

一种室内导盲机器人的 RFID-语音交互系统设计

Design of the RFID-based Voice Interaction System of Indoor Guide Robot for the Blind

靳祖光 陈超 唐坚

(江苏科技大学机械工程学院,江苏 镇江 212003)

摘要: 以导盲机器人作为研究载体,设计了一种 RFID 和非特定人声识别技术相结合的语音交互系统。采用 125 kHz 低频 RFID 系统将标签作为语音启动的触发信标,并设计其具体布局结构;同时应用 LD3320 中文语音识别芯片,通过非特定人声搜索引擎和特征库计算出识别结果。机器人样机语音交互系统可完成对用户和环境状态信息的语音问答。不同语速、关键词的延长时间和识别率的测试和统计表明,该系统可以满足语音交互的基本需求。

关键词: 导盲机器人 RFID 语音触发标签 LD3320 语音控制 语音交互

中图分类号: TP242+.3 **文献标志码:** A

Abstract: With the guide robot for blind as the research carrier, the voice interaction system based on the combination of RFID and speaker-independent automatic speech recognition (SI-ASR) technology has been designed. By using 125 kHz LF RFID system, and the labels are adopted as the trigger beacon to start up the speech, specific layout structure is designed; in addition, the LD3320 chinese speech recognition chip is applied, through speaker-independent voice searching engine and the feature database, the recognition results are calculated. The robot voice interaction system can complete the recognition function for user command and reply environment status information. Through the tests and statistics under different speech speed, time delay of the keywords and recognition rate, it is indicated that the system satisfies the basic requirement of voice interaction.

Keywords: Guide robot for the blind RFID Voice triggered label LD3320 Voice control Voice interaction

0 引言

导盲机器人属轮式自主移动机器人^[1]的一种,其设计目的是应用语音交互的方式来弥补盲人视觉信息的缺失。至今为止,世界上许多智能导盲设备如 OAS、Laser-Cane 和 PAMM 等,在人机语音交互方面都存在一定的研究空白^[2-3]。据此,本文分别采用 RFID 无线射频识别技术^[4-5]和 SI-ASR 非特定人声识别技术^[6],通过定义室内 RFID 标签来触发导盲机器人的语音交互系统,以完成环境语音播报和用户人声识别。大多数 RFID 和语音相结合的机电仪器,如带有语音的低功耗手持读卡仪^[7]只具备单纯的报音功能。本文设计的系统除具有报音功能外,还设计有人声指令的识别功能,通过定义 125 kHz 低频 RFID^[5]触发标签的布置结构和应用 LD3320 的 SI-ASR 搜索引擎和特征库,完成非人声指令、播报地点和物品信息的语音交互功能。

1 RFID-语音系统硬件设计

1.1 系统硬件总体方案

室内导盲机器人的 RFID-语音交互部分的硬件系统结构框图如图 1 所示。该系统采用主从机的分布式组成结构,各从机采用 UART 串行通信的方式,通过 MAX485 芯片将 TXD 和 RXD 中的 TTL 电平信号转换为 RS-485 信号,从而实现主控上位机和各从机间的双向信号传递。

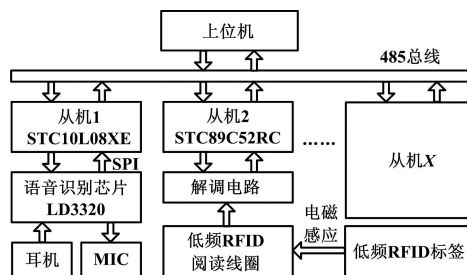


图 1 系统硬件框图

Fig. 1 Block diagram of the hardware

LD3320 语音识别芯片受控于 STC10L08XE,并采用模拟 SPI 的方式进行双向通信。耳机和麦克风接口同时集成在 LD3320 外围电路中。125 kHz 低频 RFID

江苏省高校自然科学基金资助项目(编号:09KJD46003);

江苏省研究生创新计划基金资助项目(编号:CXLX12-0687)。

修改稿收到日期:2013-05-11。

第一作者靳祖光(1988-),男,现为江苏科技大学机械电子工程专业在读硕士研究生;主要从事智能自主移动机器人的相关开发和研究工作。

阅读线圈共有四组,可分别与各自读取范围内的 RFID 标签进行非接触式通信。信号解调电路可解析标签对应的 ID 号码,并传送至从机 2 (STC89C52RC) 中与已存储的物品信息列表代码进行比较和选取,最后得到该 ID 所对应的语音触发信息。上位机负责采集从机 2 中的核对结果并进行优先级调配,最终与从机 1 完成交互工作。

1.2 LD3320 应用电路方案

导盲机器人采用了基于非特定人语音识别技术 SI-ASR 的语音识别芯片 LD3320^[6-8]。该芯片可识别中文发音,不需用户训练和录音,通过芯片内部固化的非特定人语音识别搜索引擎和特征库即可识别出最佳语音匹配的拼音串。LD3320 同时集成了 ADC、DAC、话筒和麦克接口等外围电路^[9],可辅助导盲机器人实

现语音识别、声控和人机对话功能。

LD3320 语音识别芯片接线原理图如图 2 所示。LD3320 的芯片接口结构较为复杂,LD3320 与 STC10L08XE 采用同一个 22.1184 MHz 的外部晶振, SPOP 和 SPON 为扬声器音频输出接口, MICP 和 MICN 为外部麦克风连接口, MBS 为麦克风偏置端,外部连接有 RC 电路,以保证可以提供一个浮动电压至麦克风^[10]。LD3320 与 STC10L08XE 通过 SPI 总线引脚 SDI、SDO、SDCK 和 CSB/SCS 进行双向通信。LD3320 可读取外置 Flash 内的 MP3 文件进行组合和播放,当 MP3 数据或识别结果不充分时,LD3320 通过 INTB 端触发中断信号通知 STC10L08XE 进行处理。在数据出现极端冗余的情况下,STC10L08XE 可直接将 RSTB 端置低,通知 LD3320 进行语音识别和输出功能的复位。

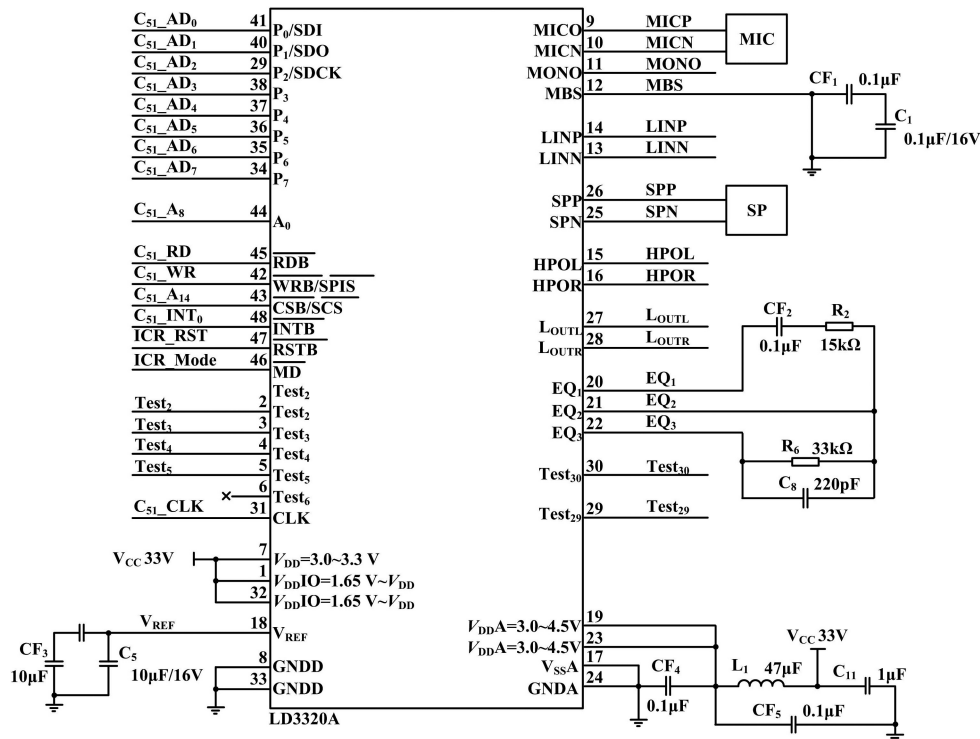


图 2 LD3320 语音识别芯片接线原理图

Fig. 2 Wiring principle of LD3320 voice recognition chip

2 语音触发标签布局设计

低频 RFID 标签具有无源、信号穿透力强、价格低和通信距离近等使用特性^[4],可以在室内环境中实现批量布置。由于 125 kHz 低频 RFID 的通信距离通常在 3 ~ 10 cm 之间,且低频阅读器和标签采用一对一的通信模式,当 1 个读写器处于 2 个标签的读写范围内时会引发读取紊乱的状况,因此低频标签必须以一定布局成组定义,才可确保机器人正确读取。

低频 RFID 阅读器和标签布局图如图 3 所示,左图为导盲机器人俯视图,阅读器分布在机器人的车体底盘上;右图为阅读器和标签重合的几种形式,标签成组设置在环境关键点处,并以边数最少的正多边形和等边三角形作为阵列的基本单位。

设标签的通信区域直径为 D ,相邻标签间距为 s ;阅读器线圈直径为 d ,阅读器线圈共 4 组 ($A_0 \sim A_3$),其中, A_0 位于底盘中心, $A_1 \sim A_3$ 分布在边长为 L 的等边三角形端点处。

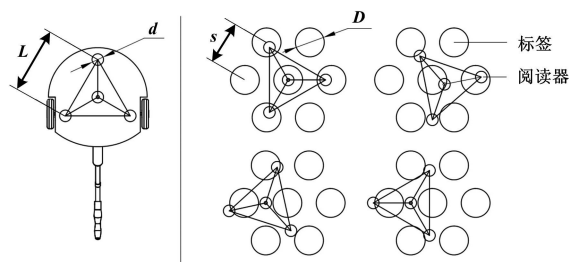


图3 低频 RFID 阅读器和标签布局图

Fig.3 Layout of LF RFID readers and lables

当机器人位于关键点处时,RFID 标签正确触发其语音系统的两个充分条件为:

① 当 A_0 不处于低频标签的信号覆盖范围内时, $A_1 \sim A_3$ 则必须同时处于 3 个不同的低频标签的信号覆盖范围内;

② 当 $A_1 \sim A_3$ 中至少有一个不处于低频标签的信号覆盖范围内时, A_0 则必须处于唯一低频标签的信号覆盖范围内。

$A_1 \sim A_3$ 之间任意两点的直线约束数学关系为:

$$\begin{cases} D+2s < L+d < 2D+s \\ d < s < D \end{cases}$$

$A_0 \sim A_3$ 组成的等边三角形与标签间的约束数学关系为:

$$\begin{cases} (d+D)/2 > \sqrt{(L \tan 30^\circ / 2)^2 + [(D+s) \sin 60^\circ - L/2]^2} \\ (d+D)/2 > \sqrt{[L \sin 60^\circ - (D+s)]^2 + [(D+s) \sin 60^\circ - L/2]^2} \\ (d+D)/2 > L \sin 60^\circ - L \tan 30^\circ / 2 - (s+D)/2 \end{cases}$$

当机器人的 4 组 RFID 阅读器和标签组的布局满足以上数学关系式时,即可确保机器人进入关键点区域时正确启动语音交互系统。

3 机器人语音交互系统设计

3.1 SI-ASR 原理和程序流程

非特定语音识别和输出原理框图如图 4 所示。

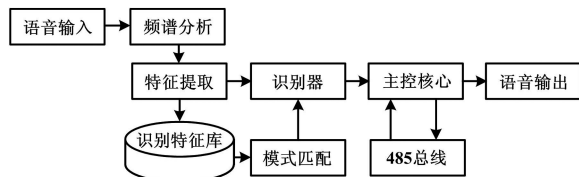


图4 语音识别及输出原理框图

Fig.4 Schematic diagram of speech recognition and output

当 LD3320 识别外部语音信号后,以拼音串的形式传入 STC10L08XE 中,由内部程序再进行拼音串和关键词的匹配,确定匹配后根据关键词属性确定是进行语音输出还是上传动作信号。

STC10L08XE 内部程序流程如下。

① 语音识别和通用初始化。对各寄存器进行初始化设置,单片机通过 RSTB 发送低电平以复位 LD3320 内各数据处理模块,并重新设定模式、始终频率和 FIFO^[8-10]。

② 将各关键词写入识别列表。关键词包括机器人的运动控制词语、目的地和目的物品词语以及简单的询问词语,每个关键词都有其所属编号,不同关键词可享有同一编号,即使用者所发出的相近语音指令都可以归属于导盲机器人同一个执行结果。关键词采用以空格相间的拼音串形式设定,在导盲机器人启动前已写入 STC10L08XE 的 EEPROM 中。

③ 准备识别外部语音。向寄存器写入麦克风音量 ADC 增益,设置寄存器清空和芯片状态,开启 ADC 和中断并准备识别外部语音。

④ 响应中断并识别。当麦克风采集到增益范围所能接收的音频信号时,LD3320 通知 STC10L08XE 产生中断并识别。在 LD3320 进行识别后,往 BA 寄存器中写入所有可能匹配结果,往 C₅ 寄存器中写入最佳分析结果。在读取寄存器结果时,若 C₅ 中没有出现所属类的结果,而 BA 寄存器中存在该结果,则单片机自动将程序跳转至语音输出动作。申请使用者重复语音指令并重复识别,若重复识别次数达到 3 次仍不匹配,则跳出识别。

⑤ 关键词匹配和动作输出。若外部语音识别结果与关键词匹配成功,则单片机 STC10L08XE 关闭识别中断,以防止多余语音信息构成中断干扰,并准备执行机器人的具体执行动作。导盲机器人的语音执行动作按形式分为两类。一类是语音输出形式,机器人应要求为使用者提供语音播报,如当前运行状况、大致位置信息以及所经过的可识别物品等。导盲机器人的 MP3 语音输出信息存放在与 STC10L08XE 相连的外置 Flash 寄存器中。为节省空间并构成更多的输出信息,这里将 MP3 语音输出信息分段存储,播放时按要求寻找所需语段,这些语段的存放首地址和末地址按顺序组合,通过扬声器一同输出。这种方法使得语音系统的输出内容变得更加智能、灵活和多变。另一类是机器人运动的输出形式,包括机器人的前进、后退、左转和右转等运动信息。当这类信息进行匹配后,单片机将运动控制字上传至 485 总线,与其他模块信息的优先级分配后,最终转化为机器人电机的具体动作。

3.2 语音交互对话单元管理

导盲机器人的语音识别状态判定程序位于导盲机器人从机 1 的主程序中,每个扫描周期都会判断当前

的识别状态,当有外部语音的中断且识别出结果时,程序执行 LD_AS_R_FOUNDOK。在 Output (nAsrRes)子程序中,单片机主要负责对识别结果的匹配和具体输出动作。为便于对应识别结果和行为的处理,本文定义了使用者和导盲机器人对话管理单元的二维数组。由于 LD3320 最多可设置 50 项候选识别关键词,每项拼音段长度不能超过 79 B^[10],因此结合该芯片特性和导盲机器人功能要求设计的对话管理数组如表 1 所示。

表 1 语音单元管理数组设计

Tab. 1 Design of dialogue unit management array

数组名称	数组定义	内容分类
识别数组	INT8U Keyword[50][79];	运动控制语句; 目的地和物品词语; 简单询问词语
动作数组	INT8U Action[50][10];	机器人运动控制; 语音状态回复; 语音询问

表 1 中,识别数组和动作数组分别存在 3 种分类属性,在识别和匹配时可以方便程序更加快速地对和寻找。动作数组和识别数组的同项目之间为一一对应的关系。为丰富机器人的语音输出信息,本文设计的动作数组中一个关键词最多可对应 10 个动作组合,运动动作和语音动作可列为同一动作数组中,以便机器人完成更复杂的任务。

4 试验测试和结果分析

本文根据研究内容自主设计了基于 RFID 的语音导盲机器人试验样机,并分别进行了 10 组功能试验测试。每组试验中,机器人进入低频信标组的角度均不同,其中机器人车体直径为 50 cm,行走速度设定为 0.5 m/s,在机器人进入信标区域的前方设定有超声波探测头。当超声波探测头探测到机器人经过时,开始计时直至识别 RFID 并触发语音交互系统。试验测得每组角度均能正常触发机器人的语音交互系统。由于需满足一定的条件,系统才可正确启动,因此每组试验的延时时间存在差异。试验中最长延时时间为 1 142 ms,最短延时时间为 272 ms,平均延时时间约为 516 ms,基本符合机器人的功能制定要求。

在 RFID 触发 LD3320 工作后,对不同语速和不同语音内容的识别率进行测试,其中每组试验中共有 1 项供识别测试的语音内容、2 种不同的语速,其测试统计结果如表 2 所示。

表 2 语音试验识别结果统计

Tab. 2 Statistics of results of experimental speech recognition

试验分组 (进入角度不同)	触发延时 /ms	语音 内容	普通语速 识别率	加快语速 识别率
第 1 组	783	前进	39/50	26/50
第 2 组	375	向前走	43/50	35/50
第 3 组	1 142	左转	36/50	28/50
第 4 组	417	向左转	42/50	38/50
第 5 组	509	右转	41/50	32/50
第 6 组	302	向右转	45/50	39/50
第 7 组	381	好的	43/50	38/50
第 8 组	272	需要	40/50	32/50
第 9 组	442	休息一会	39/50	27/50
第 10 组	532	到了吗	40/50	29/50

5 结束语

本文设计了一种 RFID 和非特定人声语音识别技术相结合的导盲机器人的语音交互系统。以特定布局的低频 RFID 标签组作为语音交互系统触发信标,触发机器人启动语音识别和输出系统,并着重设计了 LD3320 芯片的语音识别和输出程序流程,以及导盲机器人的对话管理数组及分类,最后通过研发导盲机器人样机试验验证了语音交互系统的可执行性。同时总结了关于提高系统识别率的一些方法。本语音交互系统使用方便,可扩展性强,可扩展至更多的有需求的智能化控制平台中使用。

参考文献

- [1] Roland S, Illah R N. 自主移动机器人导论[M]. 李人厚,译. 西安:西安交通大学出版社,2006:5-23.
- [2] Yuan D, Santa C, Manduchi R. Dynamic environment exploration using a virtual white cane[J]. Computer Vision and Pattern Recognition, 2005(1):243-249.
- [3] Spenko M, Yu H, Dubowsky S. Robotic personal aids for mobility and monitoring for the elderly[J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2006, 14(3):344-350.
- [4] Want R. An introduction to RFID technology[J]. IEEE Pervasive Computing, 2006, 5(1):25-33.
- [5] 游战清,刘克胜,吴翔,等. 无线射频(RFID)与条码技术[M]. 北京:机械工业出版社,2007:3-15.
- [6] 陈斌,郭大勇,施克仁. 用于拟任机器人的嵌入式语音交互系统研究[J]. 机器人, 2003, 25(5):452-455.
- [7] 赖于树,李迅波,熊燕. 一种低功耗多功能手持射频读卡仪设计[J]. 自动化仪表, 2006, 27(12):13-15.
- [8] 陈喜春. 基于 LD3320 语音识别专用芯片实现的语音控制[J]. 电子技术, 2011, 38(11):20-21.
- [9] ICRoutE INC. LD332X 开发手册[DB/OL]. [2011-10-13]. http://www.icroute.com/web_cn/LD332X_Dev.html.
- [10] ICRoutE INC. LD3320 芯片简明调试步骤[DB/OL]. [2011-10-13]. http://www.icroute.com/web_cn/LD332X_Debug-Process.html.