

基于表情和语音交互的脑瘫康复训练系统^①

卢振利^{②*} 蒋睿莹^{*} 马志鹏^{***} 波罗瓦茨 布朗尼斯拉夫^{****} 梅尼科 麦然^{****}

(^{*} 常熟理工学院电气与自动化工程学院 常熟 215500)

(^{**} 中国矿业大学信息与控制工程学院 徐州 221116)

(^{***} 诺维萨德大学技术科学学院, 诺维萨德 21000, 塞尔维亚)

(^{****} 马里博尔大学, 马里博尔 462000, 斯洛文尼亚)

摘 要 本文基于人机交互(HCI)设计了表情和语音交互的脑瘫康复训练系统。采用上位机和下位机相结合的方式,下位机采用 51 单片机进行主体驱动,语音采集模块应用 LD3320 语音芯片,并通过串口通讯的方式,将 C 语言编程的下位机与 Labview 编写的上位机相连接。语音芯片把信号辨识结果通过串口发送到上位机,上位机根据语音语法规则进行辨识匹配,对测试完成评判和统计,并将结果与表情机器人相匹配。通过对两位患者正确识别表情的次数以及识别时间进行记录,分析数据,证明该系统有助于提高脑瘫康复训练的效果。

关键词 人机交互(HCI), 语音识别, 脑瘫康复, 表情机器人, 评价分析

0 引 言

在全球范围,脑瘫发病率约为 1‰~5‰,且每年都在持续增长,小儿脑瘫是一种先天性的疾病,大部分的脑瘫患者都伴有语言障碍^[1-3]。对脑瘫患者进行语言训练是脑瘫康复训练中的重要项目^[4,5]。

现实中通常用人工方法对脑瘫患儿进行语言训练,而研究组先期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,其应用 Labview 设计了单个物体及相关名称学习进行的评价系统。除此之外,还研究了针对人动作的识别系统。

在此基础上,本设计采用上位机与下位机相结合的方式,并加入了表情机器人,对康复训练系统进行补充,配合训练患者的语言能力。

1 系统设计

系统采用 STC12C5A60S2 单片机芯片^[6]、LD3320 语音芯片^[7,8]和表情机器人。语音芯片进行语音的采集并将相关信息传到单片机系统中。单片机系统对语音信息进行判断,记录患者识别时间,并配合上位机工作。上位机负责显示患者需要识别的图片,记录患者识别的时间,以及采集患者的相关信息,同时还控制着表情机器人的动作。图 1 为脑瘫康复训练系统框图。

当患者按下下位机按钮,下位机的显示屏上会显示与上位机图片对应的拼音,上位机则显示图片,患者根据图片说出相对应的名称,若识别正确,则下位机显示屏上时间会停止计时,上位机的图片会变成默认图片,并且正确指示灯变亮,同时表情机器人做出相对应的动作。在整个训练结束后,可以通过

① 国家重大科技计划支撑(2014BAK12B01),机器人学国家重点实验室开放基金(2014-008)和中塞政府间科技合作委员会第三届例会项目(266-3-1)资助项目。
② 男,1974 年生,博士;研究方向:机器人智能控制,机器人人机交互,康复机器人;联系人,E-mail: zhenlilu@ csig. cn
(收稿日期:2018-06-04)

前面板的查询按钮,查询患者每次识别时间的变化曲线。本系统不只训练患者语言识别能力,同时依据表情机器人的对应动作,还训练患者识别生活中人的表情,让患者不仅仅能识别 2 维的表情,对 3 维动态表情也可以识别。

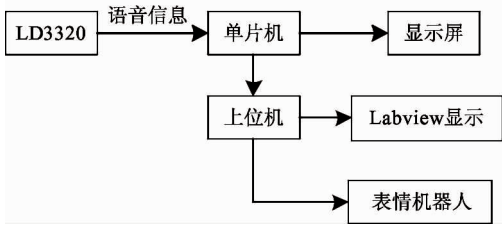


图 1 脑瘫康复训练系统框图

2 下位机系统

2.1 单片机设计

系统下位机主要包括单片机最小系统、显示屏、按键、LD3320 语音芯片和 USB to TTL 串口驱动。单片机是整个下位机系统的核心,它控制语音芯片采集、分析声音,同时将对对比后的最优结果以串口中断的方式传送至上位机,并且控制显示屏的显示信息和识别时间。

在进行下位机编程时,首先进行参数初始化,接着判断按键是否按下,如果按下,则将显示屏上出现拼音所对应的字母传送到上位机,并开始计时。然后单片机检测是否有字母传入,如果有字母传入,则判断传入的字母是否与显示屏上拼音对应的字母相同,如果相同则停止计时,并在显示屏上显示“right”,同时将对应的字母传入上位机。如果不同,则接着检测是否有字母传入单片机。下位机模块的流程图如图 2 所示。

2.2 语音识别设计

LD3320 能够识别大概 50 多个条目。编写 1 个识别列表时,每个待识别的语句都需要对应 1 个编号,这个编号的作用是识别出语音以后,通过串口返回给单片机的值^[9,10]。语音识别列表见表 1。

初始化完成后,便将关键字列表写入芯片中,芯片开始检测有无语音信号的输入,如果有语音信号的输入,则开始进行识别^[11-14]。

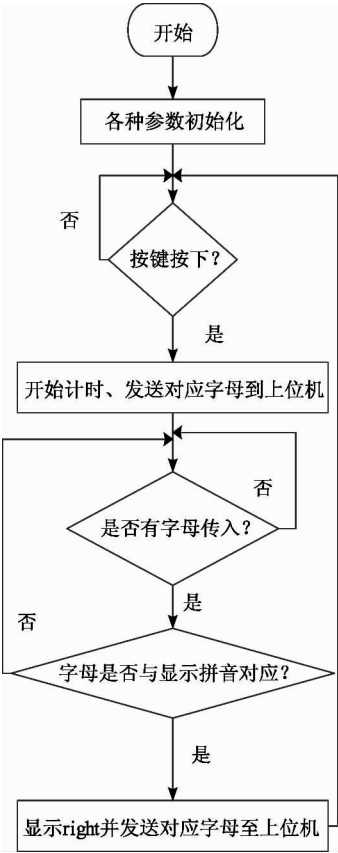


图 2 下位机整体流程图

表 1 语音识别列表

编号	字符串
0x01	weixiao
0x02	beishang
0x03	zhangzui

识别开始后,语音芯片会将识别到的结果放入寄存器内。如果麦克风采集到声音,无论这个声音正确与否,都将产生一个中断信号,并将语音信号传入到 LD3320 芯片中。接着将采集到的语音信号与芯片中的关键字列表进行比对,如果关键字列表中有相匹配的内容,则将该内容对应的返回值传回单片机。语音指令对应的串口返回值如表 2 所示。

表 2 语音指令对应的串口返回值

语音指令	对应串口返回值
微笑	0x01
悲伤	0x02
张嘴	0x03

语音芯片识别到了相关语音,则会通过串口发送对应的字母给单片机。语音指令和对应发送到单片机的字母如表 3 所示。

表 3 语音指令对应的字母

语音指令	对应字母
微笑	w
悲伤	b
张嘴	z

3 上位机系统

系统上位机利用 Labview^[15-17]来编程,可以对脑瘫患儿在训练时的情况进行实时监测,能够准确控制表情机器人的动作;可以准确判断患者识别图片是否正确,能够实时显示患者每次识别图片的时间;能够将患者的姓名、年龄、性别、识别图片信息、识别图片时间都记录在一张表格中,在前面板中能够直观看到患者识别图片每次所用时间的变化曲线。

上位机用到两个串口,一个是下位机传输数据到上位机所要使用的串口,另一个是上位机传输指令给机器人的串口。两个串口都使用 VISA 来控制。上位机读取下位机传输数据的程序如图 3 所示。

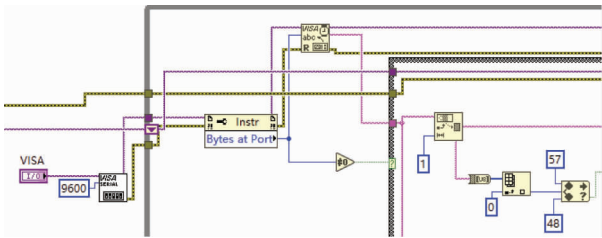


图 3 读取串口数据

3.1 机器人控制模块

控制表情机器人的是 Arduino Uno^[18],运用 Labview 对 Arduino 进行编程。上位机发送数据控制机器人的程序如图 4 所示。

利用 Labview 来控制 Arduino,使得舵机转动,其实是通过控制每个口的脉冲宽度调制 (pulse width modulation,PWM)来实现的,本设计中主要利

用了 3 个口,分别为 3、9 和 10。通过控制 3 个口的 PWM 实现 3 个舵机转动的角度,从而让表情机器人的眉毛和嘴巴进行动作。

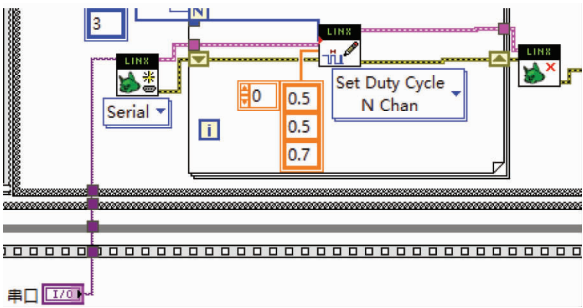


图 4 串口控制机器人

3.2 数据处理模块

患者可以通过前面板的“查询”按钮,查询到每个患者在每个图片上所用的时间,并显示出次数和每个构成的图表。查询的程序框图如图 5 所示。

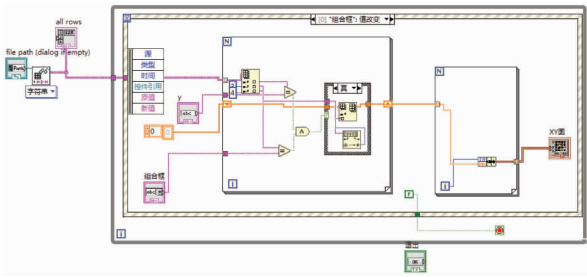


图 5 查询的程序框图

采用一个 while 循环里嵌套一个事件结构的方法,调用表格中的数据,首先筛选出相关用户的信息,接着再逐个筛选出每个患者在每个图片识别中所用的时间,生成一个图表。前面板如图 6 所示。

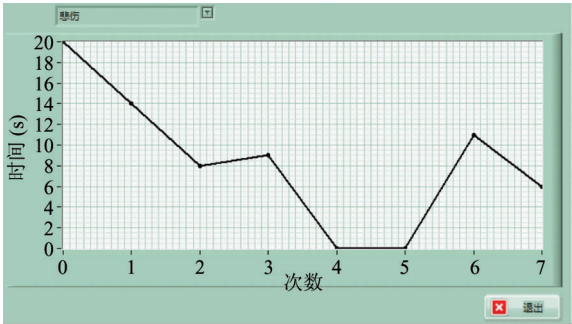


图 6 查询的前面板

3.3 计算时间模块

本设计中,下位机可以计算患者识别图片的时

间。为了便于患者可以直接地看到数据,也为了方便患者数据的统计和记录,上位机也要求可以进行时间的计算并显示在前面板上。

当下位机的按键按下时,下位机会发送一个大写的字母到上位机,当下位机完成语音识别后,下位机会发送一个小写的字母到上位机。上位机是通过计算这大写字母与小写字母间的时间来确定患者识别图片的时间。同时在开始运行系统前,将计算的时间都置零。计算时间的流程如图 7 所示。

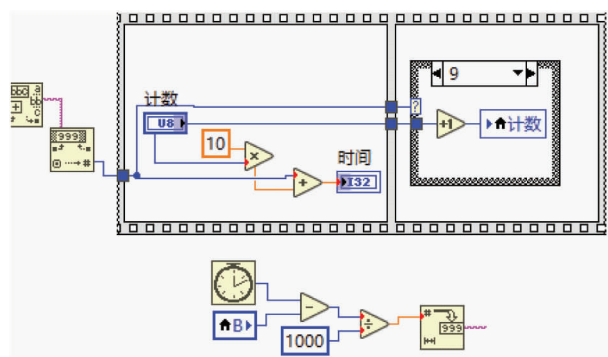


图 7 计算时间流程框图

4 系统调试

4.1 下位机模块

语音芯片连着麦克风、RX、TX、接地端和电源端。将语音芯片的 TX 端接单片机的 RX 端,语音芯片的 RX 端接单片机的 TX 端。电源线接 5 V 供电,接地端接地。

语音芯片正确接到单片机上后,语音芯片上的灯亮,语音芯片可以工作。

显示屏的开始界面,显示屏上显示“Welcome”,表示单片机准备开始工作。如图 8 所示。



图 8 显示屏开始界面

按下按键,显示屏第 2 行出现一组图片拼音,如“bei shang”,显示屏第 1 行显示正确识别语音的时间,未正确识别前,时间显示在不停地滚动。对麦克风说“悲伤”,显示屏第 2 行显示“bei shang right”,则表示识别正确,显示屏第 1 行的时间显示停止滚动。如图 9 所示。



图 9 识别完成后显示屏界面

按下按键,则“bei shang”变成“wei xiao”,重复以上步骤。

4.2 上位机模块

打开 Labview 软件,运行脑瘫康复训练系统,选择正确的串口路径和图片路径。在前面板中将信息完善。训练系统的开始界面如图 10 所示。



图 10 训练系统开始界面

按下下位机按键,下位机发送相应字母到上位机,上位机显示相对应的图片,时间显示开始滚动,患者开始识别图片。如图 11 所示。



图 11 训练系统运行界面

患者正确说出图片名字,下位机正确识别后,发送字母到上位机,训练系统中显示正确的指示灯亮,时间控件中的数字停止滚动。如图 12 所示。



图 12 识别正确后界面

表情机器人做出相应动作后,界面的图片跳转至默认图片。如图 13 所示。



图 13 识别完成后界面

在患者进行脑瘫康复训练以后,患者的相关数据会导入一个表格中,在前面板中,也可以直接查询到患者每次训练的识别时间,直接看出患者经康复训练后,对病情的治疗效果。

点击“查询”按钮,将会跳出一个生成图表的界面。如图 14 所示。

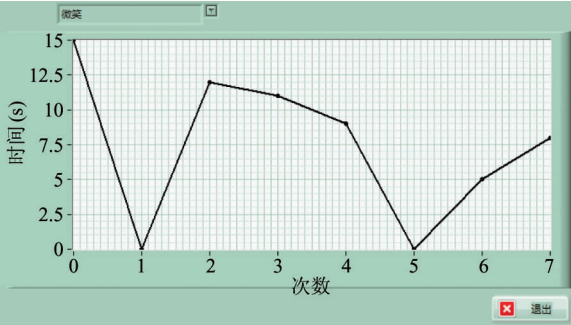


图 14 图表查看界面

4.3 机器人模块

由于单片机只有两个串口,所以表情机器人不能直接与单片机相连,只能通过 Labview 软件编程,控制表情机器人的运动。

通过调节 PWM 的值,控制表情机器人上的舵机的动作,实现“微笑”、“悲伤”、“张嘴”动作的实现。单个调节舵机和一次调节多个舵机的时候,相同的 PWM 值,会出现不同转动的角度,所以需要反复多次的调试。表情机器人微笑的动作如图 15(a)所示。表情机器人悲伤动作如图 15(b)所示。表情机器人张嘴动作如图 15(c)所示。

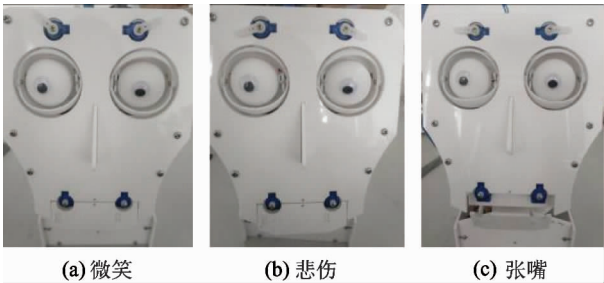


图 15 表情机器人的 3 种动作

5 系统评估

5.1 语言训练的方法

语言训练的方法有很多种,包括训练日常的交流能力、识别物体能力、语言器官运动能力等。

本系统主要是针对训练脑瘫患者识别物体的能力。在测试前,首先要对患者进行一个测试,判定患者在使用康复训练系统前处于哪个阶段,将患者所处的阶段设为训练初始状态,根据每名患者的情况,制定不同的培训方案。

5.2 语言方法介绍

本系统是采用人机交互的模式对患者进行语言训练的,通过 51 单片机、语音芯片、机器人和图片结合的方法。通过语音芯片对患者所发出的声音进行采集,并与库里已经拥有的拼音进行比对,传送相应的数据到单片机,单片机做出相应的判断,并进行计时。在按键按下和识别正确这两个状态时,传送相应的数据到上位机,使上位机可以显示与下位机显示屏上的拼音相对应的图片,以及使上位机可以进

行患者识别正确的判断和实时记录患者识别的时间。识别正确的同时,上位机会发送相应的指令给机器人,机器人会做出相应的表情。最后,所有患者的数据都被记录在一张表格中,患者的评价图将会在上位机的显示界面自动生成。

5.3 案例分析

案例 1:患者 A,女性,4 岁,被临床诊断为徐动型脑瘫。其症状为不能很好地控制自己的动作,有时会出现夸大一个动作,在一个简单动作的过程中,会伴随全身的动作以及脸部表情的扭曲。经 S-S 法检测,该患者交流能力良好,可以保持注意力集中在 8 s。语言诊断的结果为语言交流能力基本正常。对患者 A 进行了 8 次训练,每次检测的时间如表 4 所示。

表 4 患者 A 识别图片所用时间表

次数	图片	时间 (s)	图片	时间 (s)	图片	时间 (s)
1	微笑	15	悲伤	16	张嘴	14
2	微笑	13	悲伤	14	张嘴	15
3	微笑	12	悲伤	10	张嘴	12
4	微笑	11	悲伤	7	张嘴	9
5	微笑	9	悲伤	6	张嘴	9
6	微笑	7	悲伤	5	张嘴	7
7	微笑	5	悲伤	4	张嘴	5
8	微笑	5	悲伤	4	张嘴	5

从表中可以看出,患者 A 语言交流基本正常,识别图片的正确率为 100%,并且识别每张图片的时间在逐渐减少,识别速度越来越快。可见,康复训练有了一定的效果。相比较每次识别中对 3 张不同图片的识别时间,可以看出,其时间也是在逐步减少。由于脑瘫患者注意力集中较差,在最开始训练的时候,患者的注意力很难集中,所以识别的时间会比较长。但由于本系统内加入了机器人,在第一次识别正确后,机器人会做出相对应图片的表情,患者的注意力会被机器人所吸引,而且采用这一方法更容易让患者对康复系统产生兴趣,从而更有利于康复训练治疗的开展。

本设计在 Labview 的上位机界面中引入了查询的功能。在开始训练系统之前,输入患者的信息,查

询按钮按下后便可以看到该患者训练的时间变化曲线,可以看到该患者识别每个图片时所用时间的变化曲线。患者 A 的识别变化曲线如图 16 所示。

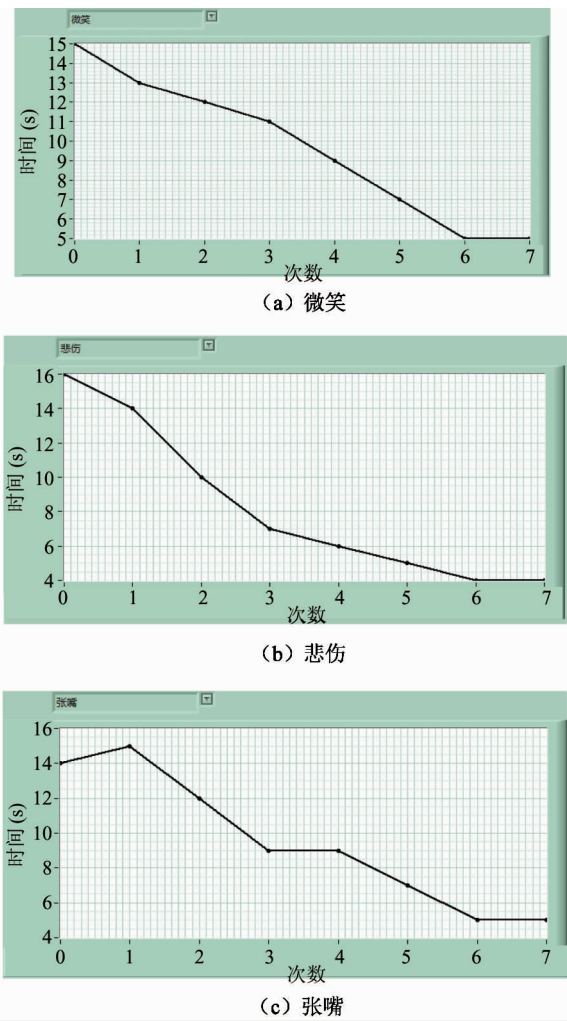


图 16 患者 A 的识别时间变化曲线

从变化曲线图可以直观地看出,患者 A 在识别图片时的总体识别时间是呈下降趋势,但有时也会有所上升,特别是在前几次的测试过程中,患者由于没有很熟练地记住所有图片对应的文字,会有个遗忘的过程,但训练到最后几次,就没有了反复的趋势,识别时间有了明显的下降,并且在良好情况下保持稳定的状态。

案例 2:患者 B,男性,5 岁,被临床诊断为运动障碍型脑瘫。其主要表现为肌肉无目的性运动,肌肉会呈现震颤,并且无法保持平衡。经 S-S 法检测,该患者可集中注意力在 10 s 左右,但发音不清晰,

并且颜色只能辨别出红色和蓝色。语言诊断的结果为语言发育缓慢。对患者 B 进行了 8 次训练,每次检测的时间如表 5 所示。

表 5 患者 B 识别图片所用时间表

次数	图片	时间 (s)	图片	时间 (s)	图片	时间 (s)
1	微笑	15	悲伤	20	张嘴	14
2	微笑	/	悲伤	14	张嘴	/
3	微笑	12	悲伤	8	张嘴	12
4	微笑	11	悲伤	9	张嘴	/
5	微笑	9	悲伤	/	张嘴	9
6	微笑	/	悲伤	/	张嘴	9
7	微笑	5	悲伤	11	张嘴	5
8	微笑	8	悲伤	6	张嘴	5

注:/表示患者未识别出图片

从表中可以看出,患者 B 的语言功能有障碍。在每张图片的识别中,都有出现识别不出图片的情况,而且患者开始识别的时间比患者 A 开始识别的时间长了很多。识别不出的情况在开始时尤为明显,在第 2 次和第 6 次测量的时候,出现了 2 张图片都识别不出的情况,但识别的时间有所下降。可见,患者识别的正确率并不与识别的时间成正比,并不是识别速度越快,识别的正确率就越高。对于有语言功能障碍的患者而言,影响识别的因素有很多。与患者接受康复训练时的心态也有很大的关系。患者 B 在前 3 次识别正确,观察到机器人做出相应动作后,情绪明显高涨,但有时情绪过于兴奋并不利于脑瘫患者的识别。患者 B 的发音不清晰,过于兴奋后,患者 B 会出现肌肉震颤,导致发音更加不清晰,语音芯片无法识别。这也是本设计一个需要改进的地方,需要改进语音芯片,提高语音芯片的识别能力。患者 B 的识别时间变化曲线如图 17 所示。

从变化曲线图可以直观地看出,患者 B 在识别图片时的识别时间是呈下降趋势,但有时也会有所上升,并且由于患者有比较严重的语言障碍,不能正常地进行语言交流,所以出现了识别不出的情况。从训练的变化趋势可以看出,患者的识别时间有了明显的下降,并且在良好情况下能保持稳定的状态。

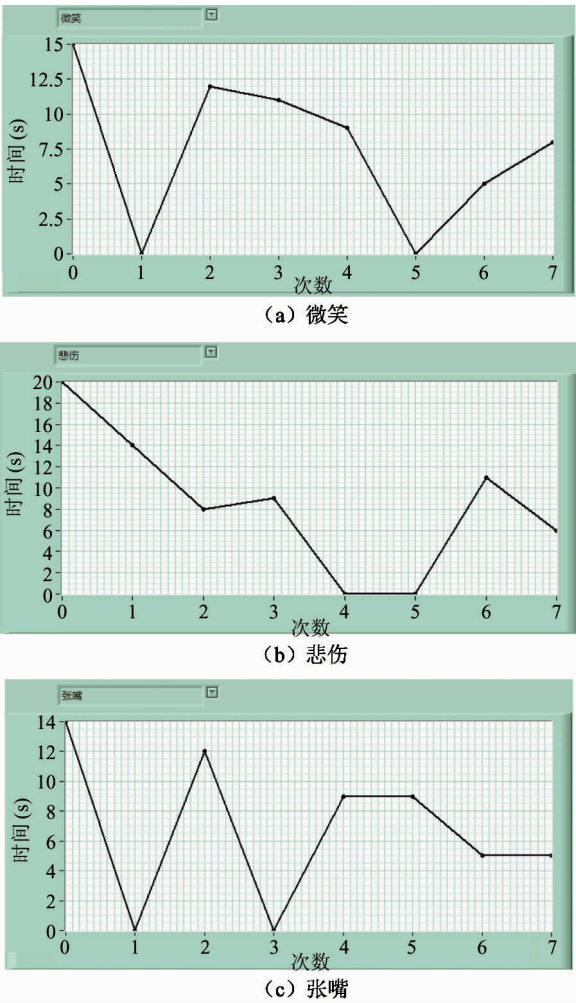


图 17 患者 B 的识别时间变化曲线

从以上 2 个案例可以看出,本系统对脑瘫患者的康复具有一定的作用,患者在初步训练的时候,识别的时间较长,但随着训练次数的增加,识别时间有了明显的下降。对于不同的患者,识别正确率也有所差别,机器人在本系统中起到辅助的作用,可以有效提高患者的注意力,也可以使患者更明白表情所对应的动作,有利于患者在日常生活中的识别和正常交流。

6 结 论

本文设计了基于语音辨识的脑瘫康复训练系统,采用上位机和下位机相结合的方式,实时记录脑瘫患者的各项数据,并以 8 次为一个疗程,便于分析训练系统对患者的治疗效果。实验结果表明,该设计为脑瘫患者康复训练提供了一种新方法。

参考文献

[1] 王秋根,张秋林主编. 脑瘫[M]. 上海:第二军医大学出版社, 2001

[2] 陈旭红主编. 图解脑瘫康复技术与管理[M]. 北京:华夏出版社, 2007

[3] 唐泽媛,李炜如编著. 幼儿疾病与智力开发[M]. 成都:四川人民出版社, 1989

[4] 李林. 小儿脑性瘫痪的教育康复[J]. 中国康复理论与实践,2007,13(12):1110-1112

[5] 王燕,张燕. 小儿脑性瘫痪的语言治疗方法[J]. 中国临床医生杂志,2003,31(4):6-8

[6] 张毅刚,彭喜元. 单片机原理与应用设计[M]. 北京:电子工业出版社, 2008

[7] 龚伟. 小词汇量孤立词语音识别系统的 DSP 实现[J]. 电器与能效管理技术, 2009(20):1-4.

[8] 韩志艳. 语音识别及语音可视化技术研究[M]. 沈阳:东北大学出版社, 2017. 01

[9] Koerner M 编著,李逸波等译. 最新语音识别技术[M]. 北京:电子工业出版社, 1998

[10] 张志霞,韩慧莲,薛宏伟. 语音识别技术分析[J]. 电脑开发与应用, 2008,21(12):33-35

[11] 洪家平. LD3320 的嵌入式语音识别系统的应用[J]. 单片机与嵌入式系统应用, 2012,12(2):47-49

[12] Alvaradomendez E, Andradelucio J A, Estudilloayala J M, et al. Implementation of a laser beam analyzer using the image acquisition card IMAQ (NI) [C]. In: Proceedings of the SPIE-International Society for Optical Engineering, Tandil, Argentina, 2001. 301-304

[13] Chiu P, Kapuskar A, Reitmeier S, et al. Room with a rear view: meeting capture in a multimedia conference room[J]. *IEEE Multimedia*, 2002, 7(4):48-54

[14] Wilpon J, Rabiner L, Lee C H. Automatic recognition of keywords in unconstrained speech using hidden Markov models [J]. *IEEE Transactions on Acoustics Speech & Signal Processing*, 1990, 38(11):1870-1878

[15] Nakmahachalasint P, Ngo K D T, Vu-Quoc L. A static hysteresis model for power ferrites[J]. *IEEE Transactions on Power Electronics*, 2002, 17(4):453-460

[16] 杨乐平著. LabVIEW 程序设计与应用[M]. 北京:电子工业出版社, 2001

[17] 郭洁,王召巴. 基于 LabVIEW 的串行通信接口设计与实现[J]. 机械工程与自动化, 2008(5):57-59

[18] 修金鹏著. Arduino 与 LabVIEW 互动设计[M]. 北京:清华大学出版社, 2014

Face-expression and speech recognition based rehabilitation training system

Lu Zhenli^{*}, Jiang Ruixuan^{*}, Ma Zhipeng^{***}, Borovac Branislav^{***}, Marjan Mernik^{****}

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

(^{**} School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116)

(^{***} Faculty of Technical Sciences, University of Novi Sad, Novi Sad 21000, Serbia)

(^{****} University of Maribor, Maribor 462000, Slovenia)

Abstract

A cerebral palsy rehabilitation training system is designed based on human-computer interaction (HCI). Using the combination of upper computer and lower computer, the lower computer adopts 51 single-chip microcomputer for main drive, the voice acquisition module uses LD3320 voice chip, and connects the lower computer of C language programming with the upper computer written by Labview through serial communication. The voice chip sends the signal identification result to the host computer through the serial port, and the host computer performs identification matching according to the voice semantic rule, judges and counts the test, and matches the result with the expression robot. By recording the number of times the two patients correctly recognized the expression and the recognition time, the data are analyzed, and the results prove that the system can improve the effect of cerebral palsy rehabilitation training.

Key words: human-computer interaction (HCI), speech recognition, cerebral palsy rehabilitation, expression robot, evaluation analysis