



肌萎缩侧索硬化症的女士能够通过大脑缓慢打字。近期举行的F8大会上播放了一段视频,一位患

## 新款人工智能语音系统 高效率复制人声

加拿大新创公司琴鸟(Lyrebird)发布了新款AI(人工智能)语音系统,其能通过分析讲话录音和对应文本以及两者之间的关联,在1分钟内模仿人类“讲话”。

让声音听起来更自然,是计算机程序在将文字转换成语音过程中面临的核心挑战,即使目前最好的语音助手,如苹果公司的Siri和亚马逊公司的Alexa,一发声给人的第一感觉仍然是“哇!这是计算机。”背后原因在于这些语音助手系统的工作原理:根据预录的声音文档整理出词汇,再通过另一个新的音频文档将这些词汇拼凑在一起发声。

而琴鸟公司的AI使用了一种全新的语音合成系统,能在“倾听”过程中“掌握”每个人说话时字母、音位和单词的发音特点,通过推理并模仿这个人声音中的情感和语调,“说”出全新的语句。

新系统使用模仿人脑思维的算法创建出一种人工神经网络,能利用深度学习技术将所听到的转换成语音,并仅靠任何人1分钟讲话内容,就能完全模仿这个人说话。其市场前景非常广阔,可用来改进个人人工智能助手、音频书籍以及残疾人语音系统等。

开发该系统的蒙特利尔大学学习算法实验室博士后亚历山大·布瑞比森表示,在学会并模仿了几个人的声音后,再模仿任何一个新对象的语音就会变得更快,因此新语音系统不需太多信息,1分钟足以捕获某个人声音的核心特点。

但美国卡内基梅隆大学语言技术研究所教授迪莫·鲍曼表示,“琴鸟的语音系统和真正的人声之间还有差距,仍然能听出其计算机语音特征。”他认为,语音系统真正令人信服地复制人声,还需再等几年。

(据《科技日报》)

## 最新研究显示 听音乐可改善自律神经功能

日本理化研究所和大阪市立大学的研究小组发现,人在聆听环境音乐时可以减轻主观疲劳感,变得舒心放松,从而改善与主观情绪变化有关的自律神经功能。从生理学观点推测,这些情绪变化可能与自律神经功能有关,但尚没有科学的共识加以解释。

此次,研究小组从钢琴、小提琴和自然音源等构成的音乐中,选择适于预防疲劳和放松恢复的乐曲制成环境音乐专辑,并以20名健康人为对象,利用数字化装置测定他们在下班后听30分钟专辑音乐与无音环境下的疲劳、不安、睡意、忧郁等主观情绪。结果发现,被测试者在听音乐时主观情绪发生了较大变化,主观疲劳感减轻,慰藉、放松情绪变动较大。环境音乐对循环系统自律神经功能的影响为心跳显著减慢。

这项研究表明,科学评价主观心情转换的效果是可能的,今后有望将音乐欣赏确立为一门与健康有关的学科。

(据《北京日报》)

## 英国科学家 揭开诚实的秘密

不知你有没有想过,为什么大多数人是诚实的,骗子只占小部分?英国科学家借助磁共振成像设备揭开了这一奥秘。

伦敦大学学院研究人员在《自然·神经学》杂志上发表文章说,他们找来56名志愿者,把他们两两分组,让这些人选择电击自己或偷偷电击对方来换取金钱报酬。研究人员发现,志愿者在做决定的时候,大脑中负责原则评估的纹状体被激活。磁共振成像扫描显示,志愿者通过电击自己换取报酬时,纹状体活动要比电击同伴换取报酬时更为剧烈。这说明,他们本能地认为电击自己的行为更可取。

研究结果显示,大脑会将他人对自己的道德评价内在化。即便知道自己的不良行为不会被发现,也会想象出侵害他人时,他人对自己的责备。

研究结果还显示,当志愿者拒绝通过侵害他人获利时,大脑中负责道德判断的侧前额叶会与纹状体产生交流。研究人员说,这意味着社会道德标准可以以神经信号的方式在人体内体现出来,这样以来,人在非法获利时就无法感受到愉悦。

(据《深圳商报》)

# 语音、触觉还是脑机接口 未来我们如何与机器交流

本报记者 任志方

## 脑电波直接输出语言 可达到每秒100个字

在F8大会上,扎克伯格表示,Facebook将在未来十年致力于虚拟现实(VR)和增强现实(AR)技术的发展,打造智能眼镜和VR社交。

Facebook旗下Oculus Research公司首席科学家迈克尔·亚伯拉什说,距离AR眼镜好到足以“飞入寻常百姓家”这一关键点,可能仅有5年时间。他说:“一款增强视觉和听觉的眼镜,能让你更加智能,通过它你可以持续不断地感知环境,与世界保持联系,而且它轻便、舒适、高效。”

1946年,电脑开始商业化,但是直到图形用户界面的出现,计算才真正变得个人化、实用化、直观。目前大家习惯的图形用户界面,自从1984年苹果公司推出麦金托什电脑之后就没怎么变过。目前的AR和VR混合现实设备就像上世纪70年代IBM的电脑,也需要一个人机交互方式的变革让其真正变成每个人都能使用的产品。

在F8大会上,Facebook公司一年前秘密组建的硬件团队Building 8首次亮相,其负责人雷吉纳·杜坎(Regina Dugan)说,Building 8的宗旨就是要创造出全新的、以社交为先的消费电子产品。她宣布了两个全新的人机交互项目,一个是通过脑电波传达信息,另一个是通过触觉来传达信息,并称之为无声的语音UI(操作界面),兼具语音UI如亚马逊智能音箱Echo的便利性和文字短信的隐私性。

第一个项目是要通过对脑电信号的解码,从脑电波直接输出语言,并达到每秒100个字,5倍于在手机上打字的速度。一旦成功,既能帮助残疾人,又能作为AR设备的新输入方式。

与其他科研机构进行的项目不同,Building 8选择通过光学造影技术,以无创的方式解析人脑中布诺卡氏区(处理语言的区域)的活动。目前在临床试验中,已经能够为肌萎缩性侧索硬化症的病人实现每分钟8个字的速度。要提高输出速度,需要更高级的传感器,以光速来探测大脑活动。当然,为了打消对侵犯隐私的顾虑,杜坎特意表示,该技术只会解析大脑语言区域的活动,而非窃听所有想法。

第二个项目是效仿盲文,通过皮肤上的神经触觉来感知和传达信息。Building 8的工程师在实验中,用绑在手腕上的发射器,感受手臂发出的不同频率,来识别同伴所表达的意思,如蓝色球体、白色方块等。如果该项目成功,不仅会造福残疾人,更能使不同母语的人之间无障碍地交流。

在过去75年里,每10年就有一次人机交互的重大革新,人类对机器的操作,从物理手柄按键,到物理键盘,再到触摸屏。10年前,从苹果推出第一代智能手机以后,键盘被取消改为触摸屏,随后改变了整个移动设备的人机交互方式。

在Facebook公司近期于美国加州圣何塞举行的F8(8小时外的创新)大会上,“掌门人”马克·扎克伯格描绘了一幅新的人机交互图景:在头脑中打字,无需通过智能手机、电脑或其他屏幕,所有计算信息都将直接投射到人眼内。这或许意味着,我们将迈入一个无屏幕时代,现在作为我们亲密助手的智能手机,最终也将被颠覆。

## 语音正在成为 新的人机交互载体

Facebook的无声语音UI仍在规划图和实验室中,但语音交互已大行其道,似乎正在成为新的交互方式。一个极具代表性的消费终端,就是亚马逊2014年推出的基于语音识别的智能音箱Echo。

Echo有着圆柱形的外观,没有屏幕。通过Wi-Fi联网,家庭用户可以对着它说话,获取各种网络资讯、点播歌曲、视频,另外也可以连接各种智能家电进行控制。相比手机而言,这种设备更适合在家庭使用,用语音的方式实现快速的操控和咨询。

目前,语音识别准确率已从2010年的70%上升到现在的90%,百度、谷歌、Hound等平台的语音识别准确率已超过90%。美国一家公司宣称专业速记员在记录对话时,转录词错率为5.1%,而该公司最新研发的语音识别系统词错率已达到5.5%,超越之前的历史最佳水平,树立了新的里程碑。

作为人工智能领域的核心问题之一,如果语音识别准确率能提高到99%,而且能够针对户外嘈杂环境进行语音识别,一定会引发交互方式的质变,0.4%的差距似乎意味着机器即将比肩人类。得益于互联网、社交媒体、移动设备和廉价的传感器,这个世界产生的数据量急剧增加。通过数据的海洋中丰富的语料以及进化中的深度学习算法,语音识别技术正不断优化。

对于消费者而言,类似Echo的

智能音箱会逐渐成为新的人机交互载体。毕竟,说话的速度相比打字输入要快得多,而且这一交互方式也不必打断人们手头正在做的事。从语音识别来看,下一个前沿是从识别走向理解,而这正是Facebook要做的事。不过,在机器能理解其所听到或看到的事物的真正含义之前,还需要很长时间的工作,有很长的路要走。

## 植入脑机接口的猴子 打出莎士比亚经典独白

人机交互方式的革新方向并不止Facebook提出的无声语音UI。人们还渴望通过脑机接口(BCI)技术将意识和机器融合到一起,从而打破人类自身的局限性。

2016年,几只猴子成为脑机接口领域的标杆性事件。2016年9月,斯坦福大学神经修复植入体实验室的研究者们往两只猴子大脑内植入了脑机接口,通过训练,其中一只猴子创造了新的大脑控制打字的纪录——1分钟内打出了10个单词和2个标点,即莎士比亚的经典独白“To be or not to be. That is the question.”(生存还是毁灭,这是个问题)。这项技术希望用于帮助包括处于肌萎缩性侧索硬化症晚期的病人的日常交流。

2016年11月,《自然》杂志上发表了一项来自于瑞士、德国、意大利、法国、中国、英国和美国的联合研究。该研究小组开发的一种脑机接口,它可以通过再现来自大脑的信号记录刺激腿部的电极,使脊髓损伤的猕猴能在无需借助任何外骨骼的帮助下重新行走。

研究小组首先绘制跑步机上行走的健康猕猴的电信号是如何从大脑发送到腿部肌肉的。然后在脊髓切断的猕猴身上再现这些信号。他们将微电极阵列植入于瘫痪的猕猴的大脑中,获取并解码与腿部运动相关的信号。这些信号被发送到位于低位脊柱的电脉冲发生装置,从而触发猕猴腿部肌肉运动,使瘫痪的猕猴再次行走。

目前,AR最常用的使用领域是建筑、工业组装、交通、体育、军事和执法,这些都是非常动态的工作场合,目前复杂繁冗的头显设备显然不够灵活机动,这种公开场合也不是特别适用语音输入,这时候脑机接口就可以大显身手了。

除了输入简单之外,脑机接口还能解决目前混合现实的另一大局限:认知超载,这时候你自己的大脑就像一个自动广告拦截器一样,通过脑机交互的自适应性界面或响应用户生物指标的软件来过滤掉不相关的信息,减少认知过载。

现在我们的手机、电脑、平板、手表都成为了我们大脑的延伸,未来它们将会变得更小、更无形,沉重的AR或VR头显会变成隐形眼镜,那时候脑机交互是此复杂交互过程唯一的解决办法。