

基于语音交互的脑瘫康复训练系统的设计^①

卢振利^② * ** 沈玄霖 *** 刘 军^③ **** 张国栋 ***** 孙慧玲 * 波罗瓦茨 布朗尼斯拉夫 ***** 李 斌 **

(* 常熟理工学院电气与自动化工程学院 常熟 215500)

(** 中国科学院沈阳自动化研究所机器人学国家重点实验室 沈阳 110014)

(*** 常熟市第二人民医院 康复医学科 常熟 215500)

(**** 浙江大学生物医学工程与仪器科学学院 杭州 310027)

(***** 大连国通电气有限公司 大连 116650)

(***** 诺维萨德大学技术科学学院 诺维萨德 21000, 塞尔维亚)

摘 要 研究了使用人机交互的方式对脑瘫患儿进行康复训练的技术,设计了基于语音交互的脑瘫康复训练系统。该系统采用 51 单片机驱动整个系统,芯片采用 LD3320 语音芯片,下位机程序采用 C 语言编程,上位机采用 LabVIEW 编程,通过串口传送的方式,把下位机接收到的命令传送到上位机,患儿通过识别上位机界面的图片,对应说出所要识别的文字,语音芯片进行识别,评价系统自动评判出患儿说的正确与否,并记录患儿每次识别所用的时间。案例实验验证了本设计的有效性。

关键词 51 单片机, LD3320, 脑瘫康复训练, 人机交互, LabVIEW

0 引 言

近年来,脑瘫康复治疗技术飞跃发展,在应用传统治疗技术的基础上,新技术不断涌现^[1]。语言训练即是脑瘫患儿的康复训练的重要项目,语音交互模式技术的出现无疑拓宽了我们的视野。即语言训练是脑瘫患儿的康复训练的重要项目,语音交互模式技术的出现很大程度上拓宽了我们的视野。

小儿脑瘫主要为运动障碍,同时也伴有姿势、语言、听觉、视觉、智力等多种障碍,是由先天性、非进行性、非感染性的大脑损伤或发育异常而导致的躯体各类残疾。从调查结果看,这 50 多年来,产科技术、围产医学、新生儿医学技术上的飞速发展,使得新生儿的死亡率、死胎率都有了明显下降,但脑瘫的

发病率却没有下降趋势,反而重症脑瘫的比例逐渐增大^[2]。全球范围内,脑瘫发病率约为 1‰ ~ 5‰,目前仍呈逐年递增趋势。因此,有效治疗脑瘫以及开展康复训练刻不容缓^[3]。

20 世纪 90 年代,人工神经网络技术得到应用,并成为一种新的方式的语音识别。该技术具有稳健性、适应性、非线性、容错性、并行性和学习特点,结构与算法显示出了巨大的潜力,在参数提取、精制的自适应模型设计和优化系统方面已取得关键进展,语音识别系统逐步从实验室走向实际的语音识别系统^[4]。本研究设计了面向脑瘫康复训练的语言系统,该系统利用 51 单片机最小系统和 LD3320 语音芯片的配合,通过串口传送过来的数据把图片载入到相应的界面中,实现下位机和上位机信息的同步。

① 中国-塞尔维亚政府间科技合作委员会第 3 届例会项目(国科外字[2015]266 号 3-1),机器人学国家重点实验室开放基金(2014-008)和校新引进教师科研启动项目(XZ1306)资助。
② 男,1974 年生,博士,教授;研究方向:机器人智能控制,蛇形机器人,机器人人机交互;E-mail: zhenlilu@ csig. n
③ 通讯作者,E-mail: zhenlilu@ csig. cn
(收稿日期:2016-07-26)

1 下位机系统架构

本课题针对脑瘫康复训练系统,选择 51 单片机 STC12C5A60S2 作为最小的系统控制,以带动整个电路,使用 LD3320 语音识别芯片,通过非特定型语音识别模块的录入和识别,根据内部输入的拼音库,将采集的声音信号转化为库里拼音,与已输入的拼音相比较,当每次按键按下去时,液晶屏显示出所要识别的文字拼音,上位机界面显示相应的图片,患儿对着芯片说“口令”,灯亮,然后对应图片说出相应的名称,即开始识别,最后对比患儿说出的名称和图片是否吻合,评判系统对识别的时间和正确率做出一定评判。语音交互系统的整体设计框图如图 1 所示。

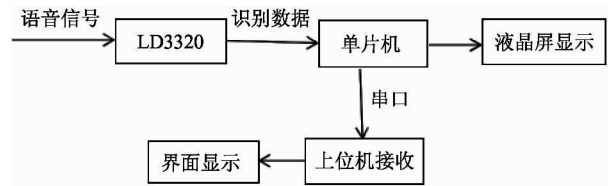


图 1 语音交互系统设计框图

1.1 硬件连接

系统中包括单片机、语音芯片、下载电路、LD3320 串口电路。本文选择 STC12C5A60S2 51 单片机模式来驱动和控制整个电路。语音芯片的型号是 LD3320。单片机是语音模块的核心控制平台,并且它控制着语音芯片去完成对语音信号的采集和对比分析,同时将对对比得出的一个最优结果以串口中断的方式传回给上位机,使上位机窗口显示对应的图片。从而读取患儿识别图片的正确率与时间的快慢。语音识别模块与单片机的通讯通过 VCC/GND/TXD2/RXD2 传送,其中 TXD2 接单片机 P3.1, RXD2 接单片机 P3.0 口,单片机与上位机通讯通过 VCC/GND/RXD/TXD 传送指令,其中 TXD 接单片机 P1.3, RXD 接单片机的 P1.2 口。

1.2 语音识别流程

语音识别的过程,是通过频谱转换成语音词汇,该词汇和条目列表中的关键词进行匹配,最佳匹配词作为识别结果。在本设计中语音识别列表是患儿

对应图片说出小猫、火车、铅笔这三个词。不管列表的条目内容对应的是什么,只需用户设置相关寄存器,就可以把适当的条目与字符识别的形式内容交付识别引擎。语音识别过程如图 2 所示。语音模块流程见图 3。

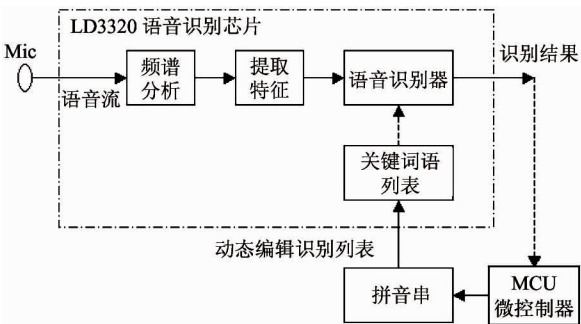


图 2 语音识别芯片识别过程

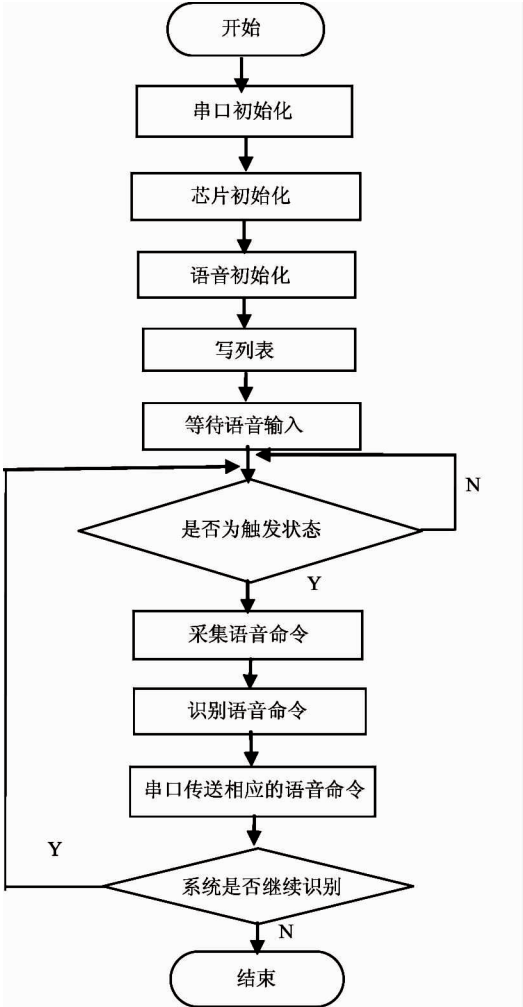


图 3 语音模块流程图

1.3 语音关键字

LD3320 识别支持多达 50 个条目,每个条目可以识别单词、短语的长度不超过 10 个字符或 79 字节拼音字符串。每个条目是一个标准的识别普通话拼音(小写),每个单词(拼音)之间有一个间隔,每一个数字可以相同,但不可以是连续的,但数值要小于 256(00H-FFH)^[5]。当编写一个列表识别时,每个待识别的语句需要一个编号。这个编号是识别出以后,通过串口返回单片机的值。值得注意的是,这个编号是要一字节的。同时列表中不同的待识别语句的编号可以相同,也可以不同。可以连续,也可以不连续。如表 1 所示。

表 1 语音识别列表

编号	字符串
0x01	kou ling
0x02	huo che"
0x03	xiao mao
0x04	qian bi

完成各初始化,将关键词列表写入芯片以后,就可以等待输入语音指令。当有信号输入时,语音芯片会调用识别程序。单片机程序,通常用一个全局变量来控制当前状态,在编程的状态组 LD3320 芯片开始识别前^[6]。

将识别到的结果存放到 C5 寄存器。如果麦克风采集到声音,不管它是否是正常的,都将产生一个中断信号,并根据寄存器分析结果的值中断程序。读取 BA 寄存器值,在 C5 寄存器中的得分最高确定为正确答案。

麦克风收集到语音信号后,运输到 LD3320 内部。与写关键字和匹配列表相比较,以获得最佳的效果。将该结果所对应的序号传递给单片机。此时模块会把该信号通过串口发送给主单片机。语音控制指令所对应的返回值如表 2 所示。

表 2 语音控制指令对应的串口返回值

语音控制指令	对应串口返回值
火车	0x01
小猫	0x02
铅笔	0x03

2 上位机系统架构

本设计中利用该芯片模拟了语音交互模式下的脑瘫康复训练系统。设计过程中,设计师通常需要释放一些短的口令,并在培训的设计环境中,如果使用键盘或鼠标输入口令,这显然与设计的初衷相反,所以使用 LD3320 控制电路和语音识别,将接收识别结果和用于串行接口的信号,输入单片机,实现语音控制的目的。

2.1 开始界面设计

本文应用 LabVIEW 设计一个脑瘫康复训练系统的开始界面,在正式测试之前,有这样一个界面可以提醒患儿马上进入测试阶段,需集中注意力。如图 4 所示。

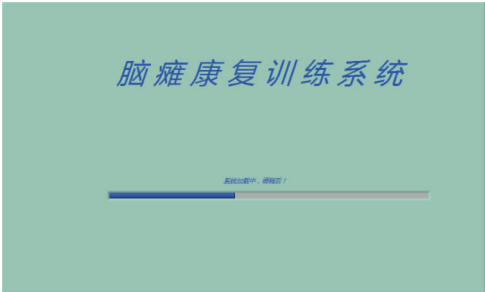


图 4 开始界面

2.2 串口通信系统的实现

LabVIEW 中使用的标准输入/输出接口应用,与 VISA 的仪器完全控制。VISA 可以控制串口、USB 和其他设备,也可以根据使用仪表类型调用相应的驱动程序、独立操作系统、总线和编程环境。也就是说,不管用什么样的设备,操作系统和编程语言,都可以使用相同的应用程序接口^[7]。在开始使用签证之前,应确保选择合适的仪表控制方法。串口数据读取和连接见图 5。

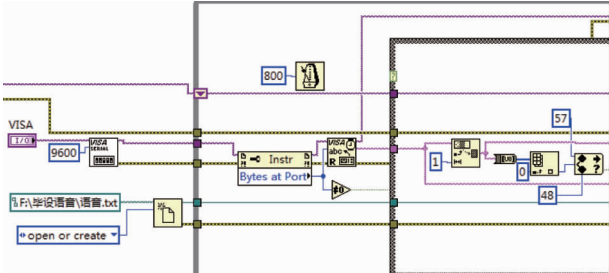


图 5 串口数据读取和连接

2.3 数据传送处理模块

(1) 数据写入二进制文件:收集到的数据存储在 一个数组,然后连接“电子表格字符串数组”功能,这样数组转换为十六进制字符串形式,然后把数据写入二进制文件。为了方便将来的数据查询,可以连接字符串“日期/时间”功能,实时图像识别时间记录,编写一个二进制文件。

(2) 数据保留:使用“打开一个文档来创建另一个”功能,数据可以存储在一个指定的 txt 文件。

(3) 数据清晰:收集到的重复的数据保存到一个集群中,再将波形图属性节点到集群的关联赋值为零,从而实现数据删除。

(4) 退出程序:使用“关闭文件”功能可以关闭打开指定的文件,并返回一个引用原始文件的位置。

2.4 评价系统测试结果

本文用 LabVIEW 设计软件,实现了对于脑瘫患儿的评价系统画面监测,经实践证明各个模块功能稳定,能准确传送和接收下位机上传的数据,实时显示患儿识别的正确率与时间,前面板布局清楚明了,可直观地检测数据,凭借多次的实验测试,可以满足本设计的需求。

在电路系统连接通路之后,计算机和单片机进行数据通信,运行 LabVIEW 通信程序,第一个串口初始化的绿灯开始闪烁,串口通信成功。若患儿识别正确,即上位机界面绿色灯亮,每次识别完显示相应的识别时间。界面如图 6 所示。



图 6 评级系统界面显示

3 系统联调实验

3.1 语音识别模块的联调

系统上电,对着语音芯片说“口令”,如果 led 灯

亮,即芯片测试正确。如图 7 所示。

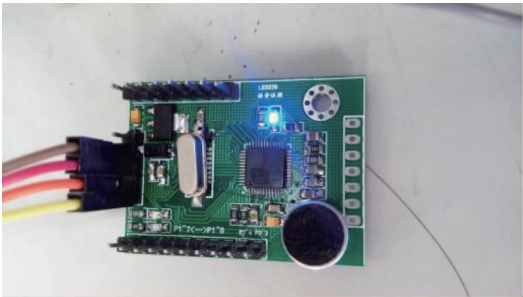


图 7 语音芯片实验

液晶屏显示一组图片拼音,如果下图显示“qian bi”,对着芯片说“口令”,紧接着说“铅笔”,液晶屏显示“qian bi right”,即识别正确。并在第一行显示识别时间。如图 8 所示。

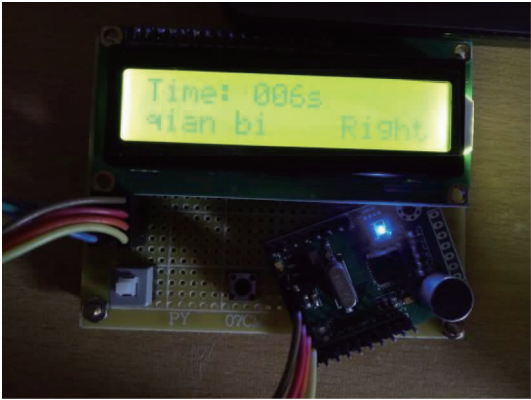


图 8 下位机测试实验一

按下按键,显示屏切换成另一组图片拼音“xiao mao”,重复步骤二,说“口令、小猫”芯片对应识别,正确即显示“right”。若不正确,显示屏即不显示 right。如图 9 所示。

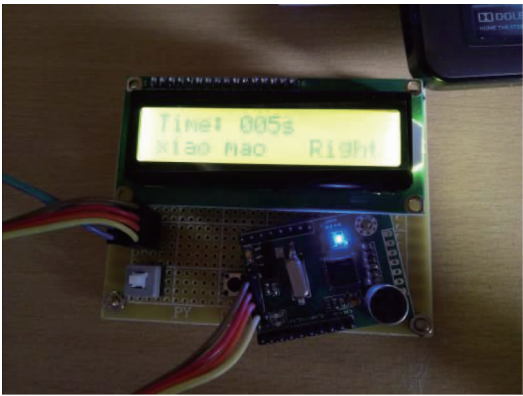


图 9 下位机测试实验二

3.2 上位机模块调试

打开 LabVIEW 软件,通过串口传送过来的数据把图片载入到相应的界面中,实现下位机和上位机的信息的同步。即下位机显示“qian bi”,上位机同时显示铅笔图片。如图 10 所示。

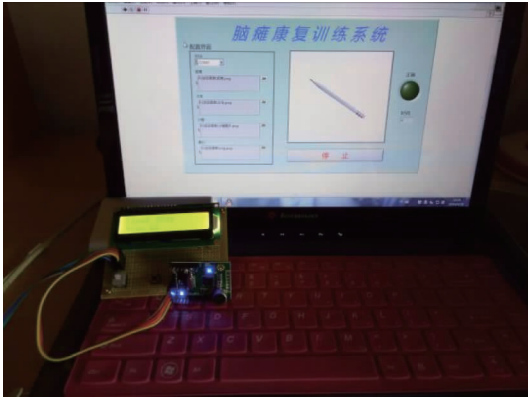


图 10 上位机测试实验一

重复语音模块调试步骤,若下位机液晶屏显示“qian bi”,上位机界面显示 qian bi 图片,对着芯片说出“口令、铅笔”液晶屏显示“right”同时上位机正确按钮灯点亮为绿色。识别结束,图片切换为原始图片。如图 11 所示。

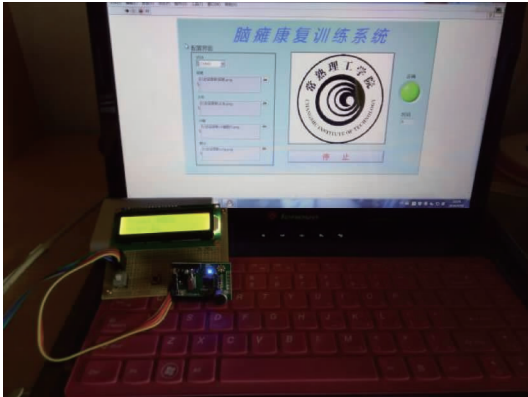


图 11 上位机测试实验二

4 言语障碍训练方法和内容

培训和指导是言语治疗的核心,包括听力的利用率,促进理解和口语,恢复或改善关节功能,提高语音清晰度和其他语言治疗。结果显示,对儿童来说,一个有效的培训计划和集成的系统的训练,可改善各个方面的语言障碍。

4.1 言语训练方法

言语训练主要包括日常生活交流能力训练、吞咽障碍训练、抑制异常姿势、反射训练、构音器官运动训练、构音训练、语言发育迟缓的训练这六大类。本文主要介绍了脑瘫儿童语言发育迟缓的训练。根据每个孩子的语言发育迟缓检查、评估结果和语言特性来拟定训练目标和方法。从测试结果来确认患儿处于哪个水平阶段,把这个阶段定为开始培训的起点,制定相应的培训内容。如果患儿通过学习已掌握了某一阶段的部分内容,接下来就可以学习在这个阶段没有掌握的其他内容,作为学习内容的逐步扩张的基础。如果横向扩张的培训已经完成,患儿达到目标,即进入下一阶段。训练方法包括未学会言语符号儿童的训练、扩大词汇量训练、语句训练、表达训练、文字训练、交流训练等^[8]。

4.2 言语训练内容

通过各种途径治疗仍不能说话,或者能说话但清晰度低的患儿,使用辅助系统进行通信的应用培训。常用的通信系统包括图片或文本通信线路板、通信设备等。设计交流板应该注意适合患儿观看,充分利用剩余的功能,随时调整。

我们采用了人机交互的方式来对患儿进行言语训练,通过将电子芯片、语音芯片与图片相结合的方式,采用 LD3320 语音识别芯片,通过非特定型语音识别模块的录入和识别,根据内部输入的拼音库,将采集的声音信号转化为库里拼音,与已输入的拼音相比较,当每次按键按下去时,液晶屏显示出所要识别的文字拼音,上位机界面显示相应的图片,患儿对着芯片说“口令”,灯亮,然后对应图片说出相应的名称,即开始识别,最后对比患儿说出的名称和图片是否吻合,评判系统对识别的时间和正确率做出一定评判。

5 康复效果评估

语言发育迟缓评定方法是指对语言表达能力有困难 的儿童,采用图片词汇测验,该测验通过听觉和视觉来了解儿童的词汇能力,评估语言智能。

本次实验中,用到的用具如表 2 所示。

表 2 检查用具表

检查用具及图片目录			数量
实物	LD3320 语音芯片		1
	计算机		1
图片	交通工具		1
	动物		1
	文具		1

6 语言发育迟缓案例分析

运行 Labview 程序,启动上位机界面;下位机系统上电,单片机初始化计算机界面播放一组图片,分别为 火车、小猫、铅笔,患儿对准语音芯片,说出每张图片所对应的是什么,评价系统对患儿识别的时间和正确率做一评判记录。所用到的图片如图 12 所示。



图 12 图片目录

案例 1: 患儿 A, 男, 4 岁, 临床诊断: 痉挛型脑瘫。现以“至今 4 周岁语言清晰度差, 两个字词语表达不流畅”为主诉就诊。经(S-S)法检查: 患儿交流态度较好, 注意力集中时间可持续 8s, 理解、表达均处于阶段 4-1 主谓水平, 基础事物概念: 自身五官指认(+)、大小(+)、颜色可识认红、蓝色, 余均不能^[9]。语言诊断: 语言发育迟缓伴构音障碍。

患儿 A: 共测试了 10 次, 每次的检测结果如表 3 所示。

从表 3 中, 根据每次识别图片的时间和正确率的记录, 我们可以画出当患儿开始识别火车这张图片时, 所花费的时间, 以及每次识别正确与否的图表。如图 13 所示。

患儿 A						
次数	火车		小猫		铅笔	
	时间(s)	正确率	时间(s)	正确率	时间(s)	正确率
1	3.1	100%	2.9	0%	2.8	0%
2	2.6	0%	2.7	100%	2.7	0%
3	2.5	100%	2.5	100%	2.9	100%
4	2.8	100%	2.4	0%	2.6	100%
5	2.4	0%	2.2	100%	2.4	100%
6	2.2	100%	1.9	0%	2.3	0%
7	2.5	0%	2	100%	2.5	100%
8	2.3	100%	2	100%	2.4	100%
9	2.3	100%	1.8	100%	2.1	100%
10	2	100%	1.7	100%	1.9	100%

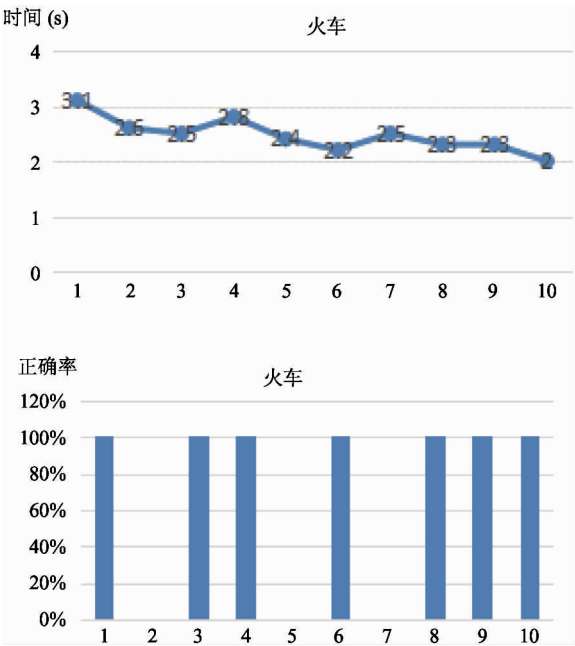


图 13 火车识别时间和正确率

根据上图表, 我们可以很直观地看出患儿 A 在识别火车这张图片时, 随着识别次数的增加, 每次识别的时间在不断地减少, 正确率在 1-5 次识别中波动较大, 因为脑瘫患儿的智力低于常人, 所以并不是速度快, 正确率就高。而是要在多次的训练和测试中, 才能进一步提高正确率, 同时提高康复训练的效率。从第 8-10 次, 基本时间和正确率都逐步稳定下来, 康复训练效果有一定体现。

识别小猫图片和铅笔图片所花的时间及正确率如图 14 和图 15 所示。

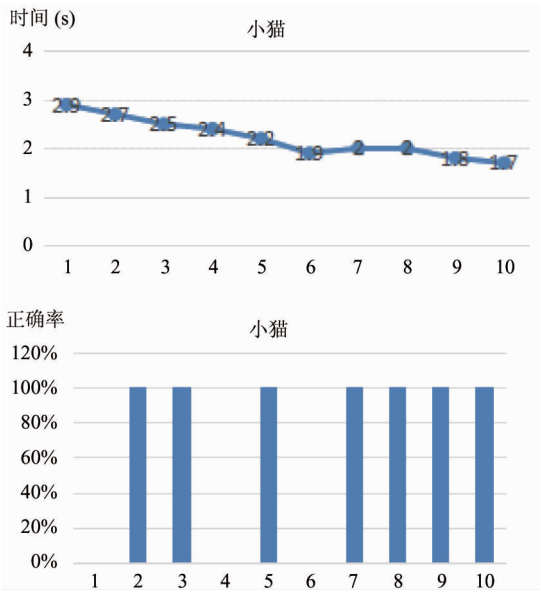


图 14 小猫识别时间和正确率

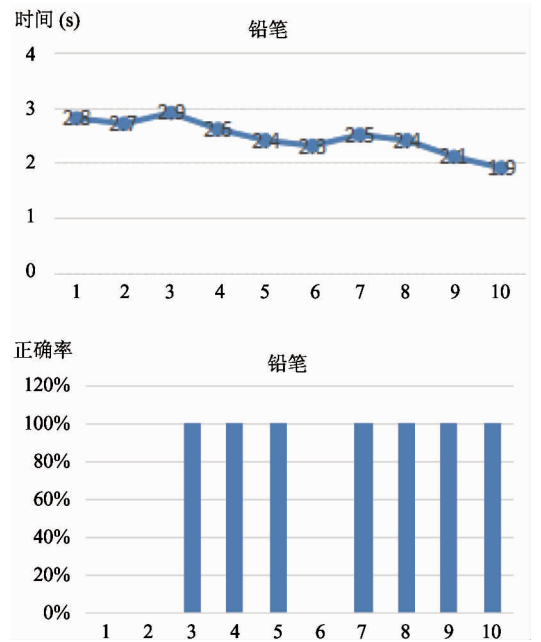


图 15 铅笔识别时间和正确率

从图 14 可看出患儿 A 在识别小猫图片时,在时间上,从第一次识别花费了 2.9s,到最后一次识别只花费了 1.7s,速度有了明显提高,正确率在最后几次也是基本稳定。

从上图可以看出,识别铅笔图片时在第 1 次、第 2 次相对正确率较低,识别时间较长,从第三次往后开始时间和正确率均有明显提高。

综合每次识别的时间和正确率,我们可以得出

患儿 A 在整个测试过程中,正确率和时间变化情况。如表 4 所示。患儿 A 在整个识别过程中的一个时间及正确率变化曲线,如下图 16 所示。从以上患儿 A 的测试结果中,我们可以看出其康复的效果是体现在多次重复测量之后,一开始可能相对识别的时间较长,正确率也不稳定,但随着测量次数的增多,不断地反复练习,每次识别的时间在相应减少,正确率也逐步稳定。也就是说,本设计系统在对脑瘫患儿康复训练过程中,多次重复训练时,有一定效果体现。

表 4 耗时和正确率表

患儿 A					
次数	时间(s)	正确率	次数	时间(s)	正确率
1	8.8	33.3%	6	6.4	33.3%
2	8	33.3%	7	7	66.7%
3	7.9	100.0%	8	6.7	100.0%
4	7.8	66.7%	9	6.2	100.0%
5	7	66.7%	10	5.6	100.0%

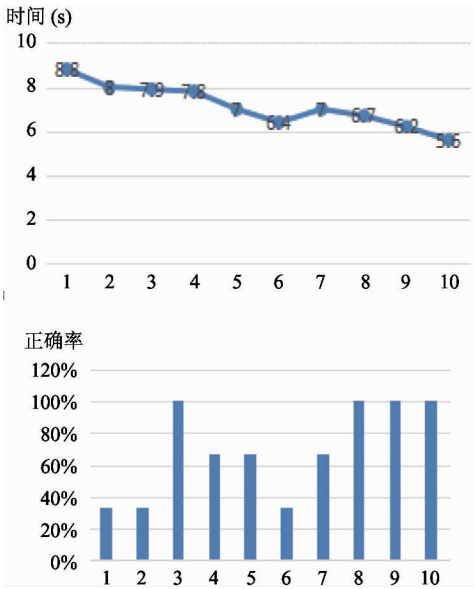


图 16 时间及正确率变化图

7 结 论

本文设计了基于 51 单片机和语言辨识芯片技术的脑瘫语音康复训练系统。采用上下位机架构实现语言与图片识别的同步训练,对脑瘫患儿的各项

数据进行实时记录。以 10 次为一个疗程,分析患儿对语言康复训练的疗效,为临床患儿诊疗提供实时数据依据。实验表明,该设计可以为脑瘫患儿的语音康复训练提供新方法。

参考文献

[1] 李庆瑜,龚远红. 我国脑瘫儿童社区康复护理现状与发展. 护理学报,2006,13(7): 30-31
[2] 史惟,李惠,王素娟等. 脑性瘫痪儿童精细运动功能评估表的心理测量学特性研究. 中华物理医学与康复杂志,2006,28(5): 320-323
[3] 董国超,寇净磊. 多 Agent 在语音识别技术中的应用. 电脑知识与技术,2009,5(13): 3541-3542

[4] 龚伟. 小词汇量孤立词语音识别系统的 DSP 实现. 低压电器,2009,20: 1-4
[5] 徐潼. 多 Agent 系统的体系结构和协作研究:[硕士学位论文]. 南京理工大学计算机系, 2003. 25-36
[6] 张毅刚,彭喜元. 单片机原理与应用设计. 北京:电子工业出版社, 2008. 4
[7] National Instruments Corporation. NI-IMAQTM User Manual, June 2003
[8] 尚鹄,杨慧霞. 脑性瘫痪与产科相关因素研究进展. 实用心脑血管病杂志, 2006,14(8): 598-599
[9] 徐芳. 重度脑瘫儿童语言能力医教结合康复的个案研究. 绥化学院学报,2013,33(4): 135-139

Design of voice interaction-based training system
for cerebral palsy rehabilitation

Lu Zhenli **, Shen Xuanlin **, Liu Jun ****, Zhang Guodong *****, Sun Huiling *,
Borovac Branislav *****, Li Bin **

(* School of Electrical Engineering and Automation, Changshu Institute of Technology, Changshu 215500)

(** State Key Laboratory of Robotics, Shenyang Institute of Automation,
Chinese Academy of Sciences, Shenyang 110014)

(*** Department of Rehabilitation Medicine ,Changshu No.2 People’s Hospital, Changshu 225500)

(**** Faculty of Biomedical Engineering & Instrumentation Science, Zhejiang University, Hangzhou 310027)

(***** Dalian Guo Tong Electric Co. LTD, Dalian 116650)

(***** Faculty of Technical Sciences, University of Novi Sad, Trg Dositeja Obradovica 6, 21000-Novı Sad, Serbia)

Abstract

The technique for rehabilitation training of the children with cerebral palsy using human-computer interaction was studied, and a rehabilitation training system for cerebral palsy based on voice interaction was designed. The system is driven by 51 single chip microcomputers. The whole system adopts the voice chip of LD3320, the lower machine uses C language programming, and the PC uses LabVIEW programming. Through the way of serial transmission, the commands the next bit machine received are sent to PC, and by identifying the picture of the PC interface, children give the corresponding names to identify the text, voice chip, which can identify the final evaluation system automatically to judge the right to say yes or not. And for every time, the identification time was recorded. A case study was carried out to verify the proposed system effectiveness.

Key words: 51 single chip, LD3320, cerebral palsy training, human computer interaction, LabVIEW