

## 移动环境下神经功能评价多通道交互模型<sup>\*</sup>

黄进<sup>1,2</sup>, 陈毅能<sup>1,2</sup>, 刘杰<sup>1,2</sup>, 田丰<sup>1</sup>, 戴国忠<sup>1</sup>, 王宏安<sup>1</sup>

<sup>1</sup>(人机交互北京市重点实验室(中国科学院 软件研究所),北京 100190)

<sup>2</sup>(中国科学院大学 计算机与控制学院,北京 100049)

通讯作者: 田丰, E-mail: tianfeng@iscas.ac.cn



**摘要:** 随着平板电脑、智能手机、智能手表等智能移动设备的普及,利用便携的智能移动设备随时随地进行健康评价受到了国内外学者的广泛关注.人机交互特有的多通道、交互式、人机协同的计算能够有效地提高移动环境下神经功能评价的准确度.然而,目前很少有研究对人机交互在这一应用场景上发挥的作用进行过充分讨论,更没有形成统一的多通道交互模型.为此,首先分析了目前移动设备上主流的神经功能评价方法,归纳总结出了一套适用该应用场景的交互原语和交互任务.然后,在此基础上提出了移动环境下神经功能评价多通道人机交互模型——MINA(multimodal human-computer interaction model for nerve function assessment in mobile environment),并对该模型的移动医学评价和多通道融合特点进行了分析.最后,依据此模型给出神经系统疾病检测的应用实例.实践证明,MINA能够较好地指导交互式神经功能评价应用的设计和开发,多通道融合的方式能够有效地提高医学评价的准确度.

**关键词:** 移动环境;神经功能评价;人机交互;多通道交互;交互模型

中文引用格式: 黄进,陈毅能,刘杰,田丰,戴国忠,王宏安.移动环境下神经功能评价多通道交互模型.软件学报,2016,27(Suppl. (2)):156-171. <http://www.jos.org.cn/1000-9825/16031.htm>

英文引用格式: Huang J, Chen YN, Liu J, Tian F, Dai GZ. Multimodal human-computer interaction model for nerve function assessment in mobile environment. Ruan Jian Xue Bao/Journal of Software, 2016,27(Suppl.(2)):156-171 (in Chinese). <http://www.jos.org.cn/1000-9825/16031.htm>

## Multimodal Human-Computer Interaction Model for Nerve Function Assessment in Mobile Environment

HUANG Jin<sup>1,2</sup>, CHEN Yi-Neng<sup>1,2</sup>, LIU Jie<sup>1,2</sup>, TIAN Feng<sup>1</sup>, DAI Guo-Zhong<sup>1</sup>, WANG Hong-An<sup>1</sup>

<sup>1</sup>(Beijing Key Laboratory of Human-Computer Interaction (Institute of Software, the Chinese Academy of Sciences), Beijing 100190, China)

<sup>2</sup>(School of Computer and Control Engineering, University of Chinese Academy of Sciences, Beijing 100049, China)

**Abstract:** With the growing popularity of smart mobile devices such as tablet PCs, smart phones and smart watches, the use of portable smart mobile devices in health evaluation has received extensive attention of scholars at home and abroad. The characteristics of multimodal, interactive and collaborative computing in human-computer interaction (HCI) can effectively improve the accuracy of nerve function assessment. However, little research has been done on the important role of human computer interaction in medical assessment, and no model of nerve function assessment in mobile environment has been established. Therefore, this paper firstly analyzes the mainstream method of nerve function assessment in mobile environment, and summarizes a set of interaction primitives and interaction tasks for the application on this scenario. Then it proposes a multimodal human-computer interaction model for nerve function assessment in mobile environment: MINA, and analyzes the characteristics of mobile health assessment and multimodal fusion in MINA. Finally, one

\* 基金项目: 国家自然科学基金(61232013,61422212)

Foundation item: National Natural Science Foundation of China (61232013,61422212)

收稿时间: 2016-11-29; 采用时间: 2016-12-05

application that builds on MINA has been demonstrated. The practice shows that MINA can better guide the development of interactive nerve assessment applications, and the multimodal fusion can effectively improve the accuracy of medical assessment.

**Key words:** mobile environment; nerve function assessment; human-computer interaction; multimodal interaction; interaction model

随着平板电脑、智能手机、智能手表等智能移动设备的普及,利用便携的智能移动设备随时随地进行健康评价受到了国内外学者的广泛关注<sup>[1-9]</sup>.移动设备的强处理能力、强通信能力、便携性和普适性使其在许多情况下比其他医疗设备更具竞争力,随着移动设备数量的快速增长,基于移动设备的医疗应用已经成为最重要的日常医疗监护手段之一.

神经系统疾病(nervous system disease,简称 NSD),例如帕金森病(parkinson's disease,简称 PD)、阿兹海默症(alzheimer's disease,简称 AD)、亨廷顿病(huntington's disease,简称 HD)、小血管病(small vessel disease,简称 SVD)等,是一种常见的导致认知功能和运动功能障碍的疾病,神经功能评价对 NSD 的检测和早期预警有重要意义.临床上主要是通过各种量表、测试、问卷调查等方法对病人的神经功能进行评价.但是这些方法依赖于专业医疗设备和医疗人员,成本较高,无法作为日常健康评价手段.

在人机交互、普适计算研究领域,已经有越来越多的学者开始利用人机交互技术在移动设备上辅助医疗诊断和健康评价.一类方法利用电子化调查表获取用户的健康状况,发现问题,及时帮助用户解决健康问题<sup>[11,10,11]</sup>.这类方法要求用户自觉填写调查表,数据的完整性和准确性得不到保证,即使收集到了完整的数据,其数据分析和结果反馈仍需要专业医疗人员的参与,这不仅没有充分利用移动设备的本地计算资源,还极大地增加了此类应用的实施成本.

随着移动互联网的快速发展,另一类利用多媒体社交网络、众包和协同计算来解决健康问题方法也得到了广泛关注<sup>[3,12,13]</sup>.这类研究充分发挥了移动计算设备的网络功能,基于社会认知理论,能够很好地发现用户群体存在的健康问题,驱动个体健康行为.然而,这类方法无法对个人健康状态或生理指标进行准确地检测,其提供的健康建议来自于社会群体认知,并非专业的医学建议,很难给患者提供专业、及时、有效的帮助.

随着移动设备计算能力的提高,越来越多的学者开始关注基于生理计算的健康评价技术<sup>[2,4,14-16]</sup>.这类技术能够充分利用移动设备的本地计算资源,快速地得到精确量化的生理参数,发现异常体征,进行健康评价和辅助诊疗.一类上述技术仅使用主体设备本身便能完成健康评价,而另一类则除了使用主体设备外,还需要借助外部设备或额外添加传感器,例如文献[2,17]中的评价方法.本文提出的模型是针对第2类技术的,其主要的原因是,外部设备的加入增加了用户认知负荷,提高了使用难度,并有可能大额增加制造成本,使得它不容易推广和普及,很难成为日常的健康评价手段.

在基于生理计算的健康评价技术中,已经有一些专门针对神经功能评价的研究.Graça 等人在文献[5]中基于智能手机开发了多项应用,提取了手部和步态等多项生理特征,使用机器学习的方法对帕金森病人和正常人进行分类,其分类的准确率达到 87.5%.Michael 等人提出了一个在智能手机上区分帕金森和特发性震颤的方法<sup>[18]</sup>,该方法利用离散小波变换和支持向量机(support vector machines,简称 SVM)对受试者进行分类,能够对两种相似的震颤症状进行准确区分,其准确率达到 96%.文献[6]中介绍了一种利用语音识别对阿兹海默类痴呆病的早期诊断方法,该研究多痴呆病人的词法进行了分析,提取了痴呆病人的词法特征,给出了各种词法特征的数据分析结果.Albert 等人在手机上检测帕金森病人的活动状态,将行走、站立、坐下、手持手机和没有携带手机 5 个状态进行区分,达到了 92.2%的准确率,为病人全天候健康监测提供支持<sup>[19]</sup>.文献[9]同样利用了智能手机感知用户的活动状态,通过对步态特征的分析 and 识别,能够在帕金森病人患病早期便提出预警,为病人尽早求医或医生早期诊断提供帮助.

然而,上述工作往往是一些单一的应用,没有形成统一的模型,研究人员在研究和开发一个新的应用时,需要大量的重复工作,提高了开发门槛和成本.为此,文献[20]提出了一种基于智能手机的新移动医疗系统模式,该研究试图找到一个能够有效整合各类信号的采集、分析和网络通信能力的系统模式,简化移动健康评价系统的设计.该研究中关于移动设备的自组织网络是具有借鉴意义的,但是,在本地数据处理方面,该研究仅讨论了脉搏一项数据,没有对多通道数据的组织和处理进行讨论,无法适应神经功能评价的要求.

归纳起来,首先,尽管人机交互技术已经在上述的许多研究中使用,但很少研究对人机交互在神经功能评价上发挥的作用进行过充分讨论,更没有形成一个统一的交互模型.其次,在研究人机交互应用时,交互原语的定义、捕获、解释和处理,是有效应用交互技术的关键因素<sup>[21,22]</sup>,然而当前尚缺乏移动环境下神经功能评价的人机交互原语和任务设计,没有统一的交互表达标准,研究者们都仅根据自身的设定和交互环境进行研究,研究成果之间可比性较差.再次,多通道交互模式已被诸多研究证明是提高交互效率和自然性的有效途径,然而先前的研究中,并没有对多通道交互在神经功能评价中的作用进行详细讨论,多通道交互的优势没有得到体现.

所以,本文认为,有必要构建一个有统一交互原语和任务标准的,能够融合多通道交互信息和评价参数的人机交互模型.通过该模型,开发者能够很好地组织各项生理参数采集和分析,充分发挥人机协同计算优势,考虑上下文信息进行更加精确的计算和分析,为移动环境下的神经功能评价应用设计提供理论依据.

为此,本文首先分析了目前移动设备上主流的神经功能评价方法,归纳总结出了一套面向神经功能评价的交互原语和交互任务,提出了移动环境下神经功能评价的多通道人机交互模型 MINA(multimodal human-computer interaction model for nerve function assessment in mobile environment),下文简称 MINA,给出此模型的交互信息处理流程,并对基于本模型的交互组件设计方法进行深入讨论,最后,给出一个使用本模型搭建的 NSD 检测实例.实践证明,MINA 能够较好的发挥移动设备的特点,满足日常神经功能评价应用的设计需求,能够为此类应用的开发与评估提供理论依据.

## 1 多通道交互原语与交互任务

交互原语(interaction primitive,简称 IP)是用户通过交互设备到计算机的一个词法输入,表示一个独立的、最小的、不可分割的操作<sup>[21]</sup>.交互任务是由原语组成,根据应用场景需求归纳总结的通用任务集.移动环境下神经功能评价的多通道交互模型与传统的基于交互范式的交互模型不同,是一种面向特定应用场景的交互模型.所以在进行原语设计时和任务归纳时,必须考虑移动设备所能提供的硬件支持以及对神经功能评价具有普遍意义的交互任务.

### 1.1 交互原语分析

目前主流的移动设备配备的硬件支持包括触摸屏、摄像头、话筒、陀螺仪和加速度传感器,一些研究在移动设备上添加新的传感器以增加对用户生理特征的感知,其中包括压力<sup>[23]</sup>、指尖空间位置<sup>[24]</sup>等.结合考虑脑神经评价应用具有普遍意义的交互任务,本文将交互原语定义为以 9 大类,见表 1,原语的表达方法见文献<sup>[21]</sup>中使用的人机交互原语描述方法,即将原语表示为四元组: $IP_A(\text{Type}, \text{Para}, \text{Time}, \text{Device})$ .其中,Type 表示交互原语 A 所对应的原语类型,Para 表示原语的参数,Time 为原语发生的时间,Device 表示产生原语的硬件设备.Device 参数的分类和描述参照 Machinlay<sup>[25]</sup>的分类方法.基于分层的思想,本文将交互原语分为基本原语和派生原语,基本原语为可由输入设备的数据结构直接表示的交互操作,其包括“点”“文本输入”“数质量输入”“单帧拍摄”及“声音”,派生原语则可能由一个或多个基本原语、其他派生原语以及上下文信息组成,其包括“手势”“多指手势”“多帧拍摄”.虽然派生原语从数据结构上可以分解,但在交互意义上是不能分割的,例如手势原语,虽然可以由触屏点序列组成,但在交互操作上不可分为若干个点选操作,所以它仍为原语的范畴.

“点”是最常用的交互原语之一,其参数的第 1 部分是三维坐标,其坐标系是以移动设备主屏幕为平面,与主平面垂直的轴为轴的笛卡尔坐标系,这样做是为了用统一的一个原语表达平面和空间的两种点交互,如需要表示平面点交互,限定轴坐标为 0 即可.“点”原语参数的第 2 部分表示压力,当设备支持压力传感,并且用户指尖位于屏幕表面时,该值会产生读数.目前,一部分主流移动设备配备了数字笔,能够支持笔交互,另一部分设备只能使用指尖进行触屏交互,然而这两种交互在原语的表达上是一致的,所以这里仅使用“点”原语即可完成对这两种交互的统一表达,同时利用原语产生设备参数加以区分.

“手势”是一个派生原语,由基本原语“点”的序列组成,“手势”原语的发生时间与其组成序列中第 1 个“点”的时间一致.

**Table 1** Interaction primitive for nerve function assessment in mobile environment

**表 1** 移动环境下神经功能评估的人机交互原语

原语类型	原语表达
点	$\langle Point, \langle \langle x, y, z \rangle, p \rangle, Time, Device \rangle$
手势	$\langle Gesture, \langle \{ Point_i   i=1, 2, \dots, n \} \rangle, Time, Device \rangle$
多指手势	$\langle MultiGesture, \langle \{ Gesture_i   i=1, 2, \dots, n \} \rangle, Time, Device \rangle$
文本输入	$\langle Text, \langle textContent \rangle, Time, Device \rangle$
数量值输入	$\langle Quantify, \langle value \rangle, Time, Device \rangle$
单帧拍摄	$\langle Image, \langle I = \{ a_{ij} = \langle r, g, b \rangle \}_{n \times m} \rangle, Time, Device \rangle$
多帧拍摄	$\langle Video, \langle \{ Image_i   i=1, 2, \dots, n \} \rangle, Time, Device \rangle$
声音	$\langle Voice, \langle soundWave \rangle, Time, Device \rangle$
空间姿态	$\langle Posture, \langle \{ p_i   i=1, 2, \dots, n \} \rangle, Time, Device \rangle$ $p = \langle \langle x, y, z \rangle, \langle w, v \rangle, t \rangle$

“多指手势”也是一个派生原语,它由至少 2 个以上的“手势”原语组成,组成它的“手势”原语的时间顺序是任意的,具体形式由识别算法的匹配模板决定,“多指手势”的发生时间为组成它的所有“手势”原语的最早发生时间。

“文本输入”和“数量值输入”是两个已经被广泛使用的传统人机交互原语,这两个交互原语通常用于获取那些已被用户掌握的生理参数,例如年龄、性别、腰围等.与“点”“手势”“多指手势”一起,这 5 个交互原语覆盖了所有基于移动设备屏幕的交互行为。

“单帧拍摄”即使用移动设备摄像头摄于一帧图像,由它派生的“多帧拍摄”即为使用摄像头拍摄一段视频.这两个原语的设计为基于摄像头的脑神经特征感知提供支持。

“声音”原语即使用移动设备配备的话筒录取一段音频,在一些神经功能评价应用中,使用话筒录取用户语音,另一些应用则录取环境声音用于感知用户的当前状态和行为。

“空间姿态”原语的参数是一个姿态的序列,姿态由 3 部分组成,第 1 部分表示移动设备在空间中的位置,第 2 部分是一个四元数,用于表示移动设备在各方向的旋转角度,第 3 部分是采样时间.本文没有和“点”原语一样,将姿态也单独定义为一个基本原语,这是因为设备在空间中单独的某一瞬时的姿态信息是没有意义的,只有当这些姿态组成一个连续序列,才对神经功能评价有意义,例如,一个保持稳定的设备姿态和一个不停震动的设备姿态。

**1.2 交互任务归纳与描述**

通过调研我们发现,目前主流的在移动环境下进行的神经功能评价方法中,使用到的交互任务可以归纳为以下 6 大类:WIMP(window, icon, menu and pointer)交互任务、平面绘图任务、空中手势任务、拍摄任务、录音任务、携带或手持设备任务,见表 2.这里,本文仅给出交互任务的顶层分类,在实际应用中,可根据上下文信息 and 应用需求,对交互任务进一步细分。

**Table 2** Summary of interaction tasks for nerve function assessment in mobile environment

**表 2** 移动环境下神经功能评价的人机交互任务归纳

描述	主要原语	举例
传统的 WIMP 范式下交互任务,例如点击、选择、拖拽、释放、放缩、旋转、输入信息等	点、手势、多指手势、文本输入、数量值输入	电子问卷的填写; 评估报表的查阅; 个人信息的输入
使用触屏或笔设备在平面上绘制指定或任意图形;使用同样方式做一个指定或任意的手势	点、手势、多指手势	绘制一个与示例一样的正方形; 比划一个“正确”(✓)手势
使指尖或笔设备在空中做一个指定或任意的手势	点、手势、多指手势	将手指保持在屏幕正上方不动; 用笔在屏幕正上方划出一条尽量直的线
使用移动设备摄像头拍摄一帧或连续多帧图像	单帧拍摄、多帧拍摄	将手指按压住摄像头,并拍摄一帧图像; 使用摄像头对着自己拍摄一个微笑的动作视频
使用移动设备麦克风录制一段声音	声音	说出“吃葡萄不吐葡萄皮”并录制下来
将移动设备固定于身体的特定位置,并执行指定或任意的动作;手持移动设备,并执行指定或任意动作	空间姿态	将移动设备至于腰间,正常行走 10 步; 手持移动设备,向前伸直,保持 10s

WIMP 交互任务是对传统图形用户界面(graphic user interface,简称 GUI)交互任务的归纳,它们是基于“点”“手势”“多指手势”“文本输入”和“数量值输入”原语的.前文提到的基于电子化调查表、多媒体社交网络的评价方法中使用的交互任务,都属于 WIMP 交互任务.基于 WIMP 范式的交互界面虽然有许多缺点,但仍是当今主流的交互手段,在神经功能评估中,我们主要利用 WIMP 获取用户信息,组织各类交互任务,发布评估反馈等.

平面绘图任务是神经功能评价中经常使用的一项交互任务,该任务对用户的认知损伤和运动损伤有很强的感知能力,用户在平面绘图任务中的总体表现、图形形态等方面能够很好的反映认知功能情况,而绘图轨迹的细节特征则能很好的刻画运动功能情况,文献[5,7,26]都利用这一交互任务对 NSD 的相关症状进行了评估.因此本文将其单独列出来进行讨论.

平面绘图任务包括绘制图形和比划手势,这两者的评价功能是类似的,仅在表现形式和交互隐喻上有所区别.需要指出,平面绘图中的手势与 WIMP 中的手势是有交集的,然而它们的设计目的和处理方式则是截然不同的:WIMP 中的手势是为了便于界面操作而提供的一组固定的简单手势集,其处理结果是 GUI 的操作结果;平面绘图中的手势是为了激发用户的认知和运动损伤症状而设计的,处理结果应当是用户执行手势的表现和细节特征.

空中手势任务是针对具有指尖空中位置感知的移动设备设计的,在对帕金森一类病症诊断时,利用该技术感知用户手部的静止性或运动性震颤是非常有效的.

与“拍摄”和“声音”原语相对应地,拍摄任务和录音任务是针对利用摄像头和麦克风进行神经功能评价而归纳的交互任务.例如,文献[6]中利用用户语音信息检测痴呆,文献[13]则利用手机拍摄的视频帮助阿兹海默病人恢复记忆力.

携带或手持设备任务设计的目的主要是面向基于“空间姿态”原语的神经功能评估应用,如文献[5,9].一些这类应用是隐式的,即不需要用户进行显示的交互行为,而只需要在身上携带,或者是手持设备进行其它交互行为,例如输入文本,移动设备即实时地采集和分析空间姿态信息,感知用户状态,例如判断用户是否跌倒.另一些这类应用是显示的,它们可能要求用户手持移动设备作一些标准动作,并从这些标准动作的空间姿态上提取信息.利用标准动作的好处是数据一致性较好,能够在很大程度上降低处理算法的复杂度.

综上所述,交互原语体现了移动设备上的硬件支持,交互任务由原语组成,通过不同的原语组合,能够形成上述通用的交互任务集,对面向的应用场景,即神经功能评价提供支持,该任务集从顶层划分的角度对目前主流的神经功能评价交互任务进行了归纳,可以适用于绝大多数的评价方法.本节分析归纳出的交互原语和交互任务集既是 MINA 中使用的原语和任务集合.

## 2 多通道交互模型

### 2.1 层次结构划分

为了将移动设备中不同通道的信息整合统一处理,利用交互原语和基本交互任务串联设备数据与应用程序,同时考虑上下文信息进行更加精确的计算和分析,本文首先对移动环境下神经功能评价的多通道交互系统的层次结构进行划分,如图 1 所示,再基于此层次结构,结合传统人机交互体系结构特点,得出本文的多通道人机交互模型——MINA.

#### (1) 硬件层

移动设备包括触摸屏、摄像头、话筒、陀螺仪和加速度传感器、压力和指尖空间位置传感器等硬件设备.

#### (2) 交互原语分析层

交互原语分析层的功能是分析硬件层输出的原始信息,生成统一的与设备无关的交互原语.本层对每个通道的信息进行词法分析,从而向任务产生层提供与设备无关的交互信息,这些信息使用原语的形式表达,它是交互信息流中的基本单元,具有不可再分性.

(3) 交互任务识别层

任务产生层的功能是根据事先规定好的任务表项,对上一层输出的交互原语进行组合与识别,得到成功执行地交互任务.本层将上一层得到的交互原语按照一定的语法规则融合成具有任务语义的交互信息,所以本层实际上是通过任务制导的语义融合得出交互任务的.在上一节中,本文对这交互任务进行了一个总的分类,但由于神经功能评价方法的多种多样,具体交互任务的识别还需要依据自定义的任务语法而定,因此,交互任务语法集应是一个可扩展的集合.

(4) 数据处理和分析层

本层的功能是对上一层得到的交互任务信息进行进一步的分析和识别,输出对神经功能评价具有特定意义的参数.这一层包含很多数据处理、统计分析和机器学习算法,在具体实现当中,本层除了提供一些通用的处理和识别算法外,还应提供用户拓展的接口,即数据处理和分析算法集合也是一个可拓展集合.

(5) 医学评价分析层

这一层的功能是综合上一层得到的若干评价参数,进行相关的医学分析,得到最终的输出结果.这些结果的获取是需要综合上一层得到的多个评价参数并结合神经功能评价的相关知识才能得到,所以本层必须对上一层得到的多个评价参数进行融合.本层输出的即可能是一个诊断结果,如“是否患有帕金森病”,也可能是一个评价指标,如“简易智能状态检查(mini-mental state examination,简称 MMSE)的得分”,具体输出哪些结果应是由应用程序决定的,所以它应与应用层最为接近.

(6) 应用层

神经功能评价软件业务及界面设计层,在前面各层的支持下,开发者在本层只需要专注于软件的业务逻辑.

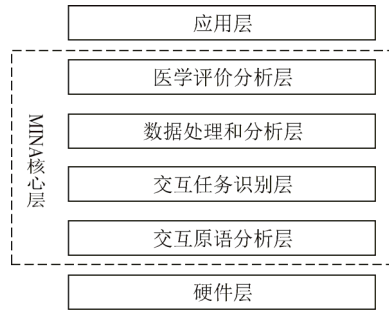


Fig.1 Layered structure of interactive system for nerve function assessment in mobile environment

图 1 移动环境下神经功能评价交互系统分层结构

2.2 MINA

在给出的分层结构基础上,依据传统多通道人机交互体系结构,结合移动设备特点,同时考虑神经功能评价这一应用场景的要求,我们提出移动计算中神经功能评价的多通道人机交互模型——MINA,如图 2 所示.

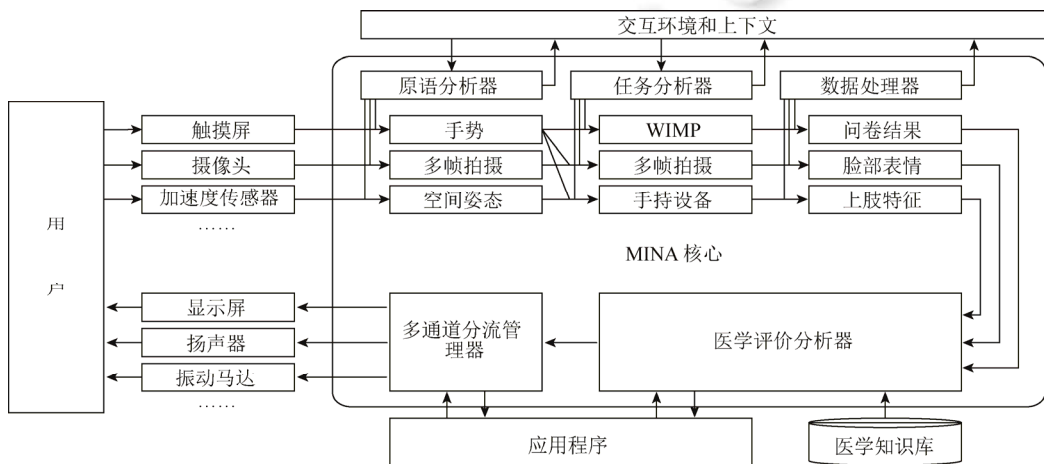


Fig.2 MINA: Multimodal human-computer interaction model for nerve assessment in mobile environment

图 2 移动计算中的神经功能评价多通道人机交互模型——MINA

从 MINA 的总体构成来看,该模型可划分为用户、输入输出设备、MINA 核心、交互上下文、应用程序及

医学知识库 6 个组成部分,它们之间的关系描述如下.

(1) 用户与输入输出设备

直接与用户进行信息交互的是输入/输出设备,由用户发出的交互信息通过不同输入设备或传感器传输到计算设备中进行处理,当处理完成后,又通过多种输出设备将信息呈现和反馈给用户.

(2) MINA 核心与输入输出设备

输入设备获取到的是原始信息,MINA 核心通过调用操作系统应用程序编程接口(application programming interface,简称 API)得到这些信息,处理完成后,MINA 核心再通过 API 控制输入输出设备完成信息的反馈.

(3) MINA 核心与交互上下文

MINA 核心在处理交互信息的过程中,会实时地将交互原语、交互任务、数据处理模块的处理结果输出到交互环境中,这些信息将作为下一时刻的交互原语与任务处理的上下文信息.

(4) MINA 核心与医学知识库

MINA 要对神经功能做出评价,必须引入医学知识库,这个知识库主要包括两部分知识,一是评价参数的规范和标准,另一部分是医学分析模型和方法,此知识库位于 MINA 核心的外部,是一个可拓展的集合.

(5) MINA 核心与应用程序

应用程序的接入点位于 MINA 信息处理过程的末端,主要负责评价结果管理和界面呈现设计两部分工作.

从 MINA 的信息处理和传递流程来看,当交互信息通过不同硬件设备传入系统后,原语分析器便对这些交互信息进行统一化处理,把意义相同而形式不同的输入统一为相同的交互原语.这个过程屏蔽了不同移动设备、不同原始数据形式之间的差异,使得后续的处理过程能够保持一致,使得 MINA 在多样化的移动设备上应用时具有鲁棒性.此外,原语分析器实时地将原语处理的结果输出到交互环境中,使其成为下一时刻词法分析上下文信息.例如,当没有任何上下文信息时触摸屏获得了一个输入点,此时原语分析器会将其暂时认定为“点”原语,并将其输出到交互环境中.下一时刻,触摸屏又获得了一个输入点,此时根据上下文中已经存在的“点”原语信息,原语分析器则会将这两个“点”原语组合成为一个“手势”原语.当原语最终被确定后便被送到任务分析器中,等待下一步的处理.

当原语分析器输出原语后,任务分析器便开始对交互原语进行融合与识别.这里实际上是 MINA 对多通道交互信息进行的第 1 次融合,我们称之为多通道交互语义融合,它是指对输入信息根据相应通道和设备的特点按照某种特定的整合算法进行整合,输出带有明确、完整语义信息的系统命令<sup>[27]</sup>,这个整合过程必须是任务制导的<sup>[28]</sup>,在实际过程中,系统不断扫描由交互原语组成的事件列表,尝试填充不同交互任务所需的任务参数槽,依据不同的任务,这些参数槽可能要求不同类型的交互原语按照特定的规则进行组合,当某一任务填充成功则识别为一个成功执行的任务实例,形成一个具有一定语法结构和任务语义的交互任务,从而为后续的数据处理做好准备.同时,与原语分析器一样,交互任务分析过程也需要实时地将识别结果输出到交互环境中,并利用上下文信息帮助交互任务的识别.

经过任务分析器处理的信息,是具有一定语法结构和任务语义的交互任务,换句话说,此时我们已经知道用户“做了什么”、“具体怎么做的”,下面,数据处理器的作用是计算得出“这样做意味着什么”,即从任务执行结果中提取相应的具有医学评价意义的特征参数.数据处理得出的评价参数主要包括任务执行结果参数,用户生理参数,设备运行参数等,例如,通过 WIMP 任务填写的调查问卷信息在这里被处理,得到用户的年龄、性别、教育程度等个人信息,也可以得到一些医学量表的评分结果;再如,“手持设备”任务信息将在里被处理,得到包括 3 个方向上的振幅、频率在内的一系列姿态特征参数.

经过数据处理器处理的信息已经具被了一定的评价意义,但它们仍是从单一交互任务中提取的,仅描述了用户某一侧面的特征,要对用户的神经功能进行评价,必须依据一定的医学评价模型,综合多个评价参数才能做出正确的评价.在 MINA 中,这个过程实际上是对多通道信息的第 2 次融合,我们称为多通道评价参数融合,它是指通过融合多通道交互任务的处理结果,根据一定的模型和规范,计算得到最终评价结果的过程.以“手持设备”交互任务为例,经过数据处理器可以得到在空间 3 个方向上的振幅与频率等评价参数,则评价参数分析器将通

过一定的模型融合这些参数,得到上肢震颤评分这一评价指标.医学评价分析器的处理过程是需要医学知识库支持的,知识库提供的支持主要包括两个部分,一是评价参数的规范和标准,二是医学分析模型和方法.其中,评价参数的规范和标准规定了进行某个评价时所需要的相关的参数以及这些参数的规范,例如单位、取值范围等.医学分析模型和方法则提供了通过特定的评价参数,计算得出评价结果的算法或规则.

通过医学评价分析器的处理,我们已经得到了最终的评价结果,下一步 MINA 需要通过多通道分流管理器将这些评价结果有效的反馈给用户.多通道分流管理器首先根据评价分析的结果,给应用程序提供可以调用控制接口在最终的应用界面上输出这些信息.这些控制接口是与一个表达意图集对应的,多通道分流管理器会根据这些意图,策划出对用户而言最高效的多通道信息呈现方式,例如在发现某一评价指标超出正常范围时,利用适当的视觉特效和震动反馈来提示用户.本文将仅给出多通道意图表达在 MINA 中的位置和作用,具体实现方法见文献[29].

### 3 MINA 分析

本文给出了 MINA 的层次划分、组织结构,并描述了它的信息处理和传递流程.本节中,我们对 MINA 的移动医学评价及多通道融合的特点展开讨论.

#### 3.1 MINA 的移动医学评价

首先,MINA 是可以适应移动计算环境的.从第 1 节的分析中可以看出,MINA 中使用的交互原语设计是充分考虑了主流移动设备上配备的硬件的,同时,统一原语的提出,能够屏蔽不同移动设备软硬件的差别,保证后续的处理流程和算法的一致.MINA 中使用的交互任务是由交互原语组成的,因此它显然是移动设备能够支持的,同时,这些交互任务又是具有一定医学评价意义,能够得出一定医学评价结果的任务集合,因此,它们能够保障后续的分析处理过程.

其次,利用 MINA 中提供的交互任务进行医学评价是可行的.WIMP 交互任务的主要作用是通过调查问卷或填写表单的形式获取用户信息,多项研究表明,利用这项交互技术能够很好地帮助 NSD 的提早发现和治疗<sup>[1,10,11]</sup>.平面绘图任务的主要作用是利用绘图中轨迹、运动状况等信息,反映用户的认知损伤和运动损伤,相关医学研究指出,书写过程不仅受上肢运动功能影响,也受到知觉、运动觉,如手眼协调、视运动整合、视运动观念等认知因素的影响<sup>[30]</sup>,使用数字笔或触屏设备进行的绘图任务,能够较好地还原书写运动,书写运动特征得到保留,这使得我们能够使用这一任务对神经功能进行评价<sup>[5,7,26]</sup>.绘图任务能够很好地反映用户的接触式运动控制能力,而空中手势则主要是利用非接触式的手部位置运动信息,反映用户的静止性、姿态性或运动性震颤病状,医学研究表明,这项任务对姿态性和静止性震颤的感知尤为敏感<sup>[26,31]</sup>.拍摄任务主要是利用图像处理或计算机视觉方法对用户的一些行为和活动进行识别,从而发现神经功能损伤,文献[19,32]表明,利用彩色图像数据能够准确地识别这些行为,并能找到病人与正常人之间的差异.患有神经功能障碍的用户说话的词频、词汇长度、名词、介词和动词出现的频率等特征表现出与正常人很大的差异<sup>[6,33]</sup>,录音任务正是利用语音处理的方式从用户的语音信息中提取这类特征来评价神经功能的.除了说话方式外,患有神经功能障碍的用户在行走姿态(即步态)上也与正常人有很大差异<sup>[34]</sup>,主流移动设备通常配备加速度传感器和陀螺仪,将其安置在用户身上的特定位置能够很好地感知这些步态上的差异<sup>[9]</sup>,从而从步态的角度对神经功能进行评价.此外,利用加速度传感器和陀螺仪,在用户手持移动设备时,也能感知其手部运动功能障碍,其方法与空中手势类似,只是数据通道有所不同.

综上所述,大量研究和实践表明,MINA 是可以适应移动医学评价需求的.下面,我们将对 MINA 的多通道融合特点展开讨论.

#### 3.2 MINA 的多通道融合

从 MINA 的结构和信息处理流程可以看出,多通道交互信息实际上进行了两次融合.其中,第 1 次融合在任务分析器里进行,是多通道交互语义融合,它将上一步骤得到的交互原语,根据交互环境和上下文的情况,组成



具有一定语法结构和任务语义的交互任务;第 2 次融合在医学评价分析器里进行,是多通道评价参数融合,它通过融合多通道交互任务的处理结果,根据一定的模型和规范,处理得到最终评价结果。

### (1) MINA 的多通道交互语义融合

下面以一个简单的按钮任务(属于 WIMP 任务)和平面绘图任务的识别为例,说明多通道交互语义融合过程。

假设画面中包含一个较大的画布,画布的左上角有一个较小的按钮.此场景包含按钮任务和平面绘图两个任务,按钮任务的参数槽包括 `BUTTON_DOWN` 和 `BUTTON_UP`,平面绘图任务的参数槽包括 `DRAW_DOWN`, `DRAW_MOVE`, `DRAW_UP`,该过程涉及到“点”和“手势”原语的融合,其识别过程可以表示为一个混合自动机<sup>[35]</sup>,如图 3 所示。

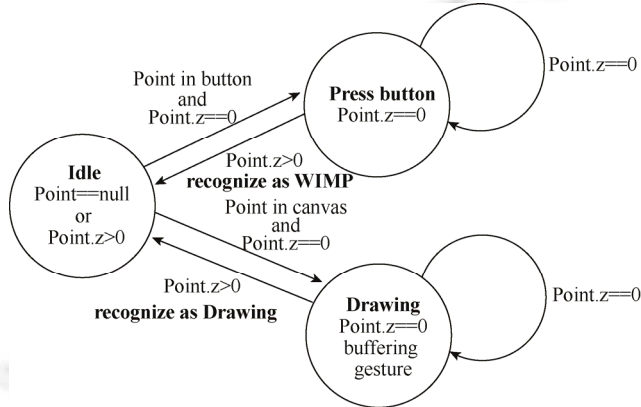


Fig.3 Identifying process of WIMP task and drawing task

图 3 一个 WIMP 和平面绘图任务的识别过程

当任务分析器没获取到“点”原语时,保持空闲状态,一旦获取到一个“点”原语,则根据上下文判断原语数据与按钮的位置关系,如果点位置在按钮内部,则按钮任务的第 1 个参数槽 `BUTTON_DOWN` 得以填充,交互状态转化为按下按钮状态(`press button`),此状态下,即使获取到新的位于画布内的点原语,也不对手势原语进行处理,等待获取到一个  $z$  轴位置大于 0 的点原语后,参数槽 `BUTTON_UP` 得以填充,按钮任务识别成功.相对应地,空闲状态下,当获取到的点原语数据位于画布内部,绘图任务的 `DRAW_DOWN` 参数得以填充,进入绘图状态(`drawing`),只要  $z$  轴位置保持为 0,则不断缓存手势原语,等待获取到一个  $z$  轴位置大于 0 的点原语后,用缓存的手势原语填充参数槽 `DRAW_MOVE`,用最后一个点原语填充参数槽 `DRAW_UP`,平面绘图任务识别成功。

可见,MINA 中的多通道交互语义融合是任务制导的,是根据某一交互任务的需求,依据当前的交互上下文,从得到的交互原语中选取恰当的信息进行融合,得到交互任务实例的,在不同的上下文中,同样的原语可能被识别为不同的交互任务。

### (2) MINA 的多通道评价参数融合

MINA 的多通道评价参数融合是通过融合多通道交互任务的处理结果,根据一定的模型和规范,计算得到最终评价结果的过程.从 MINA 结构中可以看出,这里的模型和规范来自于医学知识库,同时,应用程序可以根据需求对知识的使用进行控制.在实践中,我们通常将评价方法分为 3 类,分别为危险参数评价、基于规则的评价和基于预测方法的评价.其中,第 1 种方法一般仅使用单一交互任务即可完成,仅涉及单一评价参数,其过程主要通过临床经验制定出具有特殊意义的评价指标,并统计患者和正常人在这一指标上的分布范围,得出指标的正常和异常参考值范围,辅助医学诊断.后两种则需要多个评价参数的融合才能完成,下面分别对其进行阐述。

目前临床上应用最广泛的神经功能检查方式是量表评定法,基于规则的评价方法是量表评定法的电子化实现.量表评定法一般要求病人完成一系列规定动作,医生根据观察给出量表中各项参数的半定量或定性评价,然后根据指定规则得出最终评价结果.在电子化实现中,需要根据设备特点,对量表进行简化或者调整,采用适

当的交互通道和交互任务获取相应参数,然后再根据医学知识库提供的评定规则得出评价结果.例如,在蒙特利尔认知检查(the Montreal cognitive assessment,简称 MOCA)中,医生通过问答、绘图、语音、行为等多个方式给病人的记忆力、认知功能、运动功能等多项参数打分,用最终的总分来评价病人的神经功能.在 MINA 中实现这项测试,则可以将 MOCA 中涉及的各项问题通过 WIMP 交互、平面绘图、语音等通道分别实现,得到各项参数打分后计算总分即可.

基于预测的评价方法是利用机器学习方法,将得到的各项参数作为特征,使用事先训练好的评价模型,预测得到评价结果,即监督学习方法.多项研究显示,这类方法使用灵活,准确率高,非常适合 MINA 的交互场景.目前主流的移动设备虽然配备了多种传感器,但普遍精度较低,这使得 MINA 中处理的各项参数指标具有较高的模糊性;此外,受到移动设备的硬件限制,MINA 中得到的各项参数往往无法和传统医学评价量表一一对应,导致即便能够采集到大量数据也找不到相应评价依据的问题.机器学习方法,是处理高纬度、大数据量、不确定性数据的有效手段,因此 MINA 中采用这种方法作为预测评价的手段,处理得到的结果可以是较低级的评价结果,例如“画钟测试的总分”,也可以是较高级的诊断结果,例如“是否患有帕金森病”.

综上所述,MINA 从两个方面体现了多通道的融合,第 1 次融合是多通道交互语义融合,是传统人机交互的多通道融合,它解决了交互任务的有效实施问题.第 2 次融合是多通道评价参数融合,是为了得到更好的评价结果而进行的多通道参数融合,它解决了使用单一交互任务、单一评价参数进行医学评价的不准确和不稳定性问题.实践证明,这两次融合都是非常关键的,少了第 1 次融合则交互任务无法自然有效实施,少了第 2 次融合则无法得到准确的评价结果.下面,我们将通过一个 NSD 自动评估的实例,来探讨 MINA 在实际软件设计中的应用效果.

## 4 实例研究:笔式神经功能评价系统

### 4.1 连线测试

医学上经常使用的认知功能检测方法包括问卷量表和小测试.绘图小测试,例如连线测试(trail making test,简称 TMT)、螺旋线测试、画钟测试等,是一种有效且便于实施的方法.在绘图测试方法中,判定受试者绘制的一条线段的变形程度是大多数临床研究关注的重要特征之一.神经功能损伤包括认知功能损伤和运动功能损伤,如果受试者绘制的线段或曲线与标准曲线有严重偏差,则很可能因为认知功能损伤导致,而如果线段或曲线表现出严重的抖动,则很可能是运动功能损伤造成的.

连线测试是一种应用广泛的神经心理学测验,用于评价执行功能,由 A、B 两部分组成,其中 TMT-A 要求受试者将无序分散的数字 1→25 按从小到大的顺序依次连接,如图 4(a)所示,在传统的 TMT-A 中,受试者使用笔在印有 25 个数字的纸面上完整连线.连线测试的评分取决于完成连线的总时间,评分主要与认知处理速度相关<sup>[36]</sup>.TMT-B 与 A 类似,不同的是在每对阿拉伯数字中间,按照字母表的顺序插入英文字母.由于 TMT-A 包含的阿拉伯数字识别、排序并连线的任务相对简单,易于被绝大多数受试者接受,本文选中 TMT-A 来进行认知功能评价.

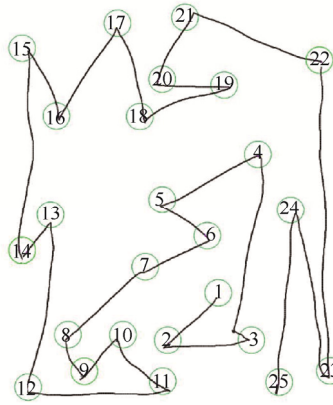
据此,本文实现了一个基连线测试的、在移动设备上利用笔手势进行神经功能评价的系统,它能自动地将不同 NSD 病人(本例中包括 PD 和 SVD)从其他样本中区分出来,辅助医学诊断.

### 4.2 交互原语和任务分析

本应用需要获取用户三类信息用于神经功能评价,它们包括绘图信息、绘图过程中设备的运动状态、用户的个人信息.首先,对于绘图信息而言,根据传统连线测试的描述,使用“笔-纸”隐喻的人机接口最为合适,其交互原语是包括“点”和“手势”(注意手势包括笔手势),其交互任务属于“平面绘图”类型.其次,为了获取更多评价特征,本应用要求用户手持设备进行测试,并在测试过程中采集设备的加速度传感器和陀螺仪数据,即移动设备的运动状态,其交互原语是“空间姿态”,其交互任务是“携带或手持设备”.最后,用户的个人信息,例如年龄、腰围、血压等信息需要用户输入,这里用到的交互原语是“点”“手势”“文本输入”和“数量值输入”,交互任务属于

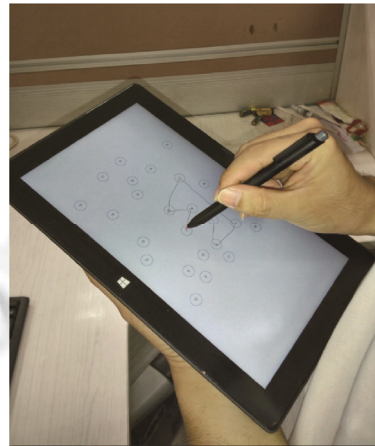
“WIMP 交互”。

基于以上考虑,本应用使用 Microsoft Surface 平板电脑作为目标移动设备进行开发,系统实际运行如图 4(b) 所示.尽管如此,通过适当修改任务和布局方案,本方法仍可容易地移植到其他移动设备上,例如智能手机。



(a) A example of trail making test completed by user

(a) 一幅由用户完成的 TMT-A 连线测试图



(b) A user is conducting trail making test

(b) 一个用户正在进行连线测试

Fig.4

图 4

### 4.3 数据处理流程

本应用的数据处理流程如图 5 所示。

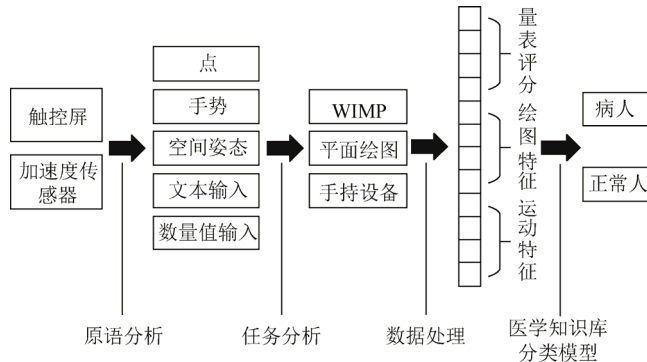


Fig.5 The data processing flow of pen-based nerve function assessment system

图 5 笔式神经功能评价系统数据处理流程

第 1 步,原语分析器从移动设备的触屏、运动传感器上获取原始交互信息,对它们进行统一化处理,生成我们需要的“点”“手势”“空间姿态”“文本输入”和“数量值输入”原语信息.这一步利用操作系统提供的 API 和基本信号处理算法即可完成。

第 2 步,任务分析器将交互原语进行组合,生成我们需要的“平面绘图”“空间姿态”和“WIMP 交互”任务.本应用首先通过简单的混合状态机区分“WIMP 交互”与“平面绘图”任务,在平面绘图任务开始时,缓存“空间姿态”原语,直到绘图结束,并用缓存的空间姿态原语填充手持设备任务参数,使其该任务同时完成。

第 3 步,数据处理器对上一步的交互任务信息进行处理,得到具有医学评价意义的参数.这些参数包括绘图特征参数、运动特征参数和量表评分参数.其中,绘图特征参数包括总时间、错误次数、连线偏离度,运笔速

度、加速度和压力的变化程度,绘图轨迹的弯曲程度,组成一个 59 维的特征向量,用于描述绘图任务体现的运动功能损伤.运动特征参数主要包括设备的空间位置特征和姿态特征,我们首先对三轴空间位置的功率谱密度和频率进行分析,形成一个 20 维的向量,再对于设备的三轴空间姿态做同样的处理,又得到一个 20 维的向量,两者一起形成一个 40 维的能够反映用户手部震颤的特征向量.量表评分参数包括年龄、受教育年限、收缩压、舒张压、腰围、臀围、是否有卒中病史、MMSE 认知评分共 8 项信息.

第 4 步,将得到的 107 项评价参数送到医学评价分析器里面,进行下一步的分析.下一节将对这一步进行详细讨论.

#### 4.4 多通道参数融合评价

##### (1) 基于医学规则的样本筛查

实践中我们发现,使用预测的方法直接对全体人群进行分类准确率较低,目前最高仅能达到 70%左右,其原因是,导致用户表现出运动或认知功能的原因可能来自其他与疾病不相关的方面,例如较低的文化程度、智力低下、神经系统曾遭受严重损伤等.因此,我们必须首先利用医学知识库里提供的相关规则,对样本进行基于规则的初筛.其具体规则为,有以下条件者<sup>[37]</sup>排除在预测模型自动诊断范围外:① 文盲(受教育年限<1 年).② MMSE 评分 $\leq 19$ 分;小学文化程度(受教育年限为 1~6 年)且 MMSE 评分 $\leq 22$ 分;中学及以上文化程度者(受教育年限 $\geq 7$ 年)且 MMSE $\leq 26$ 分.③ 有卒中病史.

##### (2) 基于预测模型的自动诊断

经过上一步骤的样本筛查,我们将部分可能表现出严重运动或认知功能障碍,但又不属于神经功能障碍的用户排除在自动诊断的范围外.之后,再利我们训练好的预测模型,对测试样本进行分类,其模型训练过程为:首先,挑选典型病例和正常对照样本,本例中使用 44 例病例,44 例年龄、性别、文化程度相匹配的正常人,共计 88 例样本.第二,将上一节处理得到的参数,除去本步骤中不使用的 MMSE、受教育年限、是否有卒中病史 3 项共计 104 项进行异常值剔除和归一化处理.第三,实施基于相关性的特征选择,得到最终特征集.最后,使用选择出来的特征训练分类器,得到预测模型.

#### 4.5 结 果

本文将采集到的 810 例样本(766 例非 NSD,44 例 NSD)按照前面的规则分为 3 个样本集,分别是训练样本集 TS (training set)、全部测试样本集 TTS(total test set)、筛查测试样本集 STS(screening test set),见表 3.选用决策树(C4.5)、贝叶斯网络 BN(Bayesian network)和线性 SVM 作为测试分类器,表 4~表 6 给出了 3 个样本集中不同分类器的平均准确率、NSD 组的查准率、查全率和 ROC 曲线面积 AUC(area under ROC curve).

**Table 3** Training set and test set used in the study

**表 3** 训练和测试样本集

样本集	样本说明
训练样本集	包括 44 个 NSD 病例样本和 44 个挑选出来的正常人对照样本
全部测试样本集	包括 44 个 NSD 病例样本和 766 个非 NSD 样本
筛查测试样本集	包括 44 个 NSD 病例样本和筛查出来的 442 个正常人样本

在 TS 中,本文通过十则交叉验证测试分类器性能.其中,BN 分类器获得了最高的平均准确率和最大 AUC 面积,SVM 分类器表现出最高的查准率,C4.5 则获得最高的查全率.

**Table 4** Classification results of TS

**表 4** 训练样本集分类结果

分类器	准确率(%)	查准率	查全率	AUC
C4.5	89.77	0.872	0.932	0.909
BN	90.91	0.929	0.886	0.969
SVM	88.64	0.972	0.795	0.886

使用 TS 训练得到的模型,对 TTS 进行分类.3 个分类器性能均大幅下降,最高的平均准确率由 C4.5 获得,仅

为 65.07%,C4.5 同时也获得了最高查准率,仅为 0.135.可见,导致分类性能下降的主要原因来自于对非 NSD 样本的错误分类.但是,3 个分类器对 NSD 样本的查全率均保持了较高水平,C4.5 达到了 1.000 的查全率,说明该分类器没有错过任何一个 NSD 样本.最大 AUC 面积由 BN 获得.

**Table 5** Classification results of TTS

**表 5** 全部测试样本集分类结果

分类器	准确率(%)	查准率	查全率	AUC
C4.5	65.07	0.135	1.000	0.815
BN	64.32	0.130	0.977	0.882
SVM	64.81	0.115	0.818	0.728

由上一节的分析可知,部分非 NSD 样本也会表现出反映迟钝、认知能力低下、手部震颤等特点,提取的特征和模型无法很好的对 TTS 中的非 NSD 样本进行分类,因此,必须利用相关医疗规则首先对全体样本进行初筛得到 STS,表 6 给出了经过筛查后的分类结果.

使用 TS 训练得到的模型,对 STS 进行分类,分类性能得到了明显提高,3 个分类器均达到了 72%以上,注意到 STS 是一个极度不均衡的样本集(正负样本比例约 1:10),在较高查全率(1.000,C4.5)和 AUC 面积(BN,0.919)的条件下,这一结果是可以接受的.换句话说,虽然有较多的正常人被误诊,但极少数 NSD 病人被漏诊,这是符合辅助临床医疗排查需求的.

**Table 6** Classification results of STS

**表 6** 筛查测试样本集分类结果

分类器	准确率(%)	查准率	查全率	AUC
C4.5	72.84	0.250	1.000	0.851
BN	72.02	0.242	0.977	0.919
SVM	72.63	0.224	0.818	0.768

尽管如此,本应用中的预测模型至少在以下两个方面还有很大提升空间:首先,NSD 病例样本不足,导致无法得到更具鲁棒性的预测模型,从结果中可以看到,虽然 3 个分类器在 TS 的十则交叉验证中表现良好,但泛化能力欠佳,更多具有代表性的病例样本是改进预测性能的关键因素.第二,本应用中没有利用空中手势、录音、摄像这几个交互任务,一些 NSD 症状没能很好地体现,相信如果能够采集这些任务相关交互数据,为分类器提供更多有价值的特征,能够进一步提升模型的预测能力.

## 5 结 论

人机交互特有的多通道、交互式、人机协同的计算能够有效地提高移动环境下神经功能评价的准确度.然而,目前人机交互在移动环境下神经功能评价的作用没有得到充分讨论,没有统一的多通道交互模型,缺少相关的人机交互原语和任务设计,多通道交互的优势没有得到体现.

为此,本文首先分析了目前移动设备上主流的神经功能评价方法,归纳总结出了一套适用该应用场景的交互原语和交互任务,然后,在此基础上提出了一种移动环境下神经功能评价的多通道人机交互模型——MINA,并对该模型的移动医学评价和多通道融合特点进行了分析,最后,依据此模型给出一个 NSD 检测的应用实例.

在实例研究中我们看到,MINA 能够较好地利用移动设备上的不同交互通道,利用医学规则与预测模型组合的方法对 NSD 病人进行有效检测和排查,基于 MINA 开发的笔式神经功能评价系统,提供了一个移动环境下对 NSD 早期预警与辅助诊疗的有效手段.

本文没有讨论以移动设备为处理终端,并借助外部设备进行数据采集的这一类方法.利用外部传感器,如肌肉电传感器、脑电传感器等,能够扩充交互通道,增强交互系统的理解感知能力.然而,外部设备加入,将给移动健康评价应用的推广和普及带来了限制,使得这类方法大多只能在实验室条件下进行.如何将更多的传感器嵌入移动设备,为移动健康评价应用提供更多的感知通道,还需要进一步研究.

**References:**

- [1] Keränen T, Liikkanen S. Medication reminder service for mobile phones: An open feasibility study in patients with Parkinson's disease. *Telemedicine and e-Health*, 2013,19(11):888–890. [doi: 10.1089/tmj.2013.0014]
- [2] Hii P, Chung W. A comprehensive ubiquitous healthcare solution on an Android™ mobile device. *Sensors*, 2011,11(7):6799–6815. [doi: 10.3390/s110706799]
- [3] Hong H, Gilbert E, Abowd GD, Arriaga RI. In-Group questions and out-group answers: Crowdsourcing daily living advice for individuals with autism. In: *Proc. of the Human Factors in Computing Systems*. Seoul: ACM, 2015. 777–786. [doi: 10.1145/2702123.2702402]
- [4] Visvanathan A, Banerjee R, Choudhury AD, Sinha A, Kundu S. Smart phone based blood pressure indicator. In: *Proc. of the Mobile Ad Hoc Networking and Computing*. Philadelphia: ACM, 2014. 19–24. [doi: 10.1145/2633651.2633657]
- [5] Graça R, Castro RSE, Cevada J. Parkdetect: Early diagnosing parkinson's disease. In: *Proc. of the Medical Measurements and Applications*. IEEE, 2014. 1–6. [doi: 10.1109/MeMeA.2014.6860027]
- [6] Baldas V, Lampiris C, Capsalis C, Koutsouris D. Early diagnosis of Alzheimer's type dementia using continuous speech recognition. In: *Proc. of the Wireless Mobile Communication and Healthcare*. 2010. 105–110. [doi: 10.1007/978-3-642-20865-2\_14]
- [7] Surangsrirat D, Intarapanich A, Thanawattano C, Bhidayasiri R, Petchrutchatachart S, Anan C. Tremor assessment using spiral analysis in time-frequency domain. In: *Proc. of the IEEE Southeastcon*. IEEE, 2013. 1–6. [doi: 10.1109/SECON.2013.6567469]
- [8] Min JK, Doryab A, Wiese J, Amini S, Zimmerman J, Hong J. Toss 'N' turn: Smartphone as sleep and sleep quality detector. In: *Proc. of the Human Factors in Computing Systems*. Toronto: ACM, 2014. 477–486. [doi: 10.1145/2556288.2557220]
- [9] Lan KC, Shih WY. Early diagnosis of Parkinson's disease using a smartphone. *Procedia Computer Science*, 2014,34(C):305–312. [doi: 10.1016/j.procs.2014.07.028]
- [10] Perrier T, Dell N, DeRenzi B, Anderson R, Kinuthia J, Unger J, John-Stewart G. Engaging pregnant women in kenya with a hybrid computer-human SMS communication system. In: *Proc. of the Human Factors in Computing Systems*. Seoul: ACM, 2015. 1429–1438. [doi: 10.1145/2702123.2702124]
- [11] Regenbrecht H, Langlotz T, Ho C, George M, Gray A, Walmsley R, Schultz M. Field test of a questionnaire-based mobile health reporting system. In: *Proc. of the Australasian Computer-Human Interaction Conference*. Sydney: ACM, 2014. 384–387. [doi: 10.1145/2686612.2686670]
- [12] Hong H, Kim JG, Abowd GD, Arriaga RI. Designing a social network to support the independence of young adults with autism. In: *Proc. of the Computer Supported Cooperative Work*. Seattle: ACM, 2012. 627–636. [doi: 10.1145/2145204.2145300]
- [13] Donnelly M, Nugent C, McClean S, Scotney B, Mason S, Passmore P, Craig D. A mobile multimedia technology to aid those with Alzheimer's disease. *IEEE Multimedia*, 2010,17(2):42–51. [doi: 10.1109/MMUL.2010.25]
- [14] Pal A, Sinha A, Choudhury AD, Chattopadhyay T, Visvanathan A. A robust heart rate detection using smart-phone video. In: *Proc. of the Mobile Ad Hoc Networking and Computing*. Paris: ACM, 2013. 44–48. [doi: 10.1145/2491148.2491156]
- [15] Larson EC, Goel M, Boriello G, Heltshe S, Rosenfeld M, Patel SN. SpiroSmart: Using a microphone to measure lung function on a mobile phone. In: *Proc. of the Ubiquitous Computing*. Pittsburgh: ACM, 2012. 280–289. [doi: 10.1145/2370216.2370261]
- [16] Fontecha J, Navarro FJ, Hervas R, Bravo J. Elderly frailty detection by using accelerometer-enabled smartphones and clinical information records. *Personal and Ubiquitous Computing*, 2013,17(6):1073–1083. [doi: 10.1007/s00779-012-0559-5]
- [17] Dell N, Francis I, Sheppard H, Simbi R, Borriello G. Field evaluation of a camera-based mobile health system in low-resource settings. In: *Proc. of the Int'l Conf. on Human-Computer Interaction*. Toronto: ACM, 2014. 33–42. [doi: 10.1145/2628363.2628366]
- [18] Woods AM, Nowostawski M, Franz EA, Purvis M. Parkinson's disease and essential tremor classification on mobile device. *Pervasive and Mobile Computing*, 2014,13:1–12. [doi: 10.1016/j.pmcj.2013.10.002]
- [19] Albert MV, Toledo SD, Shapiro MB, Kording K. Using mobile phones for activity recognition in Parkinson's patients. *Frontiers in Neurology*, 2012,158(3):1–7. [doi: 10.3389/fneur.2012.00158]
- [20] Meng J, Zhu TY. New mobile medical system mode based on smart phone. *Application Research of Computers*, 2013,30(7):2055–2060. (in Chinese with English Abstract)

- [21] Tian F, Cheng C, Chen YD, Dai GZ. Virtual assembly Toolkit: A 3D interaction platform for virtual assembly. *Journal of Computer-Aided Design & Computer Graphics*, 2002,14(3):193–198 (in Chinese with English abstract).
- [22] Tian F, Qin YY, Wang XC, Ao X, Wang HA, Dai GZ. Analysis and design on PIBG Toolkit: A pen-based user interface Toolkit. *Chinese Journal of Computers*, 2005,28(6):1036–1042 (in Chinese with English abstract).
- [23] Stewart C, Rohs M, Kratz S, Essl G. Characteristics of pressure-based input for mobile devices. In: *Proc. of the Human Factors in Computing Systems*. Atlanta: ACM, 2010. 801–810. [doi: 10.1145/1753326.1753444]
- [24] Goc ML, Taylor S, Izadi S, Keskin C. A low-cost transparent electric field sensor for 3D interaction on mobile devices. In: *Proc. of the Human Factors in Computing Systems*. Toronto: ACM, 2014. 3167–3170. [doi: 10.1145/2556288.2557331]
- [25] Mackinlay, JD, Card SK, Robertson GG. Rapid controlled movement through a virtual 3D workspace. In: *Proc. of the Computer Graphics*. ACM, 1990. 171–176. [doi: 10.1145/97880.97898]
- [26] Smits EJ, Tolonen AJ, Cluitmans L, Gils MV, Conway BA, Zietsma R, Leenders KL, Maurits NM. Standardized handwriting to assess Bradykinesia, micrographia and tremor in Parkinson's disease. *PLOS ONE*, 2014,9(5):1–8. [doi: 10.1371/journal.pone.0097614]
- [27] Wang Y, Yue WN, Wang H, Dong SH. Multi-Modal interaction in handheld mobile computing. *Ruan Jian Xue Bao/Journal of Software*, 2005,16(1):29–36. (in Chinese with English abstract) <http://www.jos.org.cn/1000-9825/16/29.htm>
- [28] 董士海,王坚,戴国忠.人机交互和多通道用户界面.北京:科学出版社,1999.
- [29] Lu L, Tian F, Dai GZ, Wang HA. A study of the multimodal cognition and interaction based on touch, audition and vision. *Journal of Computer-Aided Design & Computer Graphics*, 2014,26(4):654–661 (in Chinese with English abstract).
- [30] Werner P, Rosenblum S, Bar-On G, Heinik J, Korczyn A. Handwriting process variables discriminating mild alzheimer's disease and mild cognitive impairment. *Journal of Gerontology: Psychological Sciences*, 2006,61(4):228–236. [doi: 10.1093/geronb/61.4.p228]
- [31] Scanlon BK, Levin BE, Nation DA, Katzen HL, Guevara-Salcedo A, Singer C, Papapetropoulos S. An accelerometry-based study of lower and upper Limb tremor in Parkinson's disease. *Journal of Clinical Neuroscience*, 2013,20(6):827–830. [doi: 10.1016/j.jocn.2012.06.015]
- [32] Joumier V, Romdhane R, Bremond F, Thonnat M, Mulin E, Robert P, Derreumaux A, Piano J, Lee J. Video activity recognition framework for assessing motor behavioural disorders in Alzheimer disease patients. In: *Proc. of the Int'l Workshop on Behaviour Analysis and Video Understanding*. 2011. 1–9.
- [33] Bucks R, Singh S, Cuerden JM, Wilcock GK. Analysis of spontaneous, conversational speech in dementia of alzheimer type: Evaluation of an objective technique for analysing lexical performance. *Aphasiology*, 2000,14(1):71–91. [doi: 10.1080/02687030401603]
- [34] Bridenbaugh SA, Kressig RW. Quantitative gait disturbances in older adults with cognitive impairments. *Current Pharmaceutical Design*, 2014,20(19):3165–3172. [doi: 10.2174/13816128113196660688]
- [35] Yang L, Guan ZW, Dai GZ. Modeling post-WIMP user interfaces based on hybrid automata. *Ruan Jian Xue Bao/Journal of Software*, 2001,12(5):633–644. [http://www.jos.org.cn/ch/reader/key\\_query.aspx](http://www.jos.org.cn/ch/reader/key_query.aspx) [doi: 10.13328/j.cnki.jos.2001.05.001]
- [36] Bazner H, Oster M, Daffertshofer M, Hennerici MG. Assessment of gait in subcortical vascular encephalopathy by computerized analysis: A cross-sectional and longitudinal study. *Journal of Neurology*, 2000,247(11):841–849. [doi: 10.1007/s004150070070]
- [37] Zhang ZQ, Zahner GEP, Roman GC, Liu X, Wu C, Hong ZL, Hong X, Tang MX, Zhou B, Qu QM, Zhang X, Li H. Socio-Demographic variation of dementia subtypes in China: Methodology and results of a prevalence study in Beijing, Chengdu, Shanghai, and Xi'an. *Neuroepidemiology*, 2006,27:177–187. [doi: 10.1159/000096131]

#### 附中文参考文献:

- [20] 孟溶,朱天宇.一种基于智能手机的新移动医疗系统模式. *计算机应用研究*,2013,30(7):2055–2060.
- [21] 田丰,程成,陈由迪,戴国忠.面向虚拟装配的三维交互平台. *计算机辅助设计与图形学学报*,2002,14(3):193–198.
- [22] 田丰,秦严严,王晓春,敖翔,王宏安,戴国忠.PIBG Toolkit:一个笔式界面工具箱的分析与设计. *计算机学报*.2005,28(6):1036–1042.
- [27] 王悦,岳玮宁,王衡,董士海.手持移动计算中的多通道交互. *软件学报*,2005,16(1):29–36. <http://www.jos.org.cn/1000-9825/16/29.htm>



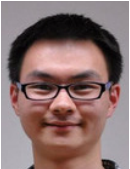
[29] 路璐,田丰,戴国忠,王宏安.融合触、听、视觉的多通道认知和交互模型.计算机辅助设计与图形学学报,2014,26(4):654-661.



黄进(1985-),男,广西玉林人,硕士,CCF 学生会员,主要研究领域为人机交互技术,图形图像处理.



田丰(1976-),男,博士,研究员,博士生导师,CCF 会员,主要研究领域为人机交互技术,虚拟现实.



陈毅能(1988-),男,博士,CCF 会员,主要研究领域为人机交互技术,生理计算.



戴国忠(1944-),男,研究员,博士生导师,CCF 会员,主要研究领域为人机交互,计算机图形学.



刘杰(1981-),男,博士,讲师,CCF 会员,主要研究领域为人机交互技术,图形图像处理.



王宏安(1963-),男,博士,研究员,博士生导师,CCF 会员,主要研究领域为实时智能,用户界面.

www.jos.org.cn