

融合语音和脑电的智慧病房控制系统^①



蔡旭刚, 王磊, 王帆, 李俊廷, 潘家辉

(华南师范大学 软件学院, 佛山 528225)

通讯作者: 潘家辉, E-mail: panjh82@qq.com

摘要: 在智能化技术革新各个传统行业的过程中, 对于传统病房护理人员提出了更高水平的服务诉求. 在传统病房实地调研的基础上, 为了提高患者日常病房生活中的自理能力, 加强医护和家属对病人生活状况的实时监控, 结合现有物联网智能控制技术, 本文提出一种融合语音和脑电的智慧病房控制系统, 实现了病房电器等基础设施的控制和云端病房实时监测. 此外, 基于眨眼动作的 ElectroEncephaloGraphy (EEG) 控制方法在一定程度上解决了对患者身体状况的更高要求的问题. 根据涉及 10 名受试者的两个实验, 语音识别的准确度达到 98%, 对健康人和患者基于脑电的眨眼识别的准确率分别为 94.3% 和 82.9%. 结果表明, 该系统能够在病房这类复杂环境下稳定运行, 为患者提供更加智能舒适的疗养环境.

关键词: 智慧病房; 脑机交互; 眨眼检测; 语音识别; 物联网; 脑电

引用格式: 蔡旭刚, 王磊, 王帆, 李俊廷, 潘家辉. 融合语音和脑电的智慧病房控制系统. 计算机系统应用, 2021, 30(11): 71-81. <http://www.c-s-a.org.cn/1003-3254/8145.html>

Smart Ward Control System Integrating Speech and EEG

CAI Xu-Gang, WANG Lei, WANG Fan, LI Jun-Ting, PAN Jia-Hui

(School of Software, South China Normal University, Foshan 528225, China)

Abstract: People put forward a higher service demand for traditional ward care in the process of intelligent technology reforming various traditional industries. According to the field investigation of traditional wards, the self-care ability of patients should be improved in daily ward life and the real-time monitoring of their living conditions by medical staff as well as families is waiting to be strengthened. To this end, this study proposes a smart ward control system integrating speech and ElectroEncephaloGraphy (EEG) and combined with the Internet of Things (IoT), thereby realizing the control of infrastructures such as ward appliances and real-time monitoring of cloud wards. Furthermore, the EEG-based control method based on eye blinks can meet the higher requirement for a patient's physical conditions. According to the two experiments involving 10 subjects, speech recognition achieved an accuracy of 98%, and the accuracy of EEG-based blink recognition for healthy people and patients is 94.3% and 82.9%. The results show that the system can operate stably in complex environments such as wards, providing patients with a more intelligent and comfortable convalescent environment.

Key words: smart ward; Brain-Computer Interface (BCI); blink recognition; speech recognition; Internet of Things (IoT); Electroencephalography (EEG)

① 基金项目: 广东省重点研发计划 (2018B030339001); 广州市科技计划重点领域研发计划 (202007030005); 国家自然科学基金 (61876067); 广东省自然科学基金 (2019A1515011375)

Foundation item: Key Research and Development Program of Guangdong Province (2018B030339001); Science and Technology Plan on Key Areas of Guangzhou Municipality (202007030005); National Natural Science Foundation of China (61876067); Natural Science Foundation of Guangdong Province (2019A1515011375)

收稿时间: 2021-01-12; 修改时间: 2021-02-07, 2021-02-23; 采用时间: 2021-03-02; csa 在线出版时间: 2021-10-22

我国医疗卫生市场庞大, 医护人员工作负荷重情况突出, 据中国国家卫生健康委员会发布的《2019年我国卫生健康事业发展统计公报》显示^[1], 2019年全国医疗卫生机构总数达1007 545个, 全国医疗卫生机构床位880.7万张, 医院医师日均担负诊疗7.1人次和住院2.5床日。伴随着新冠疫情与高新科技的冲击, 医疗卫生行业模式正加快从传统化、数字化、信息化走向智能化^[2]。医院病房作为医院医疗产业的核心支柱, 其智能化程度与病人舒适指数、医护负荷指数、医疗资源分配合理指数、管理运作效率指数等密切相关。物联网、5G通信、云计算、大数据等高新技术在医疗病房等传统行业的应用已是智能化时代背景下的大势所趋, 以“互联网+医疗”的智慧病房协同控制控制系统是当今医院医疗产业信息化研究与建设的热点方向^[3]。

智能控制的核心在于控制机制的有效性、健壮性、实时性、简易性与可维护性。在传统病房以医护为核心的人工服务机制中, 紧急呼叫与控制机制多为人工服务。应对众多患者的不同需求, 医护工作负荷沉重, 尤其是应对言语障碍(Language Disorder Syndrome, LDS)、脊髓损伤(Spinal Cord Injury, SCI)、肌萎缩性脊髓侧索硬化症(Amyotrophic Lateral Sclerosis, ALS)等生活能力较弱的病患, 难以做到针对不同病患的个性化服务, 且医护查房工作经常出现“观察查房”现象, 对病人护理与治疗缺少沟通交流, 病患缺失修养与治疗所需求的幸福感与舒适感^[4]。同时病患医疗必要生理数据获取需要消耗医护大量工作时间进行日常记录, 病房安全需要定时检查维护, 尤其对于负压病房等高要求医护病房, 大量的数据记录与机械人工预警极大浪费专业安全人员的精力与时间。

本文提出一种融合语音与脑电信号的智慧病房控制系统。语音交互方面, 基于STM32微控制器(Micro-Controller Unit, MCU)与SYN7318离线非定向语音识别模块对语音信号识别分类并转化为控制指令, 利用TCP/IP协议远程传输生理与安保数据完成云端被动预警控制; 脑电交互方面, 基于一种波峰波谷特征的波形检测算法对眨眼动作进行识别并转化控制指令。最后, 融合语音控制的高效性、脑电控制的自然性和远程终端控制的准确性设计一种互补性融合的多模态智慧病房控制系统。

本文的主要贡献如下:

(1) 根据离线语音模块的语义特性适配传输速率

并设计传输协议, 对语音信号进行指令转换, 完成语音交互控制。

(2) 根据脑电采集设备采集的脑电数据格式, 设计脑电信号传输协议, 并基于脑电信号生理特征设计一种基于波峰波谷特征的眨眼波形检测算法, 对输入的信号进行起始点检测、分帧、去噪等预处理, 经过特征提取和分析实现眨眼波形检测, 最后适配眨眼特性, 设计一种眨眼紧急呼叫机制, 完成脑电紧急呼叫。

1 国内外研究现状

1.1 脑电控制研究现状

近年来, 很多国家和组织都致力于开发实用且可靠的人机交互系统, 如语音交互系统^[5]、单一开关系统、眼球追踪系统^[6]等辅助系统, 但是对于那些完全丧失外周神经肌肉控制能力与语言沟通能力的人来说, 这些辅助系统很难使用或者无法使用, 需要借助另一种自然性更高的辅助系统, 即基于脑机接口(Brain-Computer Interface, BCI)技术的新型辅助系统。通过侵入式或非侵入式脑电信号监测技术, 将大脑中神经电信号能实时、准确地通过BCI系统转换成可被计算机识别的命令或操作信号, 为大脑与外部世界提供直接的交流方式。

随着相关研究推进, 脑机接口技术的含义进一步得到外扩, 脑机接口技术不再局限于实现对用户大脑状态的检测, 还用于实现大脑对外部设备的控制。四肢瘫痪患者利用侵入式BCI控制机械臂完成手部伸手动作与抓取动作^[7,8], 此种控制方法备受患者青睐。但安全性一直是困扰侵入式BCI发展的主要问题, 相比之下, 非侵入式BCI拥有更安全、更简易、更普适的控制优势。目前, 多名学者与研究团队将BCI控制应用到轮椅中^[9,10], 用于解决SCI、ALS患者行动不便的问题。Jafri等^[11]提出了一种基于无线脑电接口(Wireless Brain-Computer Interface, WBCI)技术的高可移植性智能家居系统, 通过P300脑机接口范式选择家电设备并完成后续的控制操作。考虑一定比例患者由于文化、学科等因素导致无法操作单一类型BCI系统^[12]。因此, “混合型”BCI系统可为患者在不同的方法中提供最优的BCI控制, Edlinger等^[13]提出了一种基于P300与SSEVP多模态的融合控制范式, 能够在虚拟智能家居环境中完成高精度的工作。Chen等^[14]结合增强现实技术、计算机视觉技术和稳态视觉诱发电位-BCI技术,

设计并实现了双臂控制系统。除脑电控制之外, 多名研究学者将眼电信号与脑电信号进行融合, 使系统拥有更高的使用效率和更广的应用范围^[15]。Venkataramanan等^[16]采集水平通道与垂直通道的眼电信号, 经由伪影和干扰电弧处理, 进行眨眼检测, 并将眨眼信号应用于医院二级报警系统控制。

1.2 语音控制研究现状

相对于其他交互方式而言, 语言是人类交往中最基础、最便捷的交流方式, 伴随着人工智能的突飞猛进, 自动语音交互 (Automatic Speech Recognition, ASR) 技术已初步实现计算机与人类之间的语音交流, 基于此项技术, 学者对语言交互的应用场景展开了研究。Portet等^[17]开发了一种针对老年人的家庭自动化语音控制系统——SWEET-HOME, 并评估了使用者及家属对该技术的接受程度。该研究表明, 语音交互技术具有解决老年人和残障者日常生活需求的巨大潜力, 比其他侵入性脑机控制解决方案更容易被接受。Jadczyk等^[18]实现了一个具有语音功能的医疗自动化平台 CardioCube, 可以从患者那里收集医疗数据, 并将其用于进一步的分析。此语音交互功能有效提高了用户操作系统的便捷性, 无论是患者还是医护人员都对该系统给予了很高的评价。Peng等^[19]使用 Google Home, Raspberry Pi 和 Smart Bluetooth Socket 构建智能家居控制系统, 通过机器学习算法融入应用程序为用户提供新型智能家居服务。在该系统下, Google Home 识别用户的语音命令, 然后将指令通过蓝牙传输到 Raspberry Pi 进行家电的控制。毛博等^[20]基于深度学习方法提出 DeepHome 智能家居管控模型, 能够基于家居环境数据预测智能设备工作状态, 并依照预测结果调整相应设备, 实现对智能家居的自动化统一管控。在真实环境数据集下, 对设备状态的预测准确率达到 98.9%。许建国等^[21]为医院病房智能监护系统的设计提出一种基于物联网和 Android 技术的解决方案。该系统由移动智能终端 SDA、ZigBee 模块、A10 平台、服务器等部分构成。

近年来, 智慧病房控制系统的升级改造一直在进行, 语音、脑机接口、物联网等交互技术在智慧家居、智慧医疗等场景取得了长足进步, 每年有大量的研究论文、专利和演示系统发表与涌现, 但实际应用仍存在诸多问题。

(1) 病房交互技术中缺乏丰富、准确、实时的控

制指令。目前大部分医院病房中护理与交互依旧是“端到端”的人力完成, 极少涉及语音交互技术与脑机交互技术在病房场合的应用, 归其原因, 语音交互系统与脑机交互系统要控制复杂对象, 需要从语音信号或脑电信号中产生多个相互独立的控制信号, 且这些控制信号要具有良好的准确性和实时性。目前的语音交互系统大多数利用大数据技术进行在线识别, 语音数据的上行传输和识别结果的下行传输需要花费一定时间, 产生时延, 应用于病房的智慧控制缺乏控制的实时性, 且语音识别对所需要的噪音环境要求较高, 在噪声很大的环境下, 其精准度受到极大影响; 在脑机交互方面, 目前的非侵入式脑机接口能提供的独立控制信号还非常有限, 大多局限于一维控制和较少类别 (一般 2 类或者 3 类) 的脑电信号检测。即使出现了少数多维/多类系统, 也基本上是一维、二维控制信号和二分类系统的扩展, 造成脑机交互系统性能和应用上有很大的局限性。

(2) 模态单一, 结合语音和脑电的多模态控制研究很少。目前一些医疗病房领域采用的智慧控制方式多为单一模态控制, 无法覆盖病房内全部患者的服务要求。控制方式中语音控制与脑电控制分属两种模态控制, 两种模态控制方式各具有其优势特色, 应对不同特征患者, 各自拥有自己差异化应用解决方案, 目前结合语音交互的高效性与脑电交互的自然性, 发挥各自差异化应用特点, 形成一款互补式的融合控制机制进行主动与被动控制, 仍然处于学术与市场的初探阶段。

2 系统总体设计

本系统将物联网技术、语音识别技术和与脑机接口技术相结合, 以传统的病房护理为场景, 构建以辅助预警及智能护理为主要功能的智慧病房控制系统, 该控制系统结构总图如图 1 所示。

语音识别交互模块能够满足病患在病房的基本控制需求与交互服务。系统中语音识别依托 SYN7318 非定向离线语音芯片, 首先采集病房内病患语音信号, 信号经由预处理后提取相应语音信号特征, 对其进行离线数据库比对, 最后识别出语音中的病患控制指令, 完成语音交互功能, 包含语音控制、离线语音聊天、语音定时、语音呼叫、音频播放等功能。

云端管理与控制模块负责预警控制。病房相关安全数据与医患常规测量数据经由 TCP/IP 协议传入云

端,以便安保人员与医护人员检测与记录,并基于阈值形成一种被动控制方式。

显示播报模块利用无线通信技术与云端管理与控制模块进行远程信息传输,将病房患者的交互需求与紧急呼叫进行可视化显示与实时动态语音提醒。

脑电采集与分析模块接收 HNNK 脑机智能生命

环原始脑电信号,对信号进行眨眼波形检测,以提供脑电控制指令紧急报警.遵照脑电信号本身生物电特性进行滤波降噪,将降噪后的连续脑电信号进行起止点检测与目标字符段进行划分,最终脑电数据经由一种基于波峰波谷特征的波形检测算法进行眨眼识别,利用患者多次眨眼信号作为紧急告警指令,实现紧急报警功能。

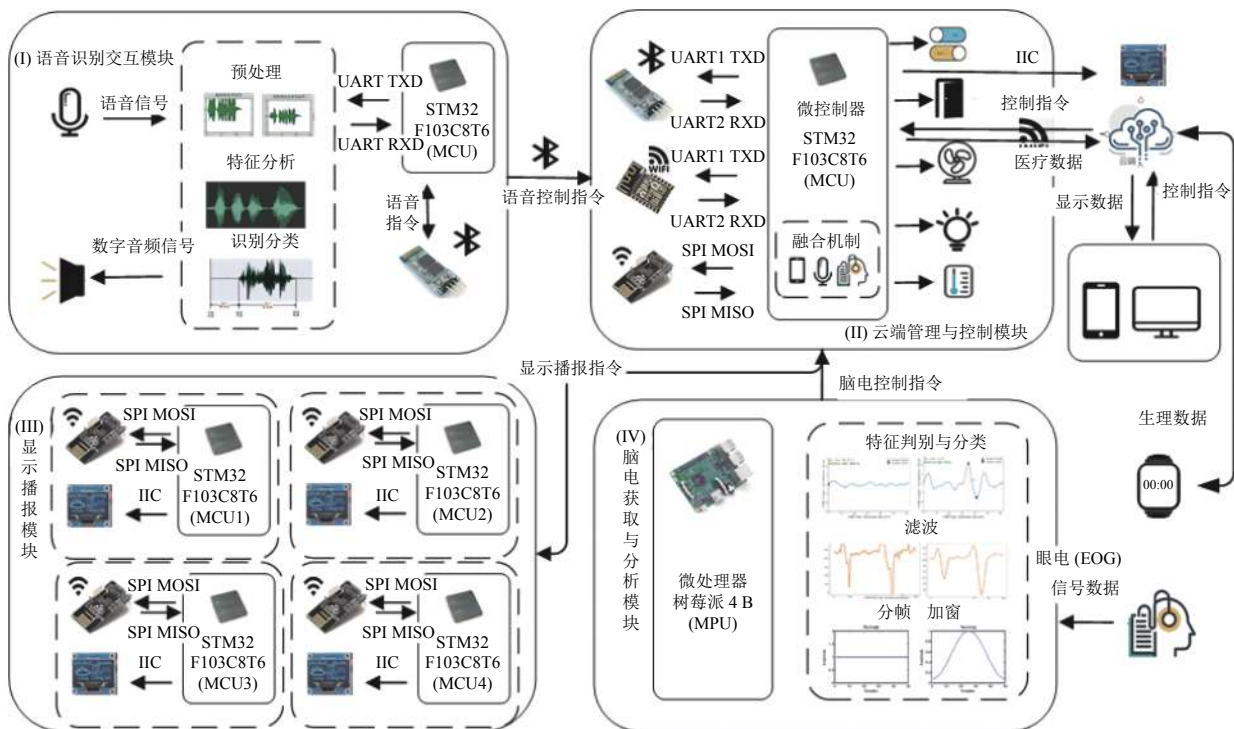


图1 智慧病房控制系统结构

3 语音识别交互模块核心技术

为了实现病患人机离线语音交互功能,语音识别处理是不可或缺的一环.有效的语音识别处理能将病患语音模拟信号实时转化为相对应的语音特征指令,并对后续系统控制提供最正确的指令特征信息.同时病房环境中识别病患语音信号能使最终所执行的控制指令更具准确性和实用性。

为实现语音信号的实时与准确识别必须完成下列任务:

- (1) 适配通信速率去保障通信传输准确性;
- (2) 分析语音命令的语义结构进行译码识别。

3.1 语音识别交互系统结构

语音识别交互系统架构包括硬件、接口、协议、功能、模式5层,如图2所示。

硬件与接口层以 STM32F103C8T6 主控器为核心,

OLED 显示屏、硬件驱动电路、离线语音识别单元、无线通信单元和外围硬件电路为辅,依托模拟信号采集传感器读取相关数据,对 STM32F103C8T6 主控单元 GPIO 引脚进行复用配置提供上层数据传输接口.在协议层依托硬件与接口层机械、电子功能部分特性,规定通信过程中的通讯逻辑、统一收发双方的数据打包与解包标准,确定双方通信信道与连接方式,保障双方有序通信,准确完成数据获取.功能层依托协议层与硬件接口层,完成语音交互、数据传输、指令融合和外部控制等模块功能.应用模式层规定顶层模式,合理组织调用业务逻辑功能与常用功能,对系统支持功能进行合理的整合与调配,保持系统的合理性、有效性和稳定性。

3.2 语音交互指令获取

(1) 波特率选择

波特率指数据信号对载波的调制速率,它用单位

时间内载波调制状态改变次数来表示,单位为 Baud. 比特率指单位时间内传输的比特数,单位 b/s^[22]. STM32F103C8T6 主控 HSE 高速外部时钟信号采用 8M 晶振为基频,经过锁相环倍频达到 72 MHz 主频,以保障系统稳定运行.其 USART1/USART2 外设时钟分别来源于 APB2/APB1 总线时钟,最大频率为 36 MHz/72 MHz,最大支持 4.5 Mb/s,考虑 SYN7318 所支持的 4800 Baud、9600 Baud、57 600 Baud 和 115 200 Baud 4 种波特率、串口 10 位传输字符格式,对双方通信设备通信速率、波特率误差、通信距离综合考虑进行如表 1 所示比较,最终选择 115 200 Baud 为通信传输波特率.计算通信速率为式 (1),计算波特率与误差为式 (2) 和式 (3):

$$I = S_{Tx/Rx\text{baud}} \times \log_2(N) \quad (1)$$

其中, I 代表比特率; $S_{Tx/Rx\text{baud}}$ 代表波特率, N 代表每个符号负载的信息量,由于 STM32F103C8T6 中单个调制状态对应 1 个二进制位,比特率等于波特率.

$$S_{Tx/Rx\text{baud}} = \frac{f_{ck}}{8 \times (2 - OVER8) \times USARTDIV} \quad (2)$$

其中, f_{ck} 为 USART 时钟, $OVER8$ 代表 16 倍过采样,用 0 表示, $USARTDIV$ 代表一个存放在波特率寄存器 (USART_BRR) 的一个无符号定点数.其中 $DIV_Mantissa[11:0]$ 位定义 $USARTDIV$ 的整数部分, $DIV_Fraction[3:0]$ 位定义 $USARTDIV$ 的小数部分,波特率的误差正是由于

对小数部分的取整赋值所造成.

$$S_{error} = \frac{(S_{real} - S_{base})}{S_{base}} \times 100\% \quad (3)$$

其中, S_{error} 代表波特率误差率, S_{base} 为设置波特率, S_{real} 为实际波特率.

(2) 命令帧解析

SYN7318 以“帧头+数据区长度+数据区”格式与 MCU 主控通信,其语音命令帧格式如表 2 所示.

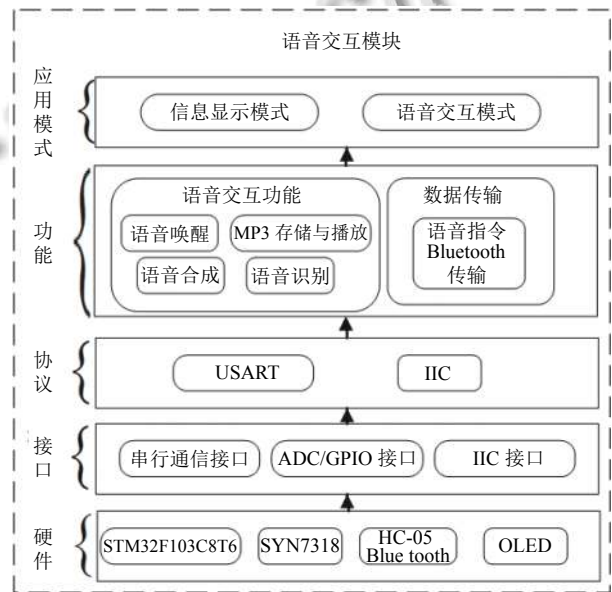


图 2 语音识别交互系统架构

表 1 通信速率与误差比较表

波特率		比特率		$f_{ck}=32\text{ MHz (16倍过采样)}$					
设定值 (Baud)	设定值 (b/s)	实际值 (Baud)	USARTDIV 寄存器编程值	理论误差 (%)	实测通信误差 (%)	理论通信速率 (KB/s)	实测通信速率 (KB/s)	通信距离 (m)	
4800	4800	4800	468.75	0	0	0.48	0.426 136	0.2	
9600	9600	9600	234.375	0	0	0.96	0.852 275	0.2	
57 600	576 00	57 600	39.0625	0	0	5.76	5.113 781	0.2	
115 200	115 200	115 384	19.5	0.15	0	11.52	10.243 802	0.2	

表 2 语音命令帧格式表

帧结构	帧头	数据区长度 (2 B)	数据区 (小于等于 4 KB+2 B)		
			命令字节 (1 B)	命令参数 (N B)	待发送文本 (<=4 KB)
数据	0xFD	0x00 0x00	0x00	0x00……	0x00……
说明	通信识别初始字节	数据总字节数 (高字节在前, 低字节在后)	“语音唤醒” “语音识别” “MP3播放” “播放控制” “综合命令”	“编码格式” “回传类型” “识别结果” “查询回传”	“回传类型” “待发送文本编码”

帧头设定“0xFD”为设备双方串口通信识别初始字节, 占用一个字节; 数据长度使用两个字节进行标识, 发送与回传指令最大长度不大于 4096 B; 数据区为包含命令字节、命令参数与待发送文本, 命令字节使用 1 B 选择模式, 同时确定对应模式具体功能, 识别模式命令参数字节长度不定, 通常为 5 B, 包含匹配度 (1 B)、词条 ID (2 B)、命令 ID (2 B), 其余模式命令参数为无, 待发送文本为存储语音合成词条的编码字节和回传类型, 字长根据语音合成文字长度设定, 字节长度小于等于 4 KB, 根据其字节接受特性与实际通信测试, 同一帧数据中, 每个字节之间的发送间隔不能超过 15 ms; 帧与帧之间的发送间隔必须超过 15 ms, 考虑通信质量不被影响, 至少留出 2 ms 时延长余量, 即帧与帧发送间隔大于 17 ms. 且在发送过程中, 发送命令帧严格控制 4096 B, 超过 4096 B 会导致发送失败.

3.3 语音指令分级编码

对 SYN7318 通信传输速率与通信协议匹配后, MCU 具备与语音模块双向通信的能力. 为完成人性化的语音交互功能, 将离线语音模块所支持的语音功能高效的综合利用, 还需要对语音指令进行分级编码传入语音、脑电与云端控制融合模块, 防止指令混传, 造成功能混乱. 编码规则如表 3 所示.

表 3 语音指令编码规则表

编码	状态编码 (1位)	模式编码 (2位)	详细功能编码 (4位)
例	激活 (1)	病房一号远程服务 (16)	呼叫护士 (0004)

编码规则采取 16 位二进制编码方式, 通过蓝牙 HC-05 近距离逐字节传入控制系统进行分析识别, 传输时间间隔为 2 s, 语音指令编码长度为 7 位, 其状态编码表示当前数据是否为有效语音指令; 模式编码为一级词条编码, 确定要选择的语音交互模式或控制指令; 详细功能编码为当前已选模式下的具体语音功能. 将语音指令进行编码化设计, 根据其编码规则进行表达与关联, 确保语音指令的准确性, 避免指令间的识别冲突, 方便后续离线语音词条的添加与删除, 降低复杂度.

4 脑电获取与分析模块核心技术

脑电控制模式针对肢体残疾且语言能力受损的特殊看护病人, 提出一种根据眨眼动作完成控制与紧急呼叫的解决方案. 其中, 眨眼波形识别处理是不可或缺

的一环. 有效的眨眼识别处理能将病患眨眼动作准确识别并转化为眨眼特征指令, 对后续脑电紧急呼叫提供最准确的指令特征信息. 为此脑电控制与紧急呼叫必须完成下列任务:

(1) 眨眼波形识别.

1) 根据 HNNK 脑机智能生命环传输协议解析并获取原始脑电信号;

2) 根据生物电特性对脑电信号进行滤波降噪, 将降噪后的连续脑电信号进行起止点检测, 对目标字符段进行划分;

3) 设计一种基于波峰波谷特征的波形检测算法进行眨眼波形检测.

(2) 脑电指令转换.

设计一种脑电紧急呼叫机制, 通过用户多次眨眼动作触发远程脑电紧急呼叫.

4.1 脑电获取与分析处理系统流程

脑电控制与紧急呼叫是本系统的一个重要功能, 其流程框图如图 3 所示.

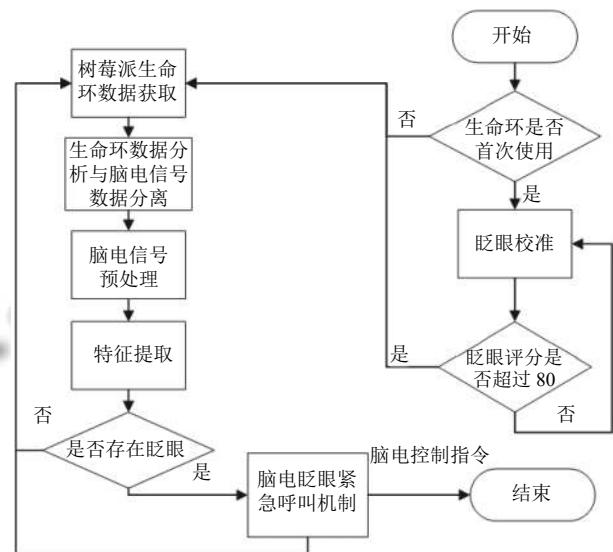


图 3 脑电紧急呼叫流程图

用户首次使用该功能, 首先需要多次眨眼校准, 待眨眼评分大于 80 后进入动态采集检测部分, 如非首次使用, 可跳过此环节. 眨眼校准完成后, 进入实时眨眼判别环节, 基于树莓派 4B 每隔 16 ms 进行一次生命环数据读取, 每次读取 20 B 数据 (包含 8 ms 时间间隔的 4 B 脑电数据), 并绘制其为实时脑电波形. 在脑电波形图中, 信号的幅值随眨眼动作的进行而改变, 为此采

用含有 n 个元素的波形向量 ($X = [x_1, x_2, x_3, \dots, x_{n-2}, x_{n-1}, x_n]$) 进行表示, 其中 x_i 为波形上的某一点幅值, 两个相邻向量点之间的时间间隔 Δt 为 8 ms, 由此组成的波形向量以脑电信号的生物电特征为基本判别标准对脑电信号进行起始段检测与分帧、滤波等预处理. 最终根据眨眼操作完成控制与紧急呼叫.

4.2 脑电信号预处理

(1) 起始段检测与分帧

根据脑电信号的非平稳特征, 即从整体来看其特征与表征本质特征的参数均为随机变化, 具有时变特性. 但在一个短时间内, 其特性基本保持不变即相对稳定, 因而可以将其看作一个准稳态过程, 即具有短时平稳性. 后续分析要进行“短时分析”^[23]. 树莓派以 8 ms 的间隔时间进行数据点存储形成波形向量, 在实际测试中, 眨眼波形的数据点远远小于未眨眼数据点, 避免浪费大量数据处理的时间, 有效捕获有眨眼动作波形, 需要对脑电信号进行起始段检测, 以排除非眨眼动作的无效脑电信号, 提高眨眼识别处理的准确度与效率.

本设计中采用移动平均法来动态检测活动段, 根据眨眼动作时, 脑电信号幅值产生波动的特点, 采用短时间信号幅值均值为预测值, 实际脑电信号幅值与预测值相减, 差值大于阈值 (阈值取值以经验值为标准) 的点为基本起始点. 其数学表达式如式 (4) 所示, 其中 x_k 为下一个时期脑电信号的预测值, n 为移动平均的时期个数, $x_{k-1}, x_{k-2}, x_{k-3}$ 等为前几期的实测脑电信号幅值. x_{real} 为下一个时期实测脑电信号幅值, T_{begin} 为起始点检测差值经验阈值. 式 (5) 中, $S=1(S=0)$ 表示检测到眨眼起始点 (未监测到眨眼起始点), 注意采用的移动平均法为简单移动平均法, 每个时期脑电信号幅值权重都同为 1.

$$x_k = [x_{k-1} + x_{k-2} + x_{k-3} + \dots + x_{k-n}] / n \quad (4)$$

$$S = \begin{cases} 0, & x_{\text{real}} - x_k \leq T_{\text{begin}} \\ 1, & x_{\text{real}} - x_k > T_{\text{begin}} \end{cases} \quad (5)$$

在实时脑电信号动态处理过程中, 在起始点确定后, 为再次减少眨眼识别的数据量与运算量, 排除脑电信号中非眨眼动作的噪音干扰, 增强数据的有效性与准确性, 对数据分帧是一个重要环节, 眨眼动作的起始端确定后, 还需确定眨眼动作的终止端.

根据经验所得, 人类绝大部分眨眼时间在 0.8 s 以内 (100 个数据点), 在此基础上对 3 名正常人 (平均年龄 23.7) 与运动功能受损但眼部功能完好的 5 名病患

共 8 人进行实验统计, 我们以 160 次眨眼动作脑电数据进行分析, 被试者绝大部分眨眼时间集中在 44–66 (0.352–0.544 s) 之间, 眨眼动作持续时间实验统计图与概率密度函数图如图 4 所示, 其均值为 56.7375, 我们规定起始端到终止端设置为眨眼识别“窗口”, 一次“窗口”为一个眨眼动作识别数据帧, 帧长为 57. 最终的处理脑电数据为 57 个元素的波形向量 ($X = [x_1, x_2, x_3, \dots, x_{55}, x_{56}, x_{57}]$).

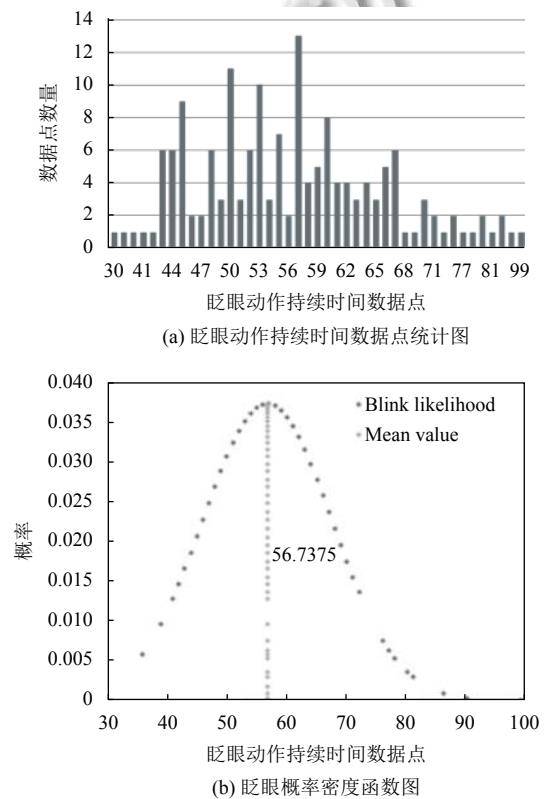


图4 眨眼动作实验统计图与概率密度函数图

(2) 滤波

脑电信号在采集过程中容易收到周围环境和自身生物电因素的干扰, 因此要对分帧后的脑电信号进行滤波去噪处理, 为后续的眨眼波形识别与处理提供可靠的数据支撑. 本设计中采用 0–10 Hz Butterworth 低通数字滤波器对波形数据进行低通滤波^[16], 最终原始波形经过分帧、加窗、滤波后的波形如图 5 所示.

4.3 特征提取与眨眼检测

通过眨眼信号和非眨眼信号对比, 如图 6 所示. 眨眼信号存在明显波峰波谷 (先波峰后波谷) 特点. 根据这种差异, 提出了一种基于波峰波谷特征的波形检测

算法. 将预处理后的波形被转换成能量波形后检测其波峰与波谷 (t_{valley} 与 t_{peak}), 并计算波峰波谷间隔时间 ($d = t_{\text{valley}} - t_{\text{peak}}$) 和能量积累 ($e = \sum_{t=t_{\text{peak}}}^{t=t_{\text{valley}}} (Vt)^2$, Vt 表示 t 时刻当前信号的电压值), 提取此眨眼信号特征阈值 (D_{min} 、 D_{max} 和 E) 进行眨眼波形检测, 其波形检测公式

如式 (6) 所示.

$$b = \begin{cases} 1, & D_{\text{min}} \leq d \leq D_{\text{max}} \text{ and } e \geq E \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

其中, $b=1$ 表示检测到眨眼动作; $b=0$ 表示未监测到眨眼动作.

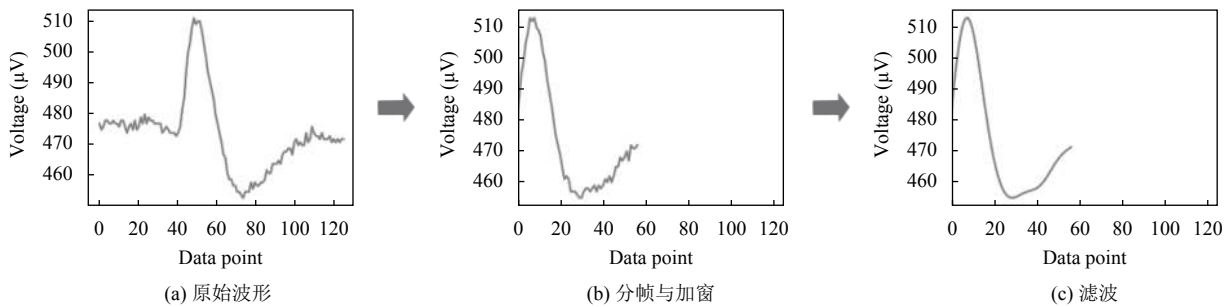


图5 原始波形经过起始段检测、分帧、滤波结果图

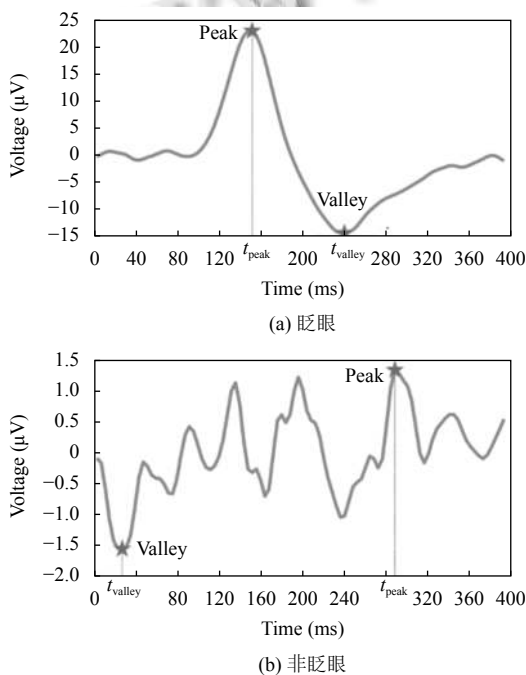


图6 眨眼与非眨眼 EOG 信号波形

待眨眼动作识别处理完成, 将识别分类情况传入脑电紧急呼叫机制, 当接收到眨眼动作为“真”情况下激活 3 s 定时函数, 并记录眨眼次数 $blink_num$, 3 s 内如果眨眼动作次数超过 3 次, 将眨眼紧急呼叫指令发送至控制系统, 反之眨眼次数清零 (防止自然眨眼所导致紧急呼叫), 等待下次眨眼动作激活. 具体眨眼紧急呼叫机制处理如式 (7) 所示.

$$eog_command = \begin{cases} 1, & blink_num \geq 3 \text{ and } time \leq 3 \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

5 系统实验

本次实验分为语音交互实验和眨眼识别实验. 实验环境基本设备: 基于树莓派 4B 脑电分析模块, Linux 环境下, 以 PyCharm 工具为平台, Python 3.8 版本; 基于 STM32F103 系列的云端管理与控制模块和显示播报模块; HNNK 脑机智能生命环; Android 5.0 以上的智能手机终端.

5.1 语音交互实验

语音交互是本设计的核心控制技术之一, 下面分别对语音交互中的语速适应性识别、抗噪音能力以及识别距离进行实验, 探究其对语音交互识别的影响. 本次实验人员为 4 名男性 (平均年龄 23.6), 实验地点为华南师范软件学院 302 实验室. 3 项实验均采用不超过 20 个音节的短句进行实验检测.

(1) 语音语速识别实验

探究不同语速对语音识别所造成的影响. 4 名实验人员在环境噪音为安静 (20–40 dB), 交互距离长度为 0–0.5 m, 语种为标准普通话进行不同语速实验. 根据人的发音速率特性^[24] 及国家标准^[25], 人在正常交流情况下发音速率在 4.0 sps 左右. 为此本实验将语速分为 2.7 sps、

3.2 sps、4.0 sps、5.3 sps 和 8.0 sps 5 个等级, 每名实验人员进行 50 次语音交互, 最终计算平均语音判决时间和识别准确率。

(2) 抗噪音实验

探究不同环境噪音对语音识别所造成的影响。本次以现实环境噪音为根据, 环境噪音等级划分为 3 个等级, 分别为安静 (20–40 dB)、较吵闹 (40–60 dB)、吵闹 (60–80 dB), 4 名实验人员在语速为正常语速 (4.0 sps), 交互距离长度为 0–0.5 m, 语种为标准普通话的基本前提下, 分别 50 次进行不同环境噪音下进行实验, 计算平均语音判决时间和识别准确率。

(3) 识别距离实验

探究不同识别距离对语音识别所造成的影响。本次实验 4 名实验人员在语速为正常语速 (4.0 sps), 噪音环境为初级 (20–40 dB), 语种为标准普通话的基本前提下及将距离分为 4 个等级, 近 (0–0.5 m)、中 (0.5–3 m)、

远 (3–5 m)、超远 (5 m 以上) 进行不同距离分别进行 25 次实验, 并计算最终平均语音判决时间和识别准确率。

(4) 实验总结

综合表 4、表 5、表 6 结果所得, 语速在 3.2–4.0 sps, 识别距离 0.5 m 以内且噪声分贝在 40 dB 以下时, 识别准确度保持在 98% 以上。但语速过快导致采样缺失与识别时间过长, 随之识别错误与识别超时数量增加, 识别准确率下降。同时外界环境噪音的大小与识别准确率成反比, 随着噪音的增加, 外界噪音被误判为人为语音引起音频采集混乱, 识别错误数与识别超时数随之增加, 准确率下降。识别距离的增加, 同样会导致准确率下降。考虑现实常规应用环境, 病房内环境噪声规定最高为 45 dB^[26], 识别距离为 3 m 之内, 语音交互识别平均判决时间小于 1 s, 识别准确率 93% 以上, 完全满足现实需求, 能够投入实际应用。

表 4 语速适应性实验

语速 (sps)	距离 (m)	环境噪音 (dB)	语种	语音平均判决时间 (s)	输入语音数 (句)	正确数 (句)	错误数 (句)	超时数 (句)	准确率 (%)
2.7	0–0.5	20–40	普通话	1.32	200	187	7	6	93.5
3.2	0–0.5	20–40	普通话	0.52	200	199	1	0	99.5
4.0	0–0.5	20–40	普通话	0.62	200	196	3	1	98
5.3	0–0.5	20–40	普通话	0.59	200	184	16	0	92
8.0	0–0.5	20–40	普通话	0.55	200	175	24	1	87.5

表 5 抗噪音实验

环境噪音 (dB)	距离 (m)	语速 (sps)	语种	语音平均判决时间 (s)	输入语音数 (句)	正确数 (句)	错误数 (句)	超时数 (句)	准确率 (%)
20–40	0–0.5	4	普通话	0.59	200	195	5	0	97.5
40–60	0–0.5	4	普通话	1.48	200	172	5	23	86
60–80	0–0.5	4	普通话	2.3	200	139	13	48	69.5

表 6 识别距离实验

距离 (m)	语速 (sps)	环境噪音 (dB)	语种	语音平均判决时间 (s)	输入语音数 (句)	正确数 (句)	错误数 (句)	超时数 (句)	准确率 (%)
0–0.5	4	20–40	普通话	0.63	100	98	2	0	98
0.5–3.0	4	20–40	普通话	0.72	100	93	5	2	93
3.0–5.0	4	20–40	普通话	0.68	100	88	3	9	88
5.0以上	4	20–40	普通话	1.53	100	69	4	27	69

5.2 眨眼识别实验

眨眼识别实验旨在检测系统从脑电波形中识别眨眼动作的能力。在本实验中, 使用 HNNK 智能脑机生命环收集脑电数据, 然后基于本文提出的一种基于波峰波谷特征的波形检测算法进行实时数据处理。实验人员共 10 人, 6 名健康受试者进行了 18 项试验, 4 名患有眼疾的患者进行了 14 项试验。每个试验包含

30 帧, 共检测 960 帧。选取信息检索和统计学分类领域中广泛应用的两个度量值: 召回率 (recall) 和准确率 (accuracy), 用于评价结果质量; 选取精确率 (precision) 来反映预测的眨眼中真正眨眼样本所占的比重。

据统计, 眨眼识别性能结果如表 7 所示, 健康受试者的准确率, 召回率和精确度分别为 94.3%, 90.6% 和 92.1%, 而患者的准确率, 召回率和精确率分别为 82.9%,

77.9%和72.7%。

从检测过程来看,随着眨眼动作的幅度增加,健康受试者和眼疾患者的眨眼波形识别率都会上升。但同时也增加由设备的抖动引起的噪声,导致精度降低。因此,寻找眨眼动作幅度的平衡点是十分必要的。

表7 算法性能结果 (%)

性能	健康受试者	患者
准确率	94.3	82.9
召回率	90.6	77.9
精确率	92.1	72.7

6 结论

本文针对医疗信息化中重要建设目标之一——智慧病房系统,提出一种融合语音和脑电的多模态智慧病房控制系统。该系统采用语音、脑电和手机终端多模态信号作为系统控制源,扩大了系统的使用受众,有效改善了当前智慧病房系统中存在的用户身体条件要求过高、受众有限的问题。通过将物联网控制、语音识别应用到病房护理场景,引入了基于语音的电器控制功能。该功能从患者角度出发,深入挖掘其在治疗护理时期实际需求,有效提高了患者在病房中的自理能力。这一设计出发点,打破了以往智慧病房设计从医院护理者出发的固有思路,值得未来智慧病房设计规划借鉴。

本文提出的智慧病房控制系统,能够提高病房护理中医护工作效率,满足病患的护理中的基本诉求。系统未来可以从以下两个方面完善优化:

(1) 丰富脑电控制功能。本文采用脑电信号进行告警,验证了脑电信号在病房护理中的可行性。但是控制功能单一,未来可以结合病房控制需求设计刺激范式,形成一套丰富的脑电控制系统。

(2) 采用基于自然语言处理的人机对话系统,提升系统语音交互的智能性。本文系统的语音交互功能,需要对固定语音词条的“回复”语音进行提前录入,这就导致语音交互过程中人机对话的单一性,交互的弱智能性。无法实现生动流畅的语音聊天、随机多媒体点播任务,因此与高智能化的人机对话结合是系统智能化不可或缺的一步。

参考文献

1 国家卫生健康委. 2019年我国卫生健康事业发展统计公报. <https://www.sxjgfyxgs.com/24/20981.html>. [2020-06-06].

- 范霁月,景慎旗,王忠民,等. 基于物联网的医院智慧病房实践. 智慧健康, 2020, 6(17): 3-5, 12. [doi: 10.19335/j.cnki.2096-1219.2020.17.002]
- 景慎旗,王忠民,单涛,等. 基于人工智能的智慧病区应用实践. 中国数字医学, 2019, 14(3): 70-72, 30. [doi: 10.3969/j.issn.1673-7571.2019.03.020]
- Nikendei C, Kraus B, Schrauth M, *et al.* Ward rounds: How prepared are future doctors? *Medical Teacher*, 2008, 30(1): 88-91. [doi: 10.1080/01421590701753468]
- Basir OA, Miners WB. Multi-participant, mixed-initiative voice interaction system. US, 8856009. 2014-10-07.
- Perrin X, Colas F, Pradalier C, *et al.* Learning user habits for semi-autonomous navigation using low throughput interfaces. *Proceedings of 2011 IEEE International Conference on Systems, Man, and Cybernetics*. Anchorage: IEEE, 2011. 1-6. [doi: 10.1109/ICSMC.2011.6083633]
- Collinger JL, Wodlinger B, Downey JE, *et al.* High-performance neuroprosthetic control by an individual with tetraplegia. *The Lancet*, 2013, 381(9866): 557-564. [doi: 10.1016/S0140-6736(12)61816-9]
- Hochberg LR, Bacher D, Jarosiewicz B, *et al.* Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. *Nature*, 2012, 485(7398): 372-375. [doi: 10.1038/485317a]
- Zhang R, Li YQ, Yan YY, *et al.* Control of a wheelchair in an indoor environment based on a brain-computer interface and automated navigation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2016, 24(1): 128-139. [doi: 10.1109/TNSRE.2015.2439298]
- Lin JS, Yang WC. Wireless brain-computer interface for electric wheelchairs with EEG and eye-blinking signals. *International Journal of Innovative Computing, Information and Control*, 2012, 8(9): 6011-6024.
- Jafri SRA, Hamid T, Mahmood R, *et al.* Wireless brain computer interface for smart home and medical system. *Wireless Personal Communications*, 2019, 106(4): 2163-2177. [doi: 10.1007/s11277-018-5932-x]
- Guger C, Daban S, Sellers E, *et al.* How many people are able to control a P300-based brain-computer interface (BCI)? *Neuroscience Letters*, 2009, 462(1): 94-98. [doi: 10.1016/j.neulet.2009.06.045]
- Edlinger G, Holzner C, Guger C. A hybrid brain-computer interface for smart home control. *Proceedings of the 14th International Conference on Human-Computer Interaction. Interaction Techniques and Environments*. Orlando: Springer, 2011. 417-426. [doi: 10.1007/978-3-642-21605-3_

- 46]
- 14 Chen XG, Huang XS, Wang YJ, *et al.* Combination of augmented reality based brain-computer interface and computer vision for high-level control of a robotic arm. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2020, 28(12): 3140–3147. [doi: [10.1109/TNSRE.2020.3038209](https://doi.org/10.1109/TNSRE.2020.3038209)]
- 15 Ma JX, Zhang Y, Cichocki A, *et al.* A novel EOG/EEG hybrid human-machine interface adopting eye movements and ERPs: Application to robot control. *IEEE Transactions on Biomedical Engineering*, 2015, 62(3): 876–889. [doi: [10.1109/TBME.2014.2369483](https://doi.org/10.1109/TBME.2014.2369483)]
- 16 Venkataramanan S, Prabhat P, Choudhury SR, *et al.* Biomedical instrumentation based on electrooculogram (EOG) signal processing and application to a hospital alarm system. *Proceedings of 2005 International Conference on Intelligent Sensing and Information Processing*. Chennai: IEEE, 2005. 535–540. [doi: [10.1109/ICISIP.2005.1529512](https://doi.org/10.1109/ICISIP.2005.1529512)]
- 17 Portet F, Vacher M, Golanski C, *et al.* Design and evaluation of a smart home voice interface for the elderly: Acceptability and objection aspects. *Personal and Ubiquitous Computing*, 2013, 17(1): 127–144. [doi: [10.1007/s00779-011-0470-5](https://doi.org/10.1007/s00779-011-0470-5)]
- 18 Jadczyk T, Kiwic O, Khandwalla RM, *et al.* Feasibility of a voice-enabled automated platform for medical data collection: CardioCube. *International Journal of Medical Informatics*, 2019, 129: 388–393. [doi: [10.1016/j.ijmedinf.2019.07.001](https://doi.org/10.1016/j.ijmedinf.2019.07.001)]
- 19 Peng CY, Chen RC. Voice recognition by google home and raspberry Pi for smart socket control. *Proceedings of the 10th International Conference on Advanced Computational Intelligence (ICACI)*. Xiamen: IEEE, 2018. 324–329. [doi: [10.1109/ICACI.2018.8377477](https://doi.org/10.1109/ICACI.2018.8377477)]
- 20 毛博, 徐恪, 金跃辉, 等. DeepHome: 一种基于深度学习的智能家居管控模型. *计算机学报*, 2018, 41(12): 2689–2701. [doi: [10.11897/SP.J.1016.2018.02689](https://doi.org/10.11897/SP.J.1016.2018.02689)]
- 21 许建国, 张佳, 郭麦成. 基于物联网的医院病房智能监护系统设计与实现. *现代电子技术*, 2018, 41(8): 83–86, 91. [doi: [10.16652/j.issn.1004-373x.2018.08.020](https://doi.org/10.16652/j.issn.1004-373x.2018.08.020)]
- 22 Frenzel L. What's the difference between bit rate and baud rate. <https://www.electronicdesign.com/technologies/communications/article/21802272/whats-the-difference-between-bit-rate-and-baud-rate>. (2012-04-28).
- 23 袁廷中. 基于眼电信号的眼部动作的分析与识别 [硕士学位论文]. 杭州: 杭州电子科技大学, 2014. [doi: [10.7666/d.D485076](https://doi.org/10.7666/d.D485076)]
- 24 彭健新. 不同信噪比条件下发音速率对室内汉语语言清晰度的影响. *振动与冲击*, 2010, 29(10): 95–98. [doi: [10.13465/j.cnki.jvs.2010.10.041](https://doi.org/10.13465/j.cnki.jvs.2010.10.041)]
- 25 国家技术监督局. GB/T 15508—1995 声学 语言清晰度测试方法. 北京: 中国标准出版社, 2006.
- 26 中华人民共和国住房和城乡建设部. GB 50118—2010 民用建筑隔声设计规范. 北京: 中国建筑工业出版社, 2011.