善, 但准确率不及本文技术, 验证了本文设计算法在高动态、低照度环境下的有效性。再对比各算法位姿校正实时性, 统计结果如表 4 所示 (时间越短表面算法的校准能力越强):

表 4 机器人作业中各算法位姿校正的实时性 ms/帧

测试次数	设计技术	对比技术 1	对比技术 2	对比技术 3
1	15	55	55	25
2	12	61	52	23
3	14	62	54	22
4	13	63	53	21
5	12	52	52	26
6	13	51	65	25
7	11	54	64	31
8	12	53	63	32
9	12	49	61	35
10	11	48	50	25
11	10	42	52	36
12	12	53	55	25
13	13	52	57	21
14	15	55	58	25
15	14	54	52	24

设计技术改进的支持向量机还能够持续从新增数据中学习并自我优化, 适应不断变化的巡检机器人作业环境。此种动态调整能力, 确保了改进支持向量机模型在面对新情况时, 依然能够保持低区域通行误判率, 实现高效准确的区域能否通行判断。由此可见, 设计技术得益于改进支持向量机的深度应用, 从而大幅度地降低了区域通行误判率及同步提升了位姿校正的实时性。

### 3.5 巡检机器人作业环境地图建立性能分析

应用设计技术、对比技术 1 与对比技术 2 构建巡检机器人作业环境地图, 如图 7 所示。

应用设计技术建立的巡检机器人作业环境地图与实际机器人作业环境平面栅格地图 (图 3) 相同, 而对比技术 1 与对比技术 2 建立的巡检机器人作业环境地图与实际机器人作业环境平面栅格地图 (图 3) 存在着较大的差异性, 各算法的环境平面栅格地图 SSIM 值统计如表 5 所示。

表 5 作业环境地图 SSIM 值统计

算法	设计技术	对比技术 1	对比技术 2
SSIM 值	1	0.84	0.72

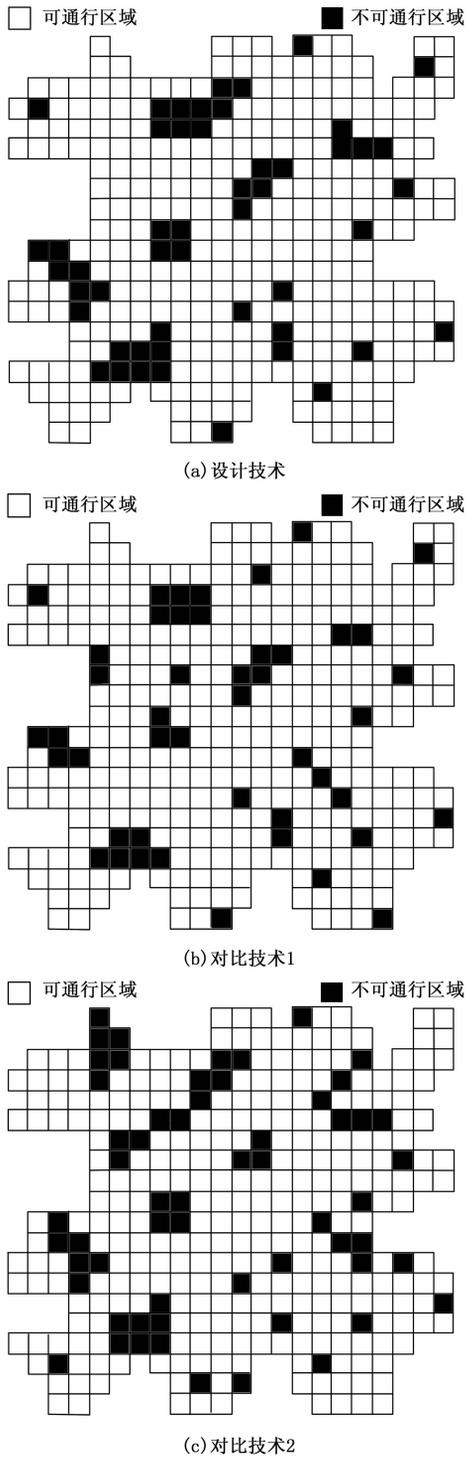


图 7 巡检机器人作业环境地图建立结果示意图

通过上述平面栅格地图对比可知：设计技术巡检机器人作业环境地图建立性能更佳，能够为巡检机器人避障路径规划提供更加精准的地图数据支撑。最后为验证在海量数据集环境下机器人的误判率和位姿校正的实时性，将测试次数提升了 500 次，结果如表 6 所示（本文设计算法的性能未随着数据集规模的增加而受到影响，进而验证了本文算法的稳定性和适用性）。

表 6 大规模数据环境下各算法性能对比

指标	本文设计技术	对比技术 1	对比技术 2	对比技术 3
误判率均值%	3.5	7.6	11.3	12.5
位姿实时性均值/ms	12.9	25.6	71.5	68.3

#### 4 结束语

位姿校正对于巡检机器人而言，是确保其能够准确执行任务的基础。在复杂多变的环境中，机器人需要不断调整自身的位置与姿态（即位姿），以精准到达目标位置并高效完成作业。定位建图是巡检机器人实现高效巡检的核心环节。构建精确的环境地图，能够为巡检机器人提供准确的导航依据，帮助其规划最优路径，避免障碍物，高效完成巡检任务。但是，已有技术由于应用算法、模型等的缺陷或者问题，无法实现精准的位姿校正与定位建图，制约着巡检机器人的发展步伐，故提出面向巡检机器人的位姿校正与定位建图技术研究。测试结果显示：设计技术大幅度地提升了巡检机器人环境图像拼接性能、巡检机器人作业区域能否通行判定性能与巡检机器人作业环境地图建立性能，其中，（1）位姿校正精度提升至 99.5%；（2）地图误判率降低至 1%；（3）图像拼接效率提升 40%，为巡检机器人的后续应用与发展提供有力的支撑与帮助。

#### 参考文献：

- [1] 徐 斌, 杨东勇. 基于边缘计算的移动机器人视觉 SLAM 方法 [J]. 高技术通讯, 2023, 33 (9): 1000 - 1008.
- [2] 王茂森, 鲍久圣, 鲍周洋, 等. 融合金字塔结构与注意力机制的煤矿井下巡检机器人 PT 目标检测算法 [J]. 煤炭科学技术, 2024, 52 (6): 206 - 215.
- [3] 杨韧韧, 吴炳龙, 彭晋民, 等. 面向机器人抓取位姿估计网络研究 [J]. 机械科学与技术, 2024, 43 (11): 1929 - 1936.
- [4] 邓 鹏, 罗 静. 复杂环境下机器人多传感器融合定位方法 [J]. 电子测量与仪器学报, 2023, 37 (12): 48 - 57.
- [5] 崔玉明, 刘送永, 吕振礼, 等. 基于级联优化和强度特征的地下退化环境机器人自主精准定位 [J]. 仪器仪表学报, 2023, 44 (12): 208 - 216.
- [6] ZHANG S, DUAN X, ZHU G, et al. Empowering robots with social cues: an initiative pose control framework for human - robot interaction [J]. Intelligent Service Robotics, 2024, 17 (5): 1005 - 1017.
- [7] XING L. Research on the positioning and recognition of an intelligent inspection robot in substations [J]. International journal of automation technology, 2023, 17 (1): 65 - 70.

（下转第 320 页）