



编者按

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。

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关注。

“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

让智能车“眼观六路耳听八方”

■本报见习记者程唯珈

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。

在自然科学基金委(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入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



有智能驾驶功能的汽车

项目组供图

“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与健康水平,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)信息处理所面临的瓶颈困难。

如何才能实现这一目标?在郑南宁看来,这

推动大脑、机器与肌体的深度融合

“人眼视觉选择性注意机制在计算机视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、动态障碍物等,并提出多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工作,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应用。这种理论与实际紧密结合的做法,值得在国家科技发展的其他领域进行推广。

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

推动大脑、机器与肌体的深度融合

■本报见习记者程唯珈

当很多人对人工智能还懵懂时,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。

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,实现人脑与外界交流,在医学康复领域的应用已逐步兴起。

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-肌信息交互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程解码技术研究中,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-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-肌-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、人机交互时延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疾病,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网络康复机器人系统“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-肌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有用武之地。团队与中国航天员中心合作,研制了空间站轨道脑-肌交互技术试验系统,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-肌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,发展新型智能生肌交互与生肌电一体化技术手段。



2018中国智能车未来挑战赛开幕式现场

项目组供图

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。

“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-肌接口比赛’,并组织了10届‘中国智能车未来挑战赛’和2届‘中国脑-肌接口比赛’。”郑南宁介绍,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应用基础研究的有机结合。

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年的摸索和攻关,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的主要研究内容。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叉——而这正是郑南宁十年来感受颇深的地方。

“我们研究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,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”两方面为切入点,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”。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:清华大学在脑机交互模型融合研究方面,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,人文社会科学和哲学学科将会大有作为。”郑南宁说。

寻求智能时代的语音空间

■本报见习记者程唯珈

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趋势。”在接《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,智能语音发展

成为国家需求。在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。

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广。

基于云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,推出了语音输入、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’重大研究计划专家组充分发挥顶层设计与专家指导作用,聚集了各领域的优秀人才,并尽最大可能发挥了有优势项目的潜力。”

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应用。”他说。