

# 基于透视校正视觉识别的小米狗智能导航系统

## ——复杂赛道环境下的自主识别与运动控制

### 1 项目背景与目标

随着机器人技术的快速发展，四足机器人在复杂环境中的自主导航能力成为了研究和应用的热点。小米CyberDog作为消费级四足机器人的代表产品，具备了良好的硬件基础和开发平台特性。然而，在面对复杂赛道环境时，传统的视觉识别算法往往受限于摄像头角度、目标姿态和光照条件等因素，难以实现稳定可靠的目标识别和路径规划。

本项目针对小米狗在复杂赛道环境中的应用需求，开发了一套基于透视校正技术的智能视觉识别系统。该系统能够在各种角度和光照条件下准确识别QR码导航标识，并结合四足机器人的运动特性，实现复杂赛道的自主通过。项目的核心目标是提升小米狗在结构化环境中的自主导航能力，使其能够完成包含多种特殊路段的复杂赛道任务。

传统的QR码识别算法通常要求目标与摄像头保持正对的角度关系，这在实际应用中往往难以满足。特别是在动态行走过程中，机器人的姿态变化、地面不平整以及目标标识的放置角度都会影响识别效果。本项目通过创新的透视校正算法，有效解决了这一技术难题，显著提升了系统在复杂环境下的鲁棒性和实用性。

项目的技术创新主要体现在多策略融合的视觉识别架构和自适应的运动控制策略两个方面。通过将计算机视觉、机器学习和机器人控制技术的有机结合，形成了一套完整的智能导航解决方案，为四足机器人在实际应用中的部署提供了重要的技术支撑。

### 2 核心技术方案

#### 2.1 透视校正视觉识别算法

本项目的核心技术创新在于开发了一套多层次的透视校正视觉识别算法。该算法采用渐进式检测策略，首先尝试标准的直接检测方法，当检测失败时自动启动透视校正流程，最后采用多角度旋转检测作为兜底策略。这种分层检测机制确保了在各种复杂情况下都能实现可靠的目标识别。

透视校正的核心算法基于四点变换原理。系统通过边缘检测和轮廓分析技术，自动识别图像中可能包含QR码的矩形区域。对于每个候选区域，算法计算最优的透视变

换矩阵，将倾斜或扭曲的QR码校正为标准的正视图。校正过程中采用双三次插值算法，保证了变换后图像的几何精度和视觉质量。

图像预处理阶段集成了多种增强技术，包括CLAHE对比度限制自适应直方图均衡化、高斯模糊降噪和锐化滤波。系统根据当前图像的统计特征自动选择最适合的预处理组合，显著改善了在不同光照条件下的识别性能。预处理模块还实现了多尺度处理能力，支持0.8到2.0倍的缩放范围，适应不同距离下的目标识别需求。

ROI区域优化策略专注于图像的中上方区域，有效减少了无关背景的干扰。通过分析小米狗的运动特性和摄像头安装位置，系统将检测区域限制在图像高度的20%到100%范围内，显著提高了检测效率和准确性。这种空间注意力机制降低了计算复杂度，同时提升了实时性能。

置信度过滤机制通过维护时序检测历史，实现了稳定的结果输出。系统保存最近3帧的检测结果，计算各候选结果的出现频率，只有当置信度超过50%阈值时才输出最终识别结果。这种时域滤波有效地减少了误检和抖动现象，提高了系统的稳定性和可靠性。

## 2.2 小米狗运动控制集成

针对小米狗的硬件特性和运动学特点，项目开发了专门的运动控制接口和步态规划算法。控制系统采用分层架构设计，上层负责任务规划和目标设定，中层实现步态生成和轨迹规划，底层完成关节控制和力矩输出。这种分层设计保证了系统的模块化和可维护性。

步态控制策略基于小米狗的实际运动能力进行了优化设计。系统实现了两种基本控制模式：姿态控制模式用于静态稳定和准备阶段，用户步态模式用于动态行走和导航。姿态控制阶段通过四足支撑实现稳定的静止状态，身体重心适当下降以增强稳定性。步态控制阶段采用保守的运动参数，前进速度设置为0.1m/s，步高参数优化为0.02米，在保证稳定性的前提下实现平稳行走。

参数调优过程结合了仿真验证和实物测试。通过大量的仿真实验确定了权重参数的初始配置，然后在实际的小米狗平台上进行微调和验证。最终确定的权重向量[500.0, 10.0, 10.0, 50.0, 2500.0, 10.0]经过了严格的稳定性测试，能够在各种地面条件下保持良好的行走效果。

安全保护机制包括碰撞检测、姿态监控和紧急停止功能。系统实时监测机器人的姿态状态，当检测到异常倾斜或碰撞风险时，立即切换到安全模式。紧急停止功能能够在100毫秒内使机器人停止运动并保持稳定姿态，为操作人员提供了必要的安全保障。

实时性优化通过多线程并行处理和算法优化实现。视觉处理模块与运动控制模块采用独立的线程，通过高效的消息传递机制实现数据同步。关键算法采用了向量化计算和查找表技术，显著提升了处理速度。整个系统的控制周期稳定在50毫秒以内，满足了实时控制的要求。

## 3 赛道任务实现

### 3.1 导航标识识别

QR码导航标识的识别是赛道任务完成的关键技术环节。系统支持['A-1', 'A-2', 'B-1', 'B-2']四种目标编码的精确识别，识别准确率达到95%以上。透视校正算法能够处理倾斜角度达到 $\pm 45^\circ$ 的QR码，适应各种安装角度和观察条件。平均识别时间控制在100毫秒以内，为实时路径规划提供了及时的导航信息。

多目标导航策略实现了复杂任务的自动化执行。系统能够根据识别到的导航标识自动更新任务状态和目标设定，支持顺序访问、条件分支等多种任务模式。当识别到特定的导航标识时，系统自动切换到相应的行为模式，执行预定义的动作序列或路径规划。

实时监控与调试功能为系统优化和问题诊断提供了重要支持。系统提供多窗口的实时图像显示，包括原始图像、ROI区域、透视校正结果等关键信息。详细的日志记录系统采用分级输出，记录系统运行的关键事件和状态变化。参数动态调整功能允许在运行过程中修改关键参数，便于现场调试和优化。