

智能语音技术让人机交互更“丝滑”

◎本报记者 都芃

6月26日,中国科学院举办学部第九届学术年会,中国科学院外籍院士盖博·施德潘作学术报告。他一边用英语讲,同传系统一边将他的报告内容实时翻译为中文,呈现在大屏幕上。内容准确,翻译迅速,即使是专业名词也不在话下。

该同传系统所采用的技术正来自于不久前获得2023年度国家科学技术进步奖一等奖的“多语种智能语音关键技术及产业化”项目。

该项目由科大讯飞牵头,联合中国科学技术大学、清华大学、华为、中移(杭州)公司等高校及行业领军企业联合完成。项目历时10余年研发成功,攻克多项关键技术,构建起具有中国自主知识产权的多语种智能语音技术体系,为我国智能产业升级发展奠定坚实基础。

复杂场景下识别准确率显著提升

语音是人类最自然便捷的沟通方式,也是智能时代人机交互的关键入口。如今我们只需张嘴,机器就能代替人类完成一系列任务。在这“丝滑”的交互体验背后,是一系列关键技术的创新突破。

要让机器听懂、听懂人类语言并不是一件容易的事。语音识别研究中有一道最具挑战性的典型难题被称为“鸡尾酒会”效应,即远场环境、噪声干扰、多人语音混叠等现象导致的语音识别困难。“比如在一个嘈杂

的鸡尾酒会上,很多人同时说话,背景音乐也很响,如何让机器在这样的环境中精准识别出需要的声音?这是很常见的应用场景,也是非常具有挑战性的难题。”科大讯飞副总裁、研究院院长刘聪介绍,针对这一典型难题,研发团队提出了多通道语音信号时空分离建模方法,即先将人声和不同的噪音分离,再进行建模。同时,团队还提出多维度语音属性解耦表征方法。“简单来说就是解耦、提取内容、韵律、音色等语音属性,以便更准确地识别和传达语音信息,解决了‘鸡尾酒会’效应难题,使得复杂场景下语音识别准确率得到显著提升。”他说。

刘聪向记者展示的一个典型应用场景中,在嘈杂的大型会议现场,三人同时就不同主题进行混叠发言,此时正常人难以分辨讲话人的具体内容。但智能语音技术不仅可以实现三人重叠语音的角色分离,还可以对每个人的发言进行实时转写,准确率在86%以上。

多语种翻译是智能语音技术另一高频、高需求应用场景,也是最能体现智能语音技术实力的场景之一。数据是语音技术训练的关键因素之一,但部分小语种训练知识匮乏、训练数据稀缺,成为困扰技术发展的瓶颈。

2014年,科大讯飞等单位开始研究包括中国部分民族语言在内的小语种智能语音技术。他们从维语和藏语等民族语言切入,与北京外国语大学、上海外国语大学等建立密切合作,积累了宝贵资源,沉淀大量技术。

在此基础上,研发团队设计出全新的多语种通用音素体系和基本语言单元,实现多语种统一音素韵律体系的构建。“我们将不同的小语种进行分类,找到同类语种的共同规律,在此基础上进行分析建模和训

练。最终在训练数据数量、质量有限的情况下,显著提升了小语种语音系统性能。”刘聪介绍。

此外,为解决复杂应用场景下语音交互、语音翻译面临的深层次语义理解困难、专业性不足等难题,研发团队还提出了多源知识增强的可信文本生成技术,提升专有词汇及领域知识引用的准确率。为解决使用国产芯片开展智能语音算法模型训练和推理面临的性能低、适配难等问题,团队通过软硬件协同优化的动态张量算子自动融合、多硬件联合的量化计算模拟等方法,解决了智能语音技术硬件平台受制于人的问题。

大模型使智能语音更具“人情味”

近年来,大模型引发的人工智能浪潮席卷多个领域。智能语音技术同样也在大模型助力下不断取得新突破。“大模型+语音”的发展模式,被认为将给智能语音技术和产业带来新一轮革命性突破。

“比如大模型可以实现语音技术的超拟人合成,让机器说话不再有浓厚的朗读腔,能够像真人一样自然对话。而全双工交互,则可以同时、瞬时进行信号的双向传输,让人机对话可以随时打断和继续等。这些都让大模型在智能语音领域备受瞩目。”科大讯飞研究院常务副院长高建清介绍说,智能语音技术的关键创新与大模型技术相结合后,二者可以相互补充、相互促进。

例如,借助语音属性解耦、语音信号时空分离等技术突破,可以将语音信号经过编码后输入到大语言模型中,显著提升语音大模型的效果。“原来的语音合成、语音识别、机器翻译等单点技术,借助大模型后台的理解能力、文本生成能力,在语义理解、指令跟随、多轮对话、情绪感知、超拟人合成等方面实现了显著提升,使语音交互体验得到极大改善。”高建清说。

着眼未来,大模型技术可以在复杂语义理解、长文本建模能力上进一步提升语音识别、合成和翻译的效果。同时基于大模型强大的语义理解、知识问答、多轮对话、多模态建模能力,它也能进一步提升智能语音技术的使用场景和应用价值,支撑实现语音同传、自动客服、辅学答疑、家庭医生、虚拟员工、陪伴机器人、服务机器人等未来智能产品创新,培育出更多产业机会,加速通用人工智能时代到来。高建清认为,大模型时代的语音技术本质上是一次推动万物互联和原有人机交互场景重构的巨大产业机会,是语音产业的全新战略机遇。

目前,科大讯飞自主研发的多语种智能语音技术已支持69个语种,除了6个联合国通用语种(汉语、英语、法语、西班牙语、俄语、阿拉伯语)之外,还包括挪威语、丹麦语等小语种63个,均达到使用门槛。搭载相关技术的智能手机海内外累计激活超10亿台,车载智能化产品累计前装超5300万套,支持奇瑞、比亚迪、长安等车企超200万套“出海”订单。

成果播报

6000米级智能电驱动深海重载采矿车辆平台研制成功

科技日报(记者刘国园)记者6月27日获悉,湖南省科技重大专项“深海矿产资源开采关键装备技术研究及示范应用”日前通过综合验收。通过该项目,我国研发出全球首台6000米级智能电驱动深海重载采矿车辆平台。

据介绍,在该项目的支持下,科研团队围绕深海矿产资源开发这一国家未来产业发展方向,聚焦3000米至6000米深海矿产的勘探、采集、输送等作业关键环节,紧跟数字化、

电动化、智能化技术发展趋势,重点攻克了多项关键核心技术和共性关键技术,研制了全球首台6000米级智能电驱动深海重载采矿车辆平台等6台(套)重要装备,部分成果填补了国际空白。项目相关研究成果共申请专利94项,其中已授权45项,获软件著作权17项。

记者了解到,该专项由湖南大学和中国五矿集团有限公司所属长沙矿冶研究院有限责任公司共同牵头,共7家在湘单位联合承担。



6000米级智能电驱动深海重载采矿车辆平台。受访单位供图

以电流画图像,“透视”大坝病灶

◎本报记者 魏依晨

解决,供电和测量电极之间的切换工作就必须人工来完成。

“网络查不到任何相关技术的具体资料,我们只能根据要实现的功能不断尝试。”邓居智介绍,从2007年开始,团队研究了多种方法。经过反复测试,他们最终在2011年研制出了拥有自主知识产权的分布式三维电阻率“CT”系统样机。

邓居智说,当时研制出的样机很笨重。如今,这套仪器已经升级到了第四代。其不同机型可分别用于中深部铀资源勘探、地下空洞探测、坝体隐患探测、垃圾填埋场渗漏探测及考古等领域。

2020年7月,江西连降暴雨,赣江、鄱阳湖流域遭受严重洪涝灾害。邓居智团队带着仪器前往抗洪一线。团队先后在赣江新区、永修县九合联圩、彭泽县芙蓉堤、庐山区、鄱阳镇邓一彭家村等危堤,开展“堤坝CT体检”工作,排查并指导修复蚁穴、孔洞和管涌泡泉通道等隐患200余处。通常情况下,完成数据收集后,该设备只需2—3个小时就能得到堤坝的三维电阻率结构图,直观地了解到内部情况。如果情况紧急,半个小时就能得到二维成像图。不仅如此,这套设备还在江西省海昏侯墓考古、贵溪市雷溪乡突发地面塌陷应急探测中发挥了重要作用。

近日,在江西一处堤坝上,东华理工大学地球物理与测控技术学院院长邓居智跟同事们,将一卷红色的电缆从车上拿下,用它们将打入土内的钢钎(电极)联接起来。

“这些是金属电极,我们根据需要布设电极后,就可以给堤坝做检查。”邓居智说,启动与电缆连接的主机,电流便可在堤坝体内传导。遇到不同的地质体,电流大小会发生变化,地面的测量电场也会随之改变。“这就如同用电流为堤坝画三维图像,从而‘透视’坝体结构,快速判断堤坝是否存在空洞、裂缝、管涌、渗流等隐患。”邓居智说。

邓居智使用的仪器,是东华理工大学自主研发的DSDT-3型双分布式三维电阻率成像系统。以往,此类仪器多依靠国外进口。邓居智深知,关键核心技术要不来、买不来、讨不来。自2000年开始,他便开始关注工程与环境地球物理的相关问题。2007年的一次国际合作项目,让他下定决心自主研发三维电阻率成像系统。

“要研制这套系统,核心在于攻克供电和测量电极之间的智能切换以及双分布式测量方式这两个难题。”邓居智回忆,如果这些问题不能

AI智能巡检机器人护航牡佳高铁运行

◎本报记者 李丽云 朱虹
通讯员 孙冰冰 张成杰

径进行巡检。当出现告警信息时,机器人可以自动采集数据进行分析判断。但由于这款机器人不具备智能避障功能,在巡检过程中不会被杂物、柜体阻挡,影响巡检效果。

2021年,根据牡丹江电务段的使用反馈,哈尔滨国铁科技集团股份有限公司开始第二代滑轨式智能机器人的研发。机器人于2024年5月投入使用。

第二代滑轨式智能机器人采用滑轨运行模式,最小转弯半径只有30毫米,还配备了升降云台。它的升降行程高达2米,可实现无死角位移。该机器人还增加了2D激光定位器,将最小对位精度提升至2毫米,确保了定位巡检的精准性。同时,它还采用了先进的TOF视觉面阵避障技术,能够敏锐感知到周围障碍物,有效解决避障难题。

同时,它还兼具监控漏水、温度、湿度、防鼠等情况的功能,如超过预设值,机器人将会发出报警信号,提醒检修人员及时处置,确保信号设备状态良好。

目前,哈尔滨国铁科技集团股份有限公司完成了牡佳高铁20台滑轨式智能巡检机器人的安装调试,并开始试运行。试运行合格后,该智能巡检机器人将在牡佳高铁线路上应用。

吊装滑轨上,智能机器人沿着预设轨迹前行,前端的红外热成像仪扫过200余台信号设备,将设备运行状态生成视频、图片,实时传回指挥中心……6月27日,在黑龙江省牡佳高铁1号中继站,由中国铁路哈尔滨局集团有限公司研发的第三代AI智能巡检机器人正在试运行。

牡佳高铁是我国最东端高寒高铁。其运行环境温度大,线路分布在山区原野。动车组运行过程中需要实时传送信号,为动车组运行提供指引。为此,铁路部门在高铁沿线建设了20个无人值守的信号机械室进行信号补偿,确保信号畅通。每个机械室有200余组信号设备,牡丹江电务段每月需要安排4名人员对这些设备进行检查。但由于中继站地处偏远,车辆无法直接抵达,检修人员要徒步前往。最近的中继站,检修人员也要走1个小时才能到达,耗时耗力。

早在2020年,哈尔滨局集团公司便开展了智能机器人的研发工作,以期用它来代替人工巡检。初代机器人以履带方式根据地面粘贴的二维码路

新建高铁汕头站“海潮奔涌”整体造型亮相

科技日报(记者孙瑜)记者6月26日从中铁建设集团有限公司获悉,粤东地区规模最大、交通接驳功能最齐全的高铁枢纽——新建汕头站站房一期工程外立面装修基本完工。“海潮奔涌”整体造型首次亮相,车站建设全面转入内部装饰装修收尾阶段。

新建汕头站设计立意“海丝腾飞,潮立汕头”,以富有海洋韵味的动感曲线模拟潮水起伏。站房高架候车层由1465块单重超700公斤的玻璃模块构成。因现场交叉作业环境复杂,传统吊装方式难以适用,项目团队引进高空摇臂吸盘车进行安装。“该设备具有毫

米级驱动系统,能在空中自由升降、旋转、移动,一天最高能安装30块玻璃,效率是传统吊装的3倍。”中铁建设新建汕头站项目幕墙专业负责人王禧年介绍。

为增强建筑抗风性和防水性,站房外立面铝板幕墙采用开缝式单元设计。“铝板背后铺设防水背板,可抵御每分钟100升流速的强烈水流冲击,相当于给幕墙穿上一层‘防水服’。”王禧年介绍,“不仅如此,开缝式铝板背后的空气层能够顺畅通,有助于幕墙内外压力保持平衡,减少狂风对幕墙的直接冲击。”

一期站房屋盖重超8000吨,支撑柱

最大跨度达96米。为保证结构稳固,项目团队研发了全国首例建构一体化树状造型钢结构柱施工技术,进一步强化结构受力安全。团队通过数十次受力模拟,将支撑柱打造成“榕树”造型,使立柱通过延展的“树枝”形成反顶之力,擎举起整个屋盖。

为确保钢结构施工安全,项目团队还开发了钢结构健康监测系统。团队通过在钢构件上布置68个监测点和185个传感器,对钢结构在安装过程、整体提升和卸载后的受力状态进行全面“体检”。传感器将监测到的结构应力、位移等参数打包发送到手机等终端平台,

高效直观呈现钢结构状态。

此外,项目团队围绕精细化管理和智慧化施工,引入喷涂机器人等20余种智能工装设备,采用物联网技术对深基坑、既有有线、配电箱、消防水压等进行数字化监测,为站房建设筑牢安全底线。“目前,工程整体施工进度已完成约70%,1600多名建设者正加速推进室内装饰装修施工进度。”中铁建设新建汕头站项目负责人马银枪说。

新建汕头站共分两期建设,一期新建东站房,二期改扩建既有西站房,建筑面积10万平方米,建成后将加快粤东地区发展和汕潮揭城市群建设。

我国首套百千瓦级电氢双向转换装置投运

科技日报(记者叶青 通讯员黄欣然 黄旭锐)6月28日,记者来到广州南沙小虎岛电氢智慧能源站,见到纵横交错的“钢管森林”旁矗立着3个银灰色集装箱,以及1个小型储氢罐。这是我

国首套自主研发的百千瓦级电氢双向转换装置。它于近日在广州正式投运。

据介绍,该装置在国内率先实现通过一套装置完成制氢和发电,实现了绿电和绿氢的双向高效可靠转换。这标

志着我国可逆固体氧化物电池技术从实验室走向了示范应用。

“这个装置基于可逆固体氧化物电池技术,具有电解池和燃料电池两种模式,集制氢与发电功能于一体。”广州供电局氢能研究中心杨怡萍介绍,“在电解池模式下,装置通过高温电解水制取‘绿氢’。将这些氢气收集到储氢罐,可随时向站内供氢。当一键切换到燃料电池模式后,装置可将储氢罐内的氢气输送回装置,通过电化学反应按需发电并网。”

研发团队先后取得大功率可逆固体氧化物电池等技术突破。目前该装置的系统电解制氢功率已达100千瓦,电堆平均每三度电可制一立方氢。相比传统电解制氢技术,该装置的效率提升了20%—30%。

广州供电局氢能研究中心技术总监区定介绍,该装置可作为储能设备,在用电低谷时将富余的绿色电能用来制氢,并储存起来,有助于粤港澳大

湾区风电等新能源的大规模就近消纳。它还可作为发电设备,在用电高峰时及各类应急情况下,将储存的绿色氢气用来发电,使电力供应更加安全可靠、绿色低碳。

该装置实现了从材料到系统的全面国产化。“作为小型、分布式的能源生产单元,未来通过模块化叠加组合,装置的制氢与发电功率可达兆瓦级。”杨怡萍说。

这套电氢转换装置处于国际先进、国内领先水平。预计到“十五五”期间,该技术将进一步应用于冶金、化工等领域,电氢转换规模将提升百倍,拓展到万千瓦级。

广州供电局副总经理龙云认为,通过“电—氢”转换,全社会用氢客户可以直接或间接地成为用电客户。这既能大幅增加电网消纳新能源的能力,也能进一步提高电能在终端能源消费的比重,最终支撑实现碳达峰碳中和。



我国首套自主研发的百千瓦级电氢双向转换装置投运。受访单位供图