

中国自动化学会通讯

COMMUNICATIONS OF CAA

第 1 期

2023 年 01 月

第44卷 总第232期

主办：中国自动化学会

<http://www.caa.org.cn>

E-mail: caa@ia.ac.cn

京内资准字2020-L0052号

恭賀
兔年



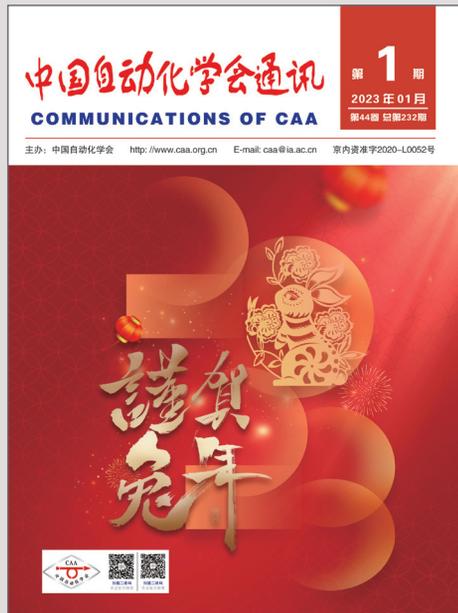
扫描二维码
关注官方微信



扫描二维码
关注官方微博



中国自动化学会通讯
Communications of CAA



主管单位 中国科学技术协会
主办单位 中国自动化学会
编辑出版 中国自动化学会办公室



关注官方微信



关注官方微博

主 编 | 郑南宁 CAA 理事长、中国工程院院士、西安交通大学教授

副主编 | 王飞跃 CAA 监事长、中国科学院自动化研究所研究员

杨孟飞 CAA 副理事长、中国科学院院士、中国空间技术研究院研究员

陈俊龙 CAA 副理事长、欧洲科学院院士、华南理工大学教授

编 委 | (按姓氏笔画排列)

丁进良 王 飞 王占山 王兆魁 王庆林
王 坛 邓 方 石红芳 付 俊 吕金虎
乔 非 尹 峰 刘成林 孙长生 孙长银
孙彦广 孙富春 阳春华 李乐飞 辛景民
张 楠 张 俊 陈积明 易建强 周 杰
赵千川 赵延龙 胡昌华 钟麦英 侯增广
姜 斌 祝 峰 高会军 黄 华 董海荣
韩建达 谢海江 解永春 戴琼海

刊名题字 | 宋 健

地 址 | 北京市海淀区中关村东路 95 号

邮 编 | 100190

电 话 | (010) 8254 4542

传 真 | (010) 6252 2248

E-mail: caa@ia.ac.cn

http: //www.caa.org.cn

印刷日期 | 2023 年 1 月 31 日

发行对象 | 中国自动化学会会员及自动化领域科技工作者

本刊声明

◆ 为支持学术争鸣, 本刊会登载学术观点彼此相左的不同文章。来稿是否采用并不反映本刊在学术分歧或争论中的立场。每篇文章只反映作者自身的观点, 与本刊无涉。

主编的话



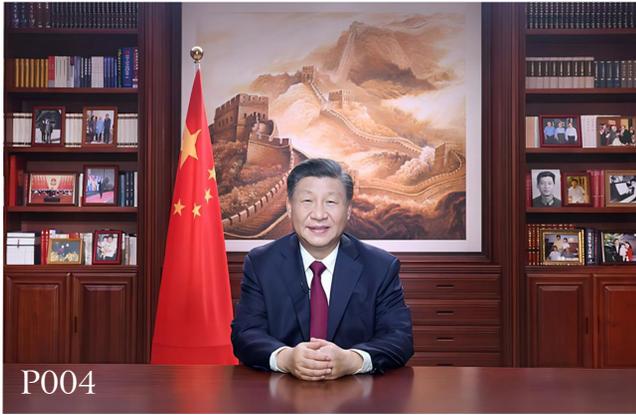
郑南军

日月开新元，万象启新篇。回首2022年，我们迎来了党的二十大胜利召开，也迎来了中国自动化学会向着第二个甲子奋进的历史新节点。这一年，学会在抗击疫情的同时稳步推进各项工作有序开展，收获满满。充满希望的2023年已经到来，首先向长期以来关心和支持《中国自动化学会通讯》发展的编委、学者、读者，以及广大会员和各界人士，致以新年最诚挚的感谢和美好的祝愿。

新岁开启新梦想，新年承载新希望！2023年，我们迈上了全面建设社会主义现代化国家的光明征程，想要实现高水平科技的自立自强，离不开广大科技工作者的智慧与汗水。中国自动化学会作为自动化领域创新的“领头雁”，将始终面向世界科技前沿、面向经济主战场、面向国家重大需求，始终把自动化科技创新摆在更加突出的战略地位，加强原创性、引领性科技攻关，奋勇攀登世界科技巅峰，坚持创新驱动，勇于跨越鸿沟，为实现国家强盛、民族复兴贡献自动化更大的力量。

为进一步满足科技工作者获取自动化前沿科技最新研究进展和成果的需求，更好地激发创新创造活力，《中国自动化学会通讯》将继续致力于为学界搭建高质量的学术成果展示和交流平台，为广大科技工作者搭建施展才华的舞台，让科技创新成果源源不断涌现出来，践行《中国自动化学会通讯》的使命与担当。

科技期刊既是科技成果的承载和传播者，也是科技强国的重要推动者。新的一年，《中国自动化学会通讯》继续立足办刊宗旨，不忘初心，以更加昂扬奋进的姿态，团结引领广大自动化人在新时期以更高的热情、更专业的能力凝心聚力办好品牌期刊，推动自动化学科取得新成绩，在建设科技强国的使命上作出新贡献。值此新春佳节之际，衷心祝愿大家身体健康，阖家幸福！



专题 / Column

- 004 国家主席习近平发表二〇二三年新年贺词
- 006 中国科协主席万钢发表二〇二三年新年贺词
- 007 中国自动化学会理事长郑南宁院士发表二〇二三年新年贺词
- 009 中国自动化学会监事长王飞跃教授发表二〇二三年新年贺词

观点 / Viewpoint

- 010 张钹院士：做负责任的人工智能
- 012 李德毅院士：认知物理学——薛定谔、图灵和维纳的启示和超越

学术前沿 / Academic Frontier

- 016 视觉语言导航研究进展
/ 司马双霖 黄岩 何科技 安东 袁辉 王亮
- 029 面向夜间疲劳驾驶检测的改进 Zero-DCE 低光增强算法 / 黄振宇 陈宇韬 林定慈 黄捷
- 039 空间机器人抓捕目标后基于任务相容性的消旋策略 / 周逸群 罗建军 王明明

学会动态 / Activities

- 054 “智行中国”系列论坛第三期“智能科学与技术交叉学科发展新机遇”线上论坛成功举办
- 062 第七届全国高校自动化教学院长（教学主任）会议顺利召开





- 064 “青少年人工智能核心素养测评” 2022 年度总结大会圆满召开
- 068 中国自动化学会发电自动化专委会召开 2022 年度工作总结交流会议

形势通报 / Voice



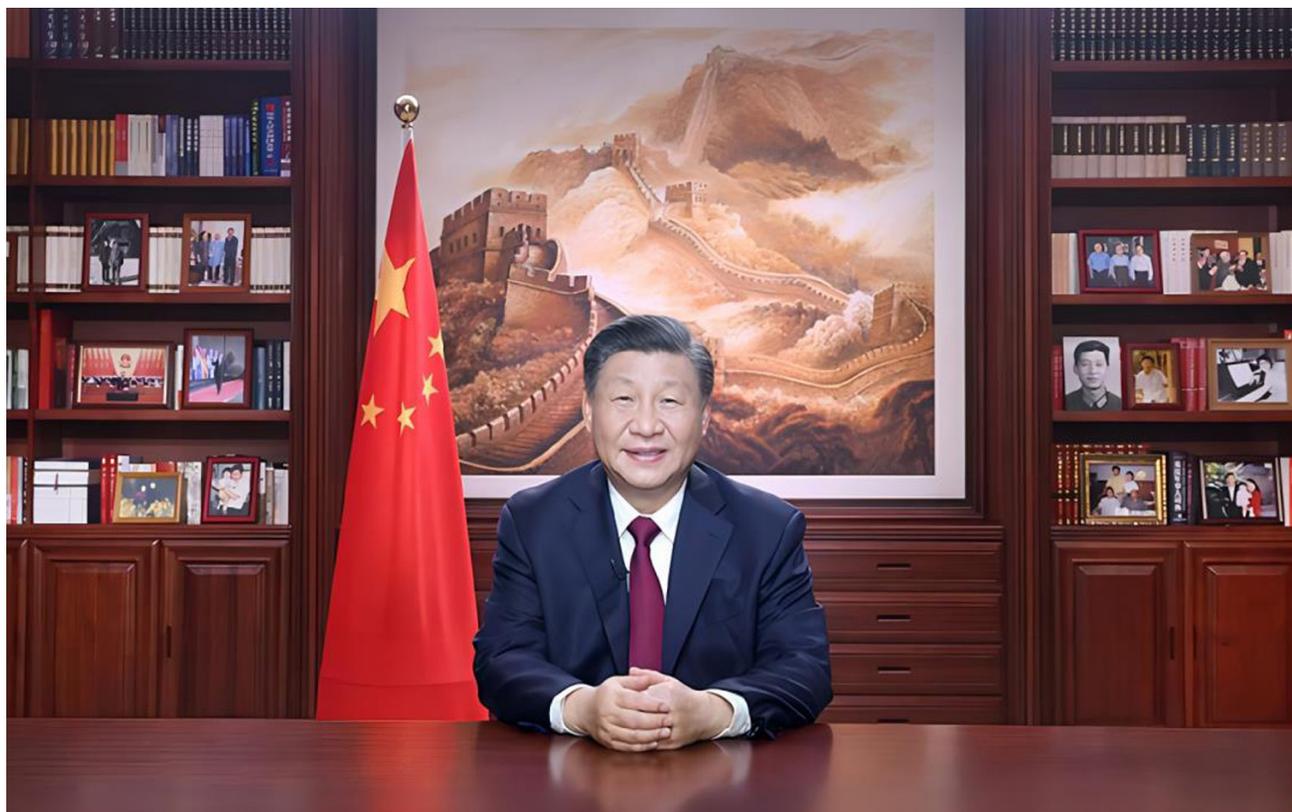
- 071 关于完善工业和信息化领域科技成果评价机制的实施方案（试行）
- 074 2023 年度国家自然科学基金改革举措

党建强会 / Party Building



- 077 习近平在二十届中央纪委二次全会上发表重要讲话
- 080 以中国式现代化全面推进中华民族伟大复兴
- 084 学会办事机构党支部组织召开 2022 年度民主生活会

国家主席习近平发表二〇二三年新年贺词



新年前夕，国家主席习近平通过中央广播电视总台和互联网，发表二〇二三年新年贺词。新华社记者 鞠鹏 摄

新年前夕，国家主席习近平通过中央广播电视总台和互联网，发表了二〇二三年新年贺词。全文如下：

大家好！2023年即将到来，我在北京向大家致以美好的新年祝福！

2022年，我们胜利召开党的二十大，擘画了全面建设社会主义现代化国家、以中国式现代化

全面推进中华民族伟大复兴的宏伟蓝图，吹响了奋进新征程的时代号角。

我国继续保持世界第二大经济体的地位，经济稳健发展，全年国内生产总值预计超过120万亿元。面对全球粮食危机，我国粮食生产实现“十九连丰”，中国人的饭碗端得更牢了。我们巩固脱贫攻坚成果，全面推进乡村振

兴，采取减税降费等系列措施为企业纾难解困，着力解决人民群众急难愁盼问题。

疫情发生以来，我们始终坚持以人民为中心、生命至上，坚持科学精准防控，因时因势优化调整防控措施，最大限度保护了人民生命安全和身体健康。广大干部群众特别是医务人员、基层工作者不畏艰辛、勇毅坚守。经过艰

苦卓绝的努力，我们战胜了前所未有的困难和挑战，每个人都不容易。目前，疫情防控进入新阶段，仍是吃劲的时候，大家都在坚忍不拔努力，曙光就在前头。大家再加把劲，坚持就是胜利，团结就是胜利。

2022年，江泽民同志离开了我们。我们深切缅怀他的丰功伟绩和崇高风范，珍惜他留下的宝贵精神财富。我们要继承他的遗志，把新时代中国特色社会主义事业不断推向前进。

历史长河波澜壮阔，一代又一代人接续奋斗创造了今天的中国。

今天的中国，是梦想接连实现的中国。北京冬奥会、冬残奥会成功举办，冰雪健儿驰骋赛场，取得了骄人成绩。神舟十三号、十四号、十五号接力腾飞，中国空间站全面建成，我们的“太空之家”遨游苍穹。人民军队迎来95岁生日，广大官兵在强军伟业征程上昂扬奋进。第三艘航母“福建号”下水，首架C919大飞机正式交付，白鹤滩水电站全面投产……这一切，凝结着无数人的辛勤付出和汗水。点点星火，汇聚成炬，这就是中国力量！

今天的中国，是充满生机活力的中国。各自由贸易试验区、海南自由贸易港蓬勃兴起，沿海地区踊跃创新，中西部地区加快发展，东北振兴蓄势待发，边疆地区兴边富民。中国经济韧性强、

潜力大、活力足，长期向好的基本面依然不变。只要笃定信心、稳中求进，就一定能实现我们的既定目标。今年我去了香港，看到香港将由治及兴十分欣慰。坚定不移落实好“一国两制”，香港、澳门必将长期繁荣稳定。

今天的中国，是赓续民族精神的中国。这一年发生的地震、洪水、干旱、山火等自然灾害和一些安全事故，让人揪心，令人难过，但一幕幕舍生取义、守望相助的场景感人至深，英雄的事迹永远铭记在我们心中。每当辞旧迎新，总会念及中华民族千年传承的浩然之气，倍增前行信心。

今天的中国，是紧密联系世界的中国。这一年，我在北京迎接了不少新老朋友，也走出国门讲述中国主张。百年变局加速演进，世界并不太平。我们始终如一珍视和平和发展，始终如一珍惜朋友和伙伴，坚定站在历史正确的一边、站在人类文明进步的一边，努力为人类和平与发展事业贡献中国智慧、中国方案。

党的二十大后我和同事们一起去了延安，重温党中央在延安时期战胜世所罕见困难的光辉岁月，感悟老一辈共产党人的精神力量。我常说，艰难困苦，玉汝于成。中国共产党百年栉风沐雨、披荆斩棘，历程何其艰辛又何其伟大。我们要一往无前、顽强拼搏，让明天的中国更美好。

明天的中国，奋斗创造奇迹。苏轼有句话：“犯其至难而图其至远”，意思是说“向最难之处攻坚，追求最远大的目标”。路虽远，行则将至；事虽难，做则必成。只要有愚公移山的志气、滴水穿石的毅力，脚踏实地，埋头苦干，积跬步以至千里，就一定能够把宏伟目标变为美好现实。

明天的中国，力量源于团结。中国这么大，不同人会有不同诉求，对同一件事也会有不同看法，这很正常，要通过沟通协商凝聚共识。14亿多中国人心往一处想、劲往一处使，同舟共济、众志成城，就没有干不成的事、迈不过的坎。海峡两岸一家亲。衷心希望两岸同胞相向而行、携手并进，共创中华民族绵长福祉。

明天的中国，希望寄予青年。青年兴则国家兴，中国发展要靠广大青年挺膺担当。年轻充满朝气，青春孕育希望。广大青年要厚植家国情怀、涵养进取品格，以奋斗姿态激扬青春，不负时代，不负华年。

此时此刻，许多人还在辛苦忙碌，大家辛苦了！新年的钟声即将敲响，让我们怀着对未来的美好向往，共同迎接2023年的第一缕阳光。

祝愿祖国繁荣昌盛、国泰民安！祝愿世界和平美好、幸福安宁！祝愿大家新年快乐、皆得所愿！

谢谢！○

来源：新华社

中国科协主席万钢发表二〇二三年新年贺词



2022年已然挥手告别，2023年正在招手致意，我谨代表中国科学技术协会，向全国广大科技工作者致以诚挚的新年祝福！

刚过去的2022年，是党和国家发展进程中具有重要里程碑意义的一年。党的二十大吹响了全面建设社会主义现代化国家、全面推进中华民族伟大复兴的时代号角。中国科协在党中央坚强领导下，团结引领广大科技工作者，深入学习宣传贯彻党的二十大精神，坚持聚焦靶心、守正创

新、赋能基层、开放协同，在加强思想政治引领中勇担使命，在发挥桥梁纽带作用中奋发有为，在深化科协系统改革中开拓创新，汇聚起众心向党、自立自强的磅礴力量。

刚过去的2022年，在中国科技史上格外闪亮。航空母舰福建舰下水、“夸父一号”开启太阳探测之旅、中国空间站全面建成、白鹤滩水电站全面投产发电、国产大飞机C919交付运营……科技对促进经济社会发展、提高国家综合实力、满足人民日益增长的美好生活需要的支撑作用显著增强，中国人民的前进动力更加强大、奋斗精神更加昂扬、必胜信念更加坚定。

在这个辞旧迎新、阖家团圆之际，偏远的试验场正在紧锣密鼓，深夜的实验室照旧灯火通明，医院的急诊室里仍在坚韧抗疫，加班的装配厂依然马达轰鸣，海外的工程现场仍然热火朝天……为了科技强国的梦想，无数攻关者挑灯夜战、迎难而上；为了民

族复兴的事业，无数科研人废寝忘食、奋勇登攀；为了人民群众的健康，无数医卫科技人员日以继夜、守护生命。广大科技工作者用心血和汗水，践行了报国为民的初心使命。

2023年是全面贯彻落实党的二十大精神开局之年。蓝图已经绘制，梦想已经起航。科协组织要进一步强化桥梁纽带职责，在系统推进科教兴国、人才强国中把握中国式现代化的历史主动，把事业发展的基点建立在亿万科技工作者和广大人民的创新创造伟大奋斗之上，将广大科技工作者团结凝聚在以习近平总书记为核心的党中央周围，以高水平科技自立自强谱写服务全面建设社会主义现代化国家的新篇章。

我们坚信，报喜的春花竞相绽放，我们将披着灿烂的曙光，迈向成功的殿堂，迎接更大的辉煌！○

来源：中国科协宣传文化部

中国自动化学会理事长郑南宁院士发表 二〇二三年新年贺词



岁将更始，时乃日新。我们即将挥别不平凡的2022年，迈向充满希望的2023年。在这辞旧迎新的时刻，中国自动化学会向自动化学科及行业的专家、学者、工程技术人员的各界同仁，致以新年最诚挚的祝福和问候，向学会所属各分支机构、期刊编辑部、省级自动化学会以及关心学会事业发展的科研院所、企事业单位、媒体等社会各界人士表示由衷的感谢。

2022年是党的二十大召开之年，是踏上全面建设社会主义现代化国家、向第二个百年奋斗目标进军新征程的开局之年，也是学会向着第二个甲子的迈进之年，

更是学会抗击疫情、与自动化人共克时艰、砥砺奋进的一年！

这一年，2021中国自动化大会暨CAA六十周年庆典活动，历时两年筹备、三地迁移、六天班师，终于在春城昆明绽放光芒，突破历史，受万众瞩目；在常态化疫情防控下，2022中国自动化大会也在所有自动化人共同努力下，再次创新会议形式，与1300万人次科技工作者相聚鹭岛厦门，开创了中国自动化大会的新历史，充分展现了我们自动化人直面困难挑战，不断勇毅前行的决心和信心。

这一年，学会坚持党建引领，组织凝聚力全面增强。学会理事会党委多次开展党史教育、知名科学家讲党课、组织召开“党的二十大精神宣讲报告会”；学会秘书处党支部进行系列主题党日、依托CAA学习平台常态化党史学习，深入贯彻学习党的二十大精神；以会员为本，分类分级实施精准会员服务，CAA云讲座、会士面对面、我和优博有个约会、CAA科普大讲堂等线上讲座活动全面开花，克服一切困难为广大会员提供优质服务；持

续优化人才成长环境，设立CAA优秀硕士学位论文奖，发布学会奖励指南，持续优化奖励管理系统，完善学会人才奖励体系，实现奖励申报总数量增长近40%，人才培育、服务与举荐三位一体，极大增强了学会对广大科技工作者的引领吸纳。

这一年，学会坚守学术本源，学术引领力显著提升。学会构建以“学术期刊-学术会议-科普活动-智库成果”为核心的CAA学术引领体系，发布《中国自动化学会学术会议推荐目录》，团结带领分支机构、省级学会以及期刊编辑部等线上线下举办千余场学术交流活动，线上观看人数突破5000万人次，其中以中国自动化大会为代表的十余个品牌学术活动连续5年被中国科协《重要学术会议指南》收录。在学科建设方面，学会积极开展智能科学与技术领域基础科学研究评估工作，编辑出版《智能控制导论》，预判学科发展趋势，推动自动化学科快速发展。加快一流期刊建设，更新《中国自动化学会推荐科技期刊目录》，连续七年参与中国科协优秀科技论文遴

选计划，学会所属期刊也再创辉煌，《自动化学报》复合影响因子6.627，影响因子和影响力指数在所属学科排名第1；《自动化学报》（英文版）入选中科院期刊分区一区Top期刊，最新SCI影响因子7.847，是自动化与控制系统领域唯一的中国主办Q1区SCI期刊。

这一年，学会秉承专业精神，社会公信力迈上新台阶。学会发挥人才优势，组建智能产业和智能制造产业科技服务团，问诊把脉十余个科创中国试点城市200余项技术需求，促成技术合作开发意向十余项，深耕长沙产业发展，落地“科创中国”服务科技经济融合高峰论坛，深度推进产学研融合；发挥智库优势，组建工业控制安全、智慧教育、智能制造决策咨询专家团队，发布《工业安全系统典型案例成果库》，入选“中国科协决策咨询专家团队试点单位”；发挥专业价值，开

展人才培训和评价工作，承接人力资源社会保障部专业技术人员知识更新工程2022年高级研修项目；制定《CAA标准化工作管理办法》，促进团体标准规范优质化发展，开展《面向网联智能驾驶的测评系统构建规范》等7项团体标准研制工作；完善科技成果评价机制，完成46项科技成果评价，以专业精神不断提供高质量、优质化的公共服务产品。

这一年，学会坚持开放合作，国际影响力开拓新局面。推荐十余位学会会员参与2026-2029年国际自动控制联合会（IFAC）等国际组织主席竞选及重要领导职务等选任；承接中国科协开放合作示范专项，进一步完善学会海外会员实施细则，建立海外融媒体宣传平台，广泛吸纳近400名海外会员；在华举办IEEE数字孪生和平行智能国际会议、网络物理与人类系统会议、2022 IEEE

智能交通系统国际会议等在华会议，其中，智能交通系统国际会议是智慧交通领域规格最高、影响力最大的顶级学术会议。

科技是国家强盛之基，创新是民族进步之魂。社会主义现代化强国建设离不开科技支撑。当今世界，自动化已成为人类文明进步和科学技术现代化的重要推动力。面对风起云涌的新一轮科技革命浪潮，自动化不仅迎来了发展新高度，也将引领中国智造迈向更广阔的舞台。

踏上新的征程，中国自动化学会将牢牢紧扣服务国家战略这一时代命题，建设世界一流学会，在向着第二个百年奋斗目标进军道路上，继续乘风破浪、一往无前！

辞旧迎新的钟声即将敲响，在新的一年里祝广大会员和自动化、信息与智能科学领域的科技工作者新年快乐！顺遂安康！○



热烈祝贺中国自动化学会被评为“2021年度科协系统财务数据汇总工作优秀单位”！

根据中国科协办公厅1月9日发布的《关于2021年度科协系统财务数据汇总工作考核情况的通报》（科协厅函财字〔2023〕1号），中国自动化学会被评为“2021年度科协系统财务数据汇总工作优秀单位”！详情请查看：<http://www.caa.org.cn/article/191/3333.html>

中国自动化学会监事长王飞跃教授发表 二〇二三年新年贺词



凯歌辞旧岁，瑞雪迎新年。值此新年即将到来之际，中国自动化学会向自动化、信息与智能科技领域的相关学者、工程技术人员的各界同仁，致以新年最诚挚的祝福和问候，向学会所属各分支机构、期刊编辑部、省级自动化学会以及关心学会事业发展的科研院所、企事业单位、媒体等社会各界人士表示由衷的感谢。

2022年是中国自动化学会新甲子的奋进起点，在这不平凡、与疫情赛跑的一年里，学会严格遵守民政部及中国科协相关规定，遵循依法办会、民主办会、科学办会的主线，瞄准时代特点，主动作为，高质量推进学会各项业务发展，实现了学会组织凝聚力、

学科引领力、社会公信力和国际影响力的大幅提升。

这一年，学会深化治理结构与治理机制改革，持续推进治理方式现代化。探索基于自治科学的智慧学会建设新思路，强化章程意识，进一步完善分工合作的决策机构和实体化建设的执行机构，推进治理结构、人事制度、运行机制和工作方式等方面的改革创新。线上线下召开7次理事长、监事长工作会议，7次秘书长工作会议，1次全体常务理事会议、理事会工作会议，形成决议20余项，落实民主选举、民主决策和民主管理，有效推进学会治理体系和治理能力现代化，逐步形成科学规范、运行有效的制度体系。

这一年，学会加强分支机构规范管理，持续推进组织体系网络化。学会不断完善优化学会分支机构管理系统，严格分支机构创建审核流程。2022年新成立分支机构1个，完成分支机构换届6个，入选中国科协全国学会分支机构示范发展专项项目。截至目前，学会共有59个专业委员会，9个工作委员会，30个省级自动化学会。

这一年，数字化建设全面融入学会治理，开创“智慧学会”新局面。学会启动CAA数字图书馆建设，持续深化官网、微信、微博等平台的广度与深度，重磅推出CAA一站式研究生招生平台、CAA云学院，接长信息化服务手臂，打造高效融媒体传播矩阵，实现学会信息平台与学会业务工作同频共振，2022学会粉丝数量增加近4万人，入选中国科协“智慧科协2.0”首批试点建设单位。

回首2022年，我们心怀感恩；展望2023年，我们笃定前行。新的一年，学会将以习近平新时代中国特色社会主义思想为指导，深入贯彻落实党的二十大精神，主动适应科技经济社会发展新形势，以自治科学探索变革学会管理格局，努力构建学会工作新格局，将学会真正建设成为深受会员喜爱的科技工作者之家，奋力推动我国自动化事业向更高质量发展！

最后，祝愿伟大的祖国繁荣昌盛！祝愿广大会员和自动化、信息与智能科学领域的科技工作者新年快乐！幸福安康！○

张钹院士：做负责任的人工智能

最早提出机器人发展中可能存在伦理风险的是，美国科幻小说家阿西莫夫在他的小说“环舞”（Runaround）中提出的，时间是1942年，早在人工智能诞生之前，为此他还提出规避风险的方法，即大家熟知的“机器人三定律”，应该说这些问题的提出具有前瞻性。后来物理学家霍金等也不断地提出类似的警告，但这些警告并没有引起大家特别是人工智能界的重视。

原因在于他们的立论依据不够充分，他们认为机器的不断进步和发展，有朝一日当它的智力超过人类，特别是机器具有主观意识时，即出现所谓的“超级智能”时，人类将会失去对机器的控制，从而带来灾难性的后果。这种“技术逻辑”对于大多数人工智能研究者来讲并不具有说服力。因为大家清楚地知道，人工智能研究工作目前还处于探索的阶段，进展缓慢、还受到很多问题的困扰，难以解决，制造“超人类”的机器人谈何容易。而且能不能通过所谓“通用人工智能”达到“超智能”的目标，也一直存在着争议。因此我们认为这些

风险只不过是未来的“远虑”而已，不急于考虑。

可是，当本世纪初基于大数据的深度学习在人工智能中崛起之后，人们的认识有了很大的变化，深切地感到人工智能的伦理风险就在眼前，治理迫在眉睫！这是为什么？大家知道，本世纪初基于大数据的深度学习被广泛地应用于各个领域，用来完成决策、预测和推荐等任务，给人类社会带来很大的影响。但是，人们很快发现基于大数据的深度学习算法具有不透明、不可控和不可靠等缺陷，导致AI技术很容易被无意误用，可能给人类社会带来严重的后果。

大家知道，根据目前人工智能的技术，我们可以通过生成式神经网络根据使用者的要求生成符合要求且质量良好的文本和图像。但同样的神经网络也可以违背用户的要求生成充满（种族、性别等）偏见、不公正和错误百出的文本与图像，完全不受使用者的控制。可以设想，如果根据这些生成的错误文本做决策或预测，就可能带来破坏公平性与公正性的严重后果。

我们原以为，只有当机器人的智能接近或超过人类之后，我们才会失去对它的控制。没有想到的是，在机器的智能还是如此低下的时候，我们已经失去对它的控制，时间居然来得这么快，这是摆在我们面前很严峻的现实。

阿西莫夫在《机器人三定律》中曾提经出规避伦理危机的方案，内容是“一，机器人不得伤害人类，或因不作为而让人类受到伤害；二，机器人必须服从人类的命令，除非这些命令与第一定律相冲突；三，机器人必须保护自己的存在，只要这种保护不违反第一或第二定律”。总之一句话，人类应该牢牢把握机器的控制权。让机器做人类的奴隶！这种办法能否解决机器的伦理危机？答案显然是否定的！

实际上，让“机器完全听从人类的指挥”，在早期“无智能”的机器中我们就是这样做的。但是如果 we 想让机器向智能化的方向发展，就不能让机器完全听候人类的“摆布”，需要赋予它一定的自由度和主动权。生成式神经网络就是根据这个原理，利用“概率”这一数学工具，使机器能

够生成丰富多样的文本和图像。但也因为这个原因，就一定存在生成不合格和有害文本与图像的概率（可能性）。这是我们在赋予机器智能的时候所必须付出的代价，难以避免。

那么我们有没有可能通过给机器规定严格的伦理准则来限制它的错误行为？实际上，这也很难！不仅因为“伦理”的准则很难准确描述，但即便可以定义，也很难执行。举一个简单的例子，比如自动驾驶车（或无人车）行驶在普通的马路上，如果我们规定自动驾驶车必须严格遵守交通规则，这个“准则”应该是很明确的。但如果路上同时还有“有意或无意违反交通规则”的有人车和行人，自动驾驶车则无法行驶去完成自身的任务。比如，自动驾驶车需要向左并线以便左拐，由于左路车道上的车辆之间没有保持规定的车距，自动驾驶车就无法实现向左并线。这恰恰说明，自动驾驶车一方面要严格遵守交通规则，另一方面要完成达到目的地的任务，在不确定的交通环境下，这两项目标是难以兼顾的。可见，人工智能的发展必然带来对伦理和传统规范的冲击。

深度学习算法的不安全、不可信与不鲁棒，同时给有意的滥用带来机会。人们可以恶意利用算法的脆弱性（不鲁棒）对算法进行攻击，导致基于该算法的人

工智能系统失效，甚至做出相反的破坏行为。深度学习还可以用来造假——即所谓“深度造假”，通过AI的“深度造假”，可以制造出大量逼真的假新闻（假视频）、假演说（假音频）等，扰乱社会的秩序、诬陷无辜的人。

人工智能无论是被有意的滥用还是被无意的误用都需要治理，不过对这两者的治理性质上完全不同。前者要靠法律的约束和社会舆论的监督，是带有强制性的治理。后一种则不同，需要通过制定相应的评估标准和规则，对人工智能的研究、开发和使用过程进行严格的科学评估和全程监管，以及问题出现之后可能采取的补救措施等，帮助大家避免AI被误用。

从根本上来讲，人工智能的研究、开发与应用都需要以人为本，从公正、公平的伦理原则出发，做负责任的人工智能。为此，我们需要努力去建立可解释、鲁棒的人工智能理论，在此基础上，才能开发出安全、可信、可控、可靠和可扩展的人工智能技术，最终推动人工智能的公平、公正和有益于全人类的应用和产业发展。这就是我们提倡的发展第三代人工智能的思路。

人工智能研究和治理都需要全世界不同领域人员的参与与合作，除从事人工智能的研发和使用人员之外，还需要法律、道德、伦理等不同领域人员的参与。我们需要明晰伦理、道德的标准，什么是

符合“道德”和“伦理”的，不同的国家、民族、团体和个人都有不尽相同的认识，因此需要全球范围的合作，共同制定出一套符合全人类利益的标准。人类是命运的共同体，我们相信通过共同的努力，一定会找到符合人类共同利益的标准。只有人工智能的研究、开发和使用人员，人人都遵守共同制定的原则，才能让人工智能健康地发展并造福于全人类。○

来源：清华大学人工智能国际治理研究院

张钹简介



张钹，中国科学院院士、中国自动化学会会士，清华大学教授计算机科学与技术专家，俄罗斯自然科学院外籍院士，清华大学人工智能研究院院长、教授、博士生导师。张钹从事人工智能理论、人工神经网络、遗传算法、分形和小波等理论研究，提出了问题求解的商空间理论，曾获国家科技进步三等奖。

李德毅院士：认知物理学

——薛定谔、图灵和维纳的启示和超越

2023年1月3日，著名人工智能学家，中国工程院院士、欧亚科学院院士，中国人工智能学会名誉理事长李德毅在Science伙伴期刊Intelligent Computing发表前瞻性观点论文《认知物理学——薛定谔、图灵和维纳的启示和超越》。论文回顾了20世纪上半叶，控制论之父维纳（1894-1964）、量子力学之父薛定谔（1887-1961）和人工智能之父图灵（1912-1954）三位杰出学者为人类留下的五篇经典之作，并受其启发，展望未来以负熵为生、可交互、会学习、自成长的智能机器，为今后机器智能的发展奠定了基础和方向。

1936年，图灵提出图灵机模型，认为图灵可计算和一般递归函数等价，说明通过暴力计算，可以用足够大去逼近无限。1943年，薛定谔撰写著作《生命是什么？》，提出生命赖负熵为生，可以预见，如果机器赖秩序为生，机器便可作为认知的载体。1948年，维纳撰写《控制论》一书，研究生命体、机器和组织的内部或彼此之间的控制和通信。控制

和通信（或者称之为交互）为实现机器的具身智能提供了支持。1950年，图灵的《计算机器与智能》发表，提出了著名的用于测试机器智能程度的图灵测试。1950年，维纳撰写《人有人的用处——控制论和社会》一书，该书描述了社会领域的控制论问题，发现人有人的智慧、机器有机器的智能，可通过交互认知得以实现。

1. 从薛定谔谈开去，机器用时序和交互来维持负熵

薛定谔在《生命是什么？》一书中探讨生命的意义时，曾经写道：生命是很多种能够决定个体未来发展完整模式的密码本（code-script），人活着就是在对抗熵增规律，生命以负熵为生。那么如果机器可被视为生命，机器如何应对熵增？我们认为机器和生物一样，遵循最基本的物理定律。生命会衰老，以负熵为生；机器则需要依靠能量，形成秩序，通过和外界交互，产生负熵。

农耕时代的工具和工业时代的机器还不能以生命观来看待。

农耕时代的工具（如轮子、大刀等）里有实体物质和虚体结构，结构直接寄生在物质上，形成硬构体。工业时代的机器（如蒸汽发动机、摆钟等）里有物质、能量和结构，结构直接寄生在物质和能量上，形成硬构体。

到了智能时代，机器里有物质、能量、结构和时间。物质和能量是物理层面的真实存在，结构和时间是认知层面的抽象思维，用结构来思维物质在空间的拓扑（几何）关系，用时间来思维物质的运动变化、能量的流动转化。结构和时间寄生在物质和能量上成为硬构体（如集成芯片），机器里的信息则是大量的软构体（如程序），体现精神，它们寄生在硬构体或者已有的其他软构体之上，可自我复用，机器的秩序显示出维持自身和产生有序事件的能力，体现在认知和行为两个方面。

机器可作为思维的载体，时间是机器认知的奠基石。机器里有物质、能量、结构和时间。时钟依赖能量，时间依赖时钟，秩序依赖时间。软构体寄生在硬构体上，机器自举实现思维自动化，

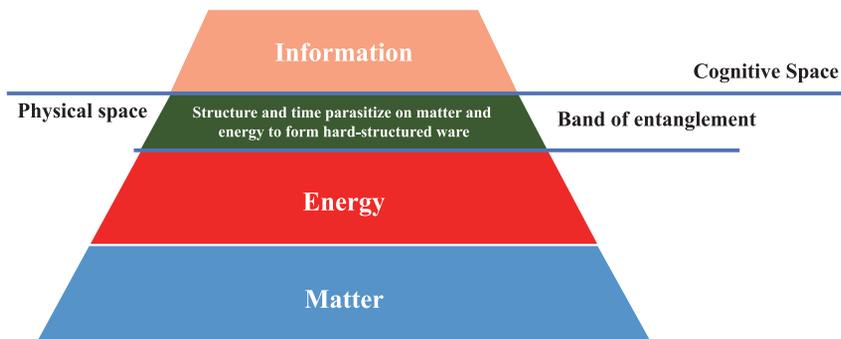


图1 智能机器的要素

自我复用实现认知自成长。机器运行靠程序，程序靠时序，软件靠交互，时序和交互产生负熵，机器赖负熵为生。时钟不停，与外界交互不息，则思维和认知不息。

2. 从图灵谈开去，图灵机的局限和图灵对人工智能的贡献

图灵究其一生研究计算机机械化，他于1936年发表的论文《论可计算数及其在判定问题中的应用》中，对可计算本质进行了严格数学定义，提出图灵机模型是一种十分简单但运算能力极强的计算装置，用来计算所有能想象得到的可计算数。后来著名的「邱奇—图灵论题」说明了「 λ 演算、递归函数和图灵可计算是等价的」，即：所有可计算或是机械程序可计算的函数，恰好就是用一般递归函数去逼近无限。可以认为「图灵可计算」是以软构体自我复用来逼近无限的过程。

图灵在1950年发表的文献《计算机与智能》中，开宗明义

地要大家考虑机器能否思维，他分析并驳斥了9种对思维机器的反对意见，主张教机器学习，只要机器在语言行为（对话）上和人没有明显差别，就应该算是能思维或有智能了，这就是后来著名的「图灵测试」。他还认为可以编制一个「儿童程序」，然后对其进行教育（指导学习、自主学习）以达到成人的智力水平。

但我们知道，学习是「感知—认知—行为」的循环过程，认知不可能独立于感知和行为，要发展机器智能就必须克服图灵机的两个局限：（1）孤立了认知，忽视了机器具身与环境的交互；（2）孤立了计算，忽视了记忆。

由此，我们可以得到更完整的关于智能的以下认识：

「智能」的定义：培养和传承解释解决预设问题的学习能力，以及解释和解决现实问题的能力。预设问题通常来自现实问题，并已经被证明存在行之有效的解决方案。学习是把未知变为可知，是解释和解决新问题的基础，解

释和解决现实问题是学习的目的，两者相互促进。

3. 从维纳谈开去，交互认知：具身行为的控制

控制就是追求熵减，通过负反馈来确保机器具有行为智能的稳定性。

这种控制在生物中的基础是神经系统。神经系统由神经细胞组成，对感知、认知起主导作用，通过和外界的信息交互获得负熵。对于人，尤其重要的是大脑皮质的进化；当然，还离不开双手的进化以及语言能力的进化，这是一个「感知—认知—行为」的闭环，对应三个区：瞬时记忆区、工作记忆区和长期记忆区。其中，瞬时记忆区处理不同感觉器官获得的感知信息，工作记忆区分析当前问题的态势，长期记忆区里留下了我们认知的结果。

机器中也有类似的控制模式，人类为机器赋予任务和行为特征。机器有感知、有认知、有行为，可交互、会学习、自成长，为了提升这些能力，就是在交互的环境里，由软构体里的算法规定的一组带有奖励和惩罚函数在内的、周而复始的复杂动作，通过迭代和交互完成在瞬时记忆区、工作记忆区和长期记忆区状态的改变。通过自动控制体现具身智能。这种强化学习，是人赋予机器使命的重要方法。

以深度学习为例，传统算法需要预写程序，通过编译才能获得智能，算法被长期困在程序里，而深度学习打破了这种局面，用标注代替人类记忆，从大数据中直接获得分类知识，使得记忆智能优于计算智能。由此可见，深度学习要在迭代和交互中完成。机器尤其擅长深度递归，无需担心它指数爆炸，排列组合试错，暴力计算是机器长项。

4. 迭代的智能：从学习到创造

机器思维离不开语言。孤立的机器自然语言处理是行不通的，需要实时的人机交互。语言能力是机器智能的杰出表现。机器用语言思维先要把语言形式化，只有形式化才可机械化，而后自动化。一旦自动化，思维的深度就一定会超越人类。当前由人赋予机器特定的语境和语用，

让机器写作，生成开幕词、闭幕词、新闻稿等已经屡见不鲜，机器写作的专业论文甚至已经被会议收录发表，机器作画、批改作业和试卷等已经常态化。图灵测试有可能区分不出来哪些是机器所为。未来机器的智能会在多样化和常态化的「图灵测试」中迭代成长。

自然选择并没有「过分」地进化人类这个物种，人类变得智能不是进化的必然，只是给了人类出生后认知二次扩张的潜力，正如同机器可以终身学习。可交互、会学习、自成长将成为智能机器的硬核（图2），智能机器要具备交互和迭代能力。自动驾驶汽车如果不能像人一样具备与时俱进的学习能力将很难获得驾照。

如果拿人和机器进行比较，人使用自然语言思维，机器使用

形式化语言思维并被机械地执行；人类思维有多样性和活跃性，机器思维有必然性和统一性；人类思维速度受限于生命的生化参数极限，而物理的机器则可暴力计算和暴力思维；人有丰富的想象力和创造力，机器有一以贯之的执着。总有一天出现机器工程师创造出新材料的配方，机器科学家提出新的假设，驱动产生新的科学发现。

最终，人工智能成为人类文明生态，使得人与机器可多通道、跨模态交互认知，让人发挥人的智慧、机器发挥机器的智能。用生命的观点看机器，人类构造了赖时序和交互为生、即赖负熵为生的智能机器；用物理学的观点看认知（图3），以物质、能量、结构和时间为基础，用软构体表达思维和认知，自我复用产生有序事件，维持不确定性中的基本

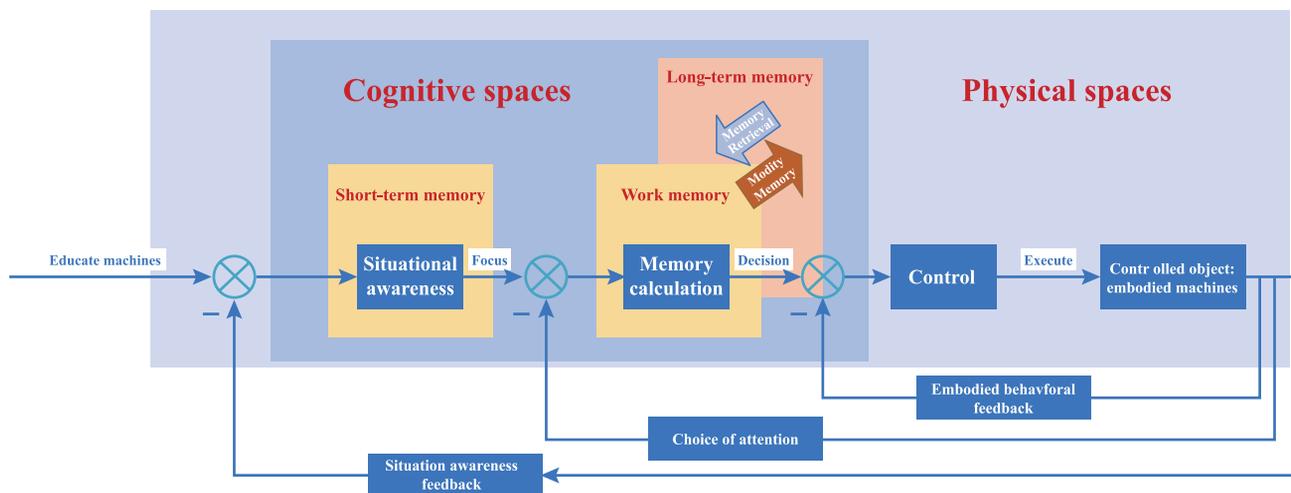


图2 可交互、会学习、自成长的机器运行流程图

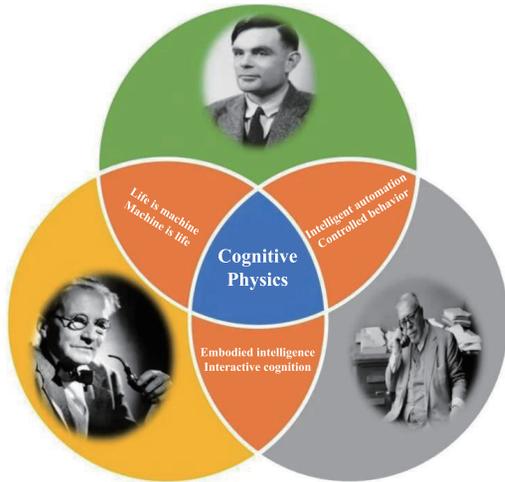


图3 认知物理学 - 薛定谔、图灵和维纳的启示

确定性，在人与机器之间，在物理空间和认知空间，形成交替认知的螺旋，会学习、自纠错、自组织、自成长，显示出解释、解决问题的认知能力，知行合一，形成迭代的智能。○

来源：Intelligent Computing

李德毅简介



中国工程院院士、欧亚科学院院士，中国自动化学会会士，指挥自动化和人工智能专家。1983年获英国爱丁堡海里奥特·瓦特大学博士学位。总参第61研究所研究员、中国指挥和控制学会名誉理事长、中国人工

智能学会理事长。

军事科学院系统工程研究院研究员，中国人工智能学会名誉理事长，中国指挥控制学会名誉理事长。中国改革开放以来第一位人工智能相关专业博士学位获得者，指挥自动化和人工智能领域专家。长期致力于信息化和智能化工作，不确定性人工智能领域的主要开拓者，中国无人驾驶的积极引领者，人工智能产学研发展的重要推动者，也是吴文俊人工智能最高成就奖获得者。

喜报
X I B A O

热烈祝贺中国自动化学会被中国科学技术协会评为“2022年度全国学会科普工作优秀单位”

根据《中国科协科普部关于开展2022年全国学会科普工作考核的通知》，中国自动化学会等85个全国学会被评为“2022年度全国学会科普工作优秀单位”。

学会将在后续工作中，积极构建全方位、多渠道的科普宣传平台，开展更多有吸引力的特色科普活动，为提高全民科学素养服务，承担起向社会公众宣传科技文化的责任。详情请查看：<http://www.caa.org.cn/article/191/3312.html>

视觉语言导航研究进展

文 / 司马双霖 黄岩 何科技 安东 袁辉 王亮

摘要：视觉语言导航，即在一个未知环境中，智能体从一个起始位置出发，结合指令和周围视觉环境进行分析，并动态响应生成一系列动作，最终导航到目标位置。视觉语言导航有着广泛的应用前景，该任务近年来在多模态研究领域受到了广泛关注。不同于视觉问答和图像描述生成等传统多模态任务，视觉语言导航在多模态融合和推理方面，更具有挑战性。然而由于传统模仿学习的缺陷和数据稀缺的现象，模型面临着泛化能力不足的问题。系统地回顾了视觉语言导航的研究进展，首先对于视觉语言导航的数据集和基础模型进行简要介绍；然后全面地介绍视觉语言导航任务中的代表性模型方法，包括数据增强、搜索策略、训练方法和动作空间四个方面；最后根据不同数据集下的实验，分析比较模型的优势和不足，并对未来可能的研究方向进行了展望。

关键词：视觉语言导航，视觉语言理解，跨模态匹配，具身智能

近年来，越来越多研究人员意识到单模态分析技术在现实中处理信息的局限性，对于自然语言、音频信息以及视觉等多模态融合方面的研究投入日^[1]益增加。视觉语言导航是智能体在第一视角下，基于真实环境下的全景图，综合处理指令和视觉信息并进行推理的多模态任务，也是智能管家等应用的核心技术之一。视觉语言导航尝试使用多模态融合的方式，为室内导航任务的研究提供了一个新的方向。如图 1 所示，智能体需要结合指令信息和视觉信息，在模拟器中完成一系列的决策，最终到达目标位置。其中主要难点在于如何学习理解

指令和视觉的信息，从而完成导航过程中的每一步决策。

Anderson 等^[1]于 2018 年首先提出视觉语言导航任务，并公开了与任务相对应的基于真实环境的 Room-to-Room (R2R) 数据集，并在 Matterport-3D^[2] 模拟器完成了导航任务的仿真。视觉语言导航一经提出便引起

了广泛的关注。随着研究的不断深入，研究人员提出了很多拓展任务，如室外视觉语言导航 Touchdown^[3]、结合导航和指称表达 (Remote embodied visual referring expression in real indoor environments, REVERIE)^[4-5] 以及视觉对话导航任务^[6]。除此之外，研究人员

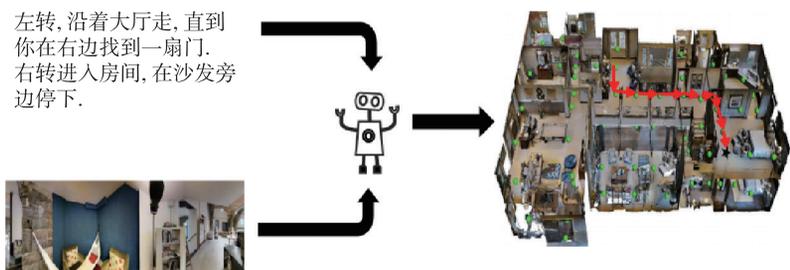


图 1 视觉语言导航过程示意图

也发现了一些视觉语言导航亟需解决的问题，如数据量级还远不能满足实际需求，数据稀缺问题愈发突出、以及模型的泛化能力低。这些问题一直阻碍着视觉语言导航的发展。另外现有的方法全是基于模拟器的仿真环境，该设定与现实场景下的导航仍存在很大差异。如何将现有的模型应用到实际环境中，也是视觉语言导航的一大难题。

为了解决以上问题，一系列的改进模型方法被相继提出。研究人员在指令集扩充、学习策略升级和多模态融合等方面进行较多探索并取得了巨大进展。本文首先介绍常用的数据集，然后按照不同的模型改进方式对现有的方法进行介绍，并结合不同数据集的实验结果分析不同模型的优势和不足，全面介绍目前视觉语言导航的研究现状。

1 视觉语言导航数据集

视觉语言导航的数据集，主要是针对 Matter-Port3D 模拟器的 90 个场景建立的。为了推动视觉语言导航任务的发展，研究人员从指令描述粒度、指令长度以及语言种类入手，收集了大量的人工指令。这一定程度上扩大了数据量，对视觉语言导航的发展，起着非常重要的作用。本节将按照不同的指令粒度和指令长度的数据集分类介绍。

1.1 R2R 数据集

R2R 数据集是由 Anderson 等^[1]构建，其中总词汇量约 3 100 个单词，构成 7189 条路径下的 21 567 条人工标注的指令，且每条指令的平均长度为 29 个单词。R2R 数据集在很大程度上覆盖了视觉环境中的大部分细节信息，具有多样性的特点。在视觉语言导航中，R2R 数据集被分成训练集、可见环境的验证集、不可见环境的验证集和测试集。其中训练集和可见环境的验证集共用 61 个真实场景，但是把相应场景下的数据集分为了两个部分：用于训练的 14 025 条指令和用于可见环境验证的 1020 条指令。不可见环境的验证集和测试集中并没有交叉重复的数据，不可见环境验证集使用 11 个真实场景和 2349 条指令，而剩余的 18 个真实场景和 4 173 条对应的指令构成测试集。

1.2 Fine-Grained R2R 数据集

由于 R2R 参考路径是由初始位置到目标位置间的最短路径构成，这在一定程度上影响了路径与指令的耦合度，同时缺少细粒度指令和视觉场景的对应关系。Hong 等^[7]提出了一种细粒度的子指令形式，对原先的 R2R 指令使用启发式算法生成相应的子指令，构造了细粒度的 Fine-grained R2R (FGR2R) 数据集。

FGR2R 训练集和验证集的每条指令平均可拆分为 3.6 条子指令，且每条子指令平均包含 7.2 个单词和 2.6 个对应的导航点。例如这条简单的指令：“左转，走上楼梯，进入卫生间”，对智能体，准确无误地理解它十分困难，必须对指令分解逐一理解每个词语的意思。这意味着将导航任务简化为多个子任务，每个子任务都有与其对应的子指令。此外，智能体不仅需要理解指令信息，而且需要对环境中的视觉物体进行识别。比如“走上楼梯”，直到检测到楼梯匹配到指令信息，才可以执行后续的动作。

1.3 R4R、R6R 和 R8R 数据集

由于 Room-for-Room (R4R)、Room-6-Room (R6R) 和 Room-8-Room (R8R) 数据集构建的思路一致，本节将介绍这 3 个数据集。在 R2R 数据集集中的路径普遍需要 4~6 个步骤完成，并且利用最短路径的方式到达目的位置。这样不利于评估指令和路径的匹配程度，因此需要一个包含更长路径的数据集来对导航中的动作与指令的一致性进行评价。文献 [8-9] 提出拼接 R2R 数据集指令的方式，由此形成更长指令长度的 R4R、R6R 和 R8R 数据集。由于指令长度和参考路径变得更长，从而增加了训练模型的难度。同时，先前的评价指标仅仅关注

是否到达目标位置，对指令和路径是否匹配并不敏感。因此针对长指令数据集，研究人员提出了一些新的评价指标和方法，来衡量和提高模型的泛化能力。

1.4 RxR 数据集

对于现有的 R2R 数据集中存在偏差、指令和路径的数量少等问题，Ku 等^[10]提出新的 Room-across-Room (RxR) 数据集，分别从指令语言种类、数据集规模、路径和指令的匹配粒度和姿态跟踪 4 个方面对数据集进行了拓展和改进。RxR 数据集包括 16500 条路径，且每条路径对应 3 条不同语种的指令，总词汇量高达 980 万条，构成 126 000 条指令。此外，He 等^[11]利用标志物信息，将 en-RxR 划分成短指令的形式，构建了 Landmark-RxR 数据集。相较于 R2R 数据集，RxR 数据集中指令对应的路径长

度更长，并且在指令和路径的匹配程度上更为一致。同时 RxR 采用对三种语言指令进行测试，可以避免对单一语种产生过拟合的现象。RxR 数据集中首次引入姿态跟踪的方式，即对比与人执行指令时所采取的动作和经过的位置。后续的工作将 RxR 数据集引入到连续环境的模拟器，以寻找更具有更加实用的模型。

除了以上的指令数据集，还有一些视觉语言导航拓展工作的数据集，如 REVERIE^[4]和 Bilingual Room-to-Room (BL-R2R)^[12]数据集。表 1 介绍了不同数据集的各项属性。

2 视觉语言导航模型

目前视觉语言导航所面临的两大难题：数据稀缺和模型的泛化性低，一直阻碍着该领域的发展。但随着越来越多研究人员投

入到视觉语言导航中，这些问题都不同程度地得到解决。我们将视觉语言导航模型分为数据增强、搜索策略、动作空间、训练策略 4 个方面来进行介绍。

2.1 基于数据增强的视觉语言导航模型

视觉语言导航是根据真实场景下的照片所构成的仿真环境和人工指令，进行一系列推理的过程。专业人员标注的指令不仅成本高，且数量十分有限，例如常用的 R2R 数据集中仅含有 21 567 条语言指令。因此，数据稀缺是视觉语言导航中的先天问题，不仅使得学习跨模态匹配更加困难，还在很大程度上限制了模型的性能。当前很多领域的研究已经证明了数据增强的有效性，特别是提升模型的性能有很大帮助。接下来，本节将从合成新指令和拼接旧指令两个方面介绍视觉语言

表 1 视觉语言导航不同数据集的对比

数据集	训练集 (条)	可见验证集 (条)	不可见验证集 (条)	测试集 (条)	平均指令长度 (单词个数)	语言种类
FGR2R ^[7]	51 377	3 775	8 481	15 385	7.2	英语
REVERIE ^[4]	10 466	4 944	3 573	6 292	18.0	英语
BL-R2R ^[12]	14 025	1 020	2 349	4 188	20.6	英语 / 中文
R2R ^[11]	14 039	1 021	2 349	4 173	29.4	英语
R4R ^[8]	233 613	1 035	45 162	—	58.4	英语
R6R ^[9]	89 632	—	35 777	—	91.2	英语
R8R ^[9]	94 731	—	43 273	—	121.6	英语
RxR ^[10]	79 467	8 813	13 625	24 164	77.8	英语 / 印地语 / 泰卢固语
Landmark-RxR ^[11]	133 602	13 591	19 547	—	21.0	英语

导航中的数据增强方法。

2.1.1 基于生成新指令的数据增强模型

Fried 等^[13]首先提出了一种数据增强的方式,如图2所示,“说话者”模型可以从视觉轨迹合成新的指令,拓展当前有限的训练指令集。而“跟随者”模型来检验生成指令,产生的轨迹可以作为“说话者”的输入,从而达到数据增强的目的。使用数据增强指令训练导航模型的方法,不仅可以显著提升模型性能,而且对提升泛化能力很有帮助。除此之外,“说话者”模型可以用来评价智能体导航路径的好坏。很多后续工作都在这种数据增强的基础上,再做相应的方法改进。Fu 等^[14]指出现有的数据增强模型性能表现并不理想,究其原因在于困难样本的导航成功率不高。针对这一问题,他们提出一种对抗训练的方式。该方法通过模型动态更新路径生成器,旨在提高困难样本的成功率,以进一步提升整体的泛化效果。而文献[15]

通过分析“说话者”模型生成的合成指令,发现大部分合成指令存在语句逻辑问题,并没有建立好和视觉环境之间的联系,反而引入了更多的误差。因此 Huang 等^[15]提出一种生成高精度指令的数据增强方式,通过设置一个判别器来评价生成数据的质量,同时引入一些负样本,以提高训练的鲁棒性。尽管模型的表现不错,但是依然没有解决有效的生成指令数量少的问题。由于缺少指令的评估指标,Zhao 等^[16]提出一个不需要参考指令的指令轨迹亲和模型。

另外,不少研究人员在导航环境方面做出新的尝试,并指出“说话者”模型训练时的有限环境数量,限制了指令的多样性。首先,Tan 等^[17]在“说话者”和“跟随者”(Speaker-follower, SF)模型的基础上,提出了基于环境的数据增强模型,即通过遮挡环境中的同类物体,进而产生新的环境;从这些环境收集新的路径,然后通过“说话者”模块

生成新的指令;最后利用这些数据微调模型。此外,An 等^[18]认为当前视角可能缺失指令中的关键物体信息,进而导致错误决策,于是提出邻近视角增强模型(Neighborview enhanced model, NvEM)。该模型使用当前视角的图像特征和相邻视角的图像特征,以扩大智能体的感受野。无论是从指令或环境入手,这些方法均是基于最短路径的原则来导航,这样会导致学习过程中出现依赖于训练时所做过动作,从而出现忽略重要语言信息和视觉信息的问题。为了解决该问题,文献[19]提出基于随机路径方式的数据增强。Yu 等^[19]基于“说话者”和“跟随者”模型,额外设置路径选择器动态地采样随机路径,并用“说话者”模块为这些路径生成相应的指令,然后再使用生成数据训练“跟随者”模块,最终达到随机路径形式的数据增强目的。这些方法都基于自主合成新指令的方式,但合成的指令与人类指令之间仍存在较大差异。主要原因是合成指令的细节不足和逻辑不通,从而导致了合成新的有效指令比较匮乏的问题。

2.1.2 基于拼接旧指令的数据增强模型

除了生成新指令的方式外,文献[8]提出拼接R2R数据集,来构成R4R数据集的方法,进

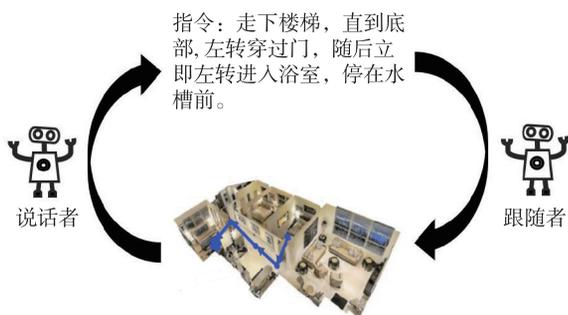


图2 “说话者”和“跟随者”^[13]模型的数据增强过程

而达到数据增强的目的。由于直接训练较长路径的模型比较困难，Jain 等^[8]提出模型先在较短路径下训练，然后再将模型迁移到较长路径的导航任务中。Zhu 等^[9]进一步将 R2R 数据集拓展到 R6R 和 R8R 长指令数据集，并提出一个记忆缓存来保存历史子指令和子轨迹对，同时使用模仿学习和课程强化学习进行两个阶段的训练。

尽管当前视觉语言导航的工作已经取得一定的进展，但是在提高视觉信息和指令耦合度方面，并没有很多突破性的工作。以往的研究验证了使用循环神经网络训练会存在长期依赖的问题，即当前状态会受一段时间之前的状态影响，这在长指令集训练过程中是无法避免的。无论是哪种数据增强手段产生的指令都存在偏差，以及有效指令和路径的数量少等问题。因此 Ku 等^[10]提出了新的 RxR 数据集，从路径轨迹采样方式、路径和指令的数量、路径和指令的粒度、语言种类四个方面对 R2R 数据集进行拓展和改进。

在视觉语言导航任务中，数据增强作为一种提升模型泛化能力的方法，一定程度上缩小模型在可见环境和不可见环境的表现差距。但是视觉语言导航中仍存在导航成功率低和过于依赖拓扑结构的问题，纯粹依赖数据增强

不能根本缓解以上问题。

2.2 基于改进搜索策略的视觉语言导航模型

早期的视觉语言导航任务采用的搜索策略是贪婪解码^[20]。因 MatterPort3D 平台将真实环境简化成离散点集，而导航过程需要连续地推理决策，来得到全局最优解，所以贪心算法的效果并不理想。Fired 等^[13]发现这一问题，提出在全景动作空间中将导航任务简化为加权无向图搜索方法。如图 3 所示，通过采用波束搜索^[21]的方式，能够选择多条备选全局路径进行打分来选择最优路径。这种搜索策略大幅提升导航成功率，但是存在路径过长、搜索效率低的缺点。为了改进以上的方法，研究人员提出带回溯的前沿搜索 (Frontier aware search with back tracking, FAST)^[22]和基于回溯机制的后悔模型^[23]，旨在降低搜索成本。回溯机制是在每次决策后及时评估，如果打分低，则选择回溯上一步，否则选择邻近未探索的节点。而 FAST 则是在此基础上，提出了一种局部信息和全局信息相结合的方式。该方法通过比较不同长度的局部路径，结合全局信号，利用异步

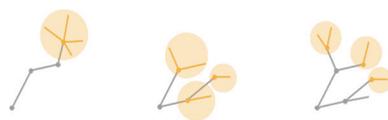


图 3 视觉语言导航任务中的不同搜索策略^[22]

搜索的方式来实现有效回溯。

由于存在指令语义模糊和局部视觉不确定性的问题，Wang 等^[24]提出一种动态决定探索方向、允许对下一步节点探索并进行预测的方法。但探索过多会导致轨迹长度过长和基于路径长度加权的成功率降低的问题，整体效果并不理想。而 Chi 等^[25]提出当智能体不知选择哪个方向时，采用辅助解惑的手段。当学习到的融合信息让智能体感到“疑惑”时，交互学习方式能帮助智能体解决导航中下一步移动的问题。而为了缩小训练和测试之间的差距，Deng 等^[26]提出可变图规划器模型 (Evolving graphical planner, EGP)，这是一种使用原始图像高效生成全局规划的方法。其通过动态地构建一个图的表示 (包括动作空间)，以便于更好决策。Hong 等^[27]则通过构建语言和视觉实体关系图模型 (Language and visual entity relationship graph, Relgraph)，更好地利用不同模态间的关系。同时他们提出一种消息传递算法，在图中的语言元素和视觉实体之间传播信息，然后将这些信息组合起来以确定下一步要采取的动作。为了更好地长期规划决策，Wang 等^[28]提出一种结构化场景记忆的模型 (Structured scene memory, SSM)，允许智能体对已探索的区域保留访问权力，然

后通过这种持久性的空间表示, 智能体在细粒度指令的辅助下, 在全局决策方面表现出色。

现有的导航策略都是尽可能地找到每步的最佳决策, 寻找一条从起始位置到目标位置的有效路径。虽然一定程度上会造成导航路径长度过长, 却对导航成功率的提升十分明显。此外, 随着研究的深入, 记忆机制和图结构的引入, 为视觉语言导航策略提供了不同的思路 and 方向。

2.3 基于不同动作空间的视觉语言导航模型

文献 [29] 首次按照动作空间划分不同的模型, 将依赖于模拟平台和导航图的模型称为高级动作模型, 对于直接预测下一个基础动作的模型称为低级动作模型。如图 4 所示, 图 4 (a) 表示低级动作空间的 5 个基础动作, 图 4 (b) 表示高级动作空间基于导航点的动作。本节根据动作空间的划分, 分别介绍高级动作空间和低级动作空间的视觉语言导航方法。

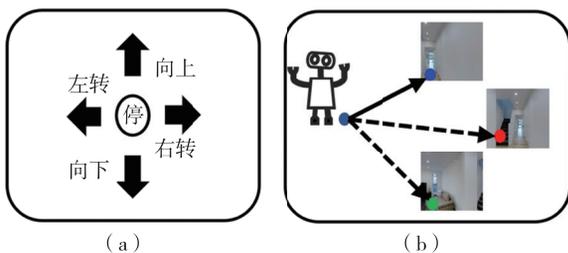


图 4 低级动作空间和高级动作空间表示^[29]

2.3.1 基于高级动作空间的视觉语言导航

早期 Fried 等^[13] 提出将 36 张不同仰角和水平偏角的图像合成一张全景图的方法, 后来该形式被通称为高级动作空间。在此高级动作空间中, 智能体只需选择邻近节点移动。高级动作空间不仅可以简化导航过程, 并且能显著提升导航成功率。在高级动作空间下, 文献 [30] 发现导航结果反馈模糊的问题, 即导航成功产生的反馈结果, 并不能反应指令和路径是否匹配。由此, Wang 等^[30] 提出强化跨模态匹配 (Reinforced cross-modal matching, RCM) 的方法来解决上述问题。利用推理导航器在局部区域内进行跨模态对齐, 再使用匹配评判器促进路径和指令

之间的全局匹配, 进一步强化模态融合效果和提高导航成功标准。Ma 等^[31] 提出自我监控智能体模型 (Self-monitor agent, SMNA)。他们根据模态匹配的关系, 认为“下一个动作的执行常常是由上一个动作完成与否决定的”, 并相应提出了视觉和语言联合对齐模型, 来监控导航进度。另外, 由于指令中含有丰富的实体描述和方向信息, Qi 等^[32] 提出物体和动作可知模型 (Object-and-action-aware model, OAAM), 分别对视觉特征和方向特征使用注意力机制, 最后再融合两部分特征。该方法充分利用指令中实体和方向信息, 来与视觉场景进行匹配, 最后设置路径损失来限制智能体仅沿着最短路径移动。

在视觉语言导航的设定中, MatterPort3D 模拟器是将场景划分为离散的可导航位置点集。这一做法简化导航过程为一个无向图的探索过程, 即每步移动都从邻近的有限点集中选择下一个目标节点。这在一定程度上减少了视觉信息对任务的影响。文献 [33] 指出在视觉语言导航和问答任务中, 不利用视觉信息的单模态模型的表现好于多模态模型, 模态融合反而造成了性能衰减。针对以上问题, Hu 等^[34] 提出在不同模态融合条件下, 对比“说话者”和“跟随者”模型^[13] 和自我监控智能体模型^[31] 的性能表现, 发现模型更容易利用几何拓扑结构信息, 而忽略了大量的视觉模态信息。模型对于拓扑结构的依赖一定程度降低了指令和视觉信息的耦合度。为了解决这个问题, Yu 等^[19] 提出改变最短路径为随机路径的移动策略, 旨在消除对于路径结构的依赖, 更多地专注语言和视觉之间的信息匹配。针对不同环境中的性能差异问题, Zhang 等^[35] 设计新的环境划分和特征替换的方案, 研究环境偏差的影响。

除了改变路径采样的方式之外, 另一个思路是回到低级动作空间。Anderson 等^[36] 尝试转移模拟环境训练的智能体到现实场景中, 并提出一个子目标模型来识别临近可达的节点。他们使用即时定位与地图构建和路径规划的方法, 建立智能体学习的高级动作

和智能体的低级动作的变换联系，将模型性能损失控制在可接受的范围内。但高级动作空间方面的迁移工作，还是受 Matterport3D 模拟器不能支持低级动作的影响，需要通过特定算法转换为低级动作，因此很多模型不便于直接由模拟环境转移现实场景中。

2.3.2 基于低级动作空间的视觉语言导航

相对于高级动作空间中选择邻近节点的移动方式，低级动作空间只包括六种基础动作：向上、向下、左转 30 度、右转 30 度、前进和停止。在这种动作空间下，模型在对于环境拓扑结构未知的情况下，直接预测智能体的动作。如图 5 所示，在视觉语言导航任务中，基于编码_解码的方法首先通过长短期记忆网络 (Long short-term memory, LSTM) [37] 编码器将指令编码 $[x_1, x_2, \dots, x_l]$ 和真实图像 $[v_1, v_2, \dots, v_l]$ ，映射到一个上下文的动作序列，再通过 LSTM 解码器融合编码后的语言特征和视觉特征预测每一步的动作 $[a_0, a_1, \dots, a_T]$ 。解码过程中额外加入了注意力机制，这一机制选择性关注视觉感知和当前指令中相关联的内容，帮助智能体结合环境选择相应的基础动作。通过建立紧密的模态间的信息联系，模型生成一系列的低级动作命令来指导完成导航任务。

高级动作空间中存在过度依赖已知的路径拓扑结构的问题，不利于未来部署在现实场景中。因此，不少研究人员开始关注更具有现实意义的模型，Landi 等 [38] 提出使用动态卷积滤波器的方法，模型基于当前指令信息动态地从视觉信息中提取相关信息，并输出低级动作空间下的动作概率。在此基础上，Landi 等 [29] 进一步提出类似 Transformer [39] 结构的感知转化移动模型 (Perceive, transform and act, PTA)，通过多次利用注意力机制的形式来融合模态间信息。实验证实了该模型同时兼容高级动作空间和低级动作空间。由于 MatterPort3D 平台环境本身的约束，新型的模拟环境平台相继被提出。首先是 FacebookAI 实验室在 2019 年提出的 Habitat 平台 [40] 和 Shen 等 [41] 在 2020 年提出的 iGibson 平台。

这些支持连续环境模拟器的推出，大大推动了基于低级动作空间方面的研究。文献 [42] 提出基于 Habitat 平台的连续环境下的视觉语言导航任务。不同于以往的高级动作空间方法，存在传送移动、依赖几何结构和精准定位的问题，Krantz 等 [42] 通过构建一个跨模态注意力机制的连续环境的视觉语言导航模型，验证了数据增强、数据聚合和进度控制对模型的积极作用。同时对比高级动作空间的视觉语言导航模型，他们发现先前的视觉语言导航模型中存在过多理想化的条件，在真实环境中的可行性有待验证。Chen 等 [43] 在基于连续环境的视觉语言导航模型的基础上，将其分解为两个阶段：计划和控制，在探索过程中，拓扑地图被建立用于导航规划。然后局部控制器接受导航规划并生成低级动作来完成导航任务。

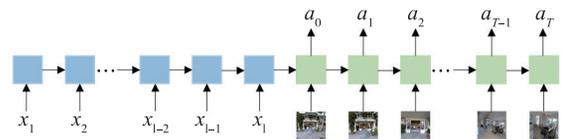


图 5 视觉语言导航中的 seq2seq 模型

无论在高级动作空间和低级动作空间中，现有的方法并未详细解释模型在模态融合后性能提升的原因。文献 [33] 对以往的模态融合方式提出了质疑，并建议以后的模型增加模态消融实验以佐证效果。为了更好地融合模态间的信息，Zhu 等 [44] 提出辅助推理导航模型 (Auxiliary reasoning navigation, AuxRN)。该模型通过四个辅助任务：动作解释、估计进度、预测方向和轨迹一致性评价，来提高模型的推理和环境感知的能力。由于指令间信息差异和指令中语义模糊的问题，Xia 等 [45] 编码相同轨迹的所有指令，其中每条指令互作补充，去提高模型的文本理解能力。在视觉语言导航中，模态间的联系并不是简单地合并指令和视觉信息，而是需要建立互为补充的关系，进一步提升模型的性能，并通过合理的实验证明在不同的动作空间下模态融合方式的有效性。

2.4 基于训练方法的视觉语言导航模型

视觉语言导航任务中常用的两种模型学习方式：监督学习和强化学习。监督学习是通过 R2R 数据集中的最短路径标注数据，学习得到一个优化的模型，进而预测不可见环境中的路径序列。而强化学习是把视觉语言导航任务看作一个马尔可夫决策过程。智能体在导航过程中观察周围环境并进行分析和反馈，并通过特定的奖励函数，尝试将学习到的经验知识应用到导航任务中，不断地进行试验，以达到完成视觉语言导航任务的目的。尽管这些方法的有效性得到很好的验证，但是各自的局限性也被揭露出来。研究人员在训练方法的选择上进行了更为深入的研究，当前预训练模型通过大规模数据训练，对视觉语言导航的性能提升十分明显。下面将对传统训练模型和预训练模型展开介绍。

2.4.1 基于传统训练的视觉语言导航模型

由于 R2R 数据集提供了参考路径，通过匹配预测动作分布和最佳路线，最初的视觉语言导航方法大多采用的是监督学习方式。文献 [1] 使用基于注意力机制的 LSTM 的序列到序列模型 (Sequence-to-sequence, seq2seq)^[46]，并结合“学生自学”^[47]的训练方法，对于先前的分布采用动作输出序列预测下一步动作，这是初期流行的一种基础方法。该方法使用交叉熵损失函数，学习标注数据的特征信息，泛化到未知环境中。但是由于人工标注的数据成本过于昂贵，最优路径并不容易获取。在后续的研究中，为了进一步提升在不可见环境下的泛化能力，Wang 等 [48] 提出使用强化规划 (Reinforced planning ahead, RPA) 的方法，将模型无关和基于模型两种强化学习联合在一起。其中展望模型结合了环境模型和策略模型，在 R2R 数据集上取得了不错的效果。最近的研究提出了很多新颖的学习方法，诸如 Wang 等^[30]提出一种自监督模仿学习的方法。通过训练，智能体可以根据过往的决策，学习产生多条可能的轨迹。模型利用最佳匹配的轨迹辅助训练，并优化轨迹的生成。文献 [17] 改进以往的方法，提

出将模仿学习和强化学习的损失结合作为一个损失函数，并用半监督学习的方式进行反向翻译和环境消除 (Environmental dropout, Envdrop)，分别为了训练额外的数据和生成未知环境。这种方法对模型的泛化能力进一步提升，如图 6 所示，模仿学习和强化学习结合的方法通过结合两种学习策略的优势，能够有效地提升模型的性能。Wang 等^[49]对于模仿学习的错误累积和强化学习中的奖励设计成本问题，尝试使用蒸馏的方法减少过拟合，提出软专家奖励学习模型 (Soft expert reward learning, SERL)。具体地，通过设置软专家蒸馏模块让模仿过程减少错误，同时利用自我感知模块让智能体一直向目的位置移动。研究人员在学习方法上的研究仍在继续，不断地优化方法策略。

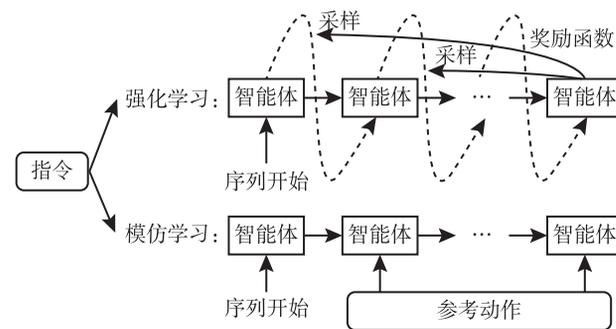


图 6 融合强化学习和模仿学习的过程

在视觉语言导航中，“学生自学”和“老师指导”^[50]训练策略的选择，是影响生成路径序列的一大因素。在视觉语言导航的设定中，因为导航路径长，并且采用离散动作的问题，所以早期的动作抽样工作是基于“学生自学”的方式。但由于全景动作空间的提出，路径长度被缩短了一大半，“老师指导”的方式开始流行起来。但是两者都存在曝光偏差^[51]的问题。一旦出现误差，则会导致大量累计误差，从而偏离正确路线。针对“学生自学”引入偏差的问题，文献^[9]提出了基于“学生自学”的模仿学习，保证轨迹和指令的一致性。为了充分发挥两者的优势，Li 等^[52]从课程抽样方式中得到灵感，提出随机动作采样的方式。具体地，

基于伯努利分布的抽样策略通过随机选择每一步的动作抽样方式,借此来保证利用“学生自学”和“老师指导”的优势,进而得到一种相对偏差较小的动作抽样方法。

目前,很多工作通过结合模仿学习和强化学习的方式,取得了较大性能提升。模仿学习学习老师的动作,而强化学习通过从奖励中采样动作,使智能体可以探索环境和提升泛化能力。如何更好地选择和利用训练策略,是提高导航模型泛化能力的一个关键因素。

2.4.2 基于大规模预训练的视觉语言导航模型

近年来,研究人员在如何提升智能体对不可见环境的泛化能力方面进行了不少尝试,包括预探索、数据增强和分析模态融合关系等方式。当前使用预训练模型提取特征的方式已经应用到各种任务中,其可以有效地提升下游任务的性能。受此启发,研究人员提出了在视觉语言导航中使用预训练模型,以解决泛化能力不足的问题。首先,文献[52]提出了使用 Bidirectional encoder representations from transformers (BERT)^[53] 等大规模预训练语言模型,来丰富指令表达。Hao 等^[54] 提出使用一种通用的预训练视觉语言导航智能体 (Pre-trained vision-and-language based navigator, PREVALENT), 并利用图像-语言-动作信息来进行预训练。实验证明预训练模型对提升模型的泛化能力很有帮助。后来 Huang 等^[55] 在 PREVALENT 的基础上,使用参数共享的方法来减低预训练模型的参数量。文献^[56] 中指出模型可以学习更多的语言知识,来提高推理的效率。此外, Hong 等^[57] 提出 Recurrent vision and-language bert for navigation (RecBERT)。这是一个多模态 BERT 模型,搭配时间感知递归函数,为智能体提供更丰富的信息。针对 RecBERT 会存在历史信息丢失的问题, Chen 等^[58] 提出 History-aware multi-modal transformer (HAMT), 将完整的历史信息编码保存,并设计了层次化的历史编码方法,降低计算复杂度。实验结果显示使用预训练的

语言模型分别在可见环境和不可见环境中的导航成功率高达 76% 和 66%, 不可见环境中的基于路径加权的成功率为 60%。进一步地证明预训练模型可以提高模型的泛化能力。

相较于传统训练方法,预训练模型引入了额外的知识表达,对视觉语言导航模型的提升十分显著。正因为高效的性能和强大的模态融合能力,如今预训练模型已经成为视觉语言导航模型的重要研究方向。

3 视觉语言导航方法的实验分析

第 2 节和第 3 节主要介绍了视觉语言导航的数据集和当前主要的模型方法,本节将对视觉语言导航的评价指标进行全面介绍,并结合 R2R 数据集、R4R 数据集和 RxR 数据集对比分析视觉语言导航模型。

3.1 视觉语言导航的评价指标

对于不同模型的评判,评价指标发挥着重要的作用,是衡量模型性能的关键性指标。随着视觉语言导航任务的发展,新的模型评价指标相继被提出。表 2 给出了视觉语言导航任务的评价指标,包括其定义和计算公式。这为第 3.2 节视觉语言导航模型性能比较提供帮助。视觉语言导航的评价指标不仅关注导航成功率 (Success rate, SR) 和路径长度 (Path length, PL), 而且需要对导航过程中路径轨迹和指令之间的一致性程度进行相应的度量评估。接下来将主要介绍目前的核心评价指标,其中基于路径加权的成功率 (Success weighted by pathlength, SPL) 的主要思想是将成功率和路径长度融合处理,来衡量导航的好坏。早期视觉语言导航模型的目标是尽可能地提高基于路径加权的成功率,来评估模型的性能。但它仅关注是否成功到达目标位置,而忽略了预测路径和参考路径的一致性问题。后续工作中提出的长度加权的覆盖分数 (Coverage weighted by length score, CLS)^[8] 和基于动态时间规整加权成功率 (Success rate weighted normalized dynamic time warping, SDTW)^[59] 两个评价指标,主要是

表 2 视觉语言导航任务中的评价指标

评价指标	定义	公式
路径长度	起始位置到停止位置的导航轨迹长度	$\sum_{v_i \in V} d(v_i, v_{i+1})$
导航误差	预测路径终点和参考路径终点的距离	$d(v_t, v_e)$
理想成功率	预测路径中任意节点距离参考路径终点的阈值距离内的概率	$\mathbb{I} \left[\left(\min_{v_i \in V} d(v_i, v_e) \right) \leq d_{th} \right]$
导航成功率	停止位置与参考路径终点的距离不大于 3 米的概率	$\mathbb{I} [NE(v_t, v_e) \leq d_{th}]$
基于路径加权的成功率	基于路径长度加权的导航成功率	$SR(v_t, v_e) \cdot \frac{d_{gt}}{\max \{PL(V), d_{gt}\}}$
长度加权的覆盖分数 ^[59]	预测路径相对于参考路径的路径覆盖率和长度分数	$PC(P, R) \cdot LS(P, R)$
基于动态时间规整加权成功率 ^[59]	由成功率加权的预测路径和参考路径的时空相似性	$SR(v_t, v_e) \cdot \exp \left(-\frac{\min_{w \in W} \sum_{(i_k, j_k) \in w} d(r_{i_k}, q_{j_k})}{ R \cdot d_{th}} \right)$

度量轨迹和指令一致性程度。长度加权的覆盖分数中包括两部分路径覆盖率 (Path coverage, PC) 和路径长度分数 (Length score, LS)。路径覆盖率表示与参考路径的一致程度, 其计算公式如下:

$$PC(P, R) = \frac{1}{|R|} \sum_{r \in R} \exp \left(-\frac{D(r, P)}{d_{th}} \right) \quad (1)$$

式中, R 代表查询路径, P 代表参考路径, r 是查询路径的位置坐标向量, d_{th} 是阈值距离。 $PC(P, R)$ 即为所计算的路径覆盖率。而路径长度分数则是评价预测路径和参考路径的一致性程度, 进而约束预测路径的长度, 产生与参考路径长度一致的预测路径, 计算公式为:

$$EPL(P, R) = PC(P, R) \cdot PL(R) \quad (2)$$

$$LS(P, R) = \frac{EPL(P, R)}{EPL(P, R) + |EPL(P, R) - PL(P)|} \quad (3)$$

式中, $EPL(P, R)$ 表示导航路径相对于参考路径覆盖范围的期望值, $PL(V)$ 表示路径长度, 表示路径覆盖率。 $LS(P, R)$ 即为所计算的路径长度得分。SDTW 是对预测路径和参考路径在时空相似性上的约束, 由导航成功率和路径一致性合并计算。

第 3.2 节将对比不同数据集下的视觉语言导航模

型, 通过以上主要的评价指标进行对比分析。

3.2 视觉语言导航模型的分析对比

表 3 和表 4 分别展示了不同模型在 R2R 数据集和 R4R 数据集上, 基于相应主要评价指标的实验结果。而表 5 以不同模型的主要创新点来划分模型方法, 包括数据增强、导航策略、动作空间和训练方法 4 个方向。表 5 中“√”表示属于对应分类的改进方向, 而—表示不属于对应分类的改进方向。

由表 3 和表 5 可知, 随着引入数据增强和改进导航策略之后, 在 R2R 数据集上, 视觉语言导航模型的 SR 和 SPL, 都较以往得到了不少的提升。文献 [13] 提出的全景动作空间形式和数据增强方法, 为视觉语言导航模型的快速发展, 提供了有力的支持。同时文献 [17] 在此基础上提出的融合强化学习和模仿学习的训练方法, 为后续的研究提供了参考模型。该方法的广泛应用对视觉语言导航任务的发展有重要的意义。此外, 最新研究发现预训练模型 BERT 和 Transformer 模型使智能体学习到更多有效的知识, 可以进一步提升导航性能。但值得一提的是, 预训练模型的训练时间和计算成本花销巨大。如何尽可能地降低计算成本, 设计一个轻量级的模型是亟待解决的问题。

表 3 在 R2R 测试数据集上的视觉语言导航方法对比

方法	路径长度 (米)	SR (%)	SPL (%)
seq2seq ^[11]	8.13	20.0	18.0
RPA ^[48]	9.15	25	23.0
SF ^[13]	14.82	35	28.0
SMNA ^[31]	18.04	48	35
PTA ^[29]	10.17	40.0	36.0
RCM ^[30]	11.97	43.0	38.0
Regretful ^[23]	13.69	48.0	40.0
FAST ^[22]	22.08	54.0	41.0
EGP ^[26]	—	53.0	42.0
PRESS ^[52]	10.77	49.0	45.0
SSM ^[28]	22.10	61.0	46.0
Envdrop ^[17]	11.66	51.0	47.0
SERL ^[49]	—	53.0	49.0
OAAM ^[32]	10.40	53.0	50.0
AuxRN ^[44]	—	55.0	51.0
PREVALENT ^[54]	10.51	54.0	51.0
RelGraph ^[27]	10.29	55.0	52.0
RecBERT ^[57]	12.35	63.0	57.0
HAMT ^[58]	12.27	65.0	60.0

表 4 在 R4R 测试数据集上的视觉语言导航方法对比

方法	SR (%)	CLS (%)	SDTW (%)
seq2seq ^[11]	25.7	20.7	9.0
SF ^[13]	23.8	29.6	9.2
FAST ^[22]	36.2	34.0	15.5
Envdrop ^[17]	29.0	34.0	9.0
Regretful ^[23]	30.1	34.1	13.5
RCM ^[30]	29.0	35.0	13.0
PTA ^[29]	27.0	35.0	8.0
OAAM ^[32]	31.0	40.0	11.0
RelGraph ^[27]	36.0	41.0	34.0
EGP ^[26]	30.0	44.0	18.0
RecBERT ^[57]	43.6	51.4	29.9
SSM ^[28]	32.0	53.0	19.0
HAMT ^[58]	44.6	57.7	31.8

表 5 视觉语言导航中的不同方法改进的对比

方法	数据增强	导航策略	动作空间	训练方法	R2R SR (%)	R4R SR (%)
Seq2seq ^[11]	—	—	—	√	20.0	25.7
RPA ^[48]	—	—	—	√	25.0	—
SF ^[31]	√	√	√	—	35.0	23.8
SMNA ^[13]	—	√	—	√	48.0	—
Regretful ^[23]	—	√	—	—	48.0	30.1
FAST ^[22]	—	√	—	—	54.0	—
PTA ^[29]	—	—	√	—	40.0	24.0
PRESS ^[52]	—	—	—	√	49.0	29.0
RCM ^[30]	—	√	—	√	43.0	29.0
Envdrop ^[17]	√	—	—	√	51.0	—
SERL ^[49]	—	—	—	√	53.0	—
OAAM ^[32]	—	—	—	√	53.0	31.0
PREVALENT ^[54]	—	—	—	√	54.0	—
EGP ^[26]	—	√	√	—	53.0	30.2
SSM ^[28]	—	√	—	√	61.0	—
RecBERT ^[57]	—	√	—	√	63.0	43.6
HAMT ^[58]	—	√	—	√	65.0	44.6

不同于 R2R 数据集的主要评价指标，R4R 数据集包括更长的轨迹，更注重指令和轨迹的一致性程度。因此，R4R 数据集将 CLS 和 SDTW 作为主要评价指标。由表 4 和表 5 可知，在导航成功率的评价指标上，R4R 数据集是明显低于 R2R 数据集。这是因为长指令的影响，导航的性能降低。由表 4 可知，在 CLS 和 SDTW 上，模型的表现并不尽如人意。其主要原因是模型过于注重是否到达目标位置，忽略了指令和轨迹一致性的比较。尽管模型在基于路径加权的成功率上有不错的表现，但这不能保证导航轨迹与指令内容一致。因此，研究人员开始转向子指令和子轨迹的研究，通过分段剪切长指令，对导航过程中的一致性问题展开一系列的研究。我们可以发现注重指令和轨迹一致性的模型，在主要评价指标上都有一定的性能提升。因此，如何更好地利用指令信息和视觉信息，是视觉

语言导航中的关键问题。

由表 5 中的对比可以看出, 早先的方法主要研究监督学习和强化学习的选择, 但整体表现并不好。随着数据增强以及模仿学习和强化学习的结合等方法的提出, 这使模型的性能得到了较大的提升。但其利用波束搜索会导致路径长度过长, 模型从而丢失现实的应用意义。导航策略的改进极大地推动了视觉语言导航的应用发展。随着研究的深入, 研究人员将眼光投入到更贴切现实的低级动作空间, 开拓新的研究方向。此外, 随着大规模预训练模型的兴起, 研究人员尝试将预训练模型引入到视觉语言导航任务中, 并取得不错的效果, 使导航性能得到大幅提升。

由表 3~5 可知, 随着各种各样的方法被提出, 视觉语言导航领域的发展更加多元化。不仅仅是模型在各项评价指标上有较大的提升, 更重要的是在细分研究方向上也得到了更多研究和关注。

4 未来展望

视觉语言导航是近年来在多模态领域中新兴的研究方向, 一经提出就受到大批研究人员的关注。随着研究的不断深入, 视觉语言导航在导航成功率和泛化能力上, 都得到了巨大的提升。研究人员通过数据增强手段, 生成新的训练数据。虽然这种方法对模型性能的提升很有帮助, 但并未解决泛化能力不足的问题。因此, 后续的研究开始着力于减少过拟合现象, 引入预训练模型。与此同时, 研究人员不断优化学习方式, 进一步地提出视觉文本对齐和回溯机制等辅助手段, 这些方法显著提升模型的泛化能力, 并取得较为理想的性能。

但是, 目前仍有一些问题亟待解决: 1) 当前提出的一些方法, 受到了仿真环境平台和数据集的限制。从低级动作空间到高级动作空间的转换, 采用波束搜索的方式, 简化了导航过程。尽管各方面的性能表现均令人信服, 但脱离现实, 模型难以迁移部署到现实环境中。如何贴近真实场景、赋予模型更多现实的应

用意义, 这需要视觉语言导航研究的重心重新转移到低级动作空间上。尽管已有基于连续环境的视觉语言导航模型, 但是其导航效果并不理想。因此视觉语言导航需要结合传统机器人技术, 进一步优化目前在模拟环境中训练的模式。2) 有实验表明视觉语言导航的模态融合方法会对模型性能产生负面作用。究其根本是模态信息之间关系模糊, 并未形成良好的互补, 模型不能有效地利用多模态信息。目前, 大规模预训练模型可以更好地利用多模态信息, 获得不错的性能表现。但是由于计算能力不足和时间开销过大, 这为视觉语言导航的发展带来新的问题。因此对多模态信息学习的研究, 仍有很大的发展空间。3) 在视觉语言导航任务中, 数据稀缺的问题尤为明显, 这是限制性能的一大阻碍。尽管研究人员提出了利用机器生成合成指令的方法, 但这些合成指令大部分是有缺陷的, 且不符合人们的语言习惯。同时不少研究人员重新搜集数据, 从规模和指令长度等方面进行拓展, 获得更接近现实场景的指令集。由此可见, 数据方面的研究工作一直都是视觉语言导航中的重要内容。

在现实场景中的导航过程是动态连续的, 而非简单的无向图探索过程, 目前不少研究人员重新投入到连续空间的视觉语言导航研究, 即在低级动作空间下, 智能体经过一系列的基础动作, 完成视觉语言导航任务。当前 Habitat 平台和 iGibson 平台都支持连续的导航。智能体可以通过低级动作完成导航, 这就为以后应用到现实场景提供了更多的可能性。此外, 由于多模态任务的输入复杂多变, 机器人和人类的理解能力差异较大, 所以 BERT 等预训练模型的引入, 为智能体提供丰富的额外知识, 有助于理解模态信息和模态间的融合。总体总之, 视觉语言导航任务无论是在现实中的应用, 以及数据获取方面的研究, 未来还有很长的路要走。

5 结束语

视觉语言导航是一种多模态理解任务, 在未来智

能家具、娱乐、养老等国计民生领域有较大应用需求。本文详细介绍了视觉语言导航任务近年来的发展，首先对于各种主流模型进行了简要介绍，然后对提升模型泛化能力的方法进行了综述，分别包括模态间的分析、指令集等拓展方式以及搜索策略、训练方式和预训练模型等辅助策略。尽管视觉语言导航任务近年来取得了快速的发展，但是随着研究的深入，也凸显出

各种约束和限制，这需要更多研究人员投入后续的工作研究。研究人员不仅要追求性能上的突破，而且要赋予模型更多的应用价值。

希望通过本文可以让更多人了解视觉语言导航任务，吸引更多人投入其中，促进其发展。○

来源：自动化学报

作者简介



司马双霖 中国科学院自动化研究所智能感知与计算研究中心硕士研究生。2020年获郑州大学学士学位。主要研究方向为视觉语言导航和具身智能。



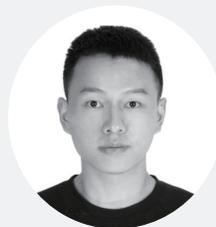
黄岩 中国科学院自动化研究所智能感知与计算研究中心副研究员。2017年获中国科学院自动化研究所博士学位。主要研究方向为计算机视觉和跨模态数据分析。



何科技 中国科学院自动化研究所智能感知与计算研究中心博士研究生。2019年获南京邮电大学学士学位。主要研究方向为视觉语言多模态和机器人



袁辉 中国科学院自动化研究所智能感知与计算研究中心机器人算法工程师。2021年获湘潭大学硕士学位。主要研究方向为视觉语言理解和机器人导航。



安东 中国科学院自动化研究所智能感知与计算研究中心博士研究生。2019年获北京大学学士学位。主要研究方向为计算机视觉和具身智能。



王亮 中国科学院自动化研究所研究员。主要研究方向为模式识别，计算机视觉，机器学习和数据挖掘。本文通信作者。

面向夜间疲劳驾驶检测的改进 Zero-DCE 低光增强算法

文 / 黄振宇 陈宇韬 林定慈 黄捷

摘要：为了提高夜间疲劳驾驶检测的准确率，在现有低光增强算法 Zero-DCE (Zero-Reference DeepCurve Estimation) 的基础上，提出改进 Zero-DCE 的低光增强算法。首先，引入上下采样结构，减少噪声影响。同时，引入注意力门控机制，提高网络对图像中人脸区域的敏感性，有效提高网络的检测率。然后，针对噪声相关问题，提出改进的核选择模块。进一步，使用 MobileNet 的深度可分离卷积替换 Zero-DCE 的标准卷积，提高网络的检测速度。最后，通过人脸关键点检测网络和分类网络，判断驾驶员的疲劳状态。实验表明，在夜间环境下，相比现有的疲劳驾驶检测算法，文中算法在人脸检测的准确率和眼睛状态的识别率上都有所提升，取得较令人满意的检测效果。

关键词：疲劳检测，低光增强，核选择模块，注意力门控机制，Zero-DCE (Zero-Reference Deep Curve Estimation)

疲劳驾驶是导致重大车祸的主要原因之一，严重威胁交通安全。美国高速公路安全管理局 (National Highway Traffic Safety Administration, NHTSA) 估计，美国每年因疲劳驾驶造成近 10 万起车祸，1 500 多人死亡和 71 000 人受伤^[1]。在中国，据保守估计，2015 年因疲劳驾驶造成约 1271 起交通事故，677 人死亡，1 600 人受伤^[2]。据美国国家安全委员会 (National Security Council, NSC) 报告，每 25 名成年驾驶员中约有 1 名驾驶员表示在过去的 30 天内出现过开车时睡着的情况，大多数驾驶员表示常在睡眠不足的状态下开车。虽然，一些机构、部门已采取措施限制疲劳驾驶现象，但是，对于像长途汽车驾驶员之类的工作者，由于工作原因，经常需要连续行车，更容易出现疲劳状况。在疲劳状态下，由于人体生理机能的下降，驾驶员往往会表现出反应过早或过慢、修正时间不准等不安全操作，对周围生命和环境造成严重威胁^[3]。因此，十分必要对驾驶员进

行疲劳检测。

现有的疲劳驾驶检测方法主要有基于车辆状态特征的方法^[4]、基于驾驶员心理生理参数的方法^[5]和基于驾驶员面部特征检测的方法^[6]。基于车辆状态特征的方法受限于特定的驾驶条件且实际使用时的检测结果不够稳定。基于驾驶员心理生理参数的方法要在驾驶员身上佩戴各种信号检测的设备，造成驾驶体验不舒服，且信号检测设备高昂。因此，这两种方法难以有效推广。近年来，随着深度学习技术的不断进步和深度学习硬件设备性能的逐步提升^[7]，基于驾驶员面部特征检测的方法得到广泛使用。然而，由于夜间环境光照强度较低，上述检测算法难以有效提取人脸关键点位置信息，导致降低检测效果。

传统的低光增强算法包括基于直方图均衡化的算法^[8-9]、基于 Log 变换的算法、基于伽玛变换的算法、基于拉普拉斯算子的算法和基于 Retinex 理论的算法^[10]。基于直方图均衡化的算法调整输入图像的

直方图分布状况以增强图像亮度，但是，由于直方图均衡化是对整幅图像进行处理，增强后的图像会存在局部过曝或过暗的问题。基于 Log 变换的算法和基于伽玛变换的算法虽然可提高低光图像的亮度，降低图像的对比度，但会造成图像的失真和噪声。基于拉普拉斯算子的算法可突出图像中的细节，但同时也会带来更多噪声。基于 Retinex 理论的算法将输入图像看作入射图像和反射图像的乘积，通过求解反射图像的方程对输入图像进行调整。

尽管传统的低光增强算法在某些情况下可产生令人较满意的结果，但图像增强后的质量往往取决于人工的先验知识和正则化，在一些复杂的情况下很难进行调优，且基于 Retinex 理论的算法会导致增强后的图像色彩失真。此外，由于传统低光增强算法复杂的优化过程，推理时间一般较长，不适合用于实时任务。

随着深度学习的发展，低光图像增强领域出现许多优秀的算法。Guo 等^[11]提出 LIME (Low-Light Image Enhancement via Illumination Map Estimation)，在图像的 RGB 三通道中找到最大值，用于估计每个像素的照度，再通过先验结构作用于像素上，完善初始图像，得到增强图像，但是经 LIME 增强后的图像色彩不太理想。Lü 等^[12]提出 MBLLEN (Multi-branch Low-Light Enhancement Network)，解决低光图像中存在的亮度、噪声、对比度等问题，但对图像的处理时间较长。近年来，研究者们开始将 Retinex 理论与深度网络结合^[13-14]。Zhang 等^[15]提出 KinD (Kind-ling the Darkness)，针对仅提高黑暗区域的亮度会带来伪影的问题，将原始图像空间分为两个较小的子空间，通过不同曝光条件下的成对图像进行训练。欧嘉敏等^[16]在 Retinex-Net 的基础上引入颜色损失函数，保留图像的低频部分，减少颜色失真的影响，但是网络参数量较大，难以实时处理图像。

同时，一些研究者开始将 GAN (Generative Adversarial Networks) 技术应用于低光图像增强

任务^[17]。基于 GAN 技术的算法利用非配对训练数据，一定程度上解决低光增强领域没有真实配对训练数据的难题，但这些非配对训练数据需要仔细选取以保证训练数据具有相同或相似的数据分布，并且基于 GAN 技术的深度学习网络模型在训练过程中十分不稳定。Guo 等^[18]提出 Zero-DCE (Zero-Reference Deep Curve Estimation)，通过深度曲线估计消除配对数据与非配对数据的需求，将低光增强问题转换为图像特定曲线估计问题，训练速度较快。但是，Zero-DCE 在训练过程中需要始终保持图像大小不变，使网络难以处理噪声，给夜间疲劳驾驶检测任务带来挑战。

综上所述，在夜间环境下进行疲劳驾驶检测，不仅需要实时处理相机拍摄的图像，还需要保证疲劳驾驶检测的准确性。因此本文在 Zero-DCE 的基础上，提出改进 Zero-DCE 的轻量级低光增强算法 (Improved Zero-DCE)。首先，使用上采样结构减少噪声影响。同时，引入注意力门控机制^[19]，使网络更关注图像中的人脸区域，提高网络的检测率。然后，针对噪声相关问题，提出改进的核选择模块。进一步，使用 MobileNet^[20]的深度可分离卷积操作替换 Zero-DCE 的标准卷积操作，减少网络的计算量。最后结合人脸关键点检测网络和分类网络，判断驾驶员的疲劳状态。

1 Zero-DCE

Zero-DCE 将一幅低光照图像作为输入，通过 DCE-Net 学习曲线参数图，然后通过亮度增强曲线对输入图像进行像素级调整，经过多次迭代获得增强后的图像。Zero-DCE 框架如图 1 所示。

图像亮度增强后的结果为：

$$LE_n(x) = LE_{n-1}(x) + A_n(x)LE_{n-1}(x)(1 - LE_{n-1}(x)),$$

其中， x 表示图像像素坐标， n 表示迭代次数， A_n 表

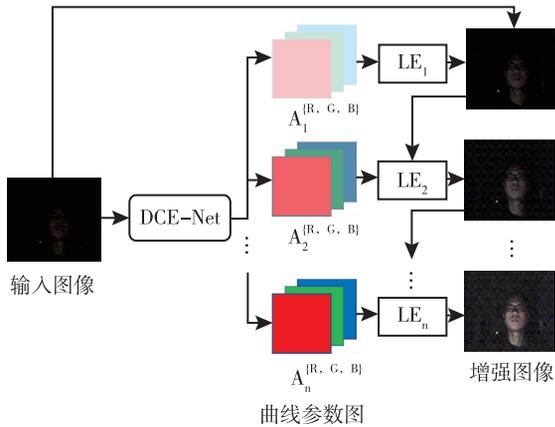


图1 Zero-DCE 框架图

示曲线参数图，大小与输入图像大小相等。通过上式，输入图像的每个像素都被赋予一条最优化的高阶曲线，使其能动态调整亮度。

为了得到合适的曲线参数图，通过神经网络模型进行训练，DCE-Net 结构如图 2 所示，整个网络模型由 7 个卷积层和激活函数组成。在整个特征提取过程中，图像的高度和宽度始终保持不变，仅在通道维度上进行改变。

首先，输入 1 幅 640×640 的图像，通过第 1 个卷积操作，将输入图像的通道数扩充到 32 维，然后再通过 3 个卷积操作，进一步提取特征。为了进一步加强特征提取，在第 5~7 个卷积操作前，先进行特征通道拼接操作，使通道维度扩充到 64 维。最后，使用卷积加激活函数操作将特征通道数压缩到 24 维，

包含 RGB 三通道各 8 个维度。

为了使网络在零参考信息的情况下完成训练，本文采用一系列非参考损失函数，包括空间一致性损失、曝光控制损失、彩色恒常损失和亮度平滑损失。下面简单介绍各个损失函数。

空间一致性损失函数是为了确保图像增强前后的局部区域与其周围像素的关系保持一致，具体公式如下：

$$L_{\text{spa}} = \frac{1}{K} \sum_{i=1}^K \sum_{j \in \Omega(i)} (|Y_i - Y_j| - |(I_i - I_j)|)^2,$$

其中， K 表示局部区域的数量， I 表示图像增强前的局部区域平均值， Y 表示图像增强后的局部区域平均值， $\Omega(i)$ 表示以区域 i 为中心的上下左右四个邻近区域，本文将每个区域 i 的大小定为 4×4 。

曝光损失函数表示平均强度值与理想曝光值间的距离，使图像增强后具有良好的曝光值。具体公式如下：

$$L_{\text{exp}} = \frac{1}{M} \sum_{k=1}^M |Y_k - E|,$$

其中， Y 表示图像增强后局部区域的平均强度值， E 表示理想状态下 RGB 颜色空间的灰度级，设置为 0.6， M 表示 16×16 的不重叠区域的数量。曝光损失函数有效控制每个局部区域的平均强度值与理想曝光值 E 之间的距离。

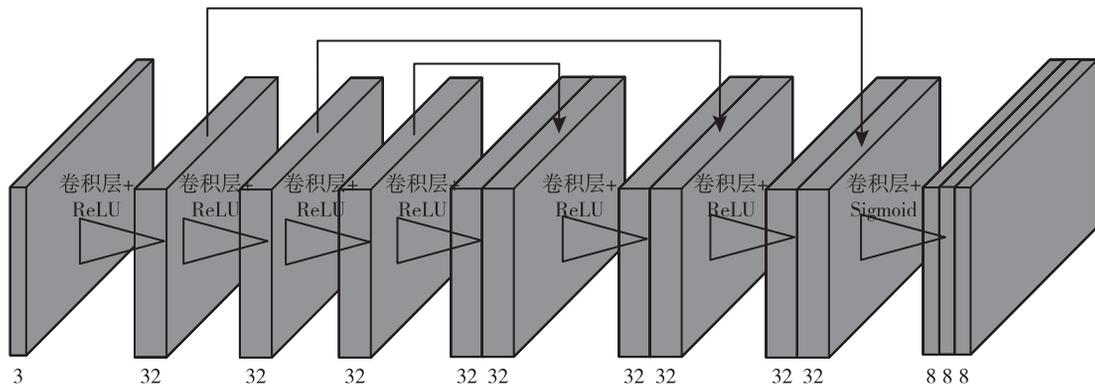


图2 DCE-Net 结构图

彩色恒常损失函数利用灰度世界的原理^[21],修正增强图像中潜在的颜色偏差,并在RGB三通道间建立调整关系,保证图像在增强后,RGB三通道的平均值尽可能接近,具体公式如下:

$$L_{col} = \sum_{\forall (p,q) \in \varepsilon} (J^p - J^q)^2,$$

$$\varepsilon = \{(R,G), (R,B), (G,B)\},$$

其中, J^p 、 J^q 分别表示通道 p 、 q 的平均强度值, (p, q) 表示属于 ε 的一组通道。

亮度平滑损失函数主要是为了保持邻近像素点间的单调关系,使相邻像素的曲线参数值尽可能接近,平滑增强后的图像。具体公式如下:

$$L_{tvA} = \frac{1}{N} \sum_{n=1}^N \sum_{c \in \xi} (|\nabla_x A_n^c| + |\nabla_y A_n^c|)^2,$$

$$\xi = \{R, G, B\},$$

其中, N 表示迭代次数, A_n^c 表示各通道的曲线参数图, ∇_x 表示图像水平梯度, ∇_y 表示图像垂直梯度, ξ 表示RGB颜色三通道。

因此,总损失函数可表示为

$$L_{total} = L_{spa} + L_{exp} + W_{col} L_{col} + W_{tvA} L_{tvA},$$

其中,权重系数 W_{col} 和 W_{tvA} 用于平衡各损失函数间的比例。

2 改进 Zero-DCE 的低光增强算法

2.1 算法总述

为了克服 Zero-DCE 在夜间环境下难以处理噪声,导致疲劳检测器出现人脸误检和漏检的问题,本

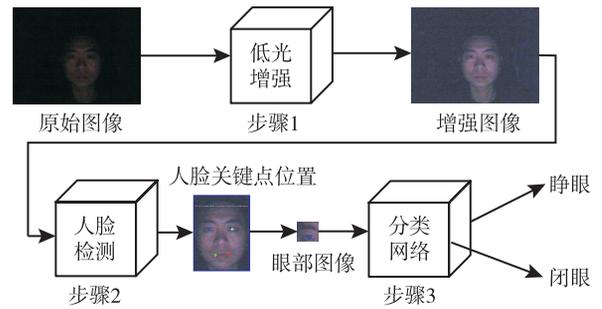


图3 本文算法流程图

文在 DCE-Net 的基础上进行改进,提出改进 Zero-DCE 的轻量级低光增强算法。整体流程如图3所示。整个过程分为3步。1) 将低光照的人脸图像输入低光增强网络,得到增强后的图像,2) 将增强后的图像通过 RetinaFace^[22] 人脸关键点检测网络检测人脸区域,并根据“三庭五眼治的先验知识截取人眼区域图像,3) 将眼部图像输入分类网络中,进行疲劳状态识别。考虑到 RetinaFace 与 MobileNetV2 检测人脸、人眼步骤的复杂性,结合 RetinaFace 与 MobileNetV2,组成人脸检测网络。然后,将低光增强网络放在人脸检测网络之前,结合二者,形成端到端的夜间疲劳驾驶检测模型。改进后的 DCE-Net 网络结构如图4所示。

2.2 核选择模块

针对噪声相关问题,当图像受到强噪声影响时,会丢失噪声周围的局部信息,此时,噪声处的上下文信息将较有用。使用不同尺度的感受野可充分利用空间特征,进行多尺度特征融合,挖掘丰富的上下文信息。在多尺度特征融合方面,现有的大多数方法

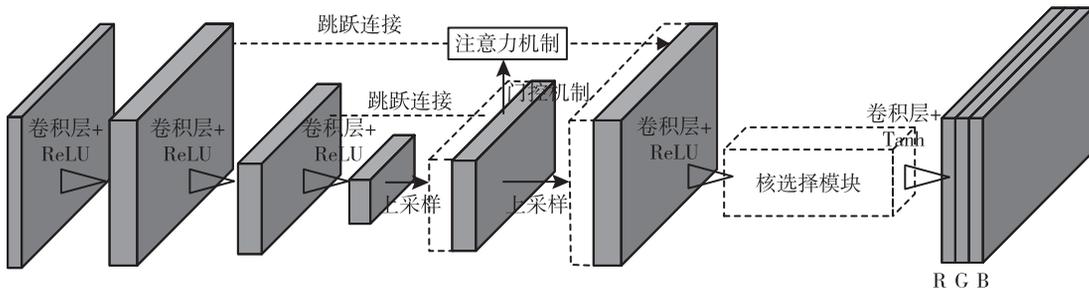


图4 改进的DCE-Net结构图

都是基于特征金字塔结构,以元素相加或串联的方式组合特征。虽然可结合不同尺度的特征图,却忽略不同尺度特征的空间和通道特异性。因此,基于特征金字塔结构的方法不能自适应地表达多尺度特征。受到 SKNet (Selective Kernel Networks)^[23]

的启发,在特征融合阶段,提出改进的核选择模块 (KernelSelecting Module),可自适应地调整感受野大小,动态选择合适的路径,特别是针对夜间场景下噪声图像中不断变化的噪声分布。改进的核选择模块结构如图 5 所示。

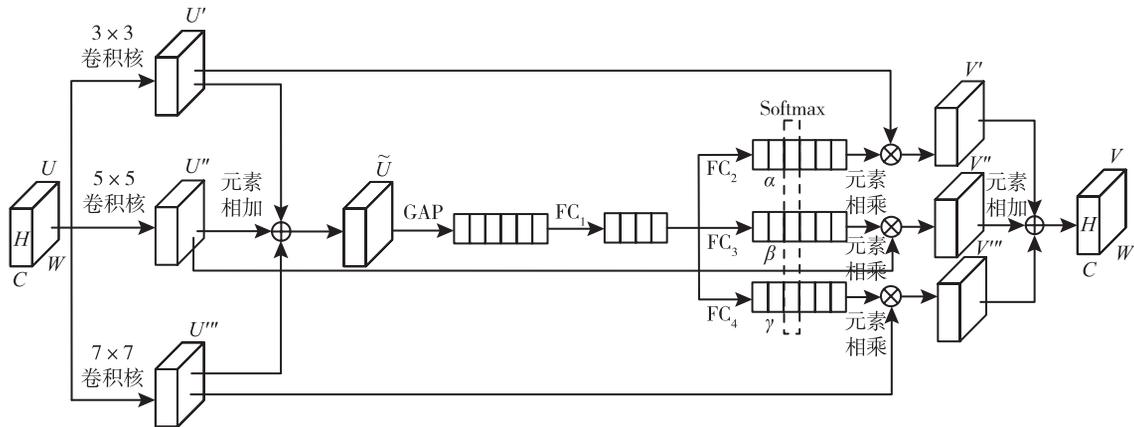


图 5 核选择模块结构图

针对输入特征图 U , 选择 3 种不同大小的卷积核进行处理, 卷积核大小分别为 3×3 , 5×5 和 7×7 , 得到 U' , U'' 和 U''' , 将三者相加得到 \tilde{U} 。然后将 \tilde{U} 通过全局平均池化 (Global Average Pooling, GAP) 和全连接层 (Fully Connected Layer, FC), 进行压缩和扩张, 得到 α , β 和 γ 三个权重系数。下一步, 使用 Softmax 激活函数, 得到各通道对应的权重值:

$$k_c = \frac{e^{\alpha_c}}{e^{\alpha_c} + e^{\beta_c} + e^{\gamma_c}}, \mathbf{k} = \alpha, \beta, \gamma,$$

其中, α , β , γ 分别表示特征图 U' , U'' , U''' 的软注意力向量, α_c 表示 α 的第 c 个元素, β_c 和 γ_c 类似。

最后将不同大小卷积核处理的特征图与其相应的软注意力向量相乘, 得到最终的输出特征图:

$$V = \alpha_c \cdot U' + \beta_c \cdot U'' + \gamma_c \cdot U'''.$$

2.3 上采样结构

针对 DCE-Net 在特征提取过程中难以处理噪声

而导致人脸检测的误检现象, 本文引入上采样结构模块, 降低噪声对输出结果的影响。网络结构如图 4 所示。首先, 输入一幅 256×256 的 RGB 三通道原始图像, 通过第 1 个卷积操作, 将图像的通道数扩展到 16 维。为了减小噪声的影响, 使用两次卷积加激活函数操作, 将特征图的大小压缩至原图大小的 $1/16$ 。之后, 通过两次上采样操作, 将特征图大小恢复成原图大小。同时, 为了弥补两次卷积加激活函数操作造成的信息丢失, 借鉴 U-Net^[24] 的跳跃连接, 连接浅层网络和深层网络, 保留浅层网络的信息, 使深层网络的特征更丰富。

2.4 注意力门控机制

在上采样结构中引入跳跃连接, 将深层网络上采样后的特征图与上一层的特征图进行拼接, 虽然可避免浅层网络在下采样过程中丢失空间信息, 但是由于浅层网络有许多冗余的特征, 在特征提取过程中会引入噪声。为了使上述的特征提取网络能提取更关键的任务目标信息, 本文引入注意力门控机制模块, 结构

如图 4 所示。门控机制用于激活图像空间信息的栅格信号，使跳跃连接可融合不同比例的特征，丰富语义信息。注意力机制通过学习的方式获取每个通道的特征权重，再根据权重值加强关键特征的效果并抑制与人脸区域不相关的特征。

门控机制结构如图 6 所示。首先，将特征图沿通道维度进行压缩，特征图的宽高保持不变，然后通过 ReLU 激活函数，激活具有正值的神经元，使神经网络更具稀疏性，加快模型的计算效率。门控机制输出特征图：

$$g = \delta (W_i^T g_i + b_i)$$

其中， b_i 表示偏差项， $\delta (\cdot)$ 表示 ReLU 激活函数， g_i 表示输入特征图， $W_i \in \mathbf{R}^{C_{int} \times C_{out}}$ 表示线性变换。门控机制对每个像素点进行操作，有助于模型动态调整输入图像的亮度。

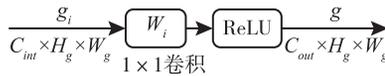


图 6 门控机制结构图

注意力机制结构如图 7 所示。首先，将通过门控机制得到的特征图与浅层特征图分别进行 1×1 卷积

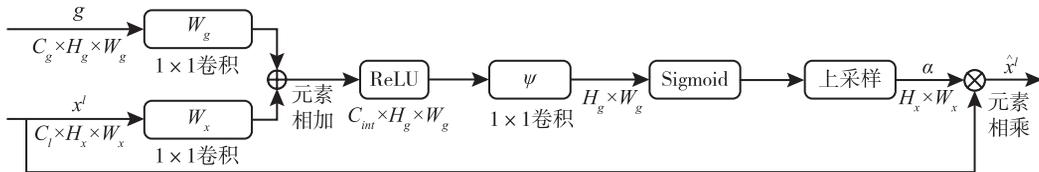


图 7 注意力机制结构图

文在实验中发现，在特征提取网络的浅层跳跃连接中引入注意力门控机制，可使网络学习到各个通道的权重系数，有效选取那些有利于低光增强任务的特征，抑制周围环境的噪声特征，降低后续夜间疲劳驾驶检测的难度。

2.5 深度可分离卷积

进一步，为了使疲劳检测网络达到实时性的需

求，转换为通道数为 C_{int} 的特征图，将 2 个特征图进行加法操作。再将相加后的特征图通过 ReLU 激活函数，舍弃全部负数的特征值，通过 1×1 卷积操作，对特征图进行通道维度上的压缩，使用 Sigmoid 激活函数，将特征值控制在 0~1 范围内。进一步，为了融合粗粒度特征图的上下文信息提取能力和细粒度特征图的纹理信息提取能力，需要对注意力图进行上采样操作，再将注意力图与浅层特征图进行乘法操作，得到各像素点的注意力系数 α ， α 的大小在一定程度上反映二维特征在特征图上的重要程度。最后，将 α 与浅层特征图上的像素点进行乘法操作，得到经过注意力机制的特征图。注意力机制的表达如下：

$$q_{att}^l = \psi^T (\delta (W_x^T x_i^l + W_g^T g_i + b_g)) + b_\psi,$$

$$\alpha_i^l = \sigma (q_{att}^l (x_i^l, g_i; \Theta_{att})),$$

其中， x_i^l 表示输入特征图， g_i 表示经门控机制的输出特征图， $\delta (\cdot)$ 表示 ReLU 激活函数， $\psi (\cdot)$ 表示 Sigmoid 激活函数， Θ_{att} 表示注意力机制的一组参数，即线性变换 $W_x \in \mathbf{R}^{C_l \times C_{int}}$ ， $W_g \in \mathbf{R}^{C_g \times C_{int}}$ ， $\psi \in \mathbf{R}^{C_{int} \times 1}$ ，偏差项 $b_\psi \in \mathbf{R}$ ， $b_g \in \mathbf{R}^{C_{int}}$ 。

注意力门控机制可添加在模型的跳跃连接上。本

求，本文引入深度可分离卷积模块。深度可分离卷积模块分为深度卷积块 (Depthwise) 和逐点卷积块 (Pointwise)，具体操作如图 8 和图 9 所示。

深度卷积块先对每个通道单独进行卷积，提取单通道的信息，再通过 1×1 的逐点卷积块对输入特征图的通道进行扩展或压缩，得到预期大小的特征图。相比标准卷积核，深度可分离卷积核在网络精度几乎

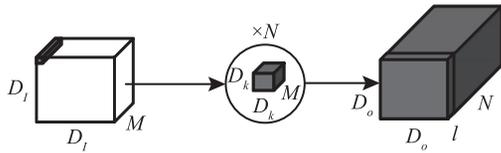


图8 标准卷积操作

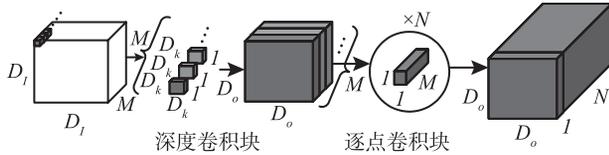


图9 深度可分离卷积操作

不降的情况下，大幅降低网络的参数量。

通常，在网络模型中使用卷积核的大小为 3×3 ，输出通道数远大于输入通道数，因此，深度可分离卷积在计算量上可近似为标准卷积的 $1/9$ ，大幅减少网络模型的参数量。

3 实验及结果分析

3.1 实验环境

本次实验选择如下实验数据集：Zero-DCE 网络数据集^[18]、笔者使用 RGB 摄像头在不同低光照环境下拍摄的人脸图像、公开的 DICM 数据集^[25]、NPE 数据集^[26]。对于 CNN 分类网络的训练，使用的数据包括 CEW 闭眼数据集^[27]和对眼部区域裁剪后的 LFPW 室外标记的人脸面部数据集^[28]，共收集 1 919 幅人眼图像，其中睁眼图像 852 幅，闭眼图像 1 067 幅。训练集和测试集的比例设为 8:2。在训练、测试过程中，为了增强数据的可靠性，将图像大小统一缩放至 256×256 。

本次实验的操作系统为 Windows10，CPU 为 AMD Ryzen 7 4800 Hz，显卡为 NVIDIA GeForce-RTX 2060，深度学习框架为 PyTorch。

在实验过程中，根据测试阶段得到的增强图像，通过图像客观评估指标和图像增强后的视觉表现，利用控制变量法，尝试不同的权重系数值。实验中发现，不使用空间一致性损失函数会使增强图像出现局

部对比度问题，不使用曝光损失函数会导致增强图像过暗，不使用彩色恒常损失函数会使增强图像出现色差，不使用亮度平滑损失函数会导致增强图像出现伪影，将图像的理想曝光值设置为 0.5 会导致欠曝现象，设置为 0.7 会导致过曝现象。最终将低光增强算法的空间一致性损失函数的权重系数设为 1，曝光损失函数的权重系数设为 1，彩色恒常损失函数的权重系数设为 5，亮度平滑损失函数的权重系数设为 1 600，光照图像理想曝光值 E 设为 0.6，分类网络的批量化大小设为 32，学习率设为 0.001，训练轮次设为 100。

为了验证本文算法的性能和优势，设计多组消融实验与对比实验，并采用如下图像客观评估指标：峰值信噪比 (Peak Signal to Noise Ratio, PSNR)、结构相似性 (Structural Similarity, SSIM)^[29]、基于感知的图像质量评估 (Perception-Based Image Quality Evaluator, PIQE)^[30] 和自然图像质量评估 (Natural Image Quality Evaluator, NIQE)^[31]。

PSNR 和 SSIM 值越高，表明增强结果图像的质量越优。PIQE 和 NIQE 值越低，表明增强结果图像的质量越优。

3.2 消融实验结果

首先，以 PSNR 和 SSIM 为性能评价指标，设计一组关于上下采样结构的消融实验，结果如表 1 所示。表中 Improved Zero-DCE 使用上下采样结构的特征提取网络代替 Zero-DCE 中的特征提取网络。相比 Zero-DCE，Improved Zero-DCE 的 PSNR 和 SSIM 值都明显提高，说明上下采样结构可有效抑制噪声，减少外界环境的干扰。由此验证上下采样结构可有效弥补 Zero-DCE 难以处理噪声的缺点。

表 1 上下采样结构的消融实验结果

算法	PSNR	SSIM
Zero-DCE	14.9333	0.2647
Improved Zero-DCE	19.2303	0.5348

此外，通过消融实验进一步验证网络是否需要注意力门控机制及注意力门控机制插入位置的问题，基础框架为 Improved Zero-DCE。实验结果如表 2 所示，表中黑体数字表示最优值。综合对比 4 种策略的 PSNR 和 SSIM 值，发现在浅层跳跃连接上添加注意力门控机制取得最优值。

表 2 注意力门控机制的消融实验结果

改进网络	PSNR	SSIM
不添加注意力门控机制	19.2303	0.5348
浅层添加注意力门控机制	19.7597	0.5524
深层添加注意力门控机制	19.1915	0.5297
浅层、深层添加注意力门控机制	19.3633	0.5288

同时，在实验中发现，相比添加注意力门控机制的低光增强网络（简记为 Attention Zero-DCE），不添加注意力门控机制的低光增强网络得到的增强图像，在步骤 2 人脸关键点检测过程中，会出现在一幅人脸图像上检测 2 个人脸框的现象，如图 10 所示，这严重影响最终的疲劳状态识别结果。由此验证注意力门控机制能使模型抑制与人脸区域不相关的特征，集中注意力激活人脸区域的特征。

最后，针对核选择模块的有效性，设计一组消融实验进行验证，基础框架为 Improved Zero-DCE。结果如表 3 所示。引入核选择模块后，PSNR 和 SSIM 值都进一步提升。由此验证核选择模块的有效

性，通过自适应调整感受野大小，进行多尺度特征融合，降低噪声影响。

表 3 核选择模块的消融实验结果

改进网络	PSNR	SSIM
不添加核选择模块	19.7597	0.5524
添加核选择模块	19.8627	0.5628

3.3 实验结果对

本节通过多个对比实验，验证本文算法的优势。为了体现本文算法的轻量性和实时性，采用如下对比算法：LIME^[11]，MBLLEN^[12]，Retinex-Net^[14]，KinD^[15]，EnlightenGAN^[17]和 Zero-DCE^[18]。各算法的性能对比如表 4 所示，表中黑体数字表示最优值。

表 4 各算法的性能对比

算法	推理时间 /S	参数量 /M	浮点运算数 /G
LIME	0.4914	—	—
MBLLEN	13.9949	0.150	301.12
Retinex-Net	0.1200	0.555	587.47
KinD	0.1480	8.160	574.95
EnlightenGAN	0.0078	8.637	273.24
Zero-DCE	0.0025	0.079	84.99
Improved Zero-DCE	0.0028	0.056	51.18

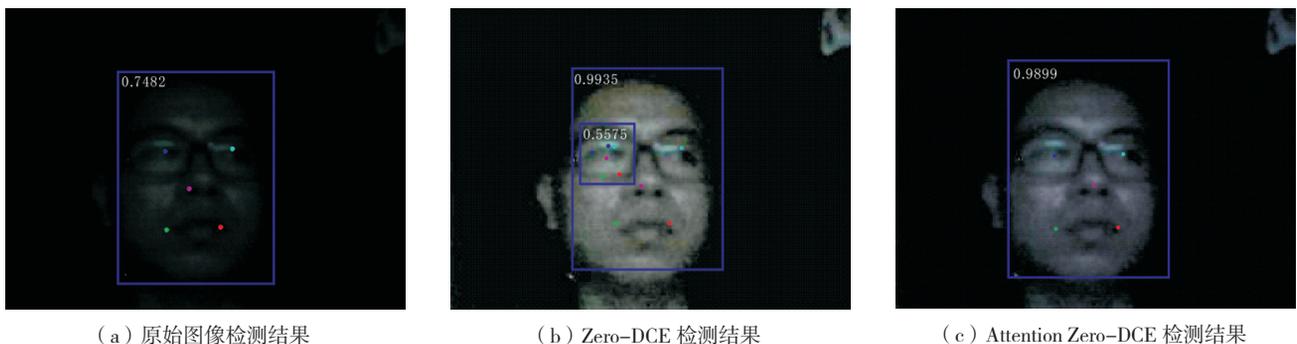


图 10 人脸关键点位置检测结果

通过替换 Zero-DCE 的特征提取网络和引入深度可分离卷积, 在模型参数量和浮点运算数上, 本文算法都优于现有的低光增强算法。在推理时间上, 本文算法优于大部分低光增强算法。

为了验证本文算法对低光照图像的增强效果, 针对低光照数据集, 首先对图像进行灰度转换, 然后统计图像中每个像素的灰度值, 最后将所有像素的灰度值相加再取均值。低光图像整体的平均亮度值为 8.83, 而增强图像整体的平均亮度值为 28.02。引入低光增强算法后, 相比原图, 增强后图像的整体亮度值大幅提升, 有利于之后的人脸检测和人脸分类步骤。

进一步, 为了综合评估本文算法的性能, 选择 DICM、NPE 数据集作为测试对象, 并采用 PIQE 和 NIOQE 作为图像客观评价指标。各算法的视觉效果对比如图 11 所示。由图可见, LIME 和 Retinex-Net 增强的结果色彩不理想, MBLLN 和 EnlightenGAN 的增强结果偏暗, KinD 的增强结果中物体边缘处产生伪影, Zero-DCE 的增强结果易受到噪声影响。

各算法在 2 个数据集上的 NIOQE、PIQE 值对比如表 5 所示, 表中黑体数字表示最优值。由表可知, 本文算法在 PIQE 指标上取得最优结果, 在 NIOQE 指

标上, 优于大部分算法。综合图像客观评价指标和图像增强结果的可视化, 本文算法的低光增强性能最佳。

表 5 各算法在公开数据集上的客观评价结果对比

方法	DICM 数据集		NPE 数据集	
	NIQE	PIQE	NIQE	PIQE
LIME	5.67	14.22	8.45	10.91
MBLLN	6.94	22.31	8.32	14.03
Retinex-Net	5.43	14.96	8.01	11.05
KinD	6.72	22.52	7.74	9.73
EnlightenGAN	6.22	14.65	8.17	9.44
Zero-DCE	6.72	14.12	8.27	9.26
Improved Zero-DCE	5.72	13.15	8.09	9.25

在低光照数据集上, 采用 RetinaFace 直接检测人脸关键点位置, 漏检率达到 6.56%, 而整体的平均准确率为 88.70%。

使用 Zero-DCE 先对原始图像进行低光增强, 再通过 RetinaFace 检测人脸关键点位置, 平均准确率为 98.55%, 虽有不小的提升, 但是漏检率升至 7.14%, 漏检更多人脸, 部分原因是由于在低光增强过程中, Zero-DCE 难以处理噪声, 导致增强后图像

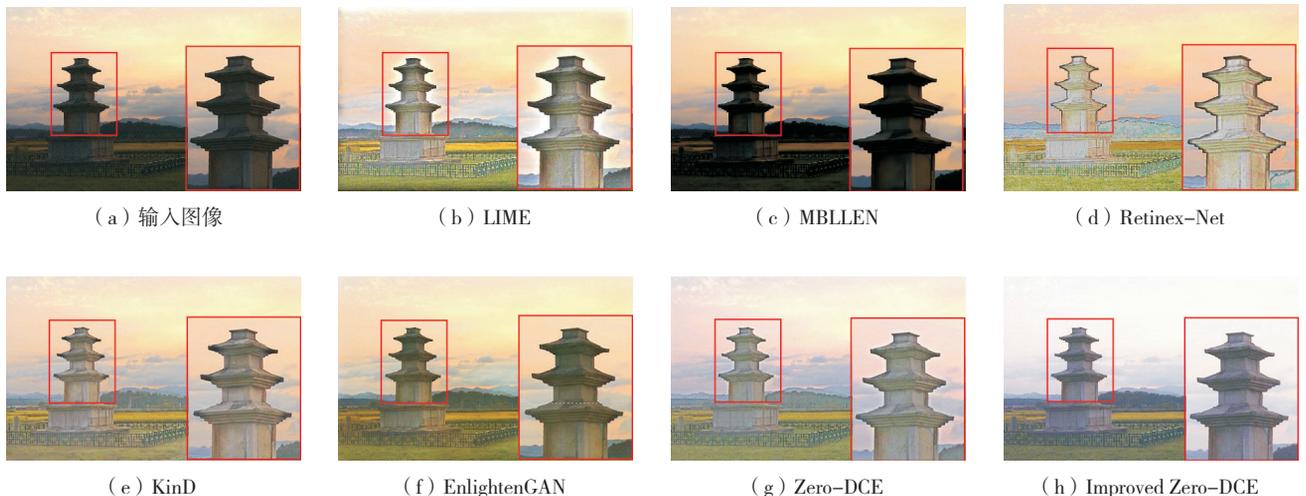


图 11 各算法在典型低光照图像上的视觉效果对比

的人脸区域夹杂着环境信息，使 RetinaFace 未检测人脸关键点位置。

而在低光增强过程中使用 Improved Zero-DCE，再通过 RetinaFace 检测人脸关键点位置，漏检率为 1.51%，平均准确率为 99.36%，都有明显改善，有利于后续对驾驶员疲劳状态的识别。

在原始图像上和使用低光增强算法后检测的人脸关键点位置如图 12 所示。由图可看出，在原始图像上的人脸检测准确率较低，人脸关键点位置都有些许偏差，而引入低光增强算法之后，人脸检测的准确率得到提高，人脸关键点位置也更精确。

综上所述，本文算法有助于提高夜间人脸关键点位置检测效果。



(a) 原始图像检测结果



(b) 低光增强算法检测结果
图 12 人脸关键点位置检测结果

为了避免低质量数据影响实验结果，经过人工筛选，本次用于眼睛状态识别实验的数据集包含睁眼、闭眼图像共计 572 幅。直接使用 RetinaFace+CNN，识别率仅有 74.13%，正确识别图像 424 幅。在人脸关键点位置检测前先使用 Zero-DCE，再使用

RetinaFace+CNN，眼睛状态的识别率得到提升，为 85.31%，正确识别图像为 488 幅。在人脸关键点位置检测前先使用本文算法得到增强图像，再使用 RetinaFace+CNN 对眼睛状态进行识别，可进一步提升识别率，为 93.53%，同时正确识别图像升至 535 幅。

实验表明，加入低光增强算法之后，疲劳驾驶检测器的性能大幅提升，主要是在夜间环境下，未使用低光增强算法的图像曝光度较低，人脸检测器难以准确检测人脸，影响对眼睛状态的判别，最终影响疲劳驾驶检测结果。

4 结束语

针对夜间疲劳驾驶检测这一任务，本文提出改进 Zero-DCE 的低光增强算法，在人脸检测任务之前引入低光增强算法，然后通过人脸关键点检测网络得到眼睛区域，最后使用 CNN 分类网络判断眼睛是睁开状态还是闭合状态。

本文算法在特征提取网络上使用上下采样结构取代 DCE-Net 的主干网络，降低噪声影响。再引入注意力门控机制，使网络有效激活图像中人脸区域的特征。然后，针对噪声相关问题，提出改进的核选择模块。进一步地，为了使夜间疲劳驾驶检测达到实时性的目的，使用 MobileNet 的深度可分离卷积核代替标准卷积核，大幅减少网络的推理时间和计算量。最后，通过多组消融实验和对比实验验证本文算法的性能。实验表明，加入低光增强算法后，相比现有的疲劳驾驶检测算法，夜间疲劳驾驶检测的准确率和识别率都有不小提升。

今后的研究重心是研究在不同场景下的疲劳驾驶检测。同时，将收集、制作更多不同场景下的疲劳驾驶图像，使疲劳驾驶检测算法应用于更广泛的实际环境中。○

来源：模式识别与人工智能

空间机器人抓捕目标后基于任务相容性的消旋策略

文 / 周逸群 罗建军 王明明

摘要：针对双臂空间机器人抓捕自旋目标后的镇定操作，在考虑机器人系统输入约束的条件下，提出了一种基于任务相容性的消旋规划与控制方法。首先，给出空间机器人抓捕目标后的组合系统的动力学模型，作为规划与控制的基础。然后，根据动力学可操作度和任务相容性设计了目标的快速消旋策略，其期望加速度的方向和大小分别取作速度的反方向和机器人系统输入约束允许的最大值。最后，基于所推导的运动学和动力学模型，通过对目标和机械臂末端分别建立柔顺度等式，提出了一种跟踪期望运动轨迹同时对末端接触力进行调节的柔顺控制方法。通过双臂 7 自由度空间机器人消除目标自旋运动的仿真结果，验证了所提方法的有效性。

关键词：双臂空间机器人；消旋策略；动力学可操作度；任务相容性；柔顺控制

1 引言 (Introduction)

空间机器人在执行卫星维护、在轨装配和空间碎片移除等在轨任务方面，与航天员相比具有诸多优势。自 20 世纪 80 年代以来，国内外航天机构已在大型航天器和在轨服务任务中实施了一系列空间机器人项目，典型实例有“工程试验卫星 7 号”（Engineering Test Satellite VII, ETS-VII）、“轨道快车”（Orbit Express, OE）、“凤凰计划”（Phoenix）、“德国轨道服务任务”（Deutsche Orbitale Servicing Mission, DEOS）等^[1]。应用机器人执行未来的空间任务已成为航天领域的一大趋势^[2]。

空间机器人对非合作目标的抓捕是在轨服务中最具挑战性的工作之一。捕获过程包括一系列操作，可分为 4 个阶段。第 1 阶段为观测和规划阶段，服务航天器获取目标卫星的运动和物理特性信息，并规划机械臂如何抓取目标。第 2 阶段是控制机器人向规划的抓捕位置移动，使机械臂准备好抓取目标。第 3 阶段

是实际捕获（物理拦截）阶段，机械臂抓取住目标卫星上的捕获设备。第 4 阶段是捕获后阶段，捕获的目标与服务系统作为一个整体稳定下来。图 1 展示了利用空间机器人抓捕目标这一典型在轨操作的示意图。

完成捕获后，空间机器人的首要任务是消除目标卫星的自旋运动。为解决这一具有挑战性的问题，需要研究相应的规划和控制方法^[3]。Dimitrov 和

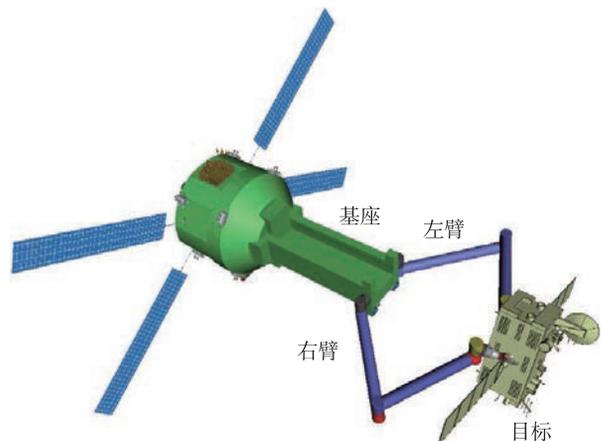


图 1 空间机器人捕获目标示意图

Yoshida^[4-5]提出了一种分布式动量控制方法,用于在抓捕后阶段吸收和管理目标的角动量,通过对目标的动量进行管理,使其不会干扰服务卫星的基座姿态。Aghili^[6-7]研究了机器人捕获翻滚卫星后的最优消旋控制策略,在末端执行器作用力和力矩满足约束的条件下使时间或能量的代价函数最小化。王明明等^[8-10]基于归一化时间或贝塞尔曲线设计了目标卫星的快速消旋策略,最优消旋时间由末端执行器处的相互作用力/力矩限制决定。Lampariello等^[11]提出了一种抑制目标相对运动的最优关节阻尼轨迹。Luo等^[12]提出了一种基于角动量分布的消旋镇定策略,在保持基座姿态稳定的前提下,将目标和机械臂的角动量传递给反作用轮,同时考虑关节速度的限制。Zong等^[13]研究了空间机械臂捕获翻滚目标后的最短时间消旋轨迹,考虑目标姿态运动边界和抓捕点处的交互力矩限制,并使用变分法求解。程靖和陈力^[14]针对空间机器人双臂捕获航天器后的姿态管理和辅助对接操作的协调控制问题,提出了基于极限学习机的自适应神经网络控制方案。此外,文[15-16]在调节服务卫星基座姿态的同时,利用服务卫星机械臂和卫星基座的协调控制来跟踪所需的消旋轨迹。

以上抓捕后的消旋策略是在目标惯性参数(质量、惯性张量和质心位置)已知的前提下发展起来的。这一假设的前提是,这些参数可以在抓捕之前通过参数辨识技术获得。参数辨识技术可分为基于视觉的^[17]、基于动量的^[18-20]和基于力的^[18, 21]技术。然而,这些参数辨识技术不能保证估计出的目标惯性参数在抓捕之前收敛到其真实值。Zhang等^[22]提出一种自适应滑模控制器来跟踪抓捕后阶段的参考轨迹,以使惯性参数不确定的非合作目标消旋。Zhu等^[23]提出了一种自适应滑模干扰观测器来估计未知目标产生的干扰,以便在跟踪期望的关节空间轨迹和调节服务卫星的基座姿态时对其进行补偿。Jayakody等^[24]应用自适应变结构控制方法,在服务卫星基座和

机械臂之间实现鲁棒的协调控制,以解决服务卫星系统中的参数不确定性。上述策略没有考虑末端执行器处的相互作用力/力矩限制。文[8, 25]使用末端执行器力/力矩测量值来实现阻抗控制,以跟踪期望的消旋轨迹。Oki等^[26]提出在不需要精确了解目标惯性参数的情况下描绘出一条消旋轨迹,并使用末端执行器处的力/力矩测量值修正轨迹,以遵守末端执行器的力/力矩限制。Nguyen-Huynh和Sharf^[27]提出了一种新的自适应算法,在空间机械臂捕获未知翻滚目标期间及捕获后产生零反作用运动。此外,Gangapersaud等^[28-30]提出一种解耦策略来调节服务位置的基座姿态,并跟踪末端执行器的参考力/力矩,从而解耦惯性参数未知但有界的非合作目标。

与单臂机器人相比,多臂机器人所能处理的载荷更多,灵活性也更高,但多臂协调操作的复杂度也更大。已有的消旋策略大多将末端执行器的力/力矩限制作为约束条件,而没有考虑服务卫星的力/力矩输入限制。在实际应用中,目标的可行运动受到空间机器人的力/力矩输入限制的影响。Yoshikawa^[31]提出动力学可操作度的概念以衡量力/力矩输入对目标加速度的映射能力。动力学可操作度可分别以操作因子和操作椭圆衡量^[32]。在此基础上,Chiu^[33]提出任务相容性以衡量机械臂构形在给定任务方向上的可操作度。Xu等^[34]研究了自由飘浮多臂空间机器人的任务相容性及其在轨迹规划和构形优化等方面的应用。

本文旨在提出一种基于空间机器人对所抓捕目标的任务相容性的消旋策略。目标的期望加速度的方向取作其速度的反方向,其大小根据当前机械臂构形的动力学可操作度和任务相容性确定,从而在机器人系统力和力矩输入限制内实现最速消旋,且无需任何优化算法。基于协调双臂运动的内外双环式结构,对目标和机械臂末端分别建立柔顺度等式,提出了一种协调稳定控制方法,以应对目标惯性参数不确定性,并降低机械臂末端与目标间的内力。

2 动力学建模 (Dynamic modeling)

在抓捕后阶段,空间机器人与所抓捕目标构成一个组合体。如图 2 所示,双臂空间机器人系统由基座和 2 条 n 自由度机械臂构成,共包含 $2n+1$ 个刚体,总自由度 $N=2n+6$ 。表 1 给出了组合体中运动学和动力学符号的含义,其中参数下标 b、m、e 和 t 分别表示与基座、机械臂、末端执行器和目标相关的项。在本文中,如无特别说明,所有参数都表示在与其固联的刚体坐标系中。 \mathbf{E}_n 和 $\mathbf{0}_n$ 分别表示 $n \times n$ 维单位矩阵和零矩阵。

P_i : 目标抓捕点
 Σ_e : 末端坐标系
 Σ_i^a : 惯性坐标系
 Σ_b : 基座坐标系

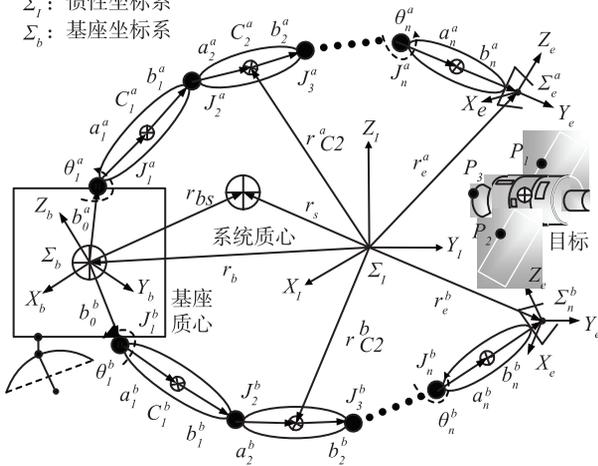


图 2 双臂空间机器人示意图

2.1 空间机器人动力学

根据拉格朗日力学原理^[17],空间机器人的动力学方程表示如下:

$$\mathbf{H}\ddot{\mathbf{q}} + \mathbf{c} = \mathbf{u} - \mathbf{J}_e^T \mathbf{f}_e \quad (1)$$

方程可以分解为以下形式:

$$\begin{bmatrix} \mathbf{H}_b & \mathbf{H}_{bm} \\ \mathbf{H}_{bm}^T & \mathbf{H}_m \end{bmatrix} \begin{bmatrix} \ddot{\mathbf{x}}_b \\ \ddot{\boldsymbol{\theta}} \end{bmatrix} + \begin{bmatrix} \mathbf{c}_b \\ \mathbf{c}_m \end{bmatrix} = \begin{bmatrix} \mathbf{f}_b \\ \boldsymbol{\tau} \end{bmatrix} - \begin{bmatrix} \mathbf{J}_b^T \\ \mathbf{J}_m^T \end{bmatrix} \mathbf{f}_e \quad (2)$$

或

$$\begin{bmatrix} \mathbf{H}_b & \mathbf{H}_{bm}^a & \mathbf{H}_{bm}^b \\ \mathbf{H}_{bm}^{aT} & \mathbf{H}_m^a & \mathbf{0}_n \\ \mathbf{H}_{bm}^{bT} & \mathbf{0}_n & \mathbf{H}_m^b \end{bmatrix} \begin{bmatrix} \ddot{\mathbf{x}}_b \\ \ddot{\boldsymbol{\theta}}^a \\ \ddot{\boldsymbol{\theta}}^b \end{bmatrix} + \begin{bmatrix} \mathbf{c}_b \\ \mathbf{c}_m^a \\ \mathbf{c}_m^b \end{bmatrix} \quad (3)$$

$$= \begin{bmatrix} \mathbf{f}_b \\ \boldsymbol{\tau}^a \\ \boldsymbol{\tau}^b \end{bmatrix} - \begin{bmatrix} \mathbf{J}_b^{aT} & \mathbf{J}_b^{bT} \\ \mathbf{J}_m^{aT} & \mathbf{0}_{n \times 6} \\ \mathbf{0}_{n \times 6} & \mathbf{J}_m^{bT} \end{bmatrix} \begin{bmatrix} \mathbf{f}_e^a \\ \mathbf{f}_e^b \end{bmatrix}$$

其中参数上标 a 和 b 分别表示与臂 a 或臂 b 相关的项。 $\mathbf{H} \in \mathbb{R}^{N \times N}$ 为空间机器人的广义惯性矩阵,由基座项 \mathbf{H}_b 、机械臂项 \mathbf{H}_m 以及耦合项 \mathbf{H}_{bm} 组成。 $\ddot{\mathbf{q}} \in \mathbb{R}^N$ 为广义加速度,包括基座加速度 $\ddot{\mathbf{x}}_b$ 和关节加速度 $\ddot{\boldsymbol{\theta}}$ 。 $\mathbf{c} \in \mathbb{R}^N$ 为广义偏差力,由基座项 \mathbf{c}_b 和机械臂项 \mathbf{c}_m 组成。 $\mathbf{u} \in \mathbb{R}^N$ 为机器人系统的广义输入,包括基座输入力和力矩 \mathbf{f}_b 以及关节扭矩 $\boldsymbol{\tau}$ 。 $\mathbf{J}_e \in \mathbb{R}^{12 \times N}$ 为雅可比矩阵,由基座项 \mathbf{J}_b 和机械臂项 \mathbf{J}_m 组成。 $\mathbf{f}_e \in \mathbb{R}^{12}$ 为末端施加的

表 1 本文所用运动学和动力学符号

符号	含义	符号	含义
J_i, C_i	关节 i 和连杆 i 的质心	$\mathbf{H}_b \in \mathbb{R}^{6 \times 6}$	基座惯性矩阵
$\mathbf{a}_i, \mathbf{b}_i \in \mathbb{R}^3$	从 J_i 到 C_i 和从 C_i 到 J_{i+1} 的位置向量	$\mathbf{H}_m \in \mathbb{R}^{n \times n}$	机械臂惯性矩阵
$\mathbf{r}_{Ci}, \mathbf{r}_i \in \mathbb{R}^3$	连杆 i 质心和目标的位置向量	$\mathbf{H}_{bm} \in \mathbb{R}^{6 \times n}$	耦合惯性矩阵
$\mathbf{r}_b, \mathbf{r}_e \in \mathbb{R}^3$	基座和末端执行器的位置向量	$\mathbf{c}_b \in \mathbb{R}^6, \mathbf{c}_m \in \mathbb{R}^n$	速度相关的非线性项
$\phi_b, \phi_e \in \mathbb{R}^3$	基座和末端执行器的姿态欧拉角	$\mathbf{f}_b, \mathbf{f}_e \in \mathbb{R}^6$	施加于基座和末端执行器的力和力矩
$\omega_b, \omega_e \in \mathbb{R}^3$	基座和末端执行器的角速度	$\boldsymbol{\tau}, \boldsymbol{\theta} \in \mathbb{R}^n$	施加于机械臂各关节的力矩和关节变量
$\mathbf{v}_b, \mathbf{v}_e \in \mathbb{R}^3$	基座和末端执行器的线速度	$\mathbf{J}_b \in \mathbb{R}^{6 \times 6}$	基座的雅可比矩阵
$m_i \in \mathbb{R}, \mathbf{I}_i \in \mathbb{R}^{3 \times 3}$	连杆 i 的质量和惯性矩阵	$\mathbf{J}_m \in \mathbb{R}^{6 \times n}$	机械臂的雅可比矩阵

力和力矩。

机械臂末端速度可以表示为

$$\dot{\mathbf{x}}_e = \begin{bmatrix} \dot{x}_e^a \\ \dot{x}_e^b \end{bmatrix} = [J_b \ J_m] \begin{bmatrix} \dot{\mathbf{x}}_b \\ \dot{\boldsymbol{\theta}} \end{bmatrix} = \begin{bmatrix} J_b^a \dot{x}_b + J_m^a \dot{\boldsymbol{\theta}}^a \\ J_b^b \dot{x}_b + J_m^b \dot{\boldsymbol{\theta}}^b \end{bmatrix} \quad (4)$$

刚体速度 $\mathbf{x} = [\boldsymbol{\omega}^T \ \mathbf{v}^T]^T$ 由角速度和线速度组成。线速度 $\mathbf{v} = \dot{\mathbf{r}}$ 即为刚体位置向量的导数，而角速度与其姿态欧拉角 $\boldsymbol{\phi} = [\alpha \ \beta \ \gamma]^T$ 通过以下公式联系（欧拉角旋转顺序为 X-Y-Z）：

$$\boldsymbol{\omega} = J_\phi \dot{\boldsymbol{\phi}} \quad (5)$$

其中：

$$J_\phi = \begin{bmatrix} \cos \beta \cos \gamma & \sin \gamma & 0 \\ -\cos \beta \sin \gamma & \cos \gamma & 0 \\ \sin \beta & 0 & 1 \end{bmatrix} \quad (6)$$

2.2 目标动力学

目标卫星的动力学方程表示如下：

$$\bar{\mathbf{I}}_t \ddot{\mathbf{x}}_t + \dot{\mathbf{x}}_t \times^* \bar{\mathbf{I}}_t \dot{\mathbf{x}}_t = J_t^T \mathbf{f}_e \quad (7)$$

其中 $\bar{\mathbf{I}}_t = \text{diag}(\mathbf{I}_t, m_t \mathbf{E}_3) \in \mathbb{R}^{6 \times 6}$ 为目标转动惯量和质量组成的广义惯性矩阵。 $\dot{\mathbf{x}}_t = [\boldsymbol{\omega}_t^T \ \mathbf{v}_t^T]^T \in \mathbb{R}^6$ 为目标质心速度。 $J_t = [J_t^{aT} \ J_t^{bT}]^T \in \mathbb{R}^{12 \times 6}$ 为目标雅可比矩阵，由目标质心到各抓捕点的转换矩阵组成。 \times^* 表示 6 维空间向量的叉乘因子，对于 $\mathbf{x} = [\mathbf{x}_1^T \ \mathbf{x}_2^T]^T \in \mathbb{R}^6$ ，其具体形式为

$$\mathbf{x} \times^* = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} \times^* = \begin{bmatrix} \mathbf{x}_1 \times & \mathbf{x}_2 \times \\ \mathbf{0}_3 & \mathbf{x}_1 \times \end{bmatrix} \quad (8)$$

对于 3 维向量 $\mathbf{x} = [x_1 \ x_2 \ x_3]^T \in \mathbb{R}^3$ ，叉乘因子 \times 的形式为

$$\mathbf{x} \times = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \times = \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix} \quad (9)$$

目标上的抓捕点速度可以表示为

$$\dot{\mathbf{x}}_g = \begin{bmatrix} \dot{x}_g^a \\ \dot{x}_g^b \end{bmatrix} = J_t \dot{\mathbf{x}}_t = \begin{bmatrix} J_t^a \dot{\mathbf{x}}_t \\ J_t^b \dot{\mathbf{x}}_t \end{bmatrix} \quad (10)$$

空间机器人完成对目标的捕获后，末端执行器和抓捕点的相对速度应为 0，即：

$$J_e \dot{\mathbf{q}} = J_t \dot{\mathbf{x}}_t \quad (11)$$

3 消旋策略 (Detumbling strategy)

空间机器人完成捕获后，末端执行器与抓捕点建立了稳固连接，如果不对目标实施消旋，目标卫星的旋转运动将会传递至空间机器人，从而引起整个组合体不可预测的运动。本节的主要内容是在满足服务卫星的力 / 力矩输入约束的条件下，如何在最优时间内使目标卫星在给定的初始角速度下实现停靠，同时需保证目标运动轨迹的连续性并尽量减小运动过程中的非期望影响。这里借助动力学可操作度及其衍生的任务相容性的概念，提出一种充分利用机器人系统对目标的操作能力的消旋策略。不失一般性，本文中变量的上标 s、d 和 f 分别代表变量的初始、期望和最终值。

3.1 动力学可操作度

动力学可操作度^[31]是衡量机械臂关节扭矩输入对操作点加速度的映射能力的指标。机器人与目标之间的加速度约束可通过对式 (11) 两边求导得到：

$$J_e \ddot{\mathbf{q}} + \dot{J}_e \dot{\mathbf{q}} = J_t \ddot{\mathbf{x}}_t \quad (12)$$

由式 (1) 和式 (7) 分别可得：

$$\ddot{\mathbf{q}} = \mathbf{H}^{-1}(\mathbf{u} - \mathbf{c} - J_e^T \mathbf{f}_e) \quad (13)$$

$$\ddot{\mathbf{x}}_t = \bar{\mathbf{I}}_t^{-1}(J_t^T \mathbf{f}_e - \dot{\mathbf{x}}_t \times^* \bar{\mathbf{I}}_t \dot{\mathbf{x}}_t) \quad (14)$$

将以上两式代入式 (12)，经整理得：

$$(J_e \mathbf{H}^{-1} J_e^T + J_t \bar{\mathbf{I}}_t^{-1} J_t^T) \mathbf{f}_e = J_e \mathbf{H}^{-1}(\mathbf{u} - \mathbf{c}) + \dot{J}_e \dot{\mathbf{q}} + J_t \bar{\mathbf{I}}_t^{-1} \dot{\mathbf{x}}_t \times^* \bar{\mathbf{I}}_t \dot{\mathbf{x}}_t \quad (15)$$

令 $\mathbf{A} = J_e \mathbf{H}^{-1} J_e^T + J_t \bar{\mathbf{I}}_t^{-1} J_t^T$ ，则有：

$$\mathbf{f}_e = \mathbf{A}^{-1}[J_e \mathbf{H}^{-1}(\mathbf{u} - \mathbf{c}) + \dot{J}_e \dot{\mathbf{q}} + J_t \bar{\mathbf{I}}_t^{-1} \dot{\mathbf{x}}_t \times^* \bar{\mathbf{I}}_t \dot{\mathbf{x}}_t] \quad (16)$$

将式 (16) 代入式 (14) 得：

$$\begin{aligned} \ddot{\mathbf{x}}_t &= \bar{\mathbf{I}}_t^{-1} J_t^T \mathbf{A}^{-1} [J_e \mathbf{H}^{-1}(\mathbf{u} - \mathbf{c}) + \dot{J}_e \dot{\mathbf{q}}] + \\ &\quad (\bar{\mathbf{I}}_t^{-1} J_t^T \mathbf{A}^{-1} J_t \bar{\mathbf{I}}_t^{-1} - \bar{\mathbf{I}}_t^{-1}) \dot{\mathbf{x}}_t \times^* \bar{\mathbf{I}}_t \dot{\mathbf{x}}_t \quad (17) \\ &= \mathbf{A}_t \mathbf{u} + \mathbf{b}_t \end{aligned}$$

此即为目标加速度的动力学可操作度公式。其中 A_t 和 b_t 分别为可操作度矩阵和偏差加速度：

$$\begin{aligned} A_t &= \bar{I}_t^{-1} J_t^T A^{-1} J_e H^{-1} \\ b_t &= (\bar{I}_t^{-1} J_t^T A^{-1} J_t \bar{I}_t^{-1} - \bar{I}_t^{-1}) \dot{x}_t \times^* \bar{I}_t \dot{x}_t + \\ &\quad \bar{I}_t^{-1} J_t^T A^{-1} (J_e \dot{q} - J_e H^{-1} c) \end{aligned} \quad (18)$$

式 (17) 可转化为

$$u = A_t^+ (\ddot{x}_t - b_t) \quad (19)$$

其中上标 + 表示矩阵的广义逆。令 $u = [f_b^T \tau^T]^T$ 满足如下输入限制：

$$\begin{aligned} |u_i| &\leq u_{i,\max} \\ u_{\max} &= [f_{b,\max}^T \tau_{b,\max}^T]^T \end{aligned} \quad (20)$$

其中 $f_{b,\max}$ 和 $\tau_{b,\max}$ 分别为基座力 / 力矩和关节扭矩的限制。考虑由各输入边界构成的权重矩阵 $W_u = \text{diag}(u_{\max})$ ，可将 u 按比例缩放为归一化向量：

$$u = W_u \hat{u}, \quad |\hat{u}_i| \leq 1 \quad (21)$$

于是有：

$$\ddot{x}_t = A_t W_u \hat{u} + b_t \quad (22)$$

以及

$$\hat{u} = W_u^{-1} A_t^+ (\ddot{x}_t - b_t) \quad (23)$$

动力学可操作度可分别以操作因子和操作椭圆衡量^[32]。首先，对 A_t 进行奇异值分解：

$$A_t = [U_1 \ U_2] \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} \quad (24)$$

其中 $U_1 \in \mathbb{R}^{6 \times r}$ (r 为 A_t 的秩)， $U_2 \in \mathbb{R}^{6 \times (6-r)}$ ， $V_1 \in \mathbb{R}^{N \times r}$ ， $V_2 \in \mathbb{R}^{N \times (N-r)}$ 。 $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_r) \in \mathbb{R}^{r \times r}$ 为 A_t 的各非零奇异值组成的对角矩阵。考虑到偏差加速度的影响，定义偏差因子：

$$\eta = \sqrt{\sum_{i=1}^6 \left(\frac{p_i}{\sigma_i} \right)^2} \quad (25)$$

其中 $p = U^{-1} b$ 。综上所述，考虑偏差的动力学可操作度因子定义为

$$ma_t = \begin{cases} (1 - \eta) \det(\Sigma(A_t) W_u), & \eta < 1 \\ 0, & \eta \geq 1 \end{cases} \quad (26)$$

假设归一化输入 \hat{u} 满足单位球约束：

$$\hat{u}^T \hat{u} \leq 1 \quad (27)$$

根据式 (23)，目标加速度满足约束：

$$\begin{aligned} &(\ddot{x}_t - b_t)^T A_t^+ W_u^{-2} A_t^+ (\ddot{x}_t - b_t) \leq 1 \\ \text{或} &(\ddot{x}_t - b_t)^T (A_t W_u^2 A_t^T)^{-1} (\ddot{x}_t - b_t) \leq 1 \end{aligned} \quad (28)$$

这一约束在几何上表示为一个球心位于 b_t 的 6 维椭球，即动力学可操作度椭球。椭球的各轴长分别对应 A_t 的各奇异值大小，在不考虑偏差加速度的情况下，椭球的体积与可操作度因子成正比。

3.2 任务相容性

任务相容性^[33]可以看作期望任务方向上的可操作度衡量。这里以单位向量 k_t 表示期望任务方向，于是目标加速度可以表示为

$$\ddot{x}_t = a k_t \quad (29)$$

其中标量 $a > 0$ 为 \ddot{x}_t 在 k_t 方向的比例值。

任务相容性有椭圆法和比例因子法两种衡量方法^[34]。将式 (29) 代入式 (28)，可得到机器人输入的单位球约束下，目标加速度比例值 a 所满足的约束：

$$(a k_t - b_t)^T (A_t W_u^2 A_t^T)^{-1} (a k_t - b_t) \leq 1 \quad (30)$$

上式可转化为如下二元一次方程：

$$\bar{\alpha} a^2 + 2\bar{\beta} a + \bar{\gamma} \leq 0 \quad (31)$$

其中 $\bar{\alpha} = k_t^T B k_t$ ， $\bar{\beta} = -k_t^T B b_t$ ， $\bar{\gamma} = b_t^T B b_t - 1$ ， $B = (A_t W_u^2 A_t^T)^{-1}$ 。求解该方程可得 a 的最大值：

$$e a_{\max} = \frac{-\bar{\beta} \pm \sqrt{\bar{\beta}^2 - \bar{\alpha} \bar{\gamma}}}{\bar{\alpha}} \quad (32)$$

$e a_{\max}$ 的几何意义为可操作度椭圆球心沿 k_t 方向与椭圆表面的距离。

与椭圆法相比，比例因子法可以更精确地描述任务的相容性。将式 (29) 代入式 (19) 和式 (20)，可得：

$$-u_{\max} \leq \mathbf{A}_t^+ ({}^s \mathbf{a} \mathbf{k}_t - \mathbf{b}_t) \leq u_{\max} \quad (33)$$

该式可改写为

$$\begin{aligned} -u_{1,\max} &\leq (\mathbf{A}_t^+ \mathbf{k}_t)_1 {}^s a - (\mathbf{A}_t^+ \mathbf{b}_t)_1 \leq u_{1,\max} \\ &\vdots \\ -u_{N,\max} &\leq (\mathbf{A}_t^+ \mathbf{k}_t)_N {}^s a - (\mathbf{A}_t^+ \mathbf{b}_t)_N \leq u_{N,\max} \end{aligned} \quad (34)$$

根据以上不等式约束组，可以得到：

$$\left\{ \begin{aligned} &\frac{-u_{i,\max} + (\mathbf{A}_t^+ \mathbf{b}_t)_i}{(\mathbf{A}_t^+ \mathbf{k}_t)_i} \leq {}^s a \leq \frac{u_{i,\max} + (\mathbf{A}_t^+ \mathbf{b}_t)_i}{(\mathbf{A}_t^+ \mathbf{k}_t)_i}, \\ &(\mathbf{A}_t^+ \mathbf{k}_t)_i > 0 \\ &{}^s a = +\infty, \quad (\mathbf{A}_t^+ \mathbf{k}_t)_i = 0, \quad (\mathbf{A}_t^+ \mathbf{b}_t)_i \leq u_{i,\max} \\ &{}^s a = -\infty, \quad (\mathbf{A}_t^+ \mathbf{k}_t)_i = 0, \quad (\mathbf{A}_t^+ \mathbf{b}_t)_i > u_{i,\max} \\ &\frac{u_{i,\max} + (\mathbf{A}_t^+ \mathbf{b}_t)_i}{(\mathbf{A}_t^+ \mathbf{k}_t)_i} \leq {}^s a \leq \frac{-u_{i,\max} + (\mathbf{A}_t^+ \mathbf{b}_t)_i}{(\mathbf{A}_t^+ \mathbf{k}_t)_i}, \\ &(\mathbf{A}_t^+ \mathbf{k}_t)_i < 0 \end{aligned} \right. \quad (35)$$

由各不等式约束获得的比例因子 ${}^s a_i$ 为

$${}^s a_i = \begin{cases} \frac{u_{i,\max} + (\mathbf{A}_t^+ \mathbf{b}_t)_i}{(\mathbf{A}_t^+ \mathbf{k}_t)_i}, & (\mathbf{A}_t^+ \mathbf{k}_t)_i > 0 \\ +\infty, & (\mathbf{A}_t^+ \mathbf{k}_t)_i = 0, \quad (\mathbf{A}_t^+ \mathbf{b}_t)_i \leq u_{i,\max} \\ -\infty, & (\mathbf{A}_t^+ \mathbf{k}_t)_i = 0, \quad (\mathbf{A}_t^+ \mathbf{b}_t)_i > u_{i,\max} \\ \frac{-u_{i,\max} + (\mathbf{A}_t^+ \mathbf{b}_t)_i}{(\mathbf{A}_t^+ \mathbf{k}_t)_i}, & (\mathbf{A}_t^+ \mathbf{k}_t)_i < 0 \end{cases} \quad (36)$$

因此，比例因子的最大值可以取作 ${}^s a_i$ 的最小值，即

$${}^s a_{\max} = \min {}^s a_i, \quad i = 1, \dots, N \quad (37)$$

${}^s a_{\max}$ 表示机器人系统输入约束所允许的期望任务方向上的最大幅值。

3.3 消旋规划

针对自旋目标卫星的消旋策略是在给定的约束条件下，寻找一条从初始状态至静止状态的最优路径。目标转动的初始状态 ϕ_i^s 和 $\dot{\phi}_i^s$ 可由星上传感器确定，消旋策略的目的是使 $\omega_i^f = \dot{\phi}_i^s = 0$ 以及 $\dot{\omega}_i^f = \dot{\phi}_i^s = 0$ ，而终止时间和状态是自由的。

已知目标在某一时刻的运动状态，为了尽快消减目标自旋，可以将目标在该时刻的期望加速度方向

(即期望任务方向) 设定为速度的反方向，即：

$$\mathbf{k}_t = -\frac{\dot{\mathbf{x}}_t}{\|\dot{\mathbf{x}}_t\|} \quad (38)$$

根据 3.2 节内容，在满足机器人系统输入约束式 (20) 的条件下，目标在期望任务方向上的最大加速度可通过比例因子法获得。为保证目标运动轨迹的连续性，并避免在接近稳定时出现运动振荡，这里令目标期望加速度幅值随速度大小成比例下降，于是目标期望加速度为

$$\ddot{\mathbf{x}}_{td} = \frac{\|\dot{\mathbf{x}}_t\|}{\|\dot{\mathbf{x}}_{ts}\|} {}^s a_i \mathbf{k}_t = -{}^s a_i \frac{\dot{\mathbf{x}}_t}{\|\dot{\mathbf{x}}_{ts}\|} \quad (39)$$

根据式 (36) ~ (39) 所确定的消旋策略，能够在机器人系统输入约束下，充分利用机器人的操作能力，以实现目标的尽快消旋，同时保证目标运动轨迹连续。目标期望运动轨迹的获取无需任何优化算法，从而能够快速计算实现。已知目标的期望运动轨迹，空间机器人的期望加速度可通过逆运动学得到：

$$\ddot{\mathbf{q}}_d = \mathbf{J}_e^+ (\mathbf{J}_t \ddot{\mathbf{x}}_{td} - \dot{\mathbf{J}}_e \dot{\mathbf{q}}) \quad (40)$$

进一步地，通过积分可得到目标和机械臂的期望速度和构形。

4 柔顺控制 (Compliant control)

目标精确的惯性参数通常很难在捕获前阶段确定。由于目标期望运动轨迹规划需要目标惯性参数的估值，如果直接跟踪这一轨迹，可能造成末端执行器施加的控制力过大而产生安全问题。当机械臂与外部环境或其他机械臂发生相互作用时，通过施加柔顺策略，顺应机器人系统的动力学特性，可以有效避免接触力和力矩过大。Hogen^[35] 提出的阻抗控制可在目标与外部环境存在相互作用力或内力的情况下应用于协调操作。其基本思想是在末端执行器与目标连接的抓捕点处加入虚拟的机械阻抗 (质量 - 弹簧 - 阻尼)。阻抗控制的优点是可以适当地设计末端执行器的阻抗参数以应对目标参数的不确定性。本文设计了一种具

上文所述的 f_{td} 为利用目标动力学方程 (7) 和期望运动轨迹得到的目标期望合外力 / 力矩。 M_t 、 D_t 和 K_t 为对称正定矩阵。由于 x_{tc} 与 f_t 直接满足上述柔顺等式，因此当目标以 x_{tc} 对应的轨迹运动时，受目标不确定性影响的末端接触力和力矩会产生相应的调整。

为控制内力对末端接触的影响，需要根据末端期望力和力矩，对经外环调整后得到的目标柔顺运动进行进一步的调整。根据运动约束 (12)，机械臂末端的期望运动为

$$\ddot{x}_{cd} = J_t \ddot{x}_{tc} \quad (47)$$

利用末端接触力 / 力矩 f_c 的测量值，构造调整末端期望运动轨迹的内环柔顺等式：

$$M_e \delta \ddot{x}_{edc} + D_e \delta \dot{x}_{edc} + K_e \delta x_{edc} = \delta f_c \quad (48)$$

其中 $\delta x_{edc} = x_{ed} - x_{ec}$ ， $\delta f_c = f_{cd} - f_{co}$ 。 x_{ec} 为经过调整后的末端柔顺运动轨迹，可通过下式的积分得到：

$$\ddot{x}_{ec} = \ddot{x}_{cd} + M_e^{-1} (D_e \delta \dot{x}_{edc} + K_e \delta x_{edc} - \delta f_c) \quad (49)$$

上文所述的 f_{ed} 可根据式 (42) ~ (44) 所制定的载荷分配方案由 f_{td} 得到。类似地，由于 x_{ec} 与 f_c 直接满足上述的柔顺等式，因此当末端沿轨迹 x_{ec} 运动时，受内力影响的末端接触力和力矩将会产生相应的调整。

4.3 跟踪控制

机器人期望运动轨迹可通过下式的积分得到：

$$\ddot{q}_d = J_c^+ (\ddot{x}_{ec} - \dot{J}_c \dot{q}) \quad (50)$$

为利用空间机器人协调地镇定被抓捕的非合作目标，基于空间机器人动力学方程 (1)，设计如下 PD (比例 - 微分) 控制器以跟踪期望运动轨迹 q_d ：

$$u = H[\ddot{q}_d + M_c^{-1} (D_c \delta \dot{q} + K_c \delta q)] + c + J_c^T f_c \quad (51)$$

其中 $\delta q = q_d - q$ 。 M_c 、 D_c 和 K_c 为 PD 控制器的控制参数。将式 (51) 代入式 (1)，可得：

$$M_c \delta \ddot{q} + D_c \delta \dot{q} + K_c \delta q = 0 \quad (52)$$

经过柔顺等式 (45) (48) 的调整，最终得到的

期望运动轨迹 q_d 考虑了目标不确定性以及内力的影响。通过跟踪该轨迹，空间机器人能够同时有针对性地控制内力和目标不确定性对末端接触所产生的影响，在机械臂末端与目标的交互过程中实现柔顺接触。

5 仿真结果 (Simulation results)

本节通过双臂空间机器人捕获自旋目标的案例来验证所提出的消旋策略。如图 2 所示，基座具有 6 个自由度，2 只抓捕机械臂都具有 7 个自由度。空间机器人的运动学与动力学参数如表 2 所示，其中 a_i 、 b_i 和 I_i 是在与其固联的本地坐标系中进行表示的。

如表 3 所示，在消旋规划和控制中，引入目标惯性参数的不确定性，以验证所提的消旋策略在此情况下的有效性。在抓捕前仅通过测量目标卫星的旋转运动是无法辨识出其惯性参数的，但可以获得目标的无量纲惯性参数。在此基础上，目标参数的估计值可通过 $\text{tr } \bar{I}_t$ 获得。机器人基座与机械臂关节输入需满足约束 $|u_i| \leq u_{i,\max}$ 。根据式 (16) (17)，较大的目标惯性参数估计值会降低机器人系统输入对目标的动力学可操作度，从而增加目标的消旋时间。因此，在目标惯性参数尚不确定的情况下，将其估计值设计得较为保守是比较合理的。

本文针对捕获后的组合体系统，设计了消旋策略及柔顺控制方法，以消除目标的自旋运动并管理接触力和力矩。目标上的 2 个抓捕点相对于目标质心坐标系的位置分别为 $r_g = [0, \pm 0.1006, 0] \text{ m}$ 。基座力和力矩输入限制为 $|f_{b,i}| \leq 10 \text{ N}$ 和 $|m_{b,i}| \leq 10 \text{ N} \cdot \text{m}$ ，关节扭矩输入限制为 $|\tau_i| \leq 5 \text{ N} \cdot \text{m}$ 。期望接触力和力矩的权重为 $k_1 = k_2 = 0.5$ 。在捕获瞬间 (即消旋运动开始的时刻)，目标相对于惯性坐标系的运动状态为

$$\begin{aligned} \phi_t^s &= [0, 0, 0]^\circ \\ \dot{\phi}_t^s &= [3, -8, 4]^\circ/\text{s} \\ r_t^s &= [3.5561, 0, 0.1680] \text{ m} \\ v_t^s &= [0, 0, 0] \text{ m/s} \end{aligned} \quad (53)$$

表 2 空间机器人的运动学与动力学参数

基座		机械臂 a 和 b						
		关节 1	关节 2	关节 3	关节 4	关节 5	关节 6	关节 7
$a/(\circ)$	—	0/0	-90/90	90/-90	90/-90	90/-90	-90/90	90/-90
a/m	—	0.0	0.0	0.0	0.0	0.0	0.0	0.0
d/m	—	0.128	0.168	1.450	0.168	1.290	0.168	0.440
θ		θ_1	θ_2	θ_3	θ_4	θ_5	θ_6	θ_7
		连杆 1	连杆 2	连杆 3	连杆 4	连杆 5	连杆 6	连杆 7
m/kg	200	3.0	8.0	2.0	6.0	2.0	2.0	4.0
I_{xx}	63.24	0.004096	1.3824	0.004704	0.8712	0.004704	0.004704	0.06453
I_{yy}	83.92	0.004096	0.0256	0.0064	0.192	0.0064	0.004704	0.06453
I_{zz}	84.68	0.0096	1.3824	0.004704	0.8712	0.004704	0.0064	0.0128
a_i/m	(0.0, 0.0, 0.0)	(0.0, 0.0, 0.064)	(0.0, -0.7, 0.084)	(0.0, 0.0, 0.09)	(0.0, -0.6, 0.084)	(0.0, 0.0, 0.09)	(0.0, 0.0, 0.084)	(0.0, 0.0, 0.168)
b_i/m	(2.52, -0.318, 0.0)	(0.0, 0.084, 0.128)	(0.0, -0.66, 0.0)	(0.0, -0.084, 0.0)	(0.0, -0.46, 0.0)	(0.0, 0.084, 0.0)	(0.0, -0.09, 0.0)	(0.0, 0.0, 0.188)

表 3 估计的目标动力学参数^[8]

符号	标准 20%	偏移 50%	偏移
$\text{tr } \bar{I}_t$	66.84	80.208	100.26
m/kg	40.00	48.00	60.00
I_{xx}	16.80	20.160	25.20
I_{yy}	24.36	29.232	36.54
I_{zz}	25.68	30.816	38.52

空间机器人的基座与机械臂的运动状态为

$$\begin{aligned}
 \phi_b^s &= [0, 0, 0]^\circ \\
 \dot{\phi}_b^s &= [2.8495, 1.0630, 0.1649]^\circ/s \\
 r_b^s &= [0, 0, 0] \text{ m} \\
 v_b^s &= [0.0142, -0.0479, -0.0414] \text{ m/s} \\
 \theta^{as} &= [0, -30, 0, -30, 0, -45, 0]^\circ \\
 \theta^{bs} &= [0, +30, 0, +30, 0, +45, 0]^\circ \\
 \dot{\theta}^{as} &= [2.3003, -2.5690, 2, 4870, -1.8419, \\
 &\quad -3.5160, 3.1081, -1.5457]^\circ/s \\
 \dot{\theta}^{bs} &= [-2.6172, -1.4544, -2.3229, -0.7601, \\
 &\quad 3.1222, 3.1408, 1.8554]^\circ/s \quad (54)
 \end{aligned}$$

外环柔顺等式、内环柔顺等式和运动跟踪控制的惯性、阻尼和刚度矩阵的选择如表 4 所示。

表 4 期望柔顺参数表

参数下标	M^*	D^*	K^*
t	$200 \cdot E_6$	$640 \cdot E_6$	$5 \cdot E_6$
e	$2000 \cdot E_{12}$	$1800 \cdot E_{12}$	$10 \cdot E_{12}$
c	$5000 \cdot E_{20}$	$3000 \cdot E_{20}$	$1000 \cdot E_{20}$

图 4 ~ 图 6 给出了标准目标参数下的仿真结果。图 4 展示了目标的角运动（欧拉角、角速度和角加速度）和线运动（位置、线速度和线加速度），其中红线表示根据本文提出的消旋策略规划的期望运动轨迹，黑线表示采用本文设计的柔顺控制得到的实际运动轨迹。以 $\|\omega_t\|_\infty \leq 10^{-3} \text{ rad/s}$ 为标准，目标的期望运动用时约 6.1 s 实现消旋稳定，稳定姿态欧拉角为 $\phi_t^f = [0.074, -0.172, 0.094] \text{ rad}$ 。在本文设计的柔顺控制方案下，目标的 α 和 β 角很快跟踪上期望值， γ 角

由开始时的振荡逐渐收敛至期望值，用约 12.9 s 实现消旋稳定，最终实现对期望状态的精确跟踪。此过程中，目标位置基本保持稳定，其线速度大小限制于 8×10^{-3} m/s 以下。图 5 展示了空间机器人的机械臂末端接触力矩和力，红线和黑线分别为由载荷分配得到的期望值和实际值。由于预计目标仅做角运动，其期望接触力为 0，接触力矩与加速度成正比。在实际情况下，由于运动约束的影响，难以同时实现对期望运动轨迹和期望力 / 力矩的精确跟踪。末端接触力矩

和力在开始时出现偏离期望值的较大值（最大值分别为 $0.96 \text{ N} \cdot \text{m}$ 和 1.52 N ），而后随着目标消旋稳定逐步收敛至 0 附近。图 6 展示了仿真过程中基座力矩和力输入分量以及各关节扭矩输入的最大值（黑线）与其输入限制（红线）。开始时基座力矩和力输入达到或接近限制值，后续输入随着目标稳定而逐渐减小。这样，在满足机器人系统力 / 力矩输入约束以及运动轨迹连续性的前提下，在最短时间内实现了对目标的镇定消旋，并避免了末端接触力和力矩过大。

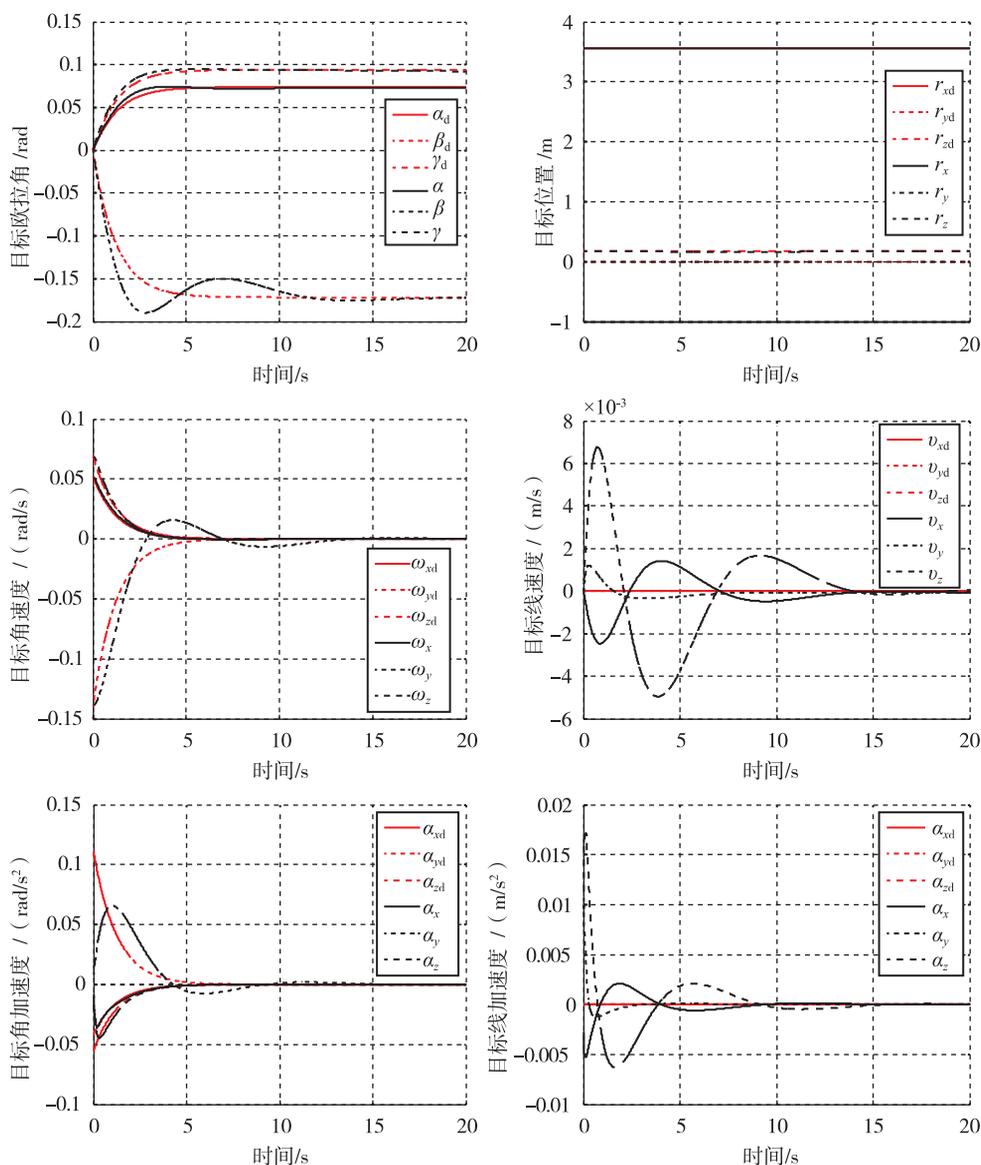


图 4 标准目标参数下目标的角运动和线运动

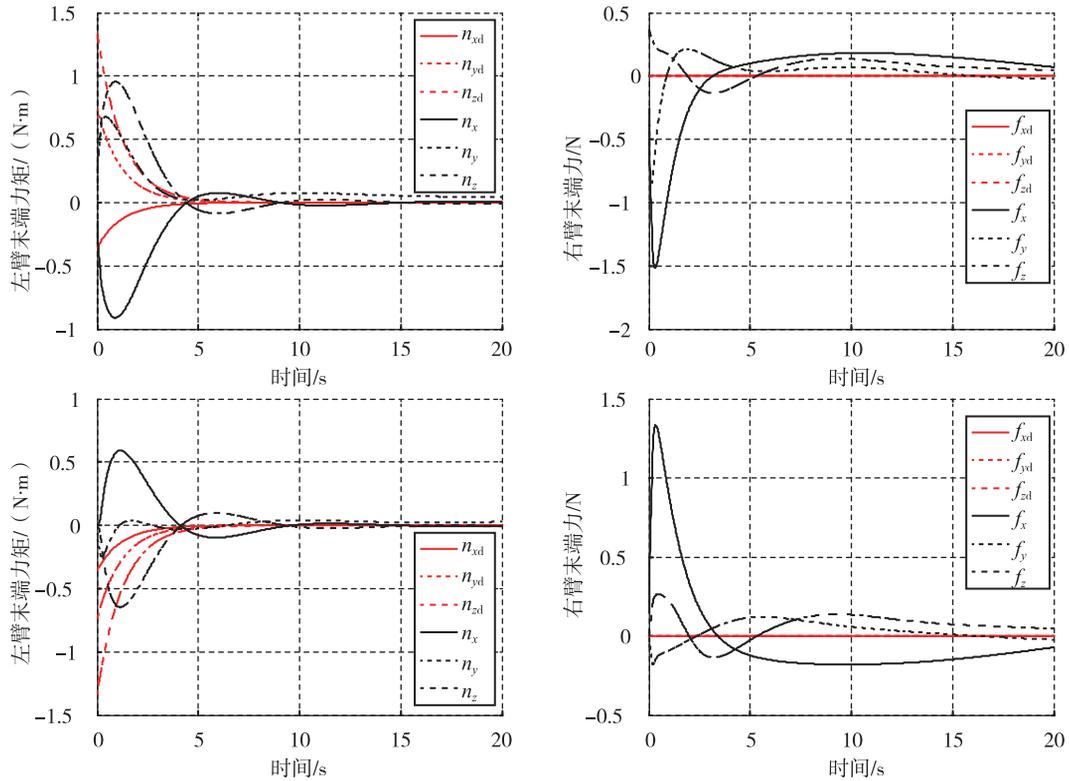


图5 标准目标参数下末端接触力矩和力

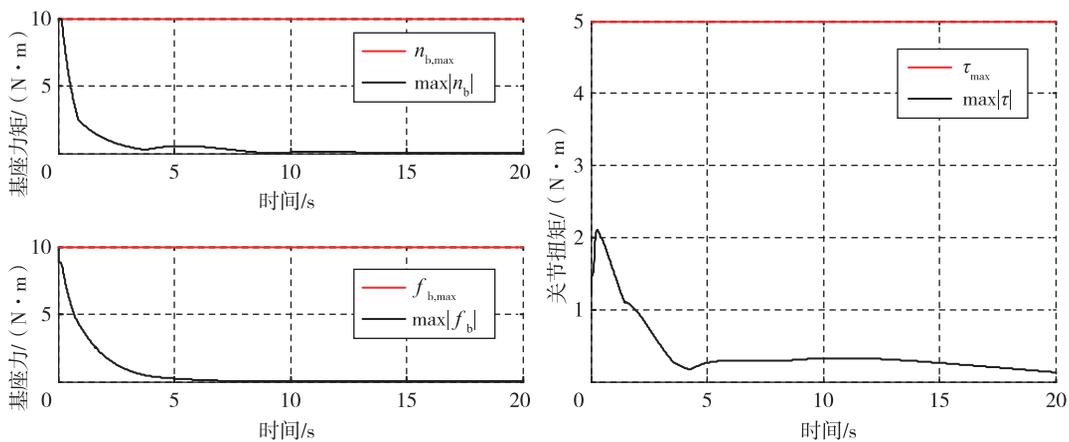


图6 标准目标参数下机器人力矩和力输入最大值

图7 ~ 图9 给出了目标参数估计存在 20% 偏差时的仿真结果。由于对目标惯性参数作了较大的估计, 因此目标期望加速度有所降低, 消旋时间有所延长。目标的期望运动用时约 6.2 s 实现消旋稳定, 稳定姿态欧拉角增大至 $\phi_1^f = [0.077, -0.178, 0.098]$ rad。在目标参数估计存在偏差的情况下, 本文的柔顺

控制方案依然能使目标运动轨迹收敛至期望值, 用时约 13.0 s 实现消旋稳定。末端接触力矩和力的最大值分别为 0.94 N·m 和 1.49 N, 而后逐步收敛至 0 附近。类似地, 机器人系统输入由开始时的达到或接近限制值, 随着目标稳定而逐渐减小。

图 10 ~ 图 12 给出了目标参数估计存在 50% 偏

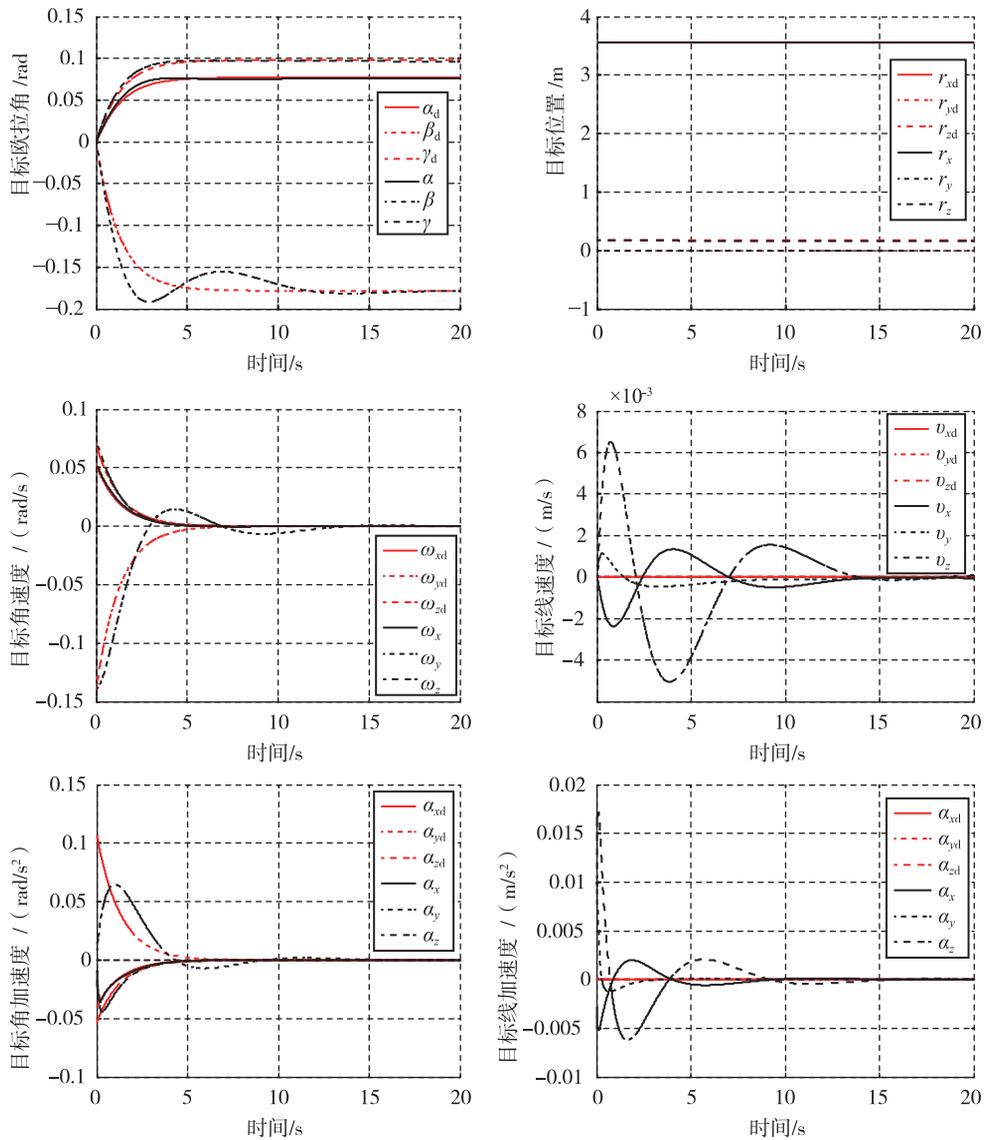


图7 目标参数估计偏移 20% 情况下目标的角运动和线运动

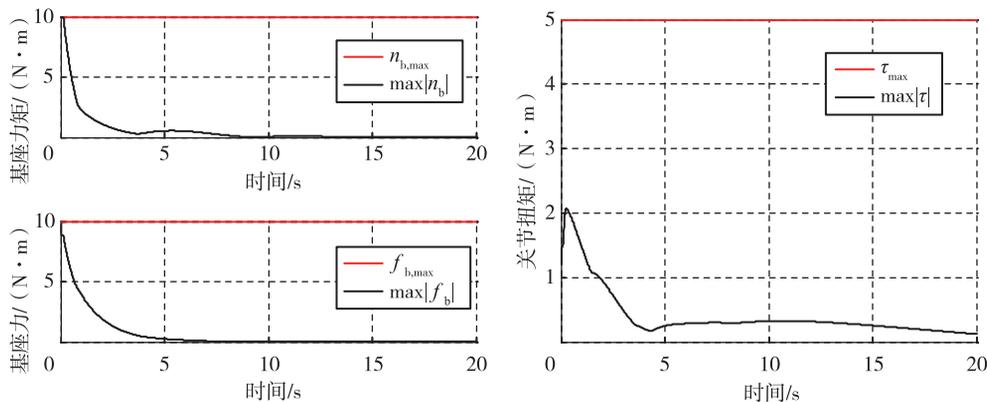


图8 目标参数估计偏移 20% 情况下机器人力矩和力输入最大值

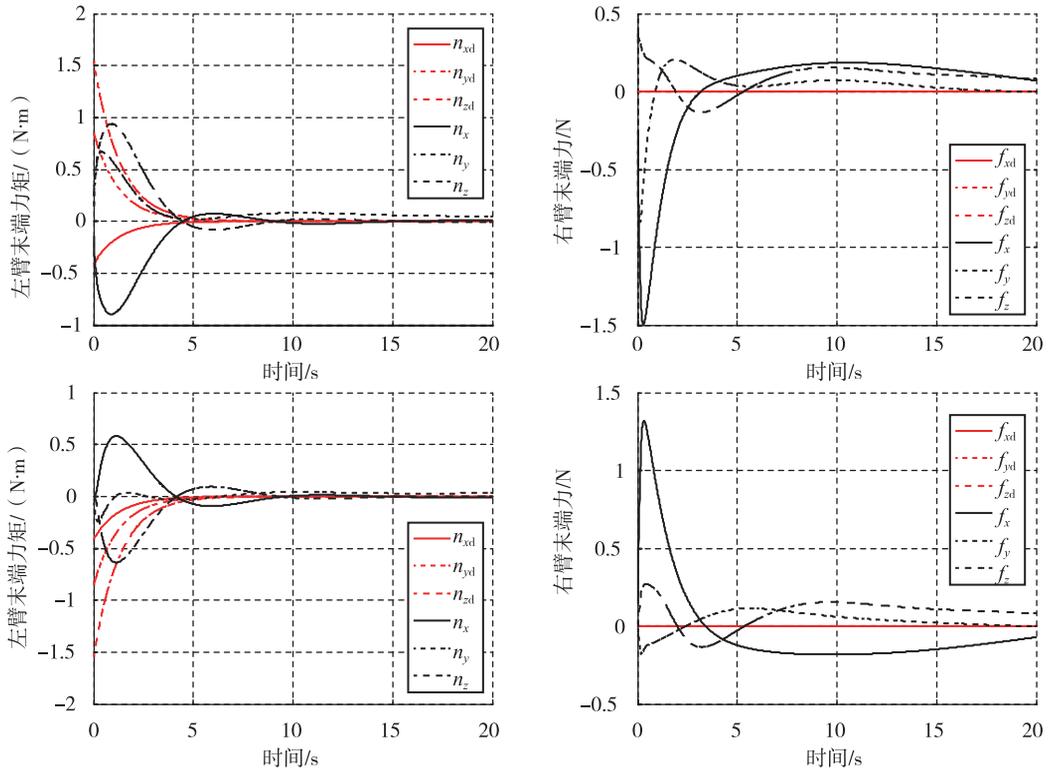


图9 目标参数估计偏移 20% 情况下末端接触力矩和力

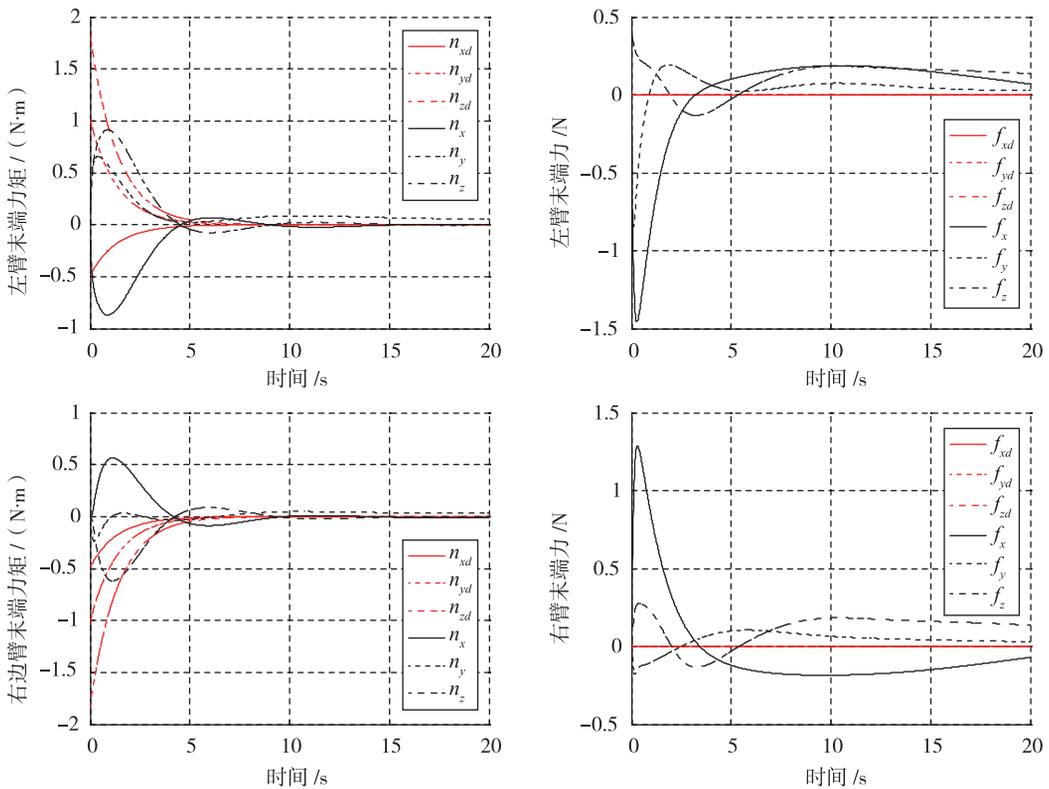


图10 目标参数估计偏移 50% 情况下末端接触力矩和力

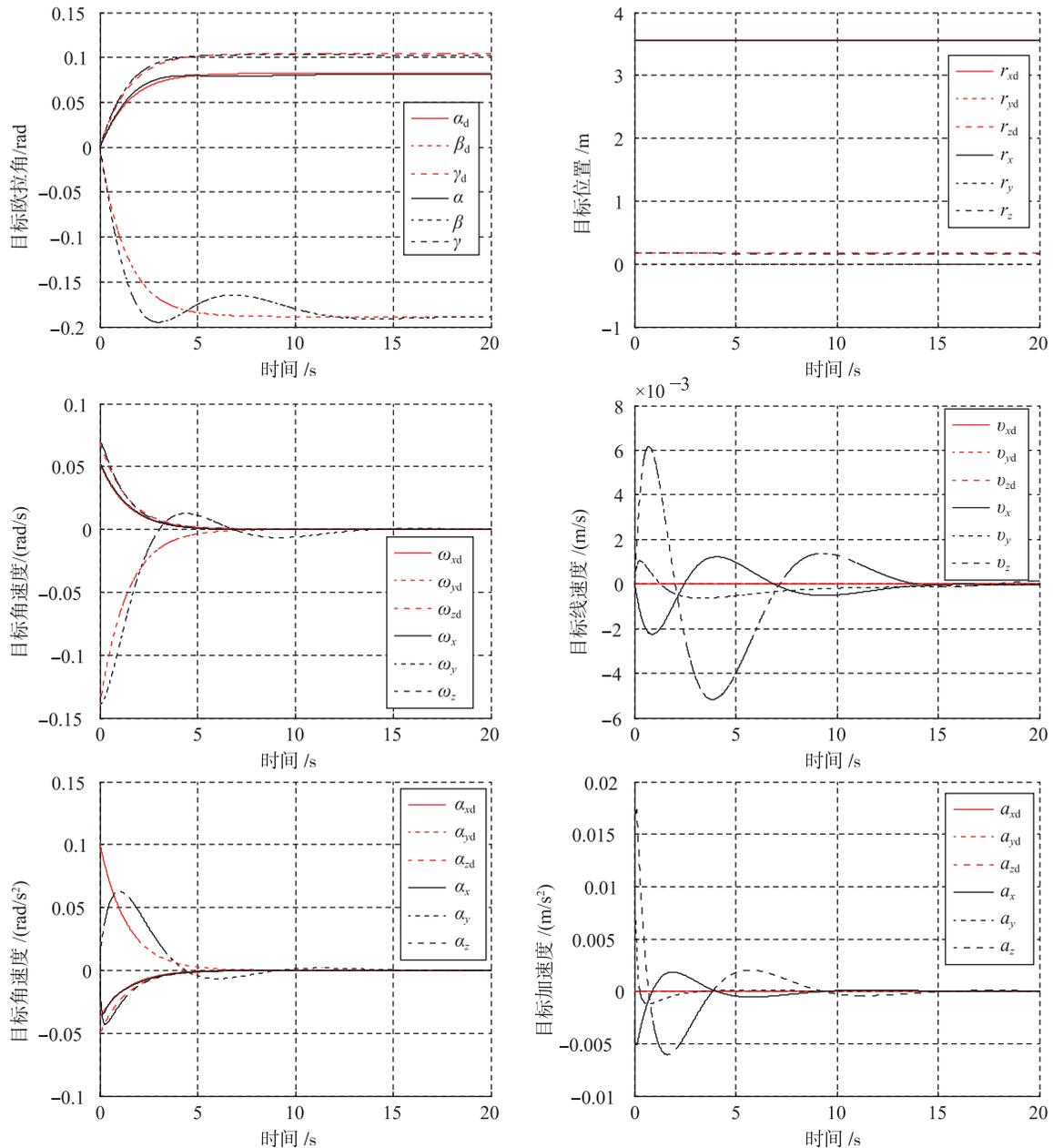


图 11 目标参数估计偏移 50% 情况下目标的角运动和线运动

差时的仿真结果。因为选择了较大的目标惯性参数估计，目标的期望运动用时约 6.7 s 实现消旋稳定，稳定姿态欧拉角进一步增大至 $\phi_1^f = [0.082, -0.189, 0.105]$ rad。目标实际运动用时约 13.1 s 实现消旋稳定，并能够使目标运动轨迹收敛至期望值。末端接触力矩和力的最大值分别为 $0.91 \text{ N} \cdot \text{m}$ 和 1.46 N ，而后逐步收敛至 0 附近。机器人系统同样满足输入限制要

求。综上所述，即使目标参数具有不确定性，所提的消旋策略和柔顺控制方法依然可以完成目标消旋与稳定控制的任务。

6 结论 (Conclusion)

在空间机器人捕获非合作目标后的阶段，为使自旋的目标卫星尽快稳定，考虑机器人系统的输入限制，

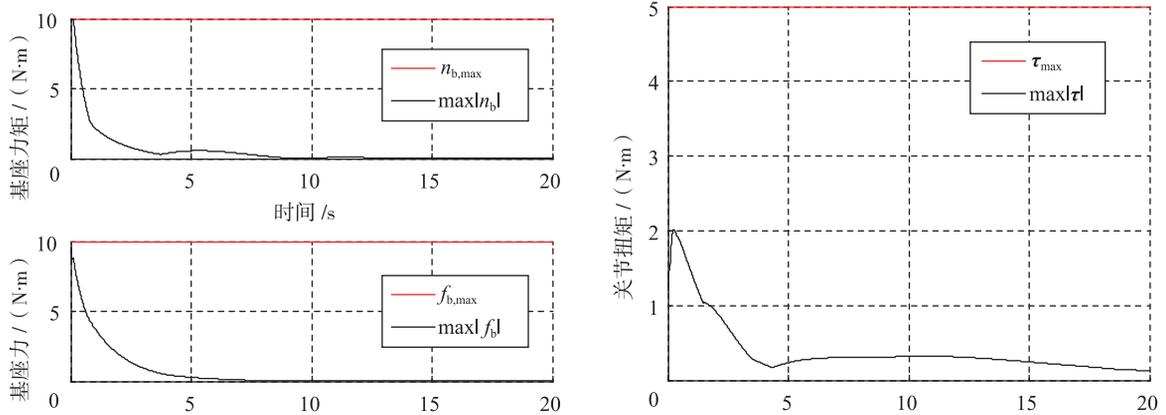


图 12 目标参数估计偏移 50% 情况下机器人力矩和力输入最大值

本文提出了一种基于任务相容性的消旋策略和柔顺控制方法。主要工作总结如下：

1) 基于任务相容性理论，提出了一种充分利用空间机器人操作能力的消旋策略。目标的期望加速度的方向取作其速度的反方向，其大小根据比例因子法确定，从而在机器人系统力和力矩输入限制内实现最速消旋，且无需任何优化算法。

2) 为应对目标不确定性和内力的影响，设计了一种内外双环式柔顺控制方法，在外环和内环分别建

立阻抗关系式，对目标和末端执行器的期望运动轨迹进行调整，最终得到机器人的柔顺运动轨迹。

仿真结果证明了消旋策略及柔顺控制的可行性与有效性，并展示了该方法能够在目标惯性参数估计存在较大偏差的情况下消除目标自旋，并避免末端接触力和力矩过大。对目标惯性参数的精确辨识与空间操作的自适应控制问题将成为未来研究工作的重点。○

来源：机器人

作者简介

周逸群（1994—），男，硕士。研究领域：空间机器人在轨操控。

罗建军（1965—），男，博士，教授。研究领域：空天飞行器导航与制导。

王明明（1984—），男，博士，副教授。研究领域：空间机器人在轨操控。

“智行中国”系列论坛第三期“智能科学与技术交叉学科发展新机遇”线上论坛成功举办

2023年1月7日，“智行中国”系列论坛第三期成功举办。本次论坛以“智能科学与技术交叉学科发展新机遇”为主题，邀请20余位人工智能领军科学家、教育家、企业家及高校学者，围绕智能科学与技术交叉学科建设、人才培养和智能科学与技术交叉学科研究前沿与应用等问题展开讨论，共同思考学科交叉和交叉学科等发展新机遇。本次线上论坛通过多个线上平台共同直播，累计在线观看人次达66.91万。

本次论坛在教育部学位管理与研究生教育司、教育部科学技术与信息化司、国家自然科学基金委信息学部、中国工程院信息与电子工程学部指导下，由中国

自动化学会、中国人工智能学会主办，中国自动化学会教育工作委员会、中国人工智能学会教育工作委员会承办，百度公司、京东集团、钢铁侠科技、浙江大学上海高等研究院、北京工业大学北京人工智能研究院支持。

论坛开幕式由中国人工智能学会教育工作委员会主任委员、浙江大学人工智能研究所所长吴飞主持。教育部学位管理与研究生教育司副司长、国务院学位委员会办公室副主任栾宗涛，中国人工智能学会副理事长刘成林，中国自动化学会副理事长李少远在开幕式上分别致辞。

教育部学位管理与研究生教育司副司长、国务院学位委员会

办公室副主任栾宗涛在致辞中强调：全球科技创新进入空前密集活跃期，新一轮科技革命和产业变革正在重构全球创新版图、重塑全球经济结构。我国正处于工业经济向数字经济、智能经济迈进的关键时期，人工智能日益成为引领。这一科技革命和产业变革的战略性技术成为推动中国经济转型、产业升级的重要推动力量。国务院学位委员会，教育部积极贯彻落实党中央、国务院部署要求，大力推进人工智能相关学科专业建设。栾司长希望论坛对交叉学科内涵、交叉学科边界和处理好学科之间关系等问题进行深入讨论。

中国人工智能学会副理事长



图1 中国人工智能学会教育工作委员会主任委员、浙江大学人工智能研究所所长吴飞主持



图2 教育部学位管理与研究生教育司副司长、国务院学位委员会办公室副主任栾宗涛致辞



图3 中国人工智能学会副理事长刘成林致辞



图4 中国自动化学会副理事长、上海交通大学教授、青岛科技大学副校长李少远致辞

刘成林在致辞中提到：中国人工智能学会一直重视智能科学与技术学科的建设工作。2010年，中国人工智能学会开始论证增设“智能科学与技术”一级学科的必要性和可行性。2017年4月，中国人工智能学会在时任学会理事长李德毅院士领导下，联合20所国家一级学会、50所高校和100位专家组织论证增设“智能科学与技术”一级学科必要性和可行性，形成了《“智能科学与技术”一级学科论证报告》。目前，加大智能科学与技术人才培养，已经成为社会共识。希望通过今天的论坛，能够对学科交叉和交叉学科中共同关心的问题进行讨论，为更高质量交叉创新人才培养提供思路，实现教育、科技和人才的统一。

中国自动化学会副理事长、上海交通大学教授、青岛科技大学副校长李少远在致辞中讲到：智能科学与技术学科的发展在很多学校进行了前期的探索，同时

也遇到了很多的问题。在中国自动化学会、中国人工智能学会以及其他学会的共同推动下，围绕着智能科学与技术这样一个新的学科目录，就如何建设好、培养高层次的人才方面，共同开展一些工作。

大会报告由中国自动化学会教育工作委员会主任委员、清华大学自动化系主任、信息科学技术学院副院长张涛，中国人工智能学会教育工作委员会主任委员、浙江大学人工智能研究所所长吴飞主持。

中国工程院院士、欧亚科学院院士、中国人工智能学会和中国指挥与控制学会名誉理事长李德毅发表题为“人工智能看教育——答图灵：机器能够思维吗？”的主旨报告。李

德毅院士表示，人类智能始于语言，人工智能始于文字，人类最伟大的智慧是发明了教育，智能植根于教育，教育让人类智能的发展从“现在进行时”变为“现在完成时”。智能机器里的物质和能量是物理层面的真实存在，结构和时间是认知层面的抽象思维。结构和时间寄生在物质和能量上，构成硬构体；机器中信息是大量的软构体，体现了精神世界，可以寄生在硬构体或者其他软构体之上，有虚有实，虚实结合，可自举和自我复用，始终至少存在下一个时间周期，使得机器能再



图5 中国自动化学会教育工作委员会主任委员、清华大学自动化系主任、信息科学技术学院副院长张涛主持大会报告



图6 中国人工智能学会教育工作委员会主任委员、浙江大学人工智能研究所所长吴飞主持



图7 中国工程院院士、欧亚科学院院士、中国人工智能学会和中国指挥与控制学会名誉理事长李德毅作报告

去“想”，其秩序显示出维持自身思维和产生有序事件的能力，人的思维和机器的思维在数学上是同构的，在物理上是同源的，都赖负熵为生。这样一来，物质和信息没有鸿沟，科技和人文不再二元，这就回答了图灵“机器能够思维吗？”的天问。智能机器的诞生冲击教育的全方位、全要素，从教“书”，演进到教“学”，再进一步演进到教“育”。智能时代教育的精髓是培育思维的想象力和创造力，人类思维的本质是抽象、联想和交互，机器亦然。机

器用软构体延伸和拓展了人的思维，不但拥有记忆智能和计算智能，把人从繁重的、重复性劳动中解脱出来，还拥有具身交互智能，更重要的是，机器可以暴力思维，人和机器互教互学，优势互补，形成迭代的智能。智能时代的教育改革问题已经迫切地摆在全人类的面前，读书改变一个人，教育改变全人类，让我们迎接这场学习的革命！

中国科学院院士、发展中国家科学院院士、司库、美国工程院外籍院士、固体力学专家杨卫

发表题为“交叉科学——无尽的前沿”的主旨报告。他从交叉学科的学理、实操、案例三个章节对交叉学科的发展进行了阐述，提出了学科发展的经纬结构、学科空间和交叉学科的衍生过程，从交叉学科部的组织、文化、领域覆盖和矩阵式扁平化管理新模式出发，论述了如何跳出传统单一学科理念，打破学科之间的壁垒束缚。通过交叉力学领域介质交叉、层次交叉、刚柔交叉和质智交叉四大交叉案例生动形象地讲述了交叉学科的实践和发展。



图8 中国科学院院士、发展中国家科学院院士、司库、美国工程院外籍院士、固体力学专家杨卫作报告



图9 中国工程院院士、中国人工智能学会理事长、国务院参事、清华大学信息科学技术学院院长戴琼海作报告

中国工程院院士、中国人工智能学会理事长、国务院参事、清华大学信息科学技术学院院长戴琼海发表题为“人工智能的几点思考”的主旨报告。他指出,近十年来,再度兴起的人工智能展现出“头雁效应”,通过与各行各业的交叉融合,赋能万业,深刻地改变时代。建议应居安思危、未雨绸缪,在着眼当下的同时,也应当以长远的眼光对人工智能的未来发展提早谋篇布局、抢占先机。面对人工智能日益增长的算力需求,提出是否有新的计算范式带来颠覆式的算力革命的问题,面对应用任务日益复杂的算法需求,是否能够以生物智能启发新一代的人工智能理论与技术?面对人与机器日益深入的交互融合,如何能够让人工智能真正拓展人类智能、服务于人?在人工智能快速发展的时代,如何围绕人工智能的交叉特性,为未来培养合格的人工智能交叉研究、应

用和治理人才?戴琼海院士围绕以上主要问题、重大变革和潜在契机,分享了对前沿研究发展和未来人才培养的一些思考和探索。

南开大学原校长、中国新一代人工智能发展战略研究院执行院长龚克发表题为“智能科技:交叉学科与学科交叉”的主旨报告。他在报告中提出了“发展交叉学科的要义是回归科学的整体性”、“人工智能为代表的智能科技是一个交叉学科群”和“学科交叉比交叉学科更重要”三个观点,并对其做了详细的阐述和深度剖析。他寄语学生要有志气,不能只为了得一个好分数、不能只为了能够推研,也不能只为了能够拿到一份好的工作,要有志向才能增强自己的智力和学习能力,要成为一个“会学”的人走向社会,才能真正跟上这个社会科学技术发展。

欧洲科学院外籍院士、俄罗斯自然科学院外籍院士、西安电

子科技大学焦李成发表题为《从人工智能本硕博一体化培养到交叉学科建设的探索与思考》的主旨报告,他指出新一轮科技革命和产业变革蓬勃兴起,一些基本科学问题孕育重大突破,可望催生新的重大科学思想和科学理论,产生颠覆性技术。人工智能技术作为当今社会的主要生产力,为社会不断赋能,在诸多行业和领域取得了非常显著的成效。我国高度重视人工智能的发展,制定了诸多战略方针,发展人工智能,引领人工智能。焦李成教授回顾了人工智能及其核心技术的内涵与发展,以及其发展过程中变革和存在的挑战,紧接着从人工智能专业建设到交叉学科的特性展开讨论,最后给出了未来的一些思考。

中国工程院院士、中国自动化学会理事长、西安交通大学人工智能与机器人研究所所长郑南宁发表题为“智能机器的行为与



图 10 南开大学原校长、中国新一代人工智能发展战略研究院执行院长龚克作报告



图 11 欧洲科学院外籍院士、俄罗斯自然科学院外籍院士、西安电子科技大学焦李成作报告



图 12 中国工程院院士、中国自动化学会理事长、西安交通大学人工智能与机器人研究所所长郑南宁作报告



图 13 中国人工智能学会副秘书长、智能感知与自主控制教育部工程中心副主任、北京工业大学马楠主持

信任”的主旨报告，他从智能学科内涵发展的问题驱动的角度联系到未来应该注重哪些研究，从机器行为研究与自动驾驶、AI 系统的信任与基本要素两个主题进行了阐述。他提出对机器行为和信任的研究，它不仅应该是透明的，它也需要能够解释自身，还要考虑到社会价值观、道德和以道德为基础的在不同文化背景下的利益相关者。对于人工智能系统适当地信任、负责任的使用尤为关键，过多甚至盲目的信任，可能会导致一系列问题，甚至

灾难。

下午进行的论坛主旨报告分别由中国人工智能学会副秘书长、智能感知与自主控制教育部工程中心副主任、北京工业大学马楠，中国人工智能学会教育工作委员会副主任委员、西安电子科技大学人工智能学院智行院长侯彪主持。

清华大学自动化系系主任，信息科学技术学院副院长张涛发表题为“自动化领域智能科学与技术发展与建设”的主旨报告，他指出，国内自动化学科历史底

蕴深厚，与智能科学与技术专业联系密切，自动化领域相关学者在智能学科非常活跃，为智能科学与技术学科的发展和建设作出了重要贡献。他从智能科学技术的内涵、发展历程、国家相关战略与政策以及国内外相关专业设置方面，国内自动化领域发展智能教育思路方面，清华大学智能学科的发展过程、主要学科方向以及清华大学自动化系智能学科发展与建设的相关基础和成果方面进行了详细报告。

中国人工智能学会教育工



图 14 中国人工智能学会教育工作委员会副主任委员、西安电子科技大学人工智能学院智行院长侯彪主持



图 15 清华大学自动化系系主任，信息科学技术学院副院长张涛作报告



图 16 中国人工智能学会教育工作委员会主任委员、浙江大学人工智能研究所所长吴飞作报告



图 17 中国人工智能学会副理事长、国家“万人计划”首批领军人才，北京大学教授刘宏作报告

作委员会主任委员、浙江大学人工智能研究所所长吴飞发表题为“教材、课程和平台促 AI+X 人才培养”的主旨报告，介绍了新一代人工智能系列教材、在线课程建设、科教创新平台智海和人工智能科普等内容，希望通过“教材、课程和平台”三位一体构建科教融合和产教协同的人工智能人才培养新生态。

中国人工智能学会副理事长、国家“万人计划”首批领军人才，北京大学教授刘宏发表题为“再思考：AI 学科建设 20 年”

的主旨报告，回顾了 2003 年第一个智能科学与技术专业创办至今，二十年来全国 AI 学科的发展历程，论述了“智能现象的生成机理与模拟机制”学科内涵与交叉属性，提出未来 20 年逐步形成“智能学科门类”的交叉学科建设发展新趋势。

中国人工智能学会教育工作委员会副主任委员、哈尔滨工业大学校长助理、哈工大计算学部主任兼计算机学院院长、自然语言处理研究所所长刘挺发表题为“对智能学科建设的思考与实践”

的主旨报告，从交叉学科、智能交叉学科的认识与哈工大的智能科学与技术、人工智能研究院的发展以及人工智能交叉和拔尖人才培养等方面进行了详细报告。

国家杰出青年科学基金获得者，教育部长江学者特聘教授，图像信息处理与智能控制教育部重点实验室主任，华中科技大学教授曾志刚发表题为“智能生物系统大脑-身体-环境共生体对类脑智能计算与控制发展的启示”的主旨报告，他指出生物系统的进化是针对群体的优化而不



图 18 中国人工智能学会教育工作委员会副主任委员、哈尔滨工业大学校长助理、哈工大计算学部主任兼计算机学院院长、自然语言处理研究所所长刘挺作报告



图 19 华中科技大学教授曾志刚作报告

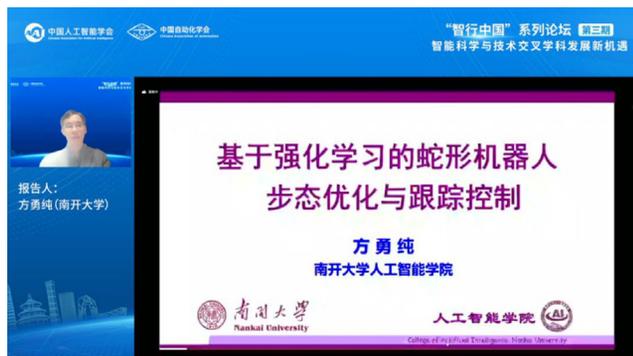


图 20 中国自动化学会理事、南开大学研究生院常务副院长方勇纯作报告

是个体的强大，从任务实现的角度来看其进化结果具有鲁棒性和最优性。他认为亿万年进化出来的所有形态的生物体都具有强大的功能潜力，而个体生物体的功能其实远没有达到“硬件”的极限。即使我们没有强大的物理算力，也能借鉴生物学为机器学习提供经验，在某些领域取得突破。进化使得智能生物体的次优解遍布超空间，使得个体生物通过简单试错就可以很容易达到比较满意的性能。大脑并不做精确计算，而是被设计用于分类和回忆复杂的信息集，智能生物体的学习和运动控制，可能就是存储足够多的次优解随时匹配，而不追求精准的最优反馈控制。

中国自动化学会理事、南开大学研究生院常务副院长方勇纯发表题为“基于强化学习的蛇形机器人步态优化与跟踪控制”的主旨报告，他指出与其他机器人相比，蛇形机器人可以在狭窄管道等环境中作业，因此具有广阔的

应用前景，步态优化与跟踪控制是蛇形机器人实现高效作业的关键。他基于强化学习等方法，就蛇形机器人的步态优化与控制问题展开论述，具体包括基于强化学习的步态优化方法，以及基于强化学习的蛇形机器人跟踪控制方法。

哈尔滨工程大学校长助理兼智能科学与工程学院院长赵玉新发表题为“航海技术智能化：机遇和挑战”的主旨报告，他指出航海技术是海洋运载器发展所需的共性基础性技术之一。随着人工智能理论与技术的快速兴起，必然对航海技术的未来产生新的认知和新的解决途径。他结合自身在海洋运载器航海技术领域的实践，探讨如何利用人工智能技术促进航



图 21 哈尔滨工程大学校长助理兼智能科学与工程学院院长赵玉新作报告

海技术的突破，特别是对针对航海技术智能化的发展方向，提出了他的观点和思考。

中国人工智能学会副秘书长、智能感知与自主控制教育部工程中心副主任，北京工业大学马楠发表题为“智能时代基于交互认知的教学模式探索与实践”的主旨报告，分为智能时代的教育特点、AI 赋能教育 - 在交互认知中成长学习、新工科人才体系培养探索 and 智能交互技术产学研建设四部分，从交互认知研究领域到科学任务引领以及如何围绕专业课程建设新形态教材、在线课程、



图 22 中国人工智能学会副秘书长、智能感知与自主控制教育部工程中心副主任，北京工业大学马楠作报告



图 23 华为计算产品线高校与科研人才发展部部长曹新方作报告



图 24 百度高校合作部副总监计湘婷作报告

教具、教学案例和虚拟仿真实践教学平台，如何培养创新型人才进行了详细报告。

华为计算产品线高校与科研人才发展部部长曹新方发表题为“产教融合，助力高校 AI+ 学科建设”的主旨报告，他从华为在人工智能产业的前沿技术融入高校教学体系，共同探索 AI+ 学科的建设，助力高校培养新时代需要的复合型人才等方面进行了详细报告。

百度高校合作部副总监计湘婷发表题为“人人皆可 AI——百度人工智能人才培养实践案例解读与思考”的主旨报告，她提出高校 AI 人才培养如何更顺滑地衔接传统产业转型的需求、培养卓越工程师，高校和企业应如何更务实地开展合作的问题。AI 赋能之下，各行各业发生了新的剧变，具备新的特点，对我国 AI 人才培养也提出了新的挑战——高校在 AI 专业建设、师资队伍、课程设置、实训实践等方面面临着教

育理念更新、知识结构重塑等各类问题。她从产业视角出发，解析 AI 人才培养的特点和实现路径，并结合百度 AI 人才培养的诸多成功案例，如松果 AI 菁英班、AI+X 课程共建、AI 微专业共建等，尝试为校企 AI 人才培养探索出一个更有效、更易落地的产教融合新范式。

京东集团副总裁何晓冬发表题为“致力于解决真实场景的问题，以产业助推智能科学发展”的主旨报告，他指出作为一家新型实体企业，京东一面连接消费互联网，一面连接产业互联网，不仅拥有前沿的 AI 技术，更连接海量的应用场景，京东也致力于解决产业里真实的问题，通过自研 + 产学研生态的方式，推动 AI 的产

业、规模化发展，让各行各业享受智能技术带来的红利。报告围绕近来火热的智能人机对话与交互、AIGC 等前沿课题，阐述京东自身在客服、数字人、政务、金融等领域的产业实践，同时介绍了京东在产学研方向的进展。

美团副总裁夏华夏发表题为“生活服务的数字化、智能化和自动化”的主旨报告，他提到美团的使命是帮大家吃得更好、生活更好，而美好生活的建造需要智能科学与技术的力量。他分享了美团在生活服务领域推进数字化、智能化和自动化的建设，结合美



图 25 京东集团副总裁何晓冬作报告

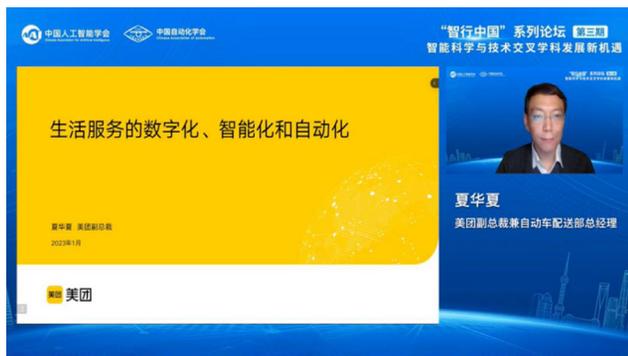


图 26 美团副总裁夏华夏作报告



图 27 联想集团副总裁王哲鹏作报告

第七届全国高校自动化教学院长（教学主任）会议顺利召开

为了促进全国高校自动化专业的交流与合作，展示和推广全国高校自动化教育教学研究和实践的优秀成果，探讨人才培养创新模式和理念，统筹谋划“十四五”自动化专业和控制科学与工程学科的建设与发展，由中国自动化学会教育工作委员会主办、北京工业大学信息学部人工智能与自动化学院承办的第七届全国高校自动化教学院长（教学主任）会议于2022年12月30日-12月31日在线上顺利召开，线上直播观看人数超1.1万人次。

本次会议主题为“新工科背景下自动化类专业人才培养与创新实践”，围绕一流专业建设、创新型人才培养、一流课程与课程思政、实践教学与工程人才培养、专业认证与自控体系建设等主题进行交流研讨，共同探讨新工科背景下自动化专业与控制科学与工程学科的高质量建设与发展，为全国高校自动化类专业教学院长（教学主任）提供了一个高层次的交流平台。会议包含开幕式、大会报告、分论坛研讨三个阶段。会议开幕式和大会报告全程在多个线上直播平台同步直播。

12月31日上午，中国工程院院士、中国自动化学会特聘顾问柴天佑教授，中国自动化学会教育工作委员会主任、清华大学自动化系主任张涛教授，北京工业大学副校长乔俊飞教授等专家学者出席开幕式。开幕式由北京工业大学信息学部副主任、人

团的实践发表了开展前沿技术探索、人才培养、技术成果转化方面的观点。

联想集团副总裁王哲鹏发表题为“人工智能推动行业智能化”的主旨报告，分享了联想研究院在人工智能推动行业智能化领域的布局、应用和探索以及在智能汽车行业自动驾驶算法上的探索。

中国自动化学会和中国人工智能学会联合主办的知行中国以“智能科学与技术交叉学科发展机遇”为主题的第三期论坛顺利举办。学科是高等教育的核心支柱，是立德树人的重要载体，在高等教育发展过程中具有全局性和技术性的影响。我们要用好学科交叉融合的催化剂，打破学科专业的壁垒，更加有效地促进学科交叉，提升创新人才培养的能力，开创智能科学与技术交叉学科发展新机遇。○

中国人工智能学会 供稿

工智能与自动化学院院长张利国教授主持。

张涛主任首先致开幕辞，感谢全体参会代表对本次大会的大力支持，表示本次会议将总结交流自动化专业建设与人才培养经验，面向“十四五”构建自动化领域高等教育高质量发展体系，提升自动化专业人才培养水平，促进自动化专业与控制科学与工程学科的高质量发展，进一步密切全国高校自动化专业的联系。乔俊飞副校长致欢迎辞，本次会议受新冠疫情影响几经周折后终于隆重召开，非常感谢中国自动化学会教育工作委员会对北京工业大学组织承办此次会议的支持，同时对与会专家的到来表示最热烈的欢迎和诚挚的感谢，相信通过本次大会的交流研讨将有助于促进全国高校自动化专业建设、展示我国高校自动化教育研究和实践的优秀成果、对高校自动化专业人才培养和创新实践产生深

远的影响，也诚挚邀请各位专家委员经常莅临北工大指导。

开幕式后，在北京工业大学信息学部主任、控制科学与工程学科负责人李晓理教授的主持下，中国工程院院士、东北大学柴天佑教授，东南大学戴先中教授，江南大学纪志成教授，湖南大学章兢教授，北京航空航天大学吕金虎教授分别作了题为“创新型综合自动化与智能化系统工程科技人才培养”、“优化工程教育专业认证中毕业要求达成评价的研究”、“国家教学成果奖培育申报与案例解析”、“专业认证：达成重要还是改进重要？”、“新时代科艺交融、自主创新的控制学科人才培养探索与实践”的大会报告。大会报告专家高屋建瓴，围绕综合自动化与智能化创新能力培养、高水平教学成果培育、毕业要求达成评价与持续改进等进行细致的解读，分享了最新的教学改革与人才培养经验。

12月31日下午，围绕人才培养与创新实践问题，大会由北京航空航天大学胡晓光教授、华中科技大学王燕舞教授、北京工业大学于乃功教授以及北京工业大学张利国教授分别组织召集了主题为“专业认证背景下自动化专业建设探索与实践”、“发展新工科课程思政，打造自动化一流金课”、“机器人工程培养方案构建及其特色专业建设”、“新时期自动化一流专业建设与创新人才培养”的四个分论坛，邀请了北京航空航天大学申功璋教授、兰州理工大学王志文教授、青岛大学刘旭东教授、东北大学孙秋野教授、北京化工大学李大学教授、南京邮电大学赵勃教授、华中科技大学王燕舞教授、南京工程学院刘娣教授、沈阳理工大学杨旗教授、东南大学周波副教授、北京工业大学于乃功教授、哈尔滨工程大学赵玉新教授、南京工业大学张广明教授、浙江大学侯迪波教授、



图1 中国自动化学会教育工作委员会主任、清华大学自动化系主任张涛教授致开幕辞



图2 北京工业大学副校长乔俊飞教授致欢迎辞

山东大学朱文兴教授就各自所在院校在一流专业建设、课程思政、实践教学改革、创新人才培养等方面的经验、方法、成果做论坛报告。与会人员就论坛主题进行了深入探讨，起到了很好的交流和宣传效果。4个分论坛累计参与人数达430余人。

12月30日晚上，第十届中国自动化学会教育工作委员会（简称“教工委”）第六次全体会议线上召开。会议向全体委员汇报了教工委全年工作，重点介绍了第七届全国高校自动化教学院长（教学主任）会议的筹备情况以及第八届全国高校自动化教学院长（教学主任）会议将由上海交通大学举办的筹备设想，并对教工委2023年的重点工作开展了详细讨论与部署。○

CAA 教育工作专委会
供稿

“青少年人工智能核心素养测评” 2022 年度总结大会圆满召开

2023年1月6日，由中国自动化学会主办，中国自动化学会普及工作委员会和中国自动化学会智慧教育专业委员会承办的“人工智能创新后备人才培养工程”之“青少年人工智能核心素养测评”在线上召开了2022年度工作总结大会。大会由领导专家致辞、人工智能主题报告、颁奖仪式、教学经验分享、测评年度工作总结、测评项目经验分享等几方面内容组成。来自测评主办单位、学术指导单位、承办单位的领导，联合发起单位、运营单位、共建示范校、测评运营中心、测评基地的专家和老师们，共计300余人参加了本次大会。

为贯彻落实国务院《新一代人工智能发展规划》和《全民科学素质行动规划纲要（2021—2035年）》、推进人工智能时代后备人才的培养、全面提升青少年人工智能核心素养，推动人工智能在基础教育阶段的规范发展，中国自动化学会于2021年发起了“青少年人工智能核心素养测评”项目（简称“AICE测评”），中国科学院大学人工智能学院和西安交通大学计算机科学与技术学院成为AICE测评的学术指导单位。

2021年11月底，AICE开启了首期测评；从2022年开始，AICE每年举行四期测评，每个季度举行一次。

2022这一年里，在疫情间歇爆发的情况下，AICE工作组同心同德、积极进取，完成了年度基本目标，组织了全年四期测评活动，丰富和健全了7个模块题库资源，实施了覆盖百所学校的公益活动，实现了测评服务全国34个省市地区全覆盖。

中国自动化学会秘书长张楠博士在总结大会上致辞，她表示，面对人工智能这个尚在发展初期、前沿技术和应用日新月异的领域，青少年人工智能教育该教什么、如何前后衔接、学生在学习上如何逐级递进，这些是中小学教师常见的困惑。“青少年人工智能核心素养测评”设计出了清晰合理、不断与时俱进、优化迭代的核心模型、测评框架及模块分级，明确了中小学阶段人工智能教育的阶梯路径，让学生和老师都有一个清晰的进阶方向。张楠秘书长说，希望青少年人工智能核心素养测评，这个工作越做越好，成为中国人工智能与自动化人才培养的一面旗帜。



图1 中国自动化学会秘书长张楠致辞



图2 中国自动化学会智慧教育专业委员会主任王万良演讲

中国自动化学会智慧教育专业委员会主任、浙江工业大学王万良教授在会上发表演讲，王教授表示，通过大家的努力，“青少年人工智能核心素养测评”已经初步建成了一个题库，题库的内容要不断完善、扩充，不断适应新的认识，适应不同年龄段学生的思维特点，让测评的大纲和题库能够和人工智能素养的培育同步，尤其是面对小学生和初中生的初级和中级题库里，要有生动的、有启发性的、能够融会贯通的题目。

中国科学院大学人工智能学院副院长肖俊教授在致辞中说，当前人工智能教育还没有特别规范化，人工智能素养测评工作的重要性不亚于人工智能教育本身。“青少年人工智能核心素养测评”的重点不在于测试我们孩子们的人工智能技术或者说水平有多高，更重要的是测评能起到很好的引导作用，引导中小学校、培训机构和广大中小学生去思考如何循序渐进地去了解和学习人工智能

的基础知识和重要技能。

中国自动化学会普及工作委员会副主任委员任思国博士在致辞中表示，要想扎实、系统地推进青少年人工智能普及教育，我们需要打造出三个体系——科学的课程体系、完备的师资培训体系，以及精准的测评体系。“青少年人工智能核心素养测评”汇聚了全国高校和中小学名校在教育、信息和人工智能方向上的教授、资深教师等一大批专家资源，为广大青少年构建起科学严谨、分级精细的人工智能核心素养测评体系，引导青少年人工智能教育的规范化、高质量发展。

领导致辞后，是专家报告环节。中国自动化学会普及工作委员会副主任委员、AICE 专家委员会负责人、中国人民大学附属中学信息技术教研组长袁中果博士以“中小学人工智能核心素养的培育”为题做了分享。袁老师指出，人工智能教育在小学阶段应该是以体验为主，重在感知；

在初中阶段是以发现为主，重在认知；高中阶段以探究为主，重在创新。高中阶段学生已经有了程序设计和人工智能原理的基本知识的积累，能够发现自身的能力和特长，形成不同学习重点的分科和分流。他认为，“青少年人工智能核心素养测评”是以评促学来促进中小学人工智能教育，让学生在人工智能领域里有潜质、有特长的孩子有机会来发挥他们的才智和潜能，让更多的孩子在中小学阶段就有机会接触人工智能。

中国自动化学会智慧教育专业委员会副秘书长、北京市东城区学科带头人、北京市第二中学的高凯老师进行了“中小学人工智能人才培养案例分享”，高老师认为，中小学的人工智能课程是一门实践性的课程，应更多以实践的方式让学生进行知识的理解。他经常在科技类课程当中提到建构主义的教育理念，希望学生在实践过程中不断地将知识和技能进行内化，真正转化成学生自己的经验；同时



图3 中国科学院大学人工智能学院副院长肖俊致辞



图4 中国自动化学会普及及工作委员会副主任委员任思国致辞

要把技术向善的理念传递给学生。

随后，大会进行了线上颁奖，中国科学院大学人工智能学院的肖俊院长和西安交通大学计算机学院的田峰教授先后宣读了各个奖项的获奖名单和获奖理由（见文末）。任赟、贾志勇、刘俊英等20位老师获得了AICE年度优秀个人贡献奖。霍颖怡、徐启发、马文杰等20位老师获得了AICE2022年度优秀辅导员；华中师范大学第一附属中学、澳门培正中学、北京市八一学校等15家学校被授予AICE2022年度共建特色校；中国儿童中心、北京猿力教育科技有限公司、腾讯科技（深圳）有限公司获得了AICE测评2022年度优秀单位贡献奖；三象教育（鲸桥编程）、剡锋教育、西安运营中心、辽宁运营中心、武汉运营中心等10单位获得AICE2022年度优秀组织单位奖；徐启发、任赟、刘俊英三位老师代表获奖老师和获奖单位发表了获奖感言。来自猿编程的刘俊英

老师表示，希望培养孩子的科技视野，让孩子去关注当前科技的发展，理解科技对于当下和未来的影响，进而具备更广阔的对于未来科技的视野。

中国自动化学会普及工作委员会副主任委员、中国自动化学会智慧教育专业委员会秘书长、AICE测评工作组组长刘希未博士在颁奖后对获奖老师和单位作了寄语，他说，期待有更多专家老师和同学们加入到测评中，共同用好测评结论，提升中小学人工智能教育水平。刘希未博士表示，AICE工作组在后续的工作中会积极听取更多专家和教学一线老师们的建议、要求和批评，在新的一年里扎实推进，持续创新，密切结合全国中小学人工智能教育的发展实际，不断的提升测评工作的科学化、系统化、标准化和规范化，切实为中小学的人工智能教育提升做好基础性、阶梯性和导向性的支持。

6日下午，大会继续进行。

CAA普及工作委员会委员张梦晗老师对2022年的AICE测评工作做了总结汇报，并对2023年的规划做了预告。随后，2022年AICE四期测评中的优秀学生代表及其辅导老师做了分享。AICE测评工作组的祝泽程老师讲解了AICE进阶提高阶段D/E项目的辅导与评审流程，上海市徐汇中学的胡奕铭同学介绍了自己参加测评的人工智能项目，上海爱砺锋科创中心的黄波曲老师分享了D/E模块的辅导经验，养蕊老师分享了中小学生学习创新思维及创新技术培养的经验和心得。

本次大会还邀请了几位在青少年及幼儿人工智能教育上富有经验的老师进行了主题分享。来自中国传媒大学的靳聪副教授分享了AI与学生艺术素养培养方面的经验，来自腾讯的刘学老师分享了基于腾讯扣叮工具的青少年人工智能活动，陈少瑜老师做了题为《助力人工智能教育标签语言》的分享，冯涛老师汇报了



图5 中国自动化学会普及工作委员会副主任委员袁中果作报告

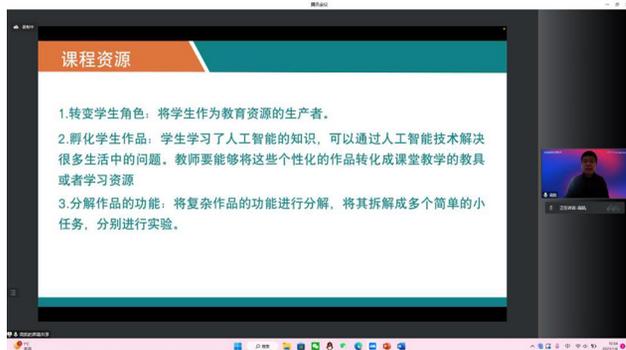


图6 中国自动化学会智慧教育专业委员会副秘书长高凯作报告

Future AI 幼儿启蒙教育的成果。一切过去，皆为序章；心之所向，素履以往。这是产学研各界汇集一堂的一次大会，是 2023

全新起点的集结誓师大会。AICE 测评工作组组长刘希未博士表示，在新的一年里，AICE 测评工作一定要密切结合全国中小学人工

智能教育的发展实际，结合 AICE 核心素养模型，针对中小学人工智能课程研发、人工智能与自动化科学普及和青少年后备创新人才培养目标，综合人工智能学科的通识性、多方向、多领域、实践性等特征，站在国家两个一百年奋斗目标的历史交汇点，服务二十大提出的建设教育强国、科技强国和人才强国的重大战略，顶天立地做好 AICE 测评的标准化设计研发与规范化落地实施工作。



图7 中国自动化学会普及工作委员会副主任委员、中国自动化学会智慧教育专业委员会秘书长、AICE 测评工作组组长刘希未讲话

青少年人工智能核心素养测评项目组 供稿



《机器人》再次入选《科技期刊世界影响力指数 (WJCI) 报告 (2022 版)》

近日，《科技期刊世界影响力指数 (WJCI) 报告》(2022 版) (以下简称“WJCI 报告”)正式发布,《机器人》成功入选,这亦是本刊自 2020 年《WJCI 报告》发布伊始连续 3 年入选。详情请查看: <http://www.caa.org.cn/article/191/3332.html>

中国自动化学会发电自动化专委会召开 2022 年度工作总结交流会议

中国自动化学会发电自动化专业委员会（以下简称“发电自动化专委会”）2022 年度工作总结、交流会议，于 2022 年 12 月 27 日以线上会议方式召开，中国自动化学会秘书长张楠、主任委员金耀华、副主任委员（尹松、许继刚、沈炯、崔志强、杨新民、刘志远、金丰、朱北恒）以及委员、顾问、受邀请专家近 80 位参加了会议。会议由尹松副主任委员主持。

中国自动化学会秘书长张楠在讲话中指出，专委会长期以来致力于发电自动化专业的安全生产和技术进步工作，团结和动员全国发电自动化专业和相关技术领域的科学技术工作者及有关

专业人员，牵头和指导解决发电生产中自动化专业出现的共性和个性问题，促进了我国发电自动化专业科学技术的繁荣发展，为行业发展和振兴乡村作出了重要贡献。

随后，会议就 2022 年专委会主要工作进行了汇报交流，崔青汝委员作了专题报告，会议中就 2023 年工作计划进行了讨论，最后金耀华主任委员进行了总结发言。

一、2022 委员会工作汇报

1. 增补马骏为第二届委员会委员

朱北恒副主任委员汇报了

第二届委员会委员增补马骏委员情况：

马骏，硕士毕业，高级工程师，学校毕业后入职上海新华控制技术（集团）有限公司，现为上海新华控制技术集团科技有限公司副总经理兼工程中心总经理。

马骏主要工作经历和主要技术、科技成果在委员会群公示后，经投票，通过了马骏为第二届委员会委员的增补。

2. 2022 年度发电自动化专业委员会工作总结报告

尹峰副秘书长代表委员会秘书处作《2022 年度发电自动化专业委员会工作总结》汇报。

2022 年在中国自动化学会的领导，各发电集团、委员单位及委员秘书处挂靠单位国网浙江省电力有限公司电力科学研究院、杭州意能电力技术有限公司的大力支持下，克服疫情带来的困难，积极、主动开展工作，主动承接政府职能，组织、协助完成了 12 项科技成果评审与鉴定，参加 3 次国家决策咨询，开展 2 项智能化建设发展专题研究；组织召开第九届中国发电自动化论坛，组



参会委员

组织、参加2次国际学术会议，组织3次国内学术会议和2次专委会工作会议；开展5项技术咨询与服务，组织完成8次自主可控或国产DCS性能与功能测试，为3家基建电厂举办热工安装调试知识培训班；组织、协助制定2项国标、2项行业标准、3项团体标准，组织、协助出版6部专业书籍；智航助学助教为乡村振兴助力取得成果，并就这些工作进行详细汇报。

3. 《2022年发电自动化专业委员会财务情况汇报》

杜永春委员代表委员会秘书处作《2022年发电自动化专业委员会财务收支情况汇报》，2022年经委员们的努力和相关电厂的支持，自主可控或国产DCS性能与功能测试、基建过程可靠性预控等咨询与服务项目取得较好业绩，财务上取得长进，为智航助学助教工作打下了良好基础。

4. 周力委员汇报2022年度智航助学助教情况

周力委员代表委员会秘书处和浙江省智航教育基金会作《智航助学助教情况2022年度总结与2023年计划》汇报，2022年中基金会延续了所有原已签协议的助学助教基地工作，对签署第一期智航助学助教协议三年到期的11所基地学校开展情况进行了评估，完成了10所第一期智航助学助教协议基地学校的第二期协议签署

和新增加6所智航助学助教基地学校的协议签署。各基地学校开展智航实验室教学情况良好，从事智航实验室教学的老师达54名，总授课课时为2161学时，接受智航实验室教学的学生达3714名。

全年举办了三次视频网络培训课程、资助弥勒三中参加全国青少年无人机云南赛区和全国竞赛。开展智航教育心理健康教师培训，组织了14次心理健康专题讲座、13次“心语驿站”个案研讨培训；扩校选点考察后增加了5所智航教育基地学校，承接了中国科协《振兴乡村学校智航科技创客教育的探索与推进》研究项目，与浙江电科院联建挂牌了浙江遂昌湖山小学智航助学助教基地、与南京科远智慧科技集团股份有限公司联建挂牌了弥勒四中智航助学助教基地。利用疫情中短暂的间隙，对一些学校进行了回访。开展年度“智航教学奖”和“智航贡献奖”的评奖活动。之后分别进行了较详细汇报。

5. 新能源控制系统与设备情况”第一阶段调研汇报

张邵清副秘书长汇报了“新能源控制系统与设备情况”第一阶段调研汇总情况。

为深入贯彻习近平总书记关于增强产业链供应链自主可控能力的重要论述、加快绿色转型发展、构建以新能源为主体的新型电力系统，中国自动化学会发电

自动化专业委员会根据有关部门意见，开展新能源发电控制系统与设备运行情况调研，于2022年11月上旬发出第一阶段函调通知。至12月初，收到华电、华能、国家能源、国家电投、皖能、晋能等6家发电集团公司所属的10家新能源企业反馈的调研资料，工作组进行了整理汇编。

10家企业中，5家区域公司控制系统实现软硬件国产化，2家区域公司部分场站实现国产化，3家区域公司对国外技术和硬件设备依赖程度较高。但控制系统中的操作系统、数据库和MMS通信软件等软件大量采用国外产品，存在潜在安全漏洞和卡脖子风险。其国产化控制系统基本满足现有使用需求，基本功能、关键性能指标和安全防护指标等达到甚至优于国家和行业标准。但与国外相比，存在稳定性不足、故障率高、个别功能欠完善等问题。工作组将于开展第二阶段的现场调研活动，收集更多的第一手资料，然后汇总存在的问题，组织专题研讨，共同探讨问题解决办法与新能源控制系统的可持续发展。

二、专题技术报告

专题技术报告阶段，崔青汝委员（国能新能源研究院副院长）作“大型能源集团数字化转型战略及实践”报告。

崔院长结合多年来国家能源

集团能源和行业数字化转型发展轨迹及实践经验，从转型的背景与态势、总体思路与技术路线、加快能源行业转型升级的建议、国家能源集团数字化转型实践四个方面，对大型能源集团数字化转型战略及未来的发展进行了报告。提出智慧能源体系将以零碳转型、能源转型、数字化转型等三大转型为核心，推动能源企业向能源流、信息流、碳流、价值流“四流合一”的能源互联网转型跨越发展。

在总体思路与技术路线方面提出了“1537”转型总体框架，其中：“1”个总体目标、“5”个转型方向、“3”个系统闭环、7个技术能力。从推动共性技术突破、推进行业转型升级、开展应用试点示范、健全发展支撑体系建议加快能源行业转型升级。并就国家能源集团数字化转型实践方面进行了具体介绍和分析。

三、2023年专委会工作计划讨论

委员们对专委会秘书处提出的2023年专委会工作计划，进行了讨论，提出一些修改意见，秘书处将根据委员提出的建议和意见完善2023工作计划。

四、金耀华主任委员总结讲话

会议结束前，金耀华主任委员就会议进行了总结发言。

金耀华主任委员首先肯定专委会2022年的工作，在中国自动化学会的领导下，专委会在疫情的大背景下出色的完成了2022年的既定工作目标，在年度常规工作、国际交流活动、咨询服务工作、智航助学助教工作等各项工作中有序的推进。即将到来的2023年，又有新的机遇与挑战，疫情的大环境的改善将为2023年

的工作带来更好的工作基础。金总从四个方面对2022年的工作进行了总结，同时对2023年的工作强调了以下四方面的工作：

1) 做好国产自主可控的跟踪与回访工作。

2) 火电厂智能化工作要有阶段性的成果。

3) 进一步开展新能源自主可控调研工作，从设备自主可控、智能化两个着力点入手开展调研工作。

4) 在智航助学助教方面，更加注重对学生的实际的帮助上，开展好心理教育咨询工作，充实更多的案例。

最后金耀华主任委员要求，秘书处应细化2023年的工作目标及工作计划，将每个工作项目落实到具体负责人员及相应的时间规划中去。○

CAA发电自动化专委会 供稿

通

知

中国自动化学会关于开展第十九届中国青年女科学家奖和第八届未来女科学家计划候选人推荐工作的通知

为深入学习贯彻习近平新时代中国特色社会主义思想，贯彻落实党的二十大精神 and 中央人才工作会议精神，激发广大女科技工作者的创新创业创造热情，大力弘扬科学家精神，加快建设世界重要人才中心和创新高地、实现高水平科技自立自强，根据《中国科学技术协会 中华全国妇女联合会 中国联合国教科文组织全国委员会关于开展第十九届中国青年女科学家奖和第八届未来女科学家计划候选人提名工作的通知》要求，中国自动化学会开展第十九届中国青年女科学家奖和第八届未来女科学家计划候选人推荐工作。详情请查看：<http://www.caa.org.cn/article/192/3331.html>

关于完善工业和信息化领域科技成果评价机制的实施方案（试行）

为推动工业和信息化领域科技成果评价工作，充分发挥科技成果评价的“指挥棒”作用，促进产业科技创新和高质量发展，根据《中华人民共和国科学技术进步法》《中华人民共和国促进科技成果转化法》等相关法规及《国务院办公厅关于完善科技成果评价机制的指导意见》（国办发〔2021〕26号），制定本方案。

一、总体要求

（一）指导思想以习近平新时代中国特色社会主义思想为指导，深入贯彻党的二十大精神，立足新发展阶段，完整、准确、全面贯彻新发展理念，构建新发展格局，坚持科技自立自强，聚焦工业和信息化领域，突出行业特色，以需求为牵引，以应用为导向，创新评价机制，完善评价体系，以评价助创新、促熟化，推动早出成果、多出成果、出好成果，推动创新链、产业链、资金链、人才链深度融合，提升产业基础高级化和产业链现代化水平，支撑制造强国和网络强国建设。

（二）基本原则全面准确，系统评价。坚持以科技创新质量、绩效、贡献为核心的评价导向，多元评价科技成果的科学价值、技术价值、经济价值、社会价值、文化价值，完整准确反映成果创新水平、转化应用绩效和对经济社会发展的实际贡献。聚焦重点，分类实施。遵循科研渐进性和成果阶段性的客观规律，依据基础研究成果、应用研究成果、技术开发和产业化成果的特点，以及工业和信息化领域关注重点，分类别、分阶段开展评价。企业主体，市场主导。强化企业科技创新主体地位，充分调动各类评价主体的积极性，提高用户、金融投资机构等在评价中的话语权、决策权。发挥市场在资源配置中的决定性作用，营造评价的良好环境。客观公正，严谨规范。坚持公平公正，推动评价工作标准化、规范化。健全完善以用户评价、市场检验和第三方评价为主的评价体系，完善监督机制。引导树立正确价值导向，防止为评而评、滥用评价结果。

（三）工作目标到2027年，

形成一套工业和信息化领域的科技成果评价规范，培育遴选一批水平高、能力强、信誉好的科技成果评价和转移转化机构，健全工业和信息化领域科技成果项目库，形成一套推广应用模式，评价或转化一批高质量科技成果，各类创新主体参与评价的行为更加科学规范、科技成果的数量和质量显著提升、推广应用成效大幅增强、产业技术进步的速度明显加快。

二、主要任务

（一）全面准确评价

1. 破“四唯”立“五元”。破除“唯论文、唯职称、唯学历、唯奖项”倾向，全面准确评价科技成果的科学、技术、经济、社会、文化等“五元”价值。科学价值重点评价在新发现、新原理、新方法方面的独创性贡献。技术价值重点评价重大技术发明，突出在解决产业关键共性技术问题、卡脖子难题、产出更多独门绝技等方面的成效。经济价值重点评价在取得经济效益、形成新的经济增长点方面的成效。社会价值

重点评价在解决人民健康、国防与公共安全、绿色低碳等方面的成效。文化价值重点评价在倡导科学家精神、工程师文化、工匠精神、企业家精神等方面的影响和贡献。

2. 分类分阶段评价。鼓励全国性行业协会、学会、研究会、专业化评估机构（以下简称行业组织和专业机构）等，制定体现行业特性的差异化评价指标和评价规范，针对科技成果在研发、工程化、产业化和商业化等不同阶段的特点，分类别分阶段开展评价。涉及军工、国防、民爆等领域的成果，可根据其非市场化特点另行制定评价指标和评价规范。

3. 推进重点项目评价。做好国家财政资金、政策支持的项目评价，加强中长期评价、后评价和成果回溯，以评价促进标志性成果产出。建立健全重大项目知识产权全流程管理制度，加强知识产权保护，完善知识产权转移转化机制，规范知识产权交易，促进知识产权价值实现。推动相关项目成果向中小企业授权、许可、转让，助力中小企业创新发展。

4. 创新评价方式方法。利用大数据、人工智能等技术，运用技术成熟度、制造成熟度等评价方法，开发智能化评价工具，提高评价质量和效率。探索“揭榜

挂帅”、“以赛代评”、概念验证、技术预测、知识产权评估和扶优式评审等评价方式。强化科技工作者的责任和信用，结合检验检测报告、用户证明、权威机构排名、专利、版权、资质等支撑材料，探索建立“自我声明”的承诺制新模式。

（二）健全评价体系

5. 优化评价机构。择优遴选一批高水平的评价机构，鼓励其牵头制定评价的国家标准、行业标准 and 团体标准，推动评价机构专业化、标准化、规范化发展。支持评价机构提升服务能力，探索线上线下相结合、客观数据分析 and 专家专业评价相结合等服务新模式，充分运用大数据检索分析等新技术赋能。鼓励评价机构延伸服务链条，提供成果到产品全生命周期服务。

6. 强化专家队伍建设。组建高水平的评价专家队伍，坚持专业性、公正性和权威性，突出职业道德、综合素质要求，科学设置专家遴选标准。推动评价机构建立评价专家库并定期动态调整，确保“有进有出”“优进劣出”，规范专家评价行为，完善评价专家的廉政风险防控机制和责任追究制度。依法保障专家权利，确保专家能够客观公正、实事求是开展评价工作。加大复合型人才培养，持续充实评价专业力量。

7. 构建多方评价体系。构建

用户、市场和第三方专业机构评价为主，权威专家参与，创新主体“自我声明”，政府引导监督相结合的多方评价体系。发挥行业组织和专业机构作用，支持其作为独立第三方开展市场化评价。引导企业家、天使投资人、创业投资机构、专业化技术转移机构等各类市场主体提早介入研发活动，对科技成果潜在经济价值、市场估值、发展前景等进行商业化评价。

8. 完善评价制度规范。加强评价制度建设，健全信用体系，规范评价机构、技术经理人、成果拥有者等各类评价参与者的行为，建立评价机构失信名单。强化行业组织和专业机构等自我管理，健全利益关联回避制度，提升评价的权威性和公信力。加强评价全过程的质量管理，推动评价方法、指标、流程等的标准化。

9. 加强公共服务。支持工业产品质量控制和技术评价实验室、产业技术基础公共服务平台、新材料重点平台等加强自身能力建设，增强公益属性，为评价提供检测、试验、分析等公共服务，支撑成果真实性和可靠性验证。支持检验检测机构“服务前移”，重点建设新技术新产品适配应用创新与验证环境，在科技成果产生、评价、产业化等全过程中提供专业服务。

（三）加速成果产业化

10. 用好评价结果。发挥好科技成果评价在技术交易、科技奖励、产业化推广等方面的作用，加强科技成果评价与项目、基地、人才等评价的有机结合，在制定和实施产业政策时，如制造业领航企业、制造业单项冠军企业、专精特新“小巨人”评选，以及首台（套）重大技术装备、首批次新材料、首版次关键软件评定等，适度采信评价结论。支持行业组织和专业机构、被评价方等在工作中科学合理运用评价结果。

11. 促进产学研结合。完善成果评价与金融机构、投资公司的联动机制，通过高质量的成果评价，引导金融机构有效支持创新型企业。充分用好国家产融合作平台，在企业“硬科技”属性评价、“科技产业金融一体化”专项目录遴选、企业上市推荐等工作中，将评价结论作为参考。支持地方、企业与高校加强合作，强化供需对接机制，围绕产业发展、企业需求建立联合式、订单式技术研发新模式，提升科技成果转化成效。

三、保障措施

（一）加强组织实施。工业和信息化部负责完善工业和信息化领域科技成果评价机制，加强对评价工作的统筹指导，会同有关部门组织成果推介和产融对接，加强评价过程的监管，对违规、

造假等行为建立黑名单制度。地方工业和信息化主管部门、行业组织按照本实施方案负责本地区、本行业的评价工作，定期向工业和信息化部推荐优秀科技成果。评价机构按照本实施方案建立规范化的评价管理办法。被评价方对评价成果的真实性负责，不得请托、干扰和影响评价过程。

（二）优化资源配置。依据评价结论，合理配置要素资源，引导企业和社会资本投入，促进高质量科技成果产出。鼓励在国家制造业高质量发展试验区、国家新型工业化产业示范基地、中国软件名城（园）、国家小型微型企业创业创新示范基地、先进制造业集群、中小企业特色产业集群等园区（基地）、各类先导区中，引进科技成果评价和转移转化机构，建设创新成果产业化中心和承载区。对工业和信息化领域重点工程、重大项目成果及制造业领航企业、制造业单项冠军企业、专精特新“小巨人”企业、国家技术创新示范企业、“揭榜挂帅”优胜企业等主体的成果评价情况纳入长期跟踪。

（三）建立信息平台。建设工业和信息化领域科技成果信息平台，分层次、分领域建立成果项目库，提高评价服务的信息化水平。加强与国家产融合作平台、制造业创新成果产业化服务平台、国家先进技术转化应用公共服务

平台等的对接，推动科技创新要素资源加速流动，加强评价结论的互认、互用。

（四）构建良好创新生态。鼓励企业、科研机构、高校和其他组织建立优势互补、分工明确、成果共享、风险共担的协同体系。面向地方产业园区（基地），强化场景驱动，挖掘符合其应用场景、技术需求的科技成果，为地方发展提供新动能。鼓励部属高校、科研院所开展校企校地对接，支持企业开展科技成果产业化应用示范。加强科技成果评价和转移转化机构建设，培育高层次专业队伍。挖掘评价典型案例和成功做法，形成示范并全面推广。

（五）释放部属单位活力。鼓励部属高校、科研院所在职称评审、绩效考核等评价过程中，加大科技成果转化等实际贡献的权重。鼓励部属单位完善科技成果转化尽职免责机制，创新职务科技成果管理体系。部属单位应建设全流程贯通的知识产权管理服务体系，推动专利导航、预警、布局及申请前评估等制度的落地实施，加大高价值专利培育，推广专利价值挖掘与二次开发，逐步完善专利放弃维持前的安全审核机制。鼓励部属单位实施专利开放许可、“先使用后付费”等新机制新模式。○

来源：工业和信息化部办公厅

2023 年度国家自然科学基金改革举措

自然科学基金委按照科学基金系统性改革总体部署，以构建理念先进、制度规范、公正高效的新时代科学基金体系为目标，在 2023 年推出以下改革举措。

一、持续开展分类评审

重点项目、面上项目和青年科学基金项目继续开展基于“鼓励探索，突出原创；聚焦前沿，独辟蹊径；需求牵引，突破瓶颈；共性导向，交叉融通”四类科学问题属性的分类评审，实施分类评审项目占科学基金项目申请比重保持稳定；同时，通过科学问题属性案例库、宣讲视频等方式，不断提高申请人和评审专家对四类科学问题属性的理解，引导广大科研人员持续提升科学问题凝练能力，提高选题质量。

二、进一步完善科学基金人才资助体系

稳步扩大青年科学基金项目资助规模，培养基础研究人才后备军。加强优秀青年科学基金项目、国家杰出青年科学基金项目与同层次国家科技人才计划的统筹衔接，提升资助效益。优化创

新研究群体项目的申请和评价机制，继续开展基础科学中心项目考核评估及延续资助工作，有意识地发现和培养更多具有战略科学家潜质的高层次复合型人才。

继续面向香港特别行政区、澳门特别行政区依托单位青年科学技术人员开放优秀青年科学基金项目（港澳）和青年科学基金项目申请，加强优秀青年人才培养。做好优秀青年科学基金项目（海外）的组织实施工作，吸引更多海外优秀青年人才回国（来华）工作。持续加大外国学者研究基金项目实施力度，分层次、全方位资助优秀外国学者来华开展高水平合作研究。

三、深入推进原创探索计划

瞄准提升原始创新能力，深入推进原创探索计划。遴选具有非共识、颠覆性、高风险等特征的原创项目，引导和激励广大科研人员投身原创性基础研究工作；具体情况详见后续发布的《2023 年度国家自然科学基金原创探索计划项目申请指南》。同时，做好资助项目跟踪管理和结题评估，鼓励探索、宽容失败，对后续有

望获得突破性原创成果的项目进行延续资助，持续完善和创新评审管理机制。

四、促进学科交叉融合

持续推进学科交叉融合和交叉科学部的学部建设工作。进一步完善交叉科学研究领域的项目设置，建立符合交叉科学研究特征的评价机制，加强顶层设计，强化交叉科学领域的多学科共性科学问题凝练。面向世界科学前沿和国家重大需求，围绕人类社会面临的综合性、复杂性重大问题，组织多学科交叉合作、协同攻关。营造有利于学科交叉和交叉科学发展的学术环境，打破学科壁垒，开拓学科前沿，产生学科生长点，探索新的研究范式，培养交叉科学创新人才和团队。

五、继续开展“负责任、讲信誉、计贡献”（RCC）评审机制试点工作

坚持正面引导，继续开展 RCC 评审机制试点工作。准确解读 RCC 评审机制政策，广泛宣传评审专家行为规范，为负责任评审营造良好氛围。认真总结试

年科学基金项目、地区科学基金项目、重点项目、优秀青年科学基金项目、国家杰出青年科学基金项目、创新研究群体项目、基础科学中心项目、国家重大科研仪器研制项目和重大项目，其研究期限由信息系统结合项目类型自动生成，申请人不可更改，为科学技术人员提供更便捷的服务。同时，继续实行填写主要参与者时不再列入学生，项目申请书主要参与者信息在线采集，填写论文等研究成果时根据其发表时的真实情况如实规范列出所有作者署名、不再标注第一作者或通

讯作者等举措，减轻科学技术人员填表负担。

十一、加强依托单位管理

完善依托单位准入和退出机制，强化依托单位动态管理，严守科学基金“入口”；建立并完善依托单位管理体系，探索依托单位分类分级的管理方法；进一步加强依托单位项目管理过程监督，深入研究并完善信誉评价与惩戒机制，压实依托单位履行管理与监督的主体责任；落实科学基金深化改革实施方案，切实推动科学基金项目申请质量提升。

十二、深入实施科学基金学风建设行动计划

深入实施“教育、激励、规范、监督、惩戒”五个方面相互支撑、有机融合、标本兼治的科学基金学风建设体系。以教育为根本，以激励为引导，以规范为准绳，以监督为抓手，以惩戒为警示，坚持远近结合、标本兼治，进一步加强“十四五”时期科学基金科研诚信建设、科技伦理治理和作风学风建设。○

来源：国家自然科学基金委员会



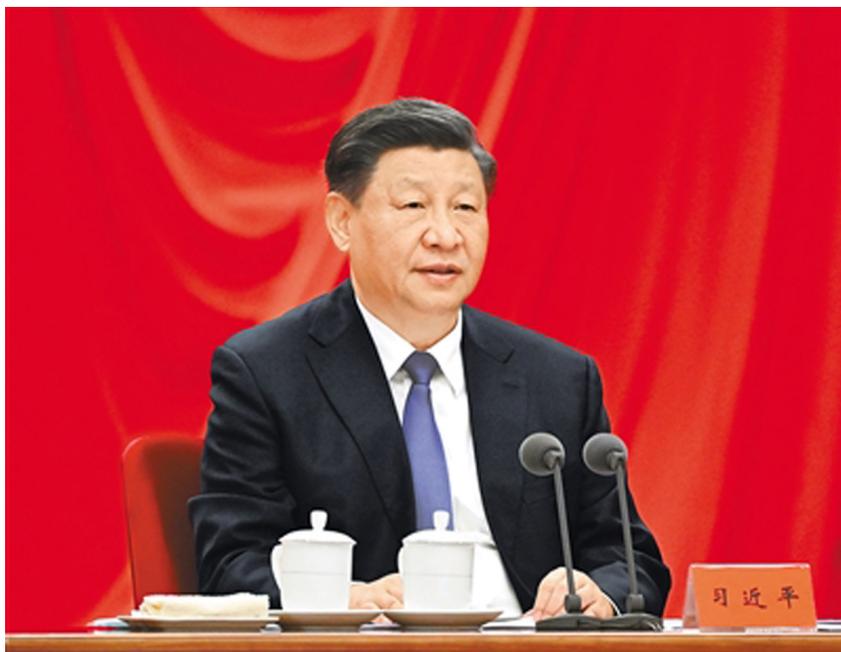
JAS 和自动化学报 3 篇论文入选中国百篇最具影响国际 / 国内论文

近日，中国科学技术信息研究所召开 2022 中国科技论文统计结果发布会，发布了 2021 年“中国百篇最具影响国际学术论文”、“中国百篇最具影响国内学术论文”。JAS 发表的 1 篇论文入选本年度“中国百篇最具影响国际学术论文”，这是 JAS 发表论文首次获此称号；《自动化学报》发表的 2 篇论文入选本年度“中国百篇最具影响国内学术论文”，学报累计 4 篇论文获此称号。详情请查看：<http://www.caa.org.cn/article/191/3323.html>

习近平在二十届中央纪委二次全会上发表重要讲话

习近平在二十届中央纪委二次全会上发表重要讲话强调

一刻不停推进全面从严治党 保障党的二十大决策部署贯彻落实



1月9日，中共中央总书记、国家主席、中央军委主席习近平在中国共产党第二十届中央纪委检查委员会第二次全体会议上发表重要讲话。新华社记者 李学仁摄

■ 要站在事关党长期执政、国家长治久安、人民幸福安康的高度，把全面从严治党作为党的长期战略、永恒课题，始终坚持问题导向，保持战略定力，发扬彻底的自我革命精神，永远吹冲锋号，把严的基调、严的措施、严的氛围长期坚持下去，把党的伟大自我革命进行到底。要坚持严管和厚爱结合、激励和约束并重，坚持“三个区分开来”，更

好激发广大党员、干部的积极性、主动性、创造性，形成奋进新征程、建功新时代的浓厚氛围和生动局面

■ 治国必先治党，党兴才能国强。新时代十年，党中央把全面从严治党纳入“四个全面”战略布局，刀刃向内、刮骨疗毒，猛药祛疴、重典治乱，使党在革命性锻造中变得更加坚强有力。全面从严治党永远在路上，要时

刻保持解决大党独有难题的清醒和坚定。如何始终不忘初心、牢记使命，如何始终统一思想、统一意志、统一行动，如何始终具备强大的执政能力和领导水平，如何始终保持干事创业精神状态，如何始终能够及时发现和解决自身存在的问题，如何始终保持风清气正的政治生态，都是我们这个党必须解决的独有难题。解决这些难题，是实现新时代新征程党的使命任务必须迈过的一道坎，是全面从严治党适应新形势新要求必须啃下的硬骨头

■ 构建全面从严治党体系是一项具有全局性、开创性的工作。新时代十年，我们党不断深化对自我革命规律的认识，不断推进党的建设理论创新、实践创新、制度创新，初步构建起全面从严治党体系。全面从严治党体系应是一个内涵丰富、功能完备、科学规范、运行高效的动态系统。健全这个体系，需要坚持制度治党、依规治党，更加突出党的各方面建设有机衔接、联动集成、协同协调，更加突出体制机制的

健全完善和法规制度的科学有效，更加突出运用治理的理念、系统的观念、辩证的思维管党治党建设党。要坚持内容上全覆盖、对象上全覆盖、责任上全链条、制度上全贯通，进一步健全全面从严治党体系，使全面从严治党各项工作更好体现时代性、把握规律性、富于创造性

中共中央总书记、国家主席、中央军委主席习近平9日上午在中国共产党第二十届中央纪律检查委员会第二次全体会议上发表重要讲话。他强调，要站在事关党长期执政、国家长治久安、人民幸福安康的高度，把全面从严治党作为党的长期战略、永恒课题，始终坚持问题导向，保持战略定力，发扬彻底的自我革命精神，永远吹冲锋号，把严的基调、严的措施、严的氛围长期坚持下去，把党的伟大自我革命进行到底。要坚持严管和厚爱结合、激励和约束并重，坚持“三个区分开来”，更好激发广大党员、干部的积极性、主动性、创造性，形成奋进新征程、建功新时代的浓厚氛围和生动局面。

中共中央政治局常委李强、赵乐际、王沪宁、蔡奇、丁薛祥出席会议。中共中央政治局常委、中央纪律检查委员会书记李希主持会议。

习近平指出，治国必先治党，党兴才能国强。新时代十年，党

中央把全面从严治党纳入“四个全面”战略布局，刀刃向内、刮骨疗毒，猛药祛疴、重典治乱，使党在革命性锻造中变得更加坚强有力。全面从严治党永远在路上，要时刻保持解决大党独有难题的清醒和坚定。如何始终不忘初心、牢记使命，如何始终统一思想、统一意志、统一行动，如何始终具备强大的执政能力和领导水平，如何始终保持干事创业精神状态，如何始终能够及时发现和解决自身存在的问题，如何始终保持风清气正的政治生态，都是我们这个党必须解决的独有难题。解决这些难题，是实现新时代新征程党的使命任务必须迈过的一道坎，是全面从严治党适应新形势新要求必须啃下的硬骨头。

习近平强调，构建全面从严治党体系是一项具有全局性、开创性的工作。新时代十年，我们党不断深化对自我革命规律的认识，不断推进党的建设理论创新、实践创新、制度创新，初步构建起全面从严治党体系。全面从严治党体系应是一个内涵丰富、功能完备、科学规范、运行高效的动态系统。健全这个体系，需要坚持制度治党、依规治党，更加突出党的各方面建设有机衔接、联动集成、协同协调，更加突出体制机制的健全完善和法规制度的科学有效，更加突出运用治理的理念、系统的观念、辩证的思

维管党治党建设党。要坚持内容上全覆盖、对象上全覆盖、责任上全链条、制度上全贯通，进一步健全全面从严治党体系，使全面从严治党各项工作更好体现时代性、把握规律性、富于创造性。

习近平指出，要以有力政治监督保障党的二十大决策部署落实见效。政治监督是督促全党坚持党中央集中统一领导的有力举措，要在具体化、精准化、常态化上下更大功夫。要推动党的二十大精神、党中央决策部署同部门、行业、领域实际紧密结合，看党的二十大关于全面贯彻新发展理念、着力推动高质量发展、主动构建新发展格局等战略部署落实了没有、落实得好不好；看党中央提出的重点任务、重点举措、重要政策、重要要求贯彻得怎么样；看属于本地区本部门本单位的职责有没有担当起来。要及时准确发现有令不行、有禁不止，做选择、搞变通、打折扣，不顾大局、搞部门和地方保护主义，照搬照抄、上下一般粗等突出问题，切实打通贯彻执行中的堵点淤点难点。要推动完善党中央重大决策部署落实机制，以有力有效日常监督促进各项政策落实落地。

习近平强调，制定实施中央八项规定，是我们党在新时代的徙木立信之举，必须常抓不懈、久久为功，直至真正化风成俗，以优良党风引领社风民风。要继

续纠治享乐主义、奢靡之风，把握作风建设地区性、行业性、阶段性特点，抓住普遍发生、反复出现的问题深化整治，推进作风建设常态化长效化。要把纠治形式主义、官僚主义摆在更加突出位置，作为作风建设的重点任务，研究针对性举措，科学精准靶向整治，动真碰硬、务求实效。

习近平指出，纪律是管党治党的“戒尺”，也是党员、干部约束自身行为的标准和遵循。要把纪律建设摆在更加突出位置，党规制定、党纪教育、执纪监督全过程都要贯彻严的要求，既让铁纪“长牙”、发威，又让干部重视、警醒、知止，使全党形成遵规守纪的高度自觉。每一个共产党员特别是领导干部都要牢固树立党章意识，更加自觉地学习党章、遵守党章、贯彻党章、维护党章，用党章党规党纪约束自己的一言一行，增强纪律意识、规矩意识，进一步养成在受监督和约束的环境中工作生活的习惯。

习近平强调，反腐败斗争形势依然严峻复杂，遏制增量、清除存量的任务依然艰巨。必须深化标本兼治、系统治理，一体推进不敢腐、不能腐、不想腐。要在不敢腐上持续加压，始终保持零容忍震慑不变、高压惩治力量常在，坚决惩治不收敛不收手、胆大妄为者，坚决查处政治问题和经济问题交织的腐败，坚决防

止领导干部成为利益集团和权势团体的代言人、代理人，坚决防止政商勾连、资本向政治领域渗透等破坏政治生态和经济发展环境。要对比较突出的行业性、系统性、地域性腐败问题进行专项整治。要在不能腐上深化拓展，前移反腐关口，深化源头治理，加强重点领域监督机制改革和制度建设，健全防治腐败滋生蔓延的体制机制。要在不想腐上巩固提升，更加注重正本清源、固本培元，加强新时代廉洁文化建设，涵养求真务实、团结奋斗的时代新风。要把不敢腐、不能腐、不想腐有效贯通起来，三者同时发力、同向发力、综合发力，把不敢腐的震慑力、不能腐的约束力、不想腐的感召力结合起来。进一步健全完善惩治行贿的法律法规，完善对行贿人的联合惩戒机制。严厉打击那些所谓“有背景”的“政治骗子”。

习近平指出，健全党统一领导、全面覆盖、权威高效的监督体系，是实现国家治理体系和治理能力现代化的重要标志。党委（党组）要发挥主导作用，统筹推进各类监督力量整合、程序契合、工作融合。要持续深化纪检监察体制改革，做实专责监督，搭建监督平台，织密监督网络，协助党委推动监督体系高效运转。要把巡视利剑磨得更光更亮，勇于亮剑，始终做到利剑高悬、震慑常在。

习近平指出，纪检监察机关是推进全面从严治党的重要力量，使命光荣、责任重大，必须忠诚于党、勇挑重担，敢打硬仗、善于斗争，在攻坚战持久战中始终冲锋在最前面。要坚持以党性立身做事，弘扬伟大建党精神，勇于自我革命，在党风廉政建设和反腐败斗争一线砥砺品格操守，在围绕中心、服务大局中彰显担当作为，在各种风险挑战中筑牢坚强屏障。要增强法治意识、程序意识、证据意识，不断提高纪检监察工作规范化、法治化、正规化水平。要完善内控机制，自觉接受各方面监督，对纪检监察干部从严管理，对系统内的腐败分子从严惩治，坚决防治“灯下黑”。要结合即将在全党开展的主题教育，把纯洁思想、纯洁组织作为突出问题来抓，切实加强政治教育、党性教育，严明法纪，坚决清除害群之马，以铁的纪律打造忠诚干净担当的铁军。

李希在主持会议时指出，习近平总书记发表的重要讲话，深刻分析大党独有难题的形成原因、主要表现和破解之道，深刻阐述健全全面从严治党体系的目标任务、实践要求，对坚定不移深入推进全面从严治党作出战略部署。讲话高屋建瓴、思想深邃、内涵丰富、论述精辟，具有很强的政治性、指导性、针对性，是深入推进全面从严治党的根本遵循，

是新时代新征程纪检监察工作高质量发展的根本指引。要深入学习贯彻习近平总书记重要讲话精神，深刻领悟“两个确立”的决定性意义，进一步增强“四个意识”、坚定“四个自信”、做到“两个维护”，在新时代新征程上一刻不停推进全面从严治党，深入推进新时代党的建设新的伟大工程，为全面建设社会主义现代化国家开好局起好步提供坚强保障。

中共中央政治局委员、中央书记处书记，是二十届中央委员的其他党和国家领导同志、中央军委委员出席会议。

中央纪律检查委员会委员，中央和国家机关各部门、各人民团体主要负责同志，军队有关单位主要负责同志等参加会议。会议以电视电话会议形式举行，各省、自治区、直辖市和新疆生产建设兵团以及军队有关单位设分会场。

中国共产党第二十届中央纪律检查委员会第二次全体会议于1月9日在北京开幕。中央纪律检查委员会常务委员会主持会议。9日下午李希代表中央纪律检查委员会常务委员会作题为《深入学习贯彻党的二十大精神，在新征程上坚定不移推进全面从严治党》的工作报告。○

来源：人民日报

以中国式现代化全面推进中华民族伟大复兴

党的二十大鲜明提出新时代新征程党的使命任务，发出全面建设社会主义现代化国家、全面推进中华民族伟大复兴的动员令。习近平总书记郑重宣示：“从现在起，中国共产党的中心任务就是团结带领全国各族人民全面建成社会主义现代化强国、实现第二个百年奋斗目标，以中国式现代化全面推进中华民族伟大复兴。”

在新征程上，必须全面把握中国式现代化的中国特色、本质要求和重大原则，锚定全面建成社会主义现代化强国的使命任务，按照党的二十大作出的重大战略部署，坚定信心、锐意进取，不断夺取全面建设社会主义现代化国家新胜利。

中国式现代化是实现中华民族伟大复兴的光明大道

筑铁路、铺公路、建大港，连通大江南北、内陆边疆……100多年前，孙中山先生在《建国方略》中设想着中国现代化景象。如今，高铁飞驰，公路密布，“神舟”问天，“嫦娥”落月，“祝融”探火，“羲和”逐日……中国的现

代化程度远不是孙中山先生当年能够设想的。

中国式现代化道路是我们党团结带领全国各族人民开创和形成的。近代以来，历经磨难中华民族苦苦探索民族复兴的现代化之路，但在中国共产党诞生之前，都没有取得成功。中国共产党自成立起团结带领中国人民所进行的一切奋斗，就是为了把我国建设成为现代化强国，实现中华民族伟大复兴。新中国成立以后，我们党孜孜以求，带领人民对中国现代化建设进行了艰辛探索。社会主义革命和建设时期，我们党提出努力把我国逐步建设成为一个具有现代农业、现代工业、现代国防和现代科学技术的社会主义强国目标。改革开放和社会主义现代化建设新时期，我们党提出“中国式的现代化”论断，制定了到21世纪中叶分三步走、基本实现社会主义现代化的发展战略。党的十八大以来，中国特色社会主义进入新时代。党的十九大站在新的更高的历史起点上，对实现第二个百年奋斗目标作出分两个阶段推进的战略安

排，提出到 2035 年基本实现社会主义现代化，到本世纪中叶把我国建成富强民主文明和谐美丽的社会主义现代化强国。在新中国成立特别是改革开放以来长期探索和实践基础上，经过党的十八大以来在理论和实践上的创新突破，我们党成功推进和拓展了中国式现代化，开辟了以中国式现代化全面推进中华民族伟大复兴的新境界。

在中国式现代化道路上，我国仅用几十年的时间就走完了西方发达国家几百年走过的工业化历程，经济实力、科技实力、综合国力、国际影响力持续增强，创造了世所罕见的经济快速发展和社会长期稳定两大奇迹。

在中国式现代化道路上，我们全面建成小康社会、实现第一个百年奋斗目标，中华民族迎来了从站起来、富起来到强起来的伟大飞跃，实现中华民族伟大复兴进入了不可逆转的历史进程。

在中国式现代化道路上，我们创造了人类文明新形态，拓展了发展中国家走向现代化的途径，给世界上那些既希望加快发展又希望保持自身独立性的国家和民族提供了全新选择。

历史和实践充分证明，中国式现代化扎根中国大地，切合中国实际，这条道路不仅走得对、走得通，而且走得稳、走得好。在新征程上，坚持以中国式现代

化全面推进中华民族伟大复兴，一定能够不断创造新的发展奇迹，为发展自身和造福世界作出新的更大贡献。2022 年 11 月习近平总书记在会见法国总统马克龙时，马克龙表示，中国式现代化发展模式令人钦佩。

深刻领会中国式现代化的中国特色和本质要求

“中国式现代化，是中国共产党领导的社会主义现代化，既有各国现代化的共同特征，更有基于自己国情的中国特色。”习近平总书记在党的二十大报告中作出的这一论断，深刻揭示了中国式现代化的本质属性，为我们走好社会主义现代化强国之路提供了科学指导。

一切成功发展振兴的民族，都是找到了适合自己实际的道路的民族。现代化是各国人民的共同期待和目标，但每个国家都应结合自身实际作出路径选择。党的二十大报告概括了中国式现代化的中国特色，即中国式现代化是人口规模巨大的现代化，是全体人民共同富裕的现代化，是物质文明和精神文明相协调的现代化，是人与自然和谐共生的现代化，是走和平发展道路的现代化。中国式现代化，是对西方以资本为中心的现代化、两极分化的现代化、物质主义膨胀的现代化、对外扩张掠夺的现代化老路的超

越，是对世界现代化理论和实践的丰富和发展。

推进中国式现代化，要遵循现代化的一般规律，更要切实把握好基于自己国情的中国特色。

我们要深刻认识到，我国 14 亿多人口整体迈进现代化社会，规模超过现有发达国家人口的总和，这将是人类发展史上前所未有的大事，艰巨性和复杂性前所未有，发展途径和推进方式也必然具有自己的特点。

我们要深刻认识到，共同富裕是中国特色社会主义的本质要求，也是一个长期的历史过程。要坚持把实现人民对美好生活的向往作为现代化建设的出发点和落脚点，着力促进全体人民共同富裕。

我们要深刻认识到，物质富足、精神富有是社会主义现代化的根本要求。物质贫困不是社会主义，精神贫乏也不是社会主义。要不断厚植现代化的物质基础，同时大力发展社会主义先进文化，促进物的全面丰富和人的全面发展。

我们要深刻认识到，人与自然是生命共同体，无止境地向自然索取甚至破坏自然必然会遭到大自然的报复。要像保护眼睛一样保护自然和生态环境，坚定不移走生产发展、生活富裕、生态良好的文明发展道路，实现中华民族永续发展。

我们要深刻认识到，我国不走一些国家通过战争、殖民、掠夺等方式实现现代化的老路，那种损人利己、充满血腥罪恶的老路给广大发展中国家人民带来深重苦难。要坚定站在历史正确的一边、站在人类文明进步的一边，在坚定维护世界和平与发展中谋求自身发展，又以自身发展更好维护世界和平与发展。

归根结底，我们要深刻认识到，中国式现代化是中国共产党领导的社会主义现代化，是为大多数人谋利益的现代化；而不是资本主义现代化、为少数人谋利益的现代化。

中国式现代化道路怎么走？党的二十大报告明确提出了中国式现代化的本质要求，这就是：坚持中国共产党领导，坚持中国特色社会主义，实现高质量发展，发展全过程人民民主，丰富人民精神世界，实现全体人民共同富裕，促进人与自然和谐共生，推动构建人类命运共同体，创造人类文明新形态。这些要求，核心是坚持党的领导和中国特色社会主义制度，并涵盖了富强、民主、文明、和谐、美丽的奋斗目标，彰显了中国式现代化的世界意义和世界贡献。这些要求，是党深刻总结我国和世界其他国家现代化建设的历史经验，对我国这样一个东方大国如何加快实现现代化在认识上不断深入、战略上不

断完善、实践上不断丰富而形成的思想理论结晶。这些要求，是以中国式现代化全面推进中华民族伟大复兴的基本遵循，我们必须深刻领会、系统把握，并贯彻落实到各项工作中。

必须牢牢把握推进中国式现代化的重大原则

“中国式现代化是中国共产党和中国人民长期实践探索的成果，是一项伟大而艰巨的事业。惟其艰巨，所以伟大；惟其艰巨，更显荣光。”习近平总书记在二十届中央政治局常委同中外记者见面时的一席话催人奋进。

全面建设社会主义现代化国家，是一项伟大而艰巨的事业，前途光明，任重道远。必须增强忧患意识，坚持底线思维，做到居安思危、未雨绸缪，准备经受风高浪急甚至惊涛骇浪的重大考验。党的二十大报告提出了前进道路上必须牢牢把握的5条重大原则，即坚持和加强党的全面领导，坚持中国特色社会主义道路，坚持以人民为中心的发展思想，坚持深化改革开放，坚持发扬斗争精神。

我们必须牢牢把握，全面建设社会主义现代化国家、全面推进中华民族伟大复兴，关键在党。要坚决维护党中央权威和集中统一领导，把党的领导落实到党和国家事业各领域各方面各环节，

确保我国社会主义现代化建设正确方向。

我们必须牢牢把握，中国特色社会主义道路是当代中国大踏步赶上时代、引领时代发展的康庄大道。要坚持把国家和民族发展放在自己力量的基点上，坚持把中国发展进步的命运牢牢掌握在自己手中。

我们必须牢牢把握，党的根基在人民、血脉在人民、力量在人民，人民是党执政兴国的最大底气。要维护人民根本利益，增进民生福祉，让现代化建设成果更多更公平惠及全体人民。

我们必须牢牢把握，改革开放是党和人民大踏步赶上时代的重要法宝，是决定当代中国命运的关键一招。要深入推进改革创新，坚定不移扩大开放，不断增强社会主义现代化建设的动力和活力。

我们必须牢牢把握，敢于斗争、敢于胜利，是党和人民不可战胜的强大精神力量。要增强全党全国各族人民的志气、骨气、底气，不信邪、不怕鬼、不怕压，依靠顽强斗争打开事业发展新天地。

前进道路上必须牢牢把握的5条重大原则，是在全面总结党的百年奋斗历史经验特别是中国特色社会主义发展实践经验、深刻把握我国发展新的历史特点、统筹考虑当前和今后一个时期发展

目标的基础上提出来的，内涵丰富、意义重大，我们要深入学习、全面领悟、整体把握，不折不扣贯彻落实。

为全面建设社会主义现代化国家开好局起好步

在全面建成小康社会、实现第一个百年奋斗目标的基础上，党的二十大报告对全面建成社会主义现代化强国两步走战略安排进行宏观展望，进一步明确了到2035年我国发展的总体目标，强调“到本世纪中叶，把我国建设成为综合国力和国际影响力领先的社会主义现代化强国”。报告强调未来5年是全面建设社会主义现代化国家开局起步的关键时期，并提出了经济、政治、文化、社会、生态、安全、国防、外交等方面的主要目标任务。学习贯彻党的二十大精神，就要深刻理解党的二十大对全面建设社会主义现代化国家作出的战略部署，紧密联系我国发展面临的新的战略机遇、新的战略任务、新的战略阶段、新的战略要求、新的战略环境，深刻认识实现全面建设社会主义现代化国家各项目标任务的艰巨性和复杂性，增强贯彻落实的自觉性和坚定性。

高质量发展是全面建设社会主义现代化国家的首要任务。今年是全面贯彻落实党的二十大精神的开局之年，做好经济工作意义重大。当前，世界之变、时代之变、历史之变正以前所未有的方式展开，我国经济恢复的基础尚不牢固，需求收缩、供给冲击、预期转弱三重压力仍然较大，外部环境动荡不安，给我国经济带来的影响加深。但要看到，我国经济韧性强、潜力大、活力足，长期向好的基本面没有变，各项政策效果持续显现，今年经济运行有望总体回升，要坚定做好经济工作的信心。

“什么时候没有困难？一个过一个过，年年过、年年好，中华民族5000多年来都是这样。爬坡过坎，关键是提振信心。”在2022年12月召开的中央经济工作会议上，习近平总书记意味深长地说，“我们有社会主义市场经济的体制优势，有超大规模市场的需求优势，有产业体系配套完善的供给优势，有勤劳智慧的广大劳动者和企业家等人力优势，只要把各方面的优势和活力真正激发出来，就能够加快构建新发展格局，在激烈的国际市场竞争和大国战略博弈中始终立于不败

之地。”

一分部署，九分落实。中央经济工作会议明确了做好今年经济工作的总体要求、主要任务和政策举措，为我们指明了前进方向、提供了根本遵循。要把思想和行动统一到党的二十大精神和党中央关于经济工作的决策部署上来，坚持稳中求进工作总基调，完整、准确、全面贯彻新发展理念，加快构建新发展格局，着力推动高质量发展，更好统筹疫情防控和经济社会发展，更好统筹发展和安全，全面深化改革开放，大力提振市场信心，把实施扩大内需战略同深化供给侧结构性改革有机结合起来，突出做好稳增长、稳就业、稳物价工作，有效防范化解重大风险，推动经济运行整体好转，实现质的有效提升和量的合理增长，为全面建设社会主义现代化国家开好局起好步。

“党的二十大制定了当前和今后一个时期党和国家的大政方针，描绘了以中国式现代化全面推进中华民族伟大复兴的宏伟蓝图。让我们踏上新征程，向着新的奋斗目标，出发！”

来源：求是网

学会办事机构党支部组织召开 2022 年度民主生活会

2023 年 1 月 12 日，学会办事机构党支部组织召开 2022 年度民主生活会，党支部书记、秘书长张楠，党支部副书记吕爱英，副秘书长王坛、石红芳及支部全体党员与积极分子参加会议。会议由学会党支部书记、秘书长张楠主持。

张楠书记首先就一年来党支部在贯彻落实党的二十大精神、弘扬科学家精神及夯实基层党组织建设等方面工作进行总结汇报，并对 2023 年工作作具体部署。她指出，学会党支部要进一步发挥

战斗堡垒作用，推动党建与业务工作深度融合，指引学会牢记初心使命，把握政策方向，深化开放合作，强化服务意识，服务国家大局。

随后，遵循“团结一批评一团结”方针开展了党员批评与自我批评工作，党支部书记张楠带头发言，率先进行批评与自我批评，给党员带好头，党员和积极分子纷纷结合自身思想认识和工作职责，踊跃发言，坦诚交流，逐一查摆自身存在的问题和不足、

深刻剖析思想根源。

通过本次集中学习，党员们加深了对习近平新时代中国特色社会主义思想和党的二十大精神感悟与理解。通过结合学会工作实际，开展批评和自我批评，全体党员和积极分子夯实了思想基础，明确今后工作努力方向，并纷纷表示，将继续积极主动作为，为推动自动化高水平自立自强贡献力量。○

学会办事机构党支部 供稿



线下参会代表



2023中国自动化大会

CHINA AUTOMATION CONFERENCE

智联万物强实体 自强自控创未来



主办单位：中国自动化学会 承办单位：重庆邮电大学

中国自动化大会是由中国自动化学会主办的国内最高层次的自动化、信息与智能科学领域的大型综合性学术会议，2023中国自动化大会将于2023年11月在重庆召开，此次中国自动化大会由重庆邮电大学承办。

2023中国自动化大会将为全球自动化、信息与智能科学领域的专家学者和产业界的同仁提供展示创新成果、展望未来发展的高端学术平台，加强不同学科领域的交叉融合，引领自动化、信息与智能科学与技术的发展。

2023中国自动化大会

征文范围

本次大会设多个特色论坛，征文领域近30种。热忱欢迎全国各高等院校、科研院所和企业事业单位中从事相关领域研究的科技工作者积极投稿，特别希望征集能反映各单位研究特色的学术论文或长摘要(summary)。

论文投稿要求

1. 来稿未曾公开发表过，具备真实性和原创性。请勿涉及国家秘密。
2. 凡投稿论文被录用且未作特殊声明者，视为已同意授权出版。
3. 中英文论文篇幅均限制4-6页。

长摘要投稿要求

1. 长摘要需包括研究背景和意义、主要研究工作、实验或仿真、结论以上所有内容。
2. 长摘要论文将被收录进论文集，但不进IEEE Xplore、EI、CNKI等检索，已发表的结果也可以投稿。
3. 长摘要长度不超过4页。
4. 长摘要论文注册费与普通论文相同。

征文领域(包括但不限于)

1. 工业互联网与智能制造
2. 智能网联汽车与交通
3. 空地协同与自主控制
4. 大数据智能化
5. 元宇宙与平行系统
6. 能源互联与绿色制造
7. 机器人智能控制
8. 复杂系统建模、控制与优化
9. 智能控制理论与方法
10. 社会计算与社会系统管理
11. 无人系统的信息处理与控制
12. 多智能体编队与协同
13. 模式识别与人工智能
14. 脑机接口与认知计算
15. 空间飞行器控制
16. 先进传感技术与仪器仪表
17. 船舶自动控制与综合操控
18. 网络集群与网络化控制
19. 医学图像、生物信息与仿生控制
20. 复杂系统理论与方法
21. 流程工业智能优化制造
22. 类脑智能与深度学习
23. 故障诊断与系统运行安全
24. 无线传感网与数据融合
25. 海洋环境监测与仿真
26. 机电液一体化自动化控制
27. 物流系统与自动化
28. 其他

专题会议

1. 学术专题论坛
2. 产业发展论坛
3. 科技奖励论坛
4. 青年人才论坛
5. 教育专题论坛
6. 女科技工作者论坛
7. 期刊出版专题论坛
8. 展览展示

论文出版

大会将出版CAC2023论文集(U盘版)。2013年以来的历届会议英文论文全部被IEEE Xplore收录，并被EI检索。经过专家评审，本届大会部分优秀论文将被推荐到《IEEE/CAA Journal of Automatica Sinica》、《Digital Communications and Networks》、《自动化学报》、《智能科学与技术学报》和《重庆邮电大学学报<自然科学版>》等国内外SCI/EI收录权威期刊发表。

时间节点

投稿开始时间:2023.03.01
 征稿截止日期:2023.06.30
 录用通知日期:2023.07.31
 论文终稿日期:2023.08.31





中国自动化学会

中国自动化学会(Chinese Association of Automation, 缩写CAA)于1961年成立,是我国最早成立的国家一级学术团体之一,是中国科学技术协会的组成部分,是发展我国自动化科技事业的重要社会力量。学会现有个人会员近9万人,团体会员单位300余个,专业委员会59个,工作委员会9个,30个省、自治区、直辖市设有地方学会组织,覆盖了我国自动化科学技术领域的各个层面。

中国自动化学会在改革中求发展,不断加强群众组织力、学术引领力、社会公信力和国际影响力。近年来,中国自动化学会重点从学术交流与应用推广、组织建设与会员服务、科技评估与人才评价、课题研究与决策支撑、科学普及与继续教育等方面开拓创新,推动中国自动化科学和事业的发展 and 壮大,成为连接政府、产业、学术、科研、会员的重要纽带,致力于成为国内外有影响力的现代社会团体组织。

学会品牌学术活动

- 中国自动化大会 ·中国认知计算与混合智能学术大会
- 国家智能车发展论坛 ·国家机器人发展论坛 ·国家智能制造论坛
- 青年菁英系列活动 ·智能自动化学科前沿讲习班 ·钱学森国际杰出科学奖系列讲座
- 中国控制会议 ·中国过程控制会议 ·青年学术年会

学会奖励奖项

- 钱学森奖 ·杨嘉墀科技奖 ·CAA科学技术奖励 ·CAA优秀博士学位论文奖
- 中国自动化与人工智能创新团队奖 ·CAA高等教育教学成果奖 ·CAA青年科学家奖
- 企业创新示范单位 ·杰出自动化工程师 ·小微创业示范单位 ·智慧系统创新解决方案示范单位

学会主办期刊

- 中国自动化学会通讯 ·自动化学报 ·自动化学报(英文版)
- 信息与控制 ·机器人 ·模式识别与人工智能 ·电气传动
- 自动化博览 ·计算技术与自动化



官方微信



官方微博

地址:北京市海淀区中关村东路95号自动化大厦

网址:<http://www.caa.org.cn/>

电话:010-62522472

传真:010-62522248

邮箱:caa@ia.ac.cn

邮编:100190