

基于语音交互功能的医疗服务机器人控制系统设计^①

毛丽民^{②*} 卢振利^{***} 谢新明^{*} 浦宇欢^{*}

(^{*}常熟理工学院电气与自动化工程学院 常熟 215500)

(^{**}阿威罗大学电子和信息通信工程研究所 阿威罗 3810-193 葡萄牙)

摘要 根据老年人、残疾人的实际需求,设计了医疗服务型机器人移动本体和 6 自由度机械手臂,搭建了以简化隐式马科夫模型为基础的自动语音辨识控制系统。该控制系统通过人机交互系统的语音命令与键盘指令,可控制机器人运行到指定位置,控制机械手抓取物体,实现机器人的避障功能和座椅、病床两模式的自主切换。针对典型医疗服务需求,对所设计的医疗服务机器人进行了实验,实验结果表明,该机器人在人机交互指令控制下能很好地实现到达指定位置、抓取物体、避障和病床/座椅模式切换,不仅易于操作,而且性能稳定,语音系统对命令的识别率达 90% 以上,从而验证了该医疗服务型机器人控制系统设计的可行性和有效性。

关键词 自动语音辨识(ASR), 医疗服务机器人(MSR), 人机交互(HRI), 隐式马尔科夫模型(HMM), 机械手臂

0 引言

服务型机器人在当今社会日益得到广泛应用^[1]。随着人口日趋老龄化,医疗服务机构会不堪重负^[2]。越来越多的老年人和残疾人会选择轮椅来辅助自己的生活^[3]。大量报道指出,对能用中文与病人实现人机交互的服务型医疗机器人的需求会迅猛增加。本研究在广泛调研的基础上,针对中国老年人和残疾人的需求,设计了一种具有中文语音自动辨识和人机交互功能的医疗服务型机器人,操作者可以用语音命令和键盘指令等控制该机器人实现机械手辅助取物,座椅/病床模式切换等多种功能。

1 机器人设计

医疗服务型机器人除了应该具备通用智能轮椅的功能外,还必须能够辅助挂注射瓶、辅助病人取药物和其他物品以及在座椅和床之间切换的功能^[4]。为此本设计应用机构综合技术,设计了一个带有差动轮驱动的移动机构,该机构可以通过控制系统实现座椅和病床模式之间的自动转换。在移动机构上

安装了一个 6 自由度机械手,帮助病人抓取物体。设计了应用单片机融合语音等传感器信息的多功能人机交互系统,使病人可以通过语音和键盘进行操作,实现设定的服务功能。

1.1 机构设计

根据上述服务需求特性,设计了机器人各个功能模块,如图 1 所示。图中 1 为座椅,2 为机械手,3 为轮子,4 为避障装置,5 为电机驱动装置,6 为摄像头,7 为语音装置,8 为挂水架,9 为控制器,10 为电源稳压模块,11 为电源,12、13 为驱动电机。

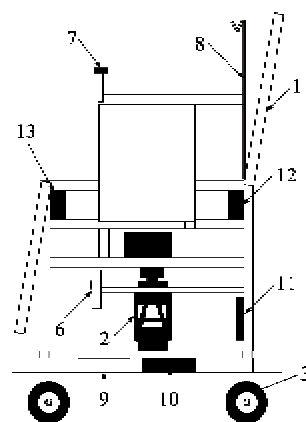


图 1 机器人模型示意图

① 江苏省高校自然科学基金(10KJD510001)和常熟理工学院校级科研(XZ1306)资助项目。

② 男,1981 年生,硕士;研究方向:机器人与目标跟踪研究;联系人,E-mail: maolimin_1981@163.com
(收稿日期:2014-02-24)

按照上述功能,搭建机器人机构如图 2 所示。其包括座椅/病床转换机构、6 自由度机械手、差动轮驱动机构、避障装置、摄像头和挂水架。

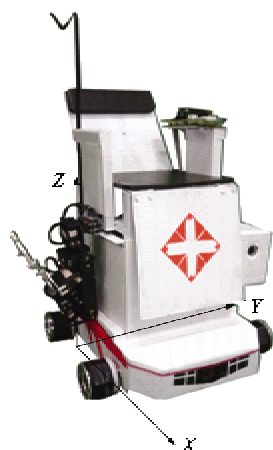


图 2 医疗服务机器人

图 2 中建立的机器人相对运动坐标系统用来控制机器人运动到指定位置。机器人前进方向为 X 轴,右侧为 Y 轴,上下为 Z 轴,机械手下方为原点。机器人根据该相对位置,建立逆运动学实现对末端手抓的位置控制。

该机器人的 6 自由度机械手如图 3 所示。各部分说明如下:1、2 为二指手抓,可以抓取物体;3 为夹持舵机,可以控制手抓的张开与闭合;4 为手腕关节,旋转驱动舵机;5 为手臂连杆;6、7 和 8 为手臂关节,控制舵机。该手臂通过电机 8 座固定在机器人移动体上。

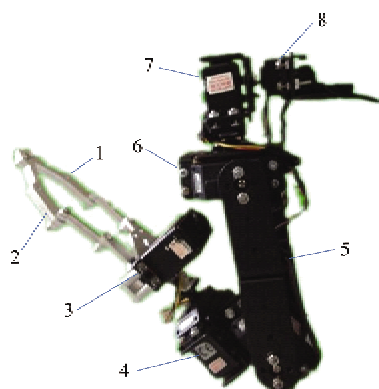


图 3 机械手

1.2 控制系统设计

该医疗服务型机器人控制系统如图 4 所示。病人可通过自动语音辨识模块发送语音命令控制机器人实现启动、前进、转弯、座椅与病床模式切换等功

能。机器人在前进过程中,如果检测到前方有人或者障碍物,机器人就会立刻停下,实现避障功能。病人通过语音也可以控制机械手实现取药、取水杯等操作,该设计的目的是最大程度方便病人。机器人一般情况下为座椅模式,当病人需要躺下时,可以通过语音将座椅模式切换成病床模式,免去移动病人的烦恼。当病人需要休息时,也可使其转换成病床模式。护士可以根据机器人上的摄像头查看现场情况,通过控制界面,更好地实现对病人的远程监护。

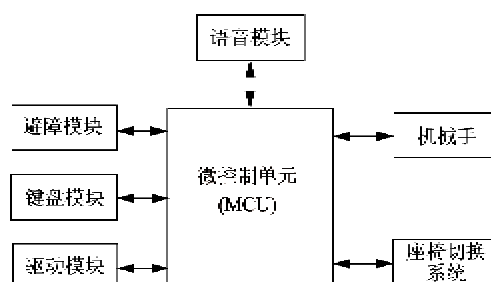


图 4 控制系统

本设计采用 STM32 处理器,基于 ARM Cortex-M3 内核。基本型时钟的频率为 36MHz。本课题应用电脑 CPU:P4 1.6GHz;内存:2G;操作系统:Windows XP;开发环境:Keil4。

2 基于语音交互功能的控制系统设计

目前的语音识别技术主流是基于概率统计模型。国内的研究主要集中于动态时间弯折算法、隐马尔可夫模型和神经网络等识别技术的研究^[5]。

2.1 语音模块的电路设计

LD3320 是一款语音识别芯片,内部集成了语音处理器和麦克风接口、DA 和 AD 转换器等,通过快速而稳定的优化算法,完成非特定人的语音识别。每次识别最多可以设置 50 项候选识别句,支持并行和串行接口^[9-11]。

本机器人语音识别系统采用统一的 3.3V 电压输入,采用的外部晶振为 22.1184MHz。本系统 WRB、RDB、CSB 和 A0 为 4 个控制信号端口实现对芯片的设置和命令,P0-P7 为该芯片的并行接口实现传送数据和接收数据,分别与单片机相接。利用 EP1、EP2、EP3 可以控制喇叭的音量增益。

图 5 为 LD3320 电路原理图,与 MCU 通信采用并行方式。

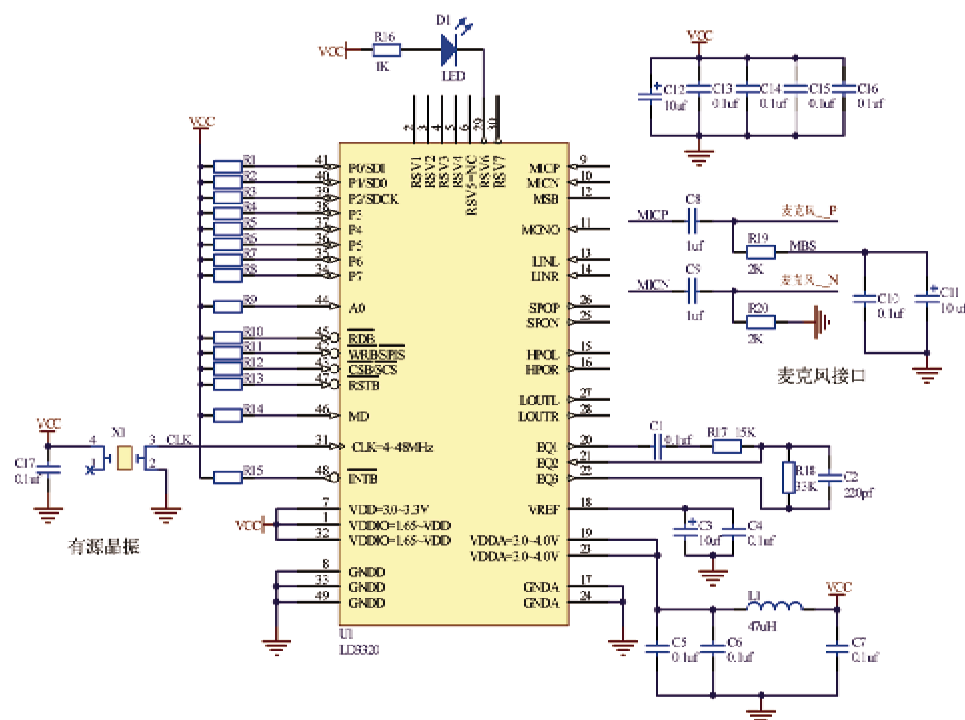


图5 语音芯片接口电路图

2.2 语音模块的软件设计

本设计根据中文发音特点,选择隐马尔可夫模型为基础的非特定语音辨识技术搭建了该机器人的人机交互系统。

非特定人语音识别技术工作原理如图6所示。

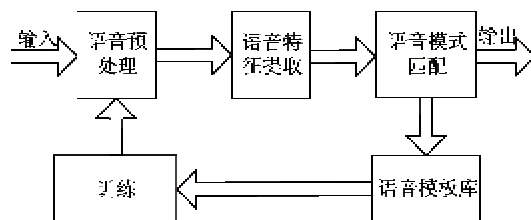


图6 语音识别原理

本系统采用基于隐式马尔可夫模型(hidden Markov model, HMM)的算法(HMM 算法),用隐含的状态对应声学层各相对稳定的发声单位;通过状态转移和状态驻留描述发音的变化;引入了概率计算模型,用概率密度函数计算语言特征参数对 HMM 模型的输出概率,通过搜索最佳状态序列,以最后验证概率为准则找到识别结果^[64]。其中,HMM 的输出概率计算表示如下:

(1) 首先定义 HMM 的前向概率 $\alpha_t(i)$ 为

$$\alpha_i(i) = P(q_1, q_2, \dots, q_i, q_i = i | \lambda) \quad (1)$$

(2) HMM 的前向概率 $\beta_t(i)$ 为

$$\beta_t(i) = P(o_{t-1}, o_{t-2}, \dots, o_t, q_t = i | \lambda) \quad (2)$$

(3) 经过分析,可以得到下面的输出概率计算公式:

$$P(O|\lambda) = \sum_{i=j}^M \alpha_i(i) \beta_i(i) = \sum_{i=1}^N \alpha_T(i),$$

$$1 \leq t \leq T-1 \quad (3)$$

本设计中所用到的模型参数解释如下: N 代表模型的状态数; $A = \{\alpha_{ij}\}$ 代表状态转移矩阵, $\alpha_{ij} = P[q_{i+1} = j | q_i = i], 1 \leq i, j \leq N$; $\pi = \{\pi_i\}$ 代表各状态的起始概率分布, $\pi_i = P[q_1 = i], 1 \leq i \leq N$; $B = \{b_j(o)\}$ 代表输出概率密度函数, $b_j(o) = \sum_{i=1}^M (c_{ji} N(o), \mu_{ji}, U_{ji}), 1 \leq j \leq N$; O 代表观察向量; M 代表每个状态包含的高斯元的个数。

本设计采用 LD3320 语音处理系统的并行通讯方式实现语音识别功能模块与控制系统主控单片机的通讯。机器人语音控制流程框图如图 7 所示。

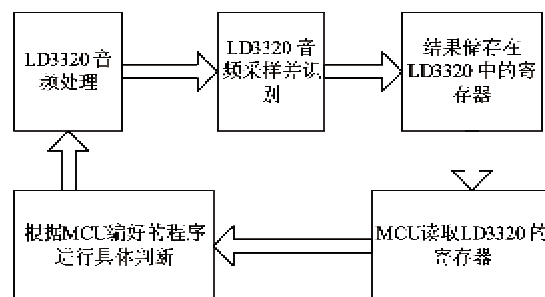


图7 语音控制流程图

语音识别分为 4 个过程,即语音识别的初始化、写入识别列表、开始语音识别及中断响应。在“开始识别”后,读取指定寄存器的值,如果为设定的值就表示有识别结果产生^[12-15]。

(1)通用初始化和语言识别用初始化。在初始化程序里,主要完成软复位、模式设定、时钟频率设定和 FIFO 设定。芯片复位是对芯片的第 47 管脚发送低电平,然后需要对片选 CS 做一次拉低至拉高的操作,以激活内部数据处理模块。

语音识别用初始化函数

{

模式设定

写 35H 对 LD3320 进行软复位

FIFO_EXT 下限低 8 位

FIFO_EXT 下限高 8 位

;

(2)写入识别列表。列表的规则是,每个识别条目对应一个特定的编号(1 个字节),不同的识别条目的编号可以相同,而且不用连续。本芯片最多支持 50 个识别条目,每个识别条目是标准普通话的汉语拼音(小写),每 2 个字(汉语拼音)之间用一个空格间隔。编号字符串见表 1。

表 1 编号字符串

1	qian jin
2	xiang qian
3	hou tui
4	zuo zhuan
5	you zhuan
6	ting zhing

#define CODE_QIANJIN 1 // 定义编号

(3)开始语音识别。设置若干相关的寄存器,即可开始语音识别,其识别流程图如图 8 所示。模数转换(ADC)通道为麦克风输入通道,ADC 增益则为麦克风音量,可设定值 00H-7Fh,建议设置值为 40H-6Fh。值越大表示麦克风音量越大,识别启动越敏感,但可能带来更多的误识别;值越小代表麦克风音量越小,需要近距离说话才能启动识别功能,其好处是对远处的干扰语音没有反应。

if(识别标志位 == 0)

{ 返回 0; 失败. }

if(识别标志位 == 1)

{ 返回 1; 成功. }

(4)识别响应。将麦克风采集到的信号传送给语音识别系统。机器人辨识系统分析结果,对各个候选结果进行评分,并应用特定寄存器存放识别结果,发送给控制器驱动机器人实现给定任务。

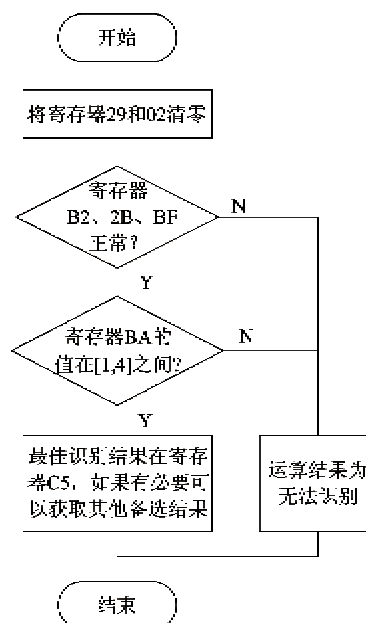


图 8 语音识别系统软件流程图

如果麦克风采集到声音,不管是否识别出正常结果,它都会产生一个中断信号。而中断程序要根据机器人的值分析结果,如图 9 所示。读取 BA 寄存器的值,可以知道有几个候选答案,而 C5 寄存器里的答案是得分最高、最可能的正确答案。

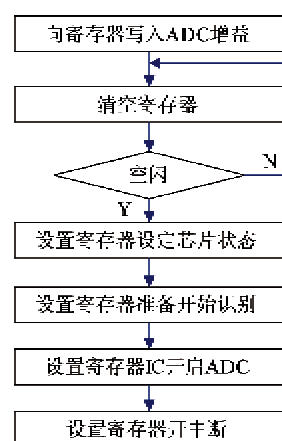


图 9 语音识别系统中断服务程序流程图

3 基于语音识别的人机交互系统训练

根据实际需要,本设计采用的语音辨识控制系统的作用是:(1)分析外部输入非特定时间的语音

数据;(2)通过端点检测技术检测出语音数据是否结束;(3)将有用的语音数据进行运算分析;(4)给出最终的识别结果。

为了检测识别效果,本研究选择多个不同音色的人在实验室环境下进行测试。每个语音命令测试50遍,测试者语音测试的正确识别数据比例如表2所示。

表2 语音命令测试数据表

语音命令	非特定人1	非特定人2	非特定人3
101 房间	47/50	45/50	45/50
104 房间	46/50	46/50	44/50
座椅模式	46/50	45/50	46/50
机械手	47/50	46/50	45/50
前进	46/50	45/50	45/50
左转	48/50	47/50	45/50
停止	46/50	45/50	45/50

由表2可知,在实验室的环境下,对于本设计中的语音系统对命令的识别率可达90%以上。为了检测端点检测技术的准确率,分别在安静的实验室和嘈杂的办公室环境下进行了测试,实验结果如图10所示。

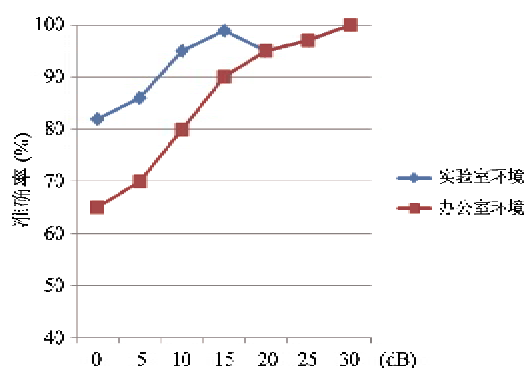


图10 不同环境下的语音辨识

由实验结果可知:(1)声音在20分贝以下,安静实验室环境中的辨识准确率明显高于有噪声的办公室环境;(2)但当声音高于20分贝时,两种环境下的声音辨识准确率基本一致;(3)当声音大于25分贝时,端点检测技术的准确率可以达到95%左右。

4 典型医疗服务功能实验

语音交互系统是根据麦克风提取到的声音信

号,经过隐式马科夫模型为基础的辨识系统得到语音指令,该指令与预定义的机器人动作相联系,进而实现基于语言命令的机器人运动控制。

机器人辅助病人实现取药、取水等操作以及该机器人可以自主实现座椅、病床两个模式的自动切换是医疗服务机器人的重要评价指标。本设计通过语音命令和键盘指令两种人机交互方式对该机器人实现上述功能进行了实验。

4.1 机器人实现基本功能

本实验中通过键盘指令/语音命令实现机器人左转、右转、避障的功能,到达某个位置。机器人左转实验结果如图11所示,右转实验如图12所示。

机器人在前进的过程中当遇到前方有障碍或者有行人时,便停止前进,直到障碍或者行人移开,机

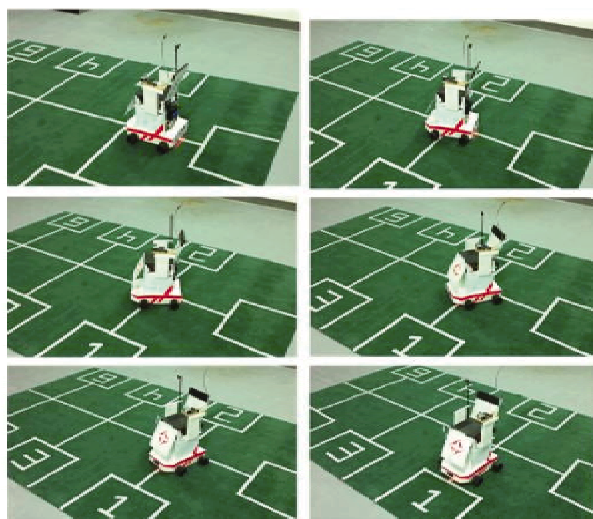


图11 机器人左转

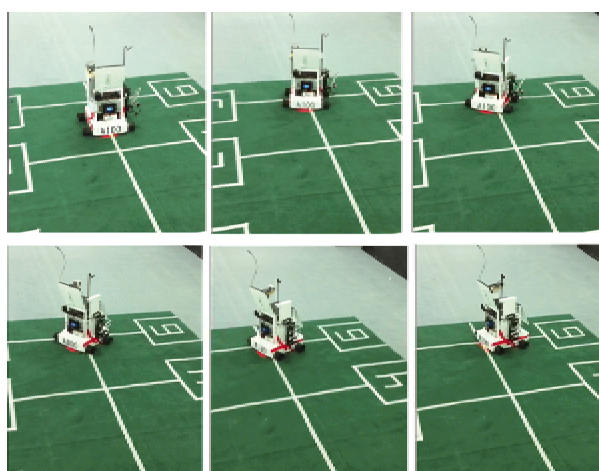


图12 机器人右转

机器人继续前进。实验结果如图 13 所示。图中:(1)为机器人前进;(2)为前方有行人;(3)为机器人停止前进;(4)为行人移开,机器人继续前进。

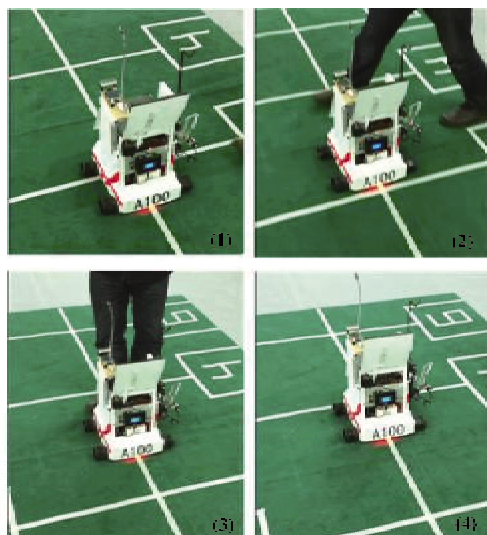


图 13 机器人避障

4.2 机器人辅助取物功能

本实验中操作者通过键盘指令/语音命令实现如下动作:(1)驱动机器人到指定位置;(2)驱动机械手抓张开;(3)靠近要抓取的物体;(4)合上手抓抓住物体;(5)抬起手臂到指定高度;(6)平移物体到病人期望的位置。实验结果如图 14 所示,机器人用其手臂成功完成辅助抓取物体动作。

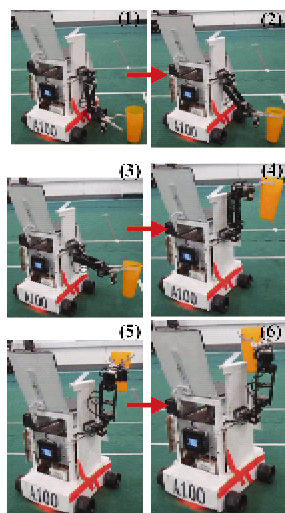


图 14 机器人辅助取物

4.3 座椅/病床模式切换功能

首先定义座椅切换标志位:0 为座椅模式,1 为病床模式。

```
if(切换按键)
{
    if(切换标记位) 病床模式
    else            座椅模式
}
```

本实验中操作者通过键盘指令/语音命令实现如下动作:(1)放下靠背到位置 a,抬起前挡板到位置 a';(2)放下靠背到位置 b,抬起前挡板到位置 b';(3)放下靠背到位置 c,抬起前挡板到位置 c';(4)放下靠背到位置 d,抬起前挡板到位置 d';(5)放下靠背到位置 e,抬起前挡板到位置 e';(6)放下靠背到位置 f,抬起前挡板到位置 f'。实验结果如图 15 所示,机器人成功由座椅模式转换到病床模式。

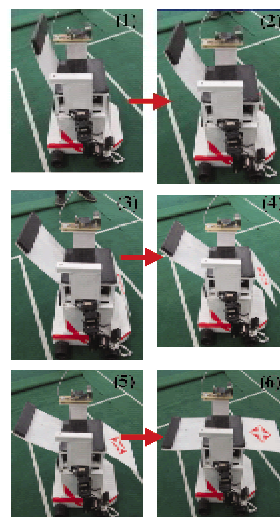


图 15 座椅/病床模式自主切换

5 结 论

依据辅老助残的功能需求,搭建了集带有差动轮驱动的移动载体和 6 自由度机械手于一体的医疗服务型机器人。在该机器人中,装备了摄像头、避障装置等多种传感器设备;设计了基于语音交互功能的机器人运动控制系统,其应用隐式马科夫模型实现非特定人群自动识别中文的语音命令。操作者可通过语音和键盘发送指令实现病床模式与座椅模式的切换以及控制机械手辅助抓取物体等功能。实验验证了该设计的有效性,这种设计可为老年人和残疾人员提供更加便捷的医疗服务型机器人。

参考文献

- [1] Beer J M, Smarr C A, Chen T L, et al. The domesticated robot: design guidelines for assisting older adults to age in

- place. In: Proceedings of HRI 12 Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction, 2012. 335-342
- [2] Kinsella K, Phillips D R. Global aging: the challenge of success. *Population Bulletin*, 2005, 60(1): 5-39
- [3] Krishnan R H, Pugazhenth S. Mobility assistive devices and self-transfer robotic systems for elderly: a review. *Intelligent Service Robotics*, 2014, 7(1): 37-49
- [4] 张西正, 侍才洪, 李瑞欣等. 医疗机器人的研究与进展. *中国医学装备*, 2009, 6(1): 7-12
- [5] 吕涛, 刘百芬, 燕贤青. 一种基于定点 DSP 的语音识别算法实现. *华东交通大学学报*, 2008, (6): 68-73
- [6] Moscovich, Luis G. Learning discrete hidden Markov models from state distribution vectors. selected topics in Louisiana State University and Agricultural&Mechanical College, 2005, 32-45
- [7] 胡钢. 汉语孤立词语音识别算法分析与研究. [硕士学位论文]. 辽宁:鞍山科技大学, 2003. 16-59
- [8] 张梅. 隐 Markov 模型的基本原理及其在语音识别中的应用. *山西科技大学学报*, 2003, 21(4): 2-9
- [9] Dailliel Jurafsky & James H. Martin. *Speech and language processing: an introduction to natural language processing. Computational Linguistics, and Speech Recognition* (2ed). Prentice-Hall, 2006. 38-42
- [10] 陈玉平, 韩纪庆, 郑铁然. 基于动态排位信息的语音关键词确认方法. *计算机工程*, 2008, 34(10): 44-45
- [11] 李如玮, 鲍长春, 窦慧晶. 基于小波变换的语音增强算法综述. *数据采集与处理*, 2009, 5: 24-31
- [12] 张从力, 雷蕾, 段其昌等. 可用于矿井紧急通讯的无线语音系统研究. *国外电子测量技术*, 2010, 29(9): 43-46
- [13] 曾定, 刘加. 基于模型融合的母语与非母语语音识别. *电子测量技术*, 2009, 32(6): 81-83
- [14] 余皓, 苏全. 语音控制机器人的设计与实现. *机器人技术*, 2007, 29(5): 29-31
- [15] 高美娟, 杨智鑫, 田景文. 移动机器人实时语音控制的实现. *电子测量技术*, 2011, 34(7): 50-53

Design of a speech interaction based control system for medical service robots

Mao Limin^{*}, Lu Zhenli^{**}, Xie Xinming^{*}, Pu Yuhuan^{*}

(^{*} Changshu Institute of Technology, Changshu 215500)

(^{**} IEETA, University of Aveiro, Aveiro, 3810-193, Portugal)

Abstract

To meet the real needs of the aged and the disabled, a medical service robot with a moving base and a 6DOF robotic arm was developed, and its control system using automatic speech recognition was designed based on an implicit hidden Markov model (HMM). The control system can control the robot to realize different functions such as moving to a given position, manipulating an object with the robotic arm, avoiding an obstacle, and performing the seat-bed switch through its human robot interaction (HRI) unit's voice instructions and keyboard commands. According to the typical requirements of the robot in the medical service, the experiments on the above mentioned functions were carried out, and the results showed that under the control of the HRI commands, the medical service robot demonstrated its advantages of high stability and easy manipulation in performing the functions, and the rate of voice command recognition was up to 90%, thus the feasibility and effectiveness of the control system design were verified.

Key words: automatic speech recognition (ASR), medical service robot (MSR), human robot interaction (HRI), hidden Markov model (HMM), robotic arm