

# 融合地方语言识别与自动跟随与机械臂辅助抓夹的超市智能购物车设计方案

韦铃馨, 杨竺伟, 农树宇, 闭淑琦, 吴茂达, 滕 煦, 梁 巍, 班 璐\*

南宁学院交通运输学院, 广西 南宁

收稿日期: 2025年12月12日; 录用日期: 2026年1月2日; 发布日期: 2026年1月16日

## 摘要

随着互联网时代的迅速发展, 网购也逐渐兴起。因为网购具有节约时间、方便快捷等特点, 所以也就成为了大多数消费者的消费渠道之一。但是这也使传统零售业遭受到了极大的打击, 而超市作为人们主要消费场所之一, 急需优化超市线下服务体系, 带给消费者更好的购物体验。为了解决传统超市购物中顾客推车费力、商品抓取不便及跨方言交互不畅等问题。本文设计并研究了一款融合多技术的超市智能购物车。该购物车的核心为嵌入式系统, 集成了地方语言识别模块、自动跟随导航模块与机械臂辅助抓夹模块。该购物车可以很大程度上解决人们在超市的不便, 最大化实现超市的智能化。

## 关键词

地方语言识别, 自动跟随, 机械臂辅助抓夹, 智能购物车

# Supermarket Smart Shopping Cart Design Scheme Integrating Local Language Recognition and Automatic Following and Robotic Arm-Assisted Grasping

Lingxin Wei, Zhuwei Yang, Shuyu Nong, Shuqi Bi, Maoda Wu, Yu Teng, Wei Liang, Lu Ban\*

School of Transportation, Nanning University, Nanning Guangxi

Received: December 12, 2025; accepted: January 2, 2026; published: January 16, 2026

\*通讯作者。

## Abstract

With the rapid development of the Internet era, online shopping is also gradually emerging. Because online shopping has the characteristics of saving time, convenience and speed, it has become one of the consumption channels of most consumers. However, this has also hit the traditional retail industry hard, and supermarkets, as one of people's main consumption places, urgently need to optimize the offline service system of supermarkets to bring consumers a better shopping experience. In order to solve the problems of strenuous customer carts, inconvenient product grabbing and poor cross-dialect interaction in traditional supermarket shopping. This paper designs and studies a supermarket smart shopping cart that integrates multiple technologies. At the heart of the shopping cart is an embedded system that integrates a local language recognition module, an auto-follow navigation module, and a robotic arm-assisted gripping module. The shopping cart can greatly solve people's inconvenience in the supermarket and maximize the intelligence of the supermarket.

## Keywords

Local Language Recognition, Auto-Follow, Arm Auxiliary Grasping Clip, Smart Shopping Cart

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

在新零售与人工智能融合背景下，超市购物需往智能化升级。传统购物车存在操作繁琐、物品搬运不便、对老年及行动不便人群友好度低等问题[1]，难以满足消费者高效便捷的需求。随着地方语言识别技术成熟、自动跟随算法优化及机械臂抓取技术突破，智能购物车研发具备了技术支撑。这类购物车可节省消费者体力、提升体验，同时帮助超市优化运营效率、增强竞争力。目前现有产品多聚焦路径规划与导航，在地方语言交互、动态场景精准跟随及柔性抓夹辅助方面存在不足，缺乏对多语言环境及不同人群的适配。因此，本文以三大核心技术为支撑，研究地方语言识别模型、自动跟随算法及机械臂柔性抓夹机制，实现语音交互、自主跟随和辅助抓取功能，为超市智能化设备研发提供新思路。

## 2. 研究的目的

本项目旨在设计一款集成地方语言识别、自动跟随及机械臂辅助抓夹功能的超市智能购物车。为了解决传统购物的痛点和需求，为不熟悉普通话的用户很大程度上解决语言沟通障碍，借助自动跟随功能用户可以解放双手、不用老盯着车，用机械臂辅助抓取减轻老年人与行动不便者弯腰踮脚的身体负担，能轻松不少。所以，最终想打造一个稳定、自动化及人性化的智能购物系统，这更能体现社会对老年群体的包容性，提升了用户购物的满意度，让购物全过程更省心省力，同时也优化了超市的运营效率。

## 3. 国内外研究与应用案例分析

### 3.1. 国内

国内在智能购物车领域的探索聚焦“成本控制”与“场景适配”，形成学术研究与商业化产品并行的格局，但存在功能碎片化、特殊人群适配不足等问题。

学术研究方面,李文豪(2024)提出的基于自动跟随的智能购物车[2],明确了自动跟随与感知避障的系统架构,模块划分清晰,还针对跟随稳定性提出优化策略以提升用户体验,但该方案仅聚焦跟随功能,无语音交互与机械臂抓取模块,也未考虑方言用户与行动不便人群的需求;冯迎宾等(2025)的基于单目视觉的机械臂抓取算法[3],通过改进视觉定位算法提升了机械臂抓取精度与鲁棒性,融合图像处理与控制理论解决了定位不准的问题,却仅针对机械臂抓取技术本身,未与购物车场景结合,也未涉及自动跟随与语音交互,功能相对孤立;张允耀黄鹤鸣,张会云(2021)针对复杂噪声环境下的语音识别问题,提出“深度学习+信号处理”的抗噪声策略[4],有效提升了嘈杂环境下的识别准确率,降低了噪声干扰,适配超市复杂声学环境,但该研究仅优化了普通话语音识别精度,未扩展至方言识别,也未与购物车的跟随、抓取等其他模块联动。

商业化产品方面,超嗨科技 S600 智能购物车(见图 1)集成了自助结算、室内定位、智能防损等 8 大功能,商业化成熟度高,支持生鲜称重与精准营销,适配超市运营需求,且已在全球投放 3 万台,具备丰富的规模化应用经验,但该产品无机械臂抓取功能,无法辅助老年人拿取商品,仅支持普通话交互,无方言识别能力,适配性有限,且单台成本高达 8000~12,000 元,中小型商超推广难度较大;银座 AI 智能购物车(见图 2)支持“边选边扫边记账”,使结算效率提升 30%,具备智能导航找货功能,降低了新用户找货难度,还能个性化推送优惠券以增强用户黏性,但同样无机械臂抓取功能,无法辅助行动不便人群,交互依赖触摸屏,无语音控制功能,老年用户操作门槛高,且仅覆盖山东区域,规模化应用不足;学术转化方案 NovelCart 智能购物车(2020)集成了 UWB 定位与 RFID 识别技术,实现了自动结算,成本低于同类商业化产品,支持传统购物车改装,但无语音交互与机械臂模块,功能局限于“定位、结算”,未考虑方言与特殊人群需求,适配性较差。



**Figure 1.** Super Hi Technology S600 smart shopping cart (Source: Baidu Website)  
**图 1.** 超嗨科技 S600 智能购物车(来源: 百度网站)



**Figure 2.** Ginza AI smart shopping cart (Source: Xiaohongshu app)  
**图 2.** 银座 AI 智能购物车(来源: 小红书 app)

### 3.2. 国外

国外研究起步早,侧重“技术前瞻性”与“服务机器人融合”,但存在本土化不足、成本过高、忽视方言需求等问题。

学术研究方面, Kanda, T., *et al.* (2010)开发的超市导购机器人[5],首次将自然语言处理引入零售场景,具备商品信息引导功能,奠定了语音交互在超市场景的应用基础,但该机器人仅为固定设备,无自动跟随与抓取功能,且仅支持英语交互,无方言适配能力,本土化适配性差; Mekonnen, A.等(2011)针对拥挤环境下的移动机器人感知需求[6],提出多模态感知(视觉+激光雷达)技术,提升了人群中跟踪稳定性,优化了动态场景下的障碍物检测能力,适配超市密集人流环境,但该研究仅聚焦感知技术,无语音交互与机械臂模块,未结合购物车场景进行设计,缺乏商品抓取与交互相关功能; Bicchi, A., & Kumar, V. (2000)对机器人抓取策略的研究[7],提出了适用于非结构化环境的多种抓取策略,覆盖多种抓取场景,鲁棒性强,为机械臂应用奠定了理论基础,但该研究未针对超市轻量级、易损品等特殊商品进行优化,且无自动跟随与语音模块,无法与购物车系统协同工作。

商业化产品方面, Simbe Robotics Tally 机器人(2010~2020)能够实现自动化货架扫描与库存盘点,提升超市运营效率,具备自主巡航避障能力,适配货架密集环境,但该产品仅聚焦超市货架管理,无用户交互与跟随功能,无机械臂抓取模块,无法辅助消费者购物,且单台成本高达 15,000~20,000 元,不面向个人用户; Amazon Scout 购物机器人(见图 3)具备高精度户外导航与自主配送能力,多模态感知技术成熟,路径规划响应快(0.3 s),适配复杂路况,但该产品应用于“末端配送”场景,并非超市购物场景,无语音交互与机械臂模块,无法满足超市购物需求,且未考虑方言与特殊人群需求,本土化适配性差。



Figure 3. Amazon Scout shopping robot (Source: Baidu Website)  
图 3. Amazon Scout 购物机器人(来源: 百度网站)

### 3.3. 对国内外现状的“针对性突破”

综合来看,国内外智能购物车研究要么功能太单一,只解决一个问题;要么压根不管方言用户、老年人这些特殊人群;要么卖得超贵,普通超市用不起;要么不符合国内的使用场景。结果就是三个核心麻烦一直没解决:不会说普通话的人跟车无法沟通、行动不方便的人拿不到货架上的东西、购物全程得自己忙前忙后,又累又麻烦。

而选择“地方语言识别+自动跟随+机械臂辅助抓夹”的组合设计,正是以“补位需求、平衡成本、适配本土”为目标,既填补了国内外方案的功能空白,又兼顾了实用性与落地性,为国内超市智能化提供了“低成本、全流程、高适配”的差异化解决方案,这也是选择该设计的根本原因。

## 4. 系统总体设计

### 4.1. 外型设计

#### (1) 基础结构

参考市面的主流超市购物车为基础框架[8]，保留储物篮用于存放商品，同时预留改装空间以集成核心模块。

#### (2) 核心组件集成

配备可快速拆卸的模块化机械臂，需兼容不同购物车，外型可能为轻量化金属材质，关节灵活以实现抓夹功能；

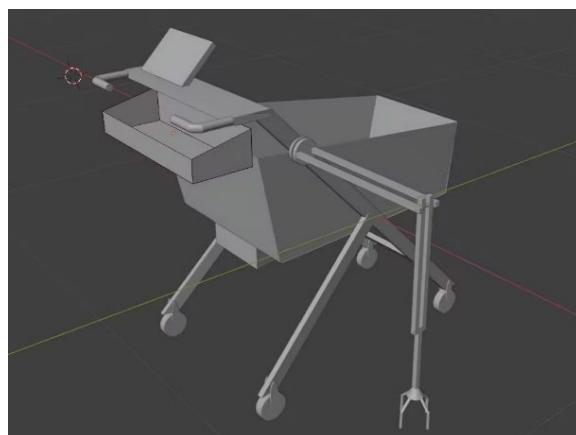
车身搭载麦克风用于方言语音识别，外型可能为小型圆形或条形，安装于车身显眼且便于收音的位置，如手柄处或储物篮侧面；

底部装有移动底盘，采用全向轮或差速轮，外型紧凑以不占用过多购物空间，同时确保灵活移动，并集成激光雷达或视觉传感器，用于避障和自动跟随，可能安装于底盘边缘或车身顶部，外型为小型探头状。

#### (3) 交互[9]界面

设有触摸屏，用于商品查询、导航显示等，外型可能为矩形，安装于购物车手柄附近或储物篮前端，方便用户操作。

具体外型见图4。



**Figure 4.** Basic appearance design  
**图 4.** 基本外型设计

### 4.2. 硬件设计

本智能购物车硬件系统采用“核心控制、感知交互、执行驱动、能量供给”四层分布式架构，以树莓派5(8 GB)为中央处理单元，通过标准接口联动四大模块，形成“感知、决策、执行、供能”完整闭环，硬件适配性强、支持扩展，覆盖超市购物全流程。

核心功能模块详解：

#### (1) 核心计算模块

树莓派5作为系统“中央大脑”，承担全模块数据处理与指令调度，高性能四核处理器算力强劲，可替代传统工业级嵌入式平台，满足多任务并行需求；配备高速内存与丰富接口，能同步运行三大核心算法，无卡顿低延迟，开发生态丰富、扩展性强。

#### (2) 感知模块

Intel RealSense D415 作为“环境感知中枢”，采用主动红外立体视觉技术，同步采集彩色图像与深度数据，实现二维和三维融合感知；可精准覆盖超市场景，复杂环境下稳定输出数据，支持百余种商品识别与高精度用户追踪，体积小、功耗低，适配嵌入式集成且开发周期短。

#### (3) 机械臂模块

MeArm 开源机械臂作为“执行终端”，是感知决策与实际操作的关键连接，4 自由度串联结构工作范围适配购物车与货架低层；搭载高扭矩舵机，轻量级商品抓取精度高、响应快、成功率高，开源架构支持二次开发，扩展性极强。

#### (4) 语音交互模块

启英泰伦 CI-D026S07J-BT 作为“人机交互入口”，接收用户指令并反馈结果，解决传统手动操作局限；内置专用芯片支持离线交互，可识别普通话与地方方言，抗干扰算法保障嘈杂环境高准确率，支持自定义指令与无线连接，无网络依赖且成本可控。

#### (5) 移动底盘模块

开源差速小车底盘(含电机、编码器)作为“移动载体”，承担移动、跟随、导航任务；金属材质承载能力强，差速驱动适配步行节奏，集成编码器与控制算法，支持高精度行驶与模式快速切换，具备急停功能，开源设计兼容 ROS 系统[10]，便于后续升级。

#### (6) 电源系统

锂电池和降压模块作为“能量供给中枢”，为全模块提供稳定供电，决定续航与运行稳定性；高能量密度锂电池满电支撑数小时连续运行，配合补电可满足单日使用，多路降压模块精准供电，具备安全保护、快速充电功能，兼顾续航与实用性。

### 4.3. 软件设计

#### (1) 系统基础与通信模块

Ubuntu 22.04 + ROS2 Humble 是整个购物车系统的基础操作系统和通信框架，具有稳定且高效率的运行环境，并支持多节点间的数据传输与功能协作等方式。使得购物车的各个硬件组件和功能模块之间能够高效联动与运行。

#### (2) 视觉感知模块

YOLOv5 [11] 主要用于识别商品和用户，他替代了 YOLOv8，对硬件算力的要求也进行了降低，能够在树莓派 5 这类嵌入式平台上较好运行。因为和 YOLOv8 版本相比，YOLOv5 不但大幅度的降低了计算的开销，并还保持着极高的精准度，实现了实时检测目标功能。为购物车的商品识别、用户定位等任务提供视觉支持。

#### (3) 导航规划模块

Navigation2 运用于购物车的路径规划与自动避障。可以使购物车在复杂环境中能够自主地进行路线的规划。行驶过程中在避开障碍物的同时，还能实现自动跟随和路径行驶等功能。他同时还是 ROS2 生态中的导航功能包，其中包含了 ROS1 中导航功能的诸多优点，支持 2D 导航，地图的构建、定位、路线的规划和运动控制等完整的流程。

#### (4) 机械臂控制模块

MoveIt! 用于规划 MeArm 开源机械臂的运动轨迹，使机械臂能精准地抓取商品、放置等动作，实现对机械臂的高效控制。同时还是 ROS 生态中成熟的机械臂运动规划框架，支持运动学求解、轨迹规划、碰撞检测等功能，多种型号的机械臂都能适配，有利于开发者快速实现机械臂的自动化控制。

### (5) 语音交互模块

启英泰伦语音 SDK 该模块由启英泰伦公司提供, 该公司专注于语音识别技术, 尤其在离线语音识别和方言支持方面有一定优势, 使购物车能够在无网络环境下仍可完成精准的语音指令识别与响应, 尤其适用于方言复杂地区。同时还免费提供 SDK, 能够降低了语音交互功能的开发门槛。

## 5. 系统的工作流程

### (1) 用户交互与任务触发

通过启英泰伦 CI-D026S07J-BT 语音模块 + 语音 SDK, 识别地方语言并解析语义, 如“找薯片并放入购物车”, 触发商品识别、抓取等核心任务, 启动视觉感知与机械臂模块。

### (2) 环境与商品感知

Intel RealSense D415 深度相机采集 RGB 图像与深度数据, 经 YOLOv5 算法识别用户和目标商品, 输出三维位置信息, 同时构建环境深度地图支撑导航。

### (3) 自动跟随导航

以树莓派 5 为核心, 结合 Navigation2 框架, 基于视觉定位规划最优避障路径, 驱动差速底盘跟随用户, 通过深度数据动态保持安全距离。

### (4) 商品抓取与放置

到达商品区域后, 树莓派 5 通过 MoveIt 框架规划 MeArm 机械臂轨迹, 精准抓取  $\leq 500\text{g}$  商品, 跟随用户至购物车区域完成放置。

### (5) 多任务并行与持续服务

系依托树莓派 5 多任务能力, 执行当前任务时可同步监听新语音指令, 如“再找一瓶可乐”, 重复上述流程实现连续购物辅助, 直至用户指令触发任务终止, 如“帮我推到收银台”。

总之就是, 系统以语音交互为入口, 视觉与深度感知为纽带, 导航与机械臂为执行载体, 形成“听、看、走、做”的全流程智能化闭环(见图 5), 大幅提升超市购物的便捷性与趣味性, 为用户带来更好的体验, 让用户走出互联网, 又身在互联网。

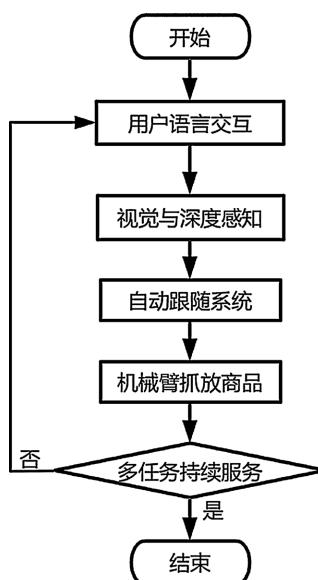


Figure 5. System workflow  
图 5. 系统工作流程

## 6. 潜在局限性

### (1) 机械臂性能受限

负载上限仅 500 g, 无法抓取大米、食用油等重物; 抓取要 1.2 到 1.5 秒与专业分拣设备(0.3~0.5 s/件)差距明显, 要是购物的人多、需要快速拿取商品的场景, 效率就不够高; 而且抓鸡蛋、薯片这类容易坏或碎的商品时, 也没法保证一定不出错。

### (2) 环境适应能力不足

方言识别在极端嘈杂环境(比如促销活动现场)准确率可能降至 82%以下, 复杂口音(混合了好几种方言的口音)识别效果有限; 自动跟随时对快速移动的障碍物(如突然跑过去的小孩)响应延迟约 0.8 s, 存在轻微安全隐患。

### (3) 效率与规模化短板

一台购物车只能跟着一个人用, 无法满足多人同时使用需求; 机械臂部件存在磨损风险, 长期规模化应用的维护成本高于无机械臂的同类产品。

### (4) 功能覆盖局限

未集成自助结算功能, 选完商品后需人工结账, 未能实现“选商品、结账”一站式完成; 货架高度适配范围仅 0.3~1.5 m, 对高层货架商品无法抓取辅助。

## 7. 研究创新与差异化分析

相较于当前智能购物车研究, 本项目在技术融合、成本控制及场景适配性上形成显著优势:

(1) 技术协同创新: 现有研究多聚焦单一功能, 国外 Simbe Robotics 的 Tally 机器人方案仅实现货架扫描, 国内 STM32 方案[12]局限于基础避障与循迹, 均未覆盖全流程。本项目首次将地方语言识别(启英泰伦 CI-D026S07J-BT 模块)、视觉和传感双模态自动跟随(Intel RealSense D415+ 差速底盘)、轻量化机械臂抓取(MeArm)三大核心技术深度耦合, 可响应方言指令完成“商品定位、自动跟随、抓取商品”闭环操作, 有效解决传统方案功能碎片化问题。

(2) 成本与性能平衡: 同类产品依赖高端平台与定制组件, 成本常破万元。本项目以树莓派 5 为核心, 搭配开源低成本组件, 控制成本的同时, 核心性能优于部分高价方案, 突破“高性能即高成本”困境。

(3) 场景精准适配: 现有方案多为标准化设计, 忽视特殊人群需求。本项目通过离线方言识别适配非普通话用户, 模块化机械臂支持改装, 围绕“减少体力消耗”设计, 贴合超市复杂人群与空间环境。

这种差异化设计不仅提升了智能购物车的实用价值, 更为零售场景智能化设备研发提供了“低成本、全流程、高适配”的可复制路径, 区别于现有研究的单一技术探索或高成本落地模式。

## 8. 结语

本研究成功研发出集成地方语言识别、自动跟随与机械臂辅助抓取的超市智能购物车, 通过开源硬件与算法协同实现全流程智能化, 有效解决传统购物痛点, 其创新在于技术深度融合、低成本适配与细分人群适配, 但机械臂负载有限, 复杂货架环境抓取精度待提升, 方言识别语种覆盖范围可拓展, 未来可优化机械臂性能、拓展方言种类并集成支付功能, 进一步完善购物链路智能化。

## 致 谢

本研究资金来源以下项目: “CartPilot 推车智驾”——基于地方语言识别和自动跟随以及机械臂辅助抓夹的超市智能购物车(202511549068)。

## 参考文献

- [1] 郭会娟, 郑翠媛, 刘思雅. 包容性视角下的老年家用智能购物车设计研究[J]. 合肥师范学院学报, 2023, 41(4): 78-80.
- [2] 李文豪. 基于自动跟随的智能购物车设计与研究[J]. 现代计算机, 2024, 30(20): 110-113.
- [3] 冯迎宾, 刘潘, 王天龙, 曾城锋. 基于单目视觉的机械臂定位与抓取算法研究[J]. 沈阳理工大学学报, 2025, 44(6): 14-22.
- [4] 张允耀, 黄鹤鸣, 张会云. 复杂噪声环境下语音识别研究[J]. 计算机与现代化, 2021(9): 68-74.
- [5] Kanda, T., Shiomi, M., Miyashita, Z., Ishiguro, H. and Hagita, N. (2010) A Communication Robot in a Shopping Mall. *IEEE Transactions on Robotics*, **26**, 897-913. <https://doi.org/10.1109/tro.2010.2062550>
- [6] Mekonnen, A.A., Lerasle, F. and Zuriarrain, I. (2011) Multi-Modal Person Detection and Tracking from a Mobile Robot in a Crowded Environment. *Proceedings of the International Conference on Computer Vision Theory and Applications*, Vilamoura, 5-7 March 2011, 511-520.
- [7] Bicchi, A. and Kumar, V. (2000) Robotic Grasping and Contact: A Review. *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings*, San Francisco, 24-28 April 2000, 348-353. <https://doi.org/10.1109/robot.2000.844081>
- [8] 郑钧丹. 超市智能购物车及应用设计研究[J]. 艺术与设计(理论), 2019, 2(8): 108-110.
- [9] 龙正江, 刘丽华, 林力, 等. 基于物联网可交互超市智能购物车设计[J]. 辽宁科技学院学报, 2019, 21(3): 17-18.
- [10] 张宸溪. 基于 ROS 的苹果采摘机器人目标检测与抓取规划研究[D]: [硕士学位论文]. 北京: 北京林业大学, 2023.
- [11] 周昊, 陆恣. 基于改进 YOLOV5 的超市商品包装老化自动识别技术[J]. 包装工程, 2025, 52(1): 78-81.
- [12] 成雷, 郑云天, 李月华, 刘旺, 金海涛, 胡世鹏. 基于 STM32 的超市智能购物车[J]. 电子世界, 2017(23): 110-112.