

# 计算科学评论

WWW.COMPUTER.ORG

2017年第1期

技术  
增强  
人类

人工智能先驱 P12

自动驾驶的致命硬伤 P74

合作机构



IEEE  
computer  
society



ISSN 1673-5153



# 云计算

人工智能

工控机

制造业

电子

嵌入计算

传感器

互联网

# 3D 打印

绿色计算

图形图像

虚拟现实

职场  
市场  
行业分析  
**创业**  
学习 科技新闻  
竞赛

找工作  
教育

互联网金融

芯片

大数据

软件架构

纳米架构

人机交互

多媒体

普适计算 MEMS



微信名：计算机人 微信号：jisuanren



# Copyright

# 版权

## 主管单位 Authorities in Charge

中华人民共和国教育部 Ministry of Education of the People's Republic of China

## 主办单位 Sponsor

中国大学出版社协会 China University Presses Association

## 出版单位 Publisher

《环球科学》杂志社有限公司 GLOBAL SCIENCE MAGAZINES Co., Ltd

社址 Address: 北京市朝阳区秀水街1号建外交公寓4-1-21 Office 4-1-21, Jianguomen Diplomatic Residence Compound, No. 1, Xiu Shui Street, Chaoyang District, Beijing, China. 邮编 100600

联系电话: 010-85325810 / 85325871

## 社长 / 总编辑 Editor-in-chief

陈宗周 ChenZongzhou

副校长 / 副总编辑 Deputy Editor-in-chief

刘芳 LiuFang

## 执行出版人 Publisher

管心宇 Xinyu Guan

## 资深编辑 Senior Editor

马法达 Falda Ma

刘妍 Yan Liu

## 特约编辑 Contributing Editor

史彦诚 Yancheng Shi

刘大明 Daming Liu

高天羽 Tianyu Gao

费楠 Yong Fei

王璇 Xuan Wang

## 运营中心 OPERATING DEPARTMENT

运营机构 Publisher

上海灵宸文化传媒有限公司

## 发行部 Circulation Department

### 发行总监 Circulation Director

谢磊 XieLei 010 - 57439192

## 市场部 Marketing Department

### 市场总监 Marketing Director

孔祥彬 KongXiangbin 010 - 85325810 - 807

## 广告部 Advertising Department

### 销售总监 sales Director

范欢 FanHuan 010-85325871-802 010-85325981

## 读者服务部 Reader Service

杜珺 Du Jun 010 - 57458982

## 印刷: 北京博海升彩色印刷有限公司

如发现本刊缺页、装订错误和损坏等质量问题, 请在当月与本刊读者服务部联系调换 (请将坏书寄回)。

国际标准刊号: ISSN 1673-5163

国内统一刊号: CN11-5480/N

广告经营许可证号: 京朝工商广字第8144号

## 知识产权声明:

IEEE, IEEE Computer, IEEE中文网站的名称和标识, 属于位于美国纽约的电气电子工程师学会有限责任公司所有的商标, 仅通过授权使用。这些材料的一部分由IEEE Computer英文版翻译而来, 版权归IEEE所有, 并经IEEE授权翻译复制。

IEEE Computer杂志的中文版权, 由美国电气电子工程师学会有限责任公司授予上海灵宸文化传媒有限公司, 并由本刊独家使用。

本刊发表的所有文章内容由作者负责, 并不代表上海灵宸文化传媒有限公司、美国电气电子工程师学会有限责任公司的立场。

本刊内容未经书面许可, 不得以任何形式转载或使用。

## 编辑团队

### 执行编辑 Carrie Clark

cclark@computer.org

### 高级编辑 Chris Nelson

### 编辑 Lee Garber, Meghan O'Dell

### 特约编辑 Christine Anthony, Rebecca Torres

### 多媒体编辑 Brian Brannon

### 设计和产品

Monette Velasco, Lead

Jennie Zhu-Mai, Lead

Mark Bartosik

Larry Bauer

Erica Hardison

### 封面设计 Andrew Baker

### 平面设计 Hector Torres

### 高级商务拓展经理 Sandy Brown

### 高级广告经理

Marian Anderson Debbie Sir

### 产品和服务总监 Evan Buttereld

### 会员总监 Eric Berkowitz

### 编辑服务高级经理 Robin Baldwin Manager

### 编辑服务内容开发 Richard Park

## 主编

Sumi Helal

University of Florida

helal@cise.u.edu

## 副主编

Ying Dar Lin

National Chiao Tung University

ydlin@cs.nctu.edu.tw

## 副主编, COMPUTING PRACTICES

Rohit Kapur Synopsis

rohit.kapur@synopsys.com

## 副主编, PERSPECTIVES

Bob Colwell

bob.colwell@comcast.net

## 副主编, SPECIAL ISSUES

George K. Thiruvathukal

gkt@cs.luc.edu

## 副主编, MULTIMEDIA EDITOR

Charles R. Severance

University of Michigan

csev@umich.edu

## 2016年IEEE计算机协会主席

Roger U. Fujii

Fujii Systems

r.fujii@computer.org

## 行业编辑

### 大数据和数据分析

Naren Ramakrishnan

Virginia Tech Ravi Kumar Google

### 云计算

Schahram Dustdar

TU Wien

### 计算机架构

David H. Albonesi

Cornell University

### Greg Byrd North

Carolina State University

### Erik DeBenedictis

Sandia National Laboratories

### 绿色和可持续计算

Kirk Cameron

Virginia Tech

### 健康信息学

Upkar Varshney

Georgia State University, Atlanta

### 高性能计算

Vladimir Getov

University of Westminster

### 识别科学和生物识别技术

Karl Ricanek

University of North Carolina

Wilmington

### 物联网

Roy Want

Google

### 安全和隐私

Rolf Oppliger

eSECURITY Technologies

### 软件

Renée Bryce

University of North Texas

Jean-Marc Jéz é quel University of Rennes

### 视觉、可可视化和增强技术

Mike J. Daily

HRL Laboratories

## 顾问委员会

Doris L. Carver

Louisiana State University (EIC Emeritus)

Carl K. Chang

Iowa State University (EIC Emeritus)

Theresa-Marie Ryne

Consultant

Bill Schilit

Google

Savitha Srinivasan

IBM Almaden Research Center

Ron Vetter

University of North Carolina

Wilmington (EIC Emeritus)

Alf Weaver

University of Virginia

**知识产权声明:**  
IEEE, IEEE Computer, IEEE中文网站的名称和标识, 属于位于美国纽约的电气电子工程师学会有限责任公司所有的商标, 仅通过授权使用。这些材料的一部分由IEEE Computer英文版翻译而来, 版权归IEEE所有, 并经IEEE授权翻译复制。  
IEEE Computer杂志的中文版权, 由美国电气电子工程师学会有限责任公司授予上海灵宸文化传媒有限公司, 并由本刊独家使用。  
本刊发表的所有文章内容由作者负责, 并不代表上海灵宸文化传媒有限公司、美国电气电子工程师学会有限责任公司的立场。  
本刊内容未经书面许可, 不得以任何形式转载或使用。

# 专帮科研人找工作



科研求职，就到科研圈招聘频道

[www.keyanquan.net/recruit](http://www.keyanquan.net/recruit) 

携手数百家科研院校与机构, 第一时间发布职位需求;  
高层次人才、博士后、教职、技术人员、企业岗位……多职位全覆盖;  
地域、领域、行业等多级筛选, 高效搜索。

# 计算科学评论



16

## 导读 人类增强技术

计算技术正给我们的世界带来翻天覆地的变化，也很有可能改变我们自己。

2017年4月刊

## 封面报道

20

### “第三眼”：视觉 障碍者的购物助手

彼得·A·津塔拉

李素妍

格斯·H·史密斯

洛瑞·布伦纳

劳伦特·依提

玛丽·B·罗森

约翰·M·卡罗尔

凯文·M·伊瑞克

维加克里施南·纳

拉亚南

30

### 定制3D打印智能 眼镜，增强日常生活现实

弗洛里安·瓦尔

张瑞

马丁·弗洛伊德

奥利弗·阿姆夫特

2017年4月刊

# 目录



40

52

## 基于人体传导声音信号的新型通信方式

张铖  
希南·海尔  
赛克  
蒲奕名  
孙丹睿  
薛秋月

## 神经技术能给我们更好的大脑吗?

萨德·E·斯  
塔那格雷戈  
里·D·艾博德  
奥默·T·伊南  
拉赞·巴特查里亚  
布莱恩·A·考夫曼  
周在勳  
马修·E·菲利普斯

## 趋势

06 超级智能：是谜团，还是紧迫的现实？

## 大师小传

12 人工智能的两位开山鼻祖  
Marvin Minsky 和 Seymour Papert

## 专家观点

67 数据分析能给科研方向带来启示吗？

中国中文信息学会社会媒体处理专业委员会

## 计算技术

74 自动驾驶的社交硬伤

巴里·布朗

所有汽车实现完全自动驾驶的那天，人们的要求就不仅仅是安全和高效了。自动驾驶汽车还必须能够理解人类驾驶员，并与之以一种自然的方式进行交流。

## 计算未来

74 构想个人未来

艾丽达·德罗特，茱莉亚·罗丝·韦斯特

两位未来学家通过教授战略前瞻方法，帮助人们更具批判性地思考未来。



# 超级智能：是谜团，还是紧迫的现实？

文 | 乔治·赫尔伯特 (George Hurlburt), STEMCorp  
译 | 刘大明

**奇**点指的是在未来人与技术合二为一的时间点。奇点意味着我们所知的人类文化的终结。随着快速发展的人工智能 (AI) 令技术具备了“超级智能”，这一天有朝一日总要来临。一些人警醒说，超级智能将在未来25年内超越人类智能。这使人们担心自己可能会在不自觉间成为其机器人的宠物。还有一些人，包括工业家伊隆·穆斯克 (Elon Musk) 和天体物理学家史蒂芬·霍金 (Stephen Hawking) 则警告说，AI是“威胁人类存在的最大威胁”，他们表达了对AI无序发展的担心，担心这最终将使人类走向毁灭。<sup>1</sup>无论如何，随着近年来AI的发展在全国范围内越发引人注目，由牛津大学研究人员尼克·博斯特罗姆 (Nick Bostrom)<sup>2</sup>所著的关于超级智能的书中提出的超级智能的末日迷因 (meme)，似乎已经获得了很大市场。

事实上，先进的智能自动化的潜力将减少对人类劳动力的需求。机器人

已经改变了制造业，取代了生产线上的许多工人。这一在制造车间中提升重复劳动和限时任务效率的大趋势，未来有可能继续有增无减。然而，这并不意味着机器人比那些精心设计的功能人更加智能。相反，机器人的精度更高。工业机器人在执行某些重复性的装配任务时，可以比人类劳动力更加可靠、速度更快。然而，这不是超级智能的强大之处，后者指的是机器在智能水平上超越人类，甚至可能显示出有意识的行为。工厂中的机器人最多只是一种简单的AI形式，执行一些反应式任务，不能回想起最近的行动，为未来的行动提供信息。

## 人工智能并非新概念

“人工智能”一词是在1955年夏天的一项研究建议中提出的，该建议提出了1956年夏天在达特茅斯大学进行的为期两个月、10人参与的最终研究。这次“达特茅斯人工智能会议”奠定了人工智能的概念：

...提出一种猜想，即关于各个方面学习或任何其他智能化特征，能在原则上进行精确描述，可以对机器进行模拟。<sup>3</sup>

会议建议中提出了实现人工智能的几种方法，包括“自动化计算机、神经网络、尺寸计算理论、自我完善、抽象性、随机性和创造性”。<sup>3</sup>

大约61年后，经历大量修正和推动，AI的定义只发生了轻微的变化。教科书中对AI的定义是：

...一个与智能行为自动化相关的计算机科学分支。

这个定义包括“基于生物、社会和代理的模型，对自动化理论和定理的证明”。其中后者包括“专家系统、自然语言处理 (NLP) 和语义、人类表现、规划和机器人技术、神经网络和遗传算法。”<sup>4</sup>因此，尽管AI的重点领域可能已经成熟，但AI的基本原则仍与1955年最

初提出时保持着惊人的相似。

然而, AI的大多数定义似乎将“智能”做字面上的理解,在关于人类智能构成上预设一定程度上的共识,就像下文中讨论第四类AI时预设的关于人类意识构成的共识一样。因此,除了列出现代AI追求的目标之外,对AI的现代定义中经常假定一般的人类智能已被充分理解。这个观点为一些基本的探索留下了空间,对于现代流行的超级智能观点持一定的批判态度。

## AI的类型

图1所示为四种类型的AI,从相当直接的I型AI到更复杂的AI变种,这一分类基于对自我意识的感知能力,并最终根据自我意识的行动能力。这个层次划分是密歇根州的AI研究人员阿伦德·辛采(Arend Hintze)在2016年提出的。<sup>5</sup>辛采的文章已经在几个流行的媒体上发表。I型AI是反应式的,不能回顾过去的行动,并基于全部固定算法来进行特定的活动。II类人工智能具备有限的记忆,可以根据发生在过去时空中的观察来采取未来行动。III型AI具备“心智理论”,系统获得对其周围环境的社会理解能力,并相应地作出反应。IV型AI表示获得了自我意识能力,可以表明自主的自由意志。本文研究了上述每种AI可能达到超级智能水平的潜力。

## 计算机只会字面理解 人类有文化底蕴

所有的二进制计算机从核心上说都是机械理解。它们说到底只能基于二进制数1或0进行操作。而人类拥有内部连接极其复杂的真正的嵌入式神经网



图1. 四种形式的人工智能(AI)<sup>5</sup>

络,往往能够不断学习文化行为。这包括在早期获得的计算机文化。对特定的背景进行归纳是人类独有的能力,这使我们能够概括出本地化的事件和话语中的意义,否则对这些事件和话语就可能有无数种合理解释。

这些心理属性使得人类得以形成广泛的社会网络,从而产生了生动的社会和持久的文化。这种历史悠久的行为倾向于强化、增长并复制人类的智力。它正强调了III型AI的本质。例如,II型AI自驾车遵守道路导航和现实规则的能力,在III型AI中成为次要因素,III型AI会理解出行的原因。在这里,心智理论会对AI进行调整,以适应更广泛的世界观。

尽管自然语言处理取得了长足进步,通用二进制机器复制人类语义行为的能力仍然是有限的。总的来说,IBM的Watson尽管占据了数字计算机的优

势,但仍然不能成功识别出语义的细微差别,而人类则对此经验丰富。作为语境知识管理手段的链接数据的质量仍然不高,这表明目前的通用计算仍处于语义缺失的状态。<sup>6</sup>

有研究进一步断言,计算机在解决人类问题上存在无法超越的极限,从而使得超级智能是否实现成为疑问,即便是在解决易处理的问题上,机器的速度要快过人类。<sup>7</sup>

## 生理学超越二元计算

审视一下人类智能的生理现实是有用的。成年人脑的平均体积约73立方英寸。它包含大约1000亿神经元突触,能够容纳100万亿左右的紧密连接的通路。<sup>8</sup>今天世界上有74亿个活的人类大脑在运转。在包括IBM“沃森”在内的已知的人造计算机中,没有任何一台能够在数量、运算速度和连通性上达到这

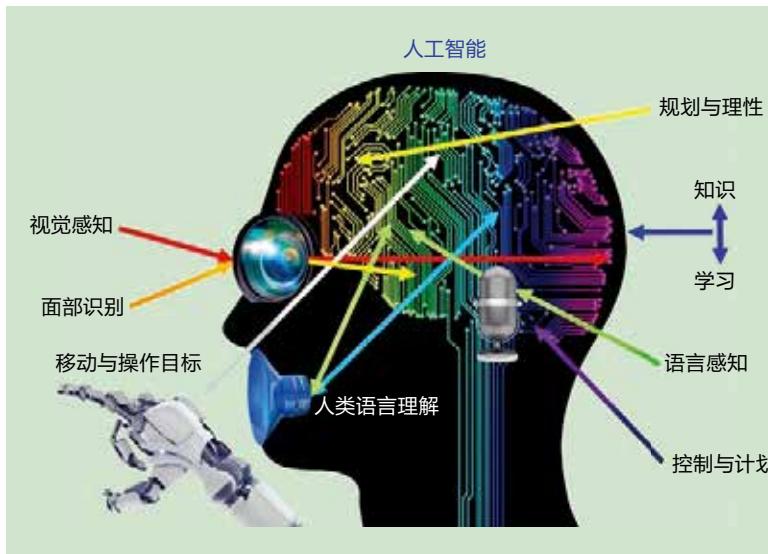


图2. 拟人化的AI构成概念

一水平。此外，正常的成年人大脑具备多模态感知能力，其范围远远超出经典意义上的“六感”，人的大脑具备学习、同步运动、语言、读写、沟通、理解、推理、概括、决策、情感、以及短期和长期记忆能力，同时具备多种自主管理功能。目前还没有人造计算机系统具有这样的回忆、感知、情境感知、自我控制，自主性和多任务处理等全能型功能。图2所示为用人类功能比喻生物智能网络各部分概览。

看起来合理的一点是，为了实现超级智能状态，所有这些认知属性将必须集成在能够胜过名义人类大脑的机器中，这里的机器在物理集成度甚至更高。这也要求机器如果要胜过人脑其运行基础就是更加生物化的，而不是二进制的。这意味着可以同时利用集成嵌入非确定性网络生成一些可预测的结果，所有结果都是基于经适当过滤后的大量的同时输入得到的。这仍然是新兴的“网络的网络”的研究内容。

一些奇点的支持者不同意上述的整体观点。他们指出，只有少数智能属性（例如IBM“沃森”的自然语言处理能力，以及它对大型知识库的访问能力）就足以描述出一台“超级智能”机器的特点。*Kurzweil*的“加速回报定律”借鉴了摩尔定律的内容，表明即使是有限的人工智能，甚至是分布式纳米级的人工智能，也可能迅速使人类智能相形见绌。<sup>9</sup>然而，这种观点一般没有考虑关于可再生能源和自我意识等更重要的一些问题。

所有生物都要消耗某种形式的能量。所有数字设备（目前被认为是无生命的非生命实体）则要消耗电力。因此，一个简单到嘴边的回答可能就是在一些超智能数字化力量成功抢夺足以武装自己的基础设施之前，迅速地拔掉电源插头。虽然这可能是科幻惊悚片里的内容，但当IV型AI的自我意识的概念发挥作用时，这个场景可能并不那么遥远。

为了理解超级智能系统对电力的渴求，该系统需要建立一个关于能源的概念，能源来自哪里，如何获得，以及如何消费。如果达尔文的观点在逻辑上同样适用于超级智能的进化，那么该系统还将设计一种保护其能量来源，并在系统的自然生命周期中保证不间断获取能量的方式。这与社会对粮食安全的需求不同，不仅提出了一些关于空间和时间上的回忆的概念，而且还强调了对自身的持续存在造成明显、即刻威胁的自卫式感知。此外，该系统还建议建立一种能够有效地与其他设备进行协调的手段，以满足其对能量的需求。然而，要想扩展出良好的外部感知能力和社会交际技能，如果没有强烈的自我保护意识，这些行为是不可能存在的。如果没有一些关于自我的概念，在自由意志下支持一些故意的、有意识的相对自主的行为，这种感知能力就说不通。这似乎就是对IV型AI特征的描述。

## 有意识的机器

IV型AI（图1）理论上可以表现出自我意识。到目前为止，没有任何机器能真正拥有类似于人类意识的东西。事实上，尽管神经科学已经取得了长足进步，关于生物体处于有意识状态的现实条件究竟是什么，这个问题的研究仍停留在理论层面。意识的产生确实存在着一个先决条件，即存在生物学上足够复杂的自适应网络。<sup>10</sup>目前的一个悬而未决的问题是，这样一个由生物领域之外运行的机器组成的网络，能否产生意识？一些人认为是可以的。<sup>11</sup>

目前可以和个人认知水平相提并论的唯一实体可能就是互联网了。这里所

说的互联网与万维网不同，后者正在迅速沿着互联网的许多分布式主干积累着世界上的知识。实际上，互联网的数据中心和节点可能可以容纳数百万亿条网络连接。然而，这些连接的相对传输速度通常受到距离的限制。事实上，互联网正受到多种形式的攻击，经常被恶意利用，并且容易分发混杂在真实信息之中的错误信息。然而，鉴于互联网的性质，它能够成为超级智能的计算实体，超越人类智能吗？

## 物联网会变得智能吗？

很快，物联网（IoT）传感器就能胜过人类感知的所有形式和模式，能够感知到的活动远远超出人类的先天感觉。假设带有传感器的智能代理可能共谋夺取基础设施资产的控制权，人类很可能确实被严重边缘化。

这种活动目前已经存在。基于代理的DDOS攻击具备相当频繁的破坏力。然而，我们必须假设新兴的先进分析技术将让有效的人为干预成为可能。在任何情况下，最初的攻击可能不是由一个故意的超级智能行为所指导的，而是在设计上存在漏洞，被其他人的有意利用，或互联网本身的非线性行为导致的意想不到的后果。互联网的现状反映出了其消费对象的社会结构，但是从核心上看仍然不具备III类或IV类AI那样足以统治这个社会的网络属性。要想实现这一目的，仍然需要人类的失误。

这可能体现了AI带来的更可怕的真正威胁。如果没有适当的保护，所有类型的AI都可以被黑客入侵。想象一下恶意入侵会对自动驾驶汽车造成的影响吧，由于许多现代汽车产品中存在的

安全漏洞，这种说法绝不是危言耸听。更糟的是，对III型和IV型AI的恶意攻击，最终可能在其毁灭性的影响范围内给社会造成重大威胁。因此，真正的风险可能不是AI本身或由AI产生的超级只能，而在于出于无益的目的对AI的过度利用。

## 2016年的形势

2016年，美国国家科学技术委员会关于AI技术趋势的报告对AI的应用前景持乐观态度，对即将到来的超级智能则轻描淡写，不过由于AI在政府部门

经典学科有关的概念性的AI研究领域（也许并不完整）。尽管应用性AI主要仍属于计算机科学领域，但仍将需要来自许多相关功能性领域从业者的额外的专业知识。

第二，随着AI在越来越多从前的人类任务中超越人类的能力越来越强，我们需要强调终身学习。这一点很有必要，以使人类不断地建立技能基础，能够在市场上竞争。要终身学习，不能有任何懈怠，这必须成为一种规范，因为AI推动了自动化的进步。例如，未来对机器人专家的需求可能会成为抑制美国

**如果没有适当的保护措施，所有类型的AI都可能被黑客入侵。想象一下恶意攻击会对自动驾驶汽车造成哪些影响。**

内部的应用越来越广泛，报告呼吁加强道德上的安全措施。<sup>12</sup>事实上，穆斯克、霍金、比尔·盖茨，史蒂夫·沃兹尼亚克以及1000多位著名的AI研究人员共同签署了一份公开信，承认AI可能实现的好处，但同时信中也呼吁采取道德保障措施，防止威胁人类的AI利用方式（公开信见<http://futureoflife.org/ai-open-letter>）。信后随附的文件阐明了与所有类型的AI相关的、涉及道德问题的具体领域。<sup>13</sup>

## 未来的方向

从目前所知的情况看，我们可以对AI的未来做出一些关键的假设。

首先，人工智能的研究需要一系列综合的多学科研究方法。图3所示为与

就业市场的因素，因为今天的大数据分析专家已经出现了短缺。

第三，如果没有坚实的道德管理框架，AI就面临着可能被人利用，对人类产生危害的风险。从前文提到的2016年总统委员会的建议和知名AI研究人员签署的公开信中可以清楚无疑地看出这一点。随着自主性武器系统的跨领域整合程度获得前所未有的加深，甚至已经涉及II类AI，对妥善管理的需求变得至关重要。

第四，为了达成IV类AI，网络将需要从当前的数字化网络物理模型演变到耦合更紧密、自适应性更高的网络生物模型。在人体中循环的基于代理的脑植入物和智能纳米机器人的概念，也许会成为这种基本进化过程中可能出现的

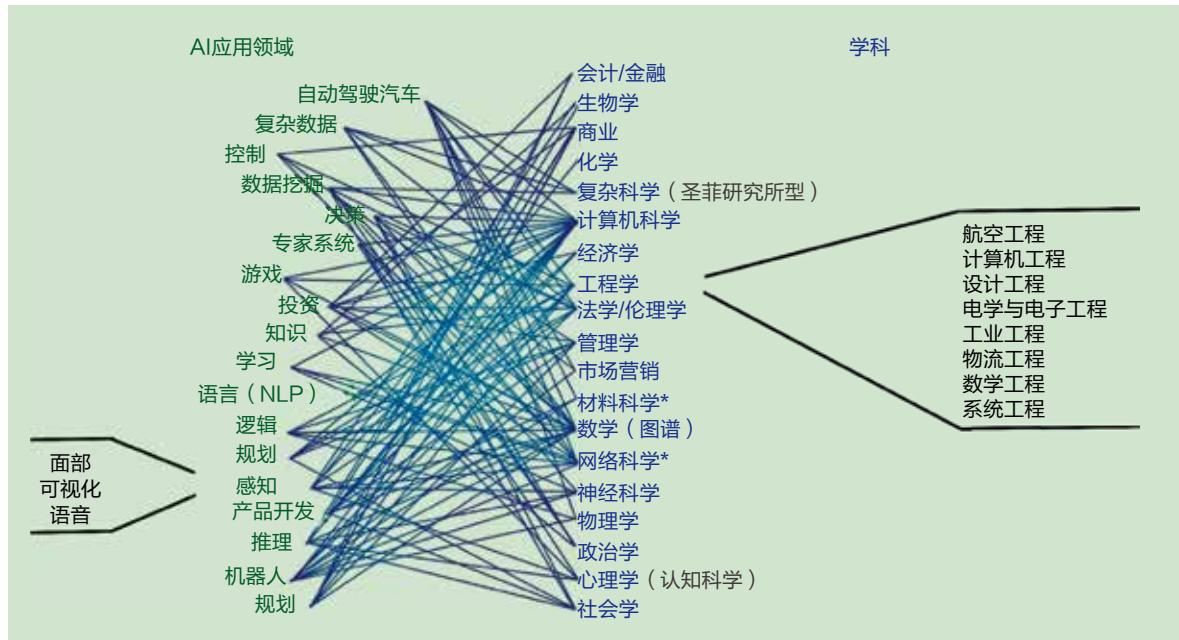


图3. AI研究与经典学术领域之间的概念性关系

示例。如果确实如此，人类将在创造IV型AI的过程中成为AI知情的合作伙伴，而不会沦为不幸的受害者。

最后，AI在互联网和基础设施管理、物联网和其他现代技术方面有很多优势。考虑到气候变化对人类的威胁比超级智能传播的奇点更加迫切，AI研究必须致力于克服这种更直接的威胁。这需要我们有足够的智慧和意愿，以针对AI制定和执行健全的道德管理方式。

# 简

而言之，在没有人类的规模帮助的情况下，AI不可能消灭人类。因此，呼吁在AI的设计和实施中实施强有力管理是至关重要的。否则，AI一旦被用于邪恶的目的，就可能对未来产生难以估量的深远影响。■

## 参考文献

1. S. Gibbs, “Mus�, Wozniak, and Hawking Urge Ban on Warfare AI and Autonomous Weapons,” *The Guardian*, 27 July 2015; [www.theguardian.com/technology/2015/jul/27/musk-wozniak-hawking-ban-ai-autonomous-weapons](http://www.theguardian.com/technology/2015/jul/27/musk-wozniak-hawking-ban-ai-autonomous-weapons).
2. N. Bostrom, *Superintelligence: Paths, Dangers, Strategies*, Oxford Univ. Press, 2014.
3. J. McCarthy et al., A Proposal for The Dartmouth Summer Research Project on Artificial Intelligence, Dartmouth Univ., 31 Aug. 1955.
4. G. Lugar, *Artificial Intelligence: Structures and Strategies for Complex Problem Solving*, 5th ed., Pearson Education, 2005.
5. A. Hintze, “Understanding the Four Types of AI, from Reactive Robots to Self-aware Beings,” *The Observer*, 18 Nov. 2016; <http://observer.com/2016/11/understanding-the-four-types-of-ai-from-reactive-robots-to-self-aware-beings/>.
6. S. Bechhofer et al., “Why Linked Data is Not Enough for Scientists,” *Future Generation Computer Systems*, vol. 29, no. 2, 2013, pp. 599–611.
7. J. Wiedermann, “A Computability Argument against Superintelligence,” *Cognitive Computation*, vol. 1, no. 1, 2009, pp. 236–245.

8. C. Zimmer, "100 Trillion Connections: New Efforts Probe and Map the Brain's Detailed Architecture," *Scientific Am.*, Jan. 2011; <http://bit.ly/2gNKg09>.
9. R. Kurzweil, *The Singularity Is Near: When Humans Transcend Biology*, Penguin Group, 2005.
10. O. Sporns, *Networks of the Brain*, MIT Press, 2011.
11. D.K. Prasad and J.A. Starzyk, "A Perspective on Machine Consciousness," *Proc. 2nd Int'l Conf. Advanced Cognitive Technologies and Applications*, 2010, pp. 109–114.
12. "Preparing for the Future of Artificial Intelligence," Executive Office of the President, US Nat'l Science and Technology Council Committee on Technology, Oct. 2016; <http://bit.ly/2hhWeTl>.
13. S. Russell, D. Dewey, and M. Tegmark, "Research Priorities for Robust and Beneficial Artificial Intelligence," *AI Magazine*, vol. 37, no. 4, 2015; [http://futureoflife.org/data/documents/research\\_priorities.pdf](http://futureoflife.org/data/documents/research_priorities.pdf).

**乔治·赫尔伯特**是非营利组织STEMCorp的首席科学家,该组织旨在利用网络科学推进经济的进一步发展,同时推动自主技术的进步,以期作为人类的有用工具。他致力于基于动态图的物联网架构研究。赫伯特是IT Professional编委会成员,南马里兰高等教育中心管理委员会委员。联系方式:[ghurlburt@change-index.com](mailto:ghurlburt@change-index.com)。



《IEEE计算机图形及应用》(IEEE Computer Graphics and Applications, 简称CG&A)把计算机图形学领域的理论和实践联系在一起。《IEEE计算机图形及应用》提供了包括从某个特定算法到全系统实现在内的同行评议的深度报道。它为那些处于计算机图形技术前沿的人们提供了必不可少的资料。无论他们处于商界还是艺术界,这本杂志都能让他们受益。

请点击: [www.computer.org/cga](http://www.computer.org/cga)



## 人工智能的 双子先驱

文 | 乔治·斯特劳 ( George Strawn )，美国国家科学、工程及医学科学院 ( US National Academies of Sciences, Engineering, and Medicine )  
坎迪斯·斯特劳 ( Candace Strawn )

今年我们失去了两位伟大的科学家：马文·明斯基 ( Marvin Minsky ) 于 2016 年 1 月 24 日去世，西摩尔·派珀特 ( Seymour Papert ) 于 7 月 31 日去世。两人都是享年 88 岁，而且两人生前在麻省理工学院的学术工作彼此密切相连，无论是在人工智能 ( AI ) 领域还是其他领域都是如此。本文将着重介绍他们的一些贡献和共同引发的一个“争议”。YouTube 上有这两位开山鼻祖的一些访谈和讲座视频，值得一看（比如 <https://youtu.be/-pb3z2w9gdg> 等）。此外，明斯基的朋友 史蒂芬·沃尔夫拉姆 ( Stephen Wolfram ) 在最近出版的《Idea Makers》一书中用一章的篇幅讲述了明斯基的生平。<sup>1</sup>

### 早年生活与教育背景

Minsky 于 1927 年出生在美国纽约的一个富裕家庭，因此能够从一开始就受到良好教育。他在布朗克斯科学高中和菲利普斯学院度过了青少年时期。

到第二次世界大战后期，他到了可以入伍的年龄，于是他 1944 年应征进入美国海军服役。直到 1945 年战争结束，他重回学校继续学业，1950 年获得哈佛大学的学士学位，1954 年获得普林斯顿大学的数学博士学位。他的学位论文题目是《神经模拟强化系统的理论及其在大脑模型问题中的应用》。这是对人工神经网络 ( ANN ) 领域所做的早期贡献，也正是在这个领域中，他和 派珀特 共同引发了前面提到的争议（关于这个争议，我们将稍后讨论）。1958 年他进入麻省理工学院任教（详见 <http://web.media.mit.edu/~minsky/minskybiog.html>）。

Papert 于 1927 年出生在南非，并在那里的威特沃特斯兰德大学接受教育，1949 年获得哲学学士学位，1952 年获得数学博士学位（详见 [https://en.m.wikipedia.org/wiki/Seymour\\_Papert](https://en.m.wikipedia.org/wiki/Seymour_Papert)）。之后他去了英国，在剑桥大学圣约翰学院工作，并于 1959 年获得第二个数学博士学位。在大学学习

期间，他是一名反对南非种族隔离政策的杰出活动家。他还曾在巴黎大学的庞加莱学院、日内瓦大学（在那里他曾与 Jean Piaget 共事，详见 [https://en.m.wikipedia.org/wiki/Jean\\_Piaget](https://en.m.wikipedia.org/wiki/Jean_Piaget)）和伦敦的国家物理实验室工作过，在此之后，他于 1963 年开始在麻省理工学院任助理研究员。

### 在麻省理工学院的学术工作

明斯基被人们尊称为人工智能之父，他于 1959 年与约翰·麦卡锡 ( John McCarthy )（此人后来去了斯坦福）共同创立了麻省理工学院人工智能实验室。不过，他其实在很多领域都做出过贡献，包括图形、符号数学计算、知识表达、常识性语义、机器感知以及符号学习和联结主义学习。他也是机器人和远程呈现领域的先驱人物。他设计并开发了一些早期的视觉扫描仪，即带有触觉传感器的机械手。他还为派珀特的 LOGO 编程语言开发了一个早期“海龟”，包括软件和硬件接口。他在 1985

年发表了一部专著《意识社会》(详见 [https://en.m.wikipedia.org/wiki/Society\\_of\\_Mind](https://en.m.wikipedia.org/wiki/Society_of_Mind))<sup>2</sup>, 对自己多年的思想进行了总结。

Papert 最为人知的可能是他对教育工作的贡献, 当然他在其他领域的贡献也很突出。明斯基在 1967 年任命他为麻省理工学院人工智能实验室主任。派珀特是 Piaget 的建构主义学习理论的主要倡导者, 并根据该理论与他人共同设计了LOGO语言, 而LOGO对后来的Smalltalk和Scratch等编程语言都产生了影响。1980 年, 他出版了《头脑风暴: 儿童、计算机和强大理念》一书(详见 [https://en.m.wikipedia.org/wiki/Mindstorms\\_\(book\)](https://en.m.wikipedia.org/wiki/Mindstorms_(book)))<sup>3</sup>。

## 明斯基的“意识社会”

“意识社会”是比人工智能更具思辨性的认知科学与哲学。也就是说, 它认为人脑是由一些独立智能体组成的“社会”, 人类在进化过程中, 随着时间的推移生成了这些智能体来处理各种需要解决的问题。这本书包含 270 个章节(每页即一章), 每一章解决一个大脑智能体、过程或互连问题。

2006 年, 明斯基又出版了该书的续集, 即《情感机器》<sup>4</sup>, 书中提出的理论可以按照多个层次的过程来解释人类的高级感觉、目标、情感和意识思维, 而且这些过程能够彼此体现。有关意识的这一观点是面向过程的, 它认为意识是 20 多个过程在人脑中作用的结果。

如果《意识社会》和《情感机器》中的想法可以在计算机上实现, 那么明斯基就确实为人工智能研究人员们做了

开山辟路的工作。如果认知神经科学能够证实他所确定的智能体和过程确实存在于大脑的结构当中, 那么他便确定了“自然智能”的主要组成部分。这些猜测是否正确? 也许在不久的将来便会一见分晓。

## 派珀特的 LOGO 语言

前面已经提到, 派珀特曾在日内瓦与 Piaget 共事过, 而且拥护 Piaget 的建构主义学习理论(详见 [https://en.m.wikipedia.org/wiki/Constructivism\\_\(philosophy\\_of\\_education\)](https://en.m.wikipedia.org/wiki/Constructivism_(philosophy_of_education)))。人们给派珀特的思想取了个相似的、易混淆的建构主义学习理论名称。在美国国家科学基金会的提案中, 他将这一理论定义如下(详见 [https://en.m.wikipedia.org/wiki/Constructionism\\_\(learning\\_theory\)](https://en.m.wikipedia.org/wiki/Constructionism_(learning_theory))):

建构主义是一种记忆法, 面向的是此项目的科学教育理论基础的两个方面。从心理学的建构主义理论出发, 我们把学习视为一种重建, 而不是知识的传播。其次, 我们把有关操纵性材料的想法扩展到认为在学习者的一部分体验活动是构建一个有意义产品时学习最有效这种想法。

(孩童)学习者构建一个有意义产品的想法是通过构建 LOGO 程序和可由 LOGO 程序控制的机器人设备实现的。与 Papert 联合发明 LOGO 编程语言的瓦里·福尔齐格(Wally Feurzeig)基于 Fortran 语言在 BBN 公司开发了

会话式编程语言, 专门用于解决数学问题。后来他们将注意力放在非数学问题上, 定义了一种简单到即使儿童也能学会的语言。而且他们将这种语言连接到“海龟”和其他机器人设备的控制装置上, 让年轻的学习者们平添动力、陡生乐趣。

Papert 在《头脑风暴: 儿童、计算机及强力理念》<sup>3</sup>中表述了这些想法。他还与建筑玩具制造商乐高公司合作开发了可用LOGO编程的“乐高头脑风暴”机器人套件, 并用他的书名来命名。第三代“乐高头脑风暴”于 2013 年推出(详见 [https://en.m.wikipedia.org/wiki/Lego\\_Mindstorms\\_EV3](https://en.m.wikipedia.org/wiki/Lego_Mindstorms_EV3))。该套件中既有软件又有硬件, 能够创建可定制、可编程的机器人。套件中包括: 一台砖状智能电脑, 可对系统进行控制; 一组模块化传感器和电机, 以及来自 Technic 生产线的乐高部件, 用于创建机械系统。(现在有多种不同的编程语言都可以在“乐高头脑风暴”中使用。)

## 人工神经网络争议与深度学习

人工神经网络的话题可追溯到 1943 年, 当时 Warren McCulloch 和 Walter Pitts 创建了一种用于神经网络的计算模型(详见 [https://en.m.wikipedia.org/wiki/Artificial\\_neural\\_network](https://en.m.wikipedia.org/wiki/Artificial_neural_network))。前面已经提到, 明斯基的博士论文对人工神经网络的早期研究做出了贡献。但在 1969 年, 明斯基和派珀特合著的一本书《感知器》<sup>5</sup>, 却引发了争议(争议的焦点是他们到底说的是什么), 这一争议直到今天

仍没有停息。在书中，他们阐明了“简单的人工神经网络”在计算能力上是相当有限的。可能他们假设“复杂的”人工神经网络（技术上称为“隐层人工神经网络”）的能力也是有限的，但在这本书中没有这样说。一位评论人士表示，“经常有人错误地引述明斯基和派珀特的话，导致人们对神经网络研究的兴趣和能够获得的经费显著下降”，他还说神经网络研究花了 10 年时间才在上世纪 80 年代得以复苏。

（《感知器》的扩充版<sup>6</sup>在 1987 年再版，指出并纠正了原版中的一些错误；详见 <https://en.m.wikipedia.org/wiki/Perceptron>）。另一位评论人士是麻省理工学院人工智能实验室的研究人员，他在 1971 年有关“岗巴网络”的“MAC 项目报告”（针对基金资助机构）中援引了明斯基和派珀特的话（详见 [https://en.m.wikipedia.org/wiki/Perceptrons\\_\(book\)](https://en.m.wikipedia.org/wiki/Perceptrons_(book))）：

人们对[这些机器]的计算能力几乎一无所知。我们相信，[它们]可以做得比低阶感知器多一点。

（要重温这段历史，请查看 [www.ucs.louisiana.edu/~isb9112/dept/phil341/histconn.html](http://www.ucs.louisiana.edu/~isb9112/dept/phil341/histconn.html)）

**到**本世纪初，隐层人工神经网络的强大能力已经得到了证明，事实上，也在为许多真正的人工智能应用程序提供了动力。隐层人工神经网络如今是系列深度学习算法的一部分（详见 [https://en.m.wikipedia.org/wiki/Deep\\_learning](https://en.m.wikipedia.org/wiki/Deep_learning)），这些算法推动着我们从“人工智能寒冬”迈入“人工智能盛夏”。

明斯基和派珀特的权威是显而易见的，他们对人工神经网络的消极看法可能是当时阻碍研究人员和研究资助者进一步开展人工神经网络研究的一个因素。当然，他们的权威源于他们对人工智能领域和联合领域所做的诸多重要贡献。人们将更多地记住他们的贡献，而不是争议。

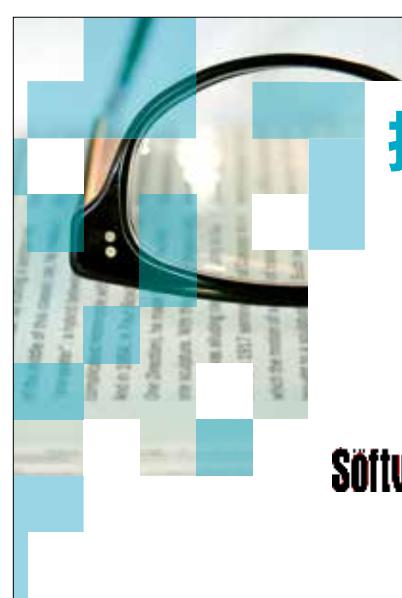
## 参考文献

1. S. Wolfram, *Idea Makers: Personal Perspectives on the Lives & Ideas of Some Notable People*, Wolfram Media, 2016.
2. M. Minsky, *The Society of Mind*, Simon & Schuster, 1988.
3. S. A. Papert, *Mindstorms: Children, Computers, and Powerful Ideas*, Basic Books, 1980.

4. M. Minsky, *The Emotion Machine: Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind*, Simon & Schuster, 2006.
5. M. Minsky and S. A. Papert, *Perceptrons*, MIT Press, 1969.
6. M. Minsky and S. A. Papert, *Perceptrons: An Introduction to Computational Geometry*, expanded ed., MIT Press, 1987.

**乔治·斯特劳**是美国国家科学、工程及医学科学院的研究数据与信息委员会主任。他是网络与信息技术研究和发展(NITRD)计划国家协调办公室的前主任。联系方式：gostrawn@gmail.com。

**坎迪斯·斯特劳**是一名退休的高中、社区学院及大学教师。联系方式：castrawn@gmail.com。



## 投稿

IEEE Software 寻找可以吸引专业和非专业读者的实用可读文章。这本杂志的目标是将可靠的信息传递给软件开发者和管理者，帮助他们站在技术浪潮之巅。投稿必须为原创，不能超过4700个词，每张表格和图片不能超过两百个词。

**Software** 投稿指南：  
[www.computer.org/software/author.htm](http://www.computer.org/software/author.htm)  
更多细节：[software@computer.org](mailto:software@computer.org)  
[www.computer.org/software](http://www.computer.org/software)



## 搜索你的工作机会

IEEE Computer Society 招聘可以帮你轻松找到IT、软件开发、计算机工程、研发、编程、架构、云计算、咨询、数据库很多其他计算机相关领域的的新工作。

**新功能：**找出那些建议或要求拥有IEEE CS CSDA或CSDP认证的工作！

点击[www.computer.org/jobs](http://www.computer.org/jobs)，  
从全世界的雇主那里搜索技术工作岗位和实习机会。

<http://www.computer.org/jobs>

IEEE  computer society | JOBS

IEEE计算机协会是AIP Career Network的合作伙伴。其他合作伙伴包括《今日物理》杂志（Physics Today），美国医学物理协会（American Association of Physicists in Medicine），美国物理教师协会（American Association of Physics Teachers），美国物理学会（American Physical Society），AVS科学和技术学会（AVS Science and Technology），物理学生协会（Society of Physics Students）和Sigma Pi Sigma。



# 人类增强技术

文 | 麦克·戴利（Mike Daily），HRL实验室（HRL Laboratories）  
安蒂·欧拉斯维尔塔（Antti Oulasvirta），芬兰阿尔托大学（Aalto University）  
历本纯一（Jun Rekimoto），东京大学（University of Tokyo）

计算技术正给我们的世界带来翻天覆地的变化，也很有可能改变我们自己。我们身体有着内在的脆弱性和限制性，所以这些技术进步可以让我们做更多麻烦的工作。本期的封面文章举例讨论了一些新的方法，可以增强人类的通信能力、感知能力和神经系统的稳定性等。



### 类 增 强 技 术(human augmentation)

是指那些可以增强人类生产力，改进或恢复人体或思维能力的计算技术。这一领域的技术进步可以加强、改进人类健康、生活质量和人体功能。人类增强技术的应用包括能加强感官能力的设备和医学植介入体，让人看到增强视觉内容的眼镜，感知处理听觉信息的下一代人工耳蜗，侵入性极低、可以增加肌肉功能的设备和义肢，以及改善学习和认知能力的神经刺激技术。

其他类型的人体增强技术需要一些特定的IT资源，包括云、大数据和移动计算。这些增强技术包括了连接人体和外部信息（视觉信息、听觉信息和文本信息）的手表和手环。在本期的封面文章中，我们将主要讨论基于计算的解决方案，而不是那些完全根植于药物

章，在专业的审稿人严格地评估之后，我们选择了4篇文章，可以代表人类增强技术的几个重要的方面。其中，有两篇文章讨论了视觉增强技术，一篇文章讨论了人体通信这一新概念，还有一篇文章描述了通过神经刺激改善脑功能的新方法。

在《定制3D打印智能眼镜，增强日常生活现实》一文中，弗洛里安·瓦尔（Florian Wahl）、张瑞、马丁·弗洛伊德（Martin Freund）和奥利弗·阿姆夫特（Oliver Amft）描述了一种为每位消费者定制智能眼镜的数字化开发过程。采用定制化模型设计，3D打印制作的内嵌传感设备的镜框在合适性和舒适度上的表现和传统眼镜相比毫不逊色。本文作者还着重介绍了其智能眼镜的若干应用，包括分析昼夜节律，监测和分析营养等。

特·依提（Laurent Itti）、玛丽·B·罗森（Mary B. Rosson）、约翰·M·卡罗尔（John M. Carroll）、凯文·M·伊瑞克（Kevin M. Irick）和维加克里施南·纳拉亚南（Vijaykrishnan Narayanan）讨论了一种创新的多模态增强技术，能够帮助视觉障碍人士（全球有2.85亿人以上）购物。他们设计的系统包含了一个可穿戴的配摄像头的眼镜，以及一个带振动反馈的手套。视障人士可以利用这套设备在商店中找到所需要的的商品，并在设备的帮助下将其从货架上取下放到购物车内。目前的测试结果让人充满信心，他们正在准备进行更大规模的评估。

在《基于人体传导声音信号的新型通信方式》一文中，作者张铖、希南·海尔赛克（Sinan Hersek）、蒲奕名、孙丹睿、薛秋月、萨德·E·斯塔那（Thad E. Starner）、格雷戈里·D·艾博德（Gregory D. Abowd）和奥默·T·伊南（Omer T. Inan）介绍了一种基于声音信号的方法，可以利用人体作为通信通道，在不同的设备间传输信息。作者们描述了如何使用这种技术，通过人体传输文本，在不同的人之间进行通信，以及让人体和体外设备通信。

在《神经技术能给我们更好的大脑吗？》一文中，作者拉赞·巴特查里亚（Rajan Bhattacharyya）、布莱恩·A·考夫曼（Brian A. Coffman）、周在勳

人类增强技术的进步可以改善人类健康、生活质量，以及人体功能。

学、心理学或者生理学的解决方法。

## 本期内容

我们收到了很多高质量的投稿文

在《“第三眼”：视觉障碍者的购物助手》一文中，彼得·A·津塔拉（Peter A. Zientara）、李素妍（Sooyeon Lee）、格斯·H·史密斯（Gus H. Smith）、洛瑞·布伦纳（Rorry Brenner）、劳伦

(Jaehoon Choe) 和马修·E·菲利普斯 (Matthew E. Phillips) 介绍了用于记录大脑信号和身体信号的非侵入式脑记录设备、非侵入式大脑刺激等辅助性神经技术如何用于增强人类大脑。他们探讨了这些技术在训练、探测险情、提高运动技能等方面的应用，以及与大脑监测、在人体上给予刺激等有关的困难和伦理问题。

**我**们希望读者能喜欢这期杂志的封面文章。我们感谢文章作者、审稿人和《计算机》杂志的主编以及编辑团队在制作过程中的辛勤工作和巨大帮助。人类增强技术将持续成为一个出现巨大进步的领域，并将对个人和社会产生巨大的价值。C

## 关于作者

**麦克·戴利**是HRL实验室系统和技术中心 (Systems and Technology Centers) 的经理。他的研究兴趣包括增强现实和虚拟现实、计算机视觉、自动系统和机器人，以及机器学习。戴利在新墨西哥大学获得计算机科学的硕士学位。他是IEEE的高级会员和《计算机》杂志的领域主编。联系方式：mjdaily@hrl.com。

**安蒂·欧拉斯维塔**是芬兰阿尔托大学 (Aalto University) 的电子工程助理教授。他领导用户界面研究组 (User Interfaces research group)。他的研究兴趣包括为人机交互建模并对其进行计算设计。他在赫尔辛基大学获得认知科学博士学位，也是《计算机》的编辑。联系方式：antti.oulasvirta@aalto.fi。

**历本纯一**是东京大学的教授，也是索尼计算科学实验室 (Sony Computer Science Laboratories) 的副主任。他的研究兴趣包括人机交互、计算机增强环境，以及用于人类的计算机增强感知。他在东京工业大学获得信息科学的博士学位。联系方式：rekimoto@acm.org。





# “第三眼”： 视觉障碍者的购物助手

文 | 彼得·A·津塔拉 (Peter A. Zientara)，李素妍 (Sooyeon Lee)，格斯·H·史密斯 (Gus H. Smith)，宾夕法尼亚州立大学 (Pennsylvania State University)

洛瑞·布伦纳 (Rorry Brenner)，劳伦特·依提 (Laurent Itti)，南加州大学 (University of Southern California)

玛丽·B·罗森 (Mary B. Rosson)，约翰·M·卡罗尔 (John M. Carroll)，宾夕法尼亚州立大学 (Pennsylvania State University)

凯文·M·伊瑞克 (Kevin M. Irick)，SiliconScapes 公司

维加克里施南·纳拉亚南 (Vijaykrishnan Narayanan)，宾夕法尼亚州立大学 (Pennsylvania State University)

译 | 马晓晓

我们结合可穿戴摄像头、硬件加速器和计算机视觉算法，设计了一款基于视觉信号的自动购物助手，帮助视力有障碍的人士轻松在商店内购物。

## 全

球约有2.85亿视觉障碍人士，对他们来说，很多正常人可以轻松完成的事（如捡起掉落的物体）也十分困难。近年来视觉感知技术领域取得

了很多突破，这些技术可以用来帮助视觉障碍人士，让他们更加独立。视觉增强技术的核心是信号处理技术，可以帮助视觉障碍人士在室内室外行走以及识别物体。<sup>1</sup>这些能力在日常生活中很重要，如购物等活动。美国一般的超市占地4.5万平方米，30多个通道内可能有35000个商品，即使是视觉正常的人，在

超市里也会被五花八门的商品搞得眼花缭乱。对视觉障碍人士来说，购物就像在足球场这么大的空间里走动，每个转角都有购物车、人流和物品。

面对如此复杂艰难的任务，视障人士通常只能向家人、朋友和社区志愿者求助。有些人还会找商场的导购帮忙，但这样的购物体验就很依赖于导购的业务熟练度及脾气性格，可能会很舒适也可能不开心。在线购物和快递服务也能在一定程度上降低视障人士的购物难度，但是它们虽然可以让顾客避

免去商场,<sup>2,3</sup>可却是以减少社会活动和购物选择为代价。另一种辅助技术是在移动设备上运行的计算机视觉程序。它们的基本功能是判断某种商品是否在货架上,但在实际复杂环境中,长相类似的多个商品往往并列放置,程序就无法正常工作。举例来说,即使是很先进的算法也很难分辨同一个谷物早餐的不同口味。在密集背景中准确识别物体需要准确的分割算法,可是这往往需要用计算量巨大的暴力搜索,这在实际应用中响应太慢。

为了解决这些问题,并给视觉障人士提供更加人性化的购物体验,我们开发了一款基于视觉的自动购物助手,能够帮助用户在一般的货架上挑选商品。本工作属于由NSF资助的“芯片视皮层”计划。亚洲文化中常常用第三只眼代表智慧和激情,所以我们把产品命名为“第三眼”。“第三眼”提供了方便的用户使用界面,解决了在繁杂的视觉场景中的导航难题。我们通过观察和采访视障人士,设计了“第三眼”的导航界面,还利用了用户原来的一些技能——如用手杖辨识障碍物等。

为了提升导航和物体选择能力,“第三眼”不仅利用了多模态的触觉和听觉交互,还需要把用户本身置入系统内部,即系统通过和用户交互实现导航和购物功能。举例来说,系统可以要求用户调整可穿戴相机的角度,使之更适

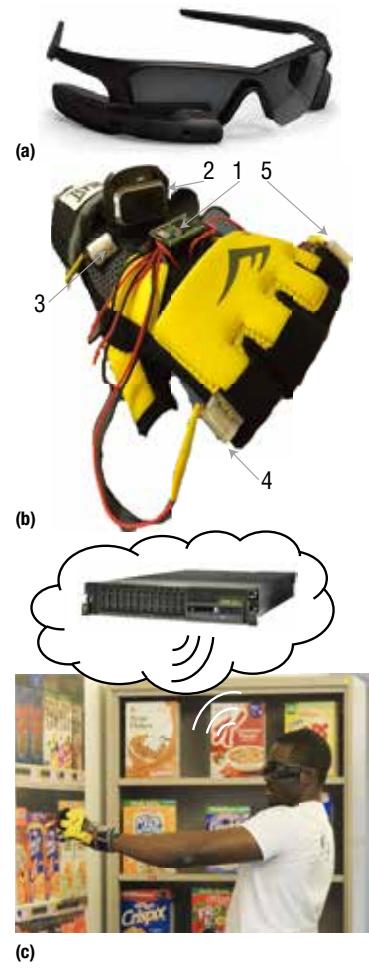
合图像分类。我们已经根据视障用户的使用反馈对系统进行了多次评估和调整,对什么是有效的沟通界面有很深刻的理解。实验参与者均对系统赞誉有加,说比起条码扫描器,“第三眼”让他们感觉更像“真正的购物”。

## 系统架构

如图1所示,“第三眼”包含了一款普通智能眼镜——配有摄像头和音频通道,眼镜与后端服务器相连,可以进行实时视频分析。相机面向用户的右眼方向,智能眼镜的视野及分辨率确定了哪些物体可以被定位并识别出来。用户还会戴上一个配有摄像头的手套,用来抓取商品。手套摄像头 用来指挥手的运动,同样可以帮助用户把商品维持在视野内,同时避免取商品时手臂遮挡镜头图像。

“第三眼”通过商店内的无线设备把视频数据传回服务器,并进行算法分析再传回用户。计算结果被用来产生音频命令或不同的触觉振动模式,指导用户靠近想要的商品。为了减少图像捕捉和反馈信号之间的时间延迟,“第三眼”利用了本地的无线路由设备,确保与服务器之间的通信稳定且快速,并对硬件进行优化加速了视频处理算法的计算速度。

智能眼镜捕捉到的图像是指导用



**图1.** “第三眼”工作机制。图中智能相机(a)和手套(b)上各有摄像头捕捉视频信号,通过无线传输到服务器(c)进行分析。系统根据分析结果给用户行动指示,可能是分析眼镜图像后用于导航的语音指示,也可能是指挥取物体的手套振动反馈。振动反馈的位置和下一步要移动的方向相关。手套包含六个部分:(1)控制和通讯模块,(2)摄像头,(3)指示向上的电机,(4)指示向右的电机,(5)指示向左的电机,以及图中没有显示的手掌下侧的指示向下的电机。

户移动的信息基础,然后手套摄像机继续指导用户调整到合适的视角,给算法提供足够视觉信息。有一些研究利用触感手套辅助视觉搜索,并和智能眼镜交互。<sup>4,5</sup>我们受其启发,在手套中加入了振动反馈。手套上装有四个微电机指示运动,四个电机产生振动指导手移向目标位置的商品。尽管之前人们会使用声



图2. 版本A的SURF算法。(a)白色箭头表示方向指令，通过语音传达给用户，使眼镜的摄像头能至少看到目标物体的一部分。(b)三个产品模板在当前图像中的检测置信度。“第三眼”把这些点和模板图像比较，计算出每个物体置信度，并根据该结果决定当前样本是否是目标商品，判断用户是否需要移动到下个物体。彩线横线代表了置信度阈值。

音反馈指导用户运动，但根据我们对视障用户的观察，我们认为加上振动提供触觉反馈更为合适。我们发现仅靠声音反馈在嘈杂的环境中效果并不好，而且会给用户招来一些不必要的注意。头戴设备可能能够解决这些问题，但这类设备需要定制，结合骨头传声或是在不断环境噪声的情况下叠加上声音指令，因为这些声音对视障人士也是重要的导航信息。

## 视觉算法

从用户进入商店的那一刻开始，“第三眼”的视觉算法就开始帮助用户，直到他们离店。首先要做的是识别出正确的通道。当用户站在通道的商品前时，算法会根据摄像机捕捉到的图像的大致场景信息分析用户所处位置。<sup>6</sup>算法从图像中抽取足以做场景分

类的低维信号。它累计每帧的场景统计量——如色彩分布、亮度、不同朝向的边缘信息等。这些累计后的统计信息总结了场景内容，并被转换为一个向量——一个低维整体特征向量，再由支持向量机分类，为图像贴上场景标签，如谷物食品通道或是咖啡通道。

假如用户所在通道不正确，那么他们需要走向下个通道，并重复上述过程直到抵达目标为止。目前在通道间的导航还没能做到完全自动化，但我们的系统有效地结合了视障人士自身的导航能力。一旦用户到了正确的通道，“第三眼”用SURF(Speeded up Robust Features)算法，把图像中(可能不完整)的物体和已知物体做特征匹配。<sup>7</sup>“第三眼”还通过提供输入的置信度对SURF进行拓展。算法可以根据这个分数把摄像头引导到更好地视角。拓展后的模板匹配算法使得“第三眼”可

以把镜头中的关键点和模板图像匹配上。每个关键点匹配均返回新图像中的位置以及匹配的置信度。图2中有一些第一个版本(版本A)的SURF结果，我们后来又改进为版本B。图2a中，“第三眼”引导用户到合适的位置，使产品的至少部分图像进入视野，然后才能使用SURF进行特征匹配。

## 可信度匹配

大部分产生置信度的算法会使用一个固定阈值，用来决定目标物体是否在图像内出现，或是直接选择置信度最高的物体。大部分用固定阈值的算法都会产生一定错误，因为当某些假正样本(false positive)的置信度比某些真正样本(true positive)还高时便会出现误差。<sup>8</sup>最可疑的假正样本是目标不在图像中时置信度最高的物体。同样，最难的真正样本是当目标出现时对应的置信度最低的样本。一旦最可疑的假正样本比最难的真正样本还可信时，算法就会有问题，因为此时系统无法准确判断目标是否出现。在这种情况下，我们的算法不给视障患者提供判断，而仅仅是指导用户向提高置信度的方向移动，直到系统可以对图像做出准确分类。

例如图2b中，白色横条表示最难真正样本的置信度，黄色横条代表了最可疑的负正样本。黄色到橙色横条的距离是负正样本之上的安全缓冲带。样

本的置信度必须超过橙色横条，“第三眼”才会判断目标物体在图像中出现。确定这些置信阈值后，左侧的绿柱置信度不够，所以“第三眼”会指示用户调整视角。中间的绿柱超过了安全缓冲区，所以“第三眼”会判断目标物体已出现在视野中。最右侧绿柱比最难真正样本还低，所以系统会忽略这种情况，不会给出任何指示。

有了如图2b一样的置信度数据后，“第三眼”可以把模板图像和摄像机镜头匹配关键点，从而根据关键点相对位置计算出单应性矩阵(homography matrix)。矩阵表征了当前物体相对模板图像的朝向，我们实验中所用的模板图像均位于正面视野的中央。相对朝向信息可被用来指导用户调整视角。

## 匹配整个货架

在A版本算法中，算法的目标是让用户视角接近模板的视角。摄像机拍摄的图像越接近这个视角，置信度就越高。然而当货架上同时存在若干个相似的物体，且用户距离货架的距离还没有近到足以分辨其细节特征时，算法就无法找到部分匹配以计算单应性矩阵，也就不能提高置信度。为了解决这个问题，“第三眼”做了货架尺度的特征提取及匹配，该算法以商店的排货图表(商品在货架上是如何排列的)为基础，指示用户靠近目标商品所在的货架区域。

当算法找到目标货架后，“第三眼”利用排货图表定位所需商品的位置，并只分析对应区域的视频信息。比如一个购物者想买某个牌子的意面，“第三眼”就不会去分析整个“米、面”过道，而是仅仅分析包含面的货架及包含意面的那一区域。通过缩小分析范围，A版本算法面临的一些问题得以解决，但此时购物者还是需要等待系统判断哪些图像是不相关的，才能听到下一步语音指示。在版本B算法中，我们对整个货架进行了关键点匹配，用户甚至

据手套摄像头指导手部运动。

我们观察到根据头部朝向自动进行切换并不稳定，比如头相对物体的角度是依赖于使用者身高的。在一次视障志愿者参与的实验中，我们发现视障用户本身更加擅长决定何时进入取物模式。和蒙眼的正常人不一样，视障用户通常会使用手杖辅助行走，所以可以准确地停在货架前一臂的距离。我们发现即使“第三眼”的深度提取算法失效了，用户依然能知道何时止步，所以我们让用户决定何时从智能眼镜切换到手套

**“第三眼”做了货架尺度的特征提取及匹配，该算法以商店的排货图表(商品在货架上是如何排列的)为基础，指示用户靠近目标商品所在的货架区域。**

可以站在12英尺之外。因此算法会找到更多特征点，可以准确区分货架上外观相似的物品。

摄像头，而不是由系统计算决定，如图3所示。

## 模式切换

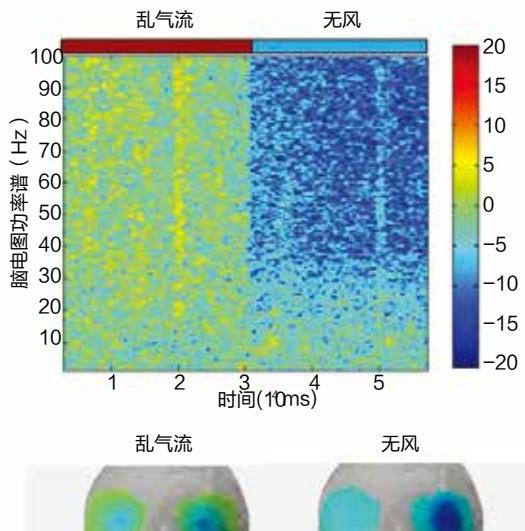
当用户走到货架面前时，系统便将信号输入源切换为手套摄像头，即不再指挥用户如何行走，而是指导其手的运动。要想执行这一切换，我们首先根据物体在镜头中的相对大小判断用户和货架的距离。在一定距离时“第三眼”会不再依赖智能眼镜的视频数据，而是根

## 系统优化

我们对“第三眼”做了反复优化调整，发现了问题主要由两个：一个是反馈延迟问题，我们最终解决了；另一个是耗电问题，我们也找到了减轻办法。

## 反馈延迟

我们的算法在IBM POWER8服务



**图3.**由用户切换两个检测模式，从智能眼镜转为手套摄像头。(a)“第三眼”将商店的排货安排作为先验信息，根据货架匹配结果指挥用户走到正确位置。(b)当用户来到目标商品跟前时，“第三眼”利用物体识别算法指导手的移动。当用户的的手杖碰到货架后，用户手动切换到取物模式，同时关闭了“第三眼”的语音指示。系统继而转用手套摄像头作为视频输入，并用手套的振动反馈指导用户移动手臂。

器上运行，服务器有160个3.6GHz逻辑核心，即使如此，对于一张 $1920 \times 1080$ 的过道图像，算法在产生9000个特征点时依然有375ms延迟。通过把视野缩小到单个货架，延迟被降低为170ms，产生约4000个特征点。除此以外，我们发现当购物者动作突然变化时，系统的延迟反馈会导致摄像机丢失物体。为了解决这个问题，我们使用了NVIDIA GPU K1200加速物体检测（货架或物品），可以把延迟降低到约110ms，这足以给用户使用了。基于K1200的测试结果，我们还设计了在FPGA上运行的加速算法，可以进一步减低延迟，实现真

正实时处理。我们目前还在测试该加速框架。

### 能耗问题

在早期版本的“第三眼”中，智能眼镜的电池寿命很短。一开始我们把原始视频信号全部无线传回服务器，不仅耗电还费带宽。于是我们尝试在智能眼镜中的处理器上运行基于软件的压缩算法。但这并没能解决问题，因为压缩算法十分耗时。所以我们最后直接调整智能眼镜硬件，结合了硬件加速的视频压缩技术，使智能眼镜的电池寿命达到了20分钟。

还有一个可能的解决方案是利用智能眼镜能提供的其他信息。不是使用耗电的视频反馈，而是通过加速度计或陀螺仪跟踪购物者的运动。这样“第三眼”可以根据排货图表，只在相关位置分析视频图像。我们已经把类似于视频处理中的技术应用于加速度计信号分析，可以分辨用户是否在步行、右转或左转等。

## 结果展示

我们首先确认整个系统是否可行，在不同难度的环境中测试了“第三眼”的表现，并同时探索可能的改进。在用完整的产品测试之前，我们先用人类观察者代替计算机视觉算法进行测试。这些早期研究对我们设计用户反馈和指导系统提供了很多样本数据。

### 可行性研究

我们招募了42位正常人志愿者，让他们蒙上双眼后进行试验。他们戴上智能眼镜，由版本A的SURF算法（没有匹配货架，也没有排货图表信息）提供语音指示。每次试验中，我们让被试者从距离目标商品5英尺处出发。在这个距离，算法通常能够获得足以部分匹配物体的特征点，但有时候“第三眼”连部分匹配目标也找不到，这种情况下被试不得不自行尝试移动。

**实验过程:** 在测试之前, 我们首先训练被试理解语音指示, 实现一些简单动作, 如转身、前进、后退、扫视或蹲下。“第三眼”通过这种方式找到新的特征点。一旦被试找到了合适的位置, 我们会要求他们伸出手臂取下目标物体。把定位和抓取分为两个阶段可以保证用户的手不会遮挡摄像头视线。我们还发现语音指示和姿势调整在排除假正样本时也很有效。在每个被试的10次试验中该方法全部成功。如图4a所示, 从听到第一次语音指示开始到拿到指定物品的耗时在7.1秒到798.8秒之间, 大部分测试中耗时在30秒到90秒之间。<sup>9</sup>

**与条码扫描比较:** 为了和标准的辅助购物技术比较, 我们还给每让每个被试用条码扫描器寻找商品, 每个商品的下方贴有自己的条码。在这个测试中, 用户靠近货架扫描各个条码, 除了当前条码是否是目标商品以外, 没有其他任何反馈。所有被试也都成功用这种方式找到了商品, 但从图4b可以看到, 34名参与者中的32位都明确表示更偏好(打分6分以上)用“第三眼”提供的连续指示方案而不是使用条码扫描器。不止一位参与者提到, “这更像是真正的购物。”

### 自动切换

之后我们接着测试了版本B的

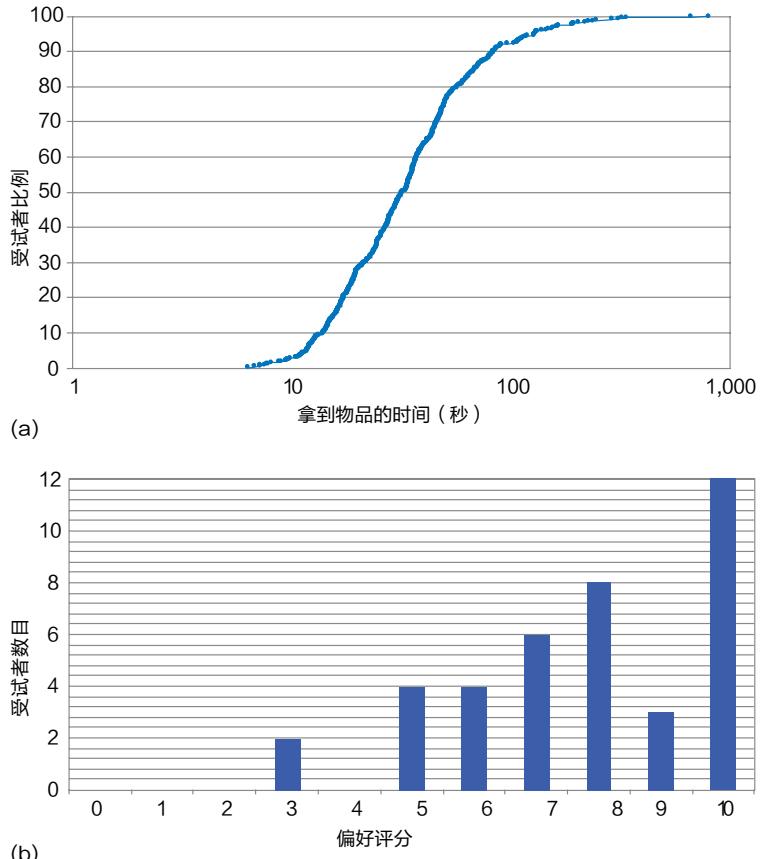


图4. 比较“第三眼”和条码扫描器的用户体验。(a) 从听到语音指示到拿到物体的耗时分布, 实验中共有42位被试, 每个被试做了10次试验。(b)39位被试提供的偏好评分。如图所示, 12位用户给“第三眼”打10分, 说明他们“十分倾向”使用“第三眼”而不是现有的条码扫描器。一半以上的用户给“第三眼”的评分在7或以上。

SURF算法的自动切换环节。在这个实验中, 我们让蒙着双眼的正常人被试从货架的10英尺开外出发, 仅提供语音反馈, 几乎没有先验指示。大部分情况下, 语音反馈系统可以成功把用户带到指定商品前方。有一些被试偶尔会误解语音指示, 比如把“移向左侧”理解为“转向左侧”。同时, “移向左侧”和“移向前方”等指示并没有明确移动多少距离, 也导致用户移动的距离有远有近。基于这些结果, 我们把语音指示改得更具体, 如“向前移动一步”或“向左移动不到一步”等。

### 仅用手套

在另一个测试中, 我们试验了仅靠手套摄像头和振动反馈被试能否在抓到3英尺内的目标物体。这个测试中使用的是版本A的SURF算法。最主要的挑战是摄像头的摆放位置, 因为我们必须找到合适的角度使摄像头视野不被手所遮挡, 因为这个镜头提供的视觉信息是系统最重要的输入来源。在“第三眼”原型机中, 摄像头被放置在平行于手臂的位置, 这样手和镜头角度方向一致。但是在一些实验中, 摄像机朝向会被一些突然的运动改变, 结果丢失视野

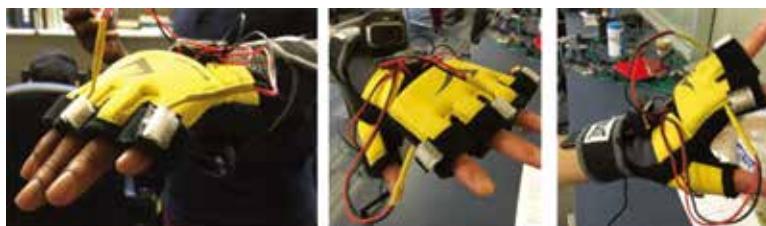


图5. 手套摄像头的位置。红色箭头是计算运动方向时算法所用的视角，绿色虚箭头是用户理解运动指令所用的朝向。(a)两个角度的差别使用户远离了目标物体。(b)角度正确对准。(c)角度没有偏差，但手掌遮挡了摄像头视野。

且指向错误物体。图5展示了我们遇到的一些设计难题。

我们还实验了一些手套振动反馈的设置。在一些试验后，我们决定把微电机安置在左手的手指上指导用户移动：拇指振动表示向右移，小指表示向左移，手掌上方表示向上移，手掌下方表示向下移。我们继续研究了不同的振动强度、时长和模式，看哪种参数最适合作为触觉反馈。<sup>10</sup>

### 视障用户测试结果

之前的试验中用的都是蒙上双眼的正常人，但我们意识到这些志愿者和真正的视障人士还是有区别，因为视障人士已经适应了视觉受损的状态。我们最近才在一位视障人士身上测试了“第三眼”的效果，这位志愿者参与过我们过去两年内的NSF项目，在一个模拟的商店货架前测试过一些基本的购物过程。这位志愿者以前还参与过人类识别设备的研究，以及视障人士购物体验的研究。这些过去的经验让他成为了给我们提供使用反馈的绝佳人选。

**自动切换。**我们给这位视障志愿者戴上智能眼镜，并只给自动切换模式（版本B算法）的声音反馈，但不给振动

反馈。志愿者的起点是距离模型货架约10英尺，货架上放了9种谷物通道的产品。然后我们用语音给志愿者提供产品目录，由她自行选择目标产品。选定后她可以按照语音指示来到目标产品附近，系统所用的信息是眼镜摄像头的视频数据。语音反馈（如“稍微向左移两步”）既包含了方向信息，也包含了距离指示，可以让志愿者保持在正确的路线上。

视障志愿者成功抵达目标物体前方后，我们发现她的头的朝向并不如蒙眼正常人稳定。我们还发现系统对用户和货架之间距离的估计并不总是可靠。尽管如此，志愿者会一直利用“第三眼”的指示向目标移动，并在自己的手杖触碰到货架时自行止步。

**用户控制切换。**基于自动切换模式的测试结果，我们又试验了由用户控制模式切换的情况，这种设置下视障用户通过手杖把系统从智能眼镜模式切换到手套摄像头模式。当用户的手杖碰到货架底部时，用户将告诉“第三眼”关闭基于智能眼镜的语音反馈。在这个实验中，视障志愿者能够在下一次试验里拿到目标物体。然而在另外一些试验中，志愿者会选中目标物体下方的商品。这些结果说明在视野上方

的物体更难处理，我们还需要对系统做进一步优化。

**手套配置。**在最后一项实验中，我们测试了不同的手套配置，看摄像头应该摆在什么角度，以及是语音反馈还是触觉反馈更为有效。视障志愿者似乎更偏好振动反馈，她说没有头戴设备的语音反馈会使购物过程招来不必要的注意，而且会干扰她用来判断方向的环境中的其他声音。

我们还发现手套摄像头的角度需要更稳定的工程设计。由于朝向没有对准，在所有试验中，志愿者最后总是和目标物体有几英寸误差。不仅如此，一旦手伸向了错误位置，要么手套摄像头的视野被遮挡，要么用户和目标距离过近，这都导致系统无法准确检测到目标商品。在这种情况下，系统会不再给出任何方向指示，因为它们超过了最高置信分数（图3b）。

经过30分钟（包括系统暂停），视障用户在系统的振动和声音指挥下成功拿到了指定的谷物早餐商品，放到了购物车里。不幸的是，我们不得不把很多其他商品放回货架，因为她经常选到紧挨着目标的其他商品。



实验结果启发，我们给系统添加了一些新的模块，保证视障用户花更多时间

在把商品放到购物车里，而不是放回货架。我们给“第三眼”添加了跟踪手位置的功能，利用眼镜摄像头的视频让最终位置更加准确，还添加了文字识别系统作为模板图像验证。我们还微调了货架深度估计界面，所以从身体移动到手部移动的两个阶段可以无缝衔接。以这些为基础才能开发出结合了基于货架图像的大尺度身体运动和基于近距离观察的手臂运动的混合系统，实现从不同视角之间的无缝切换。

我们对系统做了部分和整体的分析，这使得我们在招募更多视障志愿者进行实验前可以想清楚如何优化“第三眼”。我们已经在K12科技展上展示过目前的版本，也在国会成员面前展示过我们的产品。我们很期待看到“第三眼”后续版本的使用效果。■

## 参考文献

1. H. Guizzo and H. Goldstein, “The Rise of the Body Bots,” *IEEE Spectrum*, vol. 42, no. 10, 2005, pp. 50–56.
2. D. Dakopoulos and N.G. Bourbakis, “Wearable Obstacle Avoidance Electronic Travel Aids for Blind: A Survey,” *IEEE Trans. Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 40, no. 1, 2010, pp. 25–35.
3. J. Nicholson, V. Kulyukin, and D. Coster, “ShopTalk: Independent Blind Shopping through Verbal Route Directions and Barcode Scans,” *The Open Rehabilitation J.*, vol. 2, no. 1, 2009, pp. 11–23.
4. V. Kulyukin and A. Kutiyawala, “Accessible Shopping Systems for Blind and Visually Impaired Individuals: Design Requirements and the State of the Art,” *Open Rehabilitation J.*, vol. 3, 2010, pp. 158–168.
5. Y.T. Hsieh et al., “Designing a Willing-to-Use-in-Public Hand Gestural Interaction Technique for Smart Glasses,” *Proc. ACM CHI Conf. Human Factors in Computing Systems (CHI 16)*, 2016, pp. 4203–4215.
6. V. Lehtinen et al., “Dynamic Tactile Guidance for Visual Search Tasks,” *Proc. 25th ACM Symp. User Interface Software and Technology (UIST 12)*, 2012, pp. 445–452.
7. C. Siagian and L. Itti, “Rapid Biologically-Inspired Scene Classification using Features Shared with Visual Attention,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 29, no. 2, 2007, pp. 300–312.
8. H. Bay, T. Tuytelaars, and L. Van Gool, “SURF: Speeded Up Robust Features,” *Proc. 9th European Conf. Computer Vision (ECCV 06)*, 2006, pp. 404–417.
9. M. Merler, C. Galleguillos, and S. Belongie, “Recognizing Groceries in Situ Using in Vitro Training Data,” *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR 07)*, 2007; [www.researchgate.net/publication/221362362\\_Recognizing\\_Groceries\\_in\\_situ\\_Using\\_in\\_vitro\\_Training\\_Data](http://www.researchgate.net/publication/221362362_Recognizing_Groceries_in_situ_Using_in_vitro_Training_Data).
10. R. Brenner, J. Priyadarshi, and L. Itti, “Perfect Accuracy with Human-in-the-Loop Object Detection,” *Proc. European Conf. Computer Vision (ECCV 16)*, 2016, pp. 360–374.
11. S. Brewster and L.M. Brown, “Tactons: Structured Tactile Messages for Nonvisual Information Display,” *Proc. 5th Australasian User Interface Conf. (AUIC 04)*, 2004, pp. 15–23.

## 关于作者

**彼得·A·津塔拉**是宾夕法尼亚州立大学计算机科学与工程系的博士生。他的研究兴趣包括智能传感器、硬件加速即嵌入式视觉系统。津塔拉在宾夕法尼亚约克学院获得计算机工程本科学位。联系方式: paz117@cse.psu.edu。

**李素妍**是宾夕法尼亚州立大学信息科学与技术学院的博士生。她的研究兴趣包括和辅助技术和社区信息学相关人机交互技术。李在宾夕法尼亚州立大学信息科学与技术学院获得本科学位。联系方式: sull31@psu.edu。

**格斯·H·史密斯**是宾夕法尼亚州立大学计算机科学与工程系的硕士生。他的研究兴趣包括计算机视觉算法应用、硬件加速结构和新兴设备。联系方式: hfs5022@psu.edu。

**洛瑞·布伦纳**是南加州大学的神经科学博士生。他的研究兴趣包括与人交互的计算机视觉研究、辅助视觉系统、物体识别和人工神经网络。布伦纳在卡耐基梅隆大学获得计算机科学本科学位。联系方式: rorry.brenner@usc.edu。

**劳伦特·依提**是南加州大学计算机、心理系和神经科学教授。他最新的研究包括由生物启发的计算机视觉，特别是在视觉注意力、场景、显著性、新奇性等领域，在视频压缩、目标检测和机器人上均有应用。依提在加州理工学院获得了计算与神经系统博士学位。联系方式: itti@usc.edu。

**玛丽·B·罗森**是宾夕法尼亚州立大学信息科学和技术学院的教授和副院长。她的研究兴趣包括基于场景的交互系统设计、用户端编程与开发、网络发展社区。罗森在德克萨斯州大学奥斯汀分校获得了实验心理学博士学位。她是ACM资深科学家。联系方式: mrosson@psu.edu。

**约翰·M·卡罗尔**是宾夕法尼亚州立大学信息科学和技术学院的资深教授，同时还是脑机交互中心的主任。他的研究兴趣包括以人为中心的设计方法和理论，特别是应用在一些网络工具上，帮助合作学习、解决问题、创造并维持相应人群。卡罗尔在哥伦比亚大学拿到了实验心理学博士学位。他是IEEE和ACM会员。联系方式: jcarroll@ist.psu.edu。

**凯文·M·伊瑞克**是SiliconScapes公司的CEO和创始人。他的研究兴趣包括高性能、嵌入式的云视频分析系统，可用于安保、监控和销售应用。他还关心针对应用的硬件加速设计、硬件辅助的图像处理和识别技术、可重配置的结构中的高性能计算。伊瑞克在宾夕法尼亚州立大学获得了计算机科学和工程博士学位。联系方式: kevin.irick@siliconscares.net。

**维加克里施南·纳拉亚南**是宾夕法尼亚州立大学计算机科学和电子工程教授。他的研究兴趣包括续航监测稳定系统、嵌入式系统、纳米设备、系统架构交互、可重配置系统、计算机结构、芯片集成网络和领域特异计算。纳拉亚南在南佛罗里达大学获得计算机科学和工程博士学位。他是IEEE和ACM会员。联系方式: vijay@cse.psu.edu。

### 译者简介：

**马晓晓**，清华大学医学院BrainMatrix实验室博士生，研究对大脑认知过程的数学建模。

# DNN

A thousand words will not leave so  
**deep** an impression as one deed.

---

行胜于言



LUCKY DATA  
搜狐图文匹配算法大赛

搜狐 | IEEE-China | 清华大学



[www.luckydata.org](http://www.luckydata.org)



# 定制 3D 打印智能眼镜， 增强日常生活现实

文 | 弗洛里安·瓦尔 (Florian Wahl)，张瑞 (Rui Zhang)，马丁·弗洛伊德 (Martin Freund)  
和奥利弗·阿姆夫特 (Oliver Amft)，帕绍大学 (University of Passau)

译 | 张哲

通过使佩戴者监测自己生命体征和行为的增强现实应用程序，配有感知功能的定制 3D 打印眼镜能改善日常生活。

**大**多数人都认为眼镜既可以矫正视力又是一个时尚配饰；如果加入嵌入式技术，眼镜将来能更大程度改善佩戴者的日常生活。眼镜正好戴在连续监测身体、生理和环境参数的关键位置，因此能够支持各种监测健康的应用程序。未来的智能眼镜会有很多用途，比如追踪老年用户的日常活动，对长时间使用电脑的用户发出警告，以及帮助节食者更好地选择吃什么（参看“智能眼镜发展趋势”插文）。与智能手机不同，推动人们使用智能眼镜的是快速的数字化开发过程和定制化制作以确保佩戴的舒适性和可用性。

如果使用普通框架眼镜来制作定制化眼镜，配镜师通常会加热塑料眼镜框来使其发生形变从而适应佩戴者。但是当眼镜中有电子元件时就不能使用这种方法，因此智能眼镜只能在制作时就直接定制。全新的3D打印技术和材料是定制化智能眼镜的基础，比如镜框是数字打印的，其他功能可通过嵌入集成电子元件实现。在装上镜片后，这些眼镜就能被用于监

测佩戴者的生命体征和行为，所需数据来自传感器。

## 定制化智能眼镜

为了定制智能眼镜，我们首先分析了影响镜框形状的头部和面部特征，并得到了一组相互独立的镜框参数（这些可调节参数可以自由选择，因此代表了头部模型的自由度）。我们修改了一个常见的镜框设计，把传感器和处理器电子元件嵌入镜框和镜脚。这些镜框参数之后可适用于不同佩戴者。

## 头部建模

由于不同人的头型和脸型相差很大，我们使用了 MakeHuman 这个开源人类模拟软件来确定影响镜框设计的关键头部特征。在 146 个影响头型和脸型的参数中，我们选择了其中影响镜框设计的 26 个。舍去了其中效果重复的参

## 智能眼镜发展趋势

**智**能眼镜的多种用途表明，其在人类增强中会成为一个重要的监测组分。不过，传感器测定数据要想可靠，前提是眼镜得准确贴合佩戴者，因此定制化至关重要。接下来我们会概述已有的以及我们自己关于智能眼镜的研究。

由于认为数据监测的位置可以选在头部，奥利弗·阿姆夫特(Oliver Amft)和同事发表过一篇关于智能眼镜技术和应用机会的概述<sup>1</sup>。石丸翔也(Shoya Ishimaru)和同事使用了称作J!NS MEME的眼电图原型技术来进行基于眼球运动的活动识别<sup>2</sup>。弗洛里安·瓦尔(Florian Wahl)和同事在普通镜框上搭建了智能眼镜的原型，并评价了惯性传感器和其他传感器识别不同活动类型的效果<sup>3</sup>。张瑞(Rui Zhang)和同事使用肌电图电极研究了智能眼镜监测饮食的效果<sup>4</sup>。其他研究工作还包括谷歌眼镜等头戴式可穿戴设备的其他应用。沙哈·阿提克·拉赫曼(Shah Atiquar Rahman)和同事使用谷歌眼镜监测日常生活中的饮食活动<sup>5</sup>，哈维尔·埃尔南德斯(Javier Hernandez)和同事使用谷歌眼镜陀螺仪来测定了佩戴者不动时的心率<sup>6</sup>。

智能眼镜若想为人们接收，那么设计上就不能太另类。例如，玛丽昂·科勒(Marion Koelle)和同事发现，旁观者看到有人带着头戴式显示器会觉得不舒服，因为前者不知道后者想做什么<sup>7</sup>。此外，谷歌眼镜中的摄像头引发了很多关于隐私的讨论，降低了该设备的使用率。智能眼镜需要谨慎选择设计、传感器和彼此的交互，从而保证不引人注意并实现真正的可佩戴带。

### 参考文献

- O. Amft et al., "Making Regular Eyeglasses Smart," IEEE

Pervasive Computing, vol. 14, no. 3, 2015, pp. 32–43.

- S. Ishimaru et al., "Smarter Eyewear: Using Commercial EOG Glasses for Activity Recognition," Proc. 2014 ACM Int'l Joint Conf. Pervasive and Ubiquitous Computing: Adjunct Publication (UbiComp 14 Adjunct), 2014, pp. 239–242.
- F. Wahl, M. Freund, and O. Amft, "WISEglass: Multi-purpose Context-aware Smart Eyeglasses," Proc. 2015 ACM Int'l Symp. Wearable Computers (ISWC 15) 2015, pp. 159–160.
- R. Zhang, S. Bernhart, and O. Amft, "Diet Eyeglasses: Recognising Food Chewing using EMG and Smart Eyeglasses," Proc. Int'l Conf. Wearable and Implantable Body Sensor Networks (BSN 16), 2016; doi: 10.1109/BSN.2016.7516224.
- S.A. Rahman et al., "Unintrusive Eating Recognition Using Google Glass," Proc. 9th Int'l Conf. Pervasive Computing Technologies for Healthcare (PervasiveHealth 15), 2015, pp. 108–111.
- J. Hernandez et al., "BioGlass: Physiological Parameter Estimation Using a Head-mounted Wearable Device," Proc. 2014 EAI 4th Int'l Conf. Wireless Mobile Communication and Healthcare (Mobihealth 14), 2014, pp. 55–58.
- M. Koelle, M. Kranz, and A. Moeller, "Don't Look at Me That Way! Understanding User Attitudes Towards Data Glasses Usage," Proc. 17th Int'l Conf. Human-Computer Interaction with Mobile Devices and Services (MobileHCI 15), 2015, pp. 362–372.

数后最终留下了3个。此外，还可以使用基于光线的形状扫描设备等替代方法，但是耳朵和头发可能会影响测量结果。

### 镜框参数模型

我们使用计算机辅助设计(CAD)软件建立了一个和常规眼镜非常相像的

3D镜框模型。我们的目标是用尽可能少的头部测量值来建立一个镜框参数模型。我们首先把之前已经确认的影响头型和脸型的26个参数削减至6个相互独立的镜框参数。由于鼻托对称性这个参数实际关联性较弱，我们又忽略了这个参数，最终精简到5个。

图1描述了我们数字化开发过程的

步骤及其参数的细节。图1c展示了头型和脸型的参数以及最终使用的镜框参数。为了确保佩戴舒适，给镜腿建模时使用的参数有长度、角度和弯曲程度。用于调节鼻梁的参数有鼻子宽度和高度以及镜片高度和宽度。准确测量镜片高度非常重要，因为这可以防止镜框接触到佩戴者的颧骨。

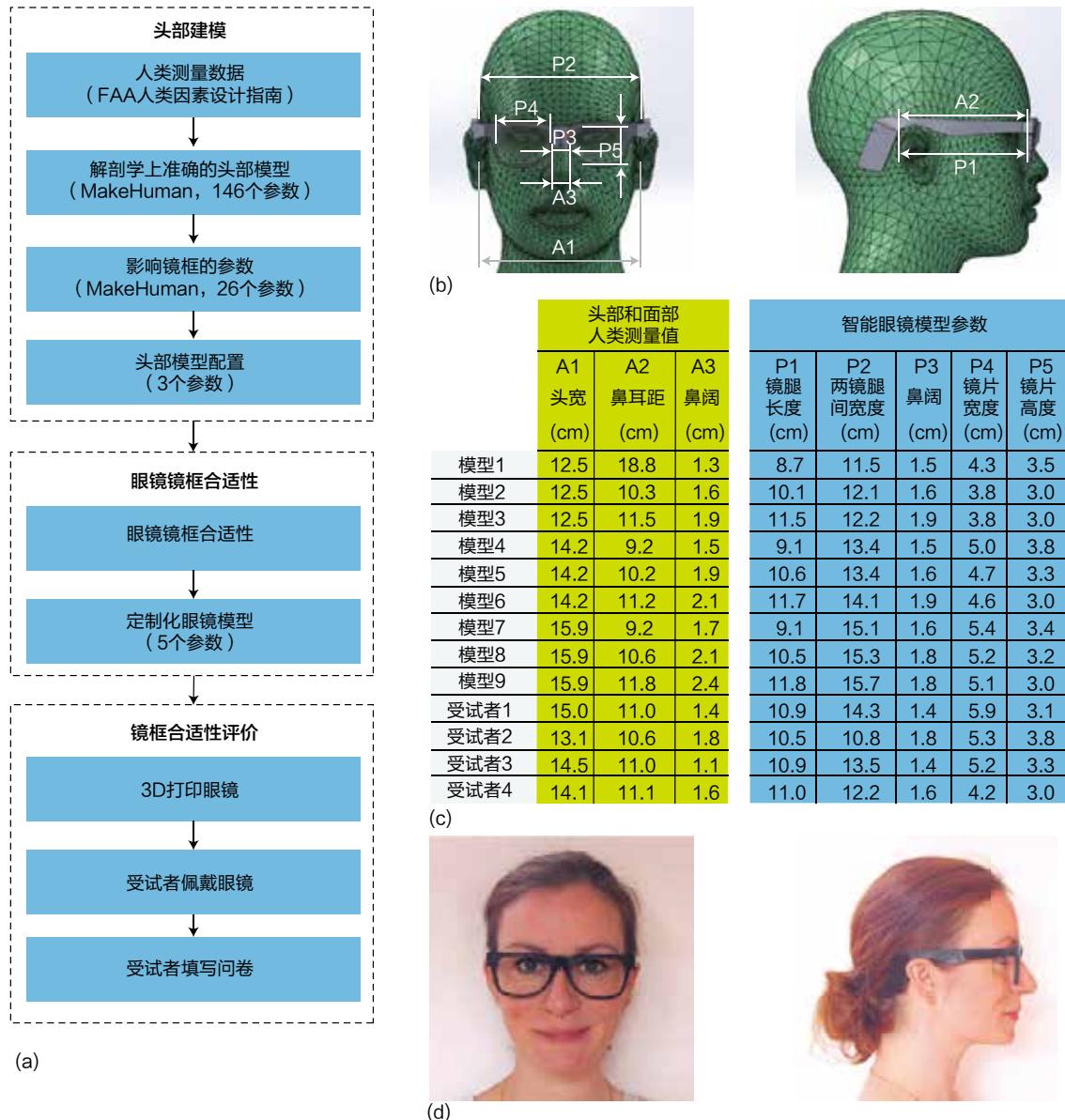


图1. 智能眼镜开发过程步骤、头部尺寸和镜框参数。(a) 使用数字化开发过程来定制智能眼镜。(b) 头部模型规格以及相匹配的镜框，前者数据来自美国联邦航空管理局 (Federal Aviation Administration) 发布的人类因素设计标准指南 (Human Factors Design Standard guide)。(c) 头部模型配制参数、受试者和匹配的 CAD 镜框模型参数。(d) 研究受试者佩戴匹配的 3D 打印眼镜。

## 镜框合适性模拟

为了验证使用这5个参数制作的镜框可以保证不同头型的人们佩戴舒适，我们使用了美国联邦航空管理局(Federal Aviation Administration)发布的人类因素设计标准指南<sup>1</sup>(Human Factors Design Standard guide)中的头型和脸型的人类测量数据。这些数据中的测量值产生了9种使用MakeHuman软件的头部模型配置。我们选择了位于人群中1%，50%和99%的头部长度和宽度。

我们把产生的头部模型导入CAD软件，之后使用一系列头部测量值来匹配镜框和头部。图1b展示了佩戴着合适镜框的头部模型配置例子，以及该头部模型的测量值。

## 硬件

我们的智能眼镜和普通镜框的最大区别在于前者的镜脚里装有电池和主处理器(包括一个微控制器，闪存和一个无线元件)。为了保持重量平衡，这些电子元件分布在两边的镜脚中，元件通过镜框中的线连接。为了使设计不显得怪异，镜框并没有加入触控面板或按键等交互功能。智能眼镜可以连接智能手机来实现交互，这样同时还可减少智能眼镜的处理负荷。

## 受试者评价镜框合适度

除了进行模拟，我们还让4位受试者(两男两女)试戴了使用相同定制程序制作的眼镜。受试者被要求佩戴这种智能眼镜一天，之后评价佩戴舒适度。所有的受试者平时都佩戴框架眼镜，只是在实验过程中换成了隐形眼镜。在得到详细的关于评价过程的信息后，所有受试者都以书面方式同意使用他们的数据和照片。

我们测量了每位受试者的镜腿长度、鼻耳距和鼻阔，之后调整CAD镜框参数以定制每副眼镜。接下来使用3D打印机(丹麦产的Diamond Hotend，所用程序为Witbox)打印出定制眼镜，不过由于本次实验的重点是评价镜框的定制化效果，因此并没有

价了他们平常戴的眼镜。我们发现，整体上来说受试者认为3D打印的镜框戴起来较为舒适(平均得分4.25)不过有点超前(平均得分3.5)。此外，受试者认为3D打印的眼镜和他们平日佩戴的眼镜在舒适度上差不多。详细实验数据请访问如下文档的附录部分([www.actlab.uni-passau.de/~fwahl/IEEE\\_Computer\\_Wahl\\_2017/appendix.pdf](http://www.actlab.uni-passau.de/~fwahl/IEEE_Computer_Wahl_2017/appendix.pdf))。

## 智能眼镜的应用

这里我们描述了智能眼镜的三种重要的用途，当然还有很多其他的潜在功能。

对很多戴眼镜的人来说，智能眼镜能够取代普通眼镜和日常活动追踪器。

有安装镜片，也没有嵌入电子元件。打印的镜框质量为28g，与普通眼镜相差无几。图1d中是一位受试者佩戴3D打印的定制眼镜。

佩戴这些镜框一天后，受试者完成了一份关于合适度和舒适度的问卷调查，其中包含6个评价项目，采用5分制的李克特量表打分。受试者也评

### 用途一： 整体健康和健康追踪

能够追踪步数、心率等日常活动的可穿戴设备已经很普遍了(比如Fitbit)，但是很少有人能长时间坚持使用<sup>2</sup>。而对很多戴眼镜的人来说，智能眼镜能够取代普通眼镜和日常活动

追踪器。

加速度计、陀螺仪等惯性传感器能够识别各种日常生活活动(ADLs)。惯性传感器能够测量人们在走路、跑步、骑车等不同重复情况下的头部运动和姿势，还能测量读书等静态活动。这些传感器并不需要和头部保持一个特定的位置，因此可以嵌入在镜框的任何地方。此外，还能够通过镜腿中脉搏血氧计来监测头部活动。在监测身体和心脏活动时，智能眼镜可以无线连接到智能手机或者网站，从而提供整体行为模式

域特征值(time-domain features)。通过主成分分析(principal component analysis)，我们把时域特征值(如加速度平均值、方差、最小值和最大值)从175个减少至78个，剩余的特征值仍可以解释99.9%的变化。使用高斯混合模型(Gaussian mixture models)对使用留一受试者交叉验证(leave-one-participant-out cross-validation)的活动聚类进行分类。当使用含三个高斯成分的高斯混合模型以及对角协方差矩阵时，除骑车外的活动聚类检测准确

到光线会推迟从昼相至夜相的切换。这种切换提前2小时至推迟3小时都是可能的<sup>5,6</sup>。节律失调和外部时间可能会降低日常行为效率，导致警觉度下降以及肠胃功能受损。人群中很多人都有节律失调或者经常性睡眠不足。此外，LED显示器(电脑和手机的屏幕)发出的蓝光能量相对较高，会抑制促进睡眠的褪黑素分泌<sup>7</sup>，从而影响昼夜时相。在傍晚和晚上过多使用这些电子设备可能会推迟向夜相的切换，从而降低睡眠质量。

### 通过不同的集成传感器，智能眼镜能够连续提供相关的食物摄入细节。

及趋势的每日或每周反馈。

在之前的工作中我们调查了使用智能眼镜从9种行为中识别走路、饮食、阅读等日常生活活动<sup>3</sup>。我们记录了9位受试者的数据，这些受试者有的平时不戴眼镜，其余人在实验中戴隐形眼镜。一位观察者使用智能手机来标记记录结果，而受试者按照一份事先写好的协议来尽量增加日常生活活动。数据总量为66小时，加速度和陀螺仪轴使用30 s滑动窗口(sliding window)和1 s的步长(step size)，从中我们得到了25个时

度均达到80%及以上。使用类别拒绝设计(class-reject design)，整体分类准确度介于70%至84%之间，平均值为77%。图2是此次日常生活活动实验的概述。

### 用途二： 昼夜节律和检测电子设备的使用

人体昼夜节律受进入眼睛的光强度和进入时间的控制。清晨看到光线能够推动从夜相切换至昼相，而在夜晚看

我们智能眼镜的一大特色是有个颜色-光线传感器(ams公司生产)，可以测量全天接收光线的量。该传感器内置在镜框的鼻梁处，离眼睛看物体的角度最近。接收光线的特征及检测电子屏幕的使用有助于预测昼夜时相变化，并对一般行为提供推荐。比如，知道了实时的总光线接收量信息，再加上蓝光检测，就可以推荐蓝光过滤软件，或者在电脑或者智能手机上自动启动这些软件。

为了分析屏幕使用并检测光线，我们进行了一项有14位受试者的实验(其中2位女性12位男性)。每位受试者被要求阅读一篇打印出来的文章，并在网上浏览信息。实验使用惠普的24英寸显示器(型号为EliteDisplay E241i)，受试者距离屏幕70厘米。这两项活动均持续20分钟。受试者佩戴了智能眼

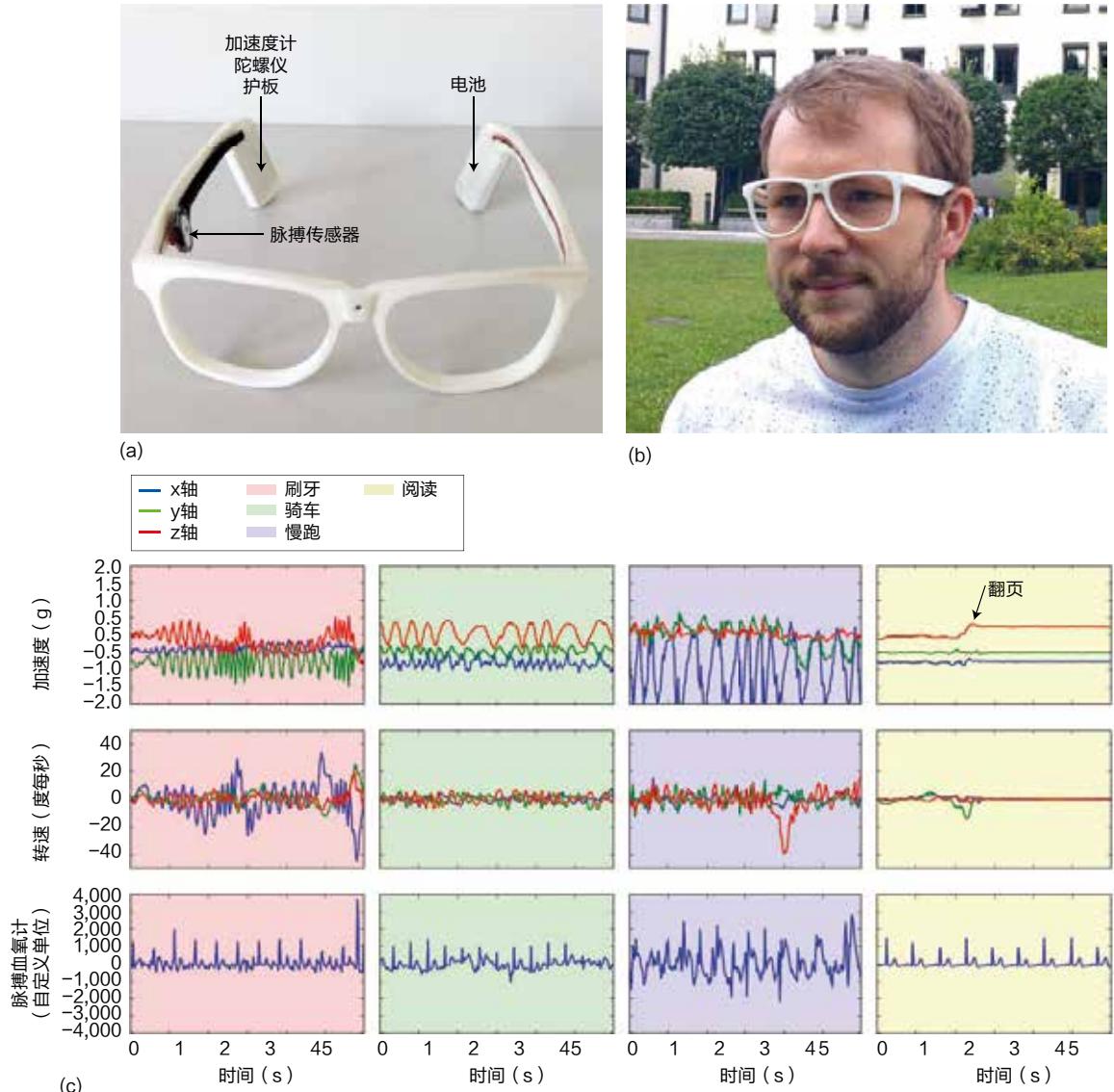
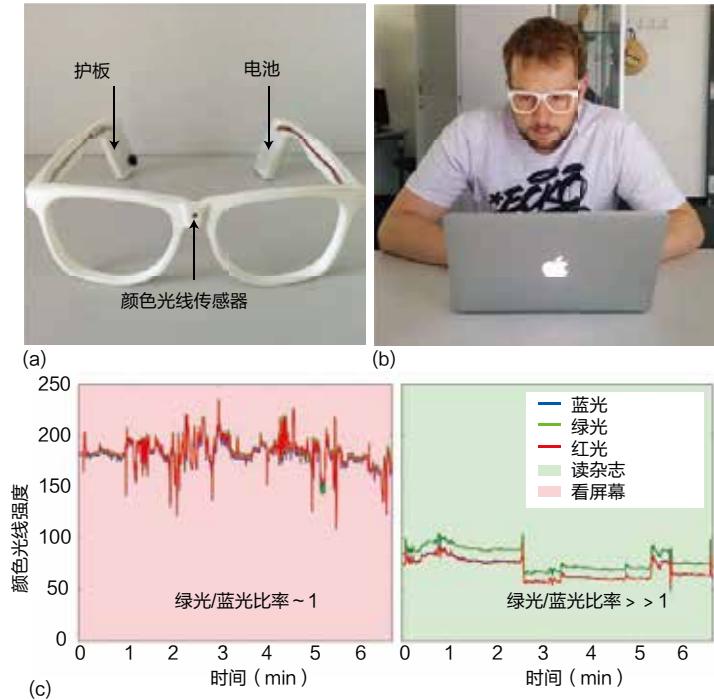


图 2. 活动识别结果。(a) 镜框原型，在镜脚安装惯性传感器的（InvenSense; [www.invensense.com](http://www.invensense.com)），此外还使用脉搏血氧计来检测日常生活活动。(b) 佩戴智能眼镜的受试者在进行日常生活活动。(c) 在从事常见日常生活活动时，来自加速度计、陀螺仪和脉搏血氧计的传感器数据的例子。比如，这些数据表明了读新一页时的头部运动，以及在刷牙、慢跑和骑自行车等活动中重复模式的不同频率。dps: 度每秒; au: 自定义单位。



**图 3. 使用智能眼镜来检测看屏幕的行为。**(a) 用于检测电子屏幕使用的智能眼镜组件。(b) 从事不同活动时佩戴智能眼镜的受试者。(c) 使用电子屏幕和阅读纸质杂志时, 红光、绿光和蓝色通道对应的光强度数据的例子。颜色通道的比率可以清晰区分不同的活动。

镜以记录红色、绿色、蓝色和无色通道, 采样率为6.5Hz(参看图3)。在每个记录单元开始时使用标准光照度仪(Amprobe; [www.amprobe.com](http://www.amprobe.com))记录环境光强度。从15.2小时的数据中, 我们找到了包括光-通道比率(light-channel ratios)在内的18个时域特征值, 该比率中滑动窗口为5s, 窗口重叠率(window overlap)为50%。在线性支持向量机的帮助下, 使用留一受试者交叉验证的方法检测看设备屏幕的时间。通过光强度与颜色通道的比值可以

很清晰地把看屏幕和其他活动分开。检测准确率平均值为90%。当环境光强度不高于500勒克斯时, 即普通夜晚的光强度, 平均准确度最高, 达95%。而当环境光强度高于500勒克斯, 即推荐的办公室室内光强度<sup>8</sup>, 检测准确率仍可达85%。

### 用途三: 监测营养摄入

基于实际摄入模式的饮食指导和营养监测可以帮助心血管疾病、糖尿病或者肥胖患者, 以及那些关心个人健康

的人们。传统的营养监测需要用户手动记录日常饮食。不过, 这种记录无法保证记录细节的准确性, 从而造成该方法准确性很低。除了忘了记录以外, 用户可能还会选择不记高热量夜宵等条目。而智能眼镜可以使用整合在眼镜中的不同传感器来提供相关的食物摄入细节, 从而解决这个不准确的问题。

尽管面部肌电图(facial surface electromyography, 简称EMG)能有效监测咀嚼行为, 但是由于这种方法要在头部放置明显的电极, 因此不适用于可穿戴营养监测设备。不过, 眼镜却是一种把基于EMG的咀嚼监测加入可穿戴配件的方案, 因为镜框会在几个关键区域和皮肤接触<sup>9</sup>。对于监测咀嚼行为来说, 双侧颞肌(bilateral temporalis muscles)非常重要, 因为这些肌肉覆盖很大一部分颅骨, 从鬓发处至耳朵, 因此在每次咀嚼周期中被用于提升颌骨。颞肌上方的大部分皮肤长有头发, 这会给EMG测量带来不便, 不过, 耳朵附近的皮肤就足够得到准确的EMG测定值。

另一个相关信源是咀嚼造成的颅骨振动。嚼碎食物产生的机械振动会从牙齿传遍整个颅骨<sup>10</sup>。这种振动能在乳突骨及耳后等其他颅骨区域检测到。对于营养监测来说, 智能眼镜的定制化是匹配EMG和振动传感器的关键, 这样方能记录关键数据以及维持和皮肤接

触。记录下的数据经智能眼镜处理后，可以及时向佩戴者提供反馈（比如请求佩戴者确认检测到的食物摄入是否与实际一致），或者通过连接的智能手机或网站提供每日或每周的饮食模式反馈。

我们调查了智能镜框附近的不同区域以期找出放置EMG电极的最佳位置。当电极放在紧挨耳朵上方的颞肌区域时，差分EMG测量时的信噪比均很高。我们把织物电极安装在镜脚，参比电极放在鼻垫。EMG数据可以拆分不同的咀嚼周期，而振动测量可以用于从物质质感中区分出不同类别的食物。

图4是整合了各种测试元件的智能眼镜原型，以及在咀嚼和其他活动中测得的EMG和振动数据的例子。我们对8位受试者（男女各半，年龄介于20至56岁）进行了实验，并记录了咀嚼5种代表不同质感食物时的振动，这5种食物是胡萝卜、吐司面包、果冻糖、香蕉和饼干。为了分析监测咀嚼周期的真实情况，受试者还进行了说话、咳漱和头部移动等其他活动。咀嚼在两侧交替进行，且速度不同。每位受试者检测到了38.5分钟的咀嚼周期数据，总共5,435个周期。我们使用平均整流和滤波EMG信号值 $x$ （滑动窗口200 ms）与阈值 $\theta$  ( $\theta = \mu + n \times \sigma$ ) 对比，其中 $n$ 是可调标量，而 $(\mu, \sigma)$  基线噪音的一个高斯模型。若 $x > \theta$ ，那么该窗口会被认为可

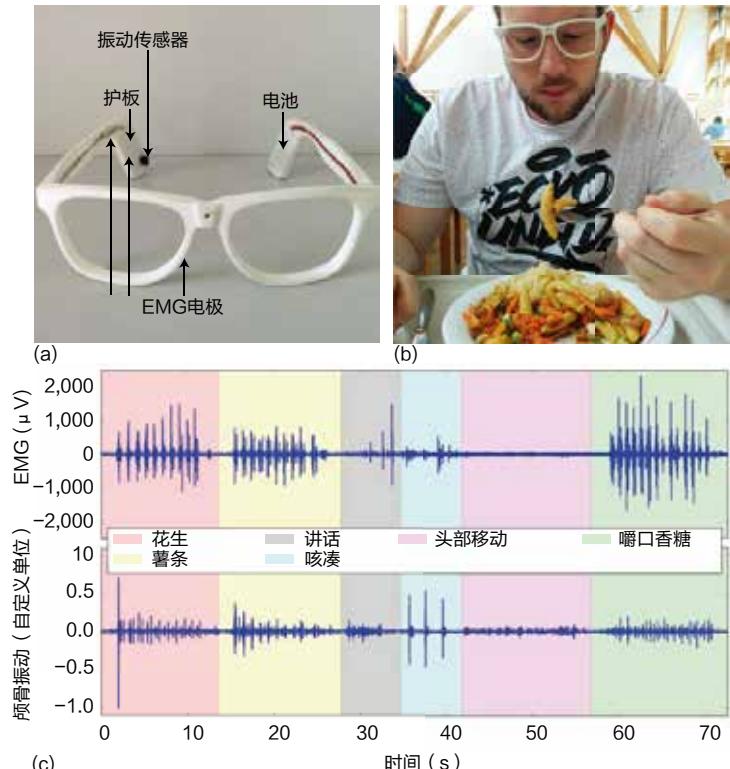


图4. 使用智能眼镜监测营养。(a) 监测咀嚼动作的智能眼镜组件。(b) 日常生活中佩戴智能眼镜的受试者。(c) 在一系列咀嚼和非咀嚼动作中，面部肌电图和颅骨振动数据的例子。不同活动的信号模式不同。

能是咀嚼周期的一部分。通过对窗口进行多数投票，并且保留400 ms至1 s间的咀嚼周期，我们确定了最终检测到的咀嚼周期。尽管监测方法有进一步优化空间，准确率和召回率仍可达80%，这确保了咀嚼时的EMG激活模式可以有效反应咀嚼周期。

尽管我们为上述每种应用场景制作了不同的镜框原型，但是他们可以被整合在一幅镜框中。但是，受到放置传感器的限制，一些应用场景下可能需要

制作专用眼镜。此外，镜框设计会影响功能组分如何嵌入以及镜框模型参数如何选择。我们预计将来为其他应用场景设计镜框时，该数字化开发过程能够重复。

**定**制化可穿戴配饰能够解决各种可穿戴计算中的开放挑战。其一，定制化物品比普通物品使用更频繁，因为前者是根据个人偏好设计并生产的。其二，量

## 关于作者

**弗洛里安·瓦尔**是德国帕绍大学(University of Passau)传感器技术中心(Chair of Sensor Technology)的在读博士研究生。主要研究方向包括泛在传感、模式识别以及感知昼夜时相。瓦尔本科毕业于荷兰埃因霍温理工大学(TU Eindhoven)，获嵌入式系统理学学士学位。联系方式: wahl@fim.uni-passau.de。

**张瑞**是德国帕绍大学传感器技术中心在读博士研究生。主要研究方向包括营养监测、生理信号处理以及可穿戴设备中的模式识别。张瑞本科毕业于德国马德格堡大学(Otto von Guericke University Magdeburg)，获医学系统工程理学学士学位。联系方式: rui.zhang@uni-passau.de。

**马丁·弗洛伊德**是德国帕绍大学传感器技术中心在读博士研究生，同时也搭建了WISEnxt嵌入式研究平台。主要研究方向包括增材制造、嵌入式系统和新型传感器技术的联用。弗洛伊德本科毕业于德国帕绍大学，获计算机科学理学学士学位。弗洛伊德是IEEE和德国工程师协会会员。联系方式: freund@fim.uni-passau.de。

**奥利弗·阿姆夫特**是德国帕绍大学全职教授，同时也是该校传感器技术中心和ACTLab研究组的负责人。主要研究方向是不易觉察的传感器技术、多模型情境模式识别和人类行为推测算法在医疗、体育和建筑技术中的应用。阿姆夫特博士毕业于瑞士苏黎世联邦理工学院(ETH Zurich)，获信息技术和电子工程博士学位。他是Frontiers in ICT以及IEEE Journal of Biomedical and Health Informatics的副主编。联系方式: amft@computer.org。

### 译者简介：

**张哲**，天津大学毕业，曾为《环球科学》等多家科技媒体翻译。

身打造的配饰穿戴起来更舒适。其三，可穿戴设备能够分析传感器数据中每天的趋势以帮助佩戴者了解行为上的改变，而不是仅仅给出可能受到随机噪音影响的单一样本。

合适的定制化加上时尚的设计，智

能眼镜或许能够替代传统眼镜。前20种眼镜设计中<sup>11</sup>，可以加入电子元件和使用我们的定制化方法的有18种。美国戴眼镜的人所占的比例为64%<sup>12</sup>。即便除了经常戴眼镜的人之外，太阳镜、运动眼镜和安全眼镜中也可以加入传感器

来增强人类。

数字打印的镜框在形状和颜色上有很多选择，并且使开发者可以快速应用于小众应用，但是如何安装电子元件以及如何量产打印组分仍是个问题。随着打印技术和材料的演变，附加功能组分或许可以直接打印出来，比如传感器、线路和基本电子元件——但是集成电子元件不大可能被完全取代。由于传感器、处理器等电子元件能嵌入在各类可穿戴配件中，他们仍可以量产，因此成本低。我们相信电子元件和打印组分混用会产生全新一代的可穿戴系统。■

### 致谢

本研究受荷兰技术基金会STW资助(项目批准号: 12184)。

### 参考文献

1. Human Factors Design Standard (HF-STD-001), Federal Aviation Administration, 2003; hf.tc.faa.gov/hfds.
2. J. Clawson et al., "No Longer Wearing: Investigating the Abandonment of Personal Health-Tracking Technologies on Craigslist," Proc. 2015 ACM Int'l Joint Conf. Pervasive and Ubiquitous Computing (UbiComp

- 15), 2015, pp. 647–658.
3. F. Wahl, M. Freund, and O. Amft, “WISEglass: Smart Eyeglasses Recognising Context,” *EAI Endorsed Trans. Pervasive Health and Technology*, vol. 16, no. 5, 2015; doi: 10.4108/eai.28-9-2015.2261470.
  4. G. Spina et al., “CRNTC+: A Smartphone-based Sensor Processing Framework for Prototyping Personal Healthcare Applications,” Proc. 7th Int'l Conf. Pervasive Computing Technologies for Healthcare (PervasiveHealth 13), 2013, pp. 252–255.
  5. V.L. Revell and C.I. Eastman, “How to Trick Mother Nature into Letting You Fly Around or Stay up All Night,” *J. Biological Rhythms*, vol. 20, no. 4, 2005, pp. 353–365.
  6. J.F. Duffy and K.P. Wright, “Entrainment of the Human Circadian System by Light,” *J. Biological Rhythms*, vol. 20, no. 4, 2005, pp. 326–338.
  7. G.C. Brainard et al., “Action Spectrum for Melatonin Regulation in Humans: Evidence for a Novel Circadian Photoreceptor,” *J. Neuroscience*, vol. 21, no. 16, 2001,
  8. “DGUV Information 215-410: Bildschirm- und Büroarbeitsplätze Leitfaden für die Gestaltung [Screen and Office Workstations Guide for Design],” Deutsche Gesetzliche Unfallversicherung [German Social Accident Insurance] (DGUV), Sept. 2015; publikationen.dguv.de/dguv/pdf/10002/215-410.pdf (in German).
  9. R. Zhang, S. Bernhart, and O. Amft, “Diet Eyeglasses: Recognising Food Chewing using EMG and Smart Eyeglasses,” Proc. Int'l Conf. Wearable and Implantable Body Sensor Networks (BSN 16), 2016; doi: 10.1109/BSN.2016.7516224.
  10. O. Amft et al., “Analysis of Chewing Sounds for Dietary Monitoring,” Proc. 7th Int'l Conf. Ubiquitous Computing (UbiComp 05), LNCS 3660, M. Beigl et al., eds., Springer, 2005, pp. 56–72.
  11. “Eyeglasses Trends 2016: What to Wear?,” *TheFashionTag.com*, 15 Mar. 2016; [thefashiontag.com/2016/03/15/eyeglasses-trends-2016-what-to-wear/](http://thefashiontag.com/2016/03/15/eyeglasses-trends-2016-what-to-wear/).
  12. The Vision Council of America, *OTC Readers Report Q4 2015, tech. report, 2016*; [www.thevisioncouncil.org/sites/default/files/research/Readers%20Index%20Q4.pdf](http://www.thevisioncouncil.org/sites/default/files/research/Readers%20Index%20Q4.pdf).





# 基于人体传导声音信号的新型通信方式

文 | 张铖，希南·海尔赛克（Sinan Hersek），蒲奕名，佐治亚理工学院（Georgia Institute of Technology）

孙丹睿，北京邮电大学

薛秋月，北京大学

萨德·E·斯塔那（Thad E. Starner），格雷戈里·D·艾博德（Gregory D. Abowd），奥默·T·伊南（Omer T. Inan），佐治亚理工学院（Georgia Institute of Technology）

译 | 薛秋月

通过调制声音信号可以传播信息，以人体为声音信号的传播媒介可以更方便更自然的在各个不同设备或物体传导信息。这个新提出的通信系统可以在人体内、人体之间或是人体和环境之间传播频率低于20kHz的声音信号。

**在**下一个计算时代中，我们可以预测一大批新的可穿戴设备将会从各方面影响我们的生活。这些可穿戴设备来自于交互系统（比如智能手表和头戴式显示设备如谷歌眼镜）中不同类型的因素，它们可以作为传感器用于监测人体活动和内在身体状况（比如心率等）。所以可以预测在未来几年内有极大的可能很多人会穿戴除了智能手表以外的各式各样的可穿戴设备。

在各种可穿戴设备被大家接受的潮流下，在这些设备之间自然且方便的通信显得尤为重要。每个设备都需要和其他设备以及终端的服务器有频繁的信息交换，大多数移动和可穿戴设备如今采用无线通信的方式例如Wi-Fi和蓝牙来发送和接收数据，但是这种无线通信技术并不是可穿戴设备通信

的最佳选择，有如下两点主要原因：

首先，无线通信技术没有完全解决可穿戴技术引进的新问题和挑战。比如，因为这些可穿戴设备是穿在身上的，它们一直在随用户移动而改变位置。因此，可穿戴设备需要随着位置的变化与环境中的不同设备通信，目前的解决办法要求在两个设备互相通信之前需要较为单调冗长设置过程，而且这些设置过程常常需要用户的输入。如果我们能让可穿戴设备直接自然的识别并联系到环境中用户想要通信的设备，这将极大的提升用户体验。

第二，现在的通信方法没有利用可穿戴设备的优势。与传统的计算设备（例如笔记本电脑）相比较，可穿戴设备是直接与人身体皮肤相连的，这为我们提供了一个从人身体获取信

号的独特机会。

我们提出利用人身体本身作为声音信号传播的物理途径。利用与人体直接相连的可穿戴设备，系统可以自动完成信息认证的过程。另外，我们介绍了这种通信方式与生俱来的在不同设备直接传递信息的安全保护：如果两个设备是连在同一个人身上的，那么这两个设备可以通信，否则这种通信不会发生。

利用人体传输的通信方式可以促进新型、自然的交互方式技术发展。例如，握手是一种世界上流行的礼仪，经常伴随着口头的相互介绍。利用我们提出的这种通信方式，很多信息例如名片等可以在握手时将名片内的信息自动交互通信。这个方法也可以用于人体和环境之间的信息交流。简单的摸一下环境的物体，用户就可以将自身的信息传递到物体上作为认证，然后环境中的物体可以在用户认证通过后将用户需要的相应信息发送给用户。这种利用人体传导信号的通信方式还有另外一些好处包括节省能量和节省成本。

利用人体来传递信息的想法目前已经被论证过用电信号<sup>1,2</sup>和磁信号<sup>3</sup>与人体结合来传播。但是，目前还没有人做过利用频率20kHz以下的声音信号在人体媒介中传递信息的研究。与利用电磁波信号的通信技术相比，发送低频率的声音信号应该是更可行的。我们设计

并实现了一个系统来证明传播20kHz以下声音信号以达到在人体内、人和人之间还有人与环境之间通信的可行性。另外通过八个参与者的实验研究了声音在

在人体各部分中传播一个相对长的距离。表征声音信号在人体中的传播可以为个人域网络和自然的人机交互带来新的契机。我们的工作主要在频率低于

## 频率低于20kHz的体内声学信号 可被很多已被商业化的传感器和 制动器支持。

不同人体之间怎样传播。实验表明我们提出的方法可以以百分之百的准确度检测接触与否。我们还利用自己设计的硬件和商用的智能手表实现了一个系统来通过人体发送文字信息。

20kHz的声音信号，这种信号是可以或在未来可以被商用的可穿戴设备中的传感器和制动器支持的。

## 在人体中发送声音 信号

声音信号在空气、固体和液体中的传播是不同的——也就是说，声音信号在不同介质中的传播速度不同。人体是由各种异质的混合物组成，包括骨头（固体），水（液体）还有肌肉（混合物）等。

研究表明骨头是非常好的传导声音信号的介质。然而，大多数利用骨头传导声音信号的方法都会受到传播距离极短的限制（几厘米），例如骨传导耳机。目前仍不清楚声音信号怎样才能

我们证明了我们提出的方法，设计实验用索尼智能手表<sup>3</sup>来通过握手在两个人体直接传递信息，和从人体到一个桌子之间传递信息。为了在人体中传播声音信号，我们的系统包含了两个子系统：一个是生成声音信号并将声音信号传递给人体的，另一个是采集人体中传递的声音信号的。最挑战的部分是找到能和人体相连并且对环境噪音（例如电磁噪音EM noise）有一定抵抗的信号传感器。例如，我们用薄片压电传感器做接收端然后发现它对环境中的电磁噪音太过敏感，因此让它来接收声音信号就很困难。

我们决定用一个骨传感器(B81, RadioEar)作为发送端，用一个极低噪音的加速传感器(356A32, PCB

## 以人体为通信传导媒介的相关工作

自从1990年可穿戴式计算设备成为一个研究热点以来，利用无线通信传输数据容易使信息被窃听便成为了人们的一个担忧。<sup>1</sup>研究者们因此开始研究怎样在可穿戴设备直接通过在靠近人体表面传播电流来传递信息。<sup>2-4</sup>基于这些前者的工作，更多人尝试研究了在人体内传播电信号的特性<sup>5-8</sup>以及这种方法的潜在应用。<sup>9</sup>另外，最近研究者们证明了用磁信号的回声来实现人体个域网络的数据交换。<sup>10</sup>

这些前人的工作都将重点放在人体内传播电信号或磁信号（兆赫兹范围频率），但人体也是一个传播声音信号的好介质。也有研究者们研究过声音信号如何在死尸内传导，<sup>11</sup>以及用声音信号在两个智能手机之间传递信息<sup>12-13</sup>或识别用户手势等。<sup>14-16</sup>莱斯大学(Rice University)的研究者们提出的OsteoConduct系统用声音信号在同一个人身体内传递信息，但是只达到了一个较低的比特率(5bit/s)。<sup>17</sup>卡耐基梅隆大学的研究者们提出的Viband系统实现了在一个腕表和一个振动马达之间传输数据，达到了165bits/s的传输速率。然而，目前仍不清楚声音信号能如何以较高的数据传输速率传播一个较长的距离或从身体中传播到一个其他物体上。

与现在已有的相关工作不同，我们的研究证明了用频率低于20kHz的声音信号在身体中传播的可能性，包括从手腕传播到身体上其他部位、在两个身体之间传播和从身体上通过接触传播到环境里其他物体的情景。

### 参考文献

1. G. Revadigar et al., "Secure Key Generation and Distribution Protocol for Wearable Devices," Proc. IEEE Int'l Conf. Pervasive Computing and Communication Workshops (PerCom Workshops 16), 2016; doi:10.1109/PERCOMW.2016.7457058.
2. T. G. Zimmerman, "Personal Area Networks: Near-Field Intrabody Communication," IBM Systems J., vol. 35, nos. 3-4, 1996, pp. 609-617.
3. E.R. Post et al., "Intrabody Buses for Data and Power," Proc. 1st Int'l Symp. Wearable Computers (ISWC 97), 1997; doi:10.1109/ISWC.1997.629919.
4. M. Fukumoto and Y. Tonomura, "Body Coupled FingerRing: Wireless Wearable Keyboard," Proc. ACM SIGCHI Conf. Human Factors in Computing Systems (CHI 97), 1997, pp. 147-154.
5. J. Bae et al., "The Signal Transmission Mechanism on the Surface of Human Body for Body Channel Communication," IEEE Trans. on Microwave Theory and Techniques, vol. 60, no. 3, 2012, pp. 582-593.
6. N. Cho et al., "The Human Body Characteristics as a Signal Transmission Medium for Intrabody Communication," IEEE Trans. Microwave Theory and Techniques, vol. 55, no. 5, 2007,

Piezotronics)作为接收端。骨传感器提供了与人体之间合适的连接耦合，尤其是当它放在骨头上面的时候。与其他类型的声音传感器(比如薄片压电传感器)相比，加速传感器是对电磁屏蔽的，它的机械设计使它可以准确的检测到小的振动，因此当它贴在人体上时对其他的环境噪音有良好的抵抗性能。更重要的是，与基于电信号的微机电系统麦克风相比，加速传感器是一

种能够只挑选出在传输媒介上直接方向的振动而不是空气中的振动信号(也即环境噪音)。除去空气中的噪音信号可以帮助我们更好的研究人体中声音信号的传播。

### 实验设置

为了理解不同人之间声音信号的传播，我们收集了在实验室环境里八个

参与者的数据，这八个参与者的平均年龄是27，身体质量指数(BMI)从17到24.2不等，平均值为21.35。所有的被试都提供了手写的告知同意书，我们的实验由佐治亚理工学院审查委员会批准。研究者用臂环将传感器紧紧贴在被试的身上。实验是为了研究描述声音信号如何通过可穿戴设备中集成好的制动器和传感器在人体之内、两个人体之间已寄人体和环境之间传播的。

- pp. 1080–1085.
7. K. Fujii, M. Takahashi, and K. Ito, “Electric Field Distributions of Wearable Devices Using the Human Body as a Transmission Channel,” *IEEE Trans. Antennas and Propagation*, vol. 55, no. 7, 2007, pp. 2080–2087.
  8. K. Hachisuka and A. Nakata, “Development and Performance Analysis of an Intra-body Communication Device,” *Proc. 12th Int'l Conf. Solid-State Sensors, Actuators, and Microsystems (Transducers 03)*, 2003, pp. 1722–1725.
  9. S. Franklin and S. Rajan, “Personal Area Network for Biomedical Monitoring Systems Using Human Body as a Transmission Medium,” *Int'l J. Bio-Science & Bio-Technology*, vol. 2, no. 2, pp. 23–28.
  10. J. Park and P.P. Mercier, “Magnetic Human Body Communication,” *Proc. Ann. Int'l Conf. IEEE Eng. Medicine and Biology Soc. (EMBC 15)*, 2015, pp. 1841–1844.
  11. S. Stenfelt and R.L. Goode, “Transmission Properties of Bone Conducted Sound: Measurements in Cadaver Heads,” *J. Acoustical Soc. of America*, vol. 118, no. 4, 2005, pp. 2373–2391.
  12. I. Hwang, J. Cho, and S. Oh, “Privacy-Aware Communication for Smartphones Using Vibration,” *Proc. IEEE Int'l Conf. Embedded and Real-Time Computing Systems and Applications (ERTS 15)*, 2015, pp. 447–452.
  13. N. Roy, M. Gowda, and R.R. Choudhury, “Ripple: Communicating through Physical Vibration,” *Proc. 12th USENIX Symp. Networked Systems Design and Implementation (NSDI 15)*, 2015, pp. 265–278.
  14. T. Deyle et al., “Hambone: A Bio-Acoustic Gesture Interface,” *Proc. 11th IEEE Int'l Symp. Wearable Computers (ISWC 07)*, 2007; doi:10.1109/ISWC.2007.4373768.
  15. C. Harrison, D. Tan, and D. Morris, “Skinput: Appropriating the Body as an Input Surface,” *Proc. ACM SIGCHI Conf. Human Factors in Computing Systems (CHI 10)*, 2010, pp. 453–462.
  16. M. Fukumoto and Y. Tonomura, “Whisper: A Wristwatch Style Wearable Handset,” *Proc. ACM SIGCHI Conf. Human Factors in Computing Systems (CHI 99)*, 1999, pp. 112–119.
  17. L. Zhong et al., “Osteoconduct: Wireless Body-Area Communication Based on Bone Conduction,” *Proc. ICST 2nd Int'l Conf. Body Area Networks (BodyNets 07)*, 2007, article 9.
  18. G. Laput, R. Xiao, and C. Harrison, “Viband: High-Fidelity Bio-Acoustic Sensing Using Commodity Smartwatch Accelerometers,” *Proc. 29th ACM Ann. Symp. User Interface Software and Technology (UIST 16)*, 2016, pp. 321–333.

图1a显示了实验的器具，实验采用了动态信号分析仪(SR785)，每一个实验中这个动态信号分析仪扫描一个频率在250Hz到20kHz的正弦波信号(对数坐标上平均分割)。这个信号传给骨传导器然后由骨传导器发送给人体。在加速传感器上检测到发送的信号再连接到一个信号放大器上，再连接到动态信号分析仪的输入端口上。信号分析仪会计算出发送信号与接收信号之间的

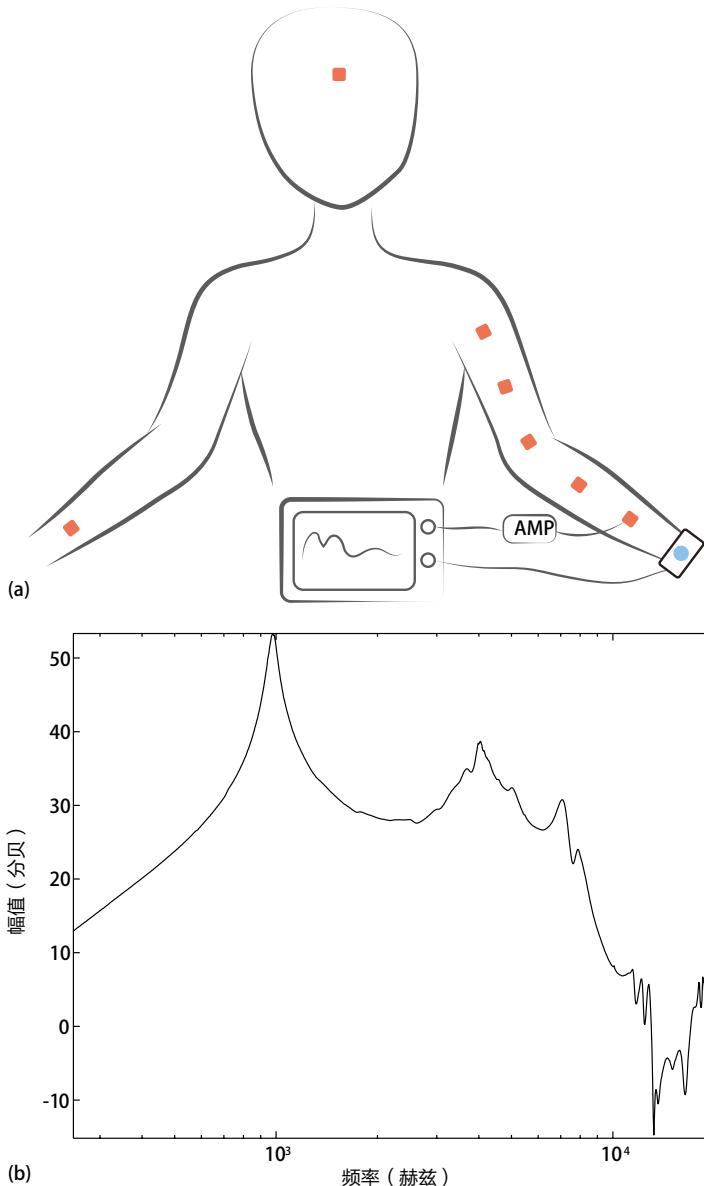
传输函数，我们记录下这个传输函数以备以后的分析之用。

为了测量实验中使用的加速传感器和骨传感器之间的频率响应，我们将骨传感器和加速传感器直接绑在一起(不经过人的身体)，然后扫描一个幅度为30mV频域上从250Hz到20kHz的正弦波，实验结果如图1b所示。可以看出这个系统在信号频率为1kHz左右响应的幅度最大。

输入到骨传感器的信号电压设置为1V。除了当加速传感器放在前额和右手腕处时，我们将输出信号电压设置为5V，因为在这两个地方接收信号会经过更大的衰减。

## 人体内传导通信

如果能实现同一人体内基于声学信号的通信，那么很多可穿戴设备，例如



**图1.** 实验设备：(a) 这张图显示了传感器的放置位置（图中红色的点），骨传感器放在左手手腕上（图中蓝色圆点处）与一个信号发生器相连，接收信号发生器产生的信号来发送声音信号。(b) 这个工作我们将骨传感器（发送端）放在空气中，然后将骨传感器与加速度计紧密的联合在一起。我们将骨传感器和加速度计紧密的直接粘在一起，然后骨传感器扫描发送频率从250Hz到20kHz的幅度为30mV的正弦波，图(b)显示了接收端在不同频率上的接收信号幅度响应值。

戴在手腕或手臂上的智能手表和臂上手环等功能可以得到进一步加强。与此对应，不同频率的发送声音信号在人体手臂上传播会经历怎样不同的损耗对可穿戴设备有很大研究意义。为了研究这个问题，我们在左手臂上将接收端（加

速传感器）放在几个离发送端不同距离的位置（发送端也在左手臂上），距离分别为7.6, 15.2, 22.9, 30.5和38.1cm。图2a显示了加速传感器在不同位置上检测到的信号响应幅度（为八个参与者的平均值）。线上的每个点表示了能达

到最大频率响应的对应频率值。在图上我们还同时用红色虚线标出了骨传感器工作在1V时的噪声值，这个噪声值是我们在不连接骨传感器的时候将加速传感器的输出与信号分析仪的原始信号比较，由信号分析仪计算而成。

图2a显示了所有频率信号接收到的幅度值都远大于噪声值，也就是说，实验中加速传感器接收到的信号是真正从骨传感器中发出，经过人体传导后接收到的信号。

另外我们也能看出不同频率的信号幅度变化很大，这里主要有两个潜在的原因影响了不同频率信号的接收幅度值。首先是因为传感器的硬件原因：理想情况下，当把发送端骨传感器和接收端加速传感器直接连在一起时，得到的频率响应应该是一条平的直线，然而，实际中很少能有这种情况，正如图1b显示，我们实验结果得到的频率响应图在频率为1和7kHz时达到峰值，图2a中也有在相同频率点观察到的峰值，我们将每个峰值都标为点，在传播距离分别为7.6, 15.2, 22.9, 30.5和38.1cm时，峰值点对应频率分别为1004, 987, 953, 1049个1054Hz。

第二个原因是因为在人体内各个不同频率的声音信号传播过程中经历的衰减不一样。因为这个原因，图2a与图1b有着明显的不同。比如说，图2a中250Hz到1kHz频段的信号比图1b中相

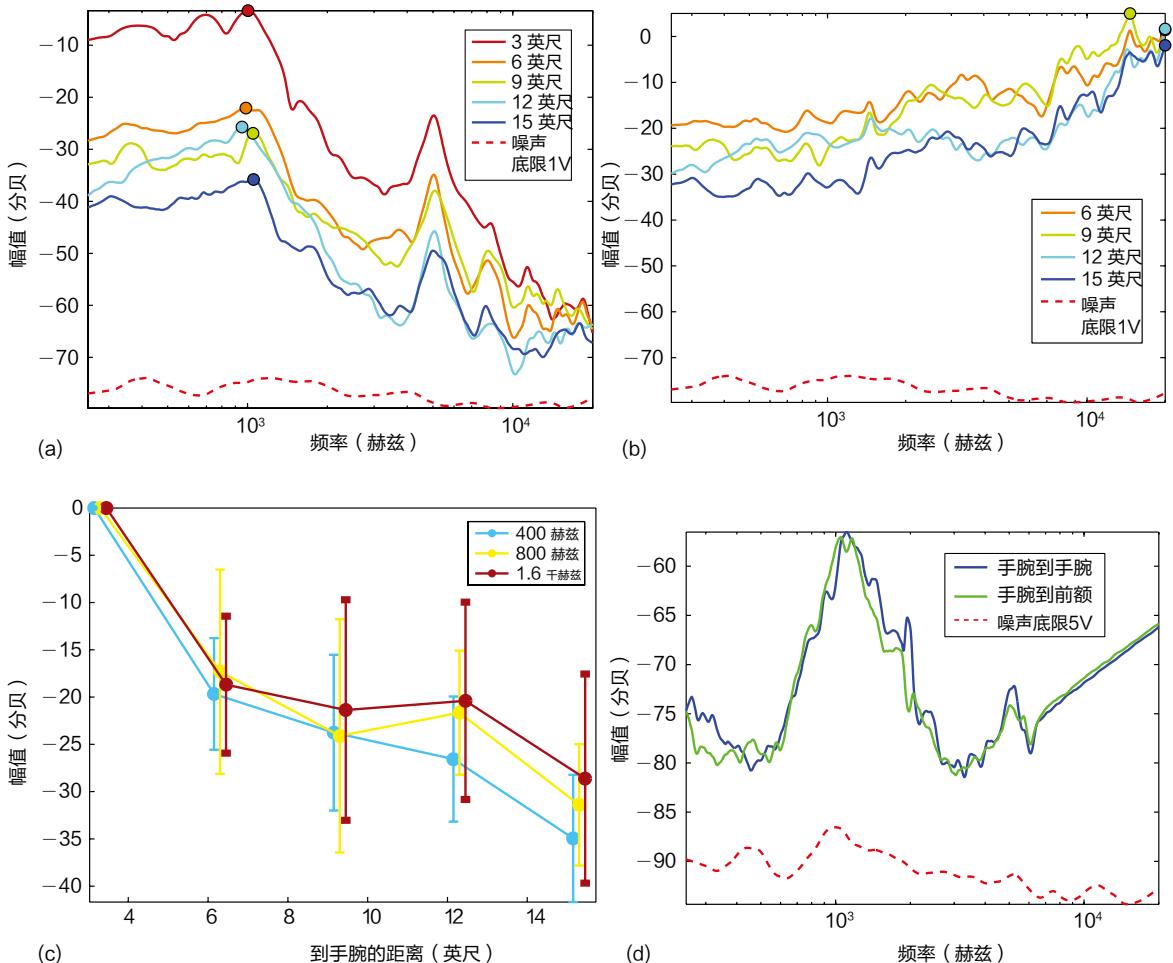


图 2. 在同一个人身体上的频率响应结果: (a) 接收端在同一个人手臂上不同位置时的频率响应结果, (b) 在手臂上的传播的路径损耗, (c) 不同频率的信号衰减, (d) 从左手手腕到右手手腕的频率响应结果以及从左手手腕到前额的频率响应结果。图中每个线上的点表示了不同频率上能达到的幅度的最大值。

应频段的结果要平坦很多。这个结果说明频率低于1kHz的声音信号在传播过程中衰减的更少。

为了更深入的研究在手臂上不同频率声音信号传播的衰减,我们将接收端距离发送端分别为7.6, 15.2, 22.9, 30.5, 38.1cm时的结果与距离为7.6cm时的结果相减,得到图2b。图2b显示出了信号在相对于7.6cm传播更远距离时的传播路径损耗,可以看出,在信号频率低于10kHz时频率响应结果相对平坦。

理论上,传播距离越长,接收端应

该观察到更多的信号衰减,图2c画出了以传播距离为横坐标,八位参与者在不同信号频率400Hz, 1kHz和1.6kHz情况下下的信号衰减。在图中,我们将传播距离7.6cm处时从手腕处接收到的信号作为参考信号(设为0dB),计算出其他传播距离的接收信号衰减值,以分贝为单位。大体上,传播距离越远,衰减越大。然而,当发送端和接收端之间传播距离从30.5cm增加到38.1cm时,接收信号的平均幅值不降反增。有趣的是,30.5cm处恰巧大概是手肘的位置。

我们将这个接收信号强度的反常

变化归结于传感器放在手肘附近,手肘附近传感器更多的与骨头相连,而不是像在其他地方例如小臂和大臂上时更多的与肌肉组织相连。正如我们所料,这个结果说明了身体不同部位传播不同频率声音信号的阻抗是不同的。当然,要得到更多的结论还需要更多的实验。

除了手臂,可穿戴设备还能戴在身体的别的地方,这些地方可能有更远的传输距离和更负责的传输路径。比如说,谷歌眼镜戴在头上,到手腕的距离可能大于50cm。为了理解信号是如何传输到更远的地方的,我们重复了先前

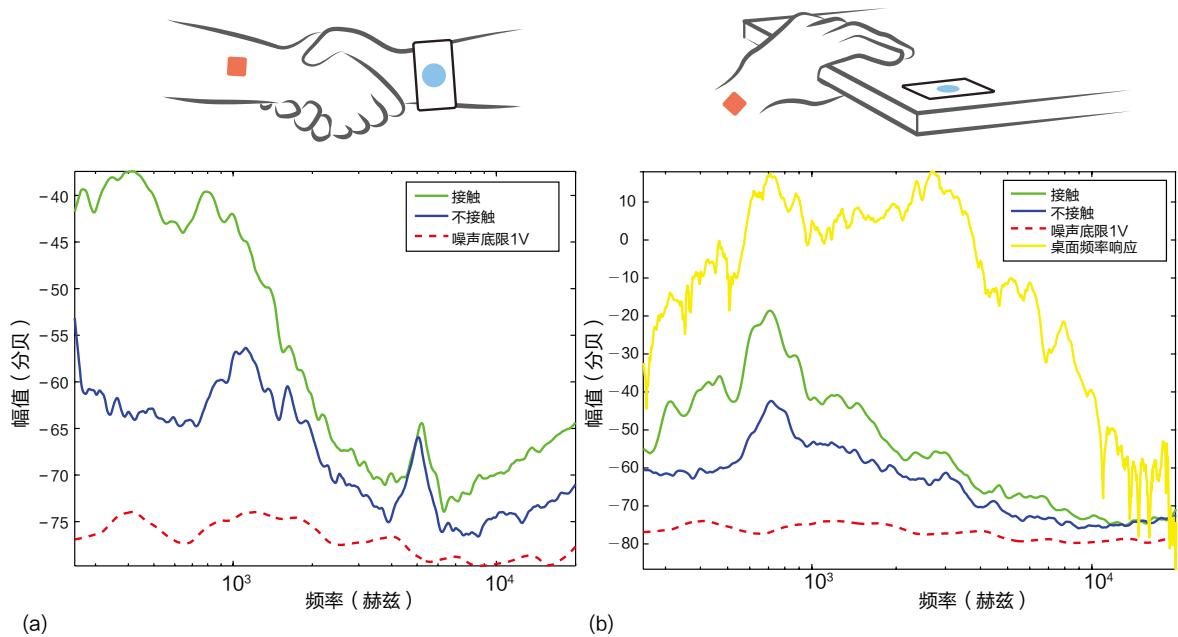


图3. 频率响应曲线: (a) 握手情景, (b) 接触桌子的情景。

的实验，但是将加速传感器（接收端）放在了前额中心和右手手腕上，发送端仍然像以前一样放在左手手腕上。图2d显示了平均的频率响应结果。仍然，接收到的信号远大于噪音值，这说明接收端接收到的结果是从发送端发送的信号，而不是噪音。当接收端分别放在前额和右手腕上时，频率响应结果分半在1045和1106Hz处达到峰值。

## 人体间传导通信

我们还研究了两个人在握手时声音信号会如何在两个人身体之间传播。我们将骨传感器系在一个参与者的左手手腕上，将加速传感器放在另一个参与者的左手手腕上。与之前的实验相似，我们记录了频率从250Hz到20kHz之间的频率响应。这个实验重复了两次。第一次，我们让两位参与者握手时

紧紧握住，我们让参与者采用同样的姿势但是手之间并不接触，大概保持1cm左右的距离。因为两次实验直接唯一的差别在于参与者手之间是否接触，我们可以通过比较两次实验频率响应结果估测出有多少声音能量通过两个手的接触传播。

图3a显示了频率响应曲线，当两手相握，频率响应峰值出现在频率为410Hz处，(为-37.4dB)，我们还能观测到在800Hz左右有另一个峰值，但这个频率峰值与系统本身频率响应峰值相近。因为系统本身没有在410Hz处产生最大的能量，我们得到的结果在410Hz处出现峰值说明410Hz的信号传输损耗更小。相反，系统本身在1kHz处有最大的能量值，但我们得到的结果在1kHz处并没有最大的能量。对于信号频率峰值的变化现象，一个可能的解释是因为低频的信号有更长的波长，使其在

传播经过两个人握手的时候能更好的穿透障碍物和经受更少的能量损耗。

## 人体与物体之间传导通信

除了人体内传导和人体与人体间传导，我们还对人体和环境中物体之间声音信号如何传导感兴趣。我们设计了一个实验来模仿一个用户可以通过接触环境中的一个物体来接收信息的场景，我们实验中这个物体是一个木制桌子。我们将骨传感器贴在桌子上，将加速传感器系在用户的左手手腕上。要求每个用户抓住桌子的边缘，桌子边缘到骨传感器距离为50.8cm。与握手实验相似，我们将这个实验重复两次，两次实验唯一的不同在于手与桌子是否接触。图3b显示了平均的结果。仍然，我们得到的结果是远大于噪音值的。然而，当手接

触桌子时，频率响应只在700Hz处有一个峰值，而不是在系统产生能量最大的1kHz。

我们将得到的频率响应峰值与系统自身频率响应的峰值归结于两个原因：桌子的频率响应和手与桌子接触处的传输损耗。为了测量桌子的频率响应，我们将加速传感器放在用户手接触桌子的位置，然后记录下这种情况的频率响应结果，我们观察到在700Hz和2.7kHz处有峰值。第一个频率峰值700Hz与用户手接触桌子时得到的频率响应峰值相同，但是当频率高于700Hz时信号幅度有明显下降。这个结果说明信号频率越高在手与桌子接触的传输路径上损耗越大。另外，这个结果与两人握手实验中结果一致，这说明低频声音信号在人体中传播的路径损耗可能更小。

## 空气传播的影响

因为骨传导器并不是与空气完全隔离的，这个系统也有可能接收到发送的经过空气传导到接收端的信号。为了检测究竟有多少信号能量从发生源传播到接收端，我们比较了手与桌面接触实验中的两个频率响应结果。

我们将接触和不接触两种情况的平均结果做比较，发现接触情况下接收到的信号强度至少比不接触的情况高

20dB。然而，图3b显示加速传感器仍然可以检测到强度最大达到-40dB的信号。我们认为是因为骨传感器功率非常大，于是将整个桌子变成了一个巨大的声音信号发生器，加大了声音经过空气传播的部分。因此，为了使以后能更好的利用这项技术，减少声音的空气传播是非常必要的。

## 检测是否接触

用基于身体的传输网络相比现在的无线网络，安全性是一个潜在的优

有发生，信号的振幅就会大幅下降。因此，我们可以通过设置接收信号强度的阈值来区分是否有接触。然而，这个接收的信号强度人与人之间是不一样的，这就要求这个阈值每次在使用前要重新调整。如果这一步没有做好的话会有信息泄露的危险。

为了解决这个问题，我们尝试了通过频率响应的曲线结果匹配来检测是否有接触，而不是简单的比较信号强度。图3显示了接触和不接触的频率响应曲线的区别，可以看出它们有明显区别，尤其是在1kHz处。基于这个观察结

为了使以后能更好的利用这项技术，减少声音的空气传播是非常必要的。

势。因为它可以只传输戴在用户身上的设备的信息。但如果发送端能从空气中传输信号，那么别人仍然可以在不接触的情况下偷听信息。特别地，如果捕捉信号的设备和信号源足够近，就可以从空气传播中获得足够强的信号能量来解码传播的信息。为了降低这一风险，在信息通信之前加是否接触的判断是必要的，即只有接触才能启动信息传递过程。

根据之前的实验结果，如果接触没

果，我们做了另外一个实验。

首先，我们将频率响应归一化以避免不同人导致的不同接收幅度影响。我们通过比较现在1kHz以下频率响应的结果和之前1kHz以下频率响应结果得到变化值曲线。我们计算了每个参与者的变化值曲线和八个参与者在接触桌面和握手两种情景下的平均结果。然后，为了检测是否有接触，我们计算了每个人的待确定测试结果和接触或不接触情况的平均结果之间的相关系数，待

确定结果与哪一类情况的相关度更高就将其划入哪一类。我们将这种方法用于接触桌面和握手两种情景的结果进行接触检测，通过八位参与者的32个频率响应曲线数据，得到了100%的准确率。

这个实验显示检测是否有接触是可以实现的，但是如果想将它运用到现实生活中则需要更多的工作。例如，我们利用了发送端到传输端之间信号的传输函数，这要求我们既知道接收端的信息，还要知道发送端的信息，但在实际应用中，发送端一般不会有接收端的接收信号，出发接收端再将接收信息送回发送端。这样的一个系统要求每个设备既包含一个发送端又包含

编码声音信号来传播文本信息，包括在身体内传播，身体间传播，还有在身体和木制桌子之间传播。在这个实验中，我们相较于之前的实验对硬件做了一点改变，接收端我们用了一个接触式麦克风 (Knowless BU-23173-000) 来记录数据，将这个麦克风连到一个我们自己设计的放大器上(见图4) 经过低噪放大和信号滤波。在低噪放大器中，接收到的信号首先经过一个截止频率为8MHz的二级低噪滤波器来去除电磁波信号影响，级联一个截止频率为0.16Hz的高通滤波器去除直流偏置；经过这一步后，信号经过一个20dB的放大，最后一级是一个截止频率为20kHz的Sallen-Key 巴特沃斯低通滤波器。最后，这个

编码过的声音信号。在实验中，输入的文本信息首先被翻译为ASCII码格式，再用FSK将ASCII码编码到声音信号中。FSK编码方法中的低频信号和高频信号分别设置为1575Hz和3150Hz。

与之前实验相似，我们将骨传感器放在手腕上，将麦克风放在距离骨传感器7.62, 22.86, 38.10cm处的手臂上。在每个位置上，我们用不同的波特率发送五个不同的字母，将接收数据记录下来然后通过一个带通滤波器，最后进行频移键控(FSK)的解码。

在上文描述的情况下，我们从接收信号中成功的解出了所有的字母。这些结果显示我们的系统可以至少以105bits/s的速率发送信息到手臂上38.10cm处。这样一个数据传输速率可以用来发送短消息，比如一个名字。我们计划以后对文字消息传输方面在更多物体和更多情景下做测试。

## 噪声影响

我们的系统可以至少以105bits/s的速率发送信息到手臂上38.10cm处。

一个接收端，这是我们计划以后继续研究的地方。

放大器的输出连接到一个音频接口设备(Fireface 800)上。

我们用一个iMac电脑接收从音频接口设备采集的数据。在发送端，和之前实验一样，我们用了一样的骨传导器来发送声音信号，唯一的区别在于我们将这个骨传导器连接到一个MacBook Pro声卡上来接收MacBook Pro生成的

如果发送的声音信号对应的频率有很强的环境噪音干扰，那么这将影响信号的传输和接收。因此，我们设计了另一个实验来评估无处不在的噪音对声音信号在人体中传播的影响。与之前实验类似，我们将骨传感器放在左手腕上，将加速传感器放在距离其7cm远处的手臂上。在第一次实验中，我们不

## Body通过身体发送文本消息

利用频移键控(FSK)方法作为编码方式，我们设计实现了一个系统通过

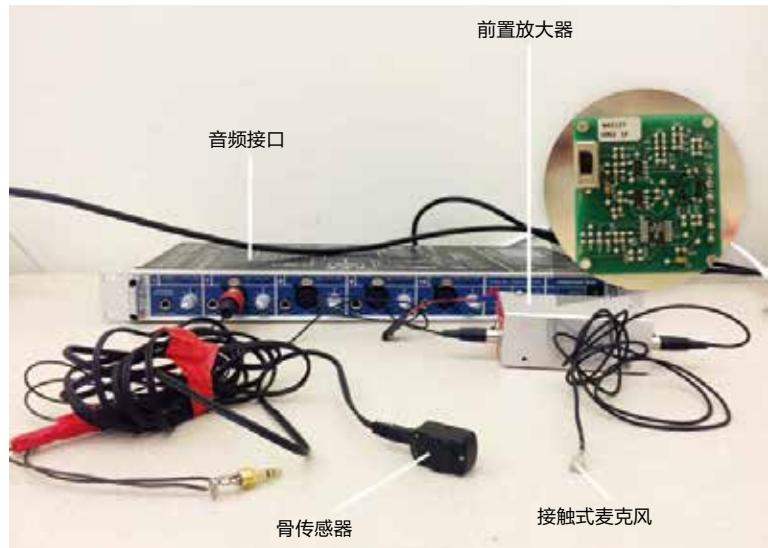


图 4. 数据传输实验的系统设置。麦克风连接到一个定制的低噪前置放大器用于放大和过滤信号。

播放噪声，记录250Hz到20kHz的频率响应结果；第二次实验中，我们在距离加速传感器58.4cm处通过一个播放器播放1kHz强度为90dB的噪声。得到的两次实验结果显示，在1kHz处信号幅值分别为-5.0dB和-4.6dB，即当有环境噪音时有0.3dB的增长。

身体的运动也可能引入噪音，因为人运动时接收端可能和皮肤摩擦产生噪音。精心选择接收器的位置和频率可以降低摩擦噪音的影响。然而，运动造成的摩擦噪音对信号传输的影响仍然是当前系统的一个限制。未来工作会在于如何消除这些摩擦的影响，无论是通过多个传感器的连接使用（将声音信号和同时的加速传感器数据结合在一起计算）或者是调控只有当运动幅度足够小时再传输数据。

## 用商用设备传输信号

传输低于20kHz的声音信号的一大好处在于大多数移动设备例如智能手表已经包含了加速传感器、陀螺仪等来接收和发送（用振动马达）信号例如声音信号。因此，我们采用了一个索尼智能手表3的200Hz振动马达来发送编码的声音信息，声音信息从手腕传导到桌子。我们发现接收信号有一个清晰的图形。在握手实验中，我们还成功的利用这个手表的麦克风作为接收端接收信

息，接收到的信号图形很明显。尽管这个实验是初级的，它证明了用商用设备来实现这个技术的可行性。

## 应用

许多潜在的应用可以受益于这项技术。正如我们之前解释的，无线通信在信息传输之前要求冗长的设置过程。用身体作为通信传输媒介，设备只有当它佩戴在身体上时才能通信。因为这项技术的设置过程可以在设备佩戴在身体上时自动完成，它可以节省不同设备之间相互通信之前设定的过程，提升用户体验。另外，它可以潜在的减小私人信息暴露到公共介质中的危险。

通过实实在在的物理接触来从可穿戴设备到环境之间传输信息还可以在很多应用中引入更自然的体验。用户身份验证就是一个例子，例如，不同于传统的输入密码或是用一个内置芯片的卡片来打开房门，用户可以只是

简单的握住门把手就能完成这个身份安全验证的过程，这个身份验证信息可以直接从手腕上戴的设备传输到门把手上。

我们的方法还可以用接触式的交互从环境中恢复信息。例如，当去博物馆时，用户可以从特定展览上获取更多的信息然后将这个信息显示在他们的可穿戴设备上例如智能手表或头戴式显示设备，简单的物理接触就可以让展览的相关信息传输到可穿戴设备上，而不是挤在人群中费力的凑近去一行一行地读，或是扫描一个二维码来获得相关的信息。



来的工作会着重于基于人体声音传导通信的其他挑战。我们这里描述的实验只探索了将传感器放在几个固定的位置上，以及参与者只做几个固定姿势。然而传感器放置的位置和传感器与皮肤之间结合的紧密度都会影响传

## 关于作者

**张铖**是美国佐治亚理工学院交互计算学院(Interactive Computing)的计算机系博士生，他的研究方向主要包括普适计算和人机交互，尤其是在通过创造新型感知方式和输入技术来提升移动设备和可穿戴设备的用户体验方面。张铖在中国科学研究院获得计算机应用硕士学位。联系方式: chengzhang@gatech.edu。

**希南·海尔赛克**是佐治亚理工学院电子与计算机工程系的博士生，他的研究兴趣主要在于非侵入型的生物医学设备用作生理监测的研究。他在佐治亚理工学院电子与计算机工程系获得硕士学位。联系方式: shersek3@gatech.edu。

**蒲奕名**是佐治亚理工学院交互计算学院的硕士研究生，她的研究兴趣主要在于新型设备和新技术与人的交互上，她在新加坡国立大学获得计算机学士学位。联系方式: amyyimingpu@gatech.edu。

**孙丹睿**是北京邮电大学通信工程系的本科生，她的研究兴趣包括交互设计和普适计算。联系方式: sundanrui@gmail.com。

**薛秋月**是北京大学信息科学技术学院的本科生，她的研究方向主要包括普适计算、人机交互和传感器网络。联系方式: qiuyuexue97@gmail.com。

**萨德·E·斯塔那**是佐治亚理工学院语境计算研究组(Contextual Computing Group)主任，是一位可穿戴设备研究的先驱，也是谷歌眼镜(Google Glass)的技术负责人。他是国际可穿戴计算设备会议(ACM/IEEE International Symposium on Wearable Computers)的创立者。他在麻省理工学院媒体实验室获得博士学位。联系方式: thad@gatech.edu。

**格雷戈里·D·艾博德**是佐治亚理工学院交互计算学院首席教授和杰出教授(J.Z. Liang Chair)。他的研究兴趣包括移动计算和普适计算，尤其是以实际应用为主导的研究。他在牛津大学获得计算机博士学位，他是IEEE会士和ACM会士，也是ACM SIGCHI Academy成员。联系方式: abowd@gatech.edu。

**奥默·T·伊南**是佐治亚理工学院电子与计算机工程系助理教授，生物工程系的教员，也是生物医学工程系的附属助理教授。他的研究兴趣包括非侵入型的生物医学检测和调制人体健康。他在斯坦福大学电子工程系获得博士学位。他是生物医学与健康信息学国际期刊(IEEE Journal of Biomedical and Health Informatics)的副主编，也是IEEE高级会员(IEEE senior member)。联系方式: inan@gatech.edu。

输出函数，这是我们目前系统的限制，导致当用户有剧烈运动时系统效果不够理想。

此外，我们现在的系统设计没有将骨传导器与空气隔离开来，因此当系统工作在5V电压、信号频率为人耳能听到的频率范围时系统发出的声音

会被用户听到。虽然实验参与者说当信号在他们身体内传播时他们没有感觉到任何不适，但是当发送的信号频率较低(例如250Hz)时用户仍然能感受到振动。

在未来工作中，我们打算研究传感器位置、传感器和制动器与人体结合的

紧密程度、信号频率对系统效率的影响以及研究相关的用户体验。我们还打算完成更多的相关示意性的相关应用，以及量化在不同情况下不同身体类型中传输数据的数据错误率。C

## 致谢

我们感谢审稿人的建设性意见反馈，以及用户实验中的参与者。这篇文章中描述的部分工作由佐治亚理工学院可穿戴计算中心资助。

## 参考文献

1. T. G. Zimmerman, "Personal Area

Networks: Near-Field Intrabody Communication," IBM Systems J., vol. 35, nos. 3–4, 1996, pp. 609–617.

2. E.R. Post et al., "Intrabody Buses for Data and Power," Proc. 1st Int'l Symp. Wearable Computers (ISWC '97), 1997; doi:10.1109/ISWC.1997.629919.

3. J. Park and P.P. Mercier, "Magnetic Human Body Communication," Proc. Ann. Int'l Conf. IEEE Eng. Medicine and Biology Soc. (EMBC '15), 2015, pp. 1841–1844.



《IEEE软件杂志》为软件从业者提供了前沿观点，专家分析和深刻的洞察，让他们跟上日新月异的技术变迁。这本杂志还为软件理论转化为应用提供了权威观点。

[www.computer.org/  
software/subscribe](http://www.computer.org/software/subscribe)

SUBSCRIBE TODAY



# 神经技术能给我们更好的大脑吗？

文 | 拉赞·巴特查里亚 (Rajan Bhattacharyya), HRL 实验室 (HRL Laboratories)  
布莱恩·A·考夫曼 (Brian A. Coffman), 匹兹堡大学 (University of Pittsburgh)  
周在勳 (Jaehoon Choe), 马修·E·菲利普斯 (Matthew E. Phillips), HRL 实验室  
译 | 朱机

用于临床的神经技术可以非侵入性地对大脑成像，但通常需要手术干预给予大脑刺激。尽管利用非侵入式刺激来增强神经系统功能正在成为越来越受欢迎的选择，但其有效性和伦理使用的问题还没未得到回答。

**随**着技术走出实验室、走向实际应用，可获取并分析大脑信号与身体信号的非侵入式记录设备开始变得越来越经济。用户现在可以使用这类设备监测自己的认知健康，或是利用监测结果来提升娱乐、游戏以及体能训练和程序培训的效果。此类记录设备在丰富娱乐、提升学习、帮助残障人士增强自主能力等方面具有无限潜力，使人们在人类增强方面获得长足进步。<sup>1</sup>举例来说，神经技术能够通过外部脑机接口实现交流，帮助个体回忆信息，整体改善认知表现。

增强现实设备可以检测人体感知系统感觉范围或视野以

外的信号帮助人提高知觉，<sup>2</sup>这种能力正在改变游戏、驾驶等活动。增强认知和神经人因工效学这两个密切相关的研究领域关注于从知觉、工作记忆和决策等角度检测和编码人的认知状态，通过调整信息传递来优化人处理信息的能力。

脑机接口用多种脑记录方法直接控制外部设备，利用外部设备辅助交流。用户可以借助设备在屏幕上挑选字词，<sup>3</sup>破解目标定位和轨迹来控制游标，<sup>4</sup>控制机器臂操纵物体，<sup>5</sup>甚至还能通过机器实现意识间的交流。<sup>6</sup>

目前，神经技术通过大脑训练和神经刺激两种途径改善认知功能。其中，大脑训练是让受试者调用局部脑回路和全脑

## 经颅电刺激具有简单、经济地改变脑功能并加快学习速度的潜力。

神经环路完成一些具有挑战性的任务，以此提高认知能力，调节神经功能。大脑训练的目的是在用户学习新技能时提高其水平，最终改变其行为，提升所能达到的表现极限。<sup>7</sup>不过，大脑训练学习到的技巧可以多大程度转化或泛化到日常任务中，研究人员对此未有定论。与大脑训练不同，神经刺激调节脑功能是直接靠外部手术改变脑功能。非侵入式神经刺激应用外部电磁场、声场或热场调节神经组织的生物物理特性，诱导神经可塑性的长时程改变。神经刺激形式多样，但目的都一样，即在不同脑区调节不同的神经功能（改变大脑可塑性）。刺激一回，效果可维持数分钟、数小时或数天不等。重复使用设备进行刺激，能够永久改变大脑的结构和功能。为了加深理解这些互补的神经技术如何增强人类大脑，我们从科学基础、设计原理到治疗效果、伦理应用等多方面考察了它们的应用潜力，并对其使用提出了一些有待讨论的问题。

### 了解经颅电刺激

神经刺激的形式之一是经颅电刺激(tES)，有时也被称作经颅电流刺激(tCS)。由于被认为具有简单、经济地改变脑功能并加快学习速度的潜力，这种神经刺激方式正在变得普及。在美国国家医学博物馆的在线信息检索平台

PubMed上，以常见的经颅电刺激形式“经颅直流电刺激(tDCS)”为关键词搜索，2000年仅有5篇文献，而在2016年已有696篇。

经颅电刺激作为一种电疗法，在各种场合应用了数百年，然而其背后的科学机制却是到了最近二十年由于脑成像工具和建模的推动才有相关的研究结果陆续发表。<sup>8</sup>经颅电刺激仪用电流来调节神经元功能。

经颅电刺激仪通过电极向头皮施加电流，由于随着脑组织深度加大电极的电场强度会陡然降低，因此用这种方法最容易获取的目标是大脑表层区域，也就是最靠近头皮的部位。此外，电极的场强关系还受个体生理结构、脑内部活动状态的影响。提高结构成像的分辨率和电极的密度或可提高定位精度。

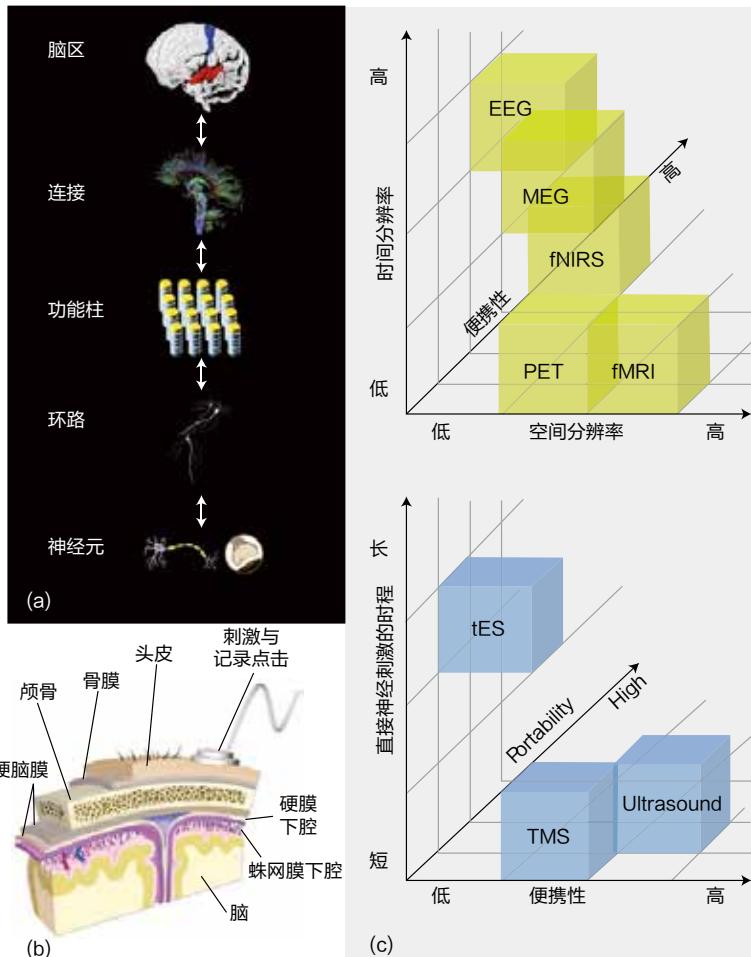
目前，研究者正借助神经影像和电磁场建模来找出经颅电刺激造成脑内变化的确切神经机制。一般认为，“阳极型”经颅直流电刺激会增强神经元兴奋性，“阴极型”刺激抑制兴奋性。不过，近来一些建模研究显示，对于不同类型（兴奋型或抑制型）、不同方位（辐射状或正切方位）以及同一个神经元的不同部位（轴突、树突还是胞体），经颅直流电刺激的效应不同。虽然在啮齿类动物和体外脑片上的实验结果显示电刺激对神经环路和单个神经元有一定程度的调制效果，但目前为止，经颅直

流电刺激如何影响灵长类的神经元和神经环路还没有结论。总体而言，经颅电刺激的有效性并未得到研究者的认可。<sup>9</sup>

### 无创检测面临的挑战

非侵入式神经刺激的最大挑战恐怕是如何在行为和所测的脑活动之间建立因果关系。在图1a显示的各个层次上，脑的复杂性都必须予以考虑。在区域层次上的复杂性是脑区间功能网络的活动模式，不同区域的连接介导了细胞组成层次上的通信动态变化。在功能柱层次上，特定结构（例如视皮层）中空间位置聚集在一起的功能柱表现出神经元之间有较强的关联和反应特性。神经元网络存在于有明确定义的计算环路中，例如呈条状的皮层。最后，在神经元层次，胞内结构和分子机制计算信息，例如树突内处理信息和受体表达通路等。

图1b是无创记录时颅骨和传感器的剖视图，其中传感器检测头皮传来的神经活动标志信号。这些信号到达头皮之前要穿过多种组织，可向空间各个方面“弥散”，并且还会受肌肉或血液等非神经组织的信号的干扰破坏。在计算上，这些伪差需要做复杂的滤波处理，这里既有积极的也有消极的因素，因为消除伪差对测量脑活动提出了重要挑



**图1.** 脑的复杂性。(a) 从神经元到脑区的各个层次。(b) 无创检测的传感器必须穿透不同组织才能检测和刺激特定脑区。电活动、磁活动、血流动力学的测量都必须消除脑组织上方干扰层(颅骨、皮肤等)导致的空间、时间伪影。(c) 不同检测和刺激技术有不同的空间和时间采样特性和便携程度。EEG: 脑电图; MEG: 脑磁图; fNIRS: 功能红外光谱; PET: 正电子发射断层扫描(放射医学或核医学成像, 能产生显示身体机能的3D图像); fMRI: 功能磁共振成像; tES: 经颅电刺激; TMS: 经颅磁刺激。

战。正因为有信号污染, 个体的神经元和环路的神经活动(电锋和局部场电位)并不容易定量和重建。无创检测要结合信号处理技术与算法去除传感器接收数据中的噪声, 降低维度, 重建出神经活动的信号源。

如图1c所示, 每种技术的空间、时间采样特性不同, 便携程度不同。例如脑电图(EEG)、脑磁图(MEG)等电磁检测手段, 是通过检测神经活动产生的电场或磁场测量初级活动。这两种技术

可以通过头皮非侵入式地记录脑代谢等次级活动。功能近红外光谱(fNIRS)利用携带方便、成本较低的红外光源和检测仪测量血氧的摄取。功能磁共振成像(fMRI)则是用大型的、通常来说很昂贵的磁共振仪器记录血红蛋白磁敏感性的变化, 测的也是血氧含量。

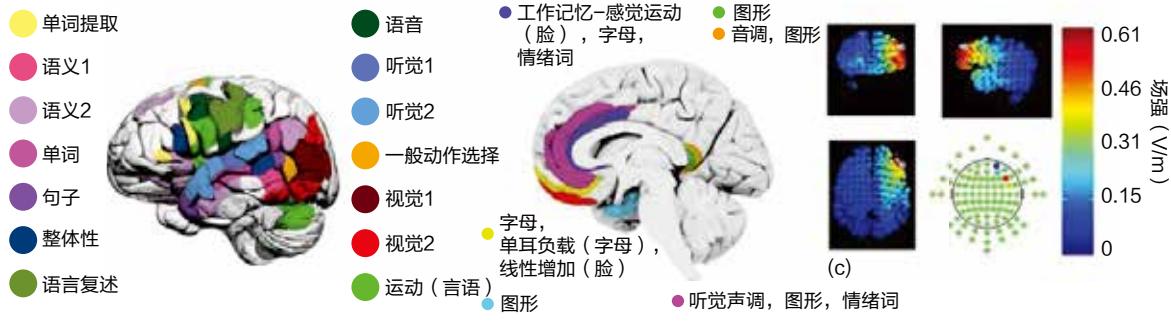
## 提高认知功能

目前有关神经增强的研究主要强

调其在调解、甚至提高基础认知功能方面的应用潜力, 因为实验范式已得到稳定的结果, 并且研究者对于这些系统如何运作有一致的基本认识。比如工作记忆, 是受到广泛研究的一项基础认知功能, 在注意选择、推理、记忆和决策制定等功能上必不可缺, 与信息的暂时储存和处理有关。图2a展示了一种区域多样性和复杂性的风格; 图2b结合了多项研究结果, 这些研究表明在工作记忆发挥作用时脑内不同区域有显著的活动。<sup>11</sup>

为提高这些功能, 非侵入式刺激必须利用不同的电极导联提供电流来定位相关区域。计算机模拟可以用于设计电极导联方式, 以使电极记录到深部脑结构的信号, 同时计算机模拟也能权衡场强、导联复杂程度与增强空间聚焦的关系。图2c显示的一种优化方案利用有限元模型(FEM)模拟了一种电极安置设计的场强。

研究人员评估工作记忆表现的一个常用范式叫作N-back任务。被试者在这一连续实施的任务中会看到一连串图像, 并被要求对该序列中匹配的图像作出判断。1-back即比较当前刺激或图像与前一个图像, 2-back则是当前刺激与之前的倒数第二个进行比较, 以此类推。刺激背外侧前额叶皮层(DLPFC)被认为可以增强基础认知功能, 包括完成N-back任务时用到的工



**图2.**不同脑区诱发的功能活动和靶刺激模型。(a)皮层表面一般认知过程的神经活动。(b)脑内视图,显示日常认知功能不可或缺的各种工作记忆任务被诱发的脑区。在图a和图b中,除非特别指明,标出部分是以n-back任务测试。(c)以有限元模型优化经颅电刺激的脑区定位。这个例子中,电极刺激部位的靶向是右半脑背外侧前额叶皮层(rDLPFC),这是与执行功能和工作记忆有关的主要区域,图中分别显示了大脑的不同方位,右下显示的是10-20电极系统放置法。

作记忆。<sup>12</sup>

然而,改善的证据在依照不同方法的不同研究中差异极大。有一项研究显示,进行N-back任务时在背外侧前额叶皮层重复施加经颅磁刺激(rTMS)可以提高正确应答的比例并缩短应答时间;但另一项实验则报告称,经颅直流电刺激只能改善反应时间。<sup>13</sup>更令人困惑的是,最近一项元研究的作者得出结论说,经颅直流电刺激在这些认知测量上全都没有实际效果。<sup>14</sup>显然,我们还需要更多研究来统一这些迥然不同的实验结论,并推动目前对各种神经增强方法背后的根本机制的认识。要想重复这些方法并优化其效果,前提正是对机制的理解。

## 语言习得

尽管有上述矛盾,一些研究还是表明刺激背外侧前额叶皮层增强了工作记忆,因而能在此基础上增强高级认知功能,比如语言的习得。非侵入式神经刺激技术尤其适用于调节接近皮层表面、特定的脑功能。在语言习得和技能习得等复杂认知功能中,学习速度是一个重

要的测量指标,而加快学习速度的能力倚赖刺激技术的空间和时间特性。语言习得包括语音的发生和理解、语义和词汇的理解等,习得过程中多个区域之间的神经网络会被激活。

除了布洛卡区、韦尼克区这两个主要脑区外(分别与语言产生和语言理解关系最密切的两个区域),涉及工作记忆的多个离散分布的区域和神经网络也参与其中。一些研究表明,语言产生和言语工作记忆有共同的神经基础,而对此施加经颅直流电刺激能够加快语言学习速度并提高掌握新单词的成功率。<sup>15</sup>

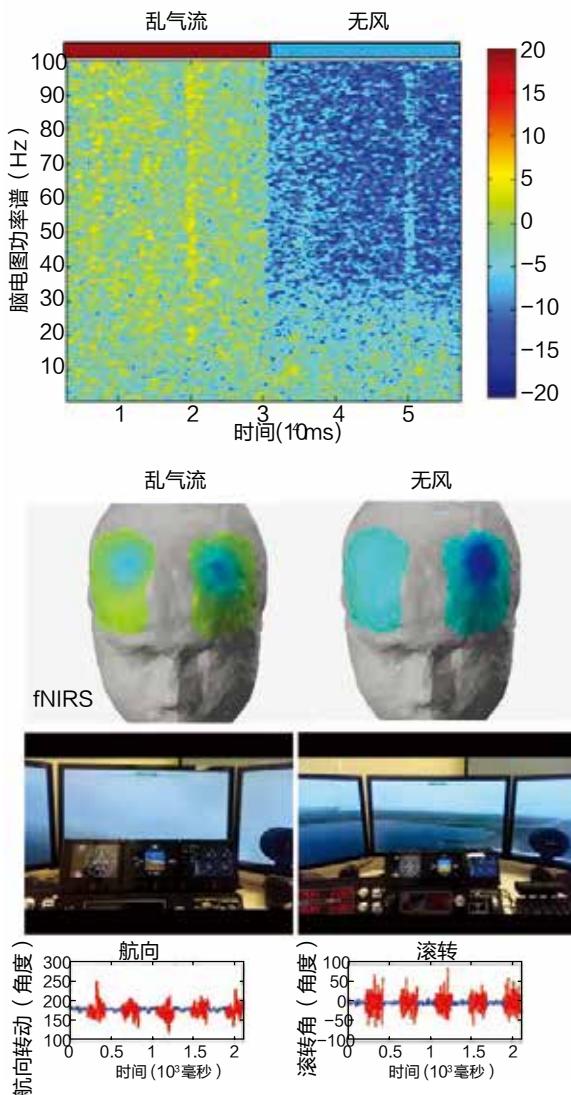
## 获得复杂技巧

尽管已有充足的研究证据表明神经刺激技术能够有效增强诸如工作记忆等基础认知功能,但这些得到增强的技能多大程度上可以推广到现实世界的任务,比如学习复杂的新程序,还是个问题。像开飞机就是这样一种任务,需要掌握多引擎、制定决策、学习知觉技能并融会贯通;探测伪装目标也是如此,需要视觉感知的学习。

**驾驶飞机。**为了加强了解神经刺激如何加快复杂技能的学习速度,我们开展了一项实验,给健康被试的背外侧前额叶皮层施加经颅直流电刺激,并指导被试在飞行模拟器上学习驾驶飞机。我们评估了两组被试使用Beechcraft Baron B-58 X-Plane飞行模拟器学习飞机着陆的学习能力。其中一组,我们采用经颅直流电刺激施加了2毫安的刺激;另一组完成相同的任务但我们给予的是假刺激,尽管被试也能感受到有刺激发放,但经颅电刺激的程序并未真正施加。

我们测了五种飞行控制试验,天气条件包括从无风到乱气流,连续进行四天。我们使用脑电图记录电位,同时用功能近红外光谱测量血氧代谢。然后,我们评估了四天的飞行学习过程中被试脑活动的功能变化。图3显示的是在乱气流天气(风速10节,乱气流10%)和无风天气时维持飞机水平飞行的结果。我们的发现指明了飞行时认知负载的一个特征标志,反映被试在维持水平飞行时全神贯注的程度。

与接受假刺激的对照组相比,从飞



**图3.** 运用经颅直流电刺激（经颅电刺激的一种常用形式，使用直流电进行刺激）调节被试在飞行模拟器上学习飞机着陆的能力。在乱气流天气条件控制飞行时，双侧背外侧前额叶皮层的脑电图神经成像（加权独立分量分析得到的功率谱）显示额叶的宽频带功率（不确定什么意思）有显著增加。与脑电图同时进行的功能近红外光谱的神经成像结果显示血红蛋白浓度有所上升。航向转动（heading）和滚转角（roll）是乱气流试验组（红色）和无风天气试验组所给的控制输入。

机着陆触地时对跑道的重力加速度测算，经颅直流电刺激帮助增强掌握驾驶飞机成功着陆技能的连贯性。<sup>16</sup>这一借助神经刺激学习真实世界技能的示范表明，这些方法或许可以运用到实验心理测试和标准心理测试以外的领域，甚至也许适用于学校教学和训练场合，用以加快学习速度和提高成绩。

神经刺激方法在其他领域和用途上的表现还需要更多研究进行测试。一个令人关切的方面是技能的保持，像我们的飞行员实验就可以开展纵向追踪研究来测试被试者掌握的新技能可以保持多久。

**探测伪装目标。** 2012年，新墨西哥大学的文森·克拉克（Vince Clark）等人

开展了一项用经颅直流电刺激增强表现的开创性研究。他们的研究报告称，在被试的右前额叶腹外侧（rVLPFC）施加30分钟2毫安阳极型经颅直流电刺激后，被试的目标探测能力表现出可观的提升。<sup>17</sup>他们给被试安排的任务是在一款用于训练中东部署军事人员的模拟装置DARWARS Ambush!上找出计算机虚拟环境中隐藏的探测目标。利用另一项的神经影像学结果，研究人员通过血液动力学反应在伪装目标探测专业水准各异的被试中确定了靶区域（即右前额叶腹外侧）。图4a显示的是2012年这项研究的结果。

利用功能核磁共振，克拉克团队在中间学习阶段的被试中找到了一组新手中没有的关键区域。这些区域包括右顶叶皮层（与视觉注意有关的一个区域），内侧前额叶皮层（动作监测时激活的一个区域），海马旁回（参与学习和记忆的区域），以及右前额叶腹侧皮层（与注意力认知控制和工作记忆有关的一个区域）。随后的实验中，该团队在找到的这些皮层区域的外表面上应用经颅直流电刺激。如图4b所示，对准右半脑腹外侧前额叶皮层和右半脑顶叶皮层的经颅直流电刺激产生了显著的表现增强现象。

有几项研究支持这一结果，强调了以神经影像学为指导的经颅直流电刺激的鲁棒性。有一项研究显示，对右前

额叶腹外侧应用经颅直流电刺激后目标探测的提高和注意力增强有相关性，这也说明利用神经影像学技术的脑刺激是有效的。<sup>18</sup>然而，尽管近来有需求，经颅直流电刺激方面的研究还远远缺乏重复结果。

## 增强运动技能

已有众多研究组证实神经刺激对增强运动技能有效，并重复出了几个先驱实验室的早期研究结果。对运动学习的调查研究表明，经颅直流电刺激可在数天内全面提高被试习得的运动技能，提升人的力量、灵活性和运动效能。不过，和很多神经调节研究一样，不同实验的结果差异很大。有一项研究显示，在一项对初级运动皮层应用经颅直流电刺激的实验中，超过半数的被试无明显反应。<sup>19</sup>还有一些研究则观察到，电流、电流密度和经颅直流电刺激诱发的运动和视觉运动学习的反应效应值呈现负相关。<sup>20</sup>经颅直流电刺激在增强运动技能方面的商业应用需要进一步针对保持能力开展更多研究，目前这一方面已发表的文献还相当有限。<sup>21-23</sup>

## 神经技术的市场发展趋势

尽管对经颅电流刺激的效果还有

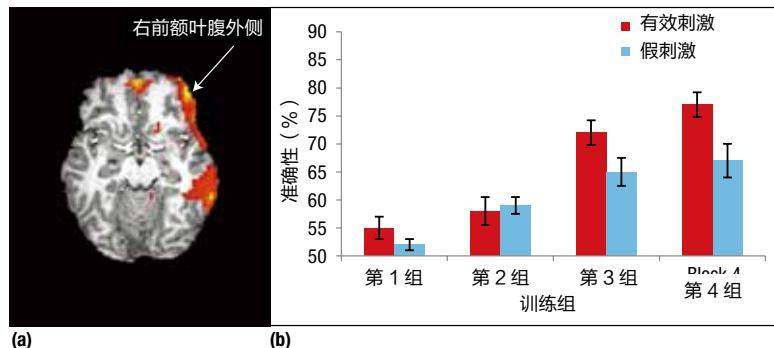


图 4. 给检测伪装目标的被试在右前额叶腹外侧皮层 (rVLPFC) 施加 30 分钟 2 毫安阳极型经颅直流电刺激，所得到的神经影像学研究结果示例和行为效果。(a) F-statistic 图显示在伪装目标检测初级水平被试和中级水平被试之间比较线索相关活动 (有线索与无线索)，右前额叶腹外侧皮层的血氧浓度依赖 (BOLD) 反应增强 (b)，其中 12 个被试接受的是有效刺激 (红色)，23 个被试接受的是假刺激 (蓝色)。误差范围用的是标准误差。

诸多问题有待回答，但这种技术仍然是目前最常用的神经刺激方式，甚至已经出现了 DIY 用户社区 ([www.diytdcs.com](http://www.diytdcs.com))，《纽约客》上还有专门的评论文章。<sup>8</sup>

我们也把经颅电刺激及一般神经技术有哪些商业发展趋势列入我们关于神经技术的科学与挑战的调查，包括当前的监管环境以及设备的简单性、经济性和实用性。

## 监管和设备特性

进入临床设备市场需要巨大投资以获得美国食品和药物管理局 (FDA) 的认证和批准。这类大规模工作有一些先例，像通过侵入式神经刺激 (如深脑刺激) 和非侵入式神经刺激 (如经颅直流电刺激) 改善帕金森病患者的运动技能。

不过，美国食品和药物管理局并不监管非临床应用的经颅电刺激技术，而这部分是一个强大的市场，各公司瞄准了神经增强的基础科研和消费者使用。供消费者使用的设备一般包含施加在头皮或前额的电极，它们与连接接口控制器的刺激电极相连。经颅电刺激仪的许多设备特征和设计难度是与其他非侵入式神经记录技术一样的。目前市场上的大部分基础技术是由一些公司在过去几十年的前期投资推动的，他们对脑的研究是为了美国国防部先进研究项目局 (DARPA) 等机构现场应用。

神经记录市场上的消费者设备正在快速成熟，主要是靠早期在游戏中的应用。此外，像 Experiential Technology 等会议 (XTech; [www.nerogamingconf.com](http://www.nerogamingconf.com)) 正受到关注。用于冥想等活动的健康相关监测和神经增

强也是神经记录设备的一大应用。

神经技术的消费市场还在继续扩大，尽管关键问题仍然存在，比如实现以易用型干电极技术记录信号。

### 实用性和用途

神经技术的实用程度主要取决于设备、应用和用户体验。十年前，神经记录主要由实验室里受过训练的技术员完成，技术员要使用传感器，校准记录设备，在被试参与任务的同时监测信号质量。现在的设备可以有用户自行操

用一次？

- › 不同的设备会有不一样的效果吗？
- › 该设备对不同人有相同的效果吗？

**设计考虑。**神经技术的应用不再局限于固定在室内的装置，比如游戏机之类。不过，要放到移动应用上就带来了新的设计考虑因素，比如环境、用户和活动条件。设计者现在必须考虑温度和湿度，用户的汗液和油脂会影响组织接口处的电极阻抗，还要考虑那些会改变接口和信号质量的动作。大部分无线

应用外，还包括在治疗抑郁症、帕金森病、脑卒中的临床应用上。经颅直流电刺激在认知增强的效果尤其显著，提高程度几乎达到一个标准差。<sup>12</sup>并且，随着新流程和新硬件的开发，分辨率提高，刺激稳定性增强，增强效果还会更大。此外，新型设备将在一个单一闭环系统中把经颅直流电刺激、脑电图和功能远红外谱组合在一起，对脑同时进行监测和刺激。<sup>16</sup>像这样的发展将拓宽非侵入式脑刺激的应用，并增加已有相当数量规模的增强应用。

### 神经技术的实用程度主要取决于设备、应用和用户体验。

作，只需微调校准。比方说大部分经颅电刺激仪，带有自动阻抗测量，可以保障安全。不过，研究人员还只是刚刚开始了解这种技术的效果，这为消费者带来了重要的问题：

- › 该设备真的会改变我的大脑吗？
- › 它有多复杂？需要花多长时间进行校准？
- › 电极安放有多重要？错误安放会有什么后果？
- › 要获得预期的增强效果必须多久

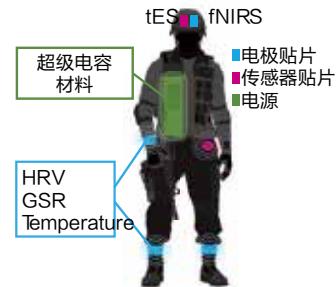
经颅电刺激仪是用电池供电的，轻便，体积小，适合带出门使用。未来的设备或许还会用在扩展活动中，比如需要在不访问基础设施的情况下间歇使用几天的军事场景。这类设备就必须耐用，有效地消耗电力。<sup>24</sup>

**应用发展。**有关非侵入式脑刺激的研究非常丰富。经颅电刺激的研究在过去几十年里也十分繁荣，除了在工作记忆、运动学习、视知觉学习、简单体感和视觉运动感知学习以及记忆等认知

### 有待讨论的问题

尽管很多研究已证明神经增强技术在多个应用领域和功能领域的有效性，但在把实验结果转换为用于工作环境和家庭环境的主流技术之前，还有许多问题和关键因素是必须先作出回答的。检测设备和刺激设备的实际限制都超出了可穿戴式传感器和设备的常用尺寸、重量、功率和成本考虑。举例来说，为军事使用而设计的神经增强设备必须考虑极端环境，战斗用途必需的耐用性，可适应于不同个体的设备，以及人的出汗程度、头发长短、服装或运动的限制等实际问题。

图5显示的是一种满足这些要求的潜在应用。检测到的生物学数据可以与刺激范式整合，以便只在个体需要的时



候才发挥神经增强的作用。这种适应性干预要基于个性化的个体神经生理学状态，以及随时出现的新任务要求。

在我们能够实现这些用途之前，有许多问题必须先行解决。挑战之一是方法上的选择，设立可以计划和指导神经增强的流程和参数。举例来说，测量或调节所需的生理或行为需要多长时间？神经调制的最佳剂量或收益可持续的时间量也没有定论。一年左右是迄今报道过的最长时间。调节神经的措施和剂量有多可靠和多有效，以及何时应当施加或停止等，都还不清楚。需要进一步实验来确定有效的测量过程和评估方法，评估相对于假刺激、距离达成目标生理状态或行为状态的进展程度。最后，持续使用神经增强对长期健康有何风险，目前也还不清楚。要把治疗开发为有市场价值的疗法，需要大量的时间和资源来开展和重复风险研究。

# 尽

管存在障碍，神经增强技术还是有着光明的未来。

目前，在可穿戴式传感器和设备、个性化医疗、可定制模型等方面进展很可能加速神经增强的研究，帮助我们理解神经增强工具的范围以及它们对表现、学习和训练的影响。有了这些理解，设计者能生产出集合多种可穿戴式传感器和调节设备的适应性系统。而这种系统的应用将只会受我们

对生物学、认知和人类物理极限的限制。■

## 参考文献

- H. Guizzo and H. Goldstein, "The Rise of the Body Bots," *IEEE Spectrum*, vol. 42, no. 10, 2005, pp. 50–56.
- S. Tachi, M. Inami, and Y. Uema, "Augmented Reality Helps Drivers See Around Blind Spots," *IEEE Spectrum*, vol. 31, 2014; [spectrum.ieee.org/transportation/advanced-cars/augmented-reality-helps-drivers-see-around-blind-spots](http://spectrum.ieee.org/transportation/advanced-cars/augmented-reality-helps-drivers-see-around-blind-spots).
- C. Guger et al., "How Many People Are Able to Control a P300-Based Brain-Computer Interface (BCI)?" *Neuroscience Letters*, vol. 462, no. 1, 2009, pp. 94–98.
- T. Aflalo et al., "Decoding Motor Imagery from the Posterior Parietal Cortex of a Tetraplegic Human," *Science*, vol. 348, no. 6237, 2015, pp. 906–910.
- L.R. Hochberg et al., "Reach and Grasp by People with Tetraplegia using a Neurally Controlled Robotic Arm," *Nature*, vol. 485, no. 7398, 2012, pp. 372–375.
- R.P. Rao et al., "A Direct
- Brain-to-Brain Interface in Humans," *PLOS ONE*, vol. 9, no. 11, 2014; [dx.doi.org/10.1371/journal.pone.0111332](https://dx.doi.org/10.1371/journal.pone.0111332).
- S. Rabipour and A. Raz, "Training the Brain: Fact and Fad in Cognitive and Behavioral Remediation," *Brain Cognition*, vol. 79, no. 2, 2012, pp. 159–179.
- E. Batuman, "Electrified: Adventures in Transcranial Direct-Current Stimulation," *The New Yorker*, 2015; [www.newyorker.com/magazine/2015/04/06/electrified](http://www.newyorker.com/magazine/2015/04/06/electrified).
- A.-K. Brem et al., "Is Neuroenhancement by Noninvasive Brain Stimulation a Net Zero-Sum Proposition?" *NeuroImage*, vol. 85, no. 3, 2014, pp. 1058–1068.
- T. Wagner, A. Valero-Cabré, and A. Pascual-Leone, "Noninvasive human brain stimulation," *Ann. Rev. of Biomedical Eng.*, vol. 9, 2007, pp. 527–565.

## 拓展阅读

**我们**们只涉及了文献中的不少研究和报告，按主题列出了以下资料来源，供有兴趣的读者进一步探究。

### 语言

D.J. Acheson et al., “A Common Neural Substrate for Language Production and Verbal Working Memory,” *J. Cognitive Neuroscience*, vol. 23, no. 6, 2011, pp. 1358–1367.

A. Flöel et al., “Noninvasive Brain Stimulation Improves Language Learning,” *J. Cognitive Neuroscience*, vol. 20, no. 8, 2008, pp. 1415–1422.

C.J. Price, “A Review and Synthesis of the First 20 Years of PET and fMRI Studies of Heard Speech, Spoken Language and Reading,” *Neuroimage*, vol. 62, no. 2, 2012, pp. 816–847.

### 运动技能

A. Antal et al., “Direct Current Stimulation over V5 Enhances Visuomotor Coordination by Improving Motion Perception in Humans,” *J. Cognitive Neuroscience*, vol. 16, no. 4, 2004, pp. 521–527.

M. Cambiaggi et al., “Brain Transcranial Direct Current Stimulation Modulates Motor Excitability in Mice,” *Eur. J. Neuroscience*, vol. 31, no. 4, 2010, pp. 704–709.

K. Monte-Silva et al., “Induction of Late LTP-Like Plasticity in the Human Motor Cortex by Repeated Noninvasive Brain Stimulation,” *Brain Stimulation*, vol. 6, no. 3, 2013, pp. 424–432.

A. Pasqualotto, B. Kobanbay, and M.J. Proulx, “Neural Stimulation Has a Long-Term Effect on Foreign Vocabulary Acquisition,” *J. Neural Plasticity*, 2015, article 671705.

E. Pavlova et al., “Transcranial Direct Current Stimulation of the Premotor Cortex: Effects on Hand Dexterity,” *Brain Research*, vol. 1576, 2014, pp. 52–62.

J. Reis et al., “Noninvasive Cortical Stimulation Enhances Motor Skill Acquisition over Multiple Days through an Effect on Consolidation,” *Proc. US Nat’l Academy of Science*, vol. 106, no. 5, 2009, pp. 1590–1595.

P.S. Williams, R.L. Hoffman, and B.C. Clark, “Preliminary Evidence that Anodal Transcranial Direct Current Stimulation Enhances Time to Task Failure of a Sustained Submaximal Contraction,” *PLOS ONE*, vol. 8, no. 12, 2013; doi: 10.1371/journal.pone.0081418.

### 临床应用

P.S. Boggio et al., “A Randomized, Double-Blind Clinical Trial on the Efficacy of Cortical Direct Current Stimulation for the Treatment of Major Depression,” *Int’l J. Neuropsychopharmacology*, vol. 11, no. 2, 2008, pp. 249–254.

F. Fregni et al., “Effects of Transcranial Direct Current Stimulation Coupled with Repetitive Electrical Stimulation on Cortical Spreading Depression,” *Experimental Neurology*, vol. 204, no. 1, 2007, pp. 462–466.

F. Fregni et al., “Noninvasive Brain Stimulation for

11. A.M. Owen et al., “N-Back Working Memory Paradigm: A Meta-Analysis of Normative Functional Neuroimaging Studies,” *Human Brain Mapping*, vol. 25, no. 1, 2005, pp. 46–59.

12. B.A. Coffman, V.P. Clark, and

13. A.R. Brunoni and M.-A. Vander-  
hasselt, “Working Memory Improvement with Noninvasive Brain Stimulation of the Dorsolateral Prefrontal Cortex: A Systematic Review and Meta-Analysis,” *Brain and Cognition*,
- R. Parasuraman, “Battery Powered Thought: Enhancement of Attention, Learning, and Memory in Healthy Adults using Transcranial Direct Current Stimulation,” *NeuroImage*, vol. 85, no. 3, 2014, pp. 895–908.

Parkinson's Disease: A Systematic Review and Meta-Analysis of the Literature," *J. Neurology, Neurosurgery, and Psychiatry*, vol. 76, no. 12, 2005, pp. 1614–1623.

B.R. Webster, P.A. Celnik, and L.G. Cohen, "Noninvasive Brain Stimulation in Stroke Rehabilitation," *NeuroRX*, vol. 3, no. 4, 2006, pp. 474–481.

### 威胁检测

B.A. Coffman, M.C. Trumbo, and V.P. Clark, "Enhancement of Object Detection with Transcranial Direct Current Stimulation Is Associated with Increased Attention," *BMC Neuroscience*, vol. 13, no. 1, 2012, p. 108.

B. Falcone et al., "Transcranial Direct Current Stimulation Augments Perceptual Sensitivity and 24-Hour Retention in a Complex Threat Detection Task," *PLOS ONE*, vol. 7, no. 4, 2012; doi: 10.1371/journal.pone.0034993.

R.A. McKinley et al., "Acceleration of Image Analyst Training with Transcranial Direct Current Stimulation," *Behavioral Neuroscience*, vol. 127, no. 6, 2013, pp. 936–946.

### 经颅直流电刺激的技术、建模和效果

D. Edwards et al., "Physiological and Modeling Evidence for Focal Transcranial Electrical Brain Stimulation in Humans: A Basis for High-Definition tDCS," *NeuroImage*, vol. 74, 2013, pp. 266–275.

J.C. Horvath, O. Carter, and J.D. Forte, "Transcranial Direct Current Stimulation: Five Important Issues We Aren't

Discussing (but Probably Should Be)," *Frontiers in Systems Neuroscience*, vol. 8, 2014, article 2.

J.C. Horvath, J.D. Forte, and O. Carter, "Evidence That Transcranial Direct Current Stimulation (tDCS) Generates Little-to-No Reliable Neurophysiologic Effect beyond MEP Amplitude Modulation in Healthy Human Subjects: A Systematic Review," *Neuropsychologia*, vol. 66, 2014, pp. 213–236.

B. Molaei-Ardekani et al., "Effects of Transcranial Direct Current Stimulation (tDCS) on Cortical Activity: A Computational Modeling Study," *Brain Stimulation*, vol. 6, no. 1, 2013, pp. 25–39.

A. Rahman et al., "Cellular Effects of Acute Direct Current Simulation: Somatic and Synaptic Terminal Effects," *J. Physiology*, vol. 591, no. 10, 2013, pp. 2563–2578.

### 记忆

F. Fregni et al., "Anodal Transcranial Direct Current Stimulation of Prefrontal Cortex Enhances Working Memory," *Experimental Brain Research*, vol. 166, no. 1, 2005, pp. 23–30.

L. Marshall et al., "Transcranial Direct Current Stimulation during Sleep Improves Declarative Memory," *J. Neuroscience*, vol. 24, no. 44, 2004, pp. 9985–9992.

vol. 86, 2014; doi: 10.1016/j.bandc.2014.01.008.

14. J. C. Horvath, J. D. Forte, and O. Carter, "Quantitative Review Finds No Evidence of Cognitive Effects in Healthy Populations from Single-Session Transcranial Direct

Current Stimulation (tDCS)," *Brain Stimulation*, vol. 8, no. 3, 2015, pp. 535–550.

15. A.R. Price et al., "A Meta-Analysis of Transcranial Direct Current Stimulation Studies Examining the Reliability of Effects on Language

Measures," *Brain Stimulation*, vol. 8, no. 6, 2015, pp. 1093–1100.

16. J. Choe et al., "Transcranial Direct Current Stimulation Modulates Neuronal Activity and Learning in Pilot Training," *Frontiers in Human Neuroscience*,

## 关于作者

**拉赞·巴特查里亚**在HRL实验室担任信息与系统科学实验室高级研究工程师，人类和机器认知研究中心主任，企业合作经理。他的研究兴趣包括神经认知建模、脑机接口和人工智能。他在加州理工大学获计算机和神经系统博士学位，IEEE成员。联系方式: [rajan@hrl.com](mailto:rajan@hrl.com)

**布莱恩·A·考夫曼**是匹兹堡大学医学院精神病学系博士后助理研究员，他的研究兴趣包括经颅直流电刺激的应用以及其他用于增强认知、减轻精神分裂等精神失常临床症状的非侵入式脑刺激技术。他在新墨西哥大学实验心理学系获博士学位。联系方式: [coffmanb@upmc.edu](mailto:coffmanb@upmc.edu)。

**周在勳**是HRL实验室信息系统科学实验室研究员。他的研究兴趣包括脑机接口、神经假体和神经建模。他在加州大学洛杉矶分校获神经科学博士学位。联系方式: [jchoe@hrl.com](mailto:jchoe@hrl.com)。

**马修·E·菲利普斯**是计算机科学家，在HRL实验室担任人类和机器认知研究中心的副主任。他的研究兴趣包括神经刺激，多模态脑成像，认识心理学，应用和系统神经科学，计算机和神经建模，以及人工智能。菲利普在耶鲁大学获物理学博士学位。联系方式: [mephillips@hrl.com](mailto:mephillips@hrl.com)。

### 译者简介：

**朱机**是科普作者和译者，中国科学院上海神经所博士。

vol. 10, 2016; doi: 10.3389/fnhum.2016.00034.

17. V.P. Clark et al., “tDCS Guided using fMRI Significantly Accelerates Learning to Identify Concealed Objects,” *NeuroImage*, vol. 59, no. 1, 2012, pp. 117–128.

18. B.A. Coffman et al., “Impact of tDCS on Performance and Learning of Target Detection: Interaction with Stimulus Characteristics

and Experimental Design,” *Neuropsychologia*, vol. 50, no. 7, 2012, pp. 1594–1602.

19. S. Wiethoff, M. Hamada, and J.C. Rothwell, “Variability in Response to Transcranial Direct Current Stimulation of the Motor Cortex,” *Brain Stimulation*, vol. 7, no. 3, 2014, pp. 468–475.  
20. H. Vollmann et al., “Anodal Transcranial Direct Current Stimulation

(tDCS) over Supplementary Motor Area (SMA) but Not Pre-SMA Promotes Short-Term Visuomotor Learning,” *Brain Stimulation*, vol. 6, no. 2, 2013, pp. 101–107.

21. A. Snowball et al., “Long-Term Enhancement of Brain Function and Cognition using Cognitive Training and Brain Stimulation,” *Current Biology*, vol. 23, no. 11, 2013, pp. 987–992.

22. S. Lefebvre et al., “Single Session of Dual-tDCS Transiently Improves Precision Grip and Dexterity of the Paretic Hand after Stroke,” *Neuro-rehabilitation and Neural Repair*, vol. 28, no. 2, 2013, pp. 100–110.

23. S. Waters-Metenier et al., “Bihemispheric Transcranial Direct Current Stimulation Enhances Effector-Independent Representations of Motor Synergy and Sequence Learning,” *J. Neuroscience*, vol. 34, no. 3, 2014, pp. 1037–1050.

24. M. Phillips and M. Ziegler, “A Neurostimulation-Based Advanced Training System for Human Performance Augmentation,” *Brain Stimulation*, vol. 2, no. 7, 2014, pp. e11–e12.

# CNN

He that would eat the **kernel** must  
crack the nut.

---

不劳无获

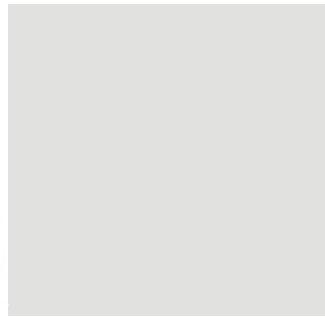


LUCKY DATA  
搜狐图文匹配算法大赛

搜狐 | IEEE-China | 清华大学



[www.luckydata.org](http://www.luckydata.org)



## 搜狐 大数据中心

搜狐集团于2014年4月成立大数据中心，旨在打通搜狐集团旗下所有产品的用户行为数据，为用户提供更好的产品体验，并提升资源的商业化变现能力。大数据中心由广告技术、用户推荐、数据仓库和商业化探索等多个团队组成，团队核心成员均来自清华、北大、中科院、北邮、北航等一流院校的数学、计算机等院系。

中心成立以来，推出了完美风暴、先声夺人等创新广告产品，并在个性化内容推荐方面与搜狐新闻、搜狐视频、海信智能电视等内外合作伙伴建立了紧密的合作关系。展望未来，在日益完善的数据基础上，搜狐大数据中心将引领品牌精准的营销趋势，同时助力搜狐集团在各产品线上打造贴心的个性化体验。

# 搜狐新闻 资讯版



## 我们的功能

搜狐新闻(资讯版),你专属的兴趣阅读管家,助你发现生活点滴乐趣!

网罗热点资讯,海量内容任你挑选;  
打造个人标签,畅享专属兴趣推荐;  
说出你的态度,你的地盘由你做主!

搜狐新闻(资讯版),有趣有用新闻资讯全覆盖,让碎片化的时间更有价值!

The screenshot displays two main sections: '个性化推荐' (Personalized Recommendations) and '态度PK' (Attitude PK). The '个性化推荐' section shows a grid of news cards with titles like '美国总统奥巴马为希拉里站台投票,谁说没有“商场”?' and '印度普吉艺术节在东南亚刮起阵阵狂飙' along with small thumbnail images. The '态度PK' section features a large circular icon with five colored dots (red, blue, green, yellow, purple) labeled '赞赏' (Appreciation), '点赞' (Like), '评论' (Comment), '收藏' (Collection), and '分享' (Share). Below these are two more news cards with titles '李连杰的故乡要拍电影了,导演是陈可辛' and '印度连续两天下雨,人们期盼已久的雨季'.

搜狐

资讯版

搜狐新闻 (资讯版)

## 我们的特色

### ——极简风格

干净的界面风格和优雅的动态水波纹效果,给你清爽的视觉体验!独特的顶部水波纹波动,有效缓解眼部疲劳和阅读压力,独具匠心的交互设计,给你灵动舒适的阅读体验。

### ——热门竞猜

未来趋势大猜想,谁是真正推理王!搜狐新闻(资讯版)开启全民娱乐模式,每天多个话题供你选,大事小情全预测,边看边猜更刺激!热门竞猜功能4月上旬上线,敬请期待哟~

The screenshot shows the '极简风格' (Minimalist Style) and '热门竞猜' (Hot Guessing) features. The minimalist style is demonstrated by the clean layout and dynamic waterwave effect at the top. The hot guessing feature is shown on the right side of the screen, featuring a large image of a woman with a rose and a poll asking '美女与野兽上映10天中国票房会破亿吗?'. Below the poll are two large buttons: '中奖规则' (Prize Rules) and '砸破10亿' (Break 100 million). At the bottom, there's a progress bar showing '76.79% 5/6'.

关注公众号追踪赛事



安卓用户扫描下载



iOS 用户扫描下载



# RNN

Time is, time was, and time is  
past.

---

只争朝夕



LUCKY DATA  
搜狐图文匹配算法大赛

搜狐 | IEEE-China | 清华大学



[www.luckydata.org](http://www.luckydata.org)



# 数据分析能给科研方向带来启示吗？

文 | 维克多·潘克拉提奥斯 (Victor Pankratius), 贾斯汀·李 (justin Li), 迈克尔·高温洛克 (Michael Gowanlock), 大卫·M·布莱尔 (David M. Blair), 科迪·鲁德 (Cody Rude), 汤姆·赫林 (Tom Herring), 弗兰克·林德 (Frank Lind), 菲利普·埃里克森 (Philip J. Erickson), 科林·朗斯代尔 (Colin Lonsdale), 麻省理工学院 (Massachusetts Institute of Technology)

**最**近的技术进步使得天文学和地球科学的数据量越来越大。<sup>1</sup>科学家现在面临一个挑战，即如何从大量数据中总结出观点。<sup>2</sup>我们现在正处在一个根本性的变革过程之中，正从数据稀缺的时代迅速进入一个新时代，现在的数据量之大已经超过了我们从中提取信息的能力，科学界面临着一堵“分析墙”。

行星和基于空间的传感器网络正在产生连续的数据流，以监测我们的环境，描述不同的现象的特征，而且在许多情况下可以预测自然灾害。现在，计算机科学和智能系统亟需开发新的系统，从大数据集和不同类型的数据集（如光学、雷达和GPS时间序列）中提取观点。此外，来自不同设备的数据融合正在发现新的自然现象，并排除误报，特别是因为要建立正确的联系，对人类来说往往是不够直观的任务。

从信息处理的角度来看，信息处理过程可以分为几层。如图1所示，数字化传感器信号数据成为表示更复杂

分析的起始点。语义层在本质上提供了语义上的有效数据，即实际有效测量的数字。顶层的下一层通常是通过数据探索、分析和开发技术来引入语义（例如，“这个数据代表特征X”）。然而，仅仅靠这些是不足以促进科学进步的，科学家们需要对语用学的支持：某个特征被识别出来意味着什么？某个发现的意义是什么，如何将这一发现融入大的理论框架？它是证明了之前建立的模型和结果，还是与之相矛盾？研究人员如何有效地对概念和想法进行测试？基于这些问题的许多实施任务目前由少数研究人员负责，他们必须巧妙地制定工具和工作流程，以寻求适当的答案。

## 为什么科学家需要机器的支持来发现搜索

我们可以将科学发现过程视为一个搜索过程。该搜索空间不仅由大量的多样化的科学数据集本身定义，而且还要由人类在工作流程中处理数据时做出

的选择而定（例如，参数选择、确定顺序、使用哪种算法，方法或过滤器来处理流程中的某个阶段）。这里假设工作流程是与图灵机表现力相当的一系列处理步骤。

我们的案例研究显示，在进行工作流程的配置中的选择可能会大大影响我们获得新发现的能力。例如，如果一组处理步骤中突出了数据产品中的大规模现象，就会对未知的小规模现象的发现产生抑制。此外，一些自然现象实际上可能是罕见的、有悖直觉的，所以人类在配置潜在的大规模参数空间上需要机器的协助，来创建数据处理和发现的作业流程。

对发现过程实现自动化的价值还在于，现在的数据集的大小已经达到PB级，不久后即将达到EB级。与数据量的增长相比，在目前和未来可预期的网络带宽限制下，将如此大量的数据直接转移到科学家的台式计算机上变得越来越不切实际，这就需要用“移动计算力”的方法来取代“移动数据”，即科学



图1. 科学数据处理中的符号层。大数据科学家在上层语用层上对相关工具和技术的需求日益增长。

家要在保存数据的地方上传并执行自己的脚本和程序，而不是去下载数据副本到本地来进行处理。这些大型数据集的存储和分析需要强大的数据中心和超级计算机的支持，这些数据中心和超级计算机本身也很复杂，即使对于简单的探索性分析也需要相当的时间、精力、专业知识和成本。

因此，我们需要建立合适的智能系统。我们的方法中的关键原则是将人纳入到系统流程的循环中，并为他们提供高级的工具支持，而不是要把整个过程都推向自动化。比如我们假设一个迭代过程，其中由人类科学家（通过查看生成的数据产品或指标）定义搜索目标，并决定是否继续（使用修正后的模型或参数）进行另一次迭代。打个比方，我们为研究人员提供了性能强大的驾驶舱、导航系统和自动驾驶仪等支持，以增强他们的认知能力，但我们将飞行的控制权交给他们掌握。

## 基于模型的计算机辅助发现方法

图2所示为我们的方法的一般性

概念设置。在图的左侧，科学家通常从现实世界中的一个问题开始，然后利用领域内专用的规则，识别出兼容和不合规的模型特征、约束、参数设置和数据输入/输出，将问题空间化为易于理解的模型表示。这就产生了理论描述，通常由一组抽象信息组成，包括以模型、公式、算法或代码为表达形式的抽象信息，所有这些都是为了寻求一致的框架来解释实证观察。因为对立理论模型也可以解释相同的经验数据，所以科学家们必须不断尝试新的理论模型，来选出最合理的模型，每一个模型都有一组可能的参数范围排列。

这种迭代搜索过程的大部分可以实现自动化。为了实现这一点，我们引入了一个天气模型表示，通过比较模型间的共同点和不同点来描述这些理论模型。从技术上来说，这是一个元模型，它描述了哪些特定的变量应该由哪些部分组成（这即天气模型表示的一个实例）。该模型与软件工程开发的特征模型和软件产品线中的生成编程社区相似。<sup>3-4</sup>例如，图3中间的模型需要一个特殊形式来包含两个子模型1和2，但是

子模型2只能是A1或A2。实际上，图中的框是通过适当准备的代码来表示的，由理论模型验证引擎生成一个有效的代码实例，在实例中，所有的代码块都被聚合编在一起（比如可以使用条件编译、装饰器或所描述的其他技术来实现<sup>3</sup>）。然后就可以在各种数据集上执行这样的表示（例如“理论模型1”），以便在人类帮助下评估假设模型，看其假设是否能够恰当地描述现实。

天气模型基本上可以让科学家在不指定所有可能的实例的情况下，表达出所有可能的算法选择和参数范围，将自动搜索限制在发现空间中。以前的数据挖掘、基于案例的推理和知识发现，<sup>5-7</sup>可以作为我们的方法的组成部分。<sup>8</sup>我们可以在不同实例中交替使用这些策略，对算法、参数范围或一般工作流顺序的不同选择进行评估。

我们可以通过修剪天气模型定义的搜索空间来加快搜索速度。灵敏度分析可以利用假设检验的历史数据来评估某些参数的变化是否导致了有趣的发现。此功能让系统能够在发现搜索空间内的特定区域提高搜索精度。此外，

人类可以参与中间发现（如通过相似度聚类），并且可以要求将“找到更多与X类似的元素”摆上更高的优先级，来定义在搜索中使用的经验相似度函数。相关社区内已经讨论了这样的搜索空间修剪问题。<sup>9</sup>

## 实现和基础设施

图3举例说明了计算机辅助发现是如何弥合理论领域与实证领域之间的鸿沟的。我们基于天气模型的方法生成了假设现象及其物理特性（例如，呈不同形状的、电子密度被电离层暴增强的羽状物）的多种模型，用来搜索各种类似匹配设备的经验数据。科学家可以根绝研究结果以迭代的方式改进模型，并将搜索收敛至目标特质上。当找到适当的匹配后，科学家可以给出一系列的解释，从实体特征（如语法或语义抽象层面上的一组图像像素）到物理意义（物理模型中的实际抽象层面上的一个参数值）。

科学家们使用Jupyter (<http://jupyter.org>) 和Python语言实现的Web浏览器中的典型使用实例与基于云的分析界面进行交互。我们的计算机辅助发现服务由服务层提供，服务层由Python和C ++使用Python包装器实现的库和算法组成。数据存储和计算功能可以托管在可扩展的云环境中，如Amazon Web Services。来自实时传感器（如NASA图像和时间序列）的传感数据通过数据集成器层进入我们的环境中。由发送服务层根据分析界面的请求，使用云中的适当数据生成作业，并将其移动至节点，并按需以透明的方式创建副本。

## 案例研究与应用

我们在这里给出四个案例研究，包括初步结果以及在多学科领域的计算机辅助发现的未来前景（见图4）。

### 案例研究1：空间与时间层面上的地球结构

地球航天科学的研究涉及地球动态等离子体和空间电离层中的复杂、相互依赖的物理特征，这些特征通常是由太阳风暴和太阳的冠状质量喷射等太空现象引起的。这种现象可能会产生重要的社会影响，比如可能将意外的电流引入电网，导致大规模停电。

对暴时增强密度（SED）羽流的发现和研究，为地球物理科学中对工作流程和分析方法的敏感性提供了一个实例，如图4a所示。在某些地磁风暴的作用下，SED羽流可以呈现更加高度的结构化，形成电离舌（TOI）。<sup>10</sup> TOI和SED特征是磁层-电离层耦合中的关键系统过程，并将整个地球等离子体层的分布形成和空间气象响应远远扩展至地球空间。<sup>11</sup> SED羽流和TOI识别对所使用的特定数据处理工作流程和可视化选择都很敏感。图4a所示为这四个工作流的相关视图，从2014年2月27日（星期二）22:40的地磁暴雨时段起，北半球黄昏时段相同的底层GPS总电子含量（TEC）数据集。图中的三个地球图使用的数据处理元素是相同的（平均时间为20分钟），但使用不同的空间平均选择来划分底层的视线TEC数据。最左边的地球图采用3×3的空间像素划分，并进一步使用20个单元的平滑函数来最大限度地提高全局TEC的覆盖范围，但这会导致TEC结构（如SED）变得严重

模糊化。图中间的地球使用3×3度的空间像素划分，与原始SED发现中使用方式相同，<sup>11,12</sup>但不做进一步平滑处理，而且最多200个TEC单位的可视化处理在很大程度上盖住了SED羽流。右侧的地球图以3×3度划分，无进一步空间平滑处理，绘图最大比例尺为100个TEC单位，这是SED识别的最佳工作流程（白色箭头）。这个例子表明，如果事先不知道空间特征形态如何，可以将不同控制参数的同一工作流输出范围应用在相同的基础数据集中，以帮助特征识别。

科学模型的变量生成可以在云计算环境中以并行化的方式扩展。例如，VariantDBSCAN算法<sup>13</sup>利用基于密度空间聚类原理来识别TEC数据点表示中的地球电离层中的相关特征。用户可以定义两个输入参数的范围：点对象周围的搜索距离m，以及在给定点距离m范围内，给定集群中所需的点数量的最小值。然后，该算法会通过高效的检索处理不同具体参数的变量，达到聚类数据量的最大化，重新使用以前的执行变量的结果来降低记忆压力和计算成本，并通过有效的调度启发方式来优化变量执行顺序。在一个专门的16核平台上，这种方法的结果比TEC数据集上的连续基准实现方式快了22倍。<sup>13</sup>

### 案例研究2：地表变形与火山现象

Plate Boundary观测站 (<http://pbo.unavco.org>) 部署在地球表面的GPS传感器网络用于研究由不同物理机制引起的各种规模的地面运动和形变。地面变形的一个特别重要的驱动因素是火山活动，会给环境造成巨大的后果，难以预测或预测。在本例中，我们利



图2. 不同来源的经验数据集(上排)与模型变量(下排)的比较。通过模型变量和实证数据集的匹配,自动找到科学家定义的未知形态。这些匹配结果也将物理模型参数与实际监测特征联系在一起。

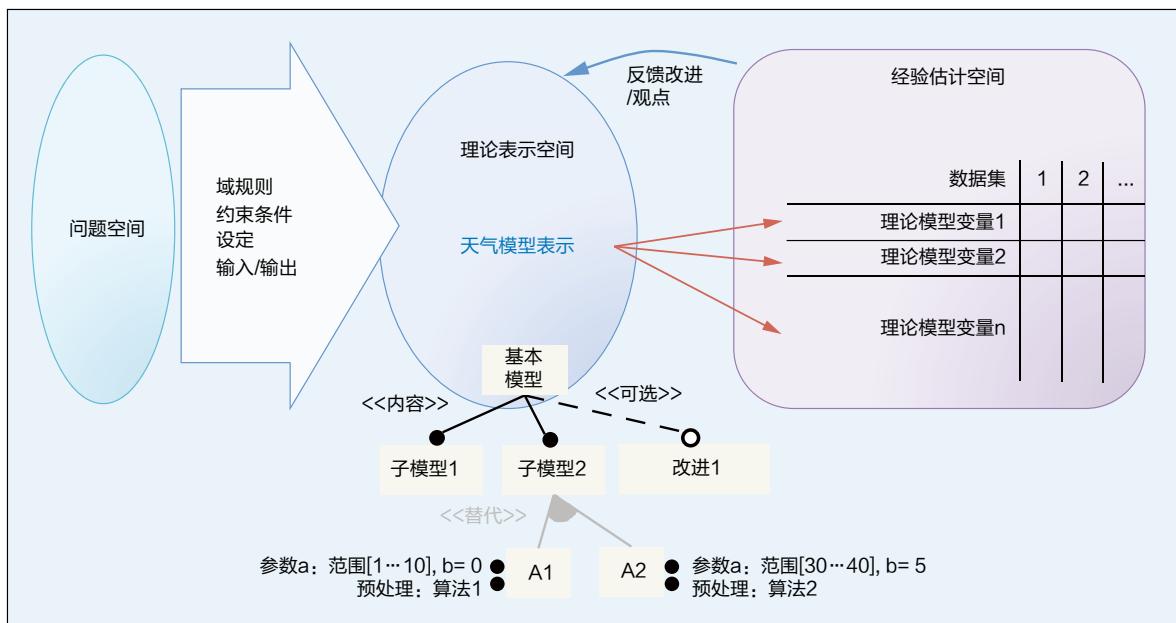


图3. 我们的计算机辅助发现中的模型表示。理论表示空间捕捉到科学家定义的对立模型中的模型变量,经验估计空间负责衡量每一种模型变量对现实世界数据的观测能力。

用计算机辅助发现方法和模块化的流程分析方法,来研究位于阿拉斯加的火山附近的GPS观测站的长期活动,以识别表明火山活动的瞬时事件,并使用阿拉斯加火山观测站的活动报告进行验证。由于GPS时间序列测量技术在

空间密度和时间精度上持续改进,测量产生的数据量越来越大,因此对发现流程进行自动化是很有必要的。

图4(b)所示为Akutan火山附近的AV06观测站的原始GPS数据,由北方,东方和垂直方向上的数据组成。

在本例中,我们所定义的流程分析首先进行过滤预处理,然后是主成分分析(PCA),来评估地面形变的可能模式。在分析流程开始运行之后,我们获得了图4(b)中的地面形变总量的PCA部分(蓝色箭头)。此外,我们可以把这

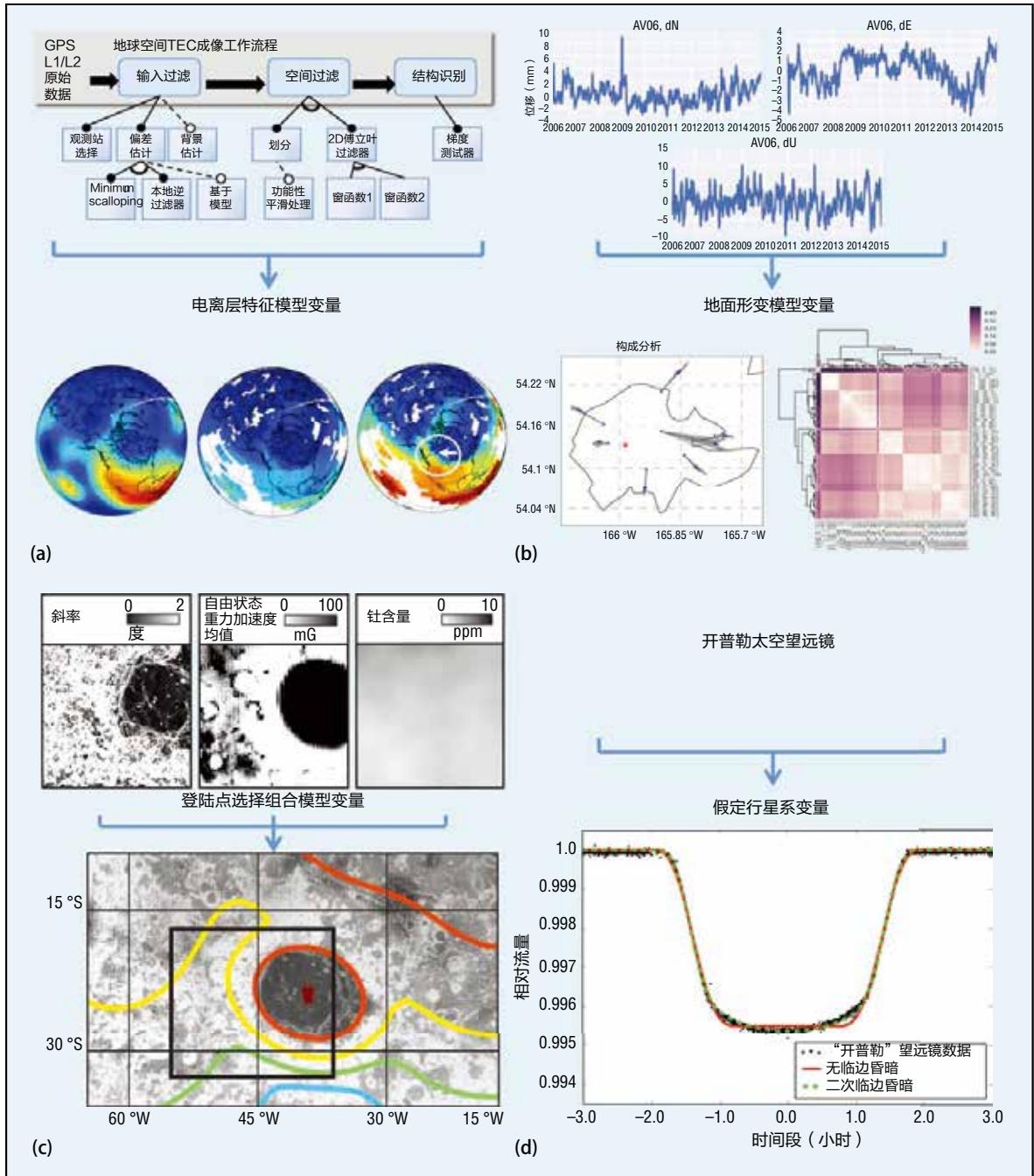


图4. 四个案例研究。(a) 电离层与空间气象, (b) 地球科学与火山活动, (c) 行星科学及人和机器人任务的着陆点选择以及 (d) 天文学和外行星探索。在(b)图中, 阿拉斯加附近的Akutan火山以红方块标注, 利用热图对形变模型进行比较。在(c)图中, 着陆点选择的发现过程按月球近地一面进行。

套流程用在不同的参数上(这里为了演示, 将其用在卡尔曼滤波器参数上), 以确定该分析流程对框架中的假设的潜在依赖性, 或者找到不同模型或不同

过滤器之间可能存在的变量, 这些变量可能会产生更合理的物理解释。我们可以用对流程中的类似迭代进行聚类的热图和树状图, 来表示导致类似的

物理模型的参数依赖关系, 其中x轴和y轴上的标签描述了特定的卡尔曼滤波器参数集, 也可以将其叠加至相同的地理图上, 产生一幅直观的物理视图。

这种对聚类的可视化可以帮助地球科学家比较不同类型的模型（例如，关于岩浆室的不同假设模型），并将模型结果与实证观察结果进行比较。因此，可以更有效地对地质构造成分进行假设检验。

### 案例研究3：用于人类和机器人任务的飞船着陆点选择

要想进行成功的行星际探索任务，需要为多种任务进行着陆点选址，例如勘探或高分辨率成像任务。这个选择过程需要对多来源的大量数据集进行同化和融合。过去，这个任务是由大量的科学家和工程师团队花费多年时间完成的。然而，随着数据量的增长和可用选择标准的数量的增加，手动选址将变得越来越不可行，因为参数空间扩大了，同时还要努力让选址保持客观和最优化。

### 案例研究4：天文和外行星搜索

现在，几个大型的天文学研究已向科学界提供开放的数据访问权限。例如，关于光的波长，斯隆数字巡天望远镜（Sloan Digital Sky Survey）的数据量大约为100TB，即将建成的大型综合巡天望远镜的可在一晚上产生超过15TB的数据。其他关于波长（如无线电、X射线和红外线）的数据量也将继续增加，这使得要对信息进行探索和分析变得很困难。计算机辅助发现模式可以让科学界享受大量天文数据的潜力。

为了详细说明这种潜力，这里太阳系（行星外）外的行星搜索为例。“开普勒”号探测器在3.5年中同时观察到

了超过10万颗恒星，在未来的任务中，将使用各种方法进一步扩大数据收集量。我们正在对我们的计算机辅助发现基础设施进行扩建，以适应和分析“开普勒”号发回的数据。对于一颗给定的星星，科学家们会分析它的光曲线，即由于一个假设的行星沿轨道运行而导致的亮度变化。因此，随着时间的推移，测量相对通量的时间序列预计会在每次行星位于观察者和星星之间时出现下降。按假定的轨道周期将这些时间序列进行折叠，就可以得到图4d这样的开普勒-13图。我们必须对每颗星重复这个搜索过程，重复过程中包含范围参数和假设。样本光曲线还显示了使用PYTRANSIT (<http://github.com/hpparvi / PyTransit>) 生成的各种匹配的行星运行模型，有没有临边昏暗现象的都有（图中分别由绿色虚线和红色实线表示<sup>15</sup>）。对各种模型进行比较，可以获得关于观测行星系统的物理性质的一些附加信息。



计算机辅助发现提供了扩展由大数据科学而生成的观点的途径，从而有可能超越人类的认知界限。我们提出的基于模型的方法，以一种新颖的方式将特定领域内的知识带入发现搜索过程中。例如，通过利用所需的物理模型和属性，来改进在大量的经验观察与现行理论之间建立联系的过程。案例研究表明，我们的概念适用性广泛，从地球研究和火山现象检测、航天器着陆点选择、天文学和外行星搜索都可以使用。由于前期取得的成功，我们认为计算机辅助发现将可能给未来科学研

究的方式带来革命性的改变。■

### 致谢

感谢国家自然科学基金会（项目号ACI-1442997）、NASA（项目号AISTNNX15AG84G）和亚马逊的支持。

### 参考文献

1. V. Pankratius and C. Mattmann, “Computing in Astronomy: To See the Unseen,” Computer, vol. 47, no. 9, 2014, pp. 23–25.
2. V. Pankratius et al., “The Rise of Computer-Aided Discovery in Geoscience,” Proc. AGU Fall Meeting, 2015; <http://agu.confex.com/agu/fm15/webprogram/Paper62710.html>.
3. K. Czarnecki and U. Eisenecker, Generative Programming: Methods, Tools, Applications, Addison Wesley, 2000.
4. V. Pankratius, Product Lines for Digital Information Products, KIT Scientific Publishing, 2007.
5. U. Fayyad and R. Uthurusamy, “Evolving Data into Mining Solutions for Insights,” Comm. ACM, vol. 45, no. 8, 2002, pp. 28–31.
6. J.B.H. Baker, J.M. Ruohoniemi, and N. Ramakrishnan, An Interactive Data Mining Framework for EarthCube, white paper, EarthCube, 2011.
7. J. Han, M. Kamber, and J. Pei, Data

- Mining: Concepts and Techniques, 3rd ed., Morgan Kaufmann, 2011.
8. K.L. Wagstaff et al., "Guiding Scientific Discovery with Explanations Using DEMUD," Proc. 27th Conf. Artificial Intelligence, 2013; [www.aaai.org/ocs/index.php/AAAI/AAAI13/paper/view/6171/7296](http://www.aaai.org/ocs/index.php/AAAI/AAAI13/paper/view/6171/7296).
  9. A. Zwinkau and V. Pankratius, "AutoTunium: An Evolutionary Tuner for General-Purpose Multicore Applications," Proc. 18th IEEE Int'l Conf. Parallel and Distributed Systems, 2012, pp. 392–399.
  10. J.C. Foster et al., "Multiradar Observations of the Polar Tongue of Ionization," J. Geophysical Research, vol. 110, no. A9, 2005; doi:10.1029/2004JA010928.
  11. B.M. Walsh et al., "Simultaneous Ground- and Space-Based Observations of the Plasmaspheric Plume and Reconnection," Science, vol. 343, no. 6175, 2014, pp. 1122–1125.
  12. J.C. Foster et al., "Ionospheric Signatures of Plasmaspheric Tails," Geophysical Research Letters, vol. 29, no. 13, 2002.
  13. M. Gowanlock, D. Blair, and V. Pankratius, "Exploiting Variant-Based Parallelism for Data Mining of Space Weather Phenomena," Proc. 30th IEEE Int'l Parallel and Distributed Processing Symp., 2016.
  14. D. Blair et al., "Improving Spacecraft Site Selection through Computer-Aided Discovery and Data Fusion," Proc. 47th Lunar and Planetary Science Conf., 2016; [www.hou.usra.edu/meetings/lpsc2016/pdf/1987.pdf](http://www.hou.usra.edu/meetings/lpsc2016/pdf/1987.pdf).
  15. K. Mandel and E. Agol, "Analytic Light Curves for Planetary Transit Searches," Astrophysics J.
- 维克多·潘克拉提奥斯**是麻省理工学院海斯塔克天文台天文学-地球科学信息学团队负责人,计算机辅助发现研究项目的PI。联系方式:vpankratius@haystack.mit.edu及[www.victorpankratius.com](http://victorpankratius.com)。
- 贾斯汀·李**是麻省理工学院海斯塔克天文台天文学-地球科学信息学团队博士后研究员。联系方式:jdli@mit.edu。
- 迈克尔·高温洛克**是麻省理工学院海斯塔克天文台天文学-地球科学信息学团队博士后研究员。联系方式:gowanloc@mit.edu。
- 大卫·M·布莱尔**是麻省理工学院海斯塔克天文台天文学-地球科学信息学团队博士后研究员。联系方式:dblair@haystack.mit.edu。
- 科迪·鲁德**是麻省理工学院海斯塔克天文台天文学-地球科学信息学团队博士后研究员。联系方式:cmrude@haystack.mit.edu。
- 汤姆·赫林**是麻省理工学院地球与行星科学教授。联系方式:tah@mit.edu。
- 弗兰克·林德**是麻省理工学院海斯塔克天文台研究工程师。联系方式:flind@haystack.mit.edu。
- 菲利普·埃里克森**是麻省理工学院海斯塔克天文台首席研究科学家,负责领导大气科学团队。联系方式:pje@haystack.mit.edu。
- 科林·朗斯代尔**是麻省理工学院海斯塔克天文台主任。联系方式:cjl@haystack.mit.edu。

微信名: 计算人

微信号: jisuanren





# 自动驾驶的社交硬伤

文 | 巴里·布朗（Barry Brown），斯德哥尔摩大学（Stockholm University）

译 | 刘大明

所有汽车实现完全自动驾驶的那天，人们的要求就不仅仅是安全和高效了，自动驾驶汽车还必须能够理解人类驾驶员，并与之以一种自然的方式进行交流。



前，像特斯拉S和沃尔沃XC90这样的汽车都配备了高级自动驾驶功能，成千上万辆这样的汽车正行驶在全球的道路上，并且这个数字每年都在增长。此外，包括特斯拉、德尔福和谷歌在内的公司正在美国的道路上测试全自动驾驶汽车，测试里程已达数百万英里。

自动驾驶技术在全球范围内的道路测试正在如火如荼地进行，然而测试结果只在公司手里，很少对公众披露。但是，自动驾驶技术中的缺陷偶尔会被行车记录仪和乘客拍摄的手机视频记录下来，并上传到视频社交网站YouTube上，引发媒体的讨论和争议。

举个例子，2016年12月份，Uber公司在

旧金山投放一队配有人类监管员的自动驾驶汽车的当天，一名监管员的行车记录仪拍到一辆自动驾驶汽车不仅闯了红灯，还差点撞上一位行人。这段视频上传到YouTube上几小时之后，美国加利福尼亚州政府强令Uber公司在获得适当许可之前立即停止测试。

在瑞典斯德哥尔摩大学，我们开发出一种新方法，即利用公共视频，帮助大众快速了解自动驾驶。这些视频揭示了自动驾驶汽车在适应人类社会时面临的一些挑战。

## 重新利用网络视频

YouTube是世界上最大的第三方

视频存储平台。我们做的第一个研究，就是用一系列关键词，在YouTube上搜索与半自动驾驶汽车和全自动驾驶汽车有关的视频。剔除了大部分推广视频后，我们将注意力集中在评论和行驶日志上——行驶日志通常是长时间的、无声的驾驶画面，以及系统行为的实时情况。

我们收集了分别在美国、英国、德国、法国、瑞典、香港、冰岛和加拿大录制的93段视频，总时长达到10.5小时。平均下来，每段视频将近7分钟，其中7段视频的长度超过30分钟。大多数视频跟特斯拉的驾驶辅助系统——Autopilot的讲解有关，有3段视频展示了在沃尔沃XC90和本田思域上搭载的类似的自动驾驶系统。还有总计11分钟的9段视频记录了谷歌自动驾驶汽车的行驶情形。另外，在西南偏南音乐节(South by Southwest)上，谷歌自动驾驶汽车项目也展示了一些有趣的事情。



**图1.** 变道与不变道：当特斯拉汽车司机想为一辆要上高速的轻型货车腾出空间并发出左变道的信号后，特斯拉Autopilot自动驾驶系统表现出的“粗鲁”行为。（a）后方左侧车道的银色汽车开始减速，方便特斯拉汽车向左变道（左图），因为此时，一辆货车从特斯拉汽车前方进入右车道（右图）。（b）银色汽车的司机认为特斯拉汽车的司机拒绝了他的好意，因此开始加速缩短与前车的距离（左图），但此时Autopilot自动驾驶系统依旧执行了左变道的命令，挡住了银色汽车司机的路（右图）（来源：Garth Woodworth，特斯拉Autopilot自动驾驶系统的变道行为，2015.11.12，[www.youtube.com/watch?v=el4OdwtgzNk](http://www.youtube.com/watch?v=el4OdwtgzNk)）

## 驾驶与社交

我们借鉴成熟的语言学和社会学研究方法，通过分析枯燥单调的人类驾驶视频，深入理解驾驶行为以及自动驾驶汽车带来的潜在问题。研究自动驾驶汽车如何与人类司机互动或许能够带来新的启示。

来看看特斯拉的Autopilot系统，它主要有三大功能：

- › 自适应巡航控制功能：根据与前车的距离加速或减速；

- › 自动转向功能：根据路面标识，或在没有路面标识的情况下根据前方的车辆情况，保持车辆位于车道内；
- › 变道辅助功能：在保证安全的前提下，驾驶员发出变道指令后，车辆移动到邻近车道。

虽然说这些功能还远算不上完全自动驾驶，但特斯拉自动驾驶汽车已经能够在相当长的时间内执行自动驾驶。

在我们收集的YouTube视频中，Autopilot自动驾驶系统在大部分时间

里都在正常行驶，没有出错。然而，受限于自身简单的操作机制，Autopilot自动驾驶系统有时会对人类司机的行为作出误判。通常情况下，人类可以根据汽车的速度、方向等因素来判断司机的行驶意图，以及他们的情绪或性格，是咄咄逼人还是犹豫不决。但是Autopilot自动驾驶系统缺乏这项本领。

图1描述的就是一个相当普遍的情况。在一条有隔离带的双向四车道的高速公路上，特斯拉司机处于交通拥挤的右侧车道。在他的前方，他看到右侧连接道路上出现一辆轻型货车即将要进



图 2. 对自动驾驶技术产生困惑的人类司机。一名司机驾驶红色汽车横穿马路，挤到两辆谷歌自动驾驶汽车中间狭窄的间隔里，进入左转车道。红色汽车司机或许错误地把自动驾驶汽车留出的安全跟随距离当成可以加塞的空间。（来源：西南偏南，2016 年 12 月，谷歌自动驾驶汽车项目西南偏南音乐节互动，[www.youtube.com/watch?v=Uj-rK8V-rik](http://www.youtube.com/watch?v=Uj-rK8V-rik)。）

入高速公路。特斯拉司机打出左转信号，表明他希望移动到左侧车道为卡车腾出空间。后方左侧车道的银色汽车司机看到左转信号后停止了加速，为特斯拉汽车变道留出空间。但是，Autopilot 自动驾驶系统没能理解这个礼貌的“姿态”，并继续保持在右侧车道上。银色汽车司机认定特斯拉司机拒绝了他的好意，重新开始加速。但是，Autopilot 自动驾驶系统仍然认为此时变道是安全的，并驶入左侧车道。

对另一辆自动驾驶汽车来说，特斯拉汽车的机器协调行为或许不会产生“困扰”，但银色汽车司机将其理解为双重“粗鲁”也是可以理解的：特斯拉司机首先对他的好意嗤之以鼻，紧接着又在最后一秒变道，跑到他前面。特斯拉司机在视频中承认了这个粗鲁的行为：“嗯，我们加塞到那个人前面，我

看出他的脸色并不好看。”

这个例子表明，即使自动驾驶汽车能做到安全可靠，它们却无法理解“道路社交”。虽然人类驾驶员之间并不总是“合拍”，但他们跟算法代理的区别在于，在大多数情况下，人类驾驶员能够察觉到人际互动中的微妙线索。

## 让红色汽车加塞

除了不能理解人类驾驶员在道路上的行为，自动驾驶系统也无法判断自身行为传递给人类驾驶员的信息。谷歌自动驾驶汽车项目首席工程师克里斯·厄姆森（Chris Urmson）在西南偏南音乐节上谈到了这个问题（参见“自动驾驶汽车之间”）。

在这个例子中，一名人类司机横向驶入一条四车道高速公路，并进入左转

车道，前方就是一个十字路口，问题是把他他的红色汽车强行加塞到两辆相邻的自动驾驶汽车中间，并且这两辆汽车之间的间隔很窄。厄姆森滑稽地模仿红色汽车驾驶员的行为：“啊，没错，我有时真的理解不了一些人在路上的举动，就像这个司机还能把车开到两辆自动驾驶汽车中间。想啥呢？”逗得观众哈哈大笑。

不过，尽管红色汽车司机做出了或许并非最佳的判断，他的行为也是能够被理解的。司机都清楚与前车保持安全距离的重要性，这一原则理所当然也被自动驾驶系统采用。在这个案例中，第二辆自动驾驶汽车缓慢靠近十字路口，同时在自己跟第一辆自动驾驶汽车中间留出了一个空档。然而，跟特斯拉自动驾驶汽车的那个例子一样，道路上的间隔有时不仅仅是安全跟随距离，也

可能被其他司机理解成可以加塞的空档。这里的人类司机或许就把自动驾驶汽车之间留出的间隔当成可以加塞的空档了。

但也有可能，红色汽车司机就是想强行加塞。人类司机的驾驶行为并不总是明智的、合法的，尤其是在容易导致愤怒、困感情绪的情形下，比如说行驶在施工区域或遇到交通堵塞。在拥堵的道路上，有些司机会钻进他们看到的每一个空档。这这种情形下，靠近十字路口的自动驾驶汽车的应对方式算得上中规中矩：减速慢行，让红色汽车加塞。

## 自动驾驶汽车的“恐怖谷”

人类司机同样会对其他驾驶员的不作为做出自己的理解，这也可能为自动驾驶系统带来新的问题，正如图3所示。

在这个案例中，一辆谷歌自动驾驶汽车比前方十字路口右侧的一辆白色汽车先到达十字路口。对十字路口的研究表明，在交叉路口减慢行驶速度非常重要，因为这会让其他司机明白你此刻正全神贯注，随时会穿过路口。尽管谷歌自动驾驶汽车在往前靠，但它的动作显然不足以表明它迫切地想要先通过路口。十字路口右侧的白色汽车司机认定谷歌自动驾驶汽车司机在犹豫，并提前三进入路口。这导致谷歌自动驾驶汽车跟一个新手一样紧急制动刹车，后面跟着的车辆也紧急停下来避免追尾。



图3. 十字路口的遭遇：自动驾驶技术带来的侵略性行为。（a）一辆谷歌自动驾驶汽车比路口右边的白色汽车提前到达路口。（b）自动驾驶汽车表面上的犹豫鼓励了白色汽车司机抢先通过路口，造成谷歌自动驾驶汽车紧急刹车，差点跟后面的一辆车撞上。（来源：[www.youtube.com/watch?v=FbSQm3YaAzA](http://www.youtube.com/watch?v=FbSQm3YaAzA)）

同时，视频还显示，司机们对动作缓慢的谷歌自动驾驶汽车十分不满，常常紧贴在汽车后面“逼迫”谷歌自动驾驶汽车快速通过路口。

就是发布声明的预行为。由于驾驶环境的复杂性，汽车司机采取很多预行为来避免潜在的碰撞事故，以及在交通拥挤时表明立场。

通过编程让自动驾驶汽车执行和识别通常很细微的预行为是非常困难的，这让自动驾驶汽车的设计师陷入一个窘境，类似于计算机图形和机器人领域的“恐怖谷”问题。“恐怖谷”问题是当计算机生成的人类形象和人形机器人不再卡通化，而是越来越真实的时候，它们身上残余的缺点也变得越来越明显，甚至让人引起不适或反感。同

## 自动驾驶汽车的“恐怖谷”

对人类社交的研究表明，我们的许多行为，包括抱怨、邀请等，都是先通过“预行为”来传达我们的意图，从而让他人为我们即将要做的事情做好准备。举例来说，把手放在立式麦克风

样，当自动驾驶汽车对驾驶越来越熟练，但行为仍不能跟人类司机一样，会让人产生愤怒或挫折情绪。举个例子，减速避免碰撞在统计学上是安全的，但并非总是最佳的应对措施，因为它会鼓励人类司机在遭遇自动驾驶汽车时的加塞或超车行为，而这只会增加事故的风险。

在汽车发展早期，汽车前方必须有一个人手持红旗，警示路上行人。那么在自动驾驶汽车能驾驭“社交道路”之前，是否应该给它们配备一个系统用于提醒人类司机，或在不明确的情形下（比如十字路口）传达意图？由此引发的问题是，自动驾驶汽车会对人类司机产生什么影响？人类司机是会更加谨

慎和耐心，还是会占“有点呆”的自动驾驶系统的便宜？



写这篇文章的目的不是批评当今的自动驾驶汽车，它们还处于发展的初期阶段。我想说的是，驾驶不仅仅是机械操作，同时也是一项复杂的社会活动。在所有汽车都实现完全的自动驾驶之前，自动驾驶汽车不但要追求安全高效，还必须能够理解人类驾驶员，并以一种自然的方式与人类司机交流。只要道路上大部分汽车还是由人操作，自动驾驶汽车的设计者就必须考虑他们的选择会对其他司机以及自己汽车上的乘客产生什么影响。否则，“社交道路”

就会有点颠簸。■

## 参考文献

1. J.F. Rodriguez, “Video Appears to Show Uber Self-Driving Car Running Red Light in SF,” San Francisco Examiner, 14 Dec. 2016; [www.sfexaminer.com/uber-self-driving-vehicle-appears-launch-red-light-first-day-sf](http://www.sfexaminer.com/uber-self-driving-vehicle-appears-launch-red-light-first-day-sf).
2. B. Brown and E. Laurier, “The Trouble with Autopilots: Assisted and Autonomous Driving on the Social Road,” to be published in Proc. 2017 CHI Conf. Human Factors in Computing Systems (CHI 17), 2017.



听听Diomidis Spinellis说什么

播客：“交易工具”

[www.computer.org/toolsofthetrade](http://www.computer.org/toolsofthetrade)

Software

IEEE computer society

巴里·布朗是斯德哥尔摩大学移动生活中心研发总监、人机交互教授。联系方式：[barry@dsv.su.se](mailto:barry@dsv.su.se)。

微信名：计算机  
微信号：jisuanren



# 构想个人未来

文 | 艾丽达·德罗特 (Alida Draudt)，美国第一资本金融公司及战略前瞻伙伴有限公司 (Capital One and Strategic Foresight Partners)  
茱莉亚·罗丝·韦斯特 (Julia Rose West)，战略前瞻伙伴有限公司 (Strategic Foresight Partners)

译 | 汪梅子

两位未来学家通过教授战略前瞻方法，帮助人们更具批判性地思考未来。

**我**们二人的共同兴趣之一是向大众普及“未来式思考”。人们通常都会将自己的过去投射到未来当中，因为他们并不知道还能以何种不同方式想象未来。然而，假使我们能更好地探索未来呢？如果人人都能以更具批判性的方式思考可能拥有怎样的未来，而不仅仅局限于掌握这种方法的人，想想看，这将会给我们的社会带来巨大裨益。

怀着这一目标，我们编写了一本互动式练习册：《前瞻思考助你探索个人未来——挑战预期，发掘偏好》（创造空间独立出版平台，2016年）。其中介绍了前瞻思维方式，将所需的工具与流程从艰巨复杂变得简单易懂。

我们的流程较为关注的是互动，而非研究。我们在这些流程的核心注入多种战略前瞻工具。有意识地思考可能的未来不仅有趣，也能帮助人们更好地理解当下的生活。这使他们得以发现当前状态与预期状态之间的差距，从而采取行动弥合差距，实现期待的未来。

那么，创造实实在在的未来是怎样的呢？如何才能成功？首先，必须搞清可能的未来要建立在哪些明确的价值基础之上。在整个前瞻思考过程中，这些价值都将发挥指导原则的作用。

明确价值之后，便需要理顺我们对未来的固有假设，这样才能超越它们，探寻其他可能性。

一旦确认自己的假设，便可采用各种方式探索可能的未来。其中一个方法是“另类未来”法，它可以帮助我们拓展思考，发现以前从未想象或考虑过的可能性。为了简化这一流程，我们为四种未来的可能性创建了类似“疯狂填词”(Mad Libs)游戏的句子模板：

- › 增长。生活中的一些要素有所增加。在美国，特别是在商业领域，这常常是我们能想到的唯一可能性。
- › 局限。生活中的一些要素被有意地管理起来。例如，政府法规或工作场所的规定。
- › 坍塌。熟悉的生活瓦解，从而影响

生活的方方面面。很多人可能会想到阿勒颇和整个叙利亚境内的当下状况。

- › 变革。发生巨变，我们的思维方式也随之改变。例如，很多人很难想象没有智能手机的生活，但以前的生活并不是这样的。

## 所

有人都应该体验一下为自己构想未来愿景的过程。我们帮助人们接触这一流程，推广前瞻思维方式，从而提高大家的集体能力，得以创建更为有趣和多样的未来。■

**艾丽达·德罗特**是一位未来学家和经验丰富的设计战略家，就职于美国第一资本金融公司，同时也是战略前瞻伙伴有限公司的共同创建者。联系方式：alida@wtforesight.com。

**茱莉亚·罗丝·韦斯特**是一位未来学家，同时也是战略前瞻伙伴有限公司的共同创建者。联系方式：julia@wtforesight.com。

### 译者简介：

**汪梅子**，英法双语自由职业译者，毕业于北京大学和巴黎高等翻译学校，科学松鼠会成员。翻译常涉领域包括国际组织会议文件、科普、科幻和奇幻等类型文学、摄影、艺术、电影等。联系方式：corbelleprune@gmail.com。



## 关于搜狐

搜狐是中国领先的互联网媒体、搜索、在线游戏集团，拥有搜狐公司(NASDAQ:SOHU)和畅游公司(NASDAQ:CYOU)两家美国纳斯达克上市公司，是中文世界最强劲的互联网品牌之一。“搜狐”在中国是家喻户晓的名字，也是2008北京奥运会互联网内容服务赞助商。搜狐为中国近6亿的互联网和移动互联网用户提供全面网络服务。



## 搜狐四大业务平台

媒体：搜狐网、搜狐新闻客户端、手机搜狐、搜狐微门户、搜狐焦点、搜狐汽车

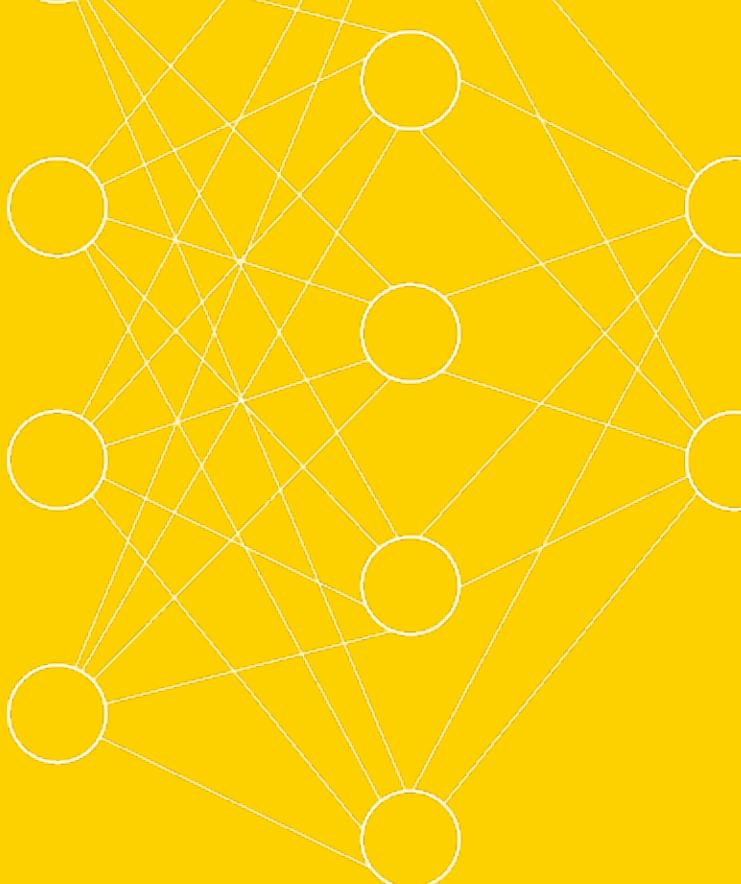
视频：搜狐视频、搜狐视频客户端、搜狐娱乐

搜索：搜狗搜索、搜狗输入法、搜狗高速浏览器、搜狗地图、搜狗号码通

游戏：畅游、17173、第七大道



# LUCKY DATA 2017 搜狐 图文匹配 算法大赛



随着信息技术的不断发展，互联网新闻资讯的生产方式正在发生深刻的变革，机器+人工的编辑方式已成为主流生产模式。搜狐作为中国新媒体界的鼻祖，也在不断探索，致力于机器学习性能的最大发挥。

新闻自动推荐配图是提升机器在新闻自动化生产中效力的重要课题，涉及自然语言处理、图形图像识别等多领域的技术，具有一定的挑战性。搜狐大数据中心联合 IEEE-China 和清华大学 keg 实验室共同举办本次新闻配图比赛，搜狐提供百万级的真实数据，邀请全球 IT 高手一同解决图文匹配这一算法课题。

## 奖励

一等奖 1 名：10000 美金

二等奖 2 名：每名 2000 美金

三等奖 3 名：每名 1000 美金

## 重要日期

3月13日：比赛正式开始，开放比赛队伍注册，同步发布训练集和验证集。

5月12日：初赛截止，公布进入复赛名单。

5月31日：复赛截止，公布进入决赛名单。

6月6日：决赛暨颁奖典礼（北京·搜狐媒体大厦）。

注意：以上日期节点均为 UTC 标准时间 00:00:00



# 保持联系。

无论你在哪里，都能紧随IEEE计算机协会的脚步。

在Twitter、Facebook、Linkedin和YouTube上关注我们。



@ComputerSociety, @ComputingNow



[facebook.com/IEEEComputerSociety](https://facebook.com/IEEEComputerSociety)  
[facebook.com/ComputingNow](https://facebook.com/ComputingNow)



IEEE Computer Society, Computing Now



[youtube.com/ieeecomputersociety](https://youtube.com/ieeecomputersociety)