

基于场景理解的AGV碰撞风险缓解方法

周润民, 杨欢

(华南师范大学人工智能学院, 广东 佛山 528225)

摘要: 针对制造业车间中自动导引车(AGV)运行时可能产生的碰撞风险问题, 提出一种基于场景理解技术的风险缓解方法。该方法利用计算机视觉技术生成场景图, 以增强AGV对周围环境的感知和理解, 并特别关注物体间的空间关系, 使AGV基于场景信息作出决策, 减少碰撞风险。实验在3种典型场景(巡线、离线、回线)下进行, 通过收集的图像数据评估AGV的环境感知能力和风险缓解能力。结果显示, 所提出的结合场景图的风险缓解框架可有效识别风险并实施缓解措施, 尤其在处理多类别单目标的静态场景时表现突出。未来研究将致力于扩展至更复杂的多类别多目标场景, 以进一步提升AGV在动态工业环境下的安全性和效率。

关键词: AGV; 场景理解; 风险缓解; 计算机视觉; 场景图生成

DOI: 10.11907/rjdk.241932

中图分类号: TP391.41

文献标识码: A

文章编号: 1672-7800(2026)002-0202-07

扫描二维码阅读全文:



AGV Collision Risk Mitigation Method Based on Scene Understanding

ZHOU Runmin, YANG Huan

(School of Artificial Intelligence, South China Normal University, Foshan 528225, China)

Abstract: To mitigate collision risks for tape-guided Automatic Guided Vehicle (AGV) in manufacturing plants, a novel risk alleviation method leveraging scene understanding is introduced. Employing computer vision, the technique generates scene graphs to enhance AGV environmental awareness, focusing on spatial object relations. Decision-making based on scene context significantly reduces collision hazards. Through experiments across patrol, off-line, and return-to-line scenarios, the AGV's perception abilities and risk management efficacy are evaluated using collected imagery. Results highlight the method's effectiveness, especially in static scenes with multiple single-object categories. Future efforts target expanding the framework to complex multi-object scenarios for improved AGV safety and efficiency in dynamic industrial environments.

Key Words: AGV; scene understanding; risk mitigation; computer vision; scene graph generation

0 引言

磁带引导的自动导引车(Automatic Guide Vehicle, AGV)在制造业车间承担着物料运输的关键任务, 频繁地在引导磁道与生产线工作台之间穿梭。在车间类场景中, 一个显著特征便是环境模型的频繁变动, 如货物的装卸与人员的流动。静态物体位置的变化对环境建模的影响以及动态障碍物的干扰, 是移动机器人在此类环境中实现长期稳定自主运行面临的核心问题^[1]。为了确保AGV在工业环境中稳定自主运行, AGV需要能够实时感知到周围环境, 理解移动障碍物的意图。传统上, AGV依赖超声波^[2]

与红外测距传感器^[3]进行障碍物检测, 但这些方案存在固有限制。距离传感器作用阈值的设定成为难题: 阈值过高, AGV无法精准接近工作台的物流放置区; 阈值过低, 则可能发生碰撞, 对协作人员造成伤害。基于此, 本文提出一种基于场景理解技术的风险缓解方法。该方法利用计算机视觉技术中的场景图建模周围环境, 评估出周围环境的风险系数后发出具体指令, 以减少突发情况下AGV所带来的危害, 保障了与AGV协同作业的工作人员或其他AGV的安全。

本文的主要贡献如下:

(1) 提出一种新的场景图范式, 以增强AGV的场景理解能力, 特别是理解AGV与所处空间中的物体关系, 便于AGV根据场景信息作出决策。

收稿日期: 2024-12-13

作者简介: 周润民(1999-), 男, 华南师范大学人工智能学院硕士研究生, 研究方向为计算机视觉与机器人; 杨欢(1986-), 男, 博士, 华南师范大学人工智能学院副研究员、硕士生导师, 研究方向为信息物理系统。本文通讯作者: 杨欢。

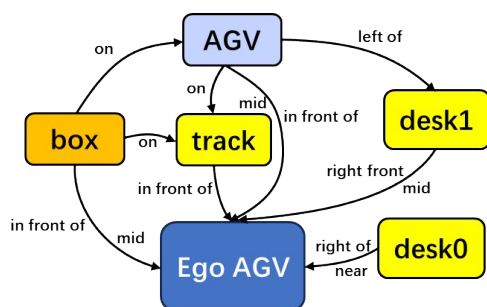
(2) 发布了一个新的场景数据集, 该数据集专注于提高场景图生成技术捕捉空间关系的能力。

Ego-AGV 场景图示例如图1所示。



(a) Scene from the perspective of an AGV

(a) AGV 视角下的场景



(b) Example of generated Ego-AGV scene graph

(b) 生成的Ego-AGV场景图示例

Fig. 1 Example of an Ego-AGV scene graph

图1 Ego-AGV 场景图示例

1 相关工作

1.1 场景图生成

场景图生成(Scene Graph Generation, SGG)是一种用于提取图像中丰富语义信息的计算机视觉技术,其不满足于仅提取出图像中的目标,还需要提取出图像中目标之间的相互关系。其以节点代表目标、以边代表目标之间的关系,通常以三元组<subject, predicate, object>的形式表达。Johnson等^[4]首次提出场景图生成用于图像检索领域。VG数据集是场景图生成领域的第一个大规模语义数据集,极大地促进了该领域的发展^[5]。Xu等^[6]提出迭代消息传递机制(Iterative Message Passing),不仅验证了上下文信息对于提升关系预测准确性的作用,而且系统定义了领域内各项任务的标准评估指标。Zellers等^[7]提出基于双向长短期记忆的方法,并指出了VG数据集中存在的长尾分布问题。Tang等^[8]提出总直接效应(Total Direct Effect),该方法基于生成谓词构建无偏场景图,以降低长尾分布问题的干扰。Sun等^[9]提出一种两阶段因果建模方法(Two-stage Causal Modeling, TsCM)来解决语义混淆和长尾分布问题,有效提

高了模型对细粒度尾部关系的预测性能。

1.2 利用场景图的机器人

随着场景图生成技术的发展,越来越多学者将场景图生成技术应用于机器人领域,以提升机器人感知环境的能力。Riaz等^[10]将场景图应用于人机协同作业(Human Robot Collaboration, HRC)进行安全分析。Kenack等^[11]将场景图用于视觉问答以辅助机器人操作。Amiri等^[12]提出一种面向机器人规划的视觉场景分析算法,提供环境上下文信息辅助机器人规划。Li等^[13]探索了在拆卸任务场景中如何通过场景图来辅助人机协作。周方波等^[14]通过构建场景图谱以提高移动机器人的搜索效率和准确性,机器人利用语义关系图中物体间的并发关系,优先在关系强度较高的地标物体附近进行搜索。

之前的研究虽然已经着眼于将场景图生成技术应用于下游领域,但未能充分挖掘场景图的内在潜力,往往无法彰显其独特优势。其中一个原因便是现有数据集有大量关系是语义关系,许多研究者为了正确捕获这种关系进行了大量研究,反而忽视了这类信息是否真的有助于下游任务来理解场景。相较于语义信息,空间信息具备更低的歧义性,更易于精确获取,为移动机器人的场景理解提供了更坚实的基础。为此,本文提出一种新的交互场景图构建方法和一个仅包含空间信息的数据集,希望能够更加充分地发挥场景图的效能。

2 本文方法

图2展示了本文方法的总体架构,旨在通过场景理解来降低AGV可能造成的风险。该架构由两部分组成:第一部分通过视觉系统获取场景图,第二部分根据感知到的场景图实现风险缓解。

2.1 空间关系场景图数据集

由于现有数据集的设计初衷并非为了满足机器人在动态和复杂环境中的感知需求,无法充分支持移动机器人在实际应用中对环境的理解。因此,本文构建了一个只有空间关系的数据集,后续研究也基于该数据集进行。该空间数据集由搭载在移动机器人上的Realsense D435摄像头采集,包含720张RGB图像。每张图像平均包含3个目标以及目标之间的空间关系。本文利用LabelImg-Rel标注工具标注该数据集^[15]。预定义的对象类Co如表1所示,依据移动性质分为了3类,分别是具有自主移动能力的人和AGV、可被人或AGV搬运的其他物品以及移动能力极差的对象。这类对象一般不会移动,可视为背景。

关系是互相且成对出现的,工人搬运纸箱这一关系也可描述为纸箱在工人的手上。因此,在三元组关系中既可标注为<person, carry, box>,也可标注为<box, on, person>。

为了减少主客体不一导致的同一关系标注重复,本文

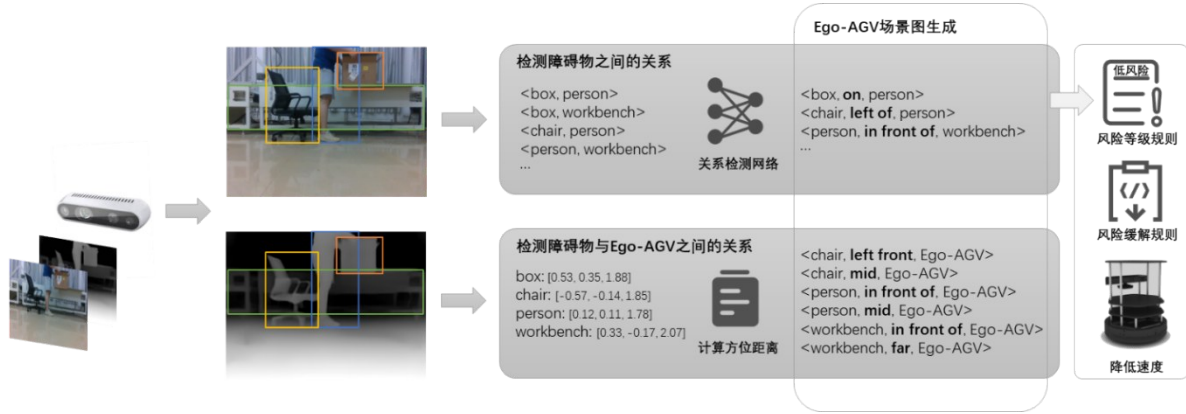


Fig. 2 Ego-AGV scene graph generation and risk mitigation process

图2 Ego-AGV 场景图生成及风险缓解流程

Table 1 Predefined object class classification

表1 预定义的对象分类

实体	类别
动态障碍物	工人、AGV
静态障碍物	扳手、纸箱、椅子、屏幕
背景障碍物	轨道、桌子、工作台、架子

根据物体的移动性质进行关系标注。在三元组关系中,静态障碍物优先于移动障碍物和背景障碍物作为主体,而移动障碍物优先于背景障碍物作为主体。当同一类别的实体之间需要标注关系时,按照类别出现顺序标注,出现顺

序靠前的优先作为主体。如图3(a)所示,箭头从主体出发,指向客体,明确了障碍物间标注的顺序。如此标注的好处是,一些主动的语义关系如搬运等无需标注,仅用空间关系即表达出了3类障碍物之间的相对关系。后续将对象组合送入模型进行关系检测时,也依照此优先级来判断物体之间的关系,从而在一定程度上降低了关系检测的复杂度。数据集中预定义的空间关系有 {in, on, in front of, behind, left of, right of, above, under} 共8种类别。图3(b)展示了一张图像对应的场景图标注示例。

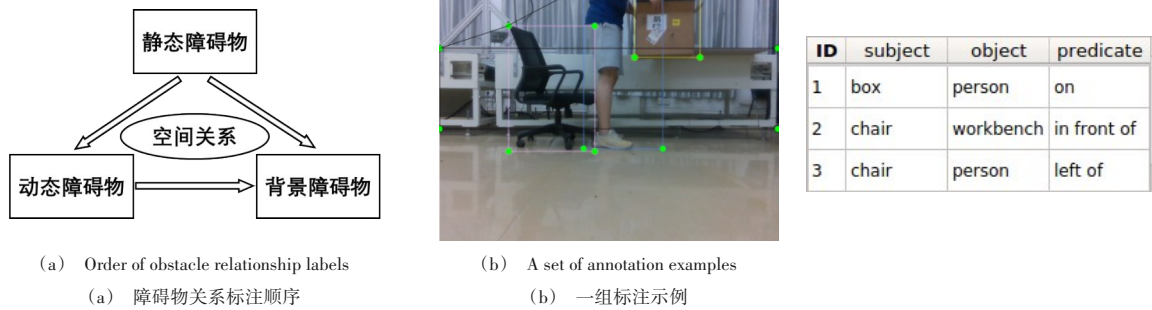


Fig. 3 The order of relation labeling in datasets and examples

图3 数据集中的关系标注顺序及示例

2.2 与空间障碍物构建关系的场景图

本文提出一种新的机器人交互场景图,命名为Ego-AGV场景图。其可以与空间中存在的障碍物构建空间关系,从而能更全面地建模感知到的障碍物信息,使AGV可以在这种复杂的环境中根据环境信息作出更合理的决策。该场景图生成方法分为4个步骤:

(1)目标检测。目标检测对于场景图生成是至关重要的,不仅决定了场景图生成的质量,而且决定了场景图生成的速度。鉴于工业应用对实时性和端侧稳定性的严苛需求,本文采用目前已在工业界广泛应用的YOLOv5。首先将通过摄像头获取的RGB图片输入到目标检测模块中,产生一系列预测对象建议 op_i :

$$op_i := \{b_i, s_{o_i}\} \quad (1)$$

每一个对象建议 op_i 包含一个目标检测边界框 b_i 和一

个对应的目标检测得分 $s_{o_i} \in (0, 1)$, 本文将目标检测到 s_{o_i} 中的最高得分视作障碍物对应的类别标签 o_i :

$$o_i := \arg \max_{o_i \in \mathcal{C}_o} s_{o_i} \quad (2)$$

式中, $\mathcal{C}_o$ 为预定义的实体对象类。场景中的实体主要有工人、AGV、架子等,详见表1。

(2)障碍物与AGV相对位置关系建立。以往的研究侧重于提取障碍物之间的关系信息,较少关注障碍物与机器人之间的关系。本文将AGV与障碍物的关系融入场景图生成领域的空间关系分析中。得益于这种关系的建立,AGV在运行过程中可以得知场景中障碍物与自身的相对位置关系,从而更好地指导AGV执行相应指令,以减少碰撞风险发生概率。本文利用RGB-D相机获取障碍物的深度信息,借此得出障碍物与移动机器人的位置关系。具体而言,将RealSense相机提供的RGB图像与深度图像对齐,

结合目标检测信息综合处理后得到障碍物的关键属性——类别、二维边界框及深度信息。通过计算障碍物中心点邻域内的平均深度,确定 AGV 至障碍物的距离。障碍物中心点像素坐标至现实世界坐标的转换遵循下述公式:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \begin{bmatrix} \frac{1}{f_x} & 0 & -\frac{c_x}{f_x} \\ 0 & \frac{1}{f_y} & -\frac{c_y}{f_y} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} Z \quad (3)$$

式中, u 和 v 为 RGB 图像的像素坐标, Z 为 RGB 图像中像素对应的深度值, f 为相机的焦距, X_c 、 Y_c 、 Z_c 为目标点在特定世界坐标系下的空间坐标。该坐标系以 RealSense 相机为原点,以相机成像平面为 XOY 平面。 f_x 和 f_y 为相机的焦距, c_x 和 c_y 为摄像头光学中心在像素坐标系中的位置。利用获取的障碍物中心点像素坐标得到 AGV 与障碍物之间的位置关系,通过以下公式计算得出障碍物与机器人的角度关系:

$$\theta = \begin{cases} \arctan\left(\frac{z}{x}\right), & x > 0 \\ \arctan\left(\frac{z}{x}\right) + \pi, & x < 0 \end{cases} \quad (4)$$

基于此角度公式,障碍物方位被划分为右方、右前方、前方、左前方和左方 5 个区域,同时根据距离定义了 3 个层级:近距离(小于 1.5 m)、中距离(1.5 ~ 3 m)和远距离(大于 3 m),如图 4 所示。此外,当障碍物位于一定高度上时,系统判定其不会对 AGV 的运行构成影响,因此不需计算该障碍物与 AGV 的空间关系。

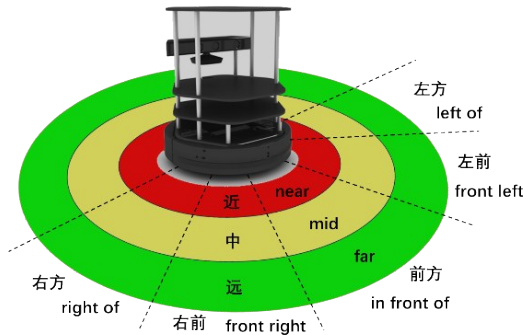


Fig. 4 The classification of the position and distance of obstacles relative to the AGV

图 4 障碍物相对于 AGV 的方位与距离划分

(3)障碍物关系检测。为了实现实时场景图生成,系统采用上下文感知视觉场景图生成(Context-Aware Visual Scene Graph Generation, CV-SGG)网络进行障碍物关系检测。该网络内置上下文场景图生成模块(Context Scene Graph Generation, C-SGG),以利用物体的位置和类别信息生成多样化的上下文描述,不依赖视觉特征而是通过上下文增强技术减少长尾偏差。之后,引入上下文引导的视觉场景图生成模块(CV-SGG),在 C-SGG 基础上进一步利用视觉信息提升对真实谓词的识别效果,专注于可能的谓

词,并忽略不可能的谓词,最终结合上下文和视觉信息生成场景图。关于关系推断过程的更多细节可以参考文献[15]。

(4)Ego-AGV 场景图生成。通过整合上述步骤获得的障碍物与 AGV 相对位置信息及障碍物间的关系,将两类关系合并构建出 Ego-AGV 场景图。借助于 Ego-AGV 场景图,AGV 具备了感知环境的能力。自此,搭载本系统的 AGV 不仅能感知场景中的所有障碍物对象,而且能明确其之间的空间关系构成。

2.3 风险缓解系统

2.3.1 风险分类

本文将 AGV 运行过程中的碰撞风险划分为 3 个等级:高风险、低风险以及安全状态。风险划分基于 Ego-AGV 场景图提供的环境信息,包括障碍物的类型、距离和方位。通过深度整合并解析这些关键参数,系统采用一套基于经验法则与逻辑推理的 if-then 规则集,以实现当前环境风险水平的即时评估,并执行相应的风险缓解策略。限于篇幅,表 2 仅展示了部分代表性场景中只存在单一障碍物时的风险等级评估规则。

Table 2 Risk level assessment rules for single obstacle

表 2 单一障碍物风险等级评估规则

障碍物类型	距离	方位	风险等级
静态	远	前方	安全
静态	近	前方	低风险
动态	近	左前	低风险
动态	近	前方	高风险
背景	中	左方	安全
背景	近	前方	低风险

在 AGV 运行过程中,遇到诸如协同作业的工人或其他 AGV 动态障碍物时,其行为意图的不确定性是一大挑战。此时若简单地将单一障碍物的风险等级评估规则组合起来,难以全面地评估场景风险等级,从而导致碰撞发生概率大幅上升。因此,本文提出一套基于场景图信息的风险评估方法,用于处理多个障碍物同时存在的场景。该方法不仅考虑障碍物的存在,而且能分析其与环境的关系,从而更准确地判断风险等级。表 3 展示了部分障碍物组合场景风险等级分类。

在序号 1 所示场景中,搭载一般场景理解策略的 AGV 可通过目标检测技术识别到工人的存在。尽管工人并不会直接阻挡 AGV 的行进,但出于防止工人突然移动可能导致的碰撞风险考虑,AGV 通常会选择减速通过。然而,借助场景图生成技术,AGV 能够进一步分析工人的状态,判断其此时处于相对静止状态,将此场景归类为安全场景,并保持正常速度运行,从而在保障安全的同时提高运输效率。另一种场景如序号 4 所示,工人将工作台前的零件(本文以箱子作为简化)搬运到工作台上进行组装时,由于零件位于 AGV 的目标放置点,AGV 易与工人发生碰撞,因此将此类场景归类为高风险场景。

Table 3 Risk level evaluation rules for scenes composed of multiple obstacles

表3 多障碍物组成场景的风险等级评估规则

序号	三元组场景	说明	风险等级
1	<chair, below, person><person, in front of, workbench>	人坐在工作台前的椅子上	安全
2	<chair, below, person><person, behind, workbench>	人坐在工作台后的椅子上	安全
3	<chair, behind, person><person, left of, desk>	人起身离开工位	低风险
4	<box, on, person><person, in front of, workbench>	人在工作台前搬运零件	高风险
5	<wrench, on, person><person, on, track>	人手拿着扳手在轨道上	高风险

2.3.2 风险缓解

风险缓解阶段基于风险评估的风险等级来降低风险,以最大限度地减少已识别的危险,并根据所提取 Ego-AGV 场景图获得的信息来发出相应指令。其中,输入是障碍物到机器人的距离、方向和多障碍物组合的风险水平,输出是整个场景的综合风险等级。根据风险等级,系统下达具体指令,以减少碰撞危害并提高效率。

在多障碍场景中,首先通过表3得到当前场景的多个障碍物组合风险等级,再根据表2得到单一障碍物的风险等级,两者结合得到该场景的风险等级。当场景为安全时,此时 AGV 恢复速度;当场景为低风险时,此时 AGV 降低速度;当场景为高风险时,此时 AGV 停止运行以避免碰撞发生。

3 实验与结果分析

3.1 实验设置

本文搭建了一个模拟工厂流水线 AGV 运行情景的实验室环境,该环境搭建了 AGV 的运行轨道、在流水线上的工作台、装载货物的架子以及代表搬运零件的纸箱,以此来实际反映 AGV 在作业时所处的环境特点。

实验采用的机器人平台为 Turtlebot2,该机器人配置了 Kobuki 自主移动底座,集成 RealSense D435 深度感知摄像头以实现三维空间感知,以及 Jetson TX2 计算模块以确保深度神经网络的流畅运行。系统内部通信则依托于成熟的机器人操作系统(Robot Operating System, ROS),以实现各硬件组件及软件模块间的高效协同与数据交换。

本文模拟 AGV 的运行过程,以不同角度拍摄原始图像数据,并标注了图像数据对应的标签,构建了用于实验的目标检测数据集。其中包含的目标如表1所示。数据集在上述实验室环境下收集,由 720 张 640*480 像素的 RGB 图像组成,并按照 80:20 的比例划分为训练集和测试集。空间关系场景图生成算法以本文之前构建的空间关系场景图数据集为基础进行训练。

实验指标根据视觉部分和策略执行部分分为两类:第一类指标为目标检测算法的平均准确率(Average Accuracy)以及场景图生成准确率(Scene Graph Generation+ Precision),这两项指标可以反映 AGV 是否可以准确地感知周围环境。平均准确率为目标检测算法在设定的实验室场景中检测出某一类别物体的检测准确率 AA_i 之和与该

类别物体被检测次数 M 的比值^[16]。具体公式如下:

$$AA = \frac{\sum_{i=1}^M AA_i}{M} \quad (5)$$

利用场景图生成准确率这一指标评估场景图构建算法的准确性,计算目标、关系以及三元组生成的准确性^[17]。具体公式如下:

$$Recall = \frac{C(O) + C(P) + C(T)}{N} \quad (6)$$

式中, $C(O)$ 表示正确识别和定位的对象数量; $C(P)$ 表示在主语和宾语正确定位的前提下正确识别的谓词数量; $C(T)$ 表示正确识别的三元组数量; N 表示真实场景图中的总条目数,即对象、谓词和关系的真值数量总和。该指标能够准确反映场景图生成系统的整体性能,即使在部分元素出现错误时也能给出合理的评估结果。为确保协作人员的安全,必须对机器人系统的风险识别和风险缓解能力进行系统评估^[19]。

第二类指标为风险识别结果以及风险缓解策略执行结果,这两项指标反映了 AGV 在动态工作环境中是否能够准确识别潜在风险,并有效执行相应的缓解策略。实验设置了 AGV 运行中需要应对的 3 类典型场景,分别为常规运行情况下的巡线场景、前往工作台运送货物的离线场景,以及完成运输后的回线场景。模拟的场景涵盖了由静态障碍物和背景障碍物构成的静态场景,也包含了引入动态障碍物后的动态场景。图5展示了 AGV 在实际测试过程中 3 类典型场景及其对应生成的 Ego-AGV 场景图。前两列展示的是静态场景及其生成的场景图,后两列则展示了动态场景及其对应的场景图。为了更清晰地区分静态障碍物、动态障碍物以及背景障碍物,本文使用了不同的颜色和线条标注图像中的目标检测结果,并在生成的场景图中对不同类型的障碍物使用不同颜色进行区分。

3.2 实验结果

通过研究目标检测算法在实验室场景中的目标检测准确率以及场景图生成准确率,以探究 AGV 是否能够全面、准确地探测周围环境。场景理解系统在实验室场景下的目标检测平均准确率详见表4。得益于数据集的数据分布均匀、标注良好、环境光照强度统一、实验室场地面积较小且障碍物尺寸较大,基本上各类别均达到了 86% 以上的准确率,为后续实验奠定了基础。

场景图生成准确率及对应的风险缓解执行实验结果见表5。在 3 个静态场景中,系统展现了良好的环境布局识别能力。其中,静态回线场景下的场景图生成准确率稍

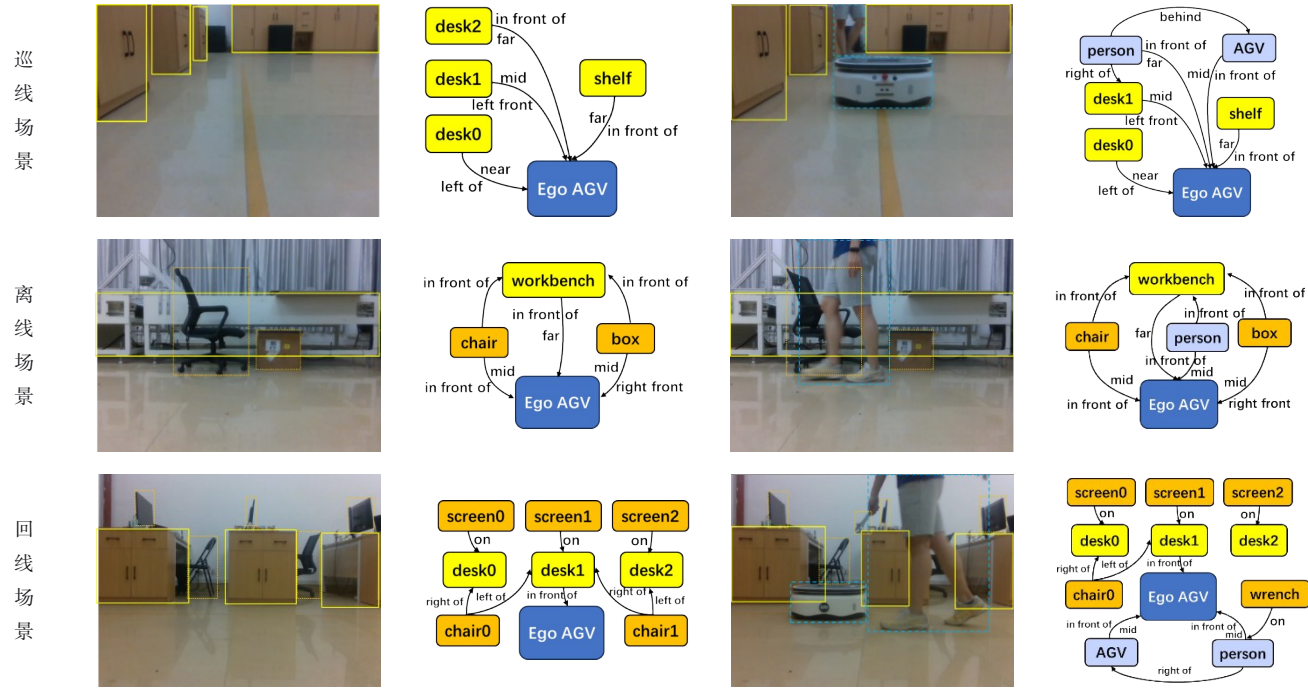


Fig. 5 Examples of scene images generated by AGV in three different scenarios

图5 AGV在3种场景下生成的场景图像示例

Table 4 Average accuracy of object detection and recognition in laboratory scenes

表 4 实验室场景下物体检测识别平均准确率			
障碍物类别	平均准确率	障碍物类别	平均准确率
工人	0.91	屏幕	0.90
AGV	0.92	轨道	0.89
扳手	0.86	桌子	0.93
纸箱	0.89	工作台	0.94
椅子	0.88	架子	0.92

Table 5 Construction of scene graph and execution results of risk mitigation strategies

表 5 场景图构建及风险缓解策略执行结果			
实验类型	场景图生成 准确率	风险识别结果 (实际/预期)	策略执行结果 (实际/预期)
静态巡线场景	0.93	安全/安全	保持速度/保持速度
静态离线场景	0.95	安全/安全	降低速度/降低速度
静态回线场景	0.90	低风险/安全	保持速度/保持速度
动态巡线场景	0.83	低风险/安全	降低速度/降低速度
动态离线场景	0.86	高风险/高风险	停止运行/停止运行
动态回线场景	0.82	低风险/安全	降低速度/降低速度

低于其他两个场景。这一现象可归因于回线场景中多元障碍物(包括但不限于多个桌子、椅子和屏幕)的存在,其构成了更复杂的空间布局,进而对系统在辨识障碍物间相对位置及相互关系方面提出了更高要求。当引入具有移动能力的障碍物至场景中,即在动态场景下,场景图生成准确率有所下降。然而,系统通过预设的综合评估机制,在若干场景中实现了更为保守的风险识别策略。此策略根据单一障碍物场景和多障碍物组合场景中评估的风险,将系统的当前风险等级设置为二者中较高的风险等级。尽管这种保守策略不可避免地会对系统的整体运行效率

产生轻微的抑制效果,但其显著提升了系统规避碰撞的能力,确保了AGV在运行过程中的安全边际。在策略执行结果上,所有实际执行的策略均严格遵循基于实时风险识别结果所制定的预期计划,未出现偏离预期的行为。

4 结语

本文针对AGV在动态工业环境中存在的感知局限性问题进行研究,设计了一种结合场景图的风险缓解框架。该框架利用所提出的Ego-AGV场景图来感知环境,不仅可以理解环境中障碍物与自身的空间关系,而且可以理解环境中障碍物的组合关系。基于感知到的环境信息,系统能够准确评估即时场景中的潜在碰撞威胁,并量化其风险等级,进而触发相应的预防措施,达到风险缓解的目的。实验结果表明,在包含多类型单个目标的相对简化的场景中,该框架能够精确地生成场景图,辨识风险并及时采取针对性的风险缓解行动。后续工作将专注于改进动态场景下的场景图生成算法,以及优化在多类别相同障碍物环境下的场景图表达,以期达到更高的准确率以及安全性与效能之间的最佳平衡。

参考文献:

[1] CAO F K, ZHUANG Y, YAN F, et al. Long-term autonomous environment adaptation of mobile robots: state-of-the-art methods and prospects [J]. Acta Automatica Sinica, 2020, 46(2): 205-221.
曹凤魁, 庄严, 闫飞, 等. 移动机器人长期自主环境适应研究进展和展望[J]. 自动化学报, 2020, 46(2): 205-221.

[2] LI W P, GENI M, AHAN J, et al. Experimental test of mobile robot ultra-

- sonic distance measurement and obstacle avoidance system[J]. Machine Tool & Hydraulics, 2022, 50(5): 1-5.
- 李卫鹏, 买买提明·艾尼, 加合甫·阿汗, 等. 移动机器人超声波测距避障系统实验测试[J]. 机床与液压, 2022, 50(5): 1-5.
- [3] MA P, ZHANG Y, SU P J, et al. Design of embedded AGV avoidance system based on ARM[J]. Machine Tool & Hydraulics, 2019, 47(5): 61-64.
- 马平, 张勇, 苏攀杰, 等. 基于 ARM 嵌入式 AGV 的避障系统设计[J]. 机床与液压, 2019, 47(5): 61-64.
- [4] JOHNSON J, KRISHNA R, STARK M, et al. Image retrieval using scene graphs[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 3668-3678.
- [5] KRISHNA R, ZHU Y, GROTH O, et al. Visual genome: connecting language and vision using crowdsourced dense image annotations[J]. International Journal of Computer Vision, 2017, 123: 32-73.
- [6] XU D F, ZHU Y K, CHOY C B, et al. Scene graph generation by iterative message passing[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 5410-5419.
- [7] ZELLERS R, YATSKAR M, THOMSON S, et al. Neural motifs: scene graph parsing with global context[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 5831-5840.
- [8] TANG K H, NIU Y L, HUANG J Q, et al. Unbiased scene graph generation from biased training[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 3716-3725.
- [9] SUN S Z, ZHI S F, LIAO Q, et al. Unbiased scene graph generation via two-stage causal modeling[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(10): 12562-12580.
- [10] RIAZ H, TERRA A, RAIZER K, et al. Scene understanding for safety analysis in human-robot collaborative operations[C]//2020 6th International Conference on Control, Automation and Robotics, 2020: 722-731.
- [11] KENFACK F K, SIDDIKY F A, BALINT-BENCZEDI F, et al. Robot-VQA—a scene-graph-and deep-learning-based visual question answering system for robot manipulation[C]//2020 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2020: 9667-9674.
- [12] AMIRI S, CHANDAN K, ZHANG S Q. Reasoning with scene graphs for robot planning under partial observability[J]. IEEE Robotics and Automation Letters, 2022, 7(2): 5560-5567.
- [13] LI S F, ZHENG P, WANG Z X, et al. Dynamic scene graph for mutual-cognition generation in proactive human-robot collaboration[J]. Procedia CIRP, 2022, 107: 943-948.
- [14] ZHOU F B, ZHAO H L, LIU H P. Indoor mobile robot target search based on the scene graphs[J]. CAAI Transactions on Intelligent Systems, 2022, 17(5): 1032-1038.
- 周方波, 赵怀林, 刘华平. 基于场景图谱的室内移动机器人目标搜索[J]. 智能系统学报, 2022, 17(5): 1032-1038.
- [15] YANG D, XU X, XIONG M C, et al. SRI-Graph: a novel scene-robot interaction graph for robust scene understanding[C]//2023 IEEE International Conference on Robotics and Automation, 2023: 8171-8178.
- [16] JIN T L, GUO F T, MENG Q W, et al. Fast contextual scene graph generation with unbiased context augmentation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 6302-6311.
- [17] TIAN P. Research on method of robot vision scene understanding[D]. Harbin: Harbin Engineering University, 2022.
- 田朋. 机器人视觉场景理解方法研究[D]. 哈尔滨: 哈尔滨工程大学, 2022.
- [18] YANG J W, LU J S, LEE S, et al. Graph R-CNN for scene graph generation[C]//Proceedings of the European Conference on Computer Vision, 2018: 670-685.
- [19] MATTHIAS B, OBERER-TREITZ S, STAAB H, et al. Injury risk quantification for industrial robots in collaborative operation with humans[C]//ISR 2010 (41st International Symposium on Robotics) and ROBOTIK 2010 (6th German Conference on Robotics), 2010: 1-6.

(责任编辑:黄健)