

# 基于图像识别的智能语音导盲杖控制系统研究

唐琪, 郭武, 刘怀海, 刘建平  
湖南应用技术学院, 湖南 常德 415100  
DOI: 10.61369/TACS.2026020020

**摘 要 :** 为了满足盲杖使用者的实际出行需求, 解决盲杖使用者出行安全环境感知缺乏、交互方式单一等问题, 设计出一种基于图像识别的智能语音导盲杖控制系统至关重要。此类系统可以很好地识别出常见的出行障碍物, 而且语音反馈及时准确, 操作方便。基于此, 本文将简述系统总体结构设计思路, 然后对系统的硬件设计、软件设计和关键功能原理及算法进行探讨, 以期为进一步改善盲杖使用者的出行体验提供理论参考。

**关 键 词 :** 图像识别; 智能语音; 导盲杖; 控制系统

## Research on the Control System of Intelligent Voice Guide Cane Based on Image Recognition

Tang Qi, Guo Wu, Liu Huaihai, Liu Jianping  
Hunan Applied Technology University, Changde, Hunan 415100

**Abstract :** To meet the actual travel needs of guide cane users and solve problems such as insufficient perception of safe travel environments and single interaction methods, it is crucial to design an intelligent voice guide cane control system based on image recognition. Such a system can effectively identify common travel obstacles, with timely and accurate voice feedback and convenient operation. Based on this, this paper will briefly describe the design idea of the overall system structure, and then discuss the system's hardware design, software design, as well as the principles and algorithms of key functions, aiming to provide theoretical reference for further improving the travel experience of guide cane users.

**Keywords :** image recognition; intelligent voice; guide cane; control system

随着人工智能、计算机视觉、嵌入式技术和语音交互技术的日新月异, 智能导盲设备的研发迎来了新的机遇。图像识别技术可以快速准确地获取周围环境中的视觉信息, 识别出障碍物、交通标志、行人等重要的物体, 语音交互技术给视障人提供了一种简单的操作方式, 不需要依靠视觉的帮助就可以完成指令的输入以及信息的获取。研究基于图像识别的智能语音导盲杖控制系统, 可以很好地克服传统导盲设备的不足, 给盲杖使用者视障人士提供更为全面、准确的出行辅助, 具有重要的社会意义和应用价值。

## 一、系统总体结构设计思路概述

### (一) 系统功能需求分析

根据视障人士的出行需求以及现有的导盲设备的不足, 智能语音导盲杖控制系统需要具备以下功能需求: 一是图像传感器能够采集周围环境图像, 并运用图像识别算法识别墙壁、台阶、坑洼、行人、车辆、井盖等常见的出行障碍物, 判断出障碍物的类别和距离, 给使用者正确的反馈信息。二是使用者可以利用语音指令来控制设备的开启、关闭以及模式的切换等操作, 也可以把障碍物信息、环境信息等用语音的形式告诉使用者, 从而达到双向语音交互的效果, 操作简单, 无需视觉辅助<sup>[1]</sup>。三是当使用者在

遇到危险或者困难的时候, 可以一键触发紧急求助功能, 设备会自动发送求助信息(包含位置信息)给预设的监护人, 并且会播放求助语音, 向周围的人发出求助信息。四是采用多传感器融合的方式获取周围环境的光照强度、温度等信息, 当光照强度低于某一阈值的时候, 就会启动辅助照明功能, 提高图像识别的准确性, 还会提醒使用者注意环境的安全状况。

### (二) 系统总体架构设计

基于图像识别的智能语音导盲杖控制系统应采用分层结构设计, 分为硬件层、软件层和应用层, 各个层次之间互相配合来完成系统各项功能。系统工作流程为: 图像采集模块实时采集周围环境图像, 传感器模块采集超声波测距数据、姿态数据、环境光

照数据等,语音采集模块采集用户的语音指令;嵌入式微控制器把采集到的数据传送到软件层,软件层用图像识别算法对图像进行处理,识别出障碍物的类别和位置,用多传感器融合算法整合各个传感器的数据,用语音识别算法识别用户的语音指令;根据处理结果,软件层产生相应的控制指令,控制语音合成模块播放反馈语音,控制辅助照明模块开启或者关闭,控制紧急求助模块发送求助信息;应用层把各项功能的执行结果反馈给使用者,实现全方位的出行辅助<sup>[2]</sup>。

## 二、硬件设计

### (一) 硬件总体设计

智能语音导盲杖控制系统硬件设计以嵌入式微控制器为主,根据图像采集、语音交互、传感器感知、紧急求助等几个方面,选择合适的硬件模块,设计出合理的电路布局,保证硬件系统稳定可靠、方便携带。硬件系统主要是核心控制模块、图像采集模块、语音交互模块、传感器模块、电源模块、紧急求助模块、辅助照明模块等,各个模块通过接口和核心控制模块相连,共同工作。

### (二) 具体硬件模块设计

首先,核心控制模块是系统的大脑,控制各个硬件模块的工作,处理采集到的各种数据,执行软件算法和用户的指令,可以使用具有图像处理功能、可以训练神经网络模型的可编程 MCU——K210 芯片。K210SOC 内部还搭载了 KPU,可以在低功耗的情况下实现卷积神经网络计算,实时获取被检测目标的大小、坐标和种类,对人脸或者物体进行检测和分类。K210 还搭载了 FPIOA、UART、SPI、TIMER 等外设,能够与 STM32 等其他 MCU 进行通信并协调运行工作<sup>[3]</sup>。其次,图像采集模块是系统执行障碍物检测和识别任务的前提,用来实时获取周围环境的图像数据,给图像识别算法赋予最初的输入数据<sup>[4]</sup>。通过高清 CMOS 摄像头进行图像采集,该摄像头分辨率高、帧率高、功耗低、体积小,可以得到清楚的环境图像,满足图像识别的要求。再者,语音交互模块主要是对语音指令进行采集、识别以及语音信息的合成、播放,是系统和用户之间进行交互的主要模块。在设计中,使用 JQ8900-16P 语音模块,驱动扬声器对紧急情况进行报警。将模块与 MCU 共地,使用 STM32 的一个串口与其进行通信,应对不同紧急情况的状态,STM32 通过指令操控语音模块发出报警命令。需要注意的是,为了满足使用者的听觉要求,扬声器要能调节音量,语音合成的语速、语调也应根据使用者的需要进行设置。

此外,在设计中运用陀螺仪 MPU6050,其内部集成了三轴 MEMS 陀螺仪、三轴 MEMS 加速度计及一个可扩展的数字运动处理器。可以先通过 IIC 通信协议与 STM32 进行数据传输,输出在直角坐标系下的 x, y, z 轴的旋转角速度和加速度数据;再通过 STM32 对数据进行算法处理,就可以判断出使用者是否摔倒以及摔倒方向<sup>[5]</sup>。

## 三、软件设计

### (一) 软件总体设计

智能语音导盲杖控制软件采用模块化的设计思想,把软件系统分成图像采集与处理模块、语音交互模块、传感器数据融合模块等几个部分。各个模块彼此独立但又互相配合,依靠操作系统任务调度机制来完成系统的所有功能。

### (二) 具体软件模块设计

图像采集与处理模块是实时采集周围环境图像数据,对图像数据进行预处理,给障碍物识别模块提供高质量图像数据的模块。其中,图像采集子模块利用图像采集模块(CMOS 摄像头)对环境进行实时的图像采集,采集得到的图像格式为 RGB 格式,分辨率根据实际情况来定。采集到的图像数据经 SPI 接口传送到微控制器上,微控制器把图像数据存入存储器里,然后等待后面进行预处理<sup>[6]</sup>。为了提高图像采集速度,使用 DMA 方式传输数据,减少微控制器占用的时间,保证图像采集的及时性。图像预处理子模块是去除原始图像中的噪声、干扰等无关信息,提升图像有用信息的特征,改善障碍物识别算法的精确度与速度。预处理主要包含灰度化、滤波去噪、图像增强、图像分割等。

语音交互模块主要完成系统和用户之间双向的语音交互,即语音指令识别和语音信息合成、播放这两个子模块。语音指令识别子模块利用语音采集单元(MEMS 麦克风阵列)获取用户发出的语音指令,对语音指令进行预处理之后,再将处理后的语音指令输入语音识别算法当中进行识别。在识别的过程中,算法会把采集到的语音特征同预设的指令特征加以对比,得出识别的结果,若是识别成功,就会执行相关的操作,若是识别失败,就会发出语音提示用户重新输入指令<sup>[7]</sup>。语音信息合成与播放子模块把系统反馈的障碍物信息、环境信息、操作提示等转换成语音信号,用扬声器播放给用户。语音合成使用的是基于深度学习的轻量化合成算法,可以进行中文语音合成,可以根据文本信息产生自然流畅的语音信号,并且可以调节语速、语调、音量等,以适应不同使用者的听觉需求。当系统得到障碍物信息、环境信息等需要反馈的信息时,把文本信息传送给语音合成算法,算法产生语音信号之后,经由语音播放单元播放给用户,从而达成信息的即时反馈。

传感器数据融合模块是将各个传感器的数据进行时间戳对齐和融合处理,消除单个传感器的测量误差,提高环境感知的准确性、可靠性。该模块采用扩展卡尔曼滤波算法,把超声波传感器、激光雷达模块、IMU 姿态传感器采集到的数据进行融合处理<sup>[8]</sup>。首先对各个传感器采集到的数据进行预处理,消除数据中的噪声和异常值,保证数据的准确性;再对预处理后的数据进行时间戳对齐,将不同传感器采集到的同一时刻的数据进行匹配,误差控制在 1ms 以内;接着初始化扩展卡尔曼滤波算法,设置滤波参数,即状态方程、观测方程、协方差矩阵等;最后用预测更新循环对融合数据进行最优估计,得到障碍物距离、设备姿态等准确的信息,给障碍物识别模块和环境感知模块提供数据支持。

## 四、关键功能原理及算法

检测图像中的红绿灯和盲道可以看作是一个对图像中的特征目标进行分类并计算出位置信息的任务。目前比较流行的图像识别方式就是运用卷积神经网络技术，模仿人脑对从眼睛传来的光信息进行处理的方式来进行识别：人脑视觉皮质由几层构成，每一层都有将一定的光信号转化为电信号的功能，而每一个阶段都将所接收的信息传送给下一级阶段，在最终整合至视觉皮质 IT 区域中形成图像的细节、方向、边缘以及颜色等信息。

卷积神经网络是模仿大脑对图像进行识别的方法而提出的，包括了输入层、卷积层、池化层、全连接层以及输出层等几层结构，在每层上可以逐步挖掘出原始图片中的特征属性。具体来说：第一，输入层接收的是已经过预处理的像素矩阵；第二，其中最重要的是卷积层，不同尺寸和权重的卷积核在图像中滑动来提取局部特征信息，从而得到图像的一些基本特征，如边缘、纹理、颜色等。例如，对于包含红绿灯以及盲道标志物的图像进行处理，则卷积核能够将图像中垂直方向及水平方向上的直线区分开来，这对于定位红绿灯以及其周围的盲道至关重要<sup>[9]</sup>。最后是卷积层之后的池化层，使用如最大池化或者平均池化等方式减少特征图尺寸的同时保留主要特征信息，并防止过拟合提高泛化能力。全连接层把前面得到的高维特征转换成样例的标记空间，用激活函数，如 softmax 函数，输出各个类的概率值来达到识别红绿灯（红灯、绿灯、黄灯）和盲道等目标。

而反向传播可以优化连接卷积核的权重。优化过程可以借助公式：

$$\theta_{i+1} = \theta - a \nabla J(\theta_i)$$

式中： $\theta_{i+1}$  为优化的权重； $\theta_i$  为初始权重；a 为学习率； $\nabla J(\theta_i)$  为损失函数的梯度。

为了更好地满足检测目标并定位，在此基础上还应运用 YOLOv2 算法。当 YOLOv2 舍去全连接层，弃用以往补全图像尺寸的方式，选择在训练过程中，让其经过几次迭代后自动改变网格尺寸并继续训练，使网络对不同分辨率的图像也能有较好的预测。在 YOLOv2 模型建立的基础上，准备好红绿灯数据集和盲道数据集，进而训练出能够检测红绿灯和盲道的模型<sup>[10]</sup>。这时便可以将训练好的模型存入 SD 卡，以便 K210 载入模型。

## 五、结语

综上所述，基于图像识别的智能语音导盲杖控制系统的开发与应用为视障人士的生活带来了极大的便利，能够让他们像正常人一样行走。未来，智能导盲设备应进一步改善系统便携性及舒适度，选用更轻质的硬件材料；改善硬件电路布置，减小系统体积和重量；改进握柄人体工学设计，增添防滑、减震性能，不断提高视障人士的使用舒适度。

## 参考文献

- [1] 夏娜, 冯彩英, 王垒. 基于 STM32 的智能导盲杖设计 [J]. 电脑与电信, 2025, (11): 20-25+33.
- [2] 边志鹏, 丛慧, 李智昆, 等. 基于 ChatGPT 的智能导盲杖设计 [J]. 机械工程与自动化, 2025, 54(05): 107-108.
- [3] 高顺强, 王智文, 赵文志, 等. 基于视觉感知的智能导盲杖系统应用研究 [J]. 物联网技术, 2025, 15(02): 147-150.
- [4] 金佳鑫, 郑皓予, 郭宇凡, 等. 智能导盲杖控制系统设计 [J]. 自动化应用, 2024, 65(19): 75-76+84.
- [5] 腾浩, 武涛, 王德帅, 等. 导盲杖研究现状及分析 [J]. 山东工业技术, 2023, (06): 50-56.
- [6] 王一帆, 王晓凤, 辛凯轩, 等. 一种多功能智能导盲杖的设计思路 [J]. 科技风, 2023, (32): 1-3.
- [7] 李凤桐, 顾雅青, 李聪. 基于 STM32 的多功能新型智能导盲杖设计 [J]. 科技资讯, 2023, 21(13): 88-91.
- [8] 杨锦, 檀玉磊, 徐郑, 等. 基于 STC89C52 单片机的智能导盲杖系统设计 [J]. 科学技术创新, 2023, (13): 87-91.
- [9] 刘庆涛, 俞志强. AI 语音交互可折叠成座椅导盲杖设计——基于 TOF 雷达测距与 GPS 定位 [J]. 信息记录材料, 2023, 24(03): 132-134+138.
- [10] 左炳辉, 范志文, 邱宇. 基于机器视觉的智能导盲杖 [J]. 自动化技术与应用, 2022, 41(03): 150-152+191.