

人类视听觉认知机理研究是认知科学的重要组成部 分,而人类视听觉信息的机器理解与计算一直是人工智能 领域的主要研究内容,在国民经济、社会发展和国家安全 等领域中扮演着十分重要的角色。

2008年,国家自然科学基金委员会(以下简称自然 科学基金委)设立了重大研究计划"视听觉信息的认知计 算",旨在发展和构建新的认知计算模型与算法,为提高计 算机对非结构感知信息与海量异构信息的理解能力和计 算效率提供科学支撑。

主编/肖洁 编辑/张双虎 校对/何工劳 Tel:(010)62580618 E-mail:news@stimes.cn

实施 10 年来,该重大研究计划取得了丰硕成果。本 期基金版将总结该重大研究计划的经验,展示其取得的 成绩。

"视听觉信息的认知计算"重大研究计划

迎接人工智能新时代

■本报见习记者 程唯珈

当前,人们对于人工智能也许并不陌生,因 为从几年前开始,相关新闻就时常见诸报端:AlphaGo 在围棋比赛中战胜人类冠军李世石和柯 洁、无人驾驶汽车获发测试牌照即将上路、越来 越多高校成立人工智能学院和研究院……

简言之,人工智能就是让机器能像人那样理 解、思考和学习,即用计算机模拟人的智能。它涵 盖认知与推理(包含各种物理和社会常识)、计算 机视觉、自然语言理解与交流(包含听觉)、机器学 习等广泛的学科领域。因此可以说,视听觉信息 的认知计算是人工智能重要研究内容,理解人类 视听觉认知并建立可计算视听觉认知模型对人 工智能的核心算法具有重大的启示意义。

来自上世纪的设想

但如果将时间回拨至二十年前,公众对人工 智能的认识既没有如此深刻,也不像如今这样对 它抱有那么大的期待。甚至在上世纪九十年代 初,面对全球范围内现代PC的出现和普及,人工 智能由于发展不及预期导致遇到资金困难等难 题,经历了一场寒冬。不过,这个在当时看似"无人 问津"的领域引起了中国工程院院士、西安交通 大学教授郑南宁的注意。

"为什么人工智能会遭遇寒冬?我们面临的 挑战是什么?"上世纪90年代初,中国工程院院 士、西安交通大学教授郑南宁对这个问题进行了 深入思考。

"1999年,'视听觉信息的认知计算'重大研 究计划立项的前期思考和顶层设计工作就开始 了。"该重大研究计划指导专家组组长郑南宁告 诉《中国科学报》,但由于科学家们最初并没有形 成基本的共识,因此历经9年艰辛的研究积累和 多次探讨,在自然科学基金委和专家们的共同努 力下,才终于在2008年正式启动了这一重大研 究计划,"这是我国在人工智能基础研究领域发 展的里程碑之一,标志着中国人工智能科学研究



有智能驾驶功能的汽车

'国家队'的正式组建"。

"本重大研究计划在立项伊始,人工智能技 术还未形成当今席卷全球范围的研究热潮,足以 体现出自然科学基金委与相关专家的学术洞察力 和战略前瞻眼光。"回首往事,郑南宁欣慰地说。通 过这一重大研究计划的资助,我国在人工智能领 域从理论、方法、技术到应用都得到了蓬勃发展。

让基础研究走出实验室

立项之初, 为确保国家安全与公共安全、推 动信息服务及相关产业发展以及提高国民生活 和健康水平,研究人员确定了"视听觉信息的认知 计算"重大研究计划的目标,即研究并构建新的计 算模型与计算方法,提高计算机对非结构化视听 觉感知信息的理解能力和海量异构信息的处理 效率,克服图像、语音和文本(语言)信息处理所面

如何才能实现这一目标?在郑南宁看来,这

需要从人类的视听觉认知机理出 发。"围绕认知过程的'表达'与 '计算'这一基本科学问题,我们 重点开展了'感知特征的提取、表 达与整合''感知数据的机器学习 与理解'和'多模态信息协同计 算'三个核心科学问题的研究。

"计算机对感知信息不能有 效处理,根本原因则是不能对真 实场景的基本特征进行可靠提 取,缺乏对真实场景基本特征的 一般表达方式以及对不同模态下 信息特征有效整合的理论。"他指 出,虽然机器学习和人工神经网 络等方法近年来在高维数据可视 化、特征提取、数据聚类与特征子 空间分析等方面取得了重要进 展,但非结构化数据本质维数的

确定仍然是一个公开的难题;现有的信息处理方 法主要是针对单模态的,对多模态信息的处理还 基本上停留在将各种单模态信息的处理结果在 决策层面上进行融合。"正是因为这些基本问题 没有得到解决,计算机才只能处理比较理想状态 下的一些简单问题,很难处理现实世界中的复杂 问题。

通过对上述三个核心科学问题的研究,十年 来,该重大研究计划在认知机理和模型、视听觉 信息处理、自然语言(汉语)理解等方面取得了一 系列标志性成果。

例如在视觉认知机理方面,研究人员提出了 知觉物体的拓扑学定义和注意瞬脱的拓扑学解 释等基础理论和模型;在视听觉信息处理与计算 方面,建立了视觉注意力统计学习计算模型和显 著性目标检测新理论;在汉语自然语言理解方 面,创建了一种新的语义计算理论框架,成功研 发了一系列面向公共安全的语言交互系统。

据不完全统计,截至2018年9月,该重大研

究计划共发表学术论文 2255 篇, 申请国家发明 专利532项。其中,在认知和信息科学相关领域 的国际权威期刊上发表论文 163 篇,包括在影响 因子 5.0 以上期刊发表论文 50 余篇。

2018 中国智能车未来挑战赛开幕式现场

"特别值得一提的是,为了进一步推动研究 工作走出实验室、产生原创性重大成果,本重大 研究计划创建了两个比赛平台,即'中国智能车未 来挑战赛'和'中国脑—机接口比赛',并组织了 10届'中国智能车未来挑战赛'和2届'中国脑— 机接口比赛'。"郑南宁介绍道,通过在真实的物理 环境中验证理论成果,解决实际环境中复杂认知 和智能行为决策等问题,改变了简单的论文汇总 或实验室成果演示的传统模式,促进了应用基础 研究与物理可实现系统的有机结合。

而这也为我国培养和造就了一大批计算机 视觉、脑机接口、无人驾驶技术和人工智能等方 面的优秀中青年人才。郑南宁说:"尤其是'中国智 能车未来挑战赛',历经10年的摸索和实践,业已 成为中国无人车研发的重要品牌,培养了一大批 本领域优秀的中青年科技骨干,是当之无愧的中 国无人车研发'黄埔军校'。

加强学科交叉共融

人类视听觉认知机理研究是认知科学的重要 组成部分, 而人类视听觉信息的机器理解与计算 一直是人工智能领域主的要研究内容。可以说,自 立项之日起,"视听觉信息的认知计算"重大研究 计划就带有明显的学科交叉属性,比如信息科学、 神经科学、认知心理学、数理科学等学科的交 一而这也是郑南宁十年来感受颇深的地方。

"我们所谈的科学问题普遍性越强,它所牵涉 的交叉性就越强。要解决基础科学问题,必须走学 科交叉这条路。"郑南宁介绍说,该重大研究计划 以"认知计算和脑机接口"和"无人驾驶与智能测 试"两方面为切人点,共部署了5个集成项目,根 据承担集成项目的10个项目组的不完全统计,论

文分别发表在信息科学、认知科学、心理学、神经 科学、物理学、生命科学等领域的国际核心学术刊 物上,"充分体现了多学科交叉的特点和我们研究 工作的学术水平,另外,自然科学基金委信息学部 在重大研究计划实施的管理机制创新方面,也为 不同领域专家的合作立项创造了宽松的环境"。

项目组供图

比如,视觉注意机制是生物视觉的一个重要 特性,早期的研究主要集中在心理学、认知科学 和神经生理学等领域,上世纪80年代后,这一课 题引起了计算机视觉、人工智能等领域学者的重 视。该重大研究计划针对自主式车辆视觉导航的 需要,多个课题组对此问题开展了深入的研究,在 计算视觉与生物视觉结合方面开展了深入的多 学科交叉,取得一批重要学术成果:清华大学在 人机驾驶模型融合研究方面,开展了驾驶员感知 信息处理与融合的认知机制研究;吉林大学模拟 真实驾驶员对预期轨迹信息的认知处理机理,研 究无人驾驶车辆的局部路径规划问题;西安交通 大学研究了视觉注意机制建模问题,成为视觉注 意力检测的代表性工作。

不过,在郑南宁看来,研究者的学科交叉还 有待进一步深入。

"一方面,学科交叉取决于学者的热情,这是 根本因素。同时,也要有自上而下的组织。"不过 他指出,从实际情况来看,这两方面都有所不足, "研究者应该更多地去主动思考科学问题背后的 学科交叉需求,对于研究中存在的一些急功近利, 也需要去改变"

实际上,不仅解决科学问题需要学科交叉, 应对人工智能所带来的深刻的社会问题,也同样 需要学科交叉。"因为人工智能模糊了物理现实、 数据和个人的界限,延伸出复杂的伦理、法律和安 全问题。人工智能的逐渐普及和深度应用一定会 给人们带来心理的影响,进而产生社会人文风险, 这已不是传统的工程安全方法能够解决的问题 了。因此在这些领域,人文社会学科和哲学学科 将会大有作为。"郑南宁说。

本报见习记者 智 能 程 覞 观 ノノ 路 耳 方

仿佛被一双无形 的手控制, 车里的方 向盘可自动左右旋 转。眼看前方有行人 或车辆通过,车子"自 觉"地停了下来……

近年来, 我国智

能车发展势头迅猛, 产业规模与市场不 断扩大。在矿山、物 流基地、码头等简易 固定路线场景中,智 能车已得到一定程度 的应用。 然而,在城市道

路和城际高速公路等 交通流量密集、行车 环境复杂的环境下, 智能车尚未具备人工 驾驶员具有的与其他 车辆进行交互和协同 的智能。

在国家自然科 学基金委员会(以下 简称自然科学基金 委)重大研究计划 "视听觉信息的认知 计算"支持下,天津军 事交通学院教授徐友 春带领团队突破了基 于视听觉信息的多车 交互协同驾驶技术 难题,为完成真实道 路开放环境下的智 能驾驶提供了必要 的技术支撑。

徐友春告诉《中国科学报》,团队 围绕实现智能车融人正常交通流自主 驾驶问题开展技术攻关,并研制了具 备智能驾驶功能的三台智能车,完成 了京津城际高速公路智能驾驶实验。

"在理论方面,我们借鉴人眼视觉 的选择性注意机制,提出变尺度栅格 法融合多源传感器信息; 揭示车辆行 驶对道路资源(路权)的竞争和共享本 质;建立多车交互策略和协同驾驶机 制。"他说。

人眼视觉选择性注意机制在计算 机视听觉认知中的实现研究是此次研 究的亮点之一。团队实现了准确、快速 识别与跟踪复杂场景中的地面标线、 动静态障碍物等,并提出多维变尺度 栅格图方法实现多源异构信息的融合 与道路环境建模。

亮点二是基于路权概念构建了 智能车交互机制与决策系统。研究人 员通过揭示车辆行驶的本质是对道 路资源的竞争和共享,提出了路权 (定义为满足车辆安全行驶条件所需 要的道路空间和时间)的概念,并基 于此实现了智能车交互协同驾驶智 能决策。

此外,团队还搭建了一个标准化、 可伸缩裁剪的线控硬件和软件平台。 在智能车改造中,实现了标准化的传 感器、车辆平台和通信接口,具有模块 化、可裁剪、可移植和可重用特点。采 用抽象分层的软件架构, 保证了程序 调试与传感器更换的便捷、快速;转 向、制动、油门等均由专用微处理器和 硬件电路实现线控,可独立于上位机 实现自动巡航控制 ACC 功能,工程化 水平较高。

徐友春表示,自然科学基金委连 续主办"中国智能车未来挑战赛",以 智能车为平台,验证技术发展水平。 在完成真实道路交通环境下的智能 驾驶科学任务过程中,项目组从中提 炼科学问题,有针对性地展开科研攻 关,取得了关键技术的突破;并且,在 与国内外同行的比拼与交流中,不断 促进研究取得进展。

同时,自然科学基金委每年的智 能车赛事,吸引了新闻媒体、汽车行业 及社会大众的广泛关注,促进了产学 研合作开展攻关,一定程度上推动了 行业和产业发展。这种理论研究与实 际应用紧密结合的做法, 值得在国家 科技发展的其他领域进行推广

据介绍,团队下一步将就听觉认 知技术,特别是视觉、听觉信息的融合 感知进行深入研究。

动 大 程 唯 脑 珈

机

器

与

体

的

深

度

融

当很多人对人工 智能还懵懵懂懂时,一 个被认为是人工智能 下一个风口的"黑科 技"——有望实现"人 机共融"的脑—机接口 技术已悄然来临。

项目组供图

什么是脑—机接 口? 脑—机接口主要是 通过解码大脑活动信 号获取思维认知等信 息,实现人脑与外界直 接交流,在医学康复领 域的应用已逐步兴起。

它除了能帮助具 有严重功能障碍的患 者建立与外界的交流 通道,还可将康复训 练中很多的被动运动 转换成患者的主动运 动,实现大脑、机器与 肌体的深度融合,靶向 诱导并强化推动中枢 神经可塑性变化,从而 提高康复效果,克服传 统康复手段被动单一 的缺陷。

在国家自然科学 基金委员会(以下简称 自然科学基金委)重大 研究计划"视听觉信息 的认知计算"支持下, 天津大学医学工程与 转化医学研究院院长、 天津神经工程国际联 合研究中心主任明东 带领团队突破了一系 列基于脑—机—肌信 息环路的卒中康复机 器人基础理论与关键

技术,有望帮助致残病人恢复正常生活、 重建其对生活和康复的信心。

"我们整个项目研究工作紧密围绕 面向高性能人机交互的训练型脑—机— 肌协同交互信息环路模型,分别开展了 脑—机、肌—机和脑—肌交互中的关键 技术研究。"明东告诉《中国科学报》,在 脑--机交互的神经信息学复合编解码技 术研究中,团队实现了反应式、主动式、 被动式的脑—机高效通信;在肌—机交互 的神经肌骨动力学复合调控技术研究中, 团队围绕神经肌骨动力学的精准调控问 题,建立了电刺激条件下的多源信息复合 模型,实现了康复过程中人体生理特征的 协同分析与运动状态的安全监测;脑一肌 交互的神经可塑性定量分析技术中,团队 发明了基于时间序列的非线性动力学信息 定量分析方法,开拓了神经系统重大疾病 皮层信息传输功能定量表征的新途径。

"研究最大的亮点特色是提出了脑-机一肌信息交互环路的基本模型,发现了 影响环路模型的三个关键要素,即信息识 别准确性、人机交互时效性及反馈形式有 效性,并在此基础上实现了大指令、精辨 识、快通信、强交互的新型卒中康复临床 应用。"他说。

脑卒中是世界致残率第一、严重危害 人类生命健康的常见高发病,探索新型康 复治疗机理与方法已成为世界范围,尤其 是我国的健康医疗研究的关注焦点。依 托该项目,课题组设计出人工神经康复机 器人系统"神工一号""神工二号",并通过 国家食品药品监督管理局检测,在天津市 人民医院、天津市第一中心医院、山东省 烟台市医院等多家三甲医院临床测试成 功,受益患者三千余例,有力推动了新兴 的脑—机交互技术在临床康复工程领域 的发展与应用。

此次相关技术成果在航天领域也大有 用武之地。团队与中国航天员中心合作,研 制了空间站在轨脑—机交互技术试验系 统,应用于2016年我国"天宫二号"和"神 舟十一号"载人飞行任务,成功完成国际首 次脑—机交互空间适应性测试。

明东告诉记者,未来的研究计划主要 包括两个方面,一方面是科研成果的转化 与落地,进一步完善助行人工神经康复机 器人系统功能与外观设计,完成系列产品 注册,实现大规模化临床示范应用。另一 方面,团队将进一步探索脑卒中神经康复 机制,发展新型智能生机交互与生机电一 体化技术手段。

求 智 程 能 时 的 垂 口 音 空

ন

和关联检索,成 为互联网音频 检索业务智能 化的竞争关键。 "基于人们 互联网连接方 式的改变,以语 音为核心的智 能交互将是大 趋势。"在接受 《中国科学报》 采访时,中科院 声学所研究员 颜永红如是说。 然而,与美国等 发达国家相比, 我国智能语音 处理技术的民 用市场普及率 较低,面对新一 轮科技竞争,智

当前,互联

网语音与影像

的内容遍布生

活各个角落,而

语言当之无愧

地成为信息最

重要的载体。如

何实现高效的

语音处理、识别

成为国家需求。

间

在国家自然科学基金委员 会(以下简称自然科学基金委) 重大研究计划"视听觉信息的认 知计算"支持下,颜永红带领团 队就多人多方对话中的语音分 离、内容分析与理解开展深入研 究,并取得了良好的市场效应。复 杂声学场景下的语音分离、内容 分析与理解是信息处理领域前沿 性和基础性的研究课题。

能化语音发展

颜永红表示,该研究最大的 亮点在于多语言建模。"我们做 的混合语音识别,比如中文里面 夹杂着英语,系统都可以有效地 识别与解码。而且通过数据共 享,我们建立了语音的统一表 达,这样在引擎构建的时候,可 以大幅削减数据量需求。

在基础理论和关键技术研 究的基础上,团队还构建了语音 交互系统和海量音频内容检索 系统,并针对国家重大需求研究 开发了相应的平台进行应用推

基于海计算和云计算环境 的音频内容识别与理解服务平 台就是其中之一。该平台重点解 决了在实际应用环境下语音识 别技术面对噪声、信道和口音的 稳健性问题及大规模并发处理 等实际应用难题。

"在互联网领域,我们先后 跟国内三大互联网企业——百 度、腾讯和阿里巴巴,进行了卓 有成效的合作,推出了语音输 人、语音和音乐检索、语音客服 等应用。

另一个则是多语种海量音 频信息自动处理平台。团队研发 出语音关键词识别、说话人/语 种识别、固定音频检索等核心技 术,有效提高了识别系统对噪 声、信道及口音的适应性和在实 际应用环境下的识别性能,构建 了符合实际业务流程的多语种 海量音频自动理解系统。

在颜永红看来,这些都离不 开自然科学基金委的统筹规划。 "视听觉的认知计算"重大研究 计划专家指导组充分发挥顶层 设计作用,聚集了各领域的优秀 人才,并尽最大可能发挥已有优

势项目的潜力。 颜永红表示,团队未来将 继续加强基础研究建设,开展 对韵律识别等领域的前瞻研究 并将项目集成与推广。"假以时 日,基于语音操作的科技产品 会有更大的市场和更广阔的用 途。"他说。