

# 基于眼电的字符输入系统<sup>①</sup>

唐秀雯, 李享运, 陈广源, 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](mailto:panjh82@qq.com)



**摘要:** 目前无障碍人机交互系统的设备普遍昂贵和体积较大, 无法真正走入残疾人的日常生活问题. 为了开发出一套低成本、高精度、便携的无障碍字符输入系统, 设计了一套基于眼电的字符输入系统, 提出了一种基于区域优化的闪烁字符交互方式, 并在识别目标字符算法上提出了动态阈值算法进行优化. 实验表明, 该系统实现了使用成本低于 1000 元的设备, 识别准确率达 97.73%, 接近到 100%, 平均字符输入速率达 1.95 个/min 的效果.

**关键词:** 字符输入系统; 低成本; 眨眼; 基于区域的刺激范式

引用格式: 唐秀雯, 李享运, 陈广源, 潘家辉. 基于眼电的字符输入系统. 计算机系统应用, 2020, 29(1): 93-98. <http://www.c-s-a.org.cn/1003-3254/7250.html>

## Character Input System Based on Eye Electricity

TANG Xiu-Wen, LI Xiang-Yun, CHEN Guang-Yuan, PAN Jia-Hui

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

**Abstract:** Nowadays, the barrier-free device in the human-computer interaction system is generally extremely expensive and its volume is large, which fail to really walk into the daily lives of people with disabilities. To develop a set of low cost, high precision, portable barrier-free character input system, this study designs a set of character input system based on EOG (Electro-OculoGram), puts forward a flashing character interaction based on regional optimization, and optimizes the dynamic threshold algorithm for identifying target character algorithm. Experiments showed that the system has realized the identification accuracy of 97.73%, nearly 100%, using equipment which costs less than RMB1000, and average character input rate of 1.95 characters/min.

**Key words:** character input system; low cost; EOG; region-based stimuli paradigm

## 1 引言

### 1.1 研究背景

重度运动障碍疾病患者, 除了视觉和眼周肌肉外, 机体的各运动系统均受到不同程序的损害, 无法进行自主运动, 该类患者基本丧失了与外界交流的能力. 尽管目前的医疗水平在不断提高, 可以帮助运动神经受损的患者延长生命, 但是由于他们不能自主地与外界进行通讯, 生活质量没有得到提高, 也给家庭和社会带来了负担. 在我国作为拥有运动障碍疾病患病人群和

病种数最多的国家之一的背景下, 让重度运动障碍患者重获与外界交流沟通的能力具有更加重要的现实意义.

### 1.2 研究现状

#### (1) 国内外研究现状

在近年来一系列医疗康复相关科学研究突破的推动下, 提出了一系列的无障碍人机交互的解决方案, 使得使用者不依赖肢体的运动实现与计算机的交互. 目前比较成熟的技术有: 基于肌电 (ElectroMyoGram,

① 收稿时间: 2019-06-10; 修改时间: 2019-07-12; 采用时间: 2019-07-24; csa 在线出版时间: 2019-12-27

EMG)的交互技术、基于脑电的脑机接口(Brain-Computer Interface, BCI)技术和基于眼电(Electro-OculoGram, EOG)的交互技术等利用生物电信号的人机交互技术。例如在肌电交互技术中,重庆邮电大学自动化学院的张毅等人实现了基于SEMG控制的智能轮椅无障碍人机交互系统<sup>[1]</sup>。在脑机接口技术方面,浙江大学的邱桐实现了基于脑机接口的智能病床运动控制器研究<sup>[2]</sup>。华南理工大学的李远清、张瑞等人将导航路径规划与脑机接口技术结合起来,实现了轮椅的自动控制<sup>[3]</sup>。在3种交互技术中,眼电交互最易于被接受,主要有3个方面的原因:(1)大部分肢体障碍患者的眼周肌肉功能并没有随之丧失,仍能灵活地使用眼睛。(2)眼电信号时域特征明显,易于被检测,识别正确率较高。(3)信号采集相对容易且采集眼电的设备相较于采集脑电的设备涉及更少电极,在部署采集装置时成本更低、更加方便。因此,在这几种无障碍人机交互方式中,基于眼电的交互系统是一种相对来说较受青睐的方式,更具实用研究价值。

#### (2) 基于眼电的无障碍人机交互研究

眨眼、眼球的运动、眼睛周围肌肉的运动、眼角膜与眼底间的电位差、视网膜上的视觉输入变化、对眼球运动的控制(例如要求受试者不眨眼)等都可以引起眼电。由于眼电图技术具有无创性、环境适应性强、易于检测且信噪比高等优点,在上个世纪90年代已经得到大范围的研究应用。例如,美国凯斯西储大学Chen等人研究出利用眼电信号来操纵机械臂的系统<sup>[4]</sup>,波士顿大学Gips教授的团队研究出利用眼电信号发送邮件,发送信息的配套软件<sup>[5]</sup>。

目前基于眼电信号的字符输入系统主要有清华大学设计的八方向虚拟键盘<sup>[6]</sup>。该系统将8个方向的眼球运动对应8个子菜单,选择了子菜单后,再通过8个方向的眼球运动选择字符输出。浙江大学设计的虚拟键盘<sup>[7]</sup>,界面中间的黄格代表当前选中的字母,使用者利用自己的眼动来控制虚拟键盘的移动,并通过有意识的眨眼来选择字母。还有重庆大学的基于眼电编码的无障碍人机交互系统<sup>[8]</sup>,使用者通过记忆字符对应的眼部动作,在注视目标示意图时通过眼球扫视运动完成字符输入。

上述系统的信息传输速率和正确率均达到了使用水准,但此类方法所开发的相关系统也存在一些问题:(1)此类系统要求使用者注意力高度集中才能准确识

别眼球运动信号,因此很容易使使用者产生疲劳;(2)该类设备的成本相对较高,通常高达上万、甚至几十万元,而重度运动障碍疾病患者家里经济条件普遍困难,很难负担的起这样昂贵费用。

在无障碍人机交互的字符输入系统研究中,很多学者强调通过改善信号处理以及分类技术的来提高字符输入系统的总体性能,例如信息传输率和字符识别的准确率,而改变系统的交互模式往往容易被人们所忽视。本文设计了一款低成本小于1000元的基于眨眼的字符输入系统,并在交互模式上提出基于区域闪烁的优化字符界面,通过动态阈值方法识别眨眼信号,完成字符的输入。

## 2 基于眼电的字符输入系统

### 2.1 系统描述

本文设计了一种基于EOG的字符输入系统,流程框架如图1,帮助患有重度运动功能疾病的病人将脑中所想的信息通过某种辅助手段传达出来。整体设计思路是设计基于区域闪烁的字符表,在目标字符闪烁时要求使用者进行主动眨眼,通过布置在眼睛周围的电极可以记录到由眨眼造成的电位差,信号采集模块采集眼电信号,然后送入微处理芯片,对眼电信号进行滤波处理和眨眼动作检测,最终将眨眼动作解析为命令控制人机交互界面进行字符输入,达到与外界无障碍进行交流的目的。

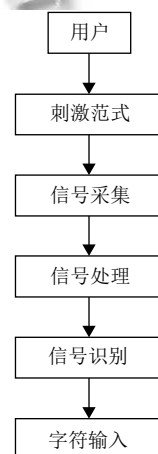


图1 字符输入流程框架图

### 2.2 使用自主眨眼产生的脑电波作为识别信号的基本原理

图2是自主眨眼时的波形图。眨眼产生的电信号高

于正常脑电信号,在额电极和其他相邻电极都可以记录到这一信号,但是其幅值随电极远离额部而迅速减弱。以佩戴 Neuroksy 的 Mindwave 设备,并使用该品牌下的 Mindrec 软件采集到的信号图为例,(横坐标从左往右为一段持续时间,纵坐标代表生物电压值),可以得出眨眼动作产生的电位变化明显、易于检测且信噪比较高。



图2 自主眨眼时的波形图

根据文献[9]显示,自然眨眼和自主眨眼在时间、幅值等方面都有差异,自主眨眼电位的幅度大于自然眨眼电位。此外,时间过程不同,自主眨眼的峰值部分持续时间更长。具体不同表现为持续时间、波值的范围和均值、高峰期开始和结束时间、半幅、幅度等。因此可以根据两者各自的特点进行区分,从而提取出自主眨眼的信号。

### 2.3 刺激范式设计的原理与实现

传统的行列式字符输入系统,其刺激范式采取了字符以行和列为单位按随机次序闪烁和重复多次的 round 行闪的方式。也有单个字符依次闪烁的方式。虽然简单易实现,但整个输入过程耗费较长的时间,同时要求受试者的注意力高度集中,导致受试者视觉疲劳、视线模糊。本文提出的闪烁方式为基于区域的、分时分区的闪烁方式,同时字符类型综合了常用的 26 个英文字母、0-9 共 10 个数字,4 个特殊标点符号。

系统开始运行后,将会进行基于区域的刺激,如图 3。刺激界面被划分为多个区域,每个区域内的字符按一定的随机次序进行闪烁,闪烁任务在区域分时执行,在整个系统分区并行执行,提高了系统的传输比特率。在不同区域并行执行闪烁任务时,相邻两个区域的边缘字符不会同时闪烁,在保证传输速率的前提下减少误差的产生。

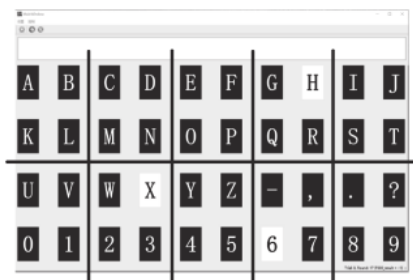


图3 基于区域的闪烁

在实现时,首先对字符表中的每个字符进行编号,在每次生成长度为 len 的闪烁序列时,在已编好号的字符中依次随机取字符编号放入闪烁序列中,并保证新序列首部长度为 interval 范围内的值与旧序列尾部不同。

另外,本文还研究了闪烁物件颜色、闪烁间隔时间、闪烁时间等交互参数对系统的影响。闪烁物件的颜色选择会在视觉诱发实验中起到突出作用,闪烁色与原色彩的色差过大则容易在短时间内产生疲劳,色差过小则可能遗漏闪烁过程。经实验验证,在系统中采用了黑白的闪烁模式。

闪烁时间是指字符处于闪烁的时长,时长太短,留给用户的反应时间也越短,可能导致用户错过目标字符闪烁。闪烁时间过长,则会增大闪烁周期,从而影响字符的传输效率。但是闪烁间隔时间略有不同,当用户选定目标字符,并处于等待字符闪烁时,主观意识是希望目标字符尽快闪烁,即要求字符闪烁间隔越小越好,但当间隔过小时,就会出现同一字符重叠闪烁或相邻字符同时闪烁的情况,不利于判断识别出目标字符。经多次实验调节参数后,选定闪烁时间为 750 ms,闪烁间隔时间为 150 ms。

人机交互的方式为:用户在安静的环境下,佩戴好设备,坐在离电脑屏幕 40 cm 左右的椅子或床上,注视屏幕中字符界面所要输出的字符,在目标字符闪烁时进行主动眨眼,等待系统界面显示出目标字符,便完成字符的输入,操作简单,方便安全。

### 2.4 眼电信号处理流程

#### 2.4.1 眼电信号的预处理

在电信号的采集方面,本文使用 Neurosky 公司开发的 MindWave 设备,尽管设备为单极导联方式,但它采用了与 MindFlex, StarWars Force Trainer, MindTune 以及 NeuroSky 的研究级产品 MindSet 相同的生物传感器。在前额部位(神经科学称之为 FP1 区)能够测量出可用于研究领域的高精度电信号。硬件采集电信号及处理的流程见图 4。

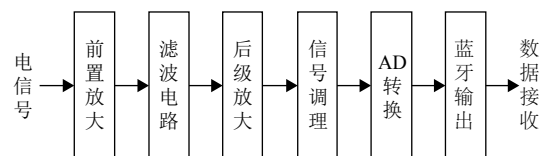


图4 硬件采集系统的框架结构图

MindWave 内置的 ThinkGear 传感器技术通过放置在前额的一个传感器以及放置在耳部的参考电极触



电进行脑电波信号测量,并通过集成芯片对这些信号进行处理.得到的数据值包括信号强弱、专注度指数、放松度指数、原始波值、和眨眼的强度,为后续对采集到的信号进行分析提供了数据支持和便利.

### 2.4.2 动态阈值算法判断目标字符

对每个字符设置标记,根据每次获取的眨眼情况给相应的字符打上标记.字符输出判定条件为所有标记中,标记数量最多的与标记数量次大的满足动态的阈值,即判断标记数量最多的字符为目标字符.动态的阈值将根据用户的个体差异性进行调整.

设标记序列为  $TA$ , 则标记最多的字符的次数为  $TA[max]$ , 次大的字符字数为  $TA[next]$ , 满足目标字符输出的情况如下:

$$TA[max] - TA[next] > Threshold$$

## 3 实验方案

### 3.1 前期准备

实验设备: NeuroSky MindWave mobile, 包含 TGAM (ThinkGear AM) 模块, 最高支持 60 Hz 的采集频率. 受试者在进行字符输入.

系统界面: 使用 Qt Creator 4.3.0 平台开发.

受试者: 男性 9 名, 女性 9 名, 年龄为 18~22 岁的青年, 视力正常或经矫正正常.

实验环境: 通风良好, 光线和温度适宜的安静室内.

每次实验开始之前, 受试者将会有一系列的训练和准备时间, 大约 5~10 分钟, 从而使受试者降低紧张感、提高注意力、了解实验规范并熟悉实验流程. 同时测试人员将对试验方式以及实验环境进行严格的控制, 以保证一些无关因素对眼电信号采集的结果带来干扰.

### 3.2 实验过程

在缺乏使用同一价位设备的参考文献, 和与应用途径相似的文獻有巨大的设备成本差异, 在精度和性能上可比性较低的情况下, 为了验证本文提出的优化交互方式和优化算法的效果, 实验依照控制变量的方法进行. 首先控制系统使用相同优化算法识别目标, 以交互方式作为变量, 受试者使用不同刺激范式的交互方式做输入相同字符的 3 组实验, 实验分为 3 天进行. 依次进行刺激范式为单次闪烁、行闪烁与列闪烁、基于区域闪烁的实验. 其次控制系统的交互方式为优化后的刺激范式, 以目标识别算法为变量, 分别采用优化

过的算法和未优化过的算法进行另外两组实验, 未优化过的识别算法为检测到目标字符立即输入, 不做重复多次的验证.

测试人员完成设备与 PC 机的蓝牙连接, 受试者佩戴好设备并坐在指定位置上后, 测试人员开启系统. 屏幕上出现上文描述的黑白字符表, 字符表中的字符开始闪烁. 受试者心中默认选择想要输入的字符, 等待目标字符闪烁, 并在目标字符闪烁时做一次自主眨眼. 等待系统界面出现这一轮的目标字符, 即可输入下一个目标字符. 在整个实验过程中, 受试者无需刻意抑制自然眨眼.

### 3.3 评估标准

无障碍人机交互系统的性能受多种因素的影响, 诸如硬件条件、系统范式、应用类型等. 因此, 不同类型的电交互系统以及同一电交互系统的不同应用, 评价标准也不尽相同.

为了全方位地评估本文研究的基于眼电的字符输入系统, 综合考虑各方面的因素, 并参照相关文献中的评价标准, 采用了如下评价指标

#### (1) 灵敏度

灵敏度用于评价 NeuroSky 的 MindWave 设备在识别眨眼产生的电位差的能力以及蓝牙传输数据的响应能力. 由于本文采用的设备需要通过蓝牙与系统连接, 那么灵敏度就是系统性能基本保障.

$$\text{灵敏度} = \frac{\text{系统识别到的眨眼次数}}{\text{用户总的眨眼次数}}$$

#### (2) 字符输入准确率

字符输入准确率是系统的基本性能指标, 用户在目标字符闪烁时自主眨眼且被识别出得字符为目标字符记为正确情况, 其他情况下输出的字符均无效.

$$\text{字符输入准确率} = \frac{\text{正确情况下输出的字符}}{\text{屏幕输出的字符}}$$

#### (3) 有效字符输入速率

有效字符输入速率是系统的重要性能指标, 它表示输入一个正确目标字符花费的平均时间. 与同类字符输入系统相似, 本文用在线实验结果的数据信息传输速率来衡量字符输入系统性能的好坏. 该指标不仅在分析实验结果有利, 在用户的真实实时使用场景下更是根本指标.

$$\text{有效字符输入速率} = \frac{\text{受试者人机交互的总时间}}{\text{正确输入的目标字符数}}$$

#### (4) 视力负荷

参考脑力负荷, 本文提出视力负荷来评价受试者完成系统的字符输入时, 视力的疲劳程度. 一方面这是用户体验的直接反映, 用户需要通过眨眼来完成字符输入, 因此在长时间的输入过程中, 会带来视觉疲惫等不可避免的影响, 视力负荷的高低可以直接反映用户体验的优良, 另外可以通过比较不同系统的视力负荷, 直观地看出本文提出的新型交互方式的在用户体验上的优越性. 本文采用 NASA 开发的 NASA-TLX 来对工作量进行多维度的评估, 主要包括脑力需求、体力需求、时间需求、绩效、努力程度、受挫程度六个指标<sup>[10]</sup>. 评价每个指标时, 将给出一条 20 等分的直线, 用户在完成实验后, 根据自己的感受, 在每条直线上勾选出相应的分数, 分值越高代表程度越大. 然后在 15 个指标对子中选出自己认为的与工作负荷关系较为密切的指标.

每一指标被选中的次数 (0-5 次) 表示该指标在总工作负荷的权重. 将直线上的位置对应数字 (0-20) 与对应权重相乘即为每个指标的评分, 将所有评分累加后除以总权重 15 即为这个字符输入任务的最后评分. 评分越高表示用户完成字符输入的视力负荷越高.

## 4 实验结果及效果分析

控制目标识别算法不变, 使用不同交互方式的效果图 5 所示, 其中  $x$  轴表示实验者编号,  $y$  轴表示准确率.

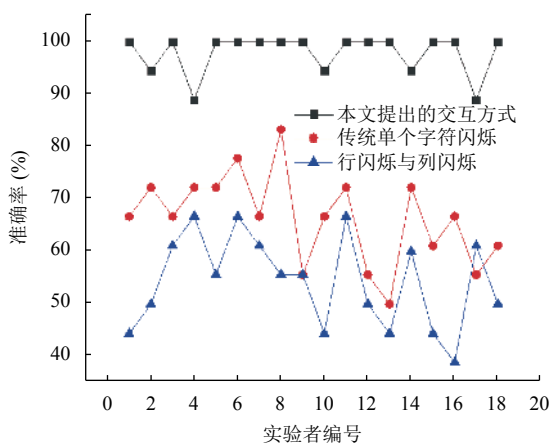


图 5 不同刺激范式下的准确率

可以看到, 本文提出的交互方式在准确率上有了很大的提升, 平均准确率达 97.73%, 虽然个别准确率

不足 90%, 但整体看来基本维持在较高准确率水平. 而传统的单个字符闪烁和行闪烁与列闪烁的交互方式则只得到较低准确率, 前者的平均准确率为 66.37%, 后者的准确率为 54.26%, 且准确率水平很不稳定, 表明容易受到外界或实验者自身干扰.

在输入速率和用户视力负荷上, 本文的交互方式也有明显优势, 见图 6 和图 7.

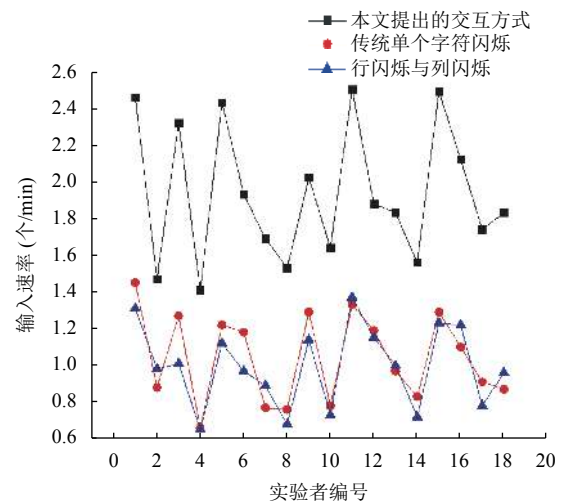


图 6 不同刺激范式下的字符输入速率

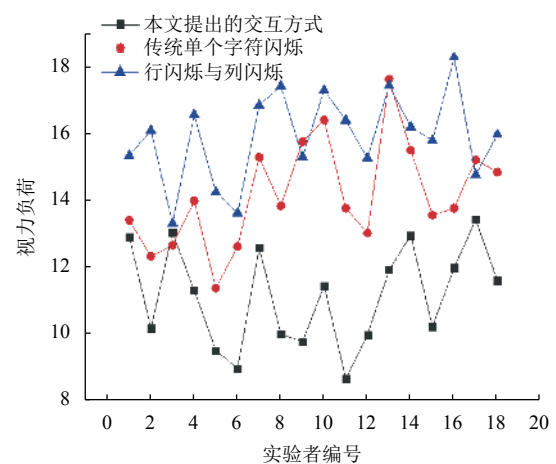


图 7 不同刺激范式下的用户视力负荷

输入速率与视力负荷基本成负相关, 输入速率高时, 用户花费更少的时间, 甚至可以更轻松完成输入, 因此视力负荷会较低. 本文提出的交互方式获得了 1.95 个/min 的平均字符输入速率, 传统单个字符闪烁则为 1.05 个/min, 行闪烁与列闪烁为 1.01 个/min, 在输入速率有了大幅提升的同时, 并不会给用户带来更重的视力负荷, 本文提出的交互方式的视力负荷为 11.16,

传统单个字符闪烁则为 14.87, 行闪烁与列闪烁为 15.99. 传统的两种交互方式的实力负荷大致在 15–20 范围内, 而新型的交互方式大致在 10–15 范围内, 依照 NASA-TLX 的评价标准来看, 新型交互方式在用户体验上实现了跨阶段的优化提升.

在应用本文提出的新型交互方式的前提下, 以目标识别算法作为变量的两组实验的效果如图 8 所示.

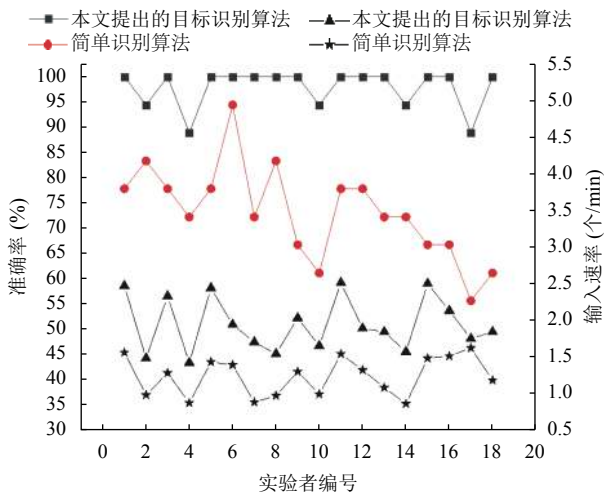


图 8 不同目标识别算法下的准确率和输入速率

简单识别算法为, 检测到目标字符立即输出, 没有任何二次的判断. 平均准确率为 73.15%, 字符输入速率为 1.23 个/min (这里指有效字符), 相较于本文的优化的识别算法获得的 97.73% 平均准确率和 1.95 个/min 平均输入速率, 可以验证优化的目标识别算法实现了更高速和更精确的字符输入.

由于上述所有组别的实验都使用同一设备进行, 因此没有单独测量每组实验的平均灵敏度, 只针对使用了优化目标识别算法和改进的交互方式的实验组做了一次测量. 平均灵敏度为 81.85%, 可知设备与系统的蓝牙连接不能保证百分百的数据传输, 这就解释了个别数据的急剧变化, 或者出现异常数据的原因. 而在这种情况下, 本文提出的字符输入系统仍能获得 97.73% 的准确率, 其优越性显而易见.

## 5 结束语

本文主要研究了基于眼电的字符输入系统. 由于

重度运动患者即使丧失了机体大部分的运动能力, 但是眼部仍然可以灵活运动, 所以选择了由眨眼产生的眼电作为识别信号, 并且眼电在时域和频域上的特征明显, 易于检测. 在交互方面, 提出了基于区域优化的字符闪烁交互方式. 对实验结果进行分析后可以看出, 该交互方式使得在采用精确度不足的低成本设备时, 系统仍能获得较高的字符输入速率和识别准确率, 并让用户疲劳程度处于可接受范围内, 对于帮助家庭困难的重度运动障碍患者获得与外界交流的能力具有重要现实意义.

## 参考文献

- 张毅, 代凌凌, 罗元. 基于 SEMG 控制的智能轮椅无障碍人机交互系统. 华中科技大学学报 (自然科学版), 2011, 39(S2): 264–267, 282.
- 邱铜. 基于脑机接口的智能病床运动控制器研究[硕士学位论文]. 杭州: 浙江大学, 2019.
- 张瑞. 面向重度残疾人的脑机接口功能辅助研究[博士学位论文]. 广州: 华南理工大学, 2016.
- Chen YX, Newman WS. A human-robot interface based on electrooculography. Proceedings of 2004 IEEE International Conference on Robotics and Automation. New Orleans, LA, USA. 2004. 243–248.
- Gips J, Olivieri P. EagleEyes: An eye control system for persons with disabilities. Proceedings of the 11th International Conference on Technology and Persons with Disabilities. Los Angeles, CA, USA. 1996. 1–15.
- 郑敏敏. 基于眼电信号检测的人机接口研究[硕士学位论文]. 北京: 清华大学, 2013.
- 李昕. 基于眼电的无障碍人机交互技术研究[硕士学位论文]. 杭州: 浙江大学, 2010.
- 温绍龙. 基于眼电编码的无障碍人机交互系统研究[硕士学位论文]. 重庆: 重庆大学, 2017.
- Lins OG, Picton TW, Berg P, et al. Ocular artifacts in EEG and event-related potentials I: Scalp topography. Brain Topography, 1993, 6(1): 51–63. [doi: 10.1007/BF01234127]
- Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. Advances in Psychology, 1988, 52: 139–183. [doi: 10.1016/S0166-4115(08)62386-9]