

# 海外学人

Overseas Scholars

2017 年特刊

ISSN: 1556-861X

信息技术专刊

Special Issue on Information Technology



中國旅美科技協會

Chinese Association for Science and Technology, USA (CAST-USA)  
[www.cast-usa.org](http://www.cast-usa.org)

暨庆祝旅美科技协会成立二十五周年特辑

*Celebrating the 25th Anniversary of CAST-USA*

## 中国旅美科技协会名誉顾问

杨振宁 清华大学教授、诺贝尔奖获得者  
周光召 中国科协名誉主席、中国科学院前院长  
朱丽兰 中国科学技术部前部长  
邓文中 美国工程院院士、美国林同琮国际公司董事长  
Dan Shechtman 2011年诺贝尔化学奖获得者

路甬祥 中国科学院前院长、全国人大常委会副委员长  
宋健 中国工程院前院长、全国政协前副主席  
陈香梅 美国国际合作委员会主席  
Mario Capecchi 2007年诺贝尔生理或医学奖获得者

## 中国旅美科技协会机构与主要负责人

总会会长

蒋为民

董事会主席

陈志雄

理事会主席

宋云明

候任会长

潘星华

### 总会负责人

万乐 副会长兼秘书长

占中华 副秘书长

沈琦 副会长(年会)

谢长春 副会长(办事处联络处)

郭光 新闻

张宣 法律顾问

王行宇 协会建设和持续发展副会长

李爱荣 学术交流和智库建设副会长

胡文闯 学术交流和智库建设副会长

薄智泉 创新双创和技术转化副会长

余国营 对外交流和资金管理副会长

程晓红 对外交流和资金管理副会长

颜为民 教育文化和教育技术副会长

陈启秀 副秘书长

黄珺兰 秘书

杨晓梅 副会长(宣传)

陆强 《海外学人》主编

曾大军 技术总监

焦德泉 协会建设和持续发展副会长

董校铭 协会建设和持续发展副会长

肖继忠 学术交流和智库建设副会长

郭勇 创新双创和技术转化副会长

孙儒杰 创新双创和技术转化副会长

郑芳林 对外交流和资金管理副会长

施京 教育文化和教育技术副会长

于娜 教育文化和教育技术副会长

# 海外学人

Overseas Scholars

## OVERSEAS SCHOLARS

2017 年信息技术专刊

——暨庆祝旅美科技协会成立二十五周年特辑

ISSN: 1556-861X

本期责任编辑：肖继忠 田英利  
张立言  
发行人：曾大军  
刊名提词：中国科学院原院长  
路甬祥院士

### 《海外学人》编辑部

名誉主编：王飞跃  
主编：陆强  
执行主编：郭光  
顾问：胡晖  
编辑：盛晓明 蔡逸强 陈志雄  
潘星华 杨晓梅 胡文阁  
李爱荣 肖继忠 蒋为民  
美术/摄影：徐璐  
公共关系：蒋为民  
排版图文设计：汤淑芳

### 杂志主办

中国旅美科技协会  
Chinese Association for Science and  
Technology, USA (CAST-USA)  
www.cast-usa.org



## 中国旅美科技协会简介

中国旅美科技协会 (Chinese Association for Science and Technology, USA, 简称旅美科协, CAST-USA) 是 1992 年在纽约成立的非政治性、非营利性的民间团体。旅美科协的三大宗旨是：促进中美之间科学和技术、教育和文化、商业与贸易等领域的交流与合作；弘扬中国传统文化、促进中美两国人民的相互了解；加强旅美学人、华人专业人士之间的团结、合作与交流。

旅美科协是一个跨地区（在美国）、跨行业的综合性科技团体，会员主要来自科技、文化、教育、法律、金融、人文等领域的中国旅美专业人士组成，现有会员近万人。许多会员在世界五百强企业或知名大公司、高等院校或研究机构从事科技开发和研究工作，许多会员已经成为中、高层管理人员。旅美科协目前在全美有十六个分会及专业学会，会员分布在美国几十个州，并在中国国内十几个省市设立了办事处和联络处。

总会设执行委员会负责日常工作，还设有董事会、理事会、学术协调委员会和顾问委员会。旅美科协成立以来的知名名誉顾问包括陈省身教授、宋健教授、杨振宁教授、朱光亚教授、陈香梅女士、田长霖教授、周光召教授、朱丽兰教授、路甬祥教授、邓文中先生，及诺贝尔奖获得者达尼埃尔·谢赫特曼教授、马里奥·卡佩奇教授等学术及社会知名人士。

旅美科协总会及各分会定期举行学术研讨活动，为会员提供学术交流的平台，并定期组织其他系列活动，为会员们提供学习、工作及生活方面的交流平台。旅美科协总会定期出版《海外学人》杂志及实时通讯，内容包括协会学术活动的介绍与中美科技界、工商界的最新动态，以及各种工作与投资机会等许多会员们切身关心的内容。

每年总会及各分会举办包括全国年会及分会年会、学术讲座等在内的几十次大中型学术研讨活动，活动中旅美科协邀请中美各界知名人士对所关心的学术及社会问题进行探讨。旅美科协注重与其他专业协会的交流与合作，加强不同学科华人的交流，同时促进中美之间科技人才的沟通和科技的发展。旅美科协各分会也注重参加当地的华人社区活动，与所在地的其它侨团建立良好的关系。

旅美科协总会现任会长蒋为民，董事会主席陈志雄，理事会主席宋云明，候任会长潘星华。旅美科协历任会长为周华康、章球、徐震春、陆重庆、马启元、周孟初、谢家叶、肖水根、石宏、邹有所、林民跃、王飞跃、李百炼、左力、沈陆、陆强、曾大军、方形、盛晓明、蔡逸强、于浩、宋云明、陈志雄。

# Table of Content

1	《海外学人：信息技术专刊》 ——暨庆祝旅美科协成立二十五周年特辑 前言.....	肖继忠, 田英利, 张立言
4	Ant Financial Technology: Ultra large-scale financial-grade computing promotes the development of global inclusive finance 超大规模金融级计算促进普惠金融发展.....	胡 喜
12	Analysis of Request Scheduling and Optimization Methods in Cloud Data Center 云数据中心请求调度与优化方法分析.....	苑海涛, 周孟初
25	Personal Archive Service System using Blockchain Technology 区块链技术在个人存档服务系统的应用.....	陈志雄, 朱仪轩
38	Sustainable Computing and Computing for Sustainability 智能计算与可持续能源发展.....	孔繁鑫, 刘学
49	Connectionism and Symbolism in Artificial Intelligence 人工智能中的联结主义和符号主义.....	顾险峰
62	Where is Artificial Intelligence Heading? 人工智能, 何去何从.....	周少华
67	Human-Machine Perception and Assistive Technology 人机感知和辅助技术研究.....	朱志刚
82	A Wearable Robotic Object Manipulation Aid for the Visually Impaired 帮助视障者抓取物体的手戴式机器人辅助工具.....	叶 苍
91	A Robot-integrated Smart Home for Elderly Care 机器人和智能家居技术在智慧养老中的应用 .....	盛卫华, Ha Manh Do, Minh Pham, 杨丹, 刘妹琴, 张森林, 袁宪锋
103	Unambiguous and Accurate Text Extraction in Natural Scenes 自然场景下的文字精确检测定位方法.....	戎雪健, 田英利
113	Lower Extremity Exoskeletons and Active Orthoses: Challenges 下肢外骨骼机器人: 研究进展与关键问题分析.....	衣淳植, 苏浩
130	《海外学人》庆祝旅美科协成立二十五周年特辑	
131	风华正茂, 继往开来.....	潘星华, 蒋为民, 万乐, 盛晓明
135	回顾科协成立初, 团聚纽约庆发展.....	周华康
137	纽带、渠道、桥梁——记旅美科协成立廿五周年.....	陆重庆
141	新的时代 新的方向——庆祝旅美科协成立二十五周年.....	谢家叶
144	中国旅美科技协会介绍	
165	旅美科协驻中国办事处/联络处负责人和联系方式	

# 《海外学人：信息技术专刊》

## 暨庆祝旅美科协成立二十五周年特辑

### 前言

肖继忠，田英利，张立言

在中国旅美科技协会（CAST-USA）成立二十五周年这个值得庆祝的日子里，《海外学人：信息技术专刊暨庆祝旅美科协成立二十五周年特辑》终于出版了。中国旅美科技协会于 1992 年在纽约成立，一直致力于推动中美在科学技术、教育和文化等方面的交流与合作，起到纽带、渠道和桥梁的作用。本特辑邀请了周华康、陆重庆、谢家叶三位老会长回顾旅美科协二十五周年的发展历史，现任会长蒋为民和候任会长潘星华总结旅美科协近几年的发展现状。展望未来，旅美科协顺应全球合作与发展的潮流，一定不辜负时代的重托，为促进中美两国的科技进步、经济发展和文化交流贡献力量。

信息技术涵盖的领域非常广阔，影响到我们生活和工作的方方面面。随着大数据和云计算、物联网、嵌入式系统、机器智能、图像及生物特征识别、网络安全等新技术的突破，以及不断进步的超大规模计算能力和计算资源虚拟化技术，云计算的力量为全球的社会、商业和金融问题提供了新的数字解决方案。本专刊收集的 11 篇文章在云计算（cloud computing）、人工智能（Artificial Intelligence）、机器人（Robotics）、辅助技术（Assistive

Technology）和智能家居（Smart Home）等几个方面展现了现代信息技术研究、应用的现状和发展趋势。

云计算是信息技术（IT）领域继个人计算机和互联网之后的第三次技术革命。作为一种新的计算模式，云计算通过网络将云计算数据中心里的 IT 资源和数据实时按需地提供给用户使用。这些 IT 资源包括存储、计算和网络等信息基础设施资源，以及操作系统、开发平台和业务应用等软件资源。云计算使得商业模式、工作模式和社会生活模式发生着重大的转型升级。蚂蚁金服一直致力普惠金融的实现，让全球个人用户与小微企业享有平等的金融服务。蚂蚁金服首席技术架构师胡喜的文章总结了他在 2017 年 5 月份美国纽约的蚂蚁金服技术论坛（Ant Financial Tech Forum，以下简称 AFTF）上的演讲内容，给大家分享了整个蚂蚁金服信息技术基础设施发展的历程和现有的能力。新泽西理工学院的周孟初教授及其合作者的文章分析了国内外云数据中心的研究现状，介绍了云数据中心请求调度与优化方法，指出了云计算技术在应用领域（物联网、智慧医疗、车联网）存在的主要挑战。美国玛希学院陈志雄教授的文章介绍了使用区块链技术进

行个人档案管理的新方法，为建立分布式、透明、不可篡改并且安全的个人档案管理和服务体系提供了一个新途径。加拿大麦吉尔大学刘学教授从能源消耗这一独特角度探讨了通过提高能耗效率以及降低温室气体排放来改善计算机和信息技术系统的可持续性，他调研了可持续性数据中心（data center），总结了增加在数据中心中清洁能源的使用从而降低温室气体排放的现行方法，并指出了一些亟待解决的问题。

人工智能是当前最热门的话题之一。纽约州立大学石溪分校顾险峰教授的文章将人类脑神经认知和人工神经网络认知进行对比，对人工智能发展的主要方向，以及目前的局限和未来的潜力进行了探讨。西门子医疗技术中心的周少华博士试图揭秘人工智能究竟何去何从？他认为，只有将人工智能技术进行踏实的产品化，才是公司的正道正解，是将人工智能进行到底的不二途径。

随着全球老龄化速度不断加剧，更多的人将面临年龄相关的视力障碍、运动障碍。辅助技术和服务机器人应用于盲人辅助导航和老人护理受到了越来越多的关注。辅助技术，顾名思义，就是用各种技术来辅助那些需要帮助的人，更好地让机器服务于人类。纽约城市大学朱志刚教授介绍了他的视觉计算实验室近几年在人机感知和辅助技术的研究成果，概述五个方面的专题：（1）辅助视觉导航技术，（2）替代感知方法，（3）面部识别和深度学习，（4）测试平台开发，和（5）教育培训项目。朱志刚教授为每一个专题配一首小诗表达核心思想，增加了文章的趣味性和可

读性。

阿肯色大学小石城分校叶苍教授介绍了一款具有三维计算机视觉及人机合作功能的手戴式视障辅助工具。它可用于帮助视障者识别并抓握一些小的物体，从而解决辅助导盲中开门及其它日常生活中抓取物件的难题。俄克拉荷马州立大学盛卫华教授及其合作者介绍了服务机器人和智能家居技术在智慧养老中的应用，研究成果为解决独居老人的护理问题提供了一种新方法。纽约城市大学田英利教授的文章比较和评判了自然场景下的文字检测与定位的多种方法，提出了一种结合最新深度学习技术与自然语言描述的无歧义文字精确检测定位方法，对理解场景内容、盲人辅助导航以及自动驾驶等应用都具有重要意义。纽约城市大学苏浩博士的文章回顾了下肢外骨骼及主动矫正器的发展历程以及取得的主要进展和仍需解决的困难与挑战。外骨骼是一种拟人性质的活动机械装置，由操作者穿戴并与其身体紧密配合、协调一致地运动，可以帮助残疾人和老年人行走或防止摔倒。

总之，信息技术日新月异，应用前景无限广阔，希望本专刊能抛砖引玉，为促进科学知识的普及与科学技术的发展贡献力量。

世界潮流浩浩荡荡，合作发展是主旋律，旅美科协定能不忘初心，为促进中美两国的科技文化交流贡献力量。

## 作者介绍



**肖继忠博士**, 现任纽约城市大学城市学院电机工程系终身教授, 机器人学实验室主任。纽约城市大学研究生院工程专业和计算机科学专业的博士生导师、感知机器人和智能仪器研究中心主任。肖继忠教授先后荣获

美国国家科学基金会“杰出青年事业奖”、德国洪堡基金资深学者科研资助奖、纽约城市大学城市学院“最佳指导教师奖”。在国内任中国科学院沈阳自动化研究所机器人学国家重点实验室特聘研究员, 博士生导师, 中科院海外评审专家。在 IEEE 汇刊等国际核心期刊和主要学术会议上发表论文一百六十余篇, 多次获得最佳论文/最佳录像奖。肖继忠教授历任中国旅美科技协会大纽约分会会长, 理事长; 现任旅美科协总会负责学术交流的副会长。



**田英利**, 纽约城市大学研究生院计算机系与城市学院电子工程系终身教授。创办并主持多媒体实验室。加盟纽约城市大学前, 曾任 IBM 沃森研究中心 (IBM T.J. Watson Research Center) 研究员和美国哥伦比亚

大学 (Columbia University) 客座教授。主要研究领域为计算机视觉和人工智能。



**张立言博士**, 现任科珍通讯公司 (全球前三大网络, 总部位于美国华盛顿特区的) 网络技术部主任, 并兼任中国旅美科技协会华盛顿分会、计算机网络协会副会长。他拥有近二十年通讯和因特网

工作经验, 九十年代于美国新泽西州郎讯贝尔实验室任高级研究员, 专攻光纤通讯和系统、网络。于最近十二年在科珍通讯公司规划、建设并营运世界最大网络之一。张博士于 1979 年考入中国科技大学并毕业留校任教, 是一个全方位高科技和网络技术专家。

# Ant Financial Technology: Ultra large-scale financial-grade computing promotes the development of global inclusive finance

## 超大规模金融级计算促进普惠金融发展

**Hu Xi**

胡喜

Chief Technical Architect of Ant Financial

蚂蚁金服首席技术架构师

### Abstract

Ant Financial is determined to provide innovative technology-driven financial services from the beginning. With the emerging of cloud computing, big data, artificial intelligence and other new technologies, Ant Financial has continuously empowered our global partners with its own technology and infrastructure capabilities, especially its ultra-large-scale financial-grade computing technology. Through empowerment and cooperation, Ant Financial and its global partners are working together to promote equal access to financial services for the general public. Access to financial services is crucial to strengthen financial sectors and domestic resource mobilization and can therefore make a significant contribution to social and economic development.

Ant Financial VP and Alibaba Partner, Mr. Hu Xi, shared technology infrastructure development process and the current status of

technology infrastructure at the Ant Financial Tech Forum in New York in May. This article summarizes Hu's presentation at AFTF.

### 摘要

蚂蚁金服长期以来一直坚持技术创新驱动金融服务的初心，随着云计算、大数据、人工智能等新技术的不断兴起，蚂蚁金服已经通过开放合作不断把多年积累沉淀的金融级大计算技术输出给合作伙伴，赋能全球金融服务的基础建设，在全球范围内与合作伙伴一起实现普通大众均等享受金融服务的机会。在五月份美国纽约的蚂蚁金服技术论坛（Ant Financial Tech Forum，以下简称 AFTF）上，蚂蚁金服首席技术架构师胡喜给大家分享了整个蚂蚁技术基础设施发展的历程和现有的能力。本文介绍了他在 AFTF 的演讲内容。

**Keywords:** Inclusive Finance, Ultra - large - scale computing, technical architecture, risk management, financial decision making, statistical analysis

**关键词：**普惠金融，超大规模计算，技术架构，风险控制，金融决策，统计分析

## 1. 引言

当面支付或者付款是日常生活中完成一笔交易最典型的场景。当用户拿出手机，打开二维码进行付款操作时，后台会经过一系列技术处理，这中间最主要的环节是风险识别。风险识别以后，是相应的营销决策，判断用户在该笔交易中应该获得多少红包或者奖励，最终完成交易，发生账户扣款并产生结果。这是今天蚂蚁金服内部最简单的一个业务场景。这样一个简单的场景背后，对于技术最大的挑战是什么？以双十一为例，每年的双十一，对于蚂蚁技术人员都是一场大考，蚂蚁金融级大计算能力在每一次大考中不断被打造而且接受了实际的大规模验证。双十一从 2010 年最开始只有每秒 500 笔的处理能力，而到 2016 年系统处理能力已提升到 12 万笔/秒，当天更是创造了 10 亿的交易笔数和 1207

亿的交易额。7 年左右的时间，蚂蚁金服交易的峰值能力、系统处理能力提升了 240 倍，未来还将向百万级/秒的交易能力快速发展。那么这么大规模的系统，而且还要能实现低成本高效率的处理，蚂蚁金服是如何解决这些核心金融级计算挑战的？最核心的三大挑战是：1. 超大规模金融交易处理，如何在小于 500ms 内完成百万级/秒的交易；2. 超大规模金融决策处理，如何在小于 100ms 内完成百万级/秒的决策；3. 超大规模金融统计分析，如何在小于 10s 内完成对 EB 级以上的数据做分析处理。

## 2. 技术挑战一：超大规模金融交易处理

蚂蚁金服的交易技术架构经历了四代架构，当前正在进行第五代架构的研发。第一代架构从 2005 年到 2007 年，以 IOE 商业软件和开源软件为主，是一种烟囱式架构。第一代架构整体比较简单，业务都在一个系统中完成，在这个阶段的交易量的处理能力，完成了从每天万笔提高到百万笔的



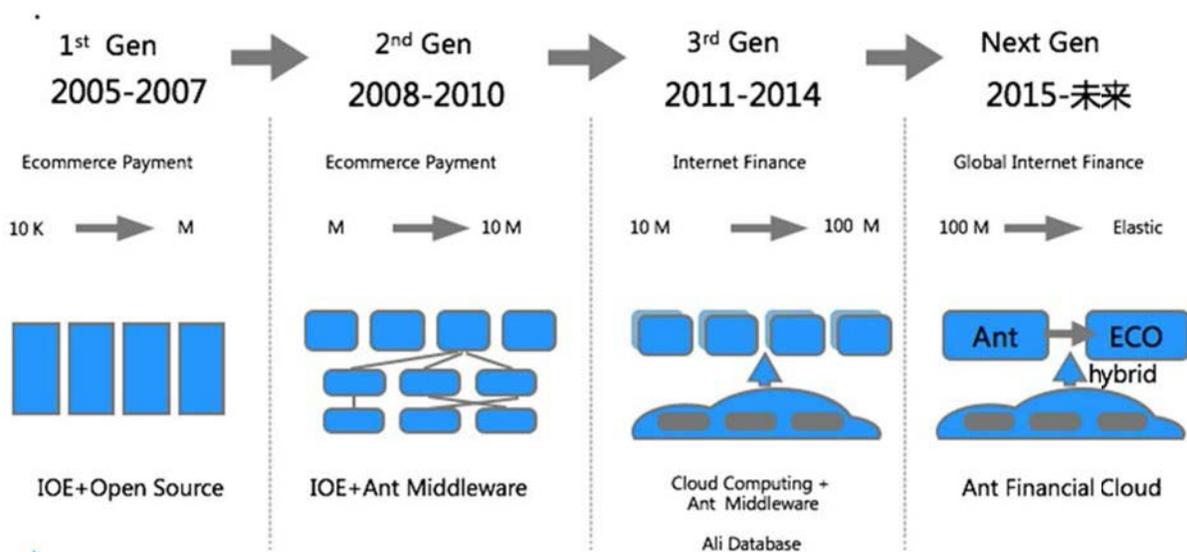


跨越。第二代架构是 2008 年到 2010 年面向服务化的 SOA 架构，所有系统经过重构并抽象出关键系统。当时的开源软件已经无法满足业务复杂性的要求，因此这代架构中很多系统是自研的，包括应用容器、消息服务等中间件组件都是在这个阶段研发出来的，第二代架构让伸缩到千万笔的交易规模成为可能。第三代架构，是在原先的 IOE 等商业软件越来越不能满足大规模、低成本、高可靠要求的情况下诞生的，必须自研新的技术平台。以 X86 服务器为基础，采用云计算的方式去支持未来更大规模、更为弹性的低成本云支付架构。这个阶段蚂蚁做了重大重构，去掉原先 IOE 相关的技术，全部采用蚂蚁自主产权的技术，包括蚂蚁的云计算技术以及自研的事务型数据库 OceanBase（以下简称 OB）等等，这次重构可以支持到亿级的交易伸缩。第四代架构的设计目标，不仅要满足蚂蚁自己的业务，还要提供给更多的外部客户使用，包括所有蚂蚁生态中的合作伙伴。这是一种混合云架构，称为“金融云”架构，为亿级之上更大规模更低成本的交易能力

铺平了道路，蚂蚁相关的网商银行、国际化业务等等，都已经实现了第四代架构的最新能力。

金融交易系统，有别于其他大规模互联网系统，对事务型的状态数据一致性处理以及交易成本的要求更高，对客户资金风险和安全性处理更加复杂。从蚂蚁四代架构的发展历程中可以看到要解决大规模交易处理，必须解决以下四个核心问题：1、如何规模化分布式服务和数据能力，也就是服务和数据处理的可伸缩性 2、当服务和数据分布后，如何保证事务中数据的强一致性，这是金融级分布式系统最大挑战之一；3、如何做到秒级快速容灾能力，做到四个 9 及以上的系统可靠性 4、实现低成本和按需提供资源配置与调度，如何做到单笔少于 1 分钱人民币的交易成本。

蚂蚁利用面向服务化的伸缩架构、机房级服务伸缩架构、多维度分布式数据伸缩、单元化的伸缩能力等手段来解决分布式服务与数据带来的挑战；利用基于 TCC 分布式事务架构、可靠事务消息服务等手



蚂蚁整体金融级大规模计算技术发展历程的一个简单概述

段来解决分布式事务与协调带来的挑战；利用 OB 多地多中心部署模式、异地多活单元化架构、机房间流量调拨等手段来保障强一致容灾以及城市级容灾；用按单元弹性扩容、按业务单元进行弹性扩容等手段来解决弹性供给与调度。

通过这些核心技术手段的实现，蚂蚁具备了超大规模的交易处理能力，而且是基于低廉的硬件去完成的。同时通过这些技术的开放，让更多的金融机构可以使用这些能力。蚂蚁相关的第一家互联网银行“网商银行”以及许多其他蚂蚁全球生态合作的金融机构，都采用了这些超大规模金融级交易技术。其中网商银行是全球第一家采用云计算技术来实现核心系统的银行。

### 3. 技术挑战二：超大规模金融决策处理

金融系统的本质是对风险的控制以及对风险的洞察与决策。回顾蚂蚁金服这十

几年的业务发展，从 2004 年开始做担保交易、到 2010 年以后开始做微贷，2013 年做余额宝，2015 年做花呗以及芝麻信用，逐渐从一个支付系统演变成全方位的超大规模的金融系统。其中对风险控制的要求，从支付系统考虑的反盗用、反欺诈、反洗钱等基本要求，到微贷考虑的实时准入、实时授信、风险甄别，到余额宝考虑的流动性风险，到花呗需要考虑的实时提额、反套现，再到芝麻考虑的实时信用评级等等，对于这些实时风险控制、洞察、与决策，都是非常大的挑战。

蚂蚁金服一直致力普惠金融的实现，让全球个人用户与小微企业享有平等的金融服务。对于金融决策结合大数据场景的例子最好的体现就是小微贷款，它以大数据为核心，基于信用，实现无抵押、无担保的高效贷款。

目前蚂蚁微贷已经可以做到 3 分钟申请、1 秒钟放贷、0 人工干预。其中的核心



是做到信用风险防控。通过客户行为、消费习惯、店铺经营状况等数据形成用户画像并构建网络模型。通过网络模型判断行业趋势，结合店铺成长轨迹给予商户放贷支持，更好地支持经营者。在信贷决策过程中，从准入、授信到欺诈再到虚拟交易、贷后监控、贷后催收等，所有环节都与大数据息息相关。在贷款的每个环节中实时发现异常，并快速解决问题，更好地服务个人用户与商户。

除了金融级风险控制以外，另一个巨大挑战是在服务蚂蚁几亿到未来几十亿消费者的过程中，如何能够真正做到精准的用户洞察以及提供个性化的金融服务。比如在实际的营销决策中，做业务和守住风险的同时，还需要知道通过什么样的奖励可以促进用户更好的消费。例如新春红包，双十二的集宝活动，还有 AR 红包，以及口碑的生活场景，在做好反作弊的风险控制同时，还需要考虑奖励或红包怎么发、给哪些消费者群体发、发多少，有可能还需要考虑用户的当前位置、以及附近不同的商业街区，等不同的营销策略。

另一个实时风险决策的例子，通过监控用户线上线下行为事件识别账户风险。例如登录与交易，通过实时数据分析、数据挖掘等技术进行风险嫌疑账户的预测，

发现异常的或有风险的操作行为，根据风险级别不同，在 100ms 内进行不同的实时风险控制处理。蚂蚁的智能实时风险监控系统是目前国内最先进的网上支付风险实时监控系統。

总体来说，金融系统最核心的挑战是能够控制风险，洞察风险，识别风险以及业务如何围绕风险进行快速决策。这对这种大规模的金融决策系统提出了四大要求：1、金融决策如何满足时效性，对在线业务系统最大的挑战是交易快速发生，因此必须在交易发生的同时或者之后快速识别风险并做出正确决策，但当交易量如此庞大，如何在背后进行快速的决策成为非常大的挑战。2、金融决策的精确性，如何确保策略、模型、特征、规则的匹配是一个非常精确的过程 3、策略的自学习，一个静态的策略永远不能防范未来的许多风险，必须让策略更为动态甚至做到自我学习，这也是蚂蚁金服一直在研发的核心方向。4、低决策成本，如何在做大规模金融决策计算的过程中，保持低成本并且实现高效率处理。这相比与低成本交易是同等甚至更复杂的技术挑战之一。

为了应对风险决策、信用决策、营销决策以及各种金融服务中实时决策相应的挑战，必须首先做好在线特征的实时计算

与学习，模型算法全生命周期的服务与管理。蚂蚁金服建立了自己的实时特征计算平台，用基于 Velocity 的特征实时计算、在线离线统一的特征计算体系、关系特征的计算、关系特征计算引擎等手段解决特征计算的时效性；同时建立了全生命周期的模型服务平台 PAI 来训练模型，兼容各种开源机器学习框架的算法学习平台 PS，通过多种最先进算法和计算手段最终达到金融级策略大规模的自学习和自适应的能力，并且提供弹性资源供给与调度的能力。

2017 年初，蚂蚁联合 10 家世界顶尖科技公司如谷歌、亚马逊、微软等与美国加州伯克利分校的 RISE 实验室形成基础研究合作，共同探索并解决大规模实时智能决策系统的世界性难题，RISE 实验室的前身是诞生 APACHE SPARK 著名大数据系统的 AMP 实验室。Michael I. Jordan 世界级机器学习教父级人物，现任蚂蚁金服科学智囊团主席，是 RISE 实验室的负责人之一。蚂

蚁希望打通世界级的基础研究的能力，未来把超大规模金融级决策能力，输出给所有蚂蚁的合作伙伴。

#### 4. 技术挑战三：超大规模金融统计分析

对于金融级系统来说除了大规模的交易处理和决策的挑战外，从超级海量（每天 EB 级以上规模）的数据中进行小于 10s 实时洞察和分析也是一个非常巨大的挑战。典型的场景之一是在知道各种金融业务运营的实时报表的同时对一些相关金融业务的宏观经济行为做出统计分析。要对海量数据进行洞察和分析，面临的挑战主要是要解决以下五大问题：

- 1、数据准备的效率如何提升？
- 2、如何使得数据产品，更实时、更快、更直观的展现
- 3、如何提供对数据的自动诊断&预测能力？



4、如何在诊断之后提供更为有效的决策建议？

5、如何快速地做出行动的指导，并对行动结果进行反馈？

这些是蚂蚁研发统计分析平台核心考虑的核心问题。但真正在做的时候面临的挑战更多，尤其在蚂蚁业务数据化的过程中，涉及的业务场景非常多也非常复杂，比如在交易场景中不仅有针对买家、还有针对卖家商户的一系列数据分析需求。

除了交易场景外、还有很多其他业务场景，比如保险、财富、信用，微贷等，面对不断拓展的业务场景，催生各种围绕场景的集合、进行数据产品流程抽象，最终形成并沉淀成统一的数据产品集合，并且反过来驱动和适应蚂蚁所有数据业务化的快速变化，这就是阿里巴巴一直在说的DT驱动业务。

为解决这些问题，蚂蚁需要做到以下几点：



1、数据的超大规模计算同时保证时效性，在小于 10s 内完成对 EB 级以上的数据做分析处理。需要能够快速计算，并把计算能力扩充到很大，实现超大规模的处理能力，以完成对数据、业务的洞察。

2、数据的诊断与洞见。如何利用数据进行洞察，在蚂蚁有财富、理财、信用、微贷等各种场景，对不同的业务场景，如何知道相关业务出了什么问题。比如余额宝，有流动性风险，如何判断它有流动性风险，如何通过数据的方式洞察到它有问题。

3、数据可视化，只有可视化才能更好地辅助运营人员确定相应数据那里出了什么问题。

4、同样的挑战如何降低成本。

## 5. 总结

总体来讲，蚂蚁金服正在逐步解决金融级基础设施中的三大技术挑战：如何处

**1** 数据准备效率如何提升？

**2** 如何更多功能更实时，更快，更直观

**3** 如何提供自动诊断&预测等功能？

**4** 如何提供有效决策建议？

**5** 如何指导行动，并对行动结果进行反馈。

理超大规模的交易；如何做超大规模的金融决策；如何处理超大规模金融统计分析。在处理大规模的交易过程中，需要解决面向金融的决策，以及洞察金融分析内部发生的问题，这些都是蚂蚁金服一直要不断解决的核心问题：1、继续推进低成本安全可靠的交易系统到百亿级伸缩规模；2、增强学习和迁移学习等新计算能力在金融实时决策系统中的应用；3、异构计算能力推进多维度计算能力的优化；4、基于一致语义的实时统计分析能力。

蚂蚁金服期待这些挑战被不断解决和优化，也期待赋能全球金融服务新基础设施能力的建设,期待与合作伙伴一起在全球范围内为普通大众均等享受金融服务机会的普惠金融愿景早日实现。

## 作者简介

**Hu Xi** – Chief Technical Architect, Mr. Hu is the head of the Infrastructure Technology Department and the Chief Technical Architect of Ant Financial. He joined Alipay in 2007 and became a main designer and developer for every generation of core payment platform infrastructure. As Ant Financial's Chief Technical Architect, Mr. Hu led the efforts to build a platform to provide financial institutions with cloud computing services. Currently, he is focused on the design and development of the next generation of TechFin infrastructure for Ant Financial.

**胡喜**，现任蚂蚁金服首席技术架构师，基础技术部负责人。2007年加入支付宝，主持支付平台基础技术的架构设计与研发工作，并且参与蚂蚁金服各代核心支付平台的架构设计和系统研发工作。作为蚂蚁金服技术框架的主要设计者和参与者，蚂蚁金服对外提供云计算服务平台的奠基人，目前专注于蚂蚁金服下一代TechFin基础设施的设计和建设。

# Analysis of Request Scheduling and Optimization Methods in Cloud Data Center

## 云数据中心请求调度与优化方法分析

Haitao Yuan<sup>1</sup>, MengChu Zhou<sup>2\*</sup>

苑海涛<sup>1</sup>, 周孟初<sup>2\*</sup>

1. School of Software Engineering, Beijing Jiaotong University, Beijing 100044, China;
2. Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, NJ 07102, USA

1. 北京交通大学, 北京市 100044, 中国;
2. 新泽西理工学院, 纽瓦克市, 新泽西州 07102, 美国

\*联系人: E-mail: zhou@njit.edu

### Abstract:

Multiple types of applications run in cloud data centers and provide services to different users through resource sharing. The construction of cloud data centers has become an important part of national information infrastructure. However, as cloud computing and its applications are widely adopted, their infrastructure, architecture and management face many new challenges. This work analyzes their recent studies, and points out several key challenges related to the application of cloud computing technologies in other areas including big data processing, Internet of Things, smart healthcare, and Internet of Vehicles. In addition, this article indicates several research directions in order to realize green, smart, and fast-responsive cloud data centers. It is hoped that this article would offer a certain theoretical significance and practical value to meet ever-increasing needs to develop

sustainable economy, society, science and technology.

### 摘要

云数据中心中运行着多种类型的应用, 并通过资源共享的方式向不同用户提供服务。建设云数据中心已成为国家信息基础设施建设的重要组成部分。然而, 随着云计算技术及应用的逐步发展, 云数据中心的基础设施、架构与管理均面临着新的挑战、变化与发展。本文分析了目前国内外关于云数据中心的研究, 并指出了云计算技术在其他领域(如大数据处理、物联网、智慧医疗、车联网)的应用仍存在的主要挑战, 希望对促进绿色、智能、快速响应的云数据中心发展、满足国家经济、社会、科技等方面发展需求具有一定的理论意义和实用价值。

**Keywords:** Cloud Computing, Data Center, Request Scheduling, Resource

**关键词：**云计算，数据中心，请求调度，资源分配，绿色计算

### 1. 建设云数据中心的必要性

云计算是信息技术（IT）领域继个人计算机（PC）和互联网之后的第三次技术革新。云计算技术是由多种技术包括分布计算、并行计算、网格计算和效用计算等融合和发展而来。作为一种新的计算模式，云计算通过网络将云计算数据中心（简称云数据中心）中的 IT 资源和数据像用电、煤气和水一样随时随地按需地提供给用户使用<sup>[1]</sup>。这些 IT 资源包括存储、计算和网络等信息基础设施资源，以及操作系统、开发平台和业务应用等软件资源。云计算因其所具有的服务化、网络化、个性化、柔性化和社会化等特性，已在国民经济、

科学研究、国家安全、国计民生和文化等领域得到广泛的应用，使得商业模式、工作模式和社会生活模式发生着重大的转型升级。随着云计算技术的不断发展，目前云数据中心已经成为一个集大数据量运算和存储为一体的高性能计算机的集中地。它由云提供商所有，并向政府、企业或公众提供信息服务。因此，各个行业的多个不同应用系统均需要云数据中心提供资源并进行各种请求的处理。

如图 1 所示，云数据中心是云计算实施与应用的基本载体，它的职能是提供基础设施资源及其高效、低成本、高质量的管理与应用。该架构包括基础设施层、平台核心软件层、资源管理层、应用接口层等，并支持多种不同类型的业务应用，如大数据处理分析<sup>[2]</sup>、移动应用<sup>[3]</sup>和社交网络<sup>[4]</sup>等热门应用。目前云数据中心中同时运行

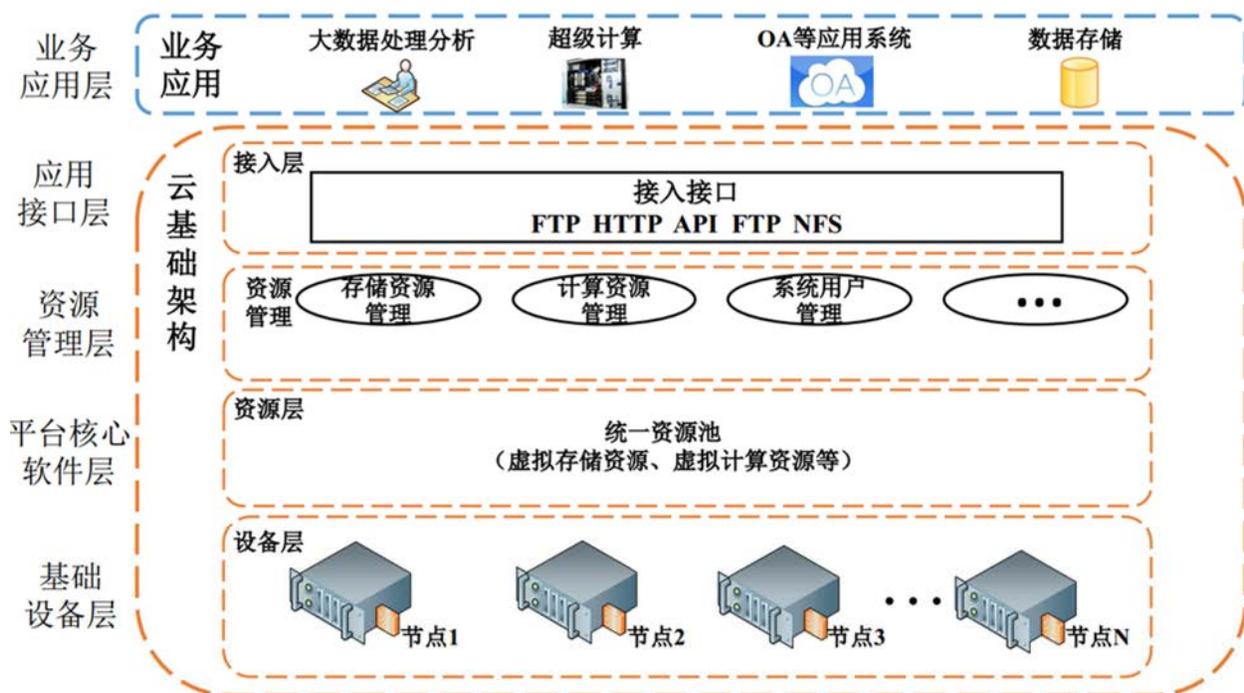


图 1 云数据中心架构图

着多种不同类型的延迟容忍型应用（如大规模仿真、高性能计算、大数据分析处理和复杂工程分析等），并通过资源共享的方式向世界各地的用户提供服务。目前云数据中心广泛采用虚拟化技术来提高服务器和网络资源的使用效率。每个物理服务器往往被虚拟化为多个虚拟机，为保证系统的稳定性和可扩展性，每个应用往往同时部署在多个同构或者异构的虚拟机中。而云数据中心管理的本质目的是在满足各类应用请求延迟要求的前提下，提高云计算的成本效益并实现其规模经济特性。随着云计算技术及应用的逐步发展，云数据中心的基础设施、架构与管理均面临着新的挑战、变化与发展。

## 2. 云数据中心国内外研究现状分析

本文从基于软件定义网络（Software-

Defined Networking, SDN）控制器的流量工程、混合云中资源优化和绿色云中请求调度等三方面对国内外的当前相关研究进行分析。

### (1) 基于 SDN 控制器的流量工程

云数据中心的流量工程和负载均衡技术是提高云数据中心系统性能和改善服务质量的关键技术。流量工程的主要目标是智能地将请求在数据中心网络中进行调度，从而对某些特定的网络性能指标进行优化。流量工程通常包括从源节点到目的节点间路径计算、多个路径间请求划分策略计算等。如图 2 所示，SDN 由网络基础设施层、控制层和业务应用层组成。SDN 控制层通过基于 OpenFlow 协议的南向接口控制网络转发设备的行为，并提供北向接口（Application Programming Interface, API）

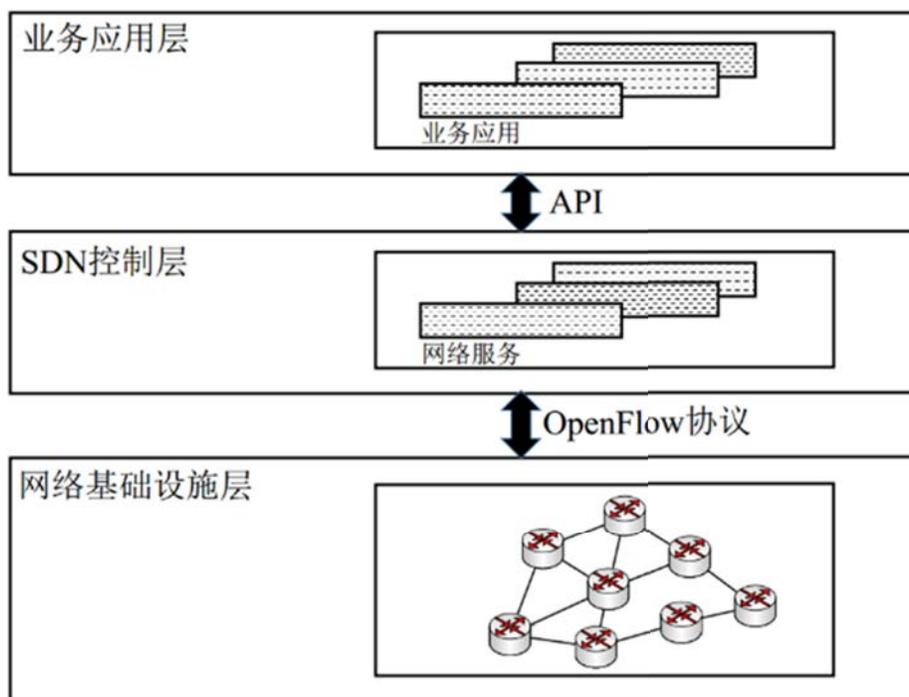


图 2 软件定义网络(SDN)的架构

以实现不同的网络业务应用。而 SDN 技术中的网络传输设备支持更加细粒度的管理和控制，能够实现更高效的请求划分策略。此外，SDN 具有逻辑上集中的控制平面，能够实时地获取网络状态和完成请求路由路径的计算，从而实现全局最优的路径规划。

目前一些研究工作使用 SDN 技术来解决数据中心网络中的流量工程和负载均衡问题<sup>[5]-[9]</sup>。为解决网络视图的不一致对网络性能的制约问题，Rothenberg 等<sup>[5]</sup>提出了一种基于请求到达的负载均衡机制，从而最小化网络的最大链路利用率。然而，对于每个应用的请求而言，最大链路利用率并不能保证最小的往返延迟时间。Agarwal 等<sup>[6]</sup>研究了集中式的 SDN 控制器在混合网络中的应用问题。该混合网络由传统的路由

器和支持 OpenFlow 协议的 SDN 交换机混合构成。研究表明即使网络中仅仅包含部分支持 OpenFlow 协议的 SDN 交换机，整个网络的性能也能得到显著的提高。然而该方法仅适用于由混合的网络转发设备（路由器和交换机等）构成的网络。Tso 等<sup>[7]</sup>提出了一种基于测量的请求流调度方法，能够根据网络利用率为每条请求流指定路由路径，从而有效利用数据中心空闲的网络能力来改善过载链路的性能。Meng 等<sup>[8]</sup>根据虚拟机之间请求流量的分析，采用流量可感知的虚拟机放置方法从而提高网络的可扩展性。然而该方法仅考虑了网络链路利用率和拥塞情况。Nahir 等<sup>[9]</sup>对请求到达规律和特点进行了分析，并提出了一种降低每条路径上请求调度开销的方法。但是该方法未考虑应用请求类型的差异。实际上高优先级的应用请求应有更高的优先

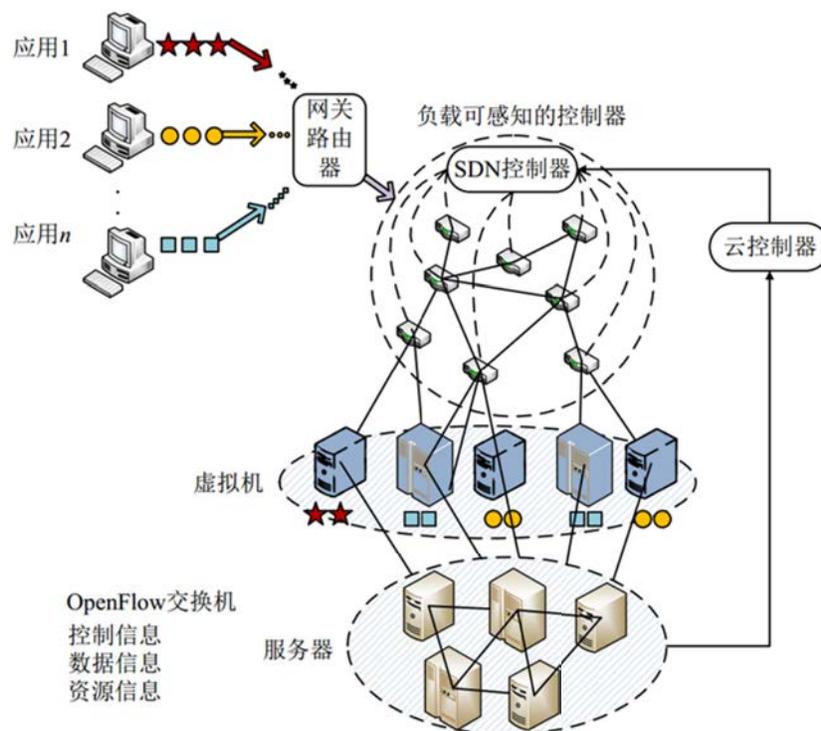


图3 负载可感知的控制器架构

权来通过整个数据中心网络。

上述研究工作仅仅考虑了数据中心网络中的流量工程问题,因此仅能对请求的网络延迟进行优化。文献<sup>[10]</sup>设计了一种负载可感知的控制器架构。如图 3 表示,该架构由 SDN 控制器和云控制器组成,分别采集请求在网络和虚拟机中的信息。如图所示,请求的总延迟由请求在网络中和虚拟机中两部分延迟组成,因此每个应用请求在虚拟机中的延迟对用户体验同样具有非常重要的影响。如果虚拟机延迟过大,将会显著增加请求的往返延迟时间,进而降低请求为云提供商带来的收益。因此目前需要设计一种负载可感知的控制器,综合考虑网络信息和虚拟机中请求队列信息,从而有效降低每个应用请求的总延迟并进而提高云提供商的收益。

## (2) 混合云中资源优化

云计算所具有的规模经济特性使得越来越多的延迟容忍型应用被部署在云数据中心中<sup>[11][12]</sup>。考虑到安全或者管理权限等因素,这些应用往往只由私有云提供商来管理和控制。然而私有云提供商的资源是有限的,应用请求的到达无特定规律、具有很大波动性。因此当私有云无法满足所有应用请求的延迟要求时需要按需动态地将部分请求调度到公有云中执行。混合云机制是当前云数据中心研究中的重点,它使得私有云能够满足所有应用请求的延迟要求。目前一些研究工作针对混合云中资源优化问题进行研究<sup>[13-16]</sup>。Rahman 等<sup>[13]</sup>提出了一种适应性的混合启发式算法来调度混合云中的数据流,从而在满足用户的成本、数据放置等要求的同时,最

小化工作流的执行成本。文献<sup>[14]</sup>提出了一种混合云环境下优化成本的调度算法,从而根据当前资源需求和资源成本确定如何从公有云获取最佳的资源,以达到在给定的时间要求内提供充足的计算能力来执行工作流的目的。Bittencourt 等<sup>[15]</sup>提出了一种适用于混合云环境的服务工作流调度策略,能够在满足工作流的延迟要求的同时最小化执行成本。该策略能够确定哪些服务工作流应该使用付费的资源,以及云提供商应该向公有云请求哪种类型的资源。Mezmaiz 等<sup>[16]</sup>提出了一种并行化的双目标混合遗传算法来解决云基础设施中有优先级约束的并行应用调度问题。通过分析应用请求的能量消耗和最大完成时间等特点,该文提出了基于动态电压调节技术的最小化能耗方法。然而,以上工作仅仅针对混合云中单一应用的请求,而忽视了云数据中心中不同应用请求的延迟要求的差异,以及多应用请求间的资源竞争。

云数据中心的建模和分析的精确性对于云计算的发展有着至关重要的作用。目前的研究工作多采用排队论的方法对云数据中心的性能进行建模分析。Nan 等<sup>[17]</sup>分别针对多媒体云中单类型服务和多类型服务的资源分配问题,基于排队模型提出了相应的资源分配优化方法。Shi 等<sup>[18]</sup>采用线性预测方法并结合 M/M/1 排队模型对云数据中心的资源利用率信息进行了分析和预测,并有针对性地对资源进行动态地分配,从而有效降低请求的响应时间和云数据中心的能量消耗。Zhang 等<sup>[19]</sup>针对单一的云提供商,采用模型预测控制方法求解受约束的离散时间优化控制问题,解决了如何满足用户需求并最大化云

提供商收益和最小化其能量成本的问题。Khazaei 等<sup>[20]</sup>提出了近似的分析模型来评价云服务器集群的性能，求解该模型可得到请求响应时间及其它重要性能指标的概率分布的估计，使得云管理者能够以此来确定所需的服务器个数。以上工作仅能够保证每个应用所有请求的平均延迟要求。然而，实际的云数据中心中每个应用的请求延迟时间具有长尾分布特征<sup>[21]</sup>，因此以上方法无法满足每个应用所有请求的延迟要求。

文献<sup>[22]</sup>提出了一种混合云中时间可感知的请求调度架构。如图 4 表示，混合云环境下私有云提供商的能量价格、私有云和公有云的虚拟机执行价格在请求延迟要求内均随着时间而变化。此外，不同应用请求的延迟要求存在差异，多应用请求间

存在资源竞争。因此，混合云中的请求调度需要综合考虑这些因素，从而能够将所有应用的请求在其对应的延迟要求内调度到私有云或者公有云中执行，并有效提高云提供商的利润和系统吞吐量。

### (3) 绿色云中请求调度

绿色云中请求调度是绿色云数据中心基础设施资源优化中的一个关键问题，其主要目标是智能地调度请求从而有效利用绿色可再生能源（如风能和太阳能等），从而在满足请求延迟要求的同时有效降低化石燃料使用对环境的污染。目前已有越来越多的研究工作关注可再生能源在大规模云数据中心的应用问题。Ghamkhari 等<sup>[23]</sup>考虑了请求的随机动态性和可再生能源的可用性，并提出了一种基于凸优化的请求调度算法来优化绿色云数据中心资源的分

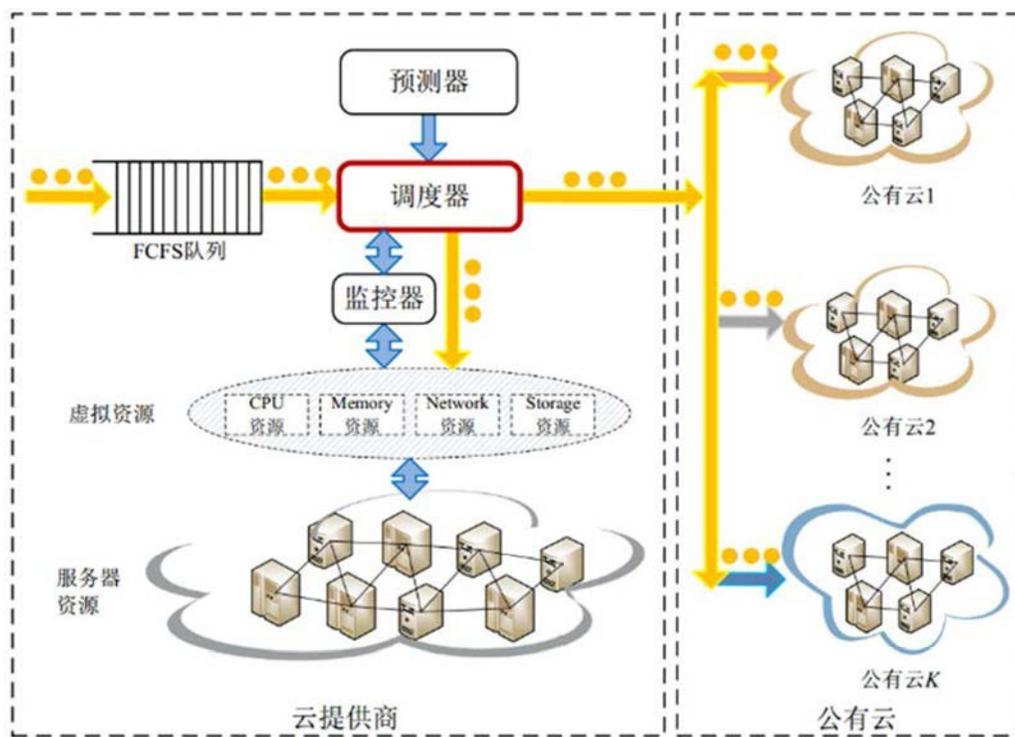


图 4 混合云中时间可感知的请求调度架构

配。Deng 等<sup>[24]</sup>设计了一种基于李雅普诺夫优化理论的在线算法，该算法针对高度动态变化的请求，能够在分布异地多个云之间实现联合的请求调度和电力管理，从而在保证请求延迟要求的同时最小化绿色云提供商的电力成本。Amokrane 等<sup>[25]</sup>提出了适用于绿色云数据中心的资源管理框架，该框架通过在多个分布异地数据中心之间分配资源，能够降低运维成本并使碳排放量不超过规定的阈值。Guo 等<sup>[26]</sup>利用可再生能源的间歇性和不可预测性，并采用热存储和负载均衡机制来改善绿色能源在云数据中心中的应用。Leslie 等<sup>[27]</sup>利用分布异地多个数据中心的电力价格和可再生能源在不同地理位置的差异，能够智能地将请求调度到电力价格较低或者可再生能源可用量较大的云数据中心的中心，以达到降低云提供商能量成本的目的。Nguyen 等<sup>[28]</sup>针对虚拟化的数据中心，考虑了环境感知的虚拟切片分配问题，结合可再生能源提出了最优的解决方案来提高数据中心的能量效率。

中心中能量优化问题，然而云计算的规模经济效应使得绿色云提供商重点关注的是应用请求执行所带来的利润问题。另外，由于实际云数据中心中请求延迟的长尾效应，现有工作仅能保证应用请求的平均延迟要求而无法严格保证每个请求的延迟要求。此外，绿色云数据中心的能量有限而请求的到达却无特定规律。为防止数据中心过载，现有工作多采用访问控制技术在绿色云数据中心能量不足时有选择地拒绝部分请求。然而，现有工作未能提供绿色云数据中心中服务器的服务率与拒绝的应用请求个数之间明确的数学关系，因此无法根据此提供精确的请求调度和资源分配方案。

文献<sup>[29]</sup>提出了一种混合云中时间可感知的请求调度架构。如图 5 表示，绿色云提供商的虚拟机使用价格、风速和太阳辐射强度在请求延迟要求内随着时间而变化。因此，绿色云数据中心需要综合考虑这些因素，从而能够在严格满足每个应用请求的延迟要求的同时，有效提高绿色云提供

上述研究工作主要集中于绿色云数据

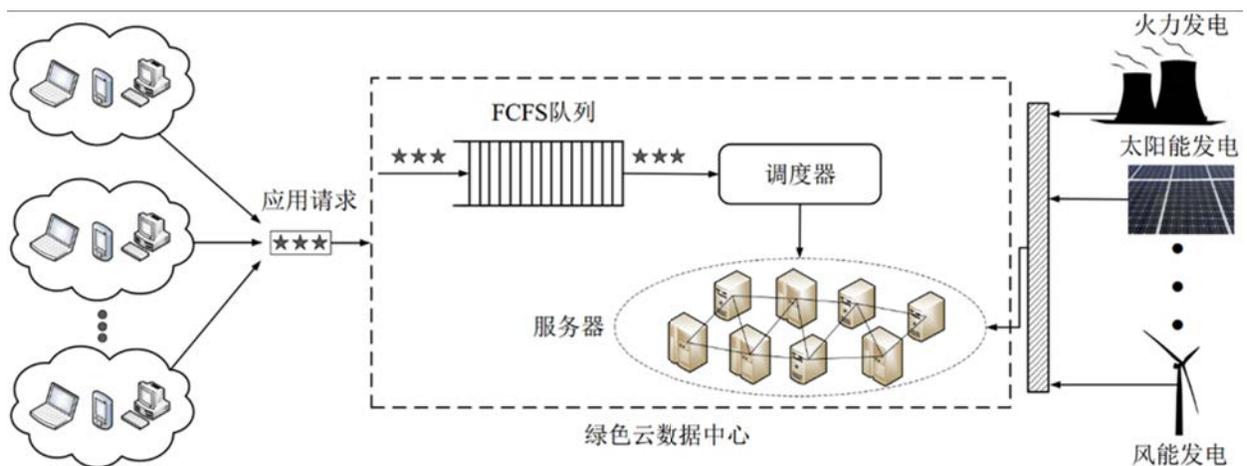


图 5 绿色云数据中心架构

商的利润和系统吞吐量，并降低环境污染。

### 3. 未来展望

云计算已经在中国、美国等世界很多国家得到了成功的部署和实施，比如中国航天科工集团的工业云制造平台、阿里云、中国联通沃云、百度开放云、青云等，以及美国的亚马逊弹性计算云（Amazon EC2）、微软的云平台 Azure 等。目前，越来越多的政府机构和企业将其应用部署在云数据中心的。而国内外关于云数据中心的请求调度与优化方法的相关研究可进一步为近年来兴起的技术包括大数据处理、物联网、智慧医疗、车联网等提供基础资源支撑。因此，需要进一步改进当前关于云数据中心请求调度与优化方法的研究，从而为以上技术的实现提供智能、绿色、高效的基础支撑。未来具体的研究方向如下。

#### (1) 大数据处理

目前针对云中大数据的研究工作尚未形成成熟可用的技术，很多现存的问题仍未得到完善的解决<sup>[30]-[33]</sup>。首先，随着云用户的显著增加，如何保证大规模数据在云中的可用性是一个目前尚未有效解决的难题。比如，需要设计高效的方法来保证移动用户在短时间内获取存储在云数据中心中的大规模数据。其次，存储在云数据中心的的数据的质量难以得到保证。以往的云数据中心中的数据往往是来自于已知的固定数据源的规范数据。然而在大数据环境下数据往往来自于多个不同的数据源，部分数据源未经确认因此其数据质量难以得到保证。这将为云数据中心提供商带来

严峻的挑战。因此，需要进一步研究如何从大量不同数据源中获取较高一致性的高质量数据。另外，如何选取合适的模型来对大规模的数据进行分析也具有至关重要的作用，基于该模型需要进一步设计具有高扩展性的分析算法，以达到在短时间内对数据进行分析处理的目的。

#### (2) 物联网

物联网技术近年来在国内外得到了广泛的关注和研究<sup>[34]</sup>。数据表明到 2020 年通过物联网相连接的设备数将达到 240 亿，这些快速增长的设备将产生大量的数据。而物联网终端设备往往是低成本和轻量级的，不能有效处理这些数据。因此，在设备本身中存储、处理和分析这些数据将远远不能满足未来的需求。而云计算技术能够为物联网技术提供海量的数据存储资源和快速处理数据的计算资源。然而，目前仍存在以下主要问题。首先，物联网和云之间的数据通信需要消耗大量的能量。尤其是视频数据的采集、加密、解密均需要大量的能量，而传统的供电设备（比如电池）远远不能满足这些能量需求。因此，需要为传感器设计从其所在环境生成能量的方法（如通过太阳能、风能等）。另外需要设计有效的休眠模式从而适时地降低传感器的能量消耗。其次，物联网中的终端设备多种多样，其对云中资源的请求也存在较大差异。因此，云提供商很难准确确定这些设备所需的资源数量。因此，需要设计有效的资源分配方法，该方法需要综合考虑传感器产生的数据的类型、大小、频率等因素。

#### (3) 智慧医疗

近年来信息技术在医疗领域<sup>[35]</sup>得到了广泛的应用，因此产生了大量的电子医疗数据。目前的医疗数据具有大规模、增长迅速、结构复杂、深层次信息多等典型特征。而云计算技术能够对大规模的数据进行快速地存储、索引、处理和分析，从而实现智慧的医疗服务。然而，目前云计算在智慧医疗领域的应用仍然存在若干难题。首先，多源异构数据之间存在融合问题。目前尚无统一的标准来对不同数据源的医疗数据进行有效的管理，因此对存储在云中的数据的数据的分析和处理就面临较大的挑战。其次，目前企业所拥有的数据与其能够处理的数据之间的差距越来越大。因此，针对现有的大规模医疗数据，如何快速有效地从海量低质量数据中快速抽取出高质量的关键信息，从而更好地利用这些医疗数据以提供个性化的健康指导、公共的健康预警等服务，仍需得到进一步的研究。

#### (4) 车联网

云计算所具有的海量的存储和计算资源同样促进了车联网<sup>[36][37]</sup>的发展，近年来出现了车载云的概念。在车载云中，具有充足资源的车辆可作为移动云服务器，向其他车辆提供资源或者网络连接服务。车联网中收集得到的数据（如红绿灯、停车计时器、摄像头、城市终端传感器等采集的数据）可以通过云数据中心进行高效的存储、处理和分析，并进一步构建得到道路交通的模型以供道路管理部门使用，从而及时做出管理决策（比如修复道路、改变交通方向、安装新的红绿灯等）。然而，目前仍有几个主要难题需要解决。首先，车辆具有高度的移动性，作为云服务器的

车辆只能在较短时间内为其覆盖区域内的其他车辆提供资源服务，这会显著影响车载云的性能和降低车载云的经济价值。因此，需要设计一种适用于高度移动环境下的高可用的资源保障机制。其次，车载云中的安全和隐私问题仍需得到解决。比如，在高度移动的环境下，在有限区域内相互通信的多个车辆之间如何建立并维持信任关系。此外，在车载云中，如何保证隐私数据的保密性以及信息的完整性等问题仍然需要进一步的研究。

#### 4. 小结

本文结合国内外相关的研究，可推断，近年来，尽管对云数据中心的研究已有一定的研究成果，但在智能、绿色、高效、经济的云数据中心的建设方面还存在诸多问题尚未解决。另外，云计算技术在其他领域（如大数据处理、物联网、智慧医疗、车联网）的应用仍然存在较多挑战，需要进一步的研究。因此，云数据中心仍需更多创新性的研究，进而提供高效成熟的资源分配和请求调度技术，最终有效提高应用请求的执行效率和云数据中心提供商的收益和利润。

#### 5. 致谢

本研究获中国博士后科学基金项目（编号：2017T100034 2016M600912）及北京市重点实验室课题（编号：HDN2017101）资助。

#### 参考文献

- [1] Armbrust M., Fox A., Griffith R., et al. Above the Clouds: A Berkeley View of Cloud computing [M]. *Technical Report No.*

- UCB/EECS-2009-28, University of California at Berkeley, USA, Feb. 10, 2009.
- [2] Wu X., Zhu X., Wu G. Q., et al. Data Mining with Big Data [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2014, 26(1): 97-107.
- [3] Lin J., Amini S., Hong J. I., et al. Expectation and Purpose: Understanding Users' Mental Models of Mobile App Privacy through Crowdsourcing, *Proceedings of the 14th ACM International Conference on Ubiquitous Computing*, Pittsburgh, Pennsylvania, USA, 2012: 501-510.
- [4] Ellison N. B., Vitak J., Gray R., et al. Cultivating Social Resources on Social Network Sites: Facebook Relationship Maintenance Behaviors and Their Role in Social Capital Processes [J]. *Journal of Computer Mediated Communication*, 2014, 19(4): 855-870.
- [5] Rothenberg C. E., Nascimento M. R., Salvador M. R., et al. Revisiting Routing Control Platforms with the Eyes and Muscles of Software-Defined Networking, *Proceedings of the 1st Workshop on Hot Topics in Software Defined Networks*, 2012, 13-18.
- [6] Agarwal S., Kodialam M., Lakshman T. V. Traffic engineering in software defined networks, *Proceedings of the 32nd IEEE International Conference on Computer Communications*, 2013: 2211-2219.
- [7] Tso F. P., Hamilton G., Weber R., et al. Longer Is Better: Exploiting Path Diversity in Data Center Networks, *Proceedings of the 2013 IEEE 33rd International Conference on Distributed Computing Systems*, 2013, 430-439.
- [8] Meng X., Pappas V., and Zhang L. Improving the Scalability of Data Center Networks with Traffic-Aware Virtual Machine Placement, *Proceedings of the 29th IEEE International Conference on Computer Communications*, 2010, 1-9.
- [9] Nahir A., Orda A., and Raz D. Schedule First, Manage Later: Network-Aware Load Balancing, *Proceedings of the 32nd IEEE International Conference on Computer Communications*, 2013, 510-514.
- [10] Yuan H., Bi J., and Li B. H. Workload-Aware Request Routing in Cloud Data Center Using Software-Defined Networking. *Journal of Systems Engineering and Electronics*, 2015, 26(1): 151-160.
- [11] Zhang P. and Zhou M. C. Dynamic Cloud Task Scheduling Based on a Two-stage Strategy. *IEEE Trans. on Automation Science and Engineering*, 2017, DOI: 10.1109/TASE.2017.2693688, on-line.
- [12] Ghahramani M. H., Zhou M. C., and Hon C. T. Toward Cloud Computing QoS Architecture: Analysis of Cloud Systems and Cloud Services. *IEEE/CAA Journal of Automatica Sinica*, 2017, 4(1): 5-17.
- [13] Rahman M., Li X., and Palit H. Hybrid Heuristic for Scheduling Data Analytics Workflow Applications in Hybrid Cloud Environment, *Proceedings of the 2011 IEEE International Symposium on Parallel and Distributed Processing Workshops and Ph.D. Forum*, 2011, 966-974.
- [14] Bittencourt L. F., and Madeira E. R. M. HCOC: A Cost Optimization Algorithm for Workflow Scheduling in Hybrid Clouds. *Journal of Internet Services and Applications*, 2011, 2(3): 207-227.
- [15] Bittencourt L. F., Senna C. R., and Madeira E. R. M. Scheduling Service Workflows for Cost Optimization in Hybrid Clouds, *Proceedings of the 2010 International Conference on Network and Service Management*, 2010, 394-397.
- [16] Mezmez M., Melab N., Kessaci Y., et al. A Parallel Bi-Objective Hybrid Metaheuristic for Energy-Aware Scheduling for Cloud Computing Systems [J]. *Journal of Parallel and Distributed Computing*, 2011, 71(11): 1497-1508.
- [17] Nan X., He Y., and Guan L. Optimal Resource Allocation for Multimedia Cloud Based on Queuing Model, *Proceedings of the 2011 IEEE 13th International Workshop on Multimedia Signal Processing*, 2011, 1-6.
- [18] Shi Y., Jiang X., and Ye K. An Energy-Efficient Scheme for Cloud Resource Provisioning Based on CloudSim, *Proceedings of the 2011 IEEE International*

- Conference on Cluster Computing*, 2011, 595-599.
- [19] Zhang Q., Zhu Q., and Boutaba R. Dynamic Resource Allocation for Spot Markets in Cloud Computing Environments, *Proceedings of the 2011 Fourth IEEE International Conference on Utility and Cloud Computing*, 2011, 178-185.
- [20] Khazaei H., Mistic J., Mistic V. B. Performance Analysis of Cloud Computing Centers Using M/G/m/m+r Queuing Systems [J]. *IEEE Transactions on Parallel and Distributed Systems*, 2012, 23(5): 936-943.
- [21] R. Ranjan, R. Buyya, P. Leitner, et al. A Note on Software Tools and Techniques for Monitoring and Prediction of Cloud Services. *Software: Practice and Experience*, 2014, 44(7), pp. 771-775.
- [22] Yuan H., Bi J., Tan W., and Li B. H. Temporal Task Scheduling with Constrained Service Delay for Profit Maximization in Hybrid Cloud. *IEEE Transactions on Automation Science and Engineering*, 2017, 14(1): 337-348.
- [23] M. Ghamkhari, and H. Mohsenian-Rad. Energy and performance management of green data centers: A profit maximization approach. *IEEE Transactions on Smart Grid*, 2013, 4(2), pp. 1017-1025.
- [24] Deng X., Wu D., Shen J., and He J. Eco-aware online power management and load scheduling for green cloud datacenters. *IEEE Systems Journal*, 2014, 10(1): 78-87.
- [25] Amokrane A., Langar R., Zhani M., Boutaba R., and Pujolle G.. Greenslater: On satisfying green SLAs in distributed clouds. *IEEE Transactions on Network and Service Management*, 2015, 12(3): 363-376.
- [26] Guo Y., Gong Y., Fang Y., Khargonekar P., and Geng X. Energy and network aware workload management for sustainable data centers with thermal storage. *IEEE Transactions on Parallel and Distributed Systems*, 2014, 25(8): 2030-2042.
- [27] Leslie L., Lee Y.C., P. Lu, and Zomaya A. Exploiting performance and cost diversity in the cloud, *Proceedings of the 6th IEEE International Conference on Cloud Computing*, 2013, 107-114.
- [28] Nguyen K. K. and Cheriet M. Environment-aware virtual slice provisioning in green cloud environment, *IEEE Transactions on Services Computing*, 2015, 8(3): 507-519.
- [29] Bi J., Yuan H., Tan W., and Li B. H. TRS: Temporal Request Scheduling with Bounded Delay Assurance in a Green Cloud Data Center. *Information Sciences*, 2016, 360(C): 57-72.
- [30] Liu B., Huang K., Li J., and Zhou M. C. An Incremental and Distributed Inference Method for Large-Scale Ontologies Based on MapReduce Paradigm. *IEEE Transactions on Cybernetics*, 2015, 45(1): 53-64.
- [31] Luo X., Zhou M. C., Shang M., Li S., and Xia Y. A Novel Approach to Extracting Non-Negative Latent Factors from Non-Negative Big Sparse Matrices. *IEEE Access*, 2016, 4: 2649-2655.
- [32] Luo X., Zhou M. C., Li S., You Z., Xia Y., and Zhu Q. A Non-negative Latent Factor Model for Large-scale Sparse Matrices in Recommender Systems via Alternating Direction Method. *IEEE Transactions on Neural Networks and Learning Systems*, 2016, 27(3): 579-592.
- [33] Sun H. C., Jiang C. J., Ding Z. J., Wang P. W., and Zhou M. C. Topic-oriented Exploratory Search Based on an Indexing Network. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2016, 46(2): 234-247.
- [34] Zhou M. C., Fortino G., Shen W., Mitsugi J., Jobin J., and Bhattacharyya R. Guest Editorial: Special Section on Advances and Applications of Internet of Things for Smart Automated Systems. *IEEE Transactions on Automation Science and Engineering*, 2016, 13(3): 1225-1229.
- [35] Liu B., Li J., Chen C., Tan W., Chen Q., and Zhou M. C. Efficient Motif Discovery for Large-scale Time Series in Healthcare. *IEEE Transactions on Industrial Informatics*, 2015, 11(3): 583-590.
- [36] J. J. Cheng, J. L. Cheng, M. C. Zhou, Q. Zhang, C. Yan, Y. Yang, and C. Liu, "Routing in Internet of Vehicles: A Review," *IEEE Trans. on Intelligent Transportation Systems*, 16(5), pp. 2339-2352, Oct. 2015.1)

- [37] X. Wang, C. Wang, J. Zhang, M. C. Zhou, C. Jiang, "Improved Rule Installation for Real-Time Query Service in Software-Defined Internet of Vehicles," *IEEE Trans. on Intelligent Transportation Systems*, 18(2), pp. 225 – 235, Feb. 2017.

## 作者介绍



**Haitao Yuan (苑海涛)** received the B.S. and M.S. degrees in Software Engineering from Northeastern University, Shenyang, China, in 2010 and 2012, respectively, and the Ph.D. degree in Control Science and

Engineering from Beihang University, Beijing, China in 2016. He is currently an Assistant Professor with the School of Software Engineering, Beijing Jiaotong University, Beijing, China. He was a visiting scholar with the New Jersey Institute of Technology (NJIT), Newark, NJ in 2015. His research interests include cloud computing, resource allocation, software-defined networking, and energy efficiency. Dr. Yuan was the recipient of the 2011 Google Excellence Scholarship.



**MengChu Zhou (周孟初)** received his B.S. degree in Control Engineering from Nanjing University of Science and Technology, Nanjing, China in 1983, M.S. degree in Automatic Control from Beijing Institute of

Technology, Beijing, China in 1986, and Ph. D. degree in Computer and Systems Engineering from Rensselaer Polytechnic Institute, Troy, NY in 1990. He joined New Jersey Institute of Technology (NJIT), Newark, NJ in 1990, and is now a Distinguished Professor of Electrical and

Computer Engineering. His research interests are in Petri nets, Internet of Things, control systems, wireless communication, and energy systems. He has over 690 publications including 12 books, 370+ journal papers (270+ in IEEE Transactions), 11 patents and 28 book-chapters. His recently co-authored/edited books include *Business and Scientific Workflows: A Web Service-Oriented Approach*, IEEE/Wiley, New Jersey, 2013 (with W. Tan) and *Contemporary Issues in Systems Science and Engineering*, IEEE/Wiley, New Jersey, 2015 (with H.-X. Li and M. Weijnen).

He was invited to lecture in Australia, Canada, China, France, Germany, Hong Kong, Italy, Japan, Korea, Mexico, Saudi Arabia, Singapore, Taiwan, and US and served as a plenary/keynote speaker for many conferences. He is the founding Editor of IEEE Press Book Series on Systems Science and Engineering. He served as Associate Editor of IEEE Transactions on Robotics and Automation from 1997 to 2000 and IEEE Transactions on Automation Science and Engineering from 2004-2007, and Editor of IEEE Transactions on Automation Science and Engineering from 2008-2013. He is Associate Editor of IEEE Transactions on Systems, Man and Cybernetics: Systems, IEEE Transactions on Industrial Informatics, IEEE Transactions on Intelligent Transportation Systems, IEEE/CAA Journal of Automatica Sinica, and Frontiers of Information Technology & Electronic Engineering. He served as a Guest-Editor for many journals including IEEE Journal of Internet of Things, IEEE Transactions on Industrial Electronics, IEEE Transactions on Semiconductor Manufacturing. He was General Chair of IEEE Conf. on Automation Science and Engineering, Washington D.C., August 23-26, 2008, General Co-Chair of 2003 IEEE International Conference on System, Man and Cybernetics (SMC), Washington DC, October

5-8, 2003, Founding General Co-Chair of 2004 IEEE Int. Conf. on Networking, Sensing and Control, Taipei, March 21-23, 2004, and General Chair of 2006 IEEE Int. Conf. on Networking, Sensing and Control, Ft. Lauderdale, Florida, U.S.A. April 23-25, 2006. He was Program Chair of 2010 IEEE International Conference on Mechatronics and Automation, August 4-7, 2010, Xi'an, China, 1998 and 2001 IEEE International Conference on SMC and 1997 IEEE International Conference on Emerging Technologies and Factory Automation. He organized and chaired over 100 technical sessions and served on program committees for many conferences. Dr. Zhou has led or participated in over 50 research and education projects with total budget over \$12M, funded by National Science Foundation,

Department of Defense, NIST, New Jersey Science and Technology Commission, and industry. He is a recipient of Humboldt Research Award for US Senior Scientists, Franklin V. Taylor Memorial Award and the Norbert Wiener Award from IEEE Systems, Man and Cybernetics Society. He has been among most highly cited scholars for years and ranked top one in the field of engineering worldwide in 2012 by Web of Science/Thomson Reuters. His Google citation count is well over 22,000 and H-index is 75. He is a life member of Chinese Association for Science and Technology-USA and served as its President in 1999. He is a Fellow of IEEE, International Federation of Automatic Control (IFAC) and American Association for the Advancement of Science (AAAS).

# Personal Archive Service System using Blockchain Technology

## 区块链技术在个人存档服务系统的应用

陈志雄, 朱仪轩

数学计算机系, Mercy College

zxchen@ieee.org; yzhu@mercymavericks.edu

### Abstract

This paper is to explore applications of the blockchain technology to the concept of “proof of X” such as proof of identity, proof of property ownership, proof of specific transaction, proof of college degree, proof of medical records, proof of academic achievements, etc. It describes a novel approach of building a decentralized transparent immutable secure personal archive management and service system. Personal archive is defined as a collection of various artifacts that reflect personal portfolio as well as personal unique identifications. Personal portfolio is beyond of a statement of personal achievement. It is an evidentiary document designed to provide qualitative and quantitative chronically documents and examples. Subjects can tag their information with proof, that is, certified by trusted entities or organizations like universities. Such proofs are associated with confidentiality levels exposed in the public domains. Personal identifications include biometrics as well as other multi-factors such as something the subject “has”, the subject “knows” or the

subject “acts”. Stack holders in a consortium oriented blockchain network serve as verifiers and /or miners that provide their trusted services the delegated proof of stake. Such personal archive based system can be exploited to various applications including professional network like Linkedin, instant credit approval like alipay or live human from social bots like internet social media. A prototype simulation shows that such personal portfolio management and service system is feasible and immune to many ID attacks.

### 摘要

本文旨在探讨利用区块链技术运用到没有第三方参与的“证明 X”的概念，如身份证明，财产所有权证明，具体交易证明，大学学位证明，医疗记录证明，学术成就证明等等。文章介绍了建立分布式、透明、不可篡改并且安全的个人档案管理和服务体系的一个新途径。个人档案包含了个人成就集以及个人身份唯一的标识。个人成就集囊括了个人成就的声明，是一种证据类文件，长期的提供对定性和定量的文件或者事件的证明。该项目可以通过可信任的实体或者机构，例如大学，来证

明所标注的信息，这些证明甚至可以在公共领域的机密信息。个人身份标识包括了生物识别以及其他多种因素，如个体“拥有”的东西，个体“知道”的信息或个体“做过”的事情。股权代表在面向联盟的区块链网络中可以作为验证者或“矿工”，为委任权益证明提供可信赖的服务。这样的个人档案系统可以被利用于多种应用中，包括专业型网络 LinkedIn，有着即时信用系统的支付宝，或者需要区分人与机器的社交媒体。模拟原型后发现，这种个人成就集的管理和服务系统是可行的，并且可以免疫许多 ID 攻击。

**关键词：**个人档案服务系统，数字成就集，鉴别，区块链，联盟链，公钥和私钥，匿名，透明，去中心化

## I. 背景介绍

个人档案服务系统（PASS）的定义是数字化的个人成就集和个人信息集，以及为这些集合服务的所有工具的集合。其中，个人数字化成就集（PDA）是个人成就的数字形式，具有可用于鉴别个人成就的唯一证据文件（PAE）和个人识别特征（PID）。PAE 包括但不限于个人的教育，经验，培训，学位，文凭，成绩单和各种证书。个人财富如财产价值，银行余额，投资也可以是 PAE 的例子。PID 的典型例子是指纹，虹膜或静脉等生物特征值的测量。其他信息比如个人所知的，个体所拥有的物品也可以成为一种 PID，此外，行为模式，个人观点和特征等现在也可以认为是一种 PID。

PAE 的集合类似于我们许多文献讨论的个人成就的集合体（PP）。它与一个人在特定的时间特定的地点等相关联。PP 与简历不同，因为它比简历中的一行语句更具包容性和扩展性。人们可以从中了解关于主体更多的信息。为了使 PDA 可证明主体并被用户信任，这些 PDA 需要由第三方进行验证。这个第三方应该对声明文件有参与、证明或颁发等权限。例如，大学可以向毕业生提供所谓官方的成绩单，业主可以向他的租户们提供房客参考书。

我们先在这个验证过程里定义各种角色。我们将“主体”或“个体”一词作为正在建立他或者她的 PDA 集合的个人或用户，“认证者”成为提供认证的机构或实体，“查询者”是提供验证服务的代理或组织。他们调查和获取任何特定主体的相关证据。“客户”是使用以上“查询者”提供验证服务的群体或个人。例如，一个公司要雇佣一个潜在的候选人。候选人是“主体”，该公司是一个“客户”，这个公司一般需要雇用第三方来验证“主体”提供的信息。第三方是“查询者”，公司是“查询者”的“客户”。“查询者”需要联系各种“认证者”，即认证机构，例如获得教育的大学，曾经工作的公司或发行各种其他认证的组织。

这种由查询者验证 PDA 的过程通常是耗时的。每当有客户提出请求时，查询者就必须验证所有的材料。绝大部分时候，认证者收到的请求都是重复的。在有些情况下，查询者有可能获得超出个体同意的查询范围和信。这会构成了对个人隐私的威胁。此外，用计算机的概念来代替以

上验证模型，可以看出其本质是一个中央系统模型，一个点的失败、滥用、混乱、延迟极其会拖累整个系统的进程甚至导致奔溃。例如，在用第三方来查询招聘的兼课教师的教学许可和资质。我们亲身体会过，在较为夸张的情况下，即使学期已经开始，还依旧在等待这些验证的完成，从而不得不临时换老师。另一个例子是主体在某一阶段把自己的名字改了，但如果不透露自己以前使用的名称，许多验证很难通过。有这样一个案例，一个主体希望他毕业的大学发给他一个新的文凭，因为他把他的名字改了，并且也不希望有人知道他以前用的名字，但通常这是不可能的，学校不会重新发布一个新名字的文凭。尽管从计算机角度来看，在这种情况下，该主体所拥有的依旧是这个 ID，只是用户名不同了。

与此同时，现阶段验证服务变得越来越重要。我们观察到“履历作假”正在变得越来越流行<sup>[1]</sup>。履历作假是指向自己的履历里添加虚假或夸张的信息，来“增强”个体的能力。超过 40%的履历夸张了自己的工资，33%的对自我工作的描述不准确，29%的对工作日期进行了作假，27%的伪造参考信息和 21%的有一定程度的欺诈。因此，可想而知的是，验证任务面临着十分严峻的挑战。

最近，首先在比特币<sup>[20]</sup>中发掘出的区块链技术作为一种创新的交易网络和一种新的货币被应用于金融相关的行业。许多其他领域，如供应链、制造、物联网也在积极地使用这个技术<sup>[21-22]</sup>。它的关键特性是其去中心化，基于共识的分布式帐本，

智能合约，高透明度，高隐私性。其交换网络中不可篡改性、信息安全性、抗攻击性，使其也开始出现在贸易，融资，运输和保险业等需要验证个体或资产的情景下。它不需要有权威的第三方、结算方或任何组织作背书，就能被用于构建一个安全匿名但透明的不可篡改的 ID 服务<sup>[33]</sup>。

本文提出的个人存档服务系统（PASS）是使用区块链技术的不可篡改的，透明的，匿名的和公众共识的理想特征。主体控制自己的 PDA，并决定向谁公布。图 1 用模块的方式显示了 PASS 的一般基础结构和体系结构。最左边的是主体，有其自己的存储库或钱包，聚合了所有与主题相关的 PDA。验证者在得到主体的要求后向主体颁发证书并告示受信任的网络参与者。这种操作只需执行一次就可以了，没有必要有第三方或者专门的查询者。网络参与者之间的信任是由他们在这个区块链网络里的授权的信托来发展的。客户向主体发出请求，主体可以授予客户以访问 PDA 的权限。

我们在这里主要的贡献是提出了一个使用区块链技术构建一个相关证书的个人档案框架。它保持了真实性，准确性和透明度，同时也保持了高度的隐私。查询者已不再被需要。一个主体可以根据客户请求的性质来决定向谁、何时、显示的内容等。比如，当请求具有学术成就时，与该标签相关联的任何 PDA 可以被解锁并显示出来；当请求是识别个体时，主体可以解锁那些相关的 PDA，例如生物特征。客户可以立即访问这些 PDA，并及时做出相应的决定。此外，PASS 提供的工具和服务可

以容易的关联那些与个人成长时间轴相关的项目。

以下章节回顾了使用区块链技术的“证明 X”概念的最新发展，其中 X 可以是身份，财产所有权，具体交易，大学学位，医疗记录和学术成就等等。接下来，下一节将讨论 PASS，个人档案服务系统。个体可以及时准确地构建个人档案。这些档案经过认证，不需要每次使用时再进行一次验证。PASS 能够保证透明度，同时也保证隐私性。最后一部分是关于机会，挑战性和未来应用的讨论。

## II. 文献讨论

构建个人数字档案（PDA）的重点是验证的过程和可信赖性。在第一节里，我们提到的统计，超过 30% 的求职者修改了文凭资料，70% 的拔高了他们取得的成就<sup>[1]</sup>。因此，围绕学术成果建立的验证和认证已有许多成果，开放徽章（Open Badge）的想法是其中之一。开放徽章的项目工作组在其网站<sup>[2]</sup>中指出，通过提供包含可验证数据和可以在网络上共享的证据的成就的徽章。开放徽章使个体能够管理自己的学习成果，项目工作组的目标是将许多证书转向更加清晰，更容易符合，维护和构建的标准。Acclaim<sup>[3]</sup>是基于开放徽章标准的数字标牌平台，由皮尔逊公司（Person）支持。Acclaim 已经从通过像 IBM 那样的信誉良好的组织发行了数百万张徽章，用于促进个人职业的向前发展。通过 Acclaim 发行的徽章是学习成果，经验或能力的数字化表现形式。这些徽章可以以简单安全的方式在线共享和验证，并包含详细的信息，提供关于实际完成的内容，哪个组织认可成

就以及获得认可的个人，提供的服务有其优点，但主要缺点是其集中模式。

索尼全球教育（Sony Global Education）宣布将在 2016 年开发并使用区块链技术于开放式、安全的学术水平与进步记录的技术<sup>[4]</sup>。它将利用区块链的安全属性来实现两个指定方之间数据的加密传输，例如个人的学术水平记录和进度的措施。实际的工作流程现在还没有公布，结果还有待观察，需要进行测试。

身份证制度是一个长期存在的问题。W3 确定了全球身份管理服务的七个一般要求<sup>[9]</sup>：可移植性、互操作性、可扩展性、协商隐私和安全性、问责制、分布式注册机构、分布式证书颁发机构和独立管理机构。此类服务必须以通用交换格式使用全局唯一标识符，支持从其他常用标识符扩展到这些标识符映射，并使用通用协议来断言和认证全局标识。它必须支持全球词汇定义以及分布式本地词汇定义，需要支持匿名和假名来保护个人隐私。因此，保护个

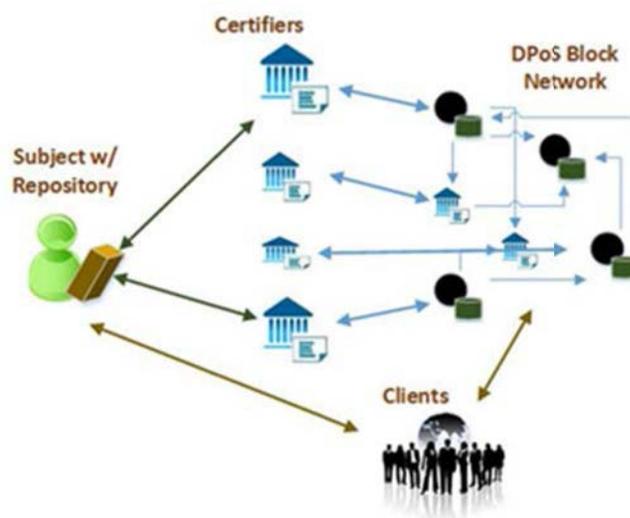


图 1 PASS

人隐私的匿名和假名属于隐私和安全的范围。

ID 系统应该支持分层和对等注册模型。管理机构应当被列为国际非营利组织，因此在各方面都是行业，供应商和政府中立的。它应该为服务设定技术和操作标准，因为两者紧密相连；应该管理通用身份属性和全局协议控制结构的全球词汇开发；应为所有代理人，包括注册和认证机构确定责任条件；应该作为分级注册或认证模型的公正的根本权威。

OpenID 和 OpenID Connect<sup>[6-11]</sup>是定义和开发开放标准和分散式身份验证协议。根据 OpenID 提供商（OP）或授权服务器执行的身份验证，依靠方（RP）可以验证终端用户的身份，并且可以在可互操作的和像 REST-风格中获取关于最终用户的基本配置文件信息。通过允许用户通过第三方服务进行身份验证，消除了网站管理员提供自己的特别登录系统和单独的身份和密码的需要。

OAuth<sup>[19]</sup>是允许网络安全授权的开放协议，使第三方应用程序能够通过 http 服务获取特定的访问权限，但不提供身份识别管理和验证。OpenID Connect 是 OAuth 2.0 协议之上的一个简单的身份层。虽然 OAuth 允许第三方授权服务，但仍然集中在 OpenID 提供者，用户必须透露许多私人信息才能进行身份验证，也容易形成单点故障。而且，这是互联网的本质，不是一个控制谁能做什么的中央权力机构。

可以使用多个生物特征信息以非常高的精度识别人类<sup>[8, 32, 33]</sup>。这在社交媒体应用

程序中尤其有用，其中用户可以匿名，但需要被标记为真实的人，而不是社交机器人或代码。因此，来自社交媒体应用的统计或投票等可以代表真实情况，而不是社交机器人可以左右的。

一些 Web 应用程序，如 taskstream<sup>[6]</sup>为用户提供了构建个人成就集的平台。它可以跟踪特定项目的进度，并记录特定主体的观点。这些应用程序可以作为一种展示个人成长时间轴的方式，可以提供特定技能的来源。

在<sup>[35]</sup>中，作者提出了使用区块链在没有中央控制数据库的情况下在数字世界中证明身份的想法，因为第三方控制自然地会导致一定程度的垄断。而分散的系统遵循通用协议，允许个人添加新的事务并使用具有不可篡改性的点对点架构进行分发。主体可以通过智能合约让银行，保险公司或政府等客户来检查文件。智能合约是在公共区块中记录的一段代码。合约可以限制查询的数量或时间，并将其全部记录在案。

### III. PASS 的结构

#### A. 成就, 成就集和档案

个人数字成就集（PDA）是一块签名的数字文档，其中包含以下标签或字段：

- PDA-ID: 由系统生成的索引
- PDA 类型: 具有证据文件（PAE）或个人身份（PID）的个人成就。它遵循文凭，学位，技能和生物测量等子类型。最终应该是标准化的代码。

- PDA 说明：描述 PDA 性质的字符串
- PDA 主体：拥有 PDA 的人的姓名
- PDA 认证人员：向该 PDA 发行认证的人员或组织
- PDA-数据-开始：PDA 的开始日期
- PDA-数据-结束：PDA 的完成日期
- PDA-日期验证：PDA 有效的可选日期
- PDA-注释：可选字段
- PDA-钥匙-证书：公钥与验证者相关联
- PDA-钥匙-子证书：为主体和这个特定的 PDA 生成的公钥。私钥由主体在其应用中保存
- PDA-唯一：验证者用来验证该 PDA 的所有权。出生日期，学生证或通用 ID，和/或仅与受试者有关的知识，可以是验证者所需的多个字段
- PDA-数据：验证者所需的数据。对于 PID 类型，可以是特定的生物特征信息；对于 PAE 类型，可以扫描原始证书（硬拷贝）

个人成就集合体 (PP) 是在 PDA 类型里的 PAE (具有证据文件的个人成就) 的集合。个人识别 (PID) 是在 PDA 类型里 PID 的集合。因此，PA 是 PP 和 PID 的联合。象征性的，我们可以用下面的符号来表示我们所描述的。

$$PA = \{P_p, P_{id}\},$$

其中

$$P_p = \{PDA_{PAE1}, PDA_{PAE2}, \dots, PDA_{PAEn}\}.$$

和

$$P_{id} = \{PDA_{PID1}, PDA_{PID2}, \dots, PDA_{PIDm}\}.$$

## B. 服务工具

服务工具是执行 PP 所需的特定任务的一段代码。它应该包括以下字段。

- ST-ID：系统生成的索引
- ST 型：咨询者，经理或其他人的服务性质。
- ST 描述：描述服务性质的字符串
- ST 输入：服务条件和参数
- ST 输出：服务结果
- ST 条件：在什么条件和/或特权下应该执行服务
- ST 创作者：发行工具的人或组织
- ST-日期：服务到位的日期
- ST-注释：可选的领域
- ST 语言：实现该工具的特定语言。

服务工具的集合表示为 T。它涵盖了使 PA 成为可用的各种服务。例如，请求证书，在本地存储库或电子钱包中汇总和同步 PP，从公共分类帐中，授予特定方检查 PDA 的权限。再此，我们可以使用以下符号来说明一组工具：

$$T = \{ST_1, ST_2, \dots, ST_n\}.$$

所以，PASS 可以抽象的表示为

$$PASS = \{PA, T\}$$

C. 主体，验证者，查询者，客户，股权代表

本文使用以下术语来描述基于区块链的分布式对等网络。

- “主体”一词是正在讨论他或她的 PDA 的人
- “验证者”一词是提供认证的机构或实体
- “查询者”一词是调查和获取相关证明的代理人或组织
- “客户”一词是使用专业服务的人员或组织（“审讯者”）。
- “股权代表”这个术语是一个面向联盟的区块链网络中的一个特殊节点。信托是以授权的证明来发展的。它就像一个比特币网络中的“采矿与验证”。

例如，对于公司潜在的候选人来说，候选人是主体，公司需要聘请第三方核实主体提供的所有信息。第三方是查询者，公司是查询者的客户。查询者需要联系各种验证者，例如获得教育的大学，曾经工作的公司或发行其他认证的组织。

#### D. 联盟网络的初始信任和区块

通常，决定在 p2p 网络中区块链的初始状态或第一区块是一项困难的任务。所有对等的参与者都需要一个可以接受的规则。PASS 采用联盟网络，其中具有某些声誉的股权所有者可以包含或投票。因此，我们可以假设这样的初始区块包括 PASS 中的数据结构和几个特殊的股权代表。验证者可以是其中一个股权代表，但不一定是必须的。为了成为一个值得信赖的验证者，验证者在成为股权代表前提供其身份。

#### E. 股份授权证明机制 (DPoS)

工作量证明 (PoW) 是比特币网络中使用的一种共识算法。PoW 是通过采矿完成的，这是分散式结算里最主要的过程，这些过程对交易进行了验证和撰写。采矿确保了比特币系统，并且在没有中央权力机构的情况下实现了全网络共识的出现。争取解决工作量证明所提出的问题是获得奖励和在区块链上获取记录交易的权利是比特币安全模型的基础。

一个更有效的共识算法是股份授权证明机制 (DPoS)。它是股权证明 (PoS) 的进一步演化，两者都是为了降低与 PoW 算力相关的成本和低效用电而开发的。PoS 允许包含硬币的每个钱包都是股权所有者，即参与验证交易和形成分布式共识的过程，并赚取硬币作为回报。在 DPoS 中，包含硬币的每个钱包都可以投票选举代表，这些代表是执行验证交易和维持块状的功能，并将交易费用作为利润，比 PoS 更有效率，更节省算力。因此，PASS 选择使用 DPoS。

#### F. 生物特征和唯一性

计算机科学中的生物特征是使用与人类特征相关的指标进行识别，认证和访问控制。指纹、掌纹、脸部识别、DNA、手型几何、虹膜识别、视网膜等，其他指标可以与诸如语音、打字节奏和行走模式之类的行为相关，并且与某些兴奋剂的反应有关<sup>[25, 26]</sup>。随着技术的进步，生物识别的准确性大大提高，获取这些信息所需的时间要快得多，获得它们的方式更容易，成本更低。

生物统计认证的一个主要方向是快速检测和识别，以便这样的系统可以用于实时认证<sup>[25, 26]</sup>。在不同的应用场景中，生物识别被用于区分人与机器人。有挑战性的是高效率地检测非人类生物识别技术，例如大量人体生物识别技术的数字修改，这在需要匿名但可识别的人性的应用中特别有用。我们正在调查各种方法，如主成分分析，小波和相关性来测试这种分类<sup>[27]</sup>。

像其他 PDA 一样，实际的生物识别技术正在使用，但不存储在网络中。单向函数使得很难获得原始的生物特征。此外，需要多个生物特征信息来减少假阳性的概率。因此，不可能有两个人在某种程度上具有相同的生物特征。

#### IV. PASS 中的通讯

我们现在来描述一个主体获得 PDA 的一般顺序。第一节中的图 1 演示了工作的流程。主要的构想是在联盟区块链里具有授权证明的注册节点可以将结果记录在全局分类帐中。这些节点可以是有信誉的组

织和公司，认证者不一定非要是网络中的一个节点，主要功能是向主体提供证书。例如，大学毕业生可以要求官方成绩单并成为其 PDA 之一，毕业生是主体，大学是认证机构。区块链网络可以由有信誉的并有动力的组织和公司来维护整个社区，大学、教科书出版社和培训学会是理想的候选人。客户可以是要向该主体提出需求的任何个体或组织。

图 2 说明了 PDA-PAE 的时序，为了更好的说明而没有详细的数据结构交换。在这里，主体向验证者要求获得一个特定成就的证书，验证者是已经注册并获得授权发放此类证书的机构。这些信息由验证者的公开密钥加密，以便没有人可以查看 PDA。当验证者获得请求时，首先验证 PDA 中的请求主体信息，并做出拒绝或验证的响应。实际上，主体和认证者之间的信息交流是相当复杂，可能会要求主体提供有关 PDA 的更多证据和具体信息；或者在签发证书之前，可以要求提供付款服务证明。验证者不仅给主体发放证书，还将

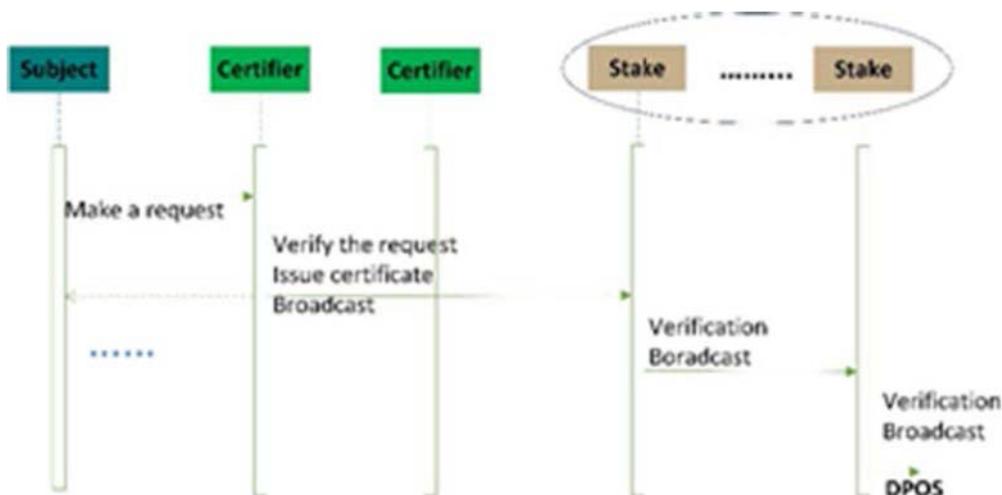


图 2 PDA-PAE 的时序

经过处理的证书代码发送到联盟区块链网络中的其他节点。该证书包括诸如具有唯一号码的认证者数字签名，PDA 公开密钥等信息，可以由其他参与者查看，任何收到证书的节点都会进行验证，并将其验证结果广播到其他节点，证书将通过授权的股份记录在全球分类帐中。

图 3 显示了 PDA-PID 时间序列的另一个例子。PID 是具有生物特征数据或其他可证明身份的相关信息，可以是所拥有物，所知信息和行为模式等。在这里，我们使用生物特征值测量来进行说明。由于使用两种或多种类型的生物特征信息进行验证，根据统计学表明可以显著减少碰撞记录，因此要求受试者呈现至少两种不同类型的生物特征测量。最终，分布式帐本通过共识拥有所有可验证的主体信息，因此需要一个计数器来区分不同类型的生物特征信息。

详细的后续操作可以在<sup>[33]</sup>中引用。记住，有许多方法实现图 2 和图 3 中时间序列。

## V. PASS 服务系统

如果客户在 PDA-PA 中需要部分或全部有关主体的信息可以直接与受试者进行谈判，而不是聘请查询者进行验证。客户与主体之间有着明确的可操作的智能合约，客户可以检查所要求的签名文件，如学位，成绩单，工作经验等。

图 4 是互联网社交媒体使用 PASS<sup>[32]</sup>的示例，互联网社交媒体对上传的信息可信度难以进行保障，部分原因是由于社交媒体无法对社交账号背后是否是人还是脚本，此人是否可信等问题进行有效的甄别。因此，对其发布内容的可信度进行评估，大部分是需要基于发布人自身及其相关信息。大多数研究集中在对外源性信息进行甄别，

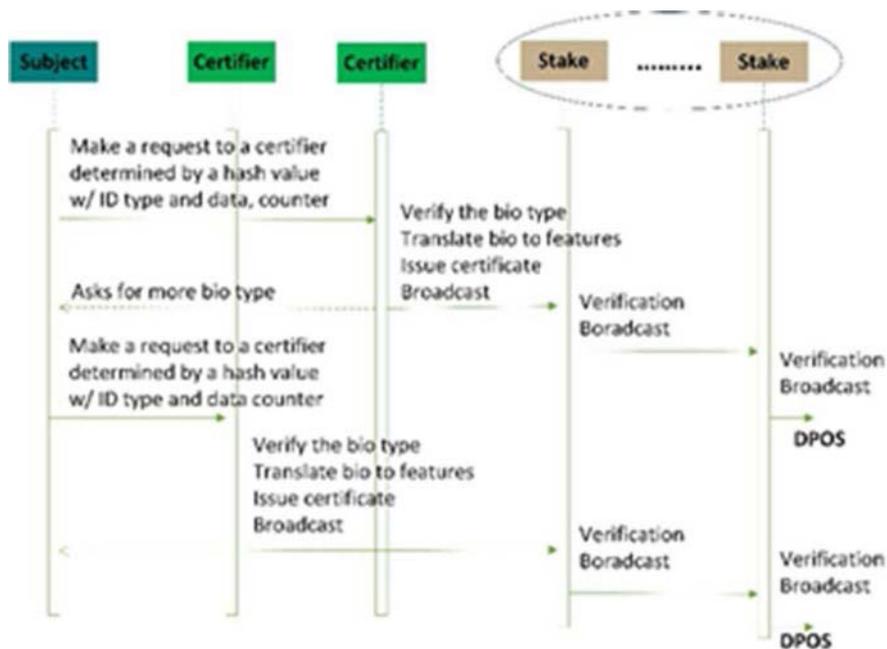


图 3 PDA-PID 的时序

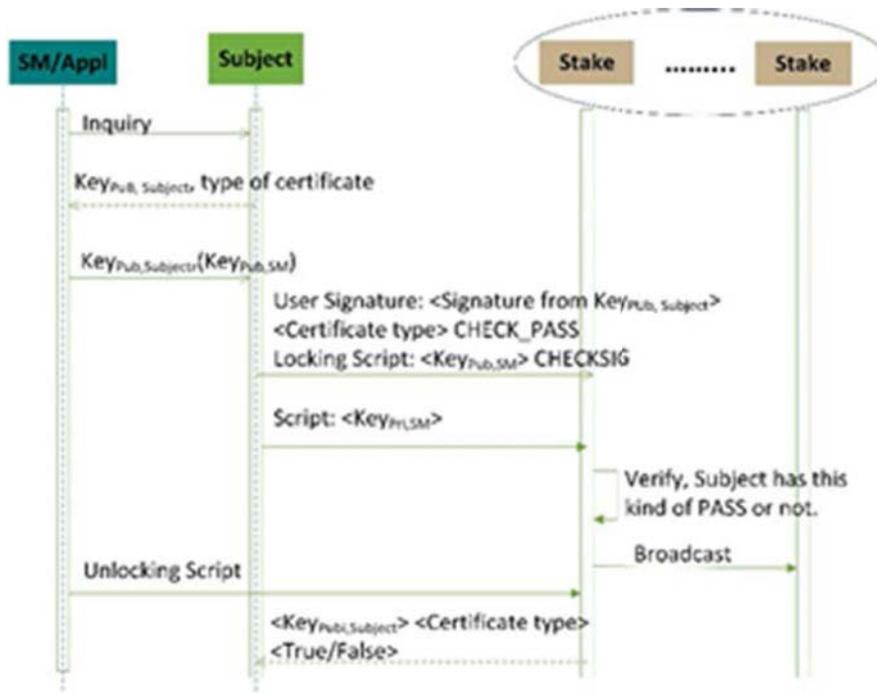


图 4 PASS 服务

如超链接结构<sup>[11-18,36]</sup>，最近的研究方向则是内源性的信息，例如来源提供的发布信息正确性<sup>[36]</sup>。这样的处理导致高质量的发布主要是由排名相对较高的专家所组成，因而普通人发布的就会丢失在信息海洋中。据观察，社交媒体对上传信息质量控制的新做法已经与个人账户相关联。这些例子反映了事物（在 Facebook<sup>[13]</sup>）、分数（在 Karma<sup>[14]</sup>）和声望（一些聊天组由声望来决定在 15 天内发表评论的权限），合并后组成了一种数字化的身份用于个人在线评分系统<sup>[7]</sup>。在登录过程中，一些应用程序使用验证码<sup>[17]</sup>，基于照片的社交认证（Facebook），访问频次限制（twitter，GitHub，Redit 等）和帐户属性判定<sup>[16]</sup>。

通过 PASS，ID 服务可以拥有以下几种特点：a) 真正的人，b) 没有别名的帐户，

c) 匿名和安全，d) 透明，e) 不可篡改的，f) 基于共识的去集中，g) 可撤销。

步骤如下：

1. 诸如互联网社交媒体之类的应用程序向该主体发出要求进行功能验证的请求；
2. 主体将其公钥和相关 PDA 发送给应用程序；
3. 应用程序返回其主体公钥签名的公钥；
4. 主体通过其签名的解锁脚本，签名，PDA 列表及其类型和方法来获取 PDA 的检查；
5. 主体还会发送一个锁定脚本，其中包含应用程序的公钥进行签名检查。仅当签名与应用程序匹配时才会解锁；

6. 联盟 P2P 网络的节点检查主体是否拥有此类 PDA；当它们存在时，它继续下去；
7. 这些节点会广播到其余的节点，所以他们都可以使用它进行验证；
8. 应用程序发送具有签名的解锁脚本；
9. 节点使用解锁脚本传回应用程序关于验证成功或失败。

如我们所见，PASS 的这种验证方式完全可以推广到需要验证 PDA 的任何其他应用程序中。

## VI. 讨论

在本文中，我们已经说明了使用区块链技术进行个人档案管理的方法，PASS 很好地利用来自区块链的特性。每当一个主体想要追加新的成就或新的特征时，这个主体可以很快地达到其目的并进行归档，而不需要等待验证者和查询者。

这种应用的机会是很广泛的，可以用于在线应用程序以及其他所需验证的程序上，如就业和升迁。完全消除了第三方，但依旧保证了其匿名性和问责性。

实行中最大的难点是找到能在区块链网络中积极运作，并且声望良好股权代表，还需要一个机制或议定书来投票支持或反对一个股权代表，股权代表的权利需要有一定的限制。

相关的难点还有征求传统知名组织或团体成为验证者的一员。这些组织需要将原本的人工流程转换成数字化流程，并以数字的形式发行证书。另外还需要制定验证人申请的门槛。

为了能广泛应用，标准化 PDA 结构和通信协议至关重要。我们知道一些 PDA 是十分容易获得认证，然而另一些 PDA 并不是。例如，获得认证的成绩单是一个标准的过程，但难以验证由个人如教授进行的个人项目，推荐信以及长期项目工作也是如此。因此，将这类信息转为可信的数字证书需要更大的验证工作量，我们可能会为某些 PDA 留下空间而无需认证。

最后一个难点似乎是永存的，就是将解锁钥匙放在安全的地方。一旦遗失了钥匙，将没有办法解锁 PDA。

## 参考文献

- [1] <http://infographicsmania.com/resume-padding-statistics/>
- [2] Open badge, <https://openbadges.org/>
- [3] Acclaim <https://www.youracclaim.com>
- [4] Sony Global Education <https://www.sonyged.com/> and <https://www.sony.net/SonyInfo/News/Press/201602/16-0222E/index.html>
- [5] <https://www1.taskstream.com/>
- [6] OpenID Connect, <http://openid.net/connect/>
- [7] A. Yasin and L. Liu, "An Online Identity and Smart Contract Management System", *2016 IEEE 40th Annual Computer Software and Applications Conference (COMPSAC)*, Atlanta, GA, 2016, pp. 192-198.
- [8] X Luna Dong, et al., Knowledge-Based Trust: Estimating the Trustworthiness of Web Sources, 02/2015, <https://arxiv.org/abs/1502.03519>,
- [9] Requirements for a Global Identity Management Service, April 2001, <https://www.w3.org/2001/03/WSWS-popa/paper57>
- [10] W Gitt, R Compton and J Fernandez, Biological Information — What is It?, [http://www.worldscientific.com/doi/pdf/10.1142/9789814508728\\_0001](http://www.worldscientific.com/doi/pdf/10.1142/9789814508728_0001)

- [11] L Page, The PageRank Citation Ranking: Bringing Order to the Web, 1998
- [12] <https://en.wikipedia.org/wiki/PageRank>
- [13] <https://www.facebook.com/help/like>
- [14] <http://lesswrong.com/>
- [15] <http://www.wechat.com/en/>
- [16] X Yang, Q Cao, M Sirivianos, SybilRank: Aiding the Detection of Fake Accounts in Large Scale Social Online Services, [https://users.cs.duke.edu/~qiangcao/sybilrank\\_project/index.html](https://users.cs.duke.edu/~qiangcao/sybilrank_project/index.html)
- [17] L v Ahn, et al, reCAPTCHA: Human-Based Character Recognition via Web Security Measures". Science, September 12, 2008. pp 1465-1468.
- [18] S Yardi, N Feamster and A Bruckman, Photo-based authentication using social networks, Proceedings of the first workshop on Online social networks, WOSN'08 PP 55-60
- [19] OAuth: <https://kantarainitiative.org/about/principles>
- [20] S Nakamoto, Bitcoin: A Peer-to-Peer Electronic Cash System, 2008, <https://bitcoin.org/bitcoin.pdf>
- [21] Hyperledger <https://www.hyperledger.org/>
- [22] R3 CEV Financial Consortium <http://www.r3cev.com/>
- [23] The Enterprise Ethereum Alliance (EEA) <http://entethalliance.org/>
- [24] DAO, decentralized autonomous organization <https://forum.daohub.org/>
- [25] Chien Le, A Survey of Biometrics Security Systems, 2011, <http://www.cse.wustl.edu/~jain/cse571-11/ftp/biomet.pdf>
- [26] L Lai, S Wai Ho and H. Vicent Poor "Privacy Security Trade-Offs in Biometric Security Systems - Part 2: Multi Use Case" IEEE Transactions on Information Forensic and Security, Vol 6, No.1, March 2011
- [27] S B Nikam, S Agarwal, Ridgelet-based fake fingerprint detection, Neurocomputing, Volume 72, Issues 10–12, June 2009, Pages 2491–2506
- [28] LF WU, et al, Non-Invertible Transformation Schemes for Face Template Protection, Signal Processing, Vol 28, No 7, 2012.
- [29] M-d Yu and Srinvas Devadas, Pervasive, Dynamic Authentication of Physical Items, COMMUNICATIONS OF THE ACM, V60, N4, 2017, pp32-39
- [30] K Gai, et al, Privacy-Preserving Adaptive Multi-channel Communications Under Timing Constraints, 2016 IEEE International Conference on Smart Cloud (SmartCloud), 2016, pp 190-195
- [31] Schneier, B. Sensible authentication. ACM Queue 1, 10(2004): 74–78
- [32] Wilson, C. et al, Fingerprint vendor technology evaluation 2003: summary of results and analysis report. NIST Internal Report 7123 (2004).
- [33] Yixuan Zhu and Zhixiong Chen, RealID: Building A Secure Anonmous Yet Transparent Immutable ID Service, 2017 IEEE 3rd International Conference on Big Data Security on Cloud, Beijing, China
- [34] Smart Contracts: The Blockchain Technology That Will Replace Lawyers, <https://blockgeeks.com/guides/smart-contracts/>
- [35] Michael Mainelli, Blockchain Will Help Us Prove Our Identities in a Digital World, Harvard Business Review, March, 2016, <https://hbr.org/2017/03/blockchain-will-help-us-prove-our-identities-in-a-digital-world>
- [36] Xin Luna Dong, et al, Knowledge-Based Trust: Estimating the Trustworthiness of Web Sources, <https://arxiv.org/abs/1502.03519> Submitted on 12 Feb 2015)

## 作者介绍



**陈志雄博士**，现任美国玛希学院（Mercy College）终身教授，数学计算机系系主任，信息安全中心主任。他的研究兴趣主要在网络社交和服务计算里的数据探索和数据保障和安全，

IT 风险分析和审计。他发表了近四十多篇的学术论文，并组织 and 担任了许多 IEEE 学术会议的主席和委员会委员。同时在权威学术期刊和专家委员会里担任评审或编委，也帮助许多会议

进行论文评审。他的研究也得到过美国国家基金的支持。陈志雄博士是信息系统安全专家（CISSP）国际注册师，美国电气和电子工程师协会（IEEE）和国际计算机学会（ACM）的高级会员。陈志雄博士获美国匹兹堡大学的数学博士和计算机科学硕士，上海交通大学数学学士和硕士。



了相关文章。

**朱仪轩**, Yixuan Zhu, 本科毕业于厦门大学嘉庚学院，现为 Mercy College 中 Cybersecurity 专业下的研究生。目前研究方向是区块链技术，曾在 IEEE IDS 2017 和 IEEE AIMS 2017 中发表了相关文章。

# Sustainable Computing and Computing for Sustainability

## 智能计算与可持续能源发展

Fanxin Kong, Xue Liu

孔繁鑫, 刘学

McGill University, Canada

加拿大麦吉尔大学

fanxin.kong@mail.mcgill.ca, xueliu@cs.mcgill.ca

### Abstract:

The growing awareness about global climate change has boosted the need to reduce greenhouse gas emissions from legacy power systems. To improve their sustainability requires much effort on management and control regarding the power demand side. This article focuses on investigating two different customers with large power loads: data centers and electric vehicles, both demanding extremely high energy consumption. And thus, their management is crucial for power systems' sustainability.

The first thrust of this article is to make computing systems more sustainable by reducing their energy cost and emissions. We look at one type of large-scale computing system, cloud data centers, as an example to address this thrust. The computing capacity and scale of data centers are increasing to meet the soaring demand for cloud applications and services. A mega data center (such as those of Google and Apple) can host thousands of

servers and require up to tens of megawatts of electricity. The high power consumption causes large energy cost for cloud providers and negative impacts to the environment. Thus, an energy-efficient, low-cost, and environment-friendly data center operation is key to the success of their applications and services.

The second thrust of this article is to leverage computing techniques to improve the sustainability and efficiency of energy systems. The second thrust of the article lies in computing for sustainability, which is to leverage computing techniques to improve the sustainability and efficiency of energy systems. The anticipated deep penetration of electric vehicles (EVs) will significantly benefit the environment, but will have a heavy load impact on the power grid, and thus jeopardizes its stability and reliability. To address this potential consequence, the article makes progressive efforts on controlling and managing EVs' charging demand to support large-scale EV deployment.

## 摘要

人们对全球气候变化的意识日益增强，对减少传统电力系统温室气体排放的需求也随之不断增长。为了提高电力系统的可持续性，需要对电力需求端的管理和控制做很多努力和创新。本论文重点介绍两种典型的电力负载：数据中心（Data Centers）和电动汽车（Electric Vehicles）。这二者的电力需求极高，因此他们的管理和控制对电力系统的可持续性至关重要。

第一个要旨是可持续计算技术（Sustainable Computing）：通过提高能效以及降低温室气体排放来改善计算机和信息技术系统的可持续性。对于该要旨，本文关注大规模数据中心（Data Centers）。为了满足云计算应用和服务需求的不断高涨，兆瓦级大规模数据中心数量持续增加。数据中心对电能的巨大需求不但给云计算提供者带来巨额的能源成本而且对环境也造成了很大的影响。因此，高能效、低成本及环境友好的数据中心至关重要。

第二个要旨是面向可持续性的计算技术（computing for sustainability）：利用计算机和信息技术来提高物理系统的可持续性。对于该要旨，本文关注电动汽车。电动汽车的节能环保等良好特性很好地满足了可持续发展这个目标。但是，大量电动汽车充电会带来很重的电力负荷，从而影响电力系统的稳定性和可靠性。为了解决这个潜在的后果，并且更好的支持电动汽车的大量部署，控制和管理电动汽车的充电至关重要。

## 1 引言

现有的电力系统严重依赖于碳密集型能源，产生大量的二氧化碳等温室气体，因此严重的影响环境。例如，在 2013 年，加拿大电力部门的温室气体排放量占有所有能源相关的排放的 14.3%<sup>[1]</sup>。对于美国，这个数字更大，大约是 40%<sup>[2]</sup>。这些气体排放是造成全球气候变暖的主要原因，对提高传统电力系统的可持续性的需求也随之变得严重。为了迎合该需求，需要对电力提供端（例如，建设更多的可再生资源发电站<sup>[3]</sup>）和电力需求端（例如，控制电力使用者的负载）都做出很多努力和创新。

本论文着重于电力需求端并研究两种不同类型的电力负载：数据中心（Data Centers）和电动汽车（Electric Vehicles）。研究二者的主要动机是他们的极高的能量需求（因而产生高排放）以及他们该需求的快速高涨。在数据中心方面，全球数据中心的用电量占在 2012 年，全球的数据中心消耗了 3400 亿千瓦时的电能<sup>[4]</sup>。这些电能足够英国这个国家使用一年。在 2013 年，美国数据中心使用了 910 亿千瓦时的电能，占全美国用电量的 1.9%，能耗成本高达 130 亿美元。据估计到 2020 年，数据中心电能需求会增加 50%。在电动车方面，数量大和充电功率高是造成其高电力需求的两个主要原因。一方面，据市场研究报告<sup>[5]</sup>，照比 2010 年，2015 年美国电动车的年销售量增加了 6 倍。到 2024 年美国每年的电动车销量将达到 240 万辆。到 2031 年，美国电动车在所有汽车市场份额中会占到 28%。到 2022 年，全球电动车数量将超过 3500 万辆。另一方面，一个快速电动车充电站的功耗可高达与整栋大型办公楼的功耗相等、与北美 200 户的家庭用电总和相等<sup>[6]</sup>。

## 2 数据中心背景

数据中心是商业，通信，学术和政府等系统正常运作的重要组成部分。较小的公司机构拥有内部数据中心，通常建在大型商业建筑物内。更大的公司通常在（可能是远程的）独立建筑基础设施上建立他们的数据中心，这些基础设施的面积可以达到几十万平方英尺，并且专用于托管计算外围设备。此外，较大的公司可能包含几十个甚至数百个数据中心。小型数据中心功耗可能是几十千瓦，而大型云计算数据中心的峰值功率可高达几十兆瓦。

### 2.1 数据中心主要组成部分

数据中心是一整套复杂的设施，包括三个主要组成部分：信息技术设备，电力输送系统，以及冷却系统<sup>[7][8]</sup>，如图 1 所示。

(a) 信息技术设备。在典型的数据中心中存在三种电子设备：(i) 用于数据处理的服务器，(ii) 用于数据存储的存储设备，以及，(iii) 用于数据通信的网络设备。

这三种类型的设备统称为信息技术设备。

(b) 电力输送系统。电力输送系统主要包括电力转换（交流到直流和直流到交流）单元，稳压器和备用设备。电源备份通常由不间断电源单元提供，防止信息技术设备遭受电源中断和可能的严重业务中断或数据丢失。

(c) 冷却系统。数据中心的冷却通常由计算机室空调单元提供。该系统排出数据中心中的热量并控制其温度和湿度。

### 2.2 数据中心功耗分解

图 1 还示出了数据中心的典型功率流。从主电网提供的电力首先传送到不间断电源。不间断电源将交流电转换为直流电以对其电池充电。来自电池的直流电进一步再转换回交流电，并通过配电单元提供给信息技术设备。一旦电源到达服务器，它将再次转换为直流电源，并提供给服务器的中央处理单元，内存，磁盘驱动器，芯片组，风扇和其他内部组件。表 1，摘自

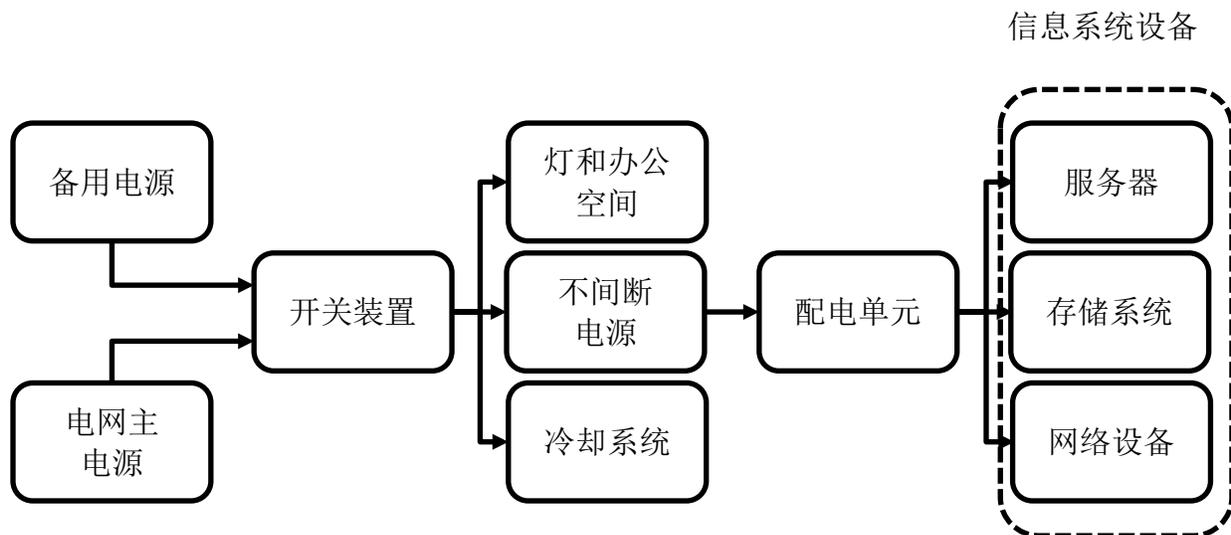


图 1. 数据中心主要组成部分

参看文献<sup>[9]</sup>，清晰地显示不同服务器组件的功耗。

表 1. 数据中心内部功耗分解

组成部分	功耗（瓦）	百分比
处理器	80	31.87%
存储	36	14.34%
磁盘	12	4.78%
外设	50	19.92%
主板	25	9.96%
风扇	10	3.98%
供电单元损失	38	15.13%
<b>总和</b>	<b>251</b>	<b>100%</b>

最近对 20 多个数据中心的电力使用进行的一项调查发现，进入数据中心的电力只有大约一半由信息技术设备使用，剩余的一半被电力输送系统和冷却系统消耗<sup>[10]</sup>。因此，传统方法主要是降低上述三个组成部分的能量消耗或是电能损失。本文从另一个角度出发，调研如何增加在数据中心中清洁能源的使用，从而降低温室气体排放。可以容易看出，该类研究与传统方法是互补的。

### 3 可持续数据中心

现代数据中心，例如 Google 和 Microsoft 运营的，已经开始采取各种措施降低运营成本和对环境的影响。他们通过减少碳排放和使用可再生能源运行变得更加环保。早期关于数据中心能耗管理的研

究工作主要集中在降低单个数据中心内的信息技术设备的功耗。他们提出了不同的硬件和软件技术，包括芯片多处理器<sup>[11]</sup>，动态电压和频率缩放技术<sup>[12]</sup>，动态功耗管理技术<sup>[13]</sup>，虚拟化视化<sup>[14]</sup>和负载平衡<sup>[15]</sup>等。

最近，许多研究工作涉及数据中心的绿色能源电源管理。它们不是降低功耗，而是专注于降低能源成本，减少排放，以及优化可再生能源利用。它们考虑了几个新的维度，包括动态定价和电网电力的时间和空间变化排放因子，使用能量存储装置缓冲能量，以及可再生能源整合到数据中心电源中等。本文重点调研该类工作。

#### 3.1 数据中心能源介绍

电力基础设施通过可以集成现场发电，电网和能量存储设备的微电网产生和分配用于信息技术设备和冷却设备的电力。传统上，储能设备用作电网和备用电源（例如柴油发电机）之间的故障切换。在电网断电时，储能设备使用存储在其中的能量给数据中心供电，直到备用电源可以启动并匹配数据中心的电力负载。近来，储能设备也被用作能量缓冲：在可再生电力充足时对它们充电和在可再生电力不足时对其进行放电，用来平滑不可控的电力（例如，风能或太阳能）。

大多数非绿色能源（例如化石燃料动力）是可调度和可控制的，也就是说，它们能够通过调节发电机来按需调整它们的输出功率，但是它们伴随着大量的温室气体排放。他们一般主要作为电源或备用电源，因为只要提供燃料源，它们就会连续发电。相比之下，绿色或可再生能源（如

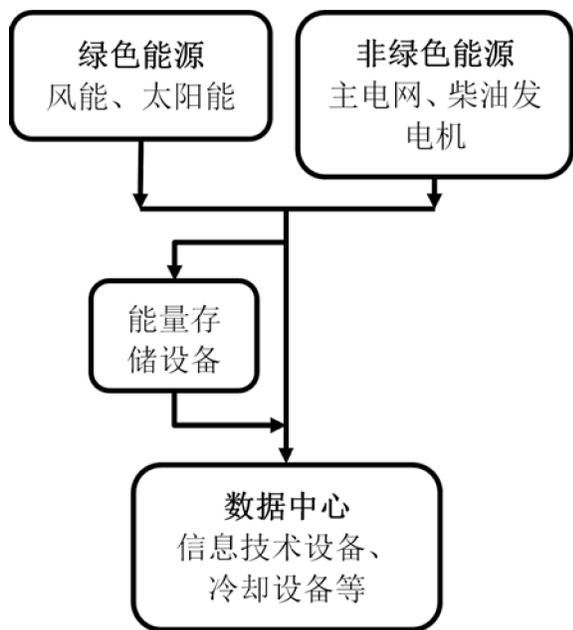


图 2. 数据中心电源设施框图

风能和太阳能) 不产生温室气体排放, 但是它们是间歇性的和不可控的。它们一般不被用作主电源, 因为它们严重依赖天气条件。

电网电力是绿色和非绿色能源的混合, 通常被算作是非绿色能源。电网具有与现场发电能源不同的特征。在正常操作期间的电网是可靠且灵活的电源, 数据中心可以从其中获取电力以填充数据中心的电力需求和现场发电之间的任何差距。这种可靠性和灵活性允许将非可控能源集成到电力基础设施中, 并为数据中心保证稳定的电力供应。当电网经历中断, 数据中心的储能设备桥接从电网到后备电源的过渡。

### 3.2 研究现状

本文重点对绿色能源敏感的现存工作进行调研, 即明确考虑可再生能源和温室气体排放。本文将这些现存工作分为以下三类。

(a) 绿色能源敏感的工作负载调度。数据中心通常支持多种类型的用户请求或工作负载, 从关键互动作业(如互联网服务)到延迟容忍的批处理作业(如科学应用)。对服务的请求可以路由到托管该服务的任何数据中心。这实现了分布式数据中心之间的地理负载平衡。批处理作业可以安排在任何时间执行, 如果他们在截止日期之前完成。这使得工作负载管理具有时间灵活性。该类研究采用作业调度技术并利用上述时间或空间灵活性进行能耗管理。典型工作包括<sup>[16][17]</sup>。

(b) 绿色能源敏感的虚拟机管理。虚拟机是像物理机一样执行工作负载的机器的软件实现。服务器可以运行多个虚拟机。每个虚拟机可以保存多个应用, 并且每个应用也可以使用多个虚拟机。虚拟机可以在不同的服务器之间迁移: 首先暂停在服务器中运行的虚拟机, 然后将其复制并传输到另一个服务器, 在该服务器上虚拟机恢复执行。假设有两个物理服务器, 每个有两个虚拟机。如果每个服务器有一个空闲虚拟机或没有正在运行的工作负载, 则可以将一个服务器中使用的虚拟机迁移到其他服务器。然后可以关闭空闲服务器以节省能量。此类工作考虑了改虚拟化环境中的数据中心的, 并利用类似上面讨论的时间和空间灵活性和多样性来管理虚拟机迁移。典型工作包括<sup>[18][19]</sup>。

(c) 绿色能源敏感的能源规划。该类工作集中讨论数据中心的建设阶段, 即设计电源管理计划, 以匹配电源与数据中心的电力需求。该计划为峰值电网功耗和电网能量使用, 不同现场/非现场能源及其使

用的能力和/或能量存储设备及其缓冲操作的能力进行选择。典型工作包括<sup>[20][21]</sup>。

### 3.3 待研究问题

尽管现今有一些现存工作对整合绿色能源的可持续性数据中心进行研究，但是还都处于起步阶段，仍有大量的问题亟待解决。以下列出三个重要问题。

(a) 精确建模问题中的各个纬度。所有的维度，例如绿色能源，数据中心电力负载，和储能设备等，对数据中心可持续性优化都有很大影响。任何维度的缺乏或不准确的建模将使得解决方案偏离最优或实际情况。例如，如果要最大限度地利用可再生能源，过度估计可再生能源发电可能会导致执行更多的工作量，并使用比实际发电更多的电力。因此，将使用更多的电网电力来填充间隙，这导致更多的排放。

(b) 联合优化时空多样性和灵活性。空间多样性意味着不同地区的数据中心具有不同的参数，例如不同数量的可再生能源和不同的电网排放因子。时间多样性意味着这些参数随时间变化。空间灵活性意味着用户请求或计算负载可以路由到不同的数据中心并在不同的数据中心处理。时间灵活性意味着延迟忍受的负载的执行可以被推迟。联合考虑这些因素会更好的提高数据中心的可持续性。

(c) 计算难度低并易于实现到现有应用中。这两个特性决定了问题的解决方法是否可以应用于实际应用。

## 4 电动汽车背景

尽管电动车早在 19 世纪中期就已经出现，但是内燃机已经是用于机动车近 100 年的主导推进方法。直到最近几十年，基于石油的运输基础设施对环境的影响，以及对石油峰值的担忧，导致了人们对电力运输基础设施再次感兴趣。内燃机驱动的传统汽车通常只能从单个或几个来源(通常是不可再生的化石燃料)获得能量。因此，会造成大量的温室气体排放。插入式充电电动车不同于传统化石燃料动力汽车，因为它们消耗的电力可以从各种来源产生，包括绿色能源和非绿色能源。电能可以通过电池存储在车辆上。电动车温室气体排放量取决于用于发电的燃料和技术。这也是插入式充电电动车的一个关键优点，即具有大大降低温室气体排放的潜力。

### 4.1 插入式充电电动车

插入式充电电动车可以从任何外部电源(例如主电网或现场发电机)充电的机动车辆，并且存储在车载可再充电电池中。该类电动车主要可分为两类：电池电动车和插电式混合动力电动车<sup>[22]</sup>。

(a) 电池电动车。该类电动车又称作纯电动车或全电动车。全电动车使用电动机和电动机控制器代替内燃机用于推进。它们从车载电池组获得所有功率，因此没有内燃机，燃料电池或燃料箱。典型例子包括 Nissan Leaf 和 Tesla Model S。

(b) 插电式混合动力电动车。该类电动车是使用可再充电电池的混合动力机动车辆，可通过插入到外部电源对其进行充电。该类电动车具有电动机和内燃机，共享常规混合动力机动车辆和具有连接到电

网充电的全电动汽车的特性。典型例子包括 Chevrolet Volt 和 Toyota Prius。

## 4.2 电动车充电负载特点

对于电动汽车的充电负载模式，需要考虑其充电特性的两个方面。第一，随着电动汽车跨越街道，地区甚至城市，其车载电能也随之移动。因此，存储在电池中的电能继承了电动车的空间移动性。电动车充电负载空间特性与真实世界的驾驶条件相关，包括路段的物理条件，交通拥堵状态和驾驶模式等。其次，由于电动车可以在一天中随机地启动或停止其充电过程，这呈现了充电负载在时间上的随机性。该时间特性不仅包括了充电过程的开始和停止时间，并且能量的使用随时间变化。这两个特性对智能电网规划和充电服务调度至关重要。

## 5 智能电动汽车充电

电动汽车的流行将给电网带来挑战性的问题。例如，在 80A 和 240V 下以 19.2kW 的充电功率对电动车进行充电的负载（称为交流 2 级充电标准）可以支持一个典型的北美家庭用电需求。当充电负载被聚集但未协调时，充电负载的影响将更为严重。这可能导致很多问题，包括电动汽车充电需求与电网电源之间的不平衡，更多的功率损耗和更大的电压偏差等。

为了实现智能电网和电动汽车的智能集成，基础设施需要全面升级。例如，当前的电网基础设施和电动车处于处理双向功率流的不成熟阶段；通信基础设施正在初级阶段。因此，随着基础设施的发展，基于电动汽车和智能电网之间的实时通信，

智能电网的协调流量控制和电动汽车充电的智能控制，可以实现双方的智能互动。

## 5.1 研究现状

本文重点调研电网和电动汽车之间的智能互动。电动汽车和电网之间的智能交互仍处于起步阶段，需要更多的大量的研究工作，提高智能网络的可靠性和安全性，电动汽车用户的自身利益等等。本段将概述该领域的最新进展，并将这些现存工作分为以下三类<sup>[23]</sup>。

(a) 面向智能电网的智能充电。如果电动车和电网之间的交互不受控制或不协调，电网将发生很多严重问题，例如，电力供求之间的不平衡，电压波动，功率损耗，能量效率降低，频率偏差等问题。然而，如果智能地控制和协调二者的交互，则可有效的保证电网的稳定性和可靠性。例如，“填谷”策略可以通过将大功率需求转移到总体负载曲线的波谷来减轻供需之间的不平衡。此外，该策略也可以缓解变压器损耗（变压器温度升高产生改损耗），而维护变压器套。这些潜在好处的实现归功于电动车充电的在时间和功率上的灵活性。调研的工作包括利用电动车充电来实现下面三个功能。（i）平整充电负载，典型工作包括<sup>[24][25]</sup>。（ii）提供频率调节服务，典型工作包括<sup>[26][27]</sup>。（iii）电压调节服务，典型工作包括<sup>[28][29]</sup>。

(b) 面向聚合器的智能充电。在很多应用中，存在用于更新和向电动车发布中间信息的集中式控制器，也称作聚合器。聚合器可以被认为是智能电网的代表。全局控制可以分解为多个本地聚合器，其管

理其局部区域内的负载。这些本地聚合器可以被认为是当地公共设施的代 表，包括变电站，充电站，停车场和住宅区等。聚合器桥接电网和电动车用户之间的电力流。更具体地，聚合器可以负责维持稳定和可靠的电力系统，同时负责满足用户的充电需求。所有三个相关方之间的智能交互有助于使未来的电力系统变得聪明。典型工作包括：直接协调控制<sup>[30][31]</sup>以及间接协调控制<sup>[32][33]</sup>。

(c) 在循来自电网或聚合器的控制的需要可能会损害电动车用户的利益，即盲目地遵循电网或聚合器的要求可能会导致对部分电动车用户不利。例如，在夜间充电的情况下，尽管用户可以使电动车完全充电，但是当充电服务价格实时变化时，可以导致用户支付更多电费。该类研究从电动车用户角度出发，智能充电管理必须尊重用户的自由，例如最小化充电成本。为了实现这一目标，用户需要包括实时服务价格，调节信号和调节价格等信息。问题是这些信息有时难以预测时。用户将面对在指定时间范围内最小化预期充电成本的情况。该类典型工作包括<sup>[34][35]</sup>。

## 5.2 待研究问题

尽管已有大量关于智能电动车充电的研究工作，下述问题仍没有得到很好地解决。

(a) 电池特性带来的充电问题。很多研究工作是基于假设电池电量是随着充电时机线性变化地，在给定恒定充电功率时。然而，实际上电池电量和充电时间之间的关系实际上是非线性的。由于具有较高电

量的电池的充电速率比具有较低电量的电池的充电速率慢，因此使用线性模型可能导致电动车用户充电成本和充电时间的增加。当用户需要到达多个中间停靠点并且多次充电时，这个问题可能更为严重。考虑正式电池充电特性的电动车充电问题会变得很具有挑战性。

(b) 电动车充电模式估计。大部分现存研究工作都假设电动汽车的充电模式（包括何时和如何充电电动车辆）已知为先验，或选择分布模型（例如高斯分布）。然而，在实际中，该假设往往不成立。聚合器或电网对电动汽车充电模式有效和准确的估计或预测是至关重要。潜在的解决方案可以结合统计分析和实时反馈，这可以通过电动车和智能电网之间的实时移动数据交换来实现。

(c) 开发更智能的电动车调度。从用户角度来看，电动车的智能集成不仅包括个人电动汽车充电的智能调度，还包括一个来自工业和学术界的共同关注：里程焦虑。电动车智能路由被定义为用于寻找最低能量消耗或最小时间的路径，或满足这两个目标之间的权衡的路径。对于混合能源电动车，很少有研究寻讨论上述问题。

## 6 结论

传统电力系统的温室气体排放是造成全球气候变化一个重要原因。为了提高其可持续性，需要在电力需求端做出大量的工作。本文包含两个要旨。第一个是可持续计算技术：通过提高能耗效率以及降低温室气体排放来改善计算机和信息技术系统的可持续性。对于该要旨，本文调研了

可持续性数据中心，介绍现存工作并指出了一些尚未解决的问题。第二个要旨是面向可持续性的计算技术：利用计算机和信息技术来提高物理系统的可持续性。对于该要旨，本文调研了智能电动车充电，简要总结了现存工作并指出了一些亟待解决的问题。

## 参考文献

- [1] Canadian Electricity Association, “Canada’s electricity industry.” <http://www.electricity.ca/media/Electricity101/Electricity101.pdf>.
- [2] EIA, “Annual energy outlook 2013”, U.S. Energy Information Administration, 2013.
- [3] F. Kong, C. Dong, X. Liu, and H. Zeng, “Quantity versus quality: Optimal harvesting wind power for the smart grid”, Proceedings of the IEEE, 2014.
- [4] ComputerWeekly, “Global census shows datacentre power demand grew 63% in 2012”, <http://www.computerweekly.com/>.
- [5] L. Liu, F. Kong, X. Liu, Y. Peng, and Q. Wang, “A review on electric vehicles interacting with renewable energy in smart grid”, Renewable and Sustainable Energy Reviews 2015.
- [6] Wikipedia, “Charging station”. [https://en.wikipedia.org/wiki/Charging\\_station](https://en.wikipedia.org/wiki/Charging_station).
- [7] F. Kong and X. Liu, “A survey on green-energy-aware power management for data-centers”, ACM Computing Surveys, 2014.
- [8] A. Rahman, X. Liu, and F. Kong, “A survey on geographic load balancing based data center power management in the smart grid environment”, IEEE Communications Surveys & Tutorials, 2014.
- [9] X. Fan, W. D. Weber, and L. A. Barroso. Power provisioning for a warehouse-sized computer, ACM SIGARCH Computer Architecture News, 2007.
- [10] S. Greenberg, E. Mills, B. Tschudi, L. Berkeley, et al. Best practices for data centers: Lessons learned from benchmarking 22 data centers, Proceedings of the ACEEE Summer Study on Energy Efficiency in Buildings in Asilomar, 2006.
- [11] L. A. Barroso, “The price of performance”, Queue, 2005.
- [12] T. Horvath, T. Abdelzaher, K. Skadron, and X. Liu, “Dynamic voltage scaling in multitier web servers with end-to-end delay control”, IEEE Transactions on Computers, 2007.
- [13] D. Meisner, B. T. Gold, and T.F. Wenisch, “PowerNap: eliminating server idle power”, International Conference on Architectural Support for Programming Languages and Operating Systems, 2009.
- [14] R. Nathuji and K. Schwan, “VirtualPower: coordinated power management in virtualized enterprise Systems”, ACM Symposium on Operating Systems Principles, 2007.
- [15] G. Chen, W. He, J. Liu, et al., “Energy-Aware Server Provisioning and Load Dispatching for Connection-Intensive Internet Services”, USENIX Symposium on Networked Systems Design and Implementation, 2008.
- [16] I. Goiri, W. Katsak, K. Le, et al., “Parasol and GreenSwitch: managing datacenters powered by renewable energy”, International Conference on Architectural Support for Programming Languages and Operating Systems, 2013.
- [17] Z. Liu, M. Lin, A. Wierman, et al., “Geographical load balancing with renewables”, Greenmetrics, 2011.
- [18] C. Li, A. Qouneh, and T. Li, “iswitch: Coordinating and optimizing renewable energy powered server clusters”, International Symposium on Computer Architecture, 2012.
- [19] S. Akoush, R. Sohan, A. Rice, et al., “Free lunch: exploiting renewable energy for computing”, HotOS, 2011.
- [20] C. Ren, D. Wang, B. Urgaonkar, and A. Sivasubramaniam, “Carbon-Aware Energy Capacity Planning for Datacenters”, International Symposium on Modeling,

- Analysis, and Simulation of Computer and Telecommunication Systems, 2012.
- [21] F. Kong and X. Liu, "Greenplanning: Optimal energy source selection and capacity planning for green datacenters", ACM/IEEE 7th International Conference on Cyber-Physical Systems, 2016.
- [22] L. Liu, F. Kong, X. Liu, et al., "A review on electric vehicles interacting with renewable energy in smart grid", Renewable and Sustainable Energy Reviews, 2015.
- [23] Q. Wang, X. Liu, J. Du, and F. Kong, "Smart Charging for Electric Vehicles: A Survey From the Algorithmic Perspective", IEEE Communications Surveys & Tutorials, 2016.
- [24] L. Gan, U. Topcu, and S. Low, "Optimal decentralized protocol for electric vehicle charging", Proc. IEEE Conf. Decision Control Eur. Control Conf., 2011.
- [25] Y. He, B. Venkatesh, and L. Guan, "Optimal scheduling for charging and discharging of electric vehicles", IEEE Trans. Smart Grid, 2012.
- [26] S. Han, S. Han, and K. Sezaki, "Development of an optimal vehicle-to-grid aggregator for frequency regulation", IEEE Trans. Smart Grid, 2010.
- [27] C. Wu, H. Mohsenian-Rad, and J. Huang, "Vehicle-to-aggregator interaction game", IEEE Trans. Smart Grid, 2012.
- [28] Y. Mitsukuri, R. Hara, H. Kita, E. Kamiya, N. Hiraiwa, and E. Kogure, "Voltage regulation in distribution system utilizing electric vehicles and communication", Proc. IEEE PES Transmiss. Distrib. Conf. Expo., 2012.
- [29] H. Wu, H. Mohsenian-Rad, and J. Huang, "PEVbased P-Q control in line distribution networks with high requirement for reactive power compensation," Proc. IEEE PES Innov. Smart Grid Technol. Conf., 2014.
- [30] F. Kong, X. Liu, Z. Sun, and Q. Wang, "Smart Rate Control and Demand Balancing for Electric Vehicle Charging", ACM/IEEE 7th International Conference on Cyber-Physical Systems, 2016.
- [31] F. Kong, and X. Liu, "On-Line Event-Driven Scheduling for Electric Vehicle Charging via Park-and-Charging", IEEE Real-Time Systems Symposium, 2016.
- [32] W. Tushar, W. Saad, H. Poor, and D. Smith, "Economics of electric vehicle charging: A game theoretic approach," IEEE Trans. Smart Grid, 2012.
- [33] F. Kong and X. Liu, "Distributed Deadline and Renewable Aware Electric Vehicle Demand Response in the Smart Grid", IEEE Real-Time Systems Symposium, 2015.
- [34] N. Rotering and M. Ilic, "Optimal charge control of plug-in hybrid electric vehicles in deregulated electricity markets", IEEE Trans. Power Syst., 2011.
- [35] J. Donadee and M. Ilic, "Stochastic optimization of grid to vehicle frequency regulation capacity bids", IEEE Trans. Smart Grid, 2014.

### 作者简介



**孔繁鑫**，先后在中国东北大学获得工学学士和工学硕士学位，在麦吉尔大学（McGill University）获得博士学位。主要研究领域包括 cyber-physical systems and Internet of Things, cloud computing and data centers, and smart grid and energy systems。在相关顶级国际会议和期刊上发表论文多篇。



**刘学**博士，清华大学应用数学系学士，清华大学自动化系硕士，美国伊利诺伊大学厄巴纳-香槟分校计算机科学系博士。现任麦吉尔（McGill University）大学正教授和特聘讲席教授。在此之前，他还在美国内布拉斯加-林肯大学（University of Nebraska-Lincoln）任特聘讲席教授和在美国惠普研究院（HP Labs in Palo Alto）任访问科学家。他曾获得多项奖项包括加拿大杰出青年计算机科学家奖。他在国际顶级学术会议和期刊上发表过 200 余篇论文，获得了多项 IEEE/ACM 的

最佳论文奖和多项国际专利。他担任 5 个主要 IEEE/ACM 汇刊的编委，并担任过 20 多个著名国际会议的主席或组委会成员。他的主要研究方向是计算机系统和网络，实时和嵌入式系统，分布式计算和大数据处理。他的多项研究成果被媒体广泛报道，包括纽约时报（New York

Times），加拿大广播公司（CBC），国际数据集团（IDG），赫芬顿邮报（Huffington Post），新科学家（New Scientist），麻省理工学院科技评论（MIT Technology Review），商业内幕（Business Insider）等。

# Connectionism and Symbolism in Artificial Intelligence

## 人工智能中的联结主义和符号主义

Xianfeng Gu

顾险峰

Computer Science Department, Stony Brook University, Stony Brook NY 11794

纽约州立大学石溪分校计算机系

### Abstract

In 2016 march, AlphaGo defeated Lee Sedol 9p in Google DeepMind Challenge Match, which shocked the whole world. Some people cheer for the development of artificial intelligence; some people worry about the future of human kind; some people contempt the brute force of machines; some people feel depressed for the loss of superiority of mankind. This article discusses the future directions of the development of artificial

intelligence, especially the limitations of the current approach and the potential for future; compares the cognition of human brain neural system and artificial neuron networks, in order to obtain a fair, objective observation and understanding of recent developments in AI.

### 摘要

2016年3月，谷歌 AlphaGo 击败了围棋九段李世石，举世震惊。有为人工智能的发展欢呼雀跃者，有为人类前途命运忧



图 1 由人工神经网络生成的 Inceptionist 图像

心忡忡者，有对机器蛮力不屑一顾者，有对人类失去优越感而沮丧彷徨者。本文对人工智能发展的主要方向，特别是目前的局限和未来的潜力进行探讨，并将人类脑神经认知和人工神经网络认知进行对比，从而对人工智能有一个公正客观，而又与时俱进的认识。

**Keywords:** Artificial Intelligence, Connectionism, Symbolism

**关键词:** 人工智能；联结主义；符号主义

人类的智能主要包括归纳总结和逻辑演绎，对应着人工智能中的联结主义（如人工神经网络，图 1 为由人工神经网络生成的 Inceptionist 图像）和符号主义（如吴文俊方法）。人类大量的视觉听觉信号的感知处理都是下意识的，基于大脑皮层神经网络的学习方法；大量的数学推导，定理证明是有强烈主观意识的，是基于公理系

统的符号演算方法。

## 1 联结主义

David Hunter Hubel 和 Torsen Wiesel 共同获得了 1981 年度诺贝尔生理学或医学奖。1959 年，Hubel 和 Wiesel 在麻醉的猫的视觉中枢上插入微电极，然后在猫的眼前投影各种简单模式，观察猫的视觉神经元的反应。他们发现，猫的视觉中枢中有些神经元对于某种方向的直线敏感，另外一些神经元对于另外一种方向的直线敏感；某些初等的神经元对于简单模式敏感，另外一些高级的神经元对于复杂模式敏感，并且其敏感度和复杂模式的位置与定向无关。证明了视觉中枢系统具有由简单模式构成复杂模式的功能，也启发了计算机科学家发明了人工神经网络。

### 1.1 层次特征（hierarchical features）

后来，通过对猴子的视觉中枢的解剖，将猴子的大脑皮层曲面平展在手术台表面



图 2 三维曲面到平面的保角映射

上，人们发现从视网膜到第一级视觉中枢的大脑皮层曲面的映射（retinotopic mapping）是保角映射（conformal mapping）<sup>[1]</sup>。如图 2 所示，保角变换的最大特点是局部保持形状，但是忽略面积大小，这说明视觉处理对于局部形状非常敏感。

人们逐步发现，人类具有多个视觉中枢，并且这些视觉中枢是阶梯级联，具有层次结构。人类的视觉计算是一个非常复杂的过程。如图 3 所示，在大脑皮层上有多个视觉功能区域（v1 至 v5 等），低级区域的输出成为高级区域的输入。低级区域识别图像中像素级别的局部的特征，例如边缘折角结构，高级区域将低级特征组合成全局特征，形成复杂的模式，模式的抽象程度逐渐提高，直至语义级别。

如图 4 所示，毕加索的名画格尔尼卡（Guernica）中充满了抽象的牛头马面，痛苦嚎哭的人脸，扭曲破碎的肢体。从图 4 中，可以毫不费力地辨认出这些夸张的几何形体。其实，图 4 中大量信息丢失，但是提供了足够的整体模式。由此可见，视觉高级中枢忽略色彩、纹理、光照等局部细节，侧重整体模式匹配和上下文关系，并可以主动补充大量缺失信息。

近年来，深度学习技术的发展，使得人们能够模拟视觉中枢的层级结构，考察每一级神经网络形成的概念。图 5 显示了一个用于人脸识别的人工神经网络经过训练后得到的各层特征。底层网络总结出各种边缘结构，中层网络归纳出眼睛、鼻子、嘴巴等局部特征，高层网络将局部特征组

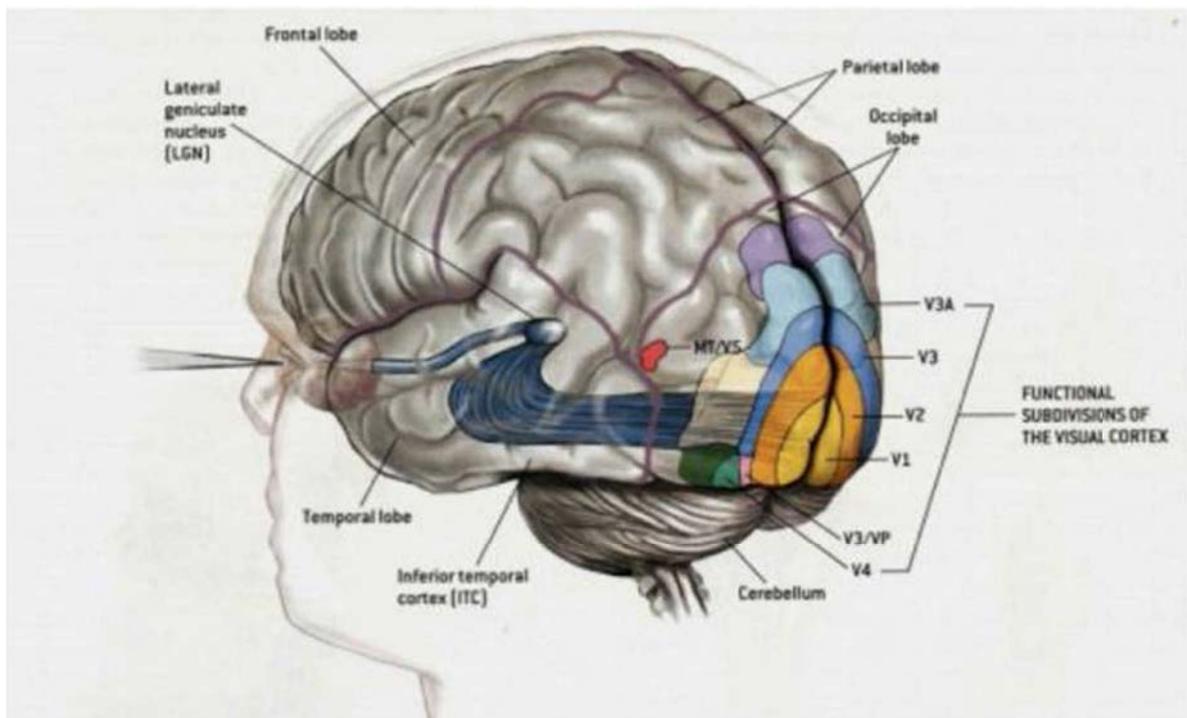


图 3 大脑皮层的视觉中枢，视觉信号的传导途径：视网膜，LGN，V1，V2，V3，V4，V5 等



图4 毕加索的 Guernica

合，得到各种人脸特征。这样，人工神经网络佐证了视觉中枢的层次特征结构。

## 1.2 专用和通用 (specific vs. general)

人工神经网络在 20 世纪 80 年代末和 90 年代初达到巅峰，随后迅速衰落，其中一个重要原因是深度神经网络的发展严重受挫。人们发现，如果网络的层数加深，那么最终网络的输出结果对于初始几层的参数影响微乎其微，整个网络的训练过程无法保证收敛。

同时，人们发现大脑具有不同的功能区域（图 6），每个区域专门负责同一类的

任务，例如视觉图像识别、语音信号处理和文字处理等。并且在不同的个体上，这些功能中枢在大脑皮层上的位置大致相同。在这一阶段，计算机科学家为不同的任务发展出不同的算法。例如，为了语音识别，人们发展了隐马尔科夫链模型；为了人脸识别，发展了 Gaber 滤波器，SIFT 特征提取算子，马尔科夫随机场的图模型。因此，在这个阶段，人们倾向于发展专用算法。

而脑神经科学的几个突破性进展使人们彻底改变了看法。2000 年左右，Jitendra Sharma 在《自然》(Nature) 上发表文章<sup>[2]</sup>，汇报了他们的一个令人耳目一新的实验。

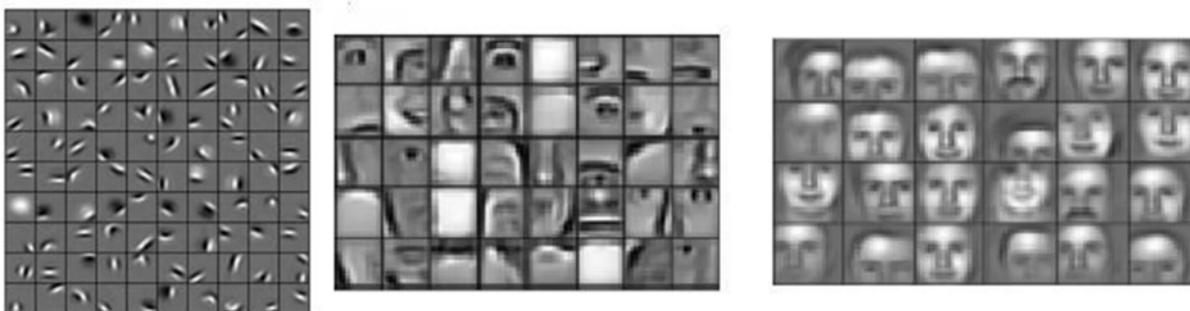


图5 深度学习神经网络学习得到的不同层次的特征 (by Andrew Ng)

Sharma 把幼年鼯鼠的视觉神经和听觉神经剪断，交换后接合，眼睛接到了听觉中枢，耳朵接到了视觉中枢。鼯鼠长大后，依然发展出了视觉和听觉。这意味着大脑中视觉和听觉的计算方法是通用的。2009 年，Vuillerme 和 Cuisinier 为盲人发明了一套装置<sup>[3]</sup>，将摄像机的输出表示成二维微电极矩阵，放在舌头表面。盲人经过一段时间的学习训练，可以用舌头“看到”障碍物。2011 年，人们发现许多盲人独自发展出一套“声纳”技术，他们可以通过回声来探测并规避大的障碍物。Thaler 等研究表明，他们“声纳”技术采用的并不是听觉中枢，而是原来被废置的视觉中枢。

种种研究表明，大脑实际上是一台“万用学习机器”（universal learning machine），同样的学习机制可以用于完全不同的应用。人类的 DNA 并不提供各种用途的算法，而只提供基本的普适学习机制，人的思维功能主要是依赖于学习所得。后天的文化和环境决定了一个人的思想和能力。换句话而言，学习的机制人人相同，

但是学习的内容决定了人的心智（mind）。

### 1.3 可塑性（plasticity）

人的大脑具有极强的可塑性，许多功能取决于后天训练。例如，不同民族语言具有不同的元音和辅音，阿拉伯语最为复杂，日语相对简单。出生不久的婴儿可以辨别听出人类能够发出的所有元音和辅音，但是在 5 岁左右，日本幼儿已经听不出很多阿拉伯语中的音素了。同样，欧洲人可以容易地辨认本民族面孔，但是非常容易混淆亚洲人面孔。人们发现，如果大脑某个半球的一个区域受损，产生功能障碍，随时间流逝，另一半球的对称区域会“接替”受损区域，掌管相应功能。这些均表明大脑神经网络具有强烈的可塑性。

长期以来，人们倾向于认为大脑神经网络通过学习和训练，其联结复杂度逐渐增加，愈来愈多的联结建立起来。近期，一些神经科学家提出了相反的看法。他们观察到，婴儿睡觉时，如果有剪刀掉到地上，婴儿的应激反应是全身的，而相对成

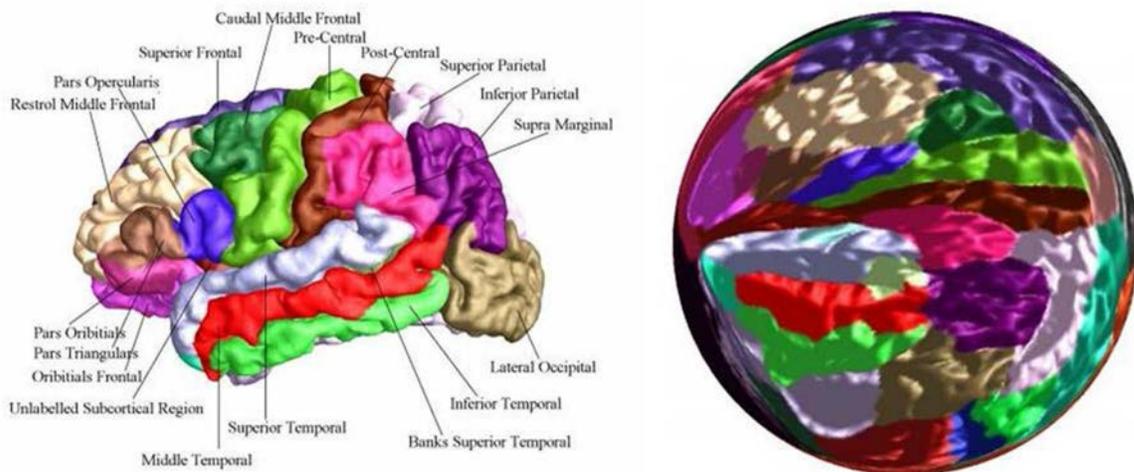


图 6 大脑皮层的不同功能区域

熟的儿童的应激反应只集中在局部肌肉上面。他们找到一些证据表明，婴儿的某些神经网络是全联通图，依随年龄的增长和学习训练的积累，许多神经联结会自行断开，从而形成简化的网络。

#### 1.4 梦的解析 (dream)

大脑学习算法的普适性和可塑性一直激励着计算机科学家不懈地努力探索。历史性的突破发生在 2006 年左右。3 位计算机科学家，Geoffrey Hinton, Yann LeCun 和 Yoshua Bengio 突破了深度学习的技术瓶颈，引领了深度学习的浪潮。相比以前的状况，主要的技术突破在于以下几点：优化方法的改进，更加简单的优化方法，特别是随机梯度下降方法的应用；使用非监督数据训练模型以达到特征自动提取；使用越来越大的数据集；深度神经网络和大数据训练需要巨大的计算能力，GPU 的普遍使用解决了这一迫切要求等。现在，深度学习方法突飞猛进，在图像识别 (image registration)、文本处理 (text)、语音处理 (speech) 等领域的基本问题上，都已经超过了传统方法。在图像识别领域，2015 年深度学习方法的识别准确率已经达到人类的水平<sup>[4]</sup>。

同时，对于深度学习神经网络的理解加深了人类对于自身智力活动的理解。长期以来，人们对于梦境一直没有很好的理解。一直解释观点如下，如图 7 所示，大脑中有一对海马体 (hipocampus)，它们和人类的长期记忆有关。如果把大脑比喻成一个数据库，那么海马体就像是索引。如果海马体有问题，那么许多存入的记忆无法被取出，同时也无法形成新的记忆。每天晚

上，海马体将当天形成短暂记忆加工成长期记忆，在这一过程中，就形成了梦。海马体和其他神经中枢相连，处理其他中枢已经处理好的数据，形成新的编码。海马体和视觉和听觉中枢直接相连，因此，在梦中能够看到并且听到；但是，海马体和嗅觉中枢并不相连，因此，在梦中无法闻到气味。在梦中经常能够看到平时看不到的奇诡景象，可以用深度学习的方法加以模拟解释。

实际上，视觉处理的过程并不只是从低级向高级传递的单向过程，高级中枢可以向低级中枢发出反馈信息，最明显的例子是高级中枢可以决定低级中枢的“注意力”和“焦点”。当看到模糊不清的图像、或一时无法辨认的图像时，高级中枢会产生各种概率上合理的解释，并且由这种猜测先入为主地影响低层中枢的判断，从而产生错觉。如图 8 所示，可以用深度学习的神经网络模拟这种先入为主的现象<sup>[5]</sup>。输入一幅白噪声，本身没有任何有意义的信息。网络由于随机涨落，在某一刹那倾向认为图中有香蕉。由此，优化这幅图像，

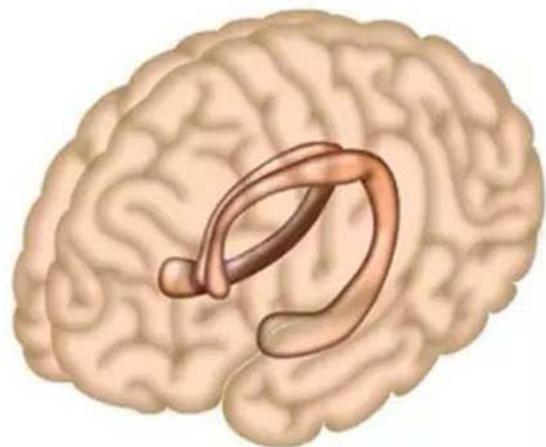


图 7 大脑的海马体 (hipocampus)

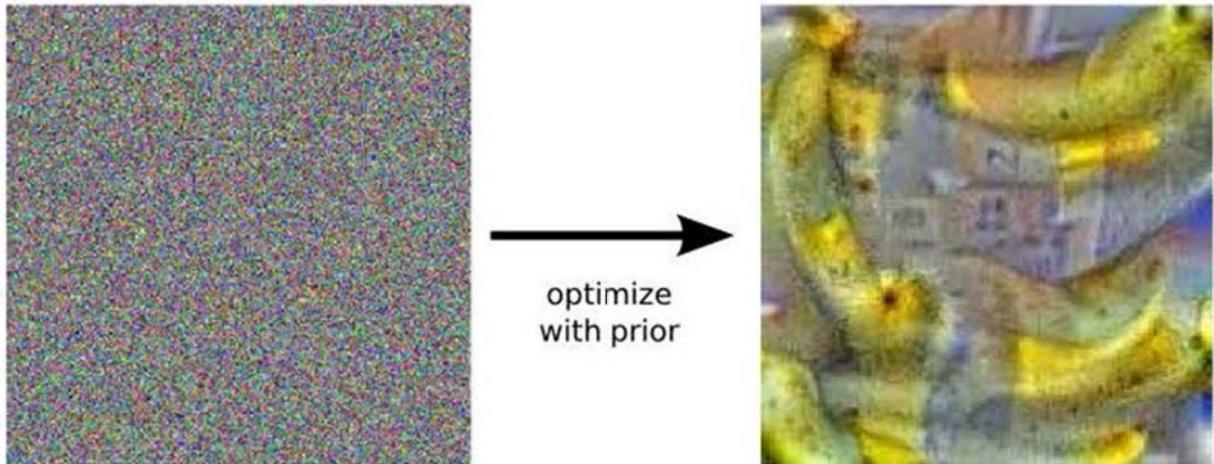


图 8 先入为主产生的错觉

使得识别香蕉的高层神经元兴奋，如此得到的图像果真看起来像香蕉。

图 9 显示了另外一个例子<sup>[5]</sup>，输入一幅羚羊图像，神经网络的低级反馈加到图像上，看到许多边缘和定向的模式出现在场景里。

许多孩子喜欢仰望蓝天白云，并且用自己丰富的想象力看到了各种奇妙的幻象。如图 10 所示<sup>[5]</sup>，将一幅蓝天白云的图像作为输入，用一个识别动物的深度学习神经网络加以处理，将高层神经元的认知模式作为反馈，优化原始图像，结果可以看到

各种山海经中才会出现的神兽：身着铠甲的将军狗，猪蜗牛，骆驼鸟，狗鱼。人在做梦时，高层神经元对于低层神经元发出各种反馈，低层神经元将图像依照高层的意图进行诠释幻化，视觉幻象由此产生。

### 1.5 美学 (aesthetics)

很久以来，人们倾向于认为机器可以理解人类的逻辑思维，但却无法理解人类的丰富感情，更无法理解人类的美学价值，当然机器也就无法产生具有美学价值的作品。事实胜于雄辩，AlphaGo 对局李世石下出石破天惊的一步，棋圣聂卫平向

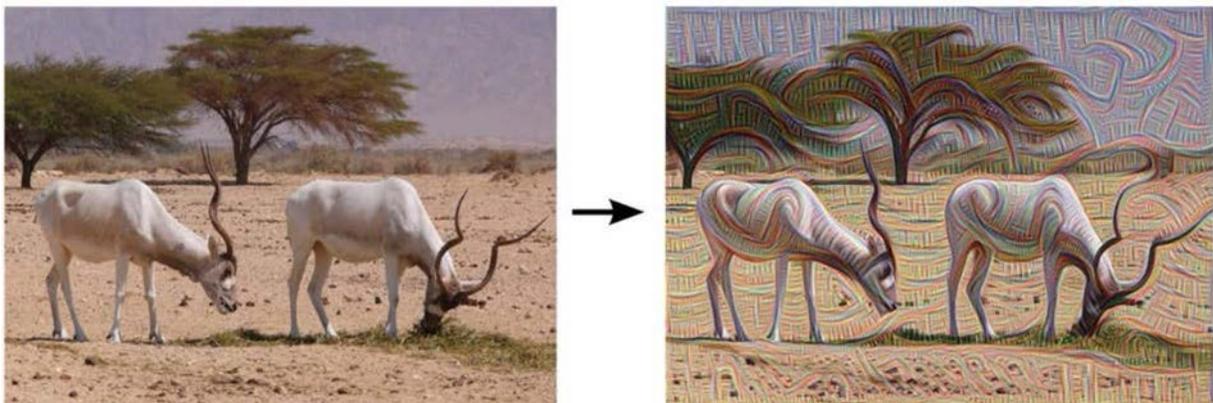


图 9 图像加入低级特征反馈

AlphaGo 的下法脱帽致敬，这说明深度学习算法已经能够自发创造美学价值。许多棋手在棋盘方寸间纵横一生，所追寻的就是美轮美奂的神机妙手。如此深邃优美、玄奥抽象，一夜间变成了枯燥平淡的神经元

参数，这令许多人心生幻灭。

其实，在视觉艺术领域，人工神经网络已经可以将一幅作品的内容和风格分开，同时向艺术大师学习艺术风格，并把艺术风格转移到另外的作品中，用不同艺术家

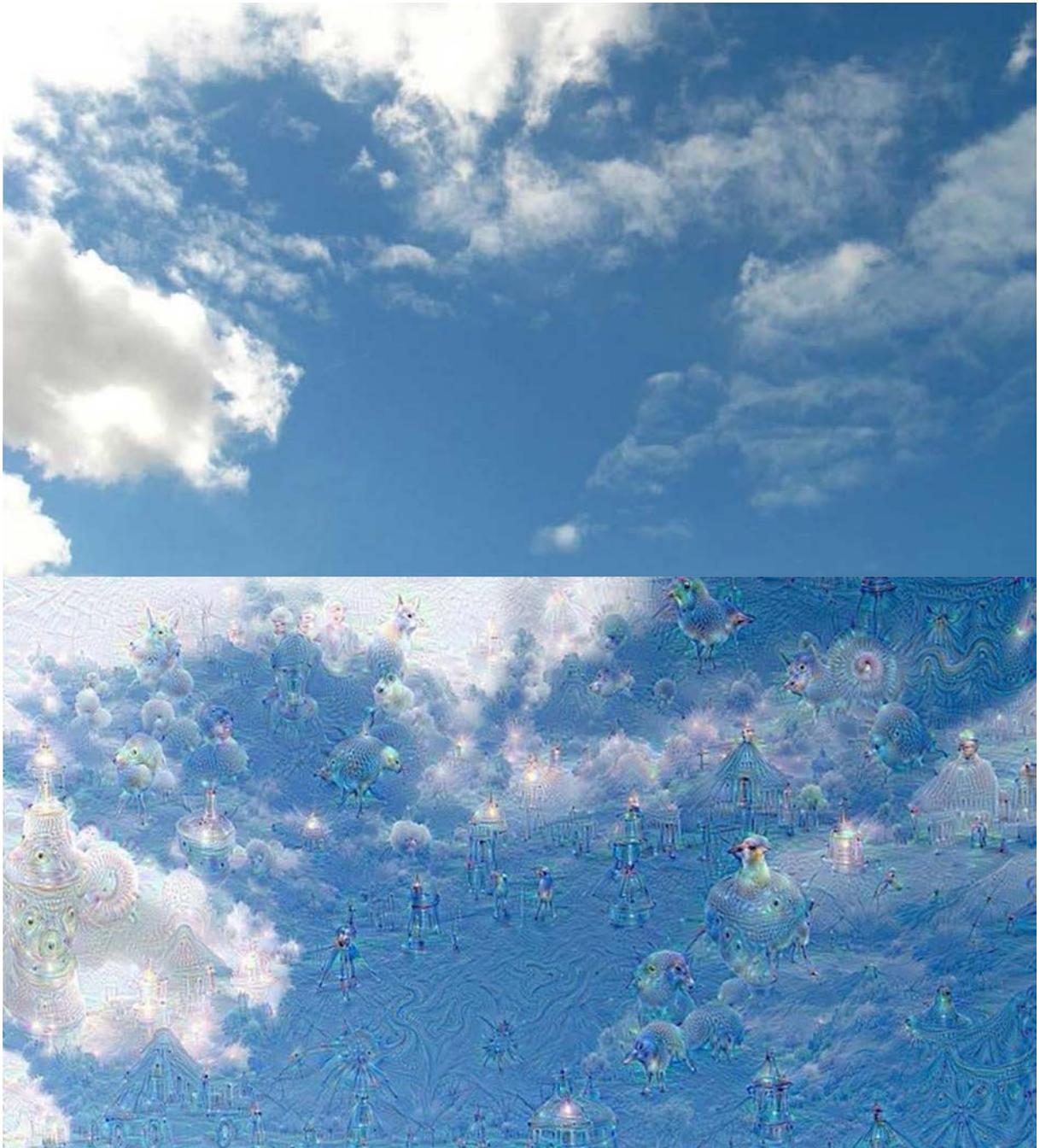




图 10 白云苍狗的机器学习解释

的风格渲染同样的内容<sup>[6]</sup>，如图 11 所示。

这意味着人工神经网络可以精确量化原本许多人文科学中模糊含混的概念，例如特定领域中的“艺术风格”，博弈中的“棋风”，并且使这些只可意会、无法言传的技巧风格变得朴实无华，容易复制和推广。

## 2 符号主义

古希腊人将欧几里得几何归纳整理成欧几里得公理体系，整个宏伟的理论大厦奠基于几条不言自明的公理，整个大厦完全由逻辑构造出来，美轮美奂，无懈可击。

这为整个人类科学发展提供了一套标准的范式。后来，牛顿编撰鸿篇巨著《自然哲学的数学原理》也遵循公理体系的范式，由公理到定义、引理、定理，再到推论。人类的现代数学和物理知识最终都被系统化整理成公理体系，比如爱因斯坦的广义相对论也是遵循公理体系的范式。当然也存在例外，虽然量子理论已经为人类科技带来天翻地覆的革命，但是量子理论的公理体系目前还没有建立起来。符号主义的主要思想就是应用逻辑推理法则，从公理出发推演整个理论体系。



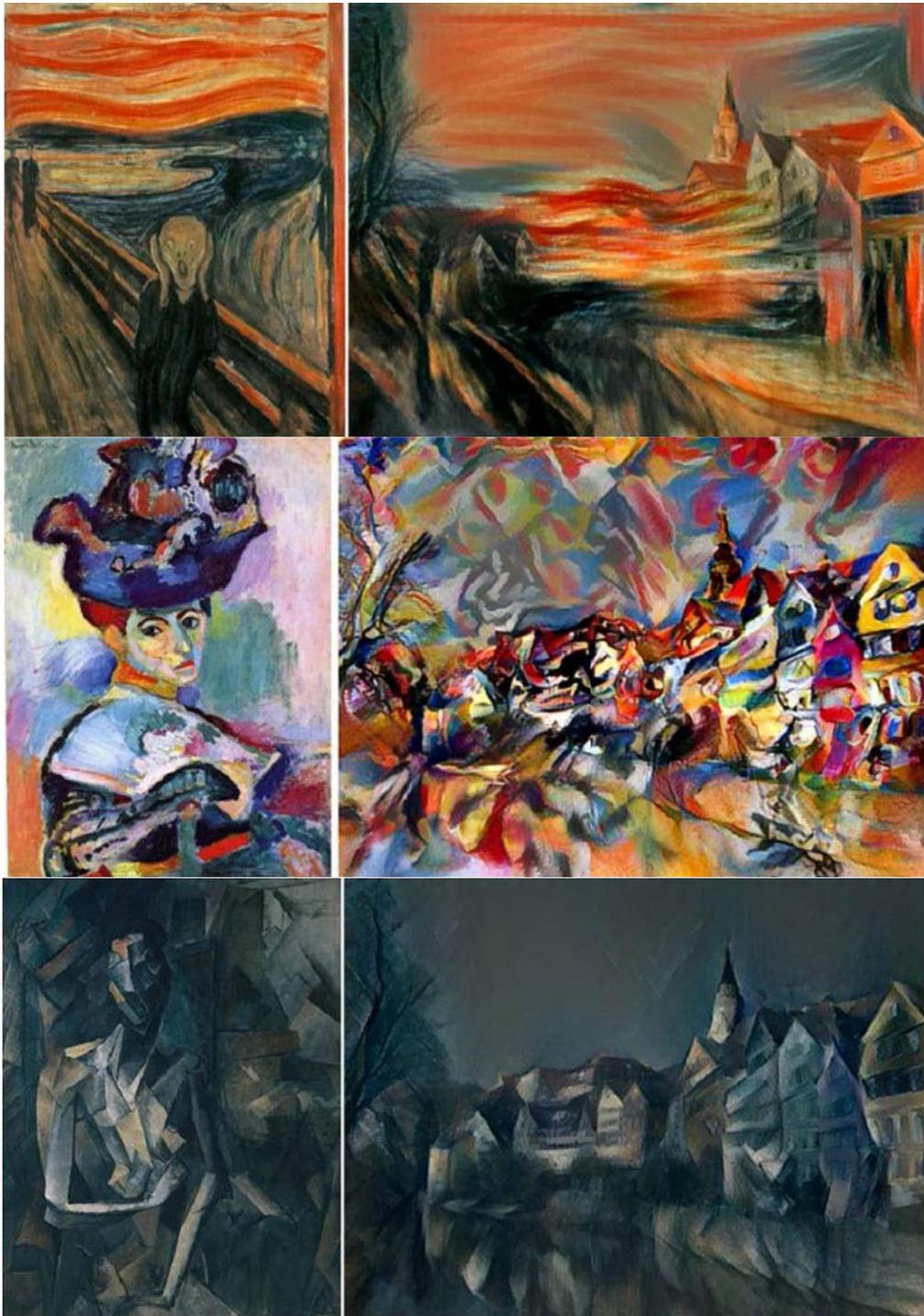


图 11 神经网络能够自动学习艺术风格，并用不同的风格渲染同样的内容

人工智能中，符号主义的一个代表就是机器定理证明，吴文俊先生创立的吴文俊方法是其巅峰之一。目前基于符号计算

的机器定理证明的理论根基是希尔伯特定理：多元多项式环中的理想都是有限生成的。首先将一个几何命题的条件转换成代

数多项式，同时把结论也转换成多项式，然后证明条件多项式生成的根理想包含结论对应的多项式，即将定理证明转换为根理想成员判定问题。一般而言，多项式理想的基底并不唯一，Groebner 基方法和吴文俊方法可以生成满足特定条件的理想基底，都可以自动判定理想成员问题。因此理论上代数范畴的机器定理证明可以被完成，但实际中这种方法有重重困难。

首先，从哲学层面上讲，希尔伯特希望用公理化方法彻底严密化数学基础。哥德尔证明了对于任何一个包含算术系统的公理体系，都存在一个命题，其真伪无法在此公理体系中判定。换言之，这一命题的成立与否都与此公理体系相容。这意味着我们无法建立包罗万象的公理体系，无论如何，总存在真理游离在有限公理体系之外；另一方面，这也意味着对于真理的探索过程永无止境。

其次，从计算角度而言，Groebner 基方法和吴文俊方法的复杂度都是超指数级别的，即便对于简单的几何命题，其机器证明过程都可能引发存储空间的指数爆炸，这揭示了机器证明的本质难度。吴文俊方法的成功有赖于大多数几何定理所涉及的代数计算问题是有结构的，因而可以快速求解。

第三，能够用理想生成的框架证明的数学命题，其本身应该是已经被代数化了。例如所有的欧几里得几何命题、初等的解析几何命题。微分几何中的许多问题的代数化，本身就是非常具有挑战性。例如黎曼流形的陈省身-高斯-博内定理：流形的总曲率是拓扑不变量。如果没有嘉当发明的

外微分和活动标架法，这一定理的证明无法被代数化。拓扑学中的许多命题的代数化本身也是非常困难的，比如众所周知的布劳威尔不动点定理：用咖啡勺缓慢均匀搅拌咖啡，然后抽离咖啡勺，待咖啡静止后，必有一个分子，其搅拌前和搅拌后的位置重合。这一命题的严格代数化是一个非常困难的问题。吴先生的高足，高小山研究员突破的微分结式理论，系统地将这种机器证明方法从代数范畴推广到微分范畴<sup>[7]</sup>。

最后，机器定理证明过程中推导出的大量符号公式，人类无法理解其内在的几何含义，无法建立几何直觉。而几何直觉和审美，实际上是指导数学家在几何天地中开疆拓土的最主要的原则。机器无法抽象出几何直觉，也无法建立审美观念，因此虽然机器定理证明经常对于已知的定理给出令人匪夷所思的新颖证明方法，但是迄今为止，机器并没有自行发现深刻的未知数学定理。

比如，人类借助计算机完成了地图四色定理的证明，但是对于这一证明的意义一直富有争议。首先，这种暴力证明方法没有提出新的概念、新的方法；其次，这个证明没有将这个问题和其他数学分支发生深刻内在的联系。数学中，命题猜测的证明本身并不重要，真正重要的是证明所引发的概念思想，内在联系和理论体系。因此，许多人认为地图四色定理的证明实际上“验证”了一个事实，而非“证明”了一个定理。目前，机器定理证明的主流逐渐演变成机器验证。

因此，和人类智慧相比，人工智能的符号主义方法依然处于相对幼稚的阶段。

### 3 展望

虽然人工智能取得了突破性进展，但是它还处于婴幼儿时期。联结主义的方法虽然摧枯拉朽、无坚不摧，但是依然没有坚实的理论基础。通过仿生学和经验积累得到的突破，依然无法透彻理解和预测。简单的神经网络学习机制加上机器蛮力，能否真正从量变到质变，这需要时间检验。

围棋是信息完全博弈游戏的巅峰，但不是人类智力的巅峰。理性思维的巅峰还是数学物理理论的创立。许多抽象的数学定理，本身的描述已经概念嵌套概念，并且在现实物理世界中找不到示例，因此机器学习方法不再适用。比如下面的命题：高斯曲率处处为负的封闭曲面无法嵌入在三维欧式空间中。这一命题的证明无法用目前符号计算的方法，也无法用机器学习

的方法。如果有朝一日，数学家开始对人工智能脱帽致敬，那么人类应该开始警醒了。

人类的孩子需要花费十数年来学习，在这漫长的学习过程中，师生和同学中的社会交往起到了至关重要的作用。目前人工智能的程序还是各自训练，彼此之间没有交互。如果人工智能程序间能够建立社会交往，彼此交换学习内容和心得，那么人工智能将会产生新的飞跃。比如一群谷歌的 Atlas 机器人（图 12），装备了 AlphaGo 的大脑，彼此自发地进行篮球游戏。通过拷贝神经网络的参数，它们可以彼此迅速交换学习成果。通过对抗演练，它们能够迅速提高手眼协调，团队协作，则很快有望进入 NBA。

也许，在不久的将来，人类和人工智能机器人和睦共处，共同学习提高。在纽约曼哈顿中央公园，有机器人在辛勤地劳



图 12 谷歌旗下的机器人

作，也有人类遛着阿尔法狗。街角，有个 Atlas 在和人类攀谈，切磋着各种幽默笑话。不远处的长椅上另一个 Atlas 默默地坐着，它头戴虚拟现实的眼镜，在孜孜不倦地学习着人类历史……

### 参考文献 (References)

- [1] Brewer A A, Liu J V, Wade A R, et. al. Visual field maps and stimulus selectivity in human ventral occipital cortex [J]. *Nature Neuroscience*, 2005; 8(10):1411-1411.
- [2] Sharma J, Angelucci A, Sur M. Induction of visual orientation modules in auditory cortex [J]. *Nature*, 404(6780): 841-847.
- [3] Vuillerme N, Cuisinier R. Sensory supplementation through tongue electrotactile stimulation to preserve head stabilization in space in the absence of vision [J]. *Investigative Ophthalmology & Visual Science*, 2009, 50(1): 476-481.
- [4] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition [R]. arXiv:1512.03385.
- [5] <http://googleresearch.blogspot.com/2015/06/inceptionism-going-deeper-into-neural.html>.
- [6] <https://github.com/jcjohnson/neural-style>
- [7] Gao X S, Li W, Yuan C M. Intersection theory in differential algebraic geometry: generic intersections and the differential

chow form [J]. *Transactions of the American Mathematical Society*, 2013, 365(9): 4575-4632.

### 作者介绍

顾险峰博士，纽约州立大学石溪分校计算机科



学系和应用数学系终身教授，清华大学丘成桐数学科学中心客座教授。2005 年获得美国国家自然科学基金 CAREER 奖，2006 年获得中国国家自然科学基金海外杰出青年学者奖，

2013 年第六届世界华人数学家大会晨兴应用数学金奖等。和丘成桐先生及其合作者共同创立了一门新兴的跨领域学科——计算共形几何。

在数学、工程和医学领域的国际顶级杂志和会议发表论文 270 多篇。担任国际期刊

《*Graphical Models*》，《*IEEE Transaction on Computer Graphics and Visualization*》的编辑和

《*Geometry, Imaging and Computation*》的主编。

其主要著作有《*计算共形几何*》、《*离散曲面的变分原理*》和《*Ricci Flow for Shape Analysis and Surface Registration*》等。

# Where is Artificial Intelligence Heading?

## 人工智能，何去何从

**Dr. S. Kevin Zhou**

周少华

Siemens Healthineers

755 College Road East, Princeton NJ 08540, USA

西门子医疗技术中心

### Abstract

Artificial intelligence (AI) is one of the hottest R&D topics. Where is AI heading? In this article, the author attempts to first uncover the AI technology, then present the status quo of AI, and finally discuss how to commercialize AI. The author firmly believes that turning AI into real products is the ultimate destination of the AI evolution.

### 引言

人工智能是当前最热门的话题之一。那么，人工智能究竟何去何从？在这里，笔者试图揭秘人工智能之技术，谈论人工智能的现状，并阐述人工智能技术如何产品化。笔者深深认为，只有将人工智能技术进行踏实的产品化，才是公司的正道正解，是将人工智能进行到底的不二途径。

### 人工智能之技术



图1 数据化和智能化

人类通过传感器把世界或环境数字化，如图 1。照相机采集可视化的图像视频，麦克风生成音频，医疗设备产生医学影像，互联网络实现文本数字化。

一般的数据信号以比特存储的，是非结构化的。当前人工智能的主要形式是基于非结构化的数据抽取结构化的内容，或作出动作并反馈给环境。

假设数据是  $X$ ，内容或动作是  $Y$ ，简单来说，人工智能就是一个函数  $Y=F(X)$ 。给定一个  $X$ ，则算出  $Y$ 。问题是：函数  $F$  是什么？在一般的问题或系统中，函数  $F$  是未知的；所以，需要“学习”得到。现在的人工智能多基于神经网络。神经网络定义了函数  $F$  的一种参数化的形式，“学习”的过程或训练神经网络则是指找到最合适的参数。

现在的人工智能核心技术是深度学习，目前的深度学习是基于监督学习的。大量的数据样本用来描述数据间的匹配关系，深度学习以此训练神经网络的海量参数。以最通俗的语言来说，神经网络其实是以无与伦比的暴力记忆能力、以海量参数生硬地记住了数据间的匹配关系，再以无可匹敌的计算能力重现了此关系。

打一比方。现在的人工智能就像是打球熟练的球手，千万次的挥拍之后，身体有了自然的肌肉反应。所谓，兵来将挡水来土掩是也。当然不同的球手会有不同的反应，正如神经网络会收敛到不同的 local minima，于是网络性能有高下之分，球手有强弱之别。

特别是，在所谓的完整信息 complete

information 系统中，也就是说系统的输入与输出有完整的定义，人工智能的确可能完胜人类。究其原因，人毕竟在记忆能力和计算能力和电脑相比有所不足。论记忆样本，电脑见到可以是完整信息体系下的全部样本（如果可以有限穷举），或极大数样本；而单个人脑所涉及的样本终究有限。同样地，电脑，既使简单如计算器，计算也比人脑快好几个数量级。

完整定义的输入输出具有很强的局限性，因而此类人工智能被称为弱人工智能。与此相对的是强人工智能，如何在信息不完整的条件下做出判断。此时，电脑不知道要算什么，或者既使知道，计算量超出现在的计算能力，从而算不出来。人类却可以利用已有的知识体系来“填补”不完整的信息，从而进行有效的推测。当然，这方面的研究也是日新月异，最近的一个成功例子是 CMU 研究者开发的 Libratus 系统赢得了 No-Limit Texas Hold'em Poker Game，但她不是基于深度学习的。

人的认知有三大区域：舒适区、学习区、恐慌区。在舒适区呆着，学不到任何新知识。出了舒适区，就进入学习区，这是人学习的主要渠道。当然离舒适区太远，进入恐慌区，人其实啥都学不到，只知道恐慌。弱人工智能其实是在舒适区学习，只有知识都定义好，数据足够多，那的确可能学得很好，甚至超过人。但一旦测试数据不在舒适区，尤其当计算机遇到突发情况、以前未曾谋面的 Outliers 时，效果马上突降，人工智能变成人工稚嫩，神经网络出“神经刀”之举，而人是保持一定的认知水准之上。

试举几例。AlphaGo 在和李世石下第四盘时，突然“发癫”，从而断送大好形势，自毁锦绣前程。Telsa Autopilot 会把横穿马路的白色大平板车当作白云，速度依旧，导致 Joshua Brown 先生命丧黄泉。大获成功的 ImageNet 的 winner algorithm，居然会把熊猫认作猩猩，哪怕只是在熊猫的原图中加入一点点微小的有结构性的干扰信号。

上述现象其实本质上不难理解的，原因来自于数据收集的局限性的：一是数据收集的不可靠性；二是数据收集的不可穷举性。更深层的原因来自于神经网络的不可解释性。黑箱性质导致只知其然，而不知其所以然。当然即使出了错误也不知道，只能是事后弥补，也就是把 outliers 收集下来，重新训练神经网络，希望迭代后的网络可以训练得更好更强。

不得不说，这距离通用的人工智能（即强人工智能）差得十万八千里。强人工智能是指机器完全复制人类智慧，包括人类的独立学习和决策能力。还是谭铁牛院士总结得非常好：“人工智能是有智能没智慧，有智商没情商，会计算不会算计，有专才无通才”。

### 人工智能之现状

随着人工智能第二春的来临，人工智能呈现缤纷繁杂、甚至互为矛盾的诸多现状。她时而显得神秘，甚至神圣，但却又被逐渐被普罗大众接受，甚至有娱乐化的倾向。虽然大小公司皆 AI+，泡沫仍在不断被放大。

**人工智能的神秘化** 人工智能击败了人类，计算机深度神经网络完胜大脑，这样

的标题不时出现在媒体报道中。横空出世的 AlphaGo 与李世石的人机对决让媒体报道达到了一个小高潮。曾经以为念念不忘的机器故事是永远不变的传说，却在 AlphaGo 的黑白棋子间成了永不褪色的背景。尤其是神经网络的黑箱性质和不可解释性，更是给人工智能披上了一层神秘面纱，抹上了一丝神秘色彩。

**人工智能的神圣化** 在人工智能屡屡战胜人类的背景下，似乎只要是人工智能出马，一切问题解决，万事搞定。甚至有人认为，只要数据足够多，网络足够深，内存足够大，可以记住全世界。什么牛顿的三大定律，爱因斯坦的相对论，全部靠边站。人工智能会很快替代低端重复性劳动（这确实很有可能），甚至连医生、律师、科学家等高端职业亦会很快消失（这其实不太可能）。

**人工智能的大众化** 李飞飞在最近一次的谷歌云产品发布会提出来了一个口号：“Democratizing AI”。意思是指要让普通大众通过谷歌云提供的接口直接获取 AI 的最新技术，占据 AI 带来的用户入口。亚马逊的 Alexa Echo 试图在语音入口之争中占得上风。

**人工智能的娱乐化** 现在不仅 VC 加大加紧投入，企图赶上“风口”，生怕错过了下一个独角兽；媒体杂志甚至娱乐节目也争相报道采访。于是，AI 专家满天飞，言必称大数据深度学习，ImageNet、AlphaGo 满嘴跑。明明是 IT 会议，活生生变成 AI 会议。于是乎，AI 专家快和娱乐明星接近等同。人工智能也一夜暴红，进入了大众视野，逐渐“深入人心”。

**人工智能的泛滥化** 与娱乐化对应的是人工智能的泛滥化。正所谓“忽如一夜春风来，千公万司被 AI。”任何公司，不论是巨头、老牌公司，还是新贵、startup，不管是生产电器的，亦或是提供服务的，都在往人工智能上靠，都要沾 AI 的光。“人工智能+”的口号也算是一个侧证；但是，“+”的重点不是为了加而加，而是为何加，如何加。

**人工智能的泡沫化** 不可否认的是，人工智能的兴起的确给笔者的同行们带来了切实的好处。最大的好处就是工作好找，工资高企。以前，如联想 CTO 芮勇博士说的，做 AI 的被认为是骗子，学术教授们的“投资回报”不成正比。现在，学术教授则成了香饽饽，各大公司争相挖人，八位数的工资亦不少见；或者，干脆直接 VC 投资创业，争取尽早实现财富自由。那怕刚毕业的博士生，也有报道收入达百万美金！这么高回报率肯定也是不合理的。随着泡沫越吹越大，最终的结果恐怕只会是破灭。

更有甚者，有人说，中国当前最火的行业有二个，一是房地产，另一是人工智能。难道人工智能已经达到了与房地产相提并论的泡沫高度？

### 人工智能之产品化

人工智能纵有“千变万化”，她真正需要的是产品化，踏踏实实的产品化，化人工智能为实实在在的生产力。

产品与技术是相辅相成的。在商业公司里，通常产品为主导，技术为驱动。也有以技术为核心，最后找到产品出口，从而实现转化。

产品是商业公司的立足之本。只有产品为客户提供了真正的价值，产品才能为客户接受，客户才会真心愿意买单，公司从而可以完成赢利。公司由此获得养分，有了进一步发展的空间，开发更新更好的产品，完成一个正向循环。正如 Steve Jobs 说的，“You got to start with the customer experience, and work backward to the technology 要从客户体验出发，倒推到技术。”

产品价值的体现是多方面的，主要是解决客户的痛点问题：省时省力，提高效率。人工智能的产业化就是要针对痛点痛下杀手；若只是隔靴搔痒，后面的想象空间就不大了。

当找到了产品的痛点场景之后，与场景匹配的数据流随之定义了数据的输入与输出，数据采集的大戏开场了。同时，这也需要产品设计时要有数据采集的意识，不然会造成核心数据的缺失，无法产生直接可用的监督信号。

在产品化时，必须要充分认识到现有人工智能技术的局限性，尤其是当场景的复杂度很高时，数据收集达不到完整程度。以无人车为例，道路桥梁千种万种，自然条件变化多端，驾驶人员粗心大意随心所欲，导致边角情况层出不穷，这就造成了无人车从辅助驾驶到真正无人驾驶的鸿沟迟迟无法被飞越。

很多情况下，人工智能系统更适合做个安静的贴心助手，不知疲倦地观测一些特定场景，并及时给出提示。这要求人机有很好的交互设计，以免人类的自尊心受

损。

就创业而言，正如李开复博士说的，技术大牛+产品经理是最佳拍档。产品经理负责场景定义，技术大牛负责技术实现。运营也是不可或缺的一部分；没有运营，产品滞销。虽说是“酒香不怕巷子深”、“金子迟早会发光”，只怕巷子太深闻不到酒香，发光太迟湮没在土中，熬不到雨过天晴阳光灿烂之时。

### 结束语

在指定的任务中，人工智能超过人类是既成的事实。但是，机器和人类并不是对立关系，机器应该会成为人类的好助手。做好产品化，让机器更好地为人类服务，而不是成为人类的对手，这是我们人工智能从业者的终极任务。

### 作者简介



**周少华**，本科毕业于中国科技大学，博士毕业于马里兰大学，专业电气工程。现于西门子医疗技术中心任职首席图像分析专家，致力于研究计算机视觉和图像分析并开发相关的创新产品。他已经撰写了两本学术专著，编辑了三本学术书籍，发表了超过近 50 篇学术期刊论文和著作章节和超过 130 篇学术会议论文，并拥有超过 80 项批准专利和超过 300 项公开发明。他多次获奖以表彰其学术成就和创新贡献，具体奖项包括研发百佳奖（即科技创新奥斯卡奖）、西门子年度发明家、西门子研究技术大奖、爱迪生专利奖、强生公司的供应商创新奖、马里兰大学电气工程系杰出校友奖等。同时，他热心于奉献给专业社区，是《IEEE 医学影像汇刊》和《医学图像分析》等期刊的副主编、美国医学与生物工程院的会员、计算机视觉和模式识别大会（CVPR）和医学图像计算和计算机辅助介入大会（MICCAI）的领域主席、公众号《视觉求索》的主编。

# Human-Machine Perception and Assistive Technology

## 人机感知和辅助技术研究

Zhigang Zhu

朱志刚

The City College, City University of New York

160 Convent Ave, New York, NY 10031, USA

纽约城市大学城市学院计算机系

Email: zhu@cs.cny.cuny.edu

### Abstract

Human vision is truly delicate, and the human brain is more amazing. But if either of them has problems, daily lives are very challenging. Assistive living requires emerging technology, and integrative research needs collaboration. For this end, funding support is important, and user engagement is even more critical. Further, the goals of research at higher education are not only to inspire new ideas, but also to train the next generation workforce, to do good to society. Under the support of an NSF program Emerging Frontier in Research & Innovation (EFRI), in the topic area of Man, Machine and Motor Control (M3C), human and machine vision research has been carried out at the City College of New York. This article describes some of the results and findings. It has been found that multimodal perception is the key for assistive technology, and integration of mobile and cloud computing is the backbone. Both virtual reality and

gaming can be more positive and beneficial, and if we are not constrained by conventional wisdom, substitute perception can be developed and well accepted. One of our objectives is for location-based services go indoors, with multimodal sensing and deep learning techniques. It also becomes evident that public transportation needs to be improved, to be both smart and accessible. Together with technology development and application-driven implementations, the hope is that the mystery of the human brain can be further explored as well.

### 引言

辅助技术（Assistive Technology），顾名思义，就是用各种技术来辅助那些需要帮助的人，特别是在感知或交流有挑战的朋友，如视障<sup>[1]</sup>，自闭<sup>[2]</sup>。本文着重在感知辅助技术上。人机感知，就是要结合人的感知和计算特长和机器（包括计算机、机器人、传感器等）的特长，更好地让机器服务于人类。本文以一个美国自然科学基

金 (NSF) 资助的跨学科人机辅助视觉研究的项目为中心, 介绍一下纽约城市学院 (CCNY) 视觉计算实验室近几年人机视觉和辅助技术的研究。这个 NSF 资助的研究创造前沿 (EFRI) 项目, 是两所院校五个实验室的联合: 城市学院的视觉计算实验室、媒体实验室、和脑科学实验室 [3, 4, 5], 佐治亚理工的机器视觉实验室和生理实验室 [6, 7]。研究的内容涵盖网膜和脑皮层 TMS, 大脑分析, 舌头读图, 标识识别, 深度触觉, 网膜移植, 运动分析等, 主要应用对象为视障人士 (图 1)。本文只介绍和城市学院视觉视觉实验室密切相关的几

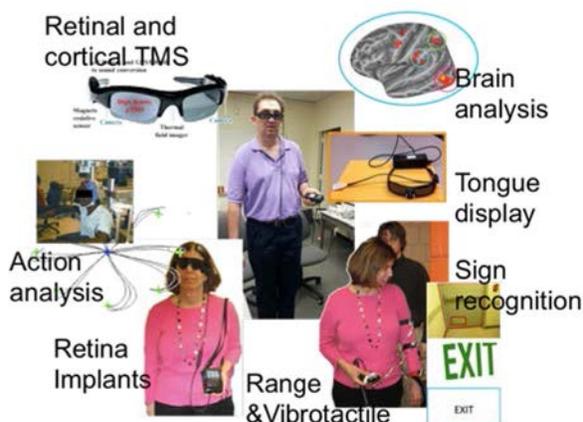


图 1 CCNY-Georgia Tech 人机感知及视觉辅助技术研究<sup>[8]</sup>。涉及得技术包括 (自左上顺时针): 网膜和脑皮层 TMS, 脑分析, 舌头读图, 标识识别, 深度触觉, 网膜移植, 动作分析。

个方面, 并试图从辅助视觉扩展道涵盖更广的感知辅助技术的研究和教育<sup>[8-11]</sup>, 在应用方面服务于视力障碍、听力障碍, 运动障碍、失忆症、老年人、自闭症等, 在科学研究方面探索人脑的奥秘, 包括感知、表情和社会交流。所报道的研究结果尚离期望有不少差距, 但求本文所描述的科研

探索能启发更多的研究思路并导致更成熟的研究成果。

下面, 本文先以一个人机视觉研究的学术研讨会为主, 讨论视觉辅助技术研究和开发方面的最新进展点滴和一点思考。然后, 再介绍视觉实验室关于人机感知和辅助技术的五个主要专题:

- (1) 视觉导航技术: 包括双目、单目、全相和三维视觉的辅助视障应用。
- (2) 替代感知方法: 包括全身目设计, 群帮眼尝试, 肌电辨声初探, 舌头“看”图案测试。
- (3) 面部计算研究: 包括人脸识别、表情识别、深度学习和人脑感知研究。
- (4) 测试应用平台: 包括一个幻境与游戏测试和训练平台, 和一个智能无障碍交通平台。
- (5) 教育和培训项目: 包括一个暑期的培训活动和一個本科的联合实习课程。

为讨论简洁起见, 在概述以上五个方面的专题时, 每一个专题配一首小诗表达核心思想, 并以图示和参考文献做补充, 从而读者一方面无需一下子进入细节, 另一方面, 若有兴趣深入了解可参阅有关文章。如若词不达意, 先请见谅, 但希望耐心细读, 领会大意, 必得些许启发。

### 跨学科视觉辅助技术研究

根据“2012 年世界卫生组织报告”<sup>[1]</sup>, 有超过 2.85 亿视障人士, 其中 3900 万人失

明。视力受损人士中，约有 65% 是 50 岁以上，而这一年龄组占世界人口的 20% 左右。随着许多国家老年人口的增加，更多的人将面临年龄相关的视力障碍。对多模式和替代感知的研究将对社会的健康和福利产生长期影响，不仅是对视力有挑战者，而且对于经常在危险环境中工作的人员，如消防员，司机和士兵也将如此。由 NSF 资助、本文作者和多位各个领域专家主持的 2013 年 IEEE 视力障碍多模态与替代感知研讨会 (MAP4VIP 2013, 图 2)<sup>[12]</sup>，是关于这个话题的第一个 IEEE 研讨会。MAP4VIP 2013 旨在使来自多个学科 (包括计算机视觉，神经科学，多媒体计算，传感器技术和辅助技术应用) 的研究人员和从业者讨论视觉感知、计算智能、神经科学和视觉假体移植的基本问题，以帮助盲人和视障

人士和在有视觉挑战的环境中工作的人，并进一步促进人脑感知功能的科学研究。研讨会的主题包括：(1) 帮助视觉上有挑战者的计算机视觉、感知和学习算法，包括电子旅行援助，电子导向辅助，标志检测，触觉显示，脸部识别，道路和障碍物检测；(2) 视觉假肢和人机交互，包括视网膜植入物，视觉皮层刺激，生物模拟和生物系统工程探索；(3) 神经科学/生物力学/心理/社会研究，包括人类感知，神经计算，视觉编码，感觉运动模型，多感觉整合，系统评估，科学技术政策和未来趋势。

研讨会将包括以下内容：主题演讲 (一个侧重人的视觉，另一个侧重视觉假肢)；专案讨论 (包括约 6-7 名来自学术界，行业，政府和社区的专家)；口头报导

## MAP4VIP IEEE Workshop on Multimodal & Alternative Perception for Visually Impaired People



### Sponsors



National Science Foundation  
WHERE DISCOVERIES BEGIN

Microsoft  
Research

the City College of New York

General Chairs: Zhigang Zhu, Zhengyou Zhang, Kok-Meng Lee, Yann LeCun  
Program Chairs: Shawn Kelly, Tony Ro, Yingli Tian

图 2 IEEE /NSF 多模式替代感知视障辅助国际研讨会<sup>[12]</sup>(资助单位：IEEE，美国自然科学基金，微软研究院，纽约城市学院)

（包括特邀演讲和公开征文）；以及看版/演示（年轻/学生研究人员的研究结果）。专案讨论《视觉辅助：需求，挑战和希望》的议题包括：（1）您认为在未来 5 年内可能开发的视障辅助装置的期望的质量是什么？这种设备目前能允许视障人士做些什么？（2）视障人士在日常生活中面临的一些最大挑战和障碍是什么？（3）开发视觉辅助技术当前的一些限制是什么？（4）有哪些注意事项来确定辅助技术设备是否将被采用并被广泛使用（例如成本，尺寸）？（5）我们可以和视障人士可以进行哪些类型的标准化和科学测试，以如何量化视觉装置的性能提高以及各种辅助设备对提升生活质量的影响？

研讨会各行专家发表了自己的见解；详见<sup>[12]</sup>。一些有代表性的认识包括：在不熟悉或复杂环境里的视觉导航技术仍有很大的挑战，需要跨领域的研究，包括电脑、人脑并生理诸方面；视觉是一个眼见、脑算后行动的综合过程，计算视觉读标识、寻找日常生活中的物件大有需要；辅助视障的研究需要比研讨会更进一步的具体合作；视觉网膜移植装置需要更高的分辨率，比如在五年内达到 16x16，硬件方面需要有突破，软件方面也需要密切配合；计算机视觉和群帮（Crowdsourcing）技术为基础的智能手机和穿戴装置的应用软件（Apps）可能会大配用场；有可持续性的政府和民间的资助必不可少（如城市学院视觉实验室的研究就是个例子<sup>[8-11]</sup>）；研究不仅需要跨学科、并结合各方面的优势，更需要用户的密切参与；虽然辅助技术需求很大，但由于视障用户大都属于低收入的群体，新的技术难以拓展市场，所以可及性

（accessibility）的真正实现需要和提升各级政府以及民众的意识相配合；辅助视觉装置的实用功能的可靠性和使用的容易度是视障用户最关心的两大因素，但系统的性能价格比和人机智能的有效结合也是重要的考虑因素；装置的标准化、科学测试，和大规模的用户试用和反馈非常重要：我们随后的调研也发现，替代视觉的途经很多，包括物理、数字和医学等三方面的途经<sup>[13]</sup>。现有的研究和开发是百花齐放，包括由耳、舌<sup>[14]</sup>、身、手感知的替代方法，更有机器视觉理解的方法和网膜移植<sup>[15]</sup>，但科学测试标准和用户试用的方法和平台仍然缺少<sup>[13]</sup>。

另一个人机视觉可以大显用场的领域是帮助自闭患者的社交和学习。统计表明<sup>[2]</sup>，在美国 68 个人中，就有一个在自闭谱系障碍（ASD）上，这主要是大脑工作的特异性问题。和视障者类似，他们也有严重的行动和交流上的困难。当然还有其他有挑战者和老人，辅助技术特别是视觉辅助可以提供很大的帮助，包括导航、读脸、找地方等。在这些方面，多种新兴技术都可用上，包括多种传感、定位技术、移动计算、深度学习等。总之，人眼功能甚精细，大脑运作更奥秘；千万精英齐努力，成果频出需继续！

## 机器视觉技术

辅助视觉的第一个主要方面，就是机器视觉传感器的选择和视觉算法的实现。由于需要人身穿戴或携带，而且视障人士大都是低收入，所以小型化、轻便、省电、便宜是主要的考虑。当然从性能方面，实时和可靠是基本的要求。这两方面往往是

相互矛盾的，所以折衷可行的方案便是研究的重点。视觉辅助可有多个层次、多个方面。从导航来说，有近距离的安全要求（可行通路和避障）和远距离的方向指示（地方和路标识别等）。我们下面的三个技术着重与近、中距离的安全导航，即定位、定向和避障。我们首先介绍一个室内导航研究的各种方案中<sup>[16-18]</sup>的一个特别的、经济的方法，就是在智能手机上装载一个简易的 360 度全向镜头（图 3），建立一个室内场景的图像数据库，然后用图像检索的方法定位<sup>[16, 19]</sup>。为了解决实时计算速度和大数据量的要求，我们采用了移动计算和云端计算相结合的方案<sup>[16]</sup>，同时在建模时融合了其它的技术，包括地图预读，路径规划等<sup>[20]</sup>。第二个技术是双目视觉的智能采样方法（图 4），主要是为解决现今低分辨率视觉替代装置有效视觉信息的提供问题（如网膜移植产品 Argus II<sup>[15]</sup> 只有 6x10，Brainport 的舌头读图板<sup>[14]</sup> 只有



图 3 智能手机全向辅助视觉（左：头盔系统；右：手持系统）

20x20）。其中核心是针对室内场景纹理少的情况的一个基于彩色图像分割的立体视觉匹配算法<sup>[21, 22]</sup>和在其基础上的一个智能采样和感兴趣目标增强的算法<sup>[23, 24]</sup>。最后一个辅助视觉方案是用单幅图像进行三维推理的方法，目标是用智能手机的摄像和计算功能，即可进行楼道或街道方向的三维估计和近、中甚至远距离通道（如门口）

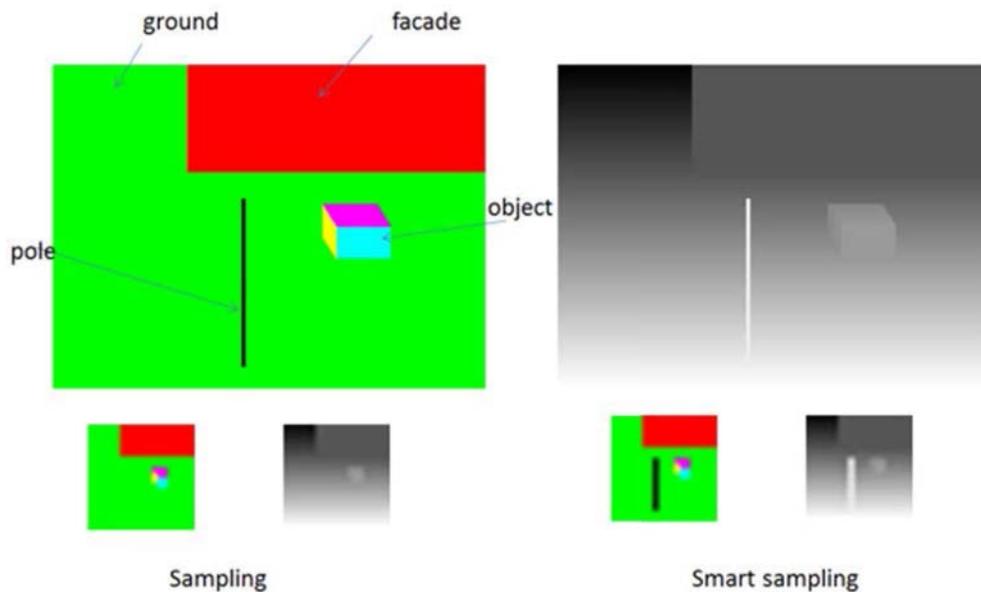


图 4 双目视觉智能采样（左上：模拟彩色图像，包括地面，楼面，方形物，细杆；右上：深度图像；左下：20x20 普通采样细杆丢失；右下：20x20 智能采样细杆保持）

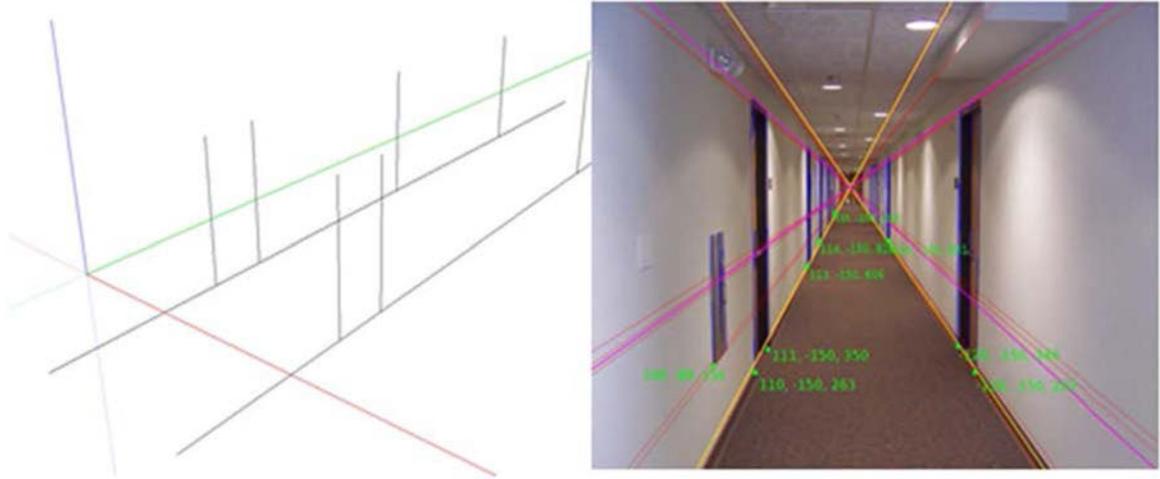


图 5 霍夫变换单目测距<sup>[25]</sup>（右：原图上标注走廊和房门的结果，并三维定位数据；左：检测结果的三维线画图模型）

的检测和距离估算。我们尝试过两种方法：<sup>[25]</sup>和基于级联（J-LINKAGE）的方法<sup>[26]</sup>。

- 智能手机观全向，图像检索定方位；<sup>[16]</sup>  
移动三维更快捷，建模定位已实时。<sup>[17, 18]</sup>
- 室内简洁少图案，立体匹配有难处；  
图像分为自然块，灵巧采样数据精。<sup>[21-24]</sup>
- 智能手机拍照易，计算功能也超强；  
人眼单目可测距，缘何不让机器做？<sup>[25, 26]</sup>
- 单项技术总有限，多种融合更完善；  
地图预读先规划，智能手机联云端。<sup>[18-20]</sup>

### 替代感知技术

辅助视觉的第二个主要方面，就是如何将采集和处理的视觉信息传递给视障用户，以及探索有无除了机器视觉以外其它

有效的方法。上面讨论的三个方案得到的定位、方向信息，除了可以用网膜移植的方法直接传递给用户的视神经外，用语音反馈恐怕是最容易想到也最有效的方法。但是太频繁占用语音频道会影响视障用户行之有效的听力导航功能，所以其它的替代感知传递方法也曾广为研究。我们尝试或测验了四个方面的替代感知技术：（1）“全身目”的设计和最优优化：尝试用皮肤各个部位来感知震动的强弱为用户提供四周的距离感知信息<sup>[29]</sup>，并决定传感器的多少和位置<sup>[30]</sup>。为此我们设计了各种模块化的距离-震动感知单元（图 6）<sup>[31]</sup>，配以适当的穿戴部件，并用 BLE 低功耗蓝牙技术将其与一智能手机连接，从而不仅方便设置参数，而且容易收集数据。（2）“群帮眼”试验系统和测试（图 7）：机器视觉是最佳选择，但是目前小巧低成本的解决方案仍待开发。作为一个替代方案，我们尝试用目前非常流行的群帮（Crowdsourcing）技术，来帮助视障用户实时导航<sup>[32, 33]</sup>。（3）肌电图（EMG）辨音：我们也研究了用置

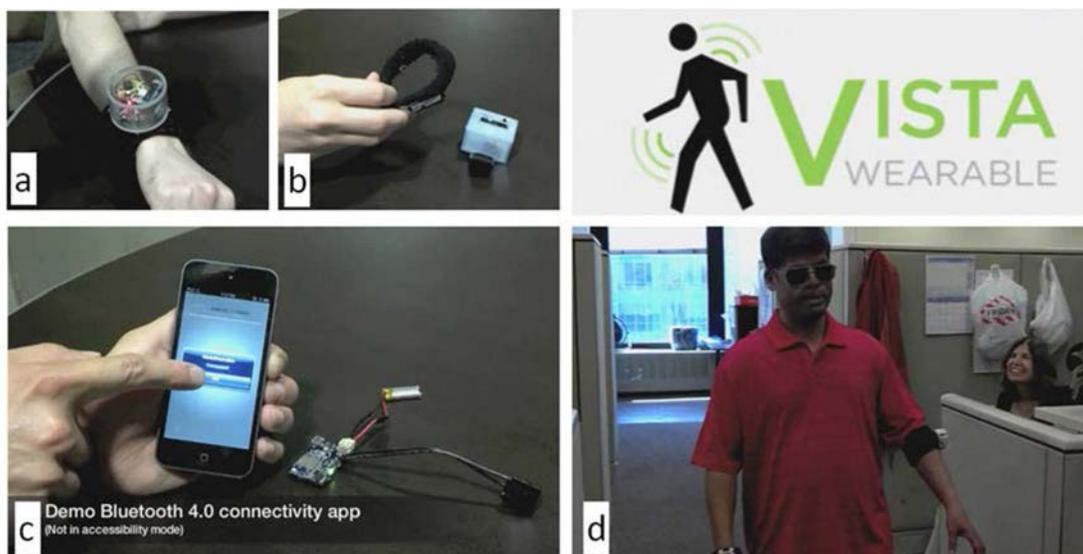


图 6 可穿戴全身目的原型设计<sup>[31]</sup>。a. 第一代圆形设计； b. 第二代扁形设计； c. 第三代智能手机蓝牙遥控设计； d. 一用户试戴。

于面部的多个 EMG 电极和深度 Bayesian 网络 (DBN) 来试验可否识别讲话者 (图 8)。从一个 20 余人的实验中, 我们发现, 讲话者只讲一个英文单字, 用 EMG 可以达到和用语音同样的识别率<sup>[34]</sup>。这个实验说明 EMG 信号采集了话者相关的信息, 将来可能为视障和其它残障用户提供感知帮助。我们正在进行同时用肌电图 (EMG) 和脑电图 (EEG), 来鉴别实验者的表情变化, 从而更好地认识人脑的工作机理。(4) 舌头“看”的测试。我们也和脑科学和感知专家合作, 测试了 Brainport 的舌头读图<sup>[14]</sup>, 虽然我们发现用其识别图案仍非常困难, 但也许可以用其实时指示用户行走的方向 (图 9)<sup>[35]</sup>, 从而可以空出用户的耳朵和手等, 以便其更自然的行走。

- 全身是目有预测, 皮肤能“看”见小说<sup>[27, 28]</sup>; 全身沉浸距离场, 振动强弱可表达<sup>[29-31]</sup>。

- 肌电信号来辨音, 深度学习做比较; 一字识得讲话者, 效果堪比用语音<sup>[34]</sup>。
- 舌头看图初测试, 识别图案待商榷; 或可巧用指方向, 空出手、耳做其它<sup>[35]</sup>。
- 群帮已成新时尚, 何不用其助导航? 用户视频传网上, 网友众目来群帮<sup>[32, 33]</sup>。

### 面部识别和深度学习

对于视障和自闭的用户, 除了满足他们基本的导航和指向需要外, 更进一步的机器视觉技术可帮助他们增强与他人交流的能力, 从而提高他们的生活质量。这可以作为辅助技术的第三个主要方面。其中, 人和人之间一个主要的交流需要是人脸识别和表情的识别。我们曾用 RGBD 传感器, 用彩色和三维信息为盲人提供人脸鉴别的功能 (图 10)<sup>[36]</sup>。我们随后更多的研究集中在人脸表情的识别上, 以便为视障和自

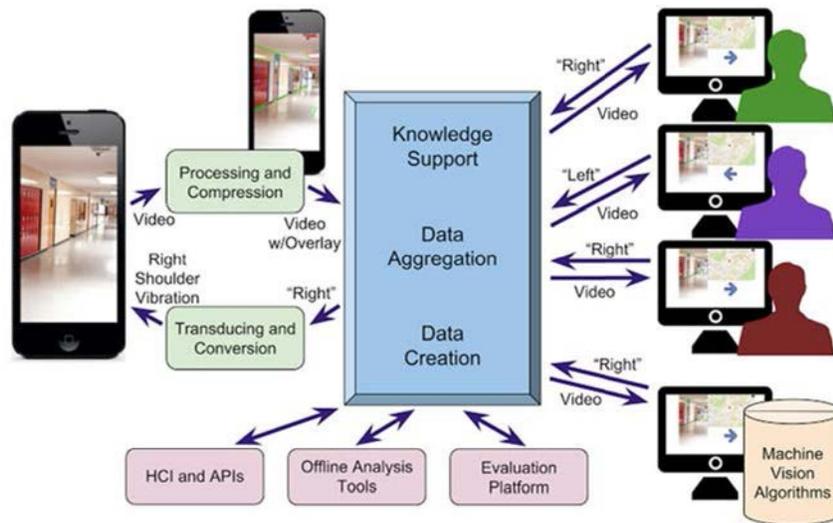


图 7 群帮眼的原理示意图<sup>[32]</sup>：视障用户的视频经预处理和压缩传往互联网，网上志愿者和机器算法提供实时行走指示，经处理系统过滤集成后，用声音或触觉等指示用户导航。收集的数据可供人机视觉研究用。

闭用户提供更有效的与他人交流的工具。比如，我们设计了一个智能手机 App，通过一个简单的用表情控制的游戏，训练自闭患者的表情识别和反馈（图 11）<sup>[37]</sup>。为了这些应用的有效性，我们进行了系列的研究：用主要的图像搜索引擎收集和清理出了一个有一万多张自然的人脸表情图像的数据库 CIFE <sup>[38]</sup>；基于该数据库训练了一个七类表情识别的深度卷积网络（CNN），以此为引擎，设计了一个简洁的表情可控

游戏，在一个月就收集到 100 多个游戏用户的超过一万五千张表情图像，生成了表情更接近日常生活、七个类别的图像数目也更平衡的 GaMo 表情图像库<sup>[39]</sup>；图像、声音、视频等多种模式融合、人工设计和

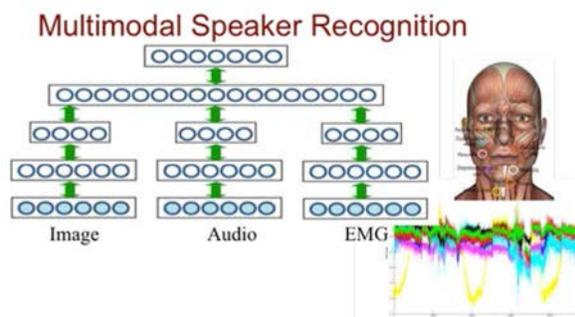


图 8. 多模式话者识别<sup>[34]</sup>：图像、语音和肌电图信号经由深度神经网络集成后，说话者讲一个字即可识别。

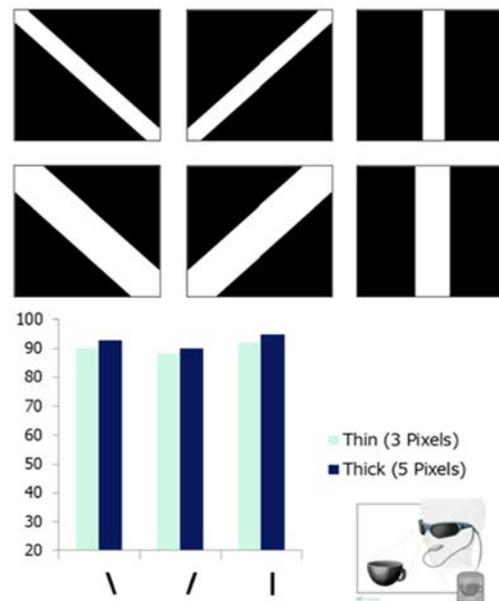


图 9 测试结果表明<sup>[35]</sup>，舌头看图较难用于形状识别。但是可以很好的辨别三个方向的线条，因此可用于导航指示。



图 10 实时三维人脸识别初步试验<sup>[36]</sup>。目标是用于视障用户的交互。

深度学习相结合的特征提取、和多核学习的方法，使得我们设计的算法在 2015 年的 ACM 表情识别挑战 (EmotiW) 的 60 个参赛中，名列第五<sup>[40]</sup>。为了进一步研究人脸表情的机理，我们研究了基于深度学习的表情单元 (Action Unit) 识别，设计了综合深度卷积学习转移 (Learning Transfer)，多层次兴趣编码 (Attention Coding) 和长短期记忆 (LSTM) 时空集成的模型，使我



图 11 帮助自闭症用户训练表情的智能手机游戏系统界面。

们的表情单元识别性能比目前最好的结果提高了 10% 以上<sup>[41, 42]</sup>。

- 喜怒哀乐写脸上，面部识别很重要；无奈视障不能见<sup>[36]</sup>，自闭症患者亦作难<sup>[37]</sup>。
- 深度学习识表情，网上收集众脸谱

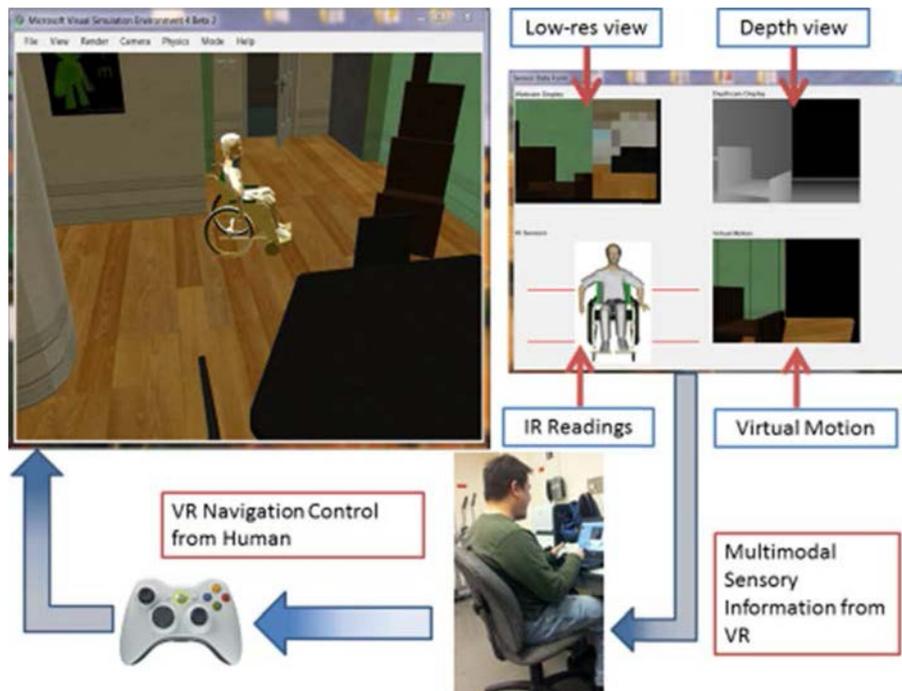


图 12 虚拟现实传感器仿真和导航测试。

[38]; 卷积网络为引擎, 设计游戏采数据 [39]。

- 多种模式来集成, 实战表情列前茅 [40]; 人机智能相结合, 多核学习最给力。
- 表情单元是核心, 学习转移省训练; 兴趣编码显效果 [41], 时空集成更提升 [42]。

### 测试和应用平台

辅助视觉和其它辅助技术的第四个主要方面是测试和应用平台。我们这里介绍两个平台: (1) GIVE-ME (Gamification In Virtual Environment for Multimodal Evaluation): 这是一个结合游戏化和虚拟现实的多模式测试和训练平台 [43, 44]。多种传感信息可实时仿真, 包括低分辨率图像, 深度图像, 红外单点测距, 运动目标模拟。传感信息经处理后, 转换成视障用户可感知的形式传给用户, 用户进而控制其替身

在虚拟环境中导航 (图 12)。(2) SAT-Hub (Smart and Accessible Transportation Hub): 这是一个由纽约城市大学(CUNY)、若歌大学 (Rutgers University) 和其它单位合作所提出的智能无障碍交通枢纽平台, 2016 年开始成为全球城市团队挑战 (GCTC) 的一个行动组 (Action Cluster)。关键技术包括: 移动定位, 人群分析, 三维语义模型, 用户定制导引。(图 13) [45-47]。

- 训练模拟在幻境, 灵活有趣无危险; 仿真各类传感器, 研究开发成本低。 [43, 44]。
- 其一模拟舌头看, 其二模拟群帮眼 [35, 33]; 其三模拟全身目, 其乐融融育博士 [31, 44]。
- 交通智能又可及, 城市、若歌双联合; 安全、交通兼服务, 集成一体利全民 [45-47]。
- 数字模型识环境, 人群分析知动态;

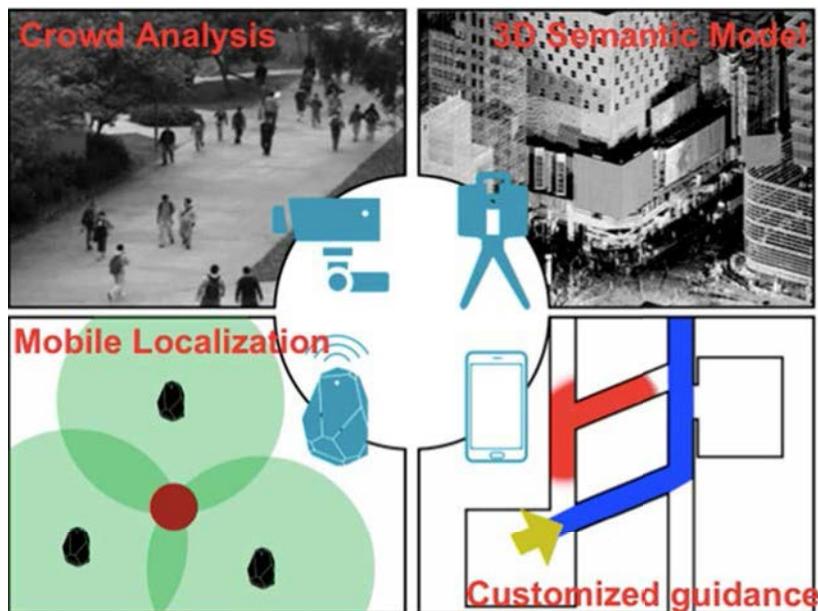


图 13 智能可及交通枢纽关键技术示意图。

多种传感定方位，户内室外均便行。

## 教育和培训

辅助技术的第五个主要方面是教育和培训。我们不仅培养博士生进行算法和系统的研究和开发，我们还开设了以下两个由 NSF 和工业界资助的教育和培训项目：

(1) 横跨三个专业——计算机科学、计算机工程和电子工程的辅助技术本科实习课（图 14）<sup>[9, 10, 47]</sup>。学生项目组不仅屡次获得创业奖，而且研究文章亦频出<sup>[17, 29, 48-53]</sup>。

(2) 主要针对少数群体、但招收从社区大学教授、大学生、高中老师到高中生的研究经历和指导（REM）项目，历经四年，训练学员 70 余人（图 15）<sup>[8, 53]</sup>。



图 14 作者（右）、Green Jacket 项目组的学生和 NYSID 总裁 Ron Romano（左）在纽约州首府 Albany 的 2016 年度 CREATE 展示会<sup>[51]</sup>

- 本科联合实战课，成功运作已六载；电子、计算和工程，培养学生逾百人<sup>[48]</sup>。
- 启于技术帮视障，智能生活渐顾及；始于用户真需要，学于实践有动力。
- 研究指导新尝试，少数群体得帮助<sup>[53]</sup>；四年培训愈七十，各种层次成果出<sup>[37, 39, 54]</sup>。

- 暑期十周密培训，日日小组细研讨；每周大组做报告，来春全国年会秀<sup>[55]</sup>。

## 结束语

人机感知和辅助技术研究，我们有三方面的思考。首先，研究的意义，我们强调科学价值、社会效益及学生培训的结合。我们认为，视觉甚精细，人脑更神奇；其一若有障，生活成问题。科研要联合，辅助需科技；多方来资助，用户亦参与。高校做研究，启发新创意；培养有用才，工作有意义。

其次，我们研究的路线，采取跨学科的联合和观念的突破：创新最前沿，机脑



图 15 2015 年 NSF 研究经历指导（REM）梯队。指导者包括项目负责人（教授）和博士生；学员包括：社区大学教授，本科生，高中老师和高三学生。

并生理；感知加计算，移动联云际。幻境与游戏，皆能更有益；不拘旧观念，感知互代替。导航到户内，深度来学习；交通在改进，智能又可及。

最后，我们在技术的各个点上，也追求有创新：双目加智能，单目可测距；全向有优势，三维更迅捷。设计全身目，尝试群帮眼；肌电能辨声，舌头“看”图案。



- Li, J. Xiao and Z. Zhu, CCNY Smart Cane, IEEE-CYBER 2017-The 7th Annual IEEE Int. Conf. on CYBER Technology in Automation, Control, and Intelligent Systems, July 31 – August 4, 2017, Hawaii, USA
- [18] H. Feng, N. Tsering, H. Tang, and Z. Zhu. Indoor Localization for the Visually Impaired Using a 3D Sensor. *Journal on Technology and Persons with Disabilities*. Vol. 4. Nov. 2016
- [19] F. Hu, H. Tang, A. Tsema, Z. Zhu, "Computer Vision for Assistive Indoor Localization", *Assistive Computer Vision*, M. Leo and G. Farinella (Eds.), the *Computer Vision and Pattern Recognition Series*, Elsevier, to appear in 2017 (invited)
- [20] H. Tang, N. Tsering, F. Hu, and Z. Zhu. Automatic Pre-Journey Indoor Map Generation Using AutoCAD Floor Plan. *Journal on Technology and Persons with Disabilities*. Vol. 4. Oct. 2016
- [21] H. Tang, Z. Zhu, A Segmentation-based Stereovision Approach for Assisting Visually Impaired People, the 13th International Conference on Computers Helping People with Special Needs (ICCHP), July 11-13, 2012, Linz, Austria.
- [22] H. Tang and Z. Zhu, Content-Based 3D Mosaics for Representing Videos of Dynamic Urban Scenes, *IEEE Transactions on Circuits and Systems for Video Technology*, 22(2), 295-308, 2012. DOI: 10.1109/TCSVT.2011.2178729
- [23] H. Tang, T. Ro, Z. Zhu, Smart Sampling And Transducing 3d Scenes For The Visually Impaired, *IEEE International Conference on Multimedia and Expo (ICME)*, July 15 to 19, 2013.
- [24] H. Tang, Z. Zhu, M. Vincent, T. Ro, From RGB-D to Low-Resolution Tactile: Smart Sampling And Early Testing, *IEEE Workshop on Multimodal and Alternative Perception for Visually Impaired People (MAP4VIP)*, in conjunction with *IEEE International Conference on Multimedia and Expo (ICME)*, San Jose, California, USA, July 15 - 19, 2013
- [25] G. Olmschenk, Z. Zhu, 3D Hallway Modeling Using A Single Image. The Fourth IEEE Workshop on Mobile Vision, in conjunction with CVPR2014
- [26] G. Olmschenk, Z. Zhu. Mobile Real-Time Single Image 3D Corridor Reconstruction Using J-Linkage. The 14th IAPR Conference on Machine Vision Applications (MVA 2015), Tokyo, May 18-22, 2015
- [27] John Apostle, Revelation, 4:8. Holy Bible, New International Version, ©2011, Biblica Inc.
- [28] Steven Millhauser, *The Wizard of West Orange*, Alfred A. Knopf, 2008
- [29] F. Palmer, Z. Zhu, T. Ro, Wearable Range-Vibrotactile Field: Design and Evaluation, the 13th International Conference on Computers Helping People with Special Needs (ICCHP), July 11-13, 2012, Linz, Austria.
- [30] W. L. Khoo, F. Plamer, J. Knapp, Z. Zhu, T. Ro, Designing and Testing Wearable Range-Vibrotactile Devices, *Journal of Assistive Technologies*, Vol. 7 Iss: 2, 2013, pp.102 – 117
- [31] E. Molina, W. L. Khoo, F. Palmer, L. Ai, T. Ro, Z. Zhu, Vista Wearable: Seeing through Whole-Body Touch without Contact, The 12th IEEE International Conference on Ubiquitous Intelligence and Computing (UIC 2015), August 10-14, 2015, Beijing China
- [32] G. Olmschenk, C. Yang, Z. Zhu, H. Tong and W. H. Seiple, Mobile Crowd Assisted Navigation for the Visually Impaired. The 12th IEEE International Conference on Ubiquitous Intelligence and Computing (UIC 2015), August 10-14, 2015, Beijing China
- [33] W. L. Khoo, G. Olmschenk, Z. Zhu and T. Ro, Evaluating Crowd Sourced Navigation for the Visually Impaired in a Virtual Environment. *IEEE 4th International Conference on Mobile Services (MS 2015)*, June 27 - July 2, 2015, New York, USA
- [34] F. Abtahi, W. Li, Z. Zhu, T. Ro, Multimodal Speaker Recognition using Deep Belief Networks, *CVPR Women in Computer Vision Workshop*, 2015
- [35] M. Vincent, H. Tang, W. L. Khoo, Z. Zhu, T. Ro. Shape Discrimination using the Tongue: Implications for a Visual-to-Tactile Sensory Substitution Device. *Multisensory Research*, 29(8), October 2016, pp. 773-798. DOI: 10.1163/22134808-00002542

- [36] W. Li, X. Li, M. Goldberg and Z. Zhu, Face Recognition by 3D Registration for the Visually Impaired Using a RGB-D Sensor, Second Workshop on Assistive Computer Vision and Robotics, Zurich, Sept 12, 2014
- [37] C. Tsangouri, W. Li, Z. Zhu, F. Abtahi and T. Ro, An Interactive Facial-Expression Training Platform for Individuals with Autism Spectrum Disorder, 2016 IEEE MIT Undergraduate Research Technology Conference (URTC), November 4-6, 2016 at MIT, Cambridge USA.
- [38] W. Li, Z. Su, M. Li, Z. Zhu. A Deep-Learning Approach to Facial Expression Recognition with Candid Images. The 14th IAPR Conference on Machine Vision Applications (MVA 2015), Tokyo, May 18-22, 2015
- [39] W. Li, F. Abtahi, C. Tsangouri and Z. Zhu, Towards an "In-the-Wild" Emotion Dataset Using a Game-based Framework. In Affect in the Wild Workshop, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR'16) Workshops (CVPRW).
- [40] W. Li, F. Abtahi, Z. Zhu, A Deep Feature Based Multi-kernel Learning Approach for Video Emotion Recognition, Emotion Recognition in the Wild (EmotiW) Challenge 2015, the 17th ACM International Conference on Multimodal Interaction (ICMI 2015), Seattle, USA. November 9-13th, 2015 (Top 5 in terms of performance among 60 submissions)
- [41] W. Li, F. Abtahi, Z. Zhu, L. Yin. EAC-Net: A Region-based Deep Enhancing and Cropping Approach for Facial Action Unit Detection. The 12th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2017), May 30 –June 3, 2017 in Washington, DC.
- [42] W. Li, F. Abtahi, Z. Zhu. Action Unit Detection with Region Adaptation, Multi-labeling Learning and Optimal Temporal Fusing, IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017), July 21-26, 2017, Honolulu, Hawaii, USA
- [43] W. L. Khoo, E. L. Seidel and Z. Zhu, Evaluating Multimodal Sensing in Virtual Environment for Assisting the Visually Impaired, the 13th International Conference on Computers Helping People with Special Needs (ICCHP), July 11-13, 2012, Linz, Austria.
- [44] W. L. Khoo, "GIVE-ME: Gamification In Virtual Environments for Multimodal Evaluation - A Framework," PhD Dissertation (advisor: Z. Zhu), Department of Computer Science, The CUNY Graduate Center, 2016.
- [45] J. Gong, C. Feeley, H. Tang, G. Olmschenk, V. Nair, Y. Yu, Z. Zhou, K. Yamamoto, and Z. Zhu. Building Smart Transportation Hubs with Internet of Things to Improve Services to People with Special Needs. Transportation Research Board (TRB) 96th Annual Meeting, January 8-12, 2017
- [46] Zhu, Z., Gong, J., Tang, H., Feeley, C., Olmschenk, G., Nair, V. and Lysicatos, M. (2016) Towards Smart Transportation Hub: Services to Persons with Special Needs Requiring Minimal New Infrastructure, GCTC Expo, June 13-14, 2016 (YouTube video: [https://youtu.be/\\_hxdXqTVG7I](https://youtu.be/_hxdXqTVG7I))
- [47] Z. Zhu, C. Kelnhofer-Feeley, J. Gong, W. Seiple, C. Brakewood. SCC-Planning: Towards Smart and Accessible Transportation Hub - Research Capacity Building and Community Engagement. NSF Award #1737533, Smart and Connected Community (S&CC) Program, Division Of Computer and Network Systems (CNS), CISE (Program Manager: Ralph Wachter). 2017-2018. [https://www.nsf.gov/awardsearch/showAward?AWD\\_ID=1737533](https://www.nsf.gov/awardsearch/showAward?AWD_ID=1737533)
- [48] Z. Zhu and J. Xiao, CCNY Joint Senior Design Program in Assistive Technology Across the Department Boundaries. The 2017 IEEE Integrated STEM Education Conference (ISEC'17), March 11, 2017, Princeton, New Jersey
- [49] F. Palmer, D. Zuleta, C. Rodriguez, J. Montesino and L. Ai (Mentors: Z. Zhu and T. Ro), V.I.S.T.A.: Vibro-Tactile Intelligent System for Travelling Aid - A Wearable Alternative. First Place Kaylie Prize of the 2012 CCNY Kaylie Entrepreneurship Competition.
- [50] L. Disla, S. Liu, O. Naeem, T. Panna. Mentors: Z. Zhu and H. Tang. GesTherapy: Using Gesture Based Therapy to Remodel Physical Rehabilitation. Second Place Dean's Prize of

the 2013 CCNY Kaylie Entrepreneurship Competition.

- [51] B. Tan, B. Xu, N. Yao and X.-B. Zhao. Mentors: J. Xiao and Z. Zhu. HAST: Health and Support Technology for Families using Open Source Wearable Devices. 2015 Social Innovation Prize Winner, Zahn Innovation Center.
- [52] A. Khan, J. Lopez, F. Moideen, W. L. Khoo and Z. Zhu, KinDectect: Kinect Detecting Objects, the 13th International Conference on Computers Helping People with Special Needs (ICCHP), July 11-13, 2012, Linz, Austria.
- [53] Z. Zhu, W. L. Khoo, C. Santistevan, Y. Gosser, E. Molina, H. Tang, T. Ro and Y. Tian, EFRI-REM at CCNY: Research Experience and Mentoring for Underrepresented Groups in Cross-disciplinary Research on Assistive Technology. The 6<sup>th</sup> IEEE Integrated STEM Education Conference (ISEC), March 6, 2016, Princeton, New Jersey.
- [54] C.-S. Lin, Z. Zhu, and T. Ro, Dynamic Project-based STEM Curriculum Model for a Small Humanities High School. The 2016 IEEE Integrated STEM Education Conference (ISEC), March 5, 2016, Princeton, New Jersey.
- [55] ERN – Emerging Researchers National Conference in STEM, <http://new.emerging-researchers.org/>

## 作者简介



**Dr. Zhigang Zhu** is currently a Herbert G. Kayser Chair Professor of Computer Science, at The City College and The Graduate Center, The City University of New York. He is Director of the City College Visual

Computing Laboratory (CvcL), and Co-Director of the Center for Perceptual Robotics, Intelligent Sensors and Machines (PRISM) at CCNY. Previously he was Associate Professor at Tsinghua

University, and a Senior Research Fellow at the University of Massachusetts, Amherst. His research interests include 3D computer vision, multimodal sensing, virtual/augmented reality, video representation, and various applications in assistive technology, education, environment, robotics, surveillance and transportation. He has published over 160 technical papers in the related fields. His research at CCNY has been supported by AFOSR, AFRL, ARO, DARPA, DHS, NSF, NY State as well as industry. He is an Associate Editor of the Machine Vision Applications Journal, Springer, and IFAC Mechatronics Journal, Elsevier. He was a Technical Editor of the ASME/IEEE Transactions on Mechatronics from 2010 to 2014.

**朱志刚博士**，现为美国纽约城市大学终身教授，赫伯特·G·凯泽 (Herbert G. Kayser) 计算机科学讲席教授。曾任清华大学计算机系副教授，美国安城麻州大学计算机系资深研究员。先后主持参与 30 余项科研项目，由美国国家科学基金、国防部、国安部以及在美工业界，和中国自然科学基金、863 计划等资助。发表论文 160 余篇，专著和译著 3 部，在世界各地做特约或专题报告 40 余次，研究的领域集中在人机视觉和多模态感知。现任计算机视觉期刊 Machine Vision and Applications 和 IFAC 机电一体化期刊副主编，曾任 ASME/IEEE 机电一体化汇刊 (Transactions on Mechatronics) 学术编辑。曾有幸分别与计算机视觉领域的两位泰斗 Thomas S. Huang 和 Takeo Kanade 合作编著《多模式监视：传感器、算法和系统》(Multimodal Surveillance: Sensors, Algorithms and Systems) 一书和《International Journal of Computer Vision》大型三维景物建模特辑。

# A Wearable Robotic Object Manipulation Aid for the Visually Impaired

## 帮助视障者抓取物体的手戴式机器人辅助工具

Cang Ye

叶苍

University of Arkansas at Little Rock, Little Rock, AR 72204

cxye@ualr.edu

阿肯色大学小石城分校, 小石城, 阿肯色州 72204, 美国

E-mail: cxye@ualr.edu

### Abstract

This paper presents a new robotic assistive device for the visually impaired. The device is called Wearable Robotic Object Manipulation Aid (W-ROMA). W-ROMA is a hand-worn assistive device with 3D computer vision and human-robot collaboration functions. It analyzes RGB-D data from an Intel Realsense F200 camera for Simultaneous Localization and Mapping, as well as object recognition. Based on the results, it computes the misalignment between the hand and the target object and determines the desired hand movement for object alignment. By using a tactile and a speech interfaces, W-ROMA guides the user's hand to approach the target object while retaining the alignment and eventually take a hold of the object. The guidance provided by the tactile interface is natural for the user to follow. W-ROMA may be used to assist a visually impaired person in grasping a small object (e.g., door/drawer

handle, toothpaste, medication bottle, soft-drink bottle) for wayfinding or other purpose in the daily life. Its application will contribute to independent living of the visually impaired and improve their quality of lives.

### 摘要

本文介绍一款新型的视障辅助工具。该工具是一个具有三维计算机视觉及人机合作功能的手戴式机器人装置。它通过分析 RGB-D 图像数据序列实现同步定位与场景构建以及物体识别, 并以此计算出对准目标物体所需的手部运动。然后, 它通过语音及触觉界面引导视障用户将手靠近目标物体并保持对准状态。该辅助工具可用于帮助视障者识别并抓握一些小的物体(比如门或抽屉把手, 杯子, 牙膏, 药瓶, 饮料瓶等日常物件), 从而解决辅助导盲中开门及其他日常生活中抓取物件的难题。该工具的使用有助于提高视障者的独立生活能力, 改善其生活质量。

**Keywords:** Simultaneous Localization

and Mapping, Object Recognition, Robotic Assistive Device

**关键词：**同步定位与场景构建, 物体识别, 机器人辅助工具。

## 1 概述

抓取物件是每个人生活中随时随地需要做的事情。这种对于健康人来说理所当然的事却是视障人士所面临的生活难题。对于视障者而言, 识别及抓取物件的能力不仅有助于他们的居家生活(如喝水, 吃药), 而且还能帮助他们的出门旅行(如开门, 挪开路障)。虽然计算机视觉及机器人技术已取得极大的进步, 目前仍然没有可行的辅助技术解决方案。视觉伺服(Visual Servoing, VS)技术可以用来帮助视障者实现手和目标物体的对准<sup>[1, 2]</sup>。这种机器人辅助视觉伺服系统(Robotic Assistive Visual Servoing, RAVS), 如果采用手戴式设计的话, 通常需要把摄像机安装在手的前端以便减少视野遮挡。其缺点是所用的摄像机不能提供可靠的距离信息以便引导手的动作。最近, 小型三维(3D)摄像机的出现为解决这个困境提供了可能。不过, 基于3D摄像机的辅助工具的开发必须解决以下技术难题: 1) 自然的, 以人为中心的人机界面(Human-Device Interface, HDI); 2) 基于有限计算资源的实时计算机视觉; 3) 紧凑轻便的设计以满足装置的便携性。其中, 人机界面是该辅助工具能否获得用户认同的根本。

为了解决这些技术挑战, 我们提出了一个双层的以人为中心的人机界面: 1) 一个手运动指示器(Hand Motion Indicator,

HMI)用来传达给用户手物对齐所需的手部运动; 2) 一个混合触觉显示系统(Hybrid Tactile Display, HTDS), 用来显示物体的二维形状。HMI的无机械部件设计和HTDS的单点双触觉阵列设计考虑了便携性。基于这个界面, 我们提出了一个新型的辅助工具—Wearable Object Manipulation Aid(W-ROMA)。W-ROMA使用单个RGB-D摄像机进行设备定位和物体检测使得系统紧凑轻巧。它通过基于高斯混合模型的物体识别方法实现了实时计算并减少存储空间的需求。

## 2 机器人辅助工具回顾

视障机器人辅助装置领域的研究主要集中在机器人导航辅助工具(Robotic Navigation Aid)上。现有的机器人导航辅助工具可以分为三类: 机器人轮椅<sup>[3, 4, 5]</sup>, 机器人导盲犬<sup>[6, 7]</sup>, 和电子导盲杖<sup>[8, 9, 10, 11]</sup>。机器人轮椅和机器人导盲犬本质上是常规的机器人系统。它们的计算机视觉系统通常使用多个传感器和强大的计算机来实现实时定位和物体检测。因此不能满足便携性的要求。这类技术不适用于可穿戴机器人装置。相比之下, 为电子导盲杖开发的一些方法可能适用于W-ROMA。在现有的电子导盲杖中, 诺丁汉障碍物检测器<sup>[8]</sup>和Navbelt<sup>[11]</sup>使用声纳进行障碍物检测。C-5激光手杖<sup>[9]</sup>使用三对激光二极管发射器和光电二极管接收器进行三角测量法测距。“虚拟盲杖”(“virtual white cane”)<sup>[10]</sup>则通过使用单点激光和小型化相机的基线三角测量系统测量距离。用户通过摆动设备来实现多点测量。尽管它们具有便携性的优点, 因使用过于简单的传感器这类装

置为用户提供的信息不足以支持物体抓取。本文作者最近开发了一种名为 Co-Robotic Cane (CRC)<sup>[12]</sup> 的机器人导盲杖。CRC 使用单个 3D 相机进行设备位姿 (pose) 估计和物体识别。位置信息及检测到的结构性物体 (如楼梯, 门等) 被用来帮助视障者在室内寻路 (wayfinding)。安装在其前端的滚轮可以实现自动转向以便引导用户的行进方向。

早期的视觉伺服系统<sup>[2, 13]</sup>用安装于手前端的摄像机的反馈来控制机器人的运动。由于缺乏 3D 物体位置及姿态信息, 它们无法实现有效的物体抓取。传统的运动恢复结构 (Structure from Motion)<sup>[14, 15]</sup> 技术可用来估计目标物体或机器人手的位姿。但是, 在我们的应用当中, 不大容易确定场

景的尺寸。基于立体视觉的视觉伺服<sup>[16, 17]</sup> 虽然不存在尺寸问题, 但却不能提供密集的 3D 点云来进行可靠的物体识别和抓握。因此, 目标物体和参考图像 (也称为模板图像) 之间的视觉特征匹配<sup>[16]</sup> 成了较为流行的方法。最近, RGB-D 摄像机在机器人视觉上得到了广泛的应用并且在物体抓取方面取得很大的成功<sup>[18, 19]</sup>。这主要归因于它能同时提供场景的彩色图像和深度信息从而避免了场景尺寸的不确定问题。而且它有体积小价格低等优点。不过, 现有的方法不能在 W-ROMA 上实现实时计算。另外, 它们产生的结果能否被 W-ROMA 用户利用则取决于是否有一个有效人机界面。

在可穿戴的视觉辅助工具领域, 文献<sup>[20, 21]</sup>和<sup>[22]</sup>分别介绍了一个头戴式和一个手持式系统, 各自用于从图像中识别物体和检测边缘。它们都使用音频反馈来引导用户接近目标或用振动信号 (由安装于手指的振动器产生) 来指示检测到的边缘。另外, 文献<sup>[23, 24]</sup>介绍了基于手机的手持式辅助装置, 分别用于帮助视障者识物<sup>[23]</sup>和寻路<sup>[24]</sup>。不过, 所有这些系统都不能解决辅助物体抓取的问题。据我们所知, W-ROMA 是第一个可用于帮助视障者抓取物体的机器人辅助工具。

### 3 W-ROMA 手戴式物体抓取辅助装置

如图 1 所示, W-ROMA 目前的设计采用了无指手套的形式。视障用户佩戴它用以解决物体识别与抓取任务。它主要由两个功能模块 (3D 计算机视觉系统和人机界面) 组成。它的计算机视觉系统 (包括一个小型计算机主板 Gumstix Overo® AirSTORM COM 和一个 RGB-D 相机) 检

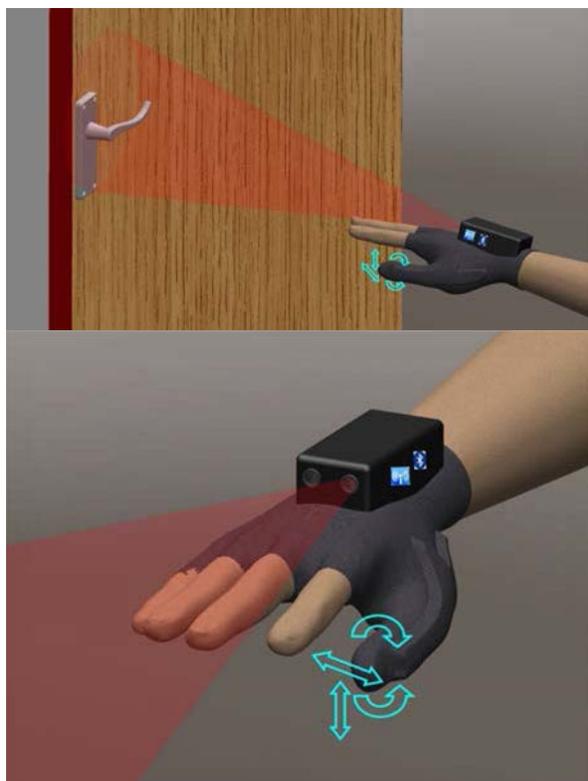


图 1 W-ROMA 手戴式物体抓取辅助装置示意图

测目标物体，计算手和物体的相对位姿，并确定手物对齐所需的手部运动（Desired Hand Movement, DHM）。然后，通过人机界面（HDI）把物体形状和所需的手部运动传送给用户。HDI 由 HMI, HTDS 及蓝牙耳机组成。HMI 通过移动拇指，以直观的方式向用户提供手部运动导引。（拇指移动/旋转的方向即为手部应该移动/旋转的方向。）同时 HTDS 通过触觉感知（在用户的手掌皮肤上显示触觉图像）向用户传达物体信息。蓝牙耳机为用户提供了一个语音界面。W-ROMA 的典型用例是：用户通过语音界面指明需要抓取的目标物体。W-ROMA 检测目标物体并用 HTDS 和 HMI 向用户传递物体形状和 DHM。基于此，用户移动其手部与目标对齐，并调整手的位姿势和形状来靠近并最终抓取物体。在此

以前，我们将用置于手端的两个震动器来指示用户向前（如前面震动器震动）或向后（如后面震动器震动）移动其手部。移动速度可以通过震动的快慢来指示。

Gumstix 的 WiFi 连接使得 W-ROMA 的资源感知计算（resource-aware computing）成为可能。如果在网络上检测到服务器计算机，Gumstix 将简单地把 RGB-D 数据中继到服务器进行处理（同步定位与场景构建以及物体识别）。然后将结果发传回 Gumstix。否则，Gumstix 将只对一帧 RGB-D 数据进行物体检测。这种计算方案是由低功耗需求和 Gumstix 有限的计算能力所决定的。

W-ROMA 的视觉传感器是一个 Intel Realsense F200 RGB-D 相机板（图 2）。该

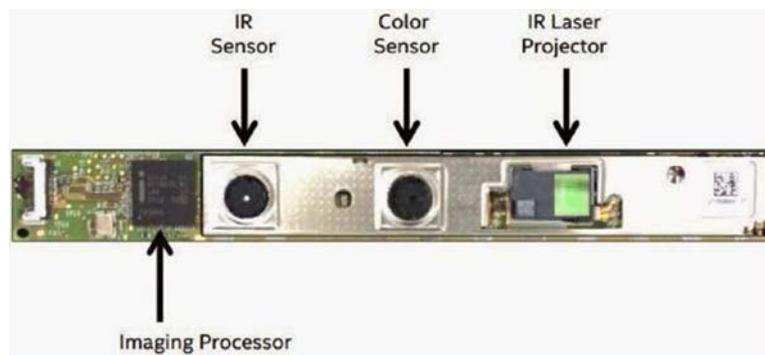


图 2 W-ROMA 示意图

过程，蓝牙耳机可以不断向用户提供目标物体的相关信息。W-ROMA 可以用作正在开发中的 CRC 机器人导盲杖的配套设备。在这种情况下，CRC 从远处检测到目标物体，引导视障用户物接近该物体，然后将抓取任务交给 W-ROMA 来实现。HTDS 的研发由内华达大学的申延涛博士承担。有关 HTDS 的具体内容从略。在 HTDS 完成

相机可同时提供彩色图像（1920×1080，30 帧/秒）和深度（Depth）数据（640×480，60 帧/秒）。其测距有效范围为 0.2-1.2 米。它由 USB 接口供电。

#### 4 W-ROMA 位姿估计

W-ROMA 的位姿（pose）包括其位置（X, Y, Z 坐标）和姿势或朝向（由三个

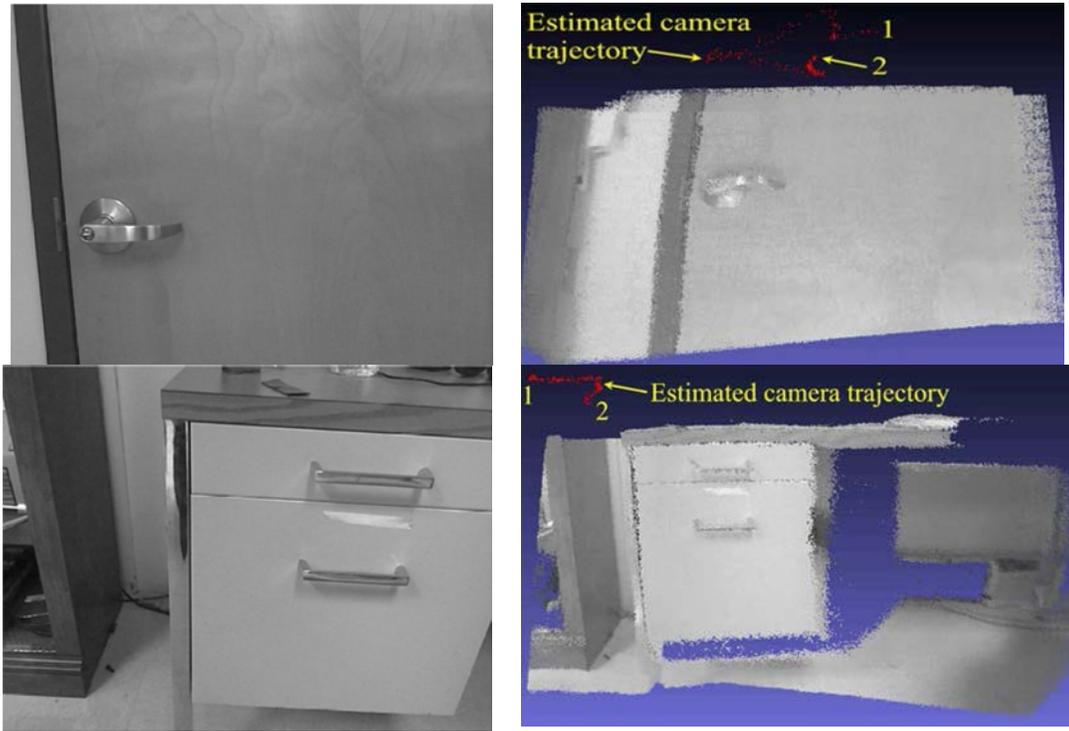


图 3 Pose-graph SLAM 产生的实验场景的点云数据。左图：灰度图像，右图：点云及估计的相机轨迹，1：相机起始点，2：相机终止点。

欧拉角描述)。该六自由度的位姿可用于三维绘图 (3D mapping) 及计算手和目标物体的偏离度。在先前的工作中，我们为移动机器人开发了一种基于 3D 闪光激光雷达相机 (flash LIDAR camera) 的位姿变化估计方法，称为 Visual-Range Odometry (VRO)<sup>[25]</sup>。最近，我们在 Realsense F200 上实现了 VRO 算法并取得了更好的测量准确度和精度。VRO 提取并匹配每两个连续图像中的视觉特征。然后，用匹配特征的 3D 坐标来计算相机的旋转和平移矩阵并由此算出两个图像帧之间相机位姿的变化。这个最小二乘拟合问题是通过奇异值分解 (Singular Value Decomposition) 方法来解决的。视觉特征匹配过程产生的异常值

(不正确的匹配) 是通过随机抽样一致 (Random Sample Consensus: RANSAC) 方法来去除的。

有了相机位姿的变化，我们就可以构建一个位姿图 (pose-graph)。该图的每个节点 (node) 代表相机的位姿，每个边 (edge) 代表位姿测量的误差及其协方差。通过图优化使图的总误差最小化，我们可以得到相机在各个点的准确位姿。用此信息我们可以把相机在各个点的 3D 点集中在一起构成一幅大的点云图 (见图 3)。这种位姿计算方法叫 pose-graph 同步定位与场景构建 (Simultaneous Localization And Mapping, SLAM)。

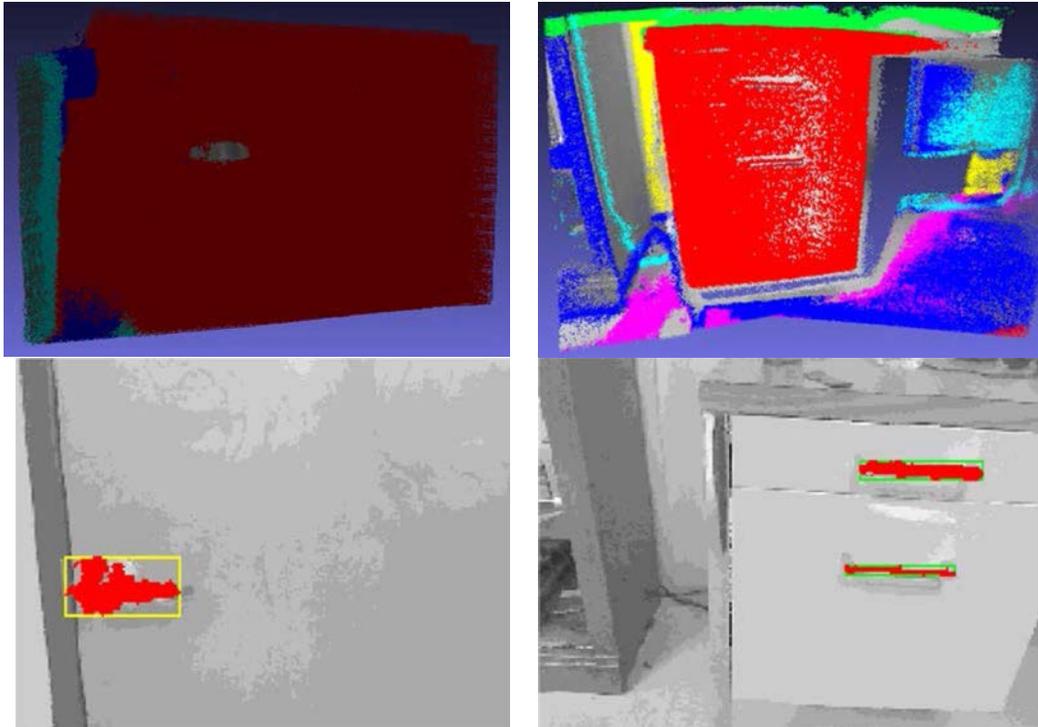


图 4 GMM 方法识别门和抽屉把手。上图：从点云中提取的平面；  
下图：检测到的门和抽屉把手被标记在图像上。

## 5 物体识别

由于 W-ROMA 主要用于室内环境。而室内环境通常会有较大的平面。我们所提出的物体识别算法从相机的 3D 点数据中提取这些平面并在图像中把相关图像区域去除，从而得到一系列隔离的图像块。然后，该方法从这些图像块中提取 SIFT (scale-invariant feature transform) 特征，并使用 SIFT 特征来识别目标物体。由于 SIFT 特征提取是在较小的图像块上进行的，该方法大大减少了计算时间。另外，为了解决基于模板图像视觉特征匹配算法带来的较高计算成本和存储空间需求的缺点，我们提出了一种基于高斯混合模型 (Gaussian Mixture Model, GMM) 的物体识别方法。用于识别特定类型物体的 GMM 将由包含

相同类型物体的图像块进行训练。从而避免了存储大量模板图像的视觉特征的需要。另外，物体识别是通过计算 GMM 对输入的 SIFT 特征的响应来实现的。由于 GMM 的单层并行结构以及较少量的 SIFT 特征 (来自较小的图像块) 输入，物体识别只需很少的时间。从而解决了 W-ROMA 有限资源条件下的实时计算的问题。

我们从各个不同的视点对相关的物体 (包括门把手，抽屉把手，杯子，胶水瓶，药瓶，牙膏) 进行大量的 RGB-D 数据采集。然后从点云数据中提取平面并生成包含目标物体的图像块。这些图像块的 SIFT 特征被用来训练六个 GMM 物体识别器。训练结束后，我们对这些物体重新采集数据。然后，用这些数据来验证所提的物体识别

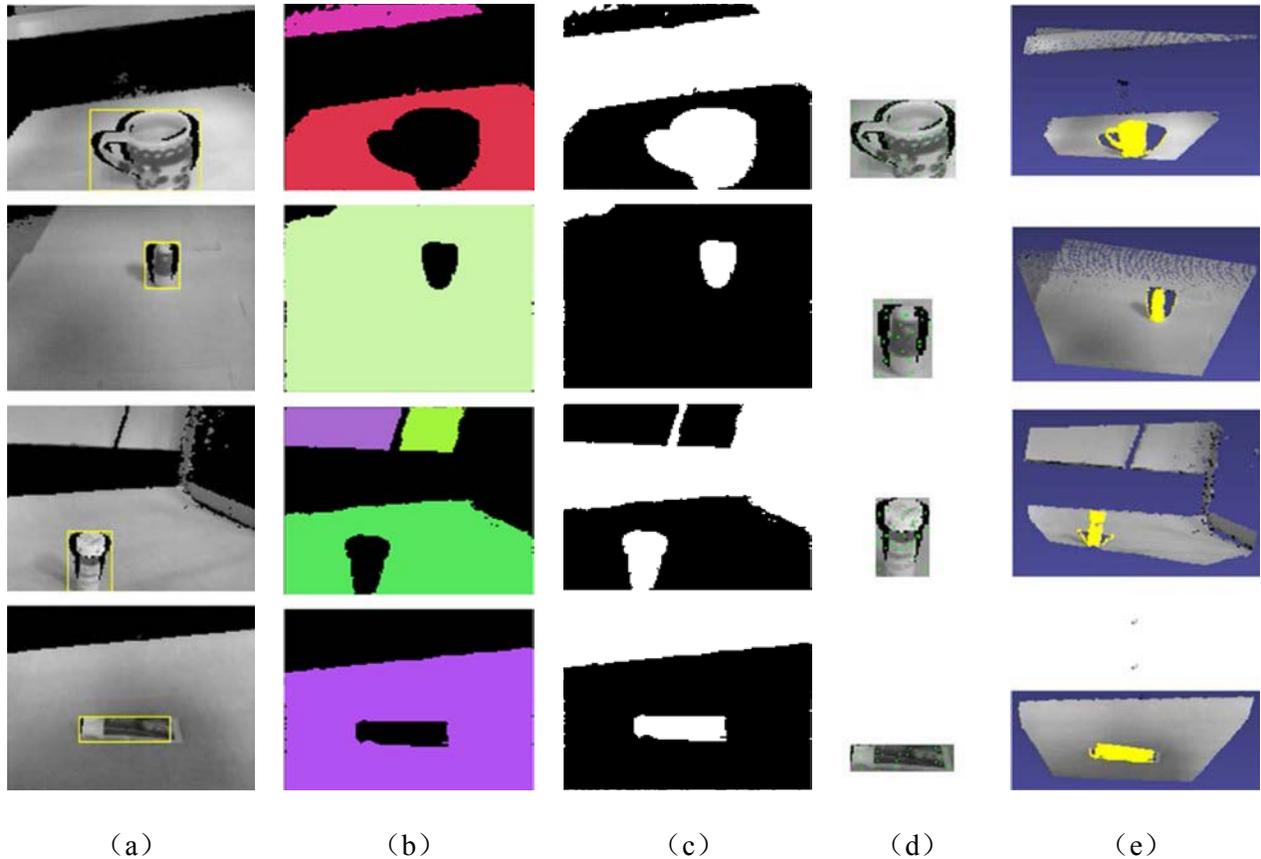


图 5 基于 GMM 方法的其它小物体识别。(a) 灰度图像；(b) 平面提取，在距离图像 (range image) 表示；(c) 白区代表图像块的区域；(d) 图像块及 SIFT 特征；(e) 检测到的物体被标注在点云图上。

方法（包含六个物体识别器）。实验结果表明该方法可以准确，实时地识别出目标物体。图 4 显示门和抽屉把手的识别结果。图 5 显示其它物体的识别结果。

## 6 总结

文本介绍一个新型的手戴式机器人辅助工具 W-ROMA。它可以用来帮助视障人士检测目标物体并通过触觉及语音界面来引导用户的手部运动使手靠近并对齐该物体。最终，完成物体的抓取。该装置采用一个 RGB-D 相机作为视觉传感器。通过处理该相机的图像和距离数据，W-ROMA 实

现了同步定位与场景构建以及物体识别。本文介绍的基于 GMM 的物体识别方法具有计算时间短，所需内存少等优点。从而，有效地解决了 W-ROMA 的有限计算资源问题。实验结果验证了上述定位及物体识别方法的有效性。

## 7 致谢

本研究获美国国立卫生研究院项目（编号：R01EY-026275）资助。

## 参考文献

- [1] G. J. Agin, "Real Time Control of a Robot with a Mobile Camera," Technical Note 179, SRI International, Feb. 1979.
- [2] S. A. Hutchinson, G. D. Hager, and P. I. Corke, "A tutorial on visual servo control," *IEEE Transaction on Robotics and Automation*, vol. 12, no. 5, pp. 651-670, 1996.
- [3] S. P. Levine, D. A. Bell, L. A. Jaros, R. C. Simpson, Y. Koren, and J. Borenstein, "The Navchair Assistive Wheelchair Navigation System," *IEEE Transactions on Rehabilitation Engineering*, vol. 7, no. 4, pp. 443-451, 1999.
- [4] C.-H. Kuo and H. W. Chen, "Human-Oriented Design of Autonomous Navigation Assisted Robotic Wheelchair for Indoor Environments," in *Proc. IEEE International Conference on Mechatronics*, 2006, pp. 230-235.
- [5] C. Galindo, J. Gonzalez, and J.-A. Fernández-Madrigal, "Control Architecture for Human-Robot Integration: Application to a Robotic Wheelchair," *IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics*, vol. 36, no. 5, pp. 1053-1067, 2006.
- [6] Ulrich, I. and Borenstein, J., 2001, "The GuideCane—Applying Mobile Robot Technologies to Assist the Visually Impaired," *IEEE Transaction on Systems, Man, and Cybernetics-Part A: Systems and Humans*, Vol. 31, No. 2, pp. 131-136, 2001.
- [7] V. Kulyukin, C. Gharpure, J. Nicholson, and G. Osborne, "Robot-Assisted Wayfinding for the Visually Impaired in Structured Indoor Environments," *Autonomous Robots*, vol. 21, no. 1, pp. 29-41, 2006.
- [8] D. Bissit and A. Heyes, "An Application of Biofeedback in the Rehabilitation of the Blind," *Applied Ergonomics*, vol. 11, no. 1, pp. 31-33, 1980.
- [9] J. M. Benjamin, N. A. Ali, and A. F. Schepis, "A Laser Cane for the Blind," in *Proc. San Diego Medical Symposium*, 1973.
- [10] D. Yuan and R. Manduchi, "A Tool for Range Sensing and Environment Discovery for the Blind," in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 2004.
- [11] S. Shoal, J. Borenstein, and Y. Koren, "Auditory guidance with the NavBelt – a computerized travel aid for the blind," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 28, no. 3, pp. 459-467, 1998.
- [12] C. Ye, S. Hong, and X. Qian, "A Co-Robotic Cane for Blind Navigation," in *Proc. IEEE International Conference on Systems, Man, and Cybernetics*, 2014, pp. 1082-1087.
- [13] C. E. Smith and N. P. Papanikolopoulos, "Vision-Guided Robotic Grasping: Issues and Experiments," *Proc. IEEE International Conference on Robotics and Automation*, 1996, pp. 3203-3208.
- [14] Z. Zhang, "Estimating Motion and Structure from Correspondences of Line Segments between Two Perspective Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 2, pp. 1129-1139, 1995.
- [15] A. Davison, I. Reid, N. Molton, and O. Stasse, "MonoSLAM: Realtime single camera SLAM," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 6, pp. 1052–1067, 2007.
- [16] D.-J. Kim, R. Lovelett, and A., Behal, "Eye-in-hand stereo visual servoing of an assistive robot arm in unstructured environments," *Proc. IEEE International Conference on Robotics and Automation*, 2009, pp. 2326-2331.
- [17] P. Allen, A. Timcenko, B. Yoshimi, and P. Michelman, "Automated tracking and grasping of a moving object with a robotic hand-eye system," *IEEE Transactions on Robotics and Automation*, vol. 9, no. 2, pp. 152-165, 1993.
- [18] Y. Jiang, S. Moseson, and A. Saxena, "Efficient grasping from RGBD images: Learning using a new rectangle representation," *Proc. IEEE International Conference on Robotics and Automation*, 2011, pp. 3304-3311.
- [19] A. Anand, H. S. Koppula, T. Joachims, and A. Saxena, "Contextually Guided Semantic

- Labeling and Search for Three-Dimensional Point Clouds,” *The International Journal of Robotics Research*, vol. 32, no. 1, pp. 19-34, 2012.
- [20] K. Thakoor, S. Marat, P. Nasiatka, B. McIntosh, F. Sahin, A. Tanguay, J. Weiland, and L. Itti, “Attention biased speeded up robust featureS (AB-SURF): A neurally-inspired object recognition algorithm for a wearable aid for the visually-impaired,” *Proc. IEEE International Conference on Multimedia and Expo Workshops*, 2013, pp. 1-6.
- [21] B. F. G. Katz, S. Kammoun, G. Parseihian, O. Gutierrez, A. Brilhault, M. Auvray, P. Truillet, M. Denis, S. Thorpe, and C. Jouffrais “NAVIG: augmented reality guidance system for the visually impaired,” *Virtual Reality*, 16:253–269, 2012.
- [22] S. Horvath, J. Galeotti, B. Wu, R. Klatzky, M. Siegel, and G. Stetten, “FingerSight: Fingertip Haptic Sensing of the Visual Environment,” *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 2, 2014.
- [23] Looktel Recognizer, “Looktel,” [online] 2009, <http://www.looktel.com/recognizer> (accessed June, 2017).
- [24] R. Tapu, B. Mocanu, A. Bursuc, T. Zaharia, “A Smartphone-Based Obstacle Detection and Classification System for Assisting Visually Impaired People,” *Proc. IEEE International Conference on Computer Vision Workshops*, 2013, pp. 444-451.
- [25] S. Hong, C. Ye, M. Bruch, and R. Halterman, “Performance Evaluation of a Pose Estimation Method Based on the SwissRanger SR4000,” *Proceedings of IEEE International Conference on Mechatronics and Automation*, 2012, pp. 499-504.

## 作者介绍

叶苍，美国阿肯色大学小石城分校系统工程系终身教授。创办并主持自治系统实验室。曾任密西根大学博士后及研究员。主要研究领域为自治机器人，计算机视觉和人工智能。

# A Robot-integrated Smart Home for Elderly Care

## 机器人和智能家居技术在智慧养老中的应用

Weihua Sheng<sup>1\*</sup>, Ha Manh Do<sup>1</sup>, Minh Pham<sup>1</sup>, Dan Yang<sup>2</sup>, Meiqin Liu<sup>3</sup>, Senlin Zhang<sup>3</sup>,  
Xianfeng Yuan<sup>4</sup>

盛卫华<sup>1\*</sup>, Ha Manh Do<sup>1</sup>, Minh Pham<sup>1</sup>, 杨丹<sup>2</sup>, 刘妹琴<sup>3</sup>, 张森林<sup>3</sup>, 袁宪锋

1. School of Electrical and Computer Engineering, Oklahoma State University,  
Stillwater, OK, 74078, USA;
2. College of Information Science and Engineering, Northeastern University, Shenyang 110819, China;
3. College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China
4. School of Control Science and Engineering, Shandong University, Jinan 250061, China

\*weihua.sheng@okstate.edu

1. 俄克拉荷马州立大学, 静水市, 俄克拉荷马州 74075, 美国;

Oklahoma State University, Stillwater, OK 74078, U.S.

2. 东北大学, 沈阳市, 辽宁省 110819, 中国;

Northeastern University, Shenyang, Liaoning Province 110819, China.

3. 浙江大学, 杭州市, 浙江省 310058, 中国;

Zhejiang University, Hangzhou, Zhejiang Province 310058, China.

4. 山东大学, 济南市, 山东省 250061, 中国;

Shandong University, Jinan, Shandong Province 250061, China.

\*联系人, E-mail: weihua.sheng@okstate.edu

### Abstract

This article presents the development of a robot-integrated smart home (RiSH) which can be used for research in assistive technologies for elderly care. The RiSH testbed consists of a smart home environment, cloud servers, remote caregivers, and an experimental infrastructure. The smart home environment integrates a home service robot, a home sensor network, a body sensor network, and a mobile device. An open and layered architecture is

proposed to guide the hardware and software design and implementation for the RiSH. We developed two low-level applications: 1) human localization and tracking by fusing wearable motion sensors and distributed binary sensors; 2) human activity recognition by combining distributed binary sensors and sound signals. Utilizing these two low-level applications, a high-level application is realized that detects and responds to human falls. Experimental results are provided to validate and evaluate the testbed and the

associated capabilities. This testbed can be employed in developing assistive technologies to the elderly who live independently at home

## 摘要

随着全球人口老龄化问题的日渐凸显,用于独居老人护理的机器人智能家居技术受到了越来越多的关注。本文设计了一种新颖的机器人智能家居实验平台,该平台主要由智能家居系统、云服务器、远程护理人员以及辅助实验设备构成,其中,智能家居系统主要包括家庭服务机器人、家庭传感器网络、穿戴传感器以及智能移动设备。本文提出了一种开放、分层的智能家居软件设计架构,利用穿戴传感器和分布式红外传感器实现了人员定位和跟踪,利用声音信号结合分布式红外传感器完成了日常活动的识别,在此基础上,提出了老人摔倒检测与应助方法。通过实验证明了本文所提方法架构及实验平台的有效性,本文研究内容为解决独居老人的护理问题提供了一种新思路。

**Keywords:** Smart Home, Elderly Care, Indoor Human Localization, Fall Detection, Sound Event Recognition

**关键词:** 智能家居, 家庭服务机器人, 老人护理, 室内人员定位, 摔倒检测, 声音事件识别

## 1 引言

近年来,全球老龄化速度不断加剧,2015年全球60岁以上人口约9亿,预计到2050年将增加至20亿<sup>[1]</sup>。虽然,养老院、老年护理中心等机构能够提供诸如健康监

护、生活护理等服务,但是这类机构中老年人的自由通常受到限制,相比而言,老年人更愿意在自己家里生活。智能家居和家庭服务机器人技术是解决上述问题最好的方式。

在现有的研究工作中,服务机器人已经开始被集成在智能家居系统中。文献<sup>[2]</sup>提出了一种远程医疗监护系统,该系统中索尼公司的AIBO机器人被用于采集现场声音图像信号。文献<sup>[3]</sup>提出了一种嵌入式智能架构,机器人本体传感器和环境中分布的传感器被视为统一、互联的感知系统。文献<sup>[4]</sup>介绍了一种泛在即插即用技术,通过该技术,不同的智能设备和监护机器人可以被集成到智能家居环境中。文献<sup>[5]</sup>介绍了一种基于云和机器人技术的健康护理系统,该系统能够提供通用的健康护理服务以及专门针对独居老人的精神健康服务。

将服务机器人技术与智能家居相结合的研究思路具有许多优势:一方面,机器人能够将计算复杂度高的任务转移至智能环境中去完成,降低机器人本体对计算能力的要求;另一方面,智能家居中的分布式传感器可以扩展机器人感知能力、降低机器人硬件成本。在集成机器人的智能家居(Robot-integrated Smart Home: RiSH)系统中,服务机器人可以利用智能家居中分布的传感器更好地与远程护理人员协作,为老年人提供更优质的护理服务。因此,RiSH是实现老年人居家护理最理想的技术手段之一,但目前尚不存在一种通用的机器人智能家居架构。

本文提出了一种将服务机器人集成至智能家居中的综合解决方案,设计了一种

RiSH 通用总体架构，基于该架构实现了高低两个层次的应用服务，包括 1) 低层次的室内人员定位和活动识别服务；2) 高层次的跌倒检测及救援服务。

## 2 用于老年人护理的 RiSH 总体设计方案

远程护理人员既可以是专业的医护人员，也可以是普通家庭成员或朋友。

## 3 RiSH 实验平台硬件系统设计

如图 2 所示，RiSH 实验平台主要包括 RiSH 系统及辅助实验设备。RiSH 通过家庭网关连接到云端服务器，远程护理人员可

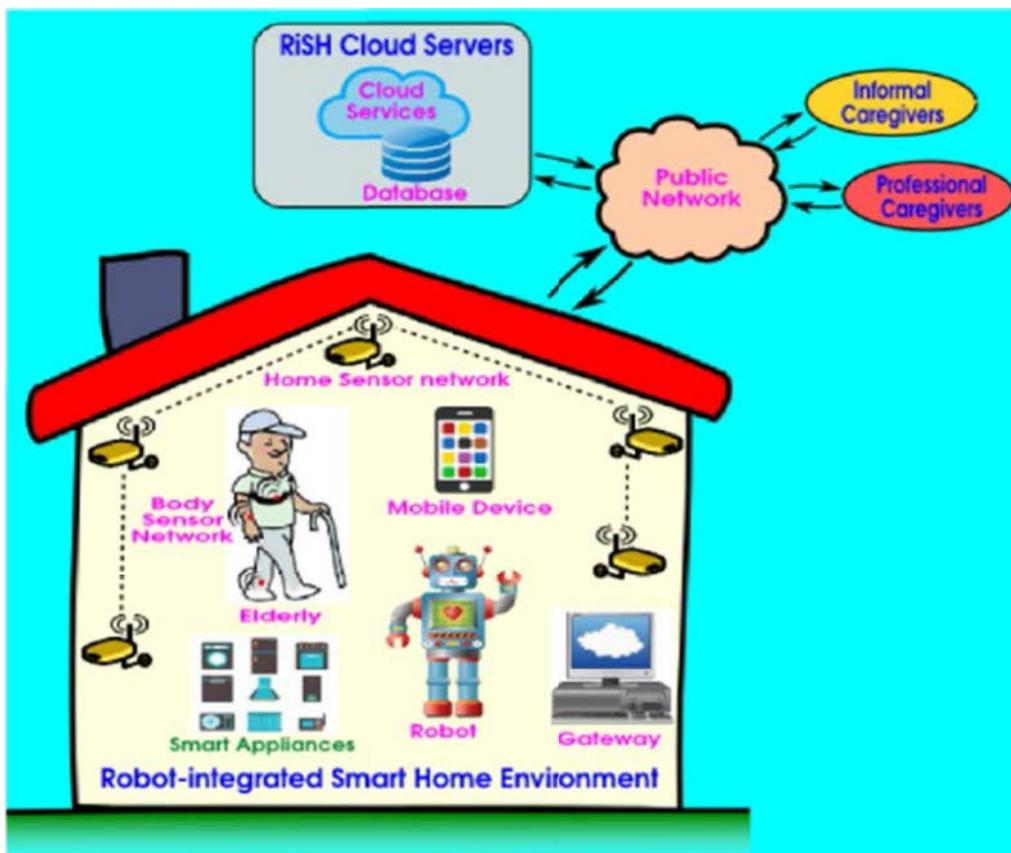


图 1 用于老年护理的 RiSH 整体方案

本文设计了如图 1 所示的 RiSH 整体方案，该方案主要由 RiSH 系统、RiSH 服务器、远程护理人员三部分构成，上述三者通过公网互联。其中 RiSH 系统主要包括分布式智能家庭传感器网络、家庭服务机器人、穿戴式传感器、智能移动设备、智能家电以及家庭网关等；RiSH 服务器主要提供诸如智能计算、系统管理等云端服务；

以通过与家庭服务机器人协作实现对老年人的护理工作。辅助实验设备主要包括声音模拟系统以及室内人员定位系统。RiSH 系统中的关键部分介绍如下：

### (1) 家庭服务机器人

基于 Pioneer P3-DX 移动平台开发的 ASCC 家庭服务机器人装备了丰富的传感器

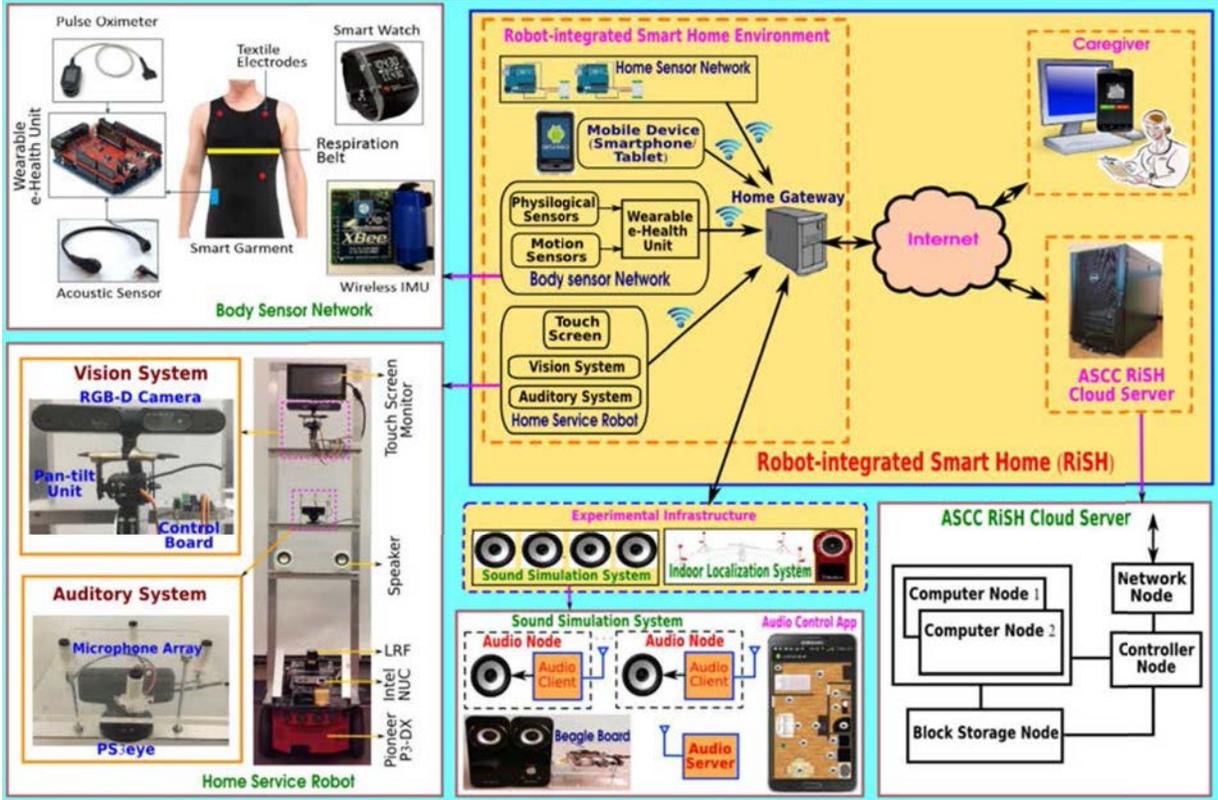


图 2 RiSH 硬件系统架构

等设备，如：激光雷达、视觉系统、听觉系统、Intel NUC 主机、电池等。视觉系统主要包括 Asus Xtion 深度相机<sup>[6]</sup>、云台以及控制电路等；听觉系统通过四麦克风阵列<sup>[7]</sup>可以实现语音识别、声音定位以及噪声环境下的声音分离等功能。

## (2) 穿戴式传感器网络

穿戴式传感器网络主要包括生理参数传感器、运动传感器、智能手表以及可佩带的 e-Health 传感器平台<sup>[8]</sup>。该传感器网络由测试人员佩戴，用于采集生理信号和运动数据，具体包括：心电数据（ECG）、血氧浓度（SpO<sub>2</sub>）、呼吸频率、颈部周围的声学信号以及身体的活动信息。其中，身体活动信息采集由定制的惯性测量单元

（IMU）完成，IMU 平台包含 VN-100 惯性传感器模组<sup>[9]</sup>、XBee 无线模块等。

## (3) 智能移动设备

智能移动设备（如智能手机或平板电脑等）可以用于控制家庭服务机器人，也可以用于穿戴式传感器数据的采集与分析。护理人员可以通过智能移动设备远程实现上述功能，也可以与家中老年人实时通信。

## (4) 家庭传感器网络

家庭传感器网络由分布在家庭环境中的各类传感器组成，用于室内人员的位置和活动信息的采集，主要包括一组分布式被动红外传感器（PIR）以及通过 XBee 协议连接的 GridEye 传感器<sup>[10]</sup>，所有传感器

节点按照一定规则安装在天花板的不同位置。

### (5) 家庭网关

家庭网关是智能家居本地数据采集与处理中心，负责与云端服务器及远程护理人员通信工作。家庭网关采集智能家居中各类传感器数据，并本地完成计算量小、实时性要求不高的任务；对于计算量大的任务需要上传到云端服务器处理。

### (6) 私有云服务器

本文基于 OSU ASCC 实验室的私有云基础架构，利用开源的云编排软件 OpenStack Juno<sup>[11, 12]</sup> 搭建智能家居云服务器。如图 2 所示，私有云使用了三个服务器节点和一个存储节点：大多数云管理服务均在控制器节点实现；两个计算节点承载为

实例创建基于内核的虚拟机管理程序和客户端服务；存储节点以块存储的方式实现运行实例的永久存储。

### (7) 辅助实验设备

辅助实验设备主要包括室内定位和声音模拟系统：Natural Point 公司的 OptiTrack 运动捕捉系统<sup>[13]</sup> 可提供机器人、测试人员或其他需要跟踪目标的位置信息；声音模拟系统用来模拟在多个典型生活场景中的声音，该系统包括音频节点、音频服务器和控制软件。音频节点基于 Beagleboard 微控制器<sup>[14]</sup> 和扬声器开发，智能手机端的音频控制软件可以触发放置在不同位置的音频节点模拟发出日常活动产生的声音。

## 4 RiSH 软件架构设计

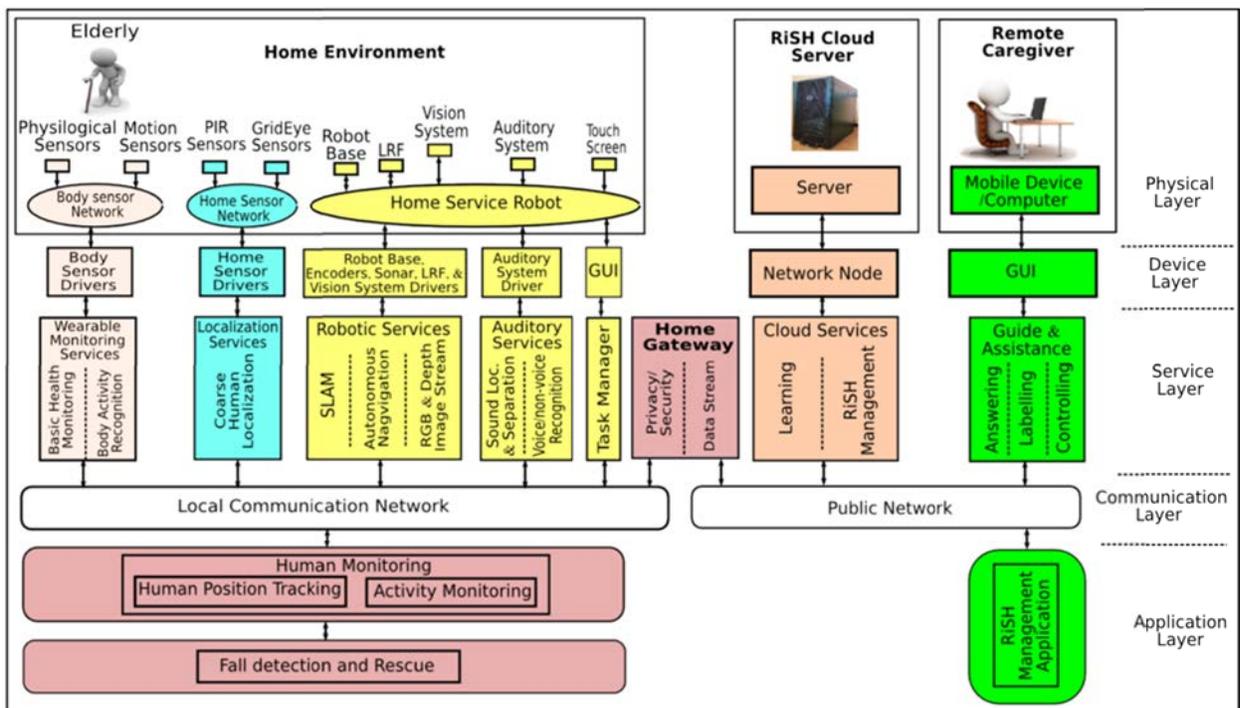


图 3 开放式层次化的 RiSH 软件架构

现有的智能家居架构通常是针对集中式同质系统设计，但 RiSH 实际是一个复杂异构系统，另外，RiSH 软件架构应具有模块化、可扩展、可重用等特性，因此，现有的架构无法直接应用在 RiSH 中。如图 3 所示，本文提出了一种分层 RiSH 软件架构，该架构主要由物理层、设备层、服务层、通信层和应用层构成。

#### 4.1 物理及设备层

物理设备层即 RiSH 的硬件系统，主要包括服务机器人、穿戴式传感器、家庭传感器网络以及智能移动设备等硬件模组。硬件系统由设备层管理，该层能够为机器人装备的传感器和执行器提供设备驱动程序。机器人的驱动程序基于机器人操作系统（Robot Operating System: ROS）<sup>[15]</sup>开发，声音系统的驱动程序基于开源声音处理软件 HARK<sup>[16]</sup>开发。

#### 4.2 服务层

家庭传感器网络中的每个 PIR 节点检测其感知区域内的活动人员，实现室内人员的粗略定位服务；穿戴式传感器网络可以提供人体行为识别和基本健康监测服务；机器人可以为使用者提供听觉感知服务、人机交互界面等。另外，与机器人相关的服务还包括基于 Rao-Blackwellized 粒子滤波<sup>[17]</sup>的 SLAM，基于自适应蒙特卡洛方法的机器人定位<sup>[18]</sup>等；听觉服务基于 GEVDMUSIC 方法实现声源定位<sup>[19]</sup>，基于 GHDSS 方法<sup>[20]</sup>实现声音分离，通过语音/非语音识别算法将分离的声音分为语音和非语音。本文使用由 MFCC（Mel-frequency 倒谱系数）、MFCC 的增量、对数能量、过零率、熵、频谱质心、频谱宽度、偏度和峰度组成的 31 维特征向量<sup>[21]</sup>计算每帧音频的统计特征；任务管理服务能够处理来自用户的命令；家庭网关在 RiSH 和护理人员之间提供数据流并保障其安全和隐私；云学习服务允许机器人将计算量大的机器学习算法转移至云端服务器中实现。

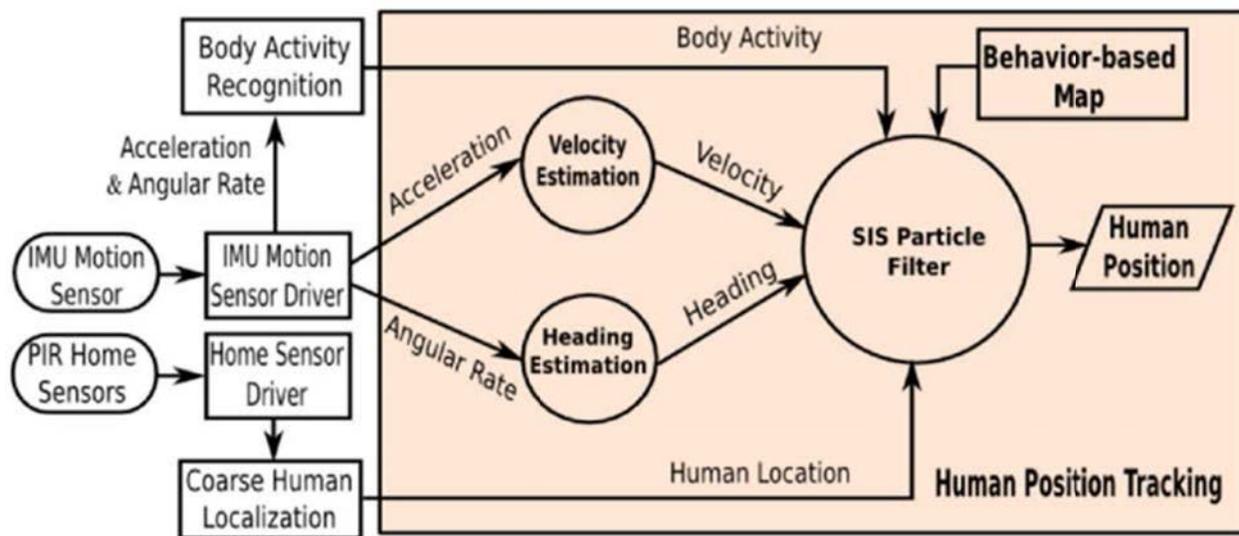


图 4 基于多传感器融合的人员定位与跟踪算法

### 4.3 通信及应用层

应用层包括初级应用层和高级应用层，初级应用层主要指 RiSH 的基本功能，如：人员位置跟踪和活动监控；基于初级应用设计的高级应用实现了老人的跌倒检测与救援。通信层通过局域网和公网在服务层和应用层的功能模块之间提供无缝连接。

### 5 RiSH 应用层实现

应用层的各类初级应用和高级应用均是基于服务层提供的各类基础服务实现的，首先，本文实现了两个初级应用，即人员位置跟踪和活动监控，然后，基于初级应用实现了摔倒检测与救援这一高级应用。

#### 5.1 人员位置跟踪

人员位置跟踪方案如图 4 所示。IMU 采集的 3 轴加速度和 3 轴角速度信息被用于估计人的行走速度和方向，PIR 传感器网络完成人的初步定位。运动数据和初步定位结果通过粒子滤波实现融合，融合后可以得到一个更加精确的人员位置信息。建立一个包含了人员位置与日常活动之间关系的广义地图，该地图确定了墙壁、家具以及其他设施的位置信息，实现了日常活动与位置关系之间的概率映射，以此提高基于贝叶斯网络的人员定位精度<sup>[22]</sup>。

#### 5.2 人员活动识别

日常活动一般会产生具有明显特征的声音事件，可以通过结合穿戴式传感器和声音事件识别算法来实现人的活动识别与监视。本文提出了一种新颖的基于上下文

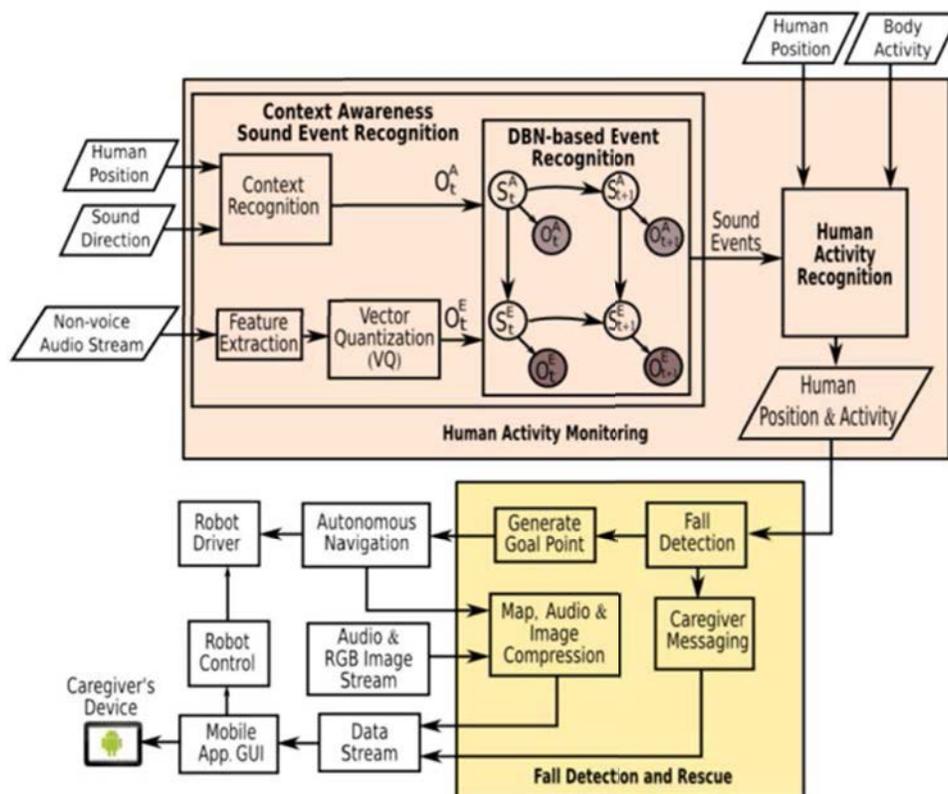


图 5 基于 DBN 的日常活动识别及跌倒检测与救助

信息的声音事件识别方法，该方法使用动态贝叶斯网络（Dynamic Bayesian Network: DBN）描述上下文信息与声音事件之间的关系。如图 5 所示，基于上下文信息的声音事件识别方法主要包括四个步骤：上下文识别、特征提取、矢量化以及基于 DBN 的声音事件识别。每帧声音信号的长度在 5 ms-150 ms 之间，相邻两帧之间的重叠长度设置为半个帧长，计算每帧声音信号的 31 维特征向量，采用 LGB（Linde-Buzo-Gray）算法设计编码字典，将声音特征向量转换为相应的码字标签，得到的标签作为声音事件的观测集，基于观测集构建如图 5 所示的双层 DBN 网络并对上下文信息与声音事件之间关系进行建模。

### 5.3 摔倒检测与紧急救援

如图 5 所示，本文提出了一种基于位置跟踪、活动识别以及声音事件识别的跌倒检测方法。如果机器人检测到跌倒事件发生，首先，以跌倒人的位置作为目标点，利用导航算法规划生成初始路径；然后，以跌倒人为中心划定安全区域，以此安全区域与初始路径的交点为最终目标点微调初始路径；最后，服务机器人自主导航至目标点检查现场情况，并将求救消息发送至远程护理人。地图和现场音视频也会发送到护理人员的移动终端，护理人员可以控制机器人与老年人进行视频通话，以更好的处理紧急情况。

## 6 实验结果及分析

### 6.1 基于穿戴式传感器网络的基本健康参数监测

如图 6 所示，测试者的健康数据可以

成功上传云服务器，护理人员可以通过智能终端查看相关数据。该功能可用于老年人体征参数监测，也可用于研究日常活动与健康之间的关系。

### 6.2 人员位置跟踪实验

为了验证所提位置跟踪算法的有效性，我们设计与日常家庭活动相对应的运动轨迹，每条轨迹都对应着一系列的日常活动，如进入厨房、走近冰箱、走近餐桌等。如图 7 所示的定位实验，绿色实线表示了 OptiTrack 捕捉到的真实轨迹，红色虚线表示估计轨迹，测试路线总长 35.6 米。仅使用 PIR、IMU 以及使用二者组合的定位均方误差分别为：2.25 米，0.411 米和 0.140 米。

### 6.3 日常活动识别实验

我们从互联网收集了浴室、厨房、客厅等场合的声音文件，建立了场景声音数



图 6 健康数据采集手机终端截图

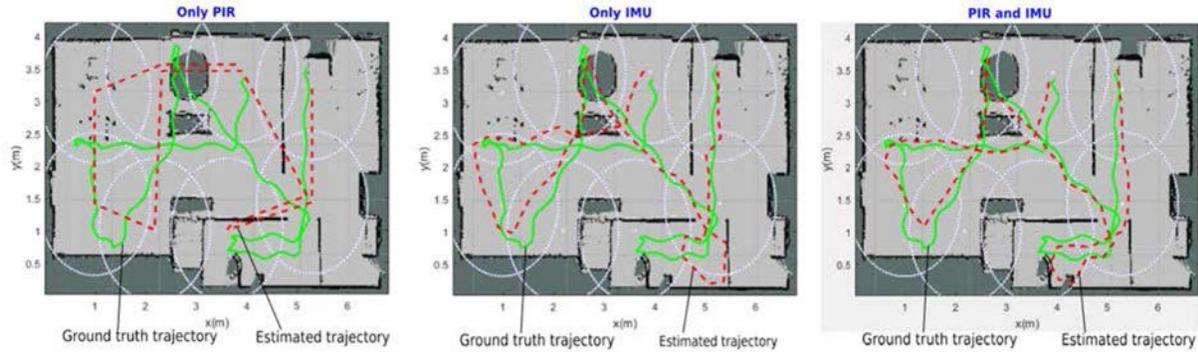


图 7 室内人员定位与跟踪实验

数据库 ASCCsoundLib，该数据库包含了 38 种已标注的声音事件：6 种餐厅声音事件、3 种客厅声音事件、8 种浴室声音事件、14 种厨房声音事件、2 种卧室声音事件以及其他 5 种类型的声音事件。

如图 8 所示，实验开始时，机器人处

于 a 点，测试人员走向洗手间 b 处，然后完成刷牙、洗手等活动，与此同时，利用声音模拟系统播放与测试人员活动相对应的声音，机器人采集到的声音信号如图 8 (f1) 所示，声音事件识别结果如图 8 (f2) 所示；接着，测试人员走到 c 处厨房烧水，随后，

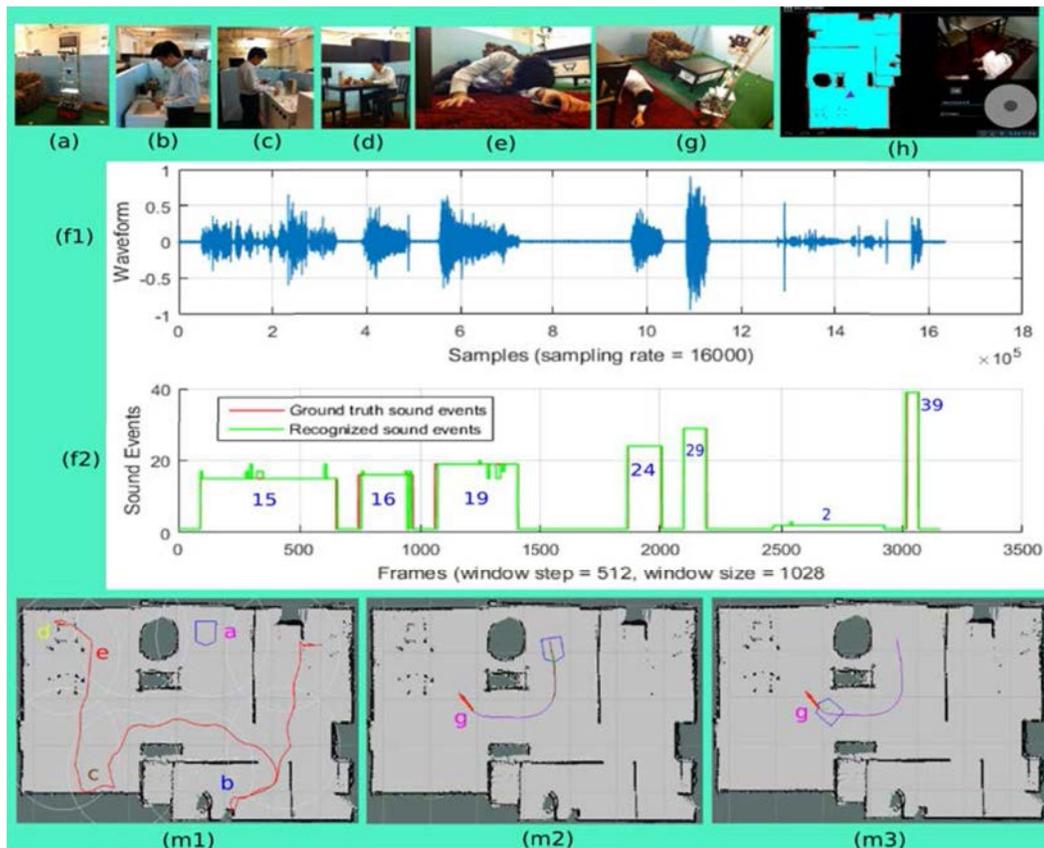


图 8 日常活动识别实验

进入 d 处餐厅用早餐, 这段活动的声音信号也被成功捕捉并且识别出了相应活动; 用餐后, 测试人员跌倒在 e 处, 机器人捕获声音信号并及时识别出跌倒事件, 根据测试人员跌倒位置和声音的方向, 机器人确定目标点 g 并自主导航至目标点, 同时将报警信息分别推送至云端和护理人员的智能终端, 护理人员可以通过如图 8 (h) 所示的交互界面远程遥控机器人实现对跌倒人员的救助。通过对比实验发现, 未考虑上下文信息的声音事件识别准确率较低, 而本文所提基于上下文信息的 DBN 声音事件识别正确率一般可达 80%以上。

## 7 结论与展望

为了解决老年人居家护理问题, 本文开发了 ASCC 家庭服务机器人平台, 提出了基于机器人的智能家居解决方案, 设计了一种分层开放式的 RiSH 软件架构及相应的软硬件测试平台。通过多传感器融合等算法, 本文所提的 RiSH 架构能够实现独居老人体征参数采集、位置跟踪、活动监测等多种功能。实验结果证明了本文所提方法架构及实验平台的有效性。下一步, 我们将重点针对人机交互界面、家庭服务机器人的拟人化服务等内容展开研究工作。

## 8 致谢

本研究获美国自然科学基金项目 (编号: CISE/IIS 1231671/IIS 1427345)、浙江大学工业控制技术国家重点实验室课题 (编号: ICT1600217)、以及深圳市海外孔雀团队项目 (编号: KQTD20140630154026047) 资助。

This work is supported by the National

Science Foundation (NSF) Grant CISE/IIS 1231671, CISE/IIS/1427345, the Open Research Project of the State Key Laboratory of Industrial Control Technology, Zhejiang University, China (No. ICT1600217) and the Shenzhen Overseas High Level Talent (Peacock Plan) Program (No. KQTD20140630154026047).

## 参考文献

- [1] WACV, 2016. WHO. World Health Organization: 10 facts on ageing and the life course. [Online]. Available: Website:<http://www.who.int/features/factfiles/ageing/en/>
- [2] P. R. Liu, M. Q.-H. Meng, P. X. Liu, F. F. Tong, and X. Chen, "A telemedicine system for remote health and activity monitoring for the elderly," *Telemedicine Journal & e-Health*, vol. 12, no. 6, pp. 622–631, 2006
- [3] M. Broxvall, M. Gritti, A. Saffiotti, B.-S. Seo, and Y.-J. Cho, "Peis ecology: Integrating robots into smart environments," in *Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on*. IEEE, 2006, pp. 212–218.
- [4] R. Borja, J. De La Pinta, A. Alvarez, and J. M. Maestre, "Integration of service robots in the smart home by means of upnp: A surveillance robot case study," *Robotics and Autonomous Systems*, vol. 61, no. 2, pp. 153–160, 2013.
- [5] M. Chen, Y. Ma, S. Ullah, W. Cai, and E. Song, "ROCHAS: robotics and cloud-assisted healthcare system for empty nester," in *Proceedings of the 8th International Conference on Body Area Networks*, 2013.
- [6] ASUS. Xtion pro live. [Online]. Available: <https://www.asus.com/us/3D-Sensor/>
- [7] Sony. Ps3 eye camera. [Online]. Available: <http://www.sony.co.in/product/playstation+eye>
- [8] C. duino Hacks. And e-health raspberry

- sensor pi. [Online] Available: <http://www.cooking-hacks.com/documentation/tutorials/ehealth-v1-biometric-sensor-platform-arduino-raspberry-pi-medical/>
- [9] The vn-100 rugged imu. [Online]. Available: <http://www.vectornav.com/products/vn100-rugged>
- [10] Infrared array sensor grid-eye. [Online]. Available: <http://industrial.panasonic.com/ww/products/sensors/built-in-sensors/grid-eye>
- [11] Open source software for creating private and public clouds. [Online]. Available: <https://www.openstack.org/>
- [12] O. Khedher, OpenStack Sahara Essentials. Packt Publishing Ltd, 2016
- [13] Motion capture systems - optitrack. [Online]. Available: <https://www.naturalpoint.com/optitrack/>
- [14] Beagleboard-xm. [Online]. Available: <http://beagleboard.org/beagleboard-xm>
- [15] Ros wiki. [Online]. Available: <http://www.ros.org/wiki/>
- [16] K. Nakadai, T. Takahashi, H. G. Okuno, H. Nakajima, Y. Hasegawa, and H. Tsujino, "Design and implementation of robot audition system 'hark' open source software for listening to three simultaneous speakers," *Advanced Robotics*, vol. 24, no. 5-6, pp. 739-761, 2010.
- [17] W. B. G. Grisetti, C. Stachniss, "Improved Techniques for Grid Mapping With Rao-Blackwellized Particle Filters," *Robotics, IEEE Transactions on*, vol. 23, no. 1, pp. 34-46, 2007.
- [18] D. Fox, "Adapting the Sample Size in Particle Filters Through KLD Sampling," *International Journal of Robotics Research*, vol. 22, 2003.
- [19] K. Nakamura, K. Nakadai, F. Asano, Y. Hasegawa, and H. Tsujino, "Intelligent sound source localization for dynamic environments," in *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on*. IEEE, 2009, pp. 664-669.
- [20] K. Nakadai, G. Ince, K. Nakamura, and H. Nakajima, "Robot audition for dynamic

environments," in *Signal Processing, Communication and Computing (ICSPCC), 2012 IEEE International Conference on*. IEEE, 2012, pp. 125-130.

- [21] K.-M. Kim, S.-Y. Kim, J.-K. Jeon, and K.-S. Park, "Quick audio retrieval using multiple feature vectors," *Consumer Electronics, IEEE Transactions on*, vol. 52, no. 1, pp. 200-205, 2006.
- [22] M. Pham, D. Yang, W. Sheng, and M. Liu, "Human localization and tracking using distributed motion sensors and an inertial measurement unit," in *2015 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2015, pp. 2127-2132.

## 作者介绍



**盛卫华**, 美国俄克拉荷马州立大学 (Oklahoma State University) 电子与计算机工程系任副教授。他于 2002 年在美国密歇根州立大学 (Michigan State University) 大学电子与计算机工程系获

得博士学位。分别于 1994 年和 1997 年在浙江大学电机工程系获得学士和硕士学位。他是美国俄克拉荷马州立大学高级传感、计算与控制实验室主任。发表了 170 篇以上的国际会议和期刊论文并取得了 1 项美国专利。承担多项美国科学基金 (NSF) 资助的研究工作以外, 还承担美国国防部 (DoD)、美国交通部 (DoT) / 俄克拉荷马州政府等资助的研究。研究领域包括移动机器人、穿戴式计算、人机交互、和智能交通系统。



**Ha Manh Do**, 俄克拉荷马州立大学电气与计算机工程系在读博士生, 1999 年于越南河内科技大学电子与通信工程专业获学士学位, 2015 年于俄克拉荷马州立大学电气与计算机工

程专业获硕士学位。主要研究领域包括移动机器人，智能家居，模式识别，深度学习等。



**Minh Pham**，俄克拉荷马州立大学电气与计算机工程系在读博士生，2007 年获越南河内科技大学计算机专业获学士学位。主要研究领域为智能家居和穿戴式计算。



**杨丹**，副教授，2002 年东北大学生物医学工程专业获学士学位；2005 年东北大学生物医学工程专业获得硕士学位，同年留校任教；2009 年东北大学检测技术与自动化装置专业获得博士学位；

2010- 2012 年在东北大学计算机科学与技术专业，进行博士后工作。2013 年受国家留学基金委资助赴美国俄克拉荷马州立大学做访问学者。主持省部级以上科研项目 4 项，横向课题 5 项；授权发明专利 3 项，实用新型专利 2 项；主编出版教材 2 部；发表学术论文 30 余篇。主要从事电磁信号检测技术与装置，有限元计算分析等领域的研究工作。



**刘妹琴**，浙江大学教授，国家优秀青年科学基金获得者、教育部新世纪优秀人才、浙江省杰出青年基金获得者，浙江省 151 人才一层次培养人员。中南大学控制理论与控制工程专业博士，华中科技大学控制系

博士后，美国 University of New Orleans 访问学者，IEEE 高级会员。长期从事信息融合技术、人工智能理论以及非线性控制等方面的理论及应用研究。近 5 年发表 100 余篇论文，其中近 70 多篇被 SCI 收录，出版学术专著 1 部；主持包括 863 计划、国家自然科学基金和浙江省自

然科学基金等在内的国家与省部级项目 18 项；以第一完成人获浙江省科学技术奖二等奖 1 项，浙江省高校科研成果奖一等奖 1 项；2017 年获中国自动化学会第三届青年科学家奖。



**张森林**，浙江大学控制理论与控制工程专业硕士（1988 年-1991 年），浙江大学电气工程学院教授（2005 年至今），长期从事智能家居、信息融合、纺织自动化的研究工作。近 5 年发表论文 50 余篇，其中 40

多篇被 SCI、EI 收录。出版专著 1 本，合编教材 1 本。主持国家自然科学基金项目 2 项，浙江省自然科学基金重点项目 2 项，国家科技支撑计划 2 项。主持或作为技术负责人的项目获国家科技进步奖 1 项，省部级科技进步奖 5 项。获得国家发明专利 5 项，实用新型专利 20 多项。2005 年获得桑麻纺织科技二等奖。2006 年入选“浙江省 151 人才工程第二层次培养人员”。



**袁宪锋**，山东大学控制理论与控制工程专业在读博士生，2007 年于山东大学自动化专业获工学学士学位，同年保送至山东大学控制学院机器人研究中心攻读博士学位（硕博连读），2016 年受国家

留学基金委资助赴美国俄克拉荷马州立大学 ASCC 实验室联合培养一年。主要研究领域包括家庭服务机器人、机器学习、智能故障诊断、嵌入式软硬件系统研发等。

# Unambiguous and Accurate Text Extraction in Natural Scenes

## 自然场景下的文字精确检测定位方法

Xuejian Rong, Yingli Tian\*

戎雪健, 田英利\*

The City College, The City University of New York, New York, NY 10031, USA;

纽约城市大学城市学院, 纽约市, 纽约州 10031, 美国;

\*联系人, E-mail: ytian@ccny.cuny.edu

### Abstract

Text instance as one category of self-described objects provides valuable information for understanding and describing cluttered scenes. In this paper, we explore the task of unambiguous text localization and retrieval, to accurately localize a specific targeted text instance in a cluttered image given a natural language description that refers to it. To address this issue, first a novel recurrent Dense Text Localization Network (DTLN) is proposed to sequentially decode the intermediate convolutional representations of a cluttered scene image into a set of distinct text instance detections. Our approach avoids repeated detections at multiple scales of the same text instance by recurrently memorizing previous detections, and effectively tackles crowded text instances in close proximity. Second, we propose a Context Reasoning Text Retrieval (CRTR)

model, which jointly encodes text instances and their context information through a recurrent network, and ranks localized text bounding boxes by a scoring function of context compatibility. Quantitative evaluations on standard scene text localization benchmarks and a newly collected scene text retrieval dataset demonstrate the effectiveness and advantages of our models for both scene text localization and retrieval.

### 摘要

近年来, 自然场景下的文字检测与定位正得到越来越多的关注。文字作为自然场景中最为直观, 精确与富含信息量的特征之一, 其检测与定位对于识别后续文字, 理解场景内容, 建立图片索引, 以及自动驾驶等应用都具有重要意义。本文针对这一领域的最新方法进行了比较与评判, 同时还提出了一种新型的文字精确检测定位方法, 在结合无歧义自然语言描述

的基础上，实现了图像文字的高精度检测与检索。文章最后总结了该领域的挑战与发展趋势。

**Keywords:** Scene Text Detection, Deep Learning

**关键词:** 自然场景文字检测, 深度学习方法

## 1 概述

近年来，复杂自然场景下的图像文本检测与定位已成为计算机视觉与文档分析领域的热点研究课题。早期的计算方法与数据集以及相关的比赛任务，如自 2003 年开始的国际文字检测分析竞赛（ICDAR Robust Reading Competition），主要关注近景下的文档文字检测定位，并取得了不错的效果。但近年来，随着数字多媒体技术与移动互联网的发展，以及数码相机与智能手机等移动设备的普及，抓拍的复杂自然场景图片越来越多，传统方法已不能很好的解决大场景，多文字，复杂背景下的文本检测定位问题。相关问题的解决，对于诸多行业的最新应用，都有着重要意义，例如：智能手机的实时拍照翻译软件，自动驾驶系统的交通标志与车牌识别，网络视频图像的自动分类与检索，穿戴式的盲人导航室内识别系统等<sup>[1,2]</sup>。

自然场景下的文字检测识别不同于文档图像中的文字识别。文档文字识别一般是白底黑字，背景颜色单一，字体规范且大小统一，因此检测率与识别率早已达到工业界与日常环境下的实用要求，如扫描复印领域常见的光学字符识别（Optical Character Recognition, OCR）应用等。而与

之不同的是，自然场景下的文本检测定位指判断不同场景下摄制的图像是否存在有效文本，这些文本文字本身是图像内容的一部分，若有，则定位所有文本所在的坐标位置。而自然场景下拍摄的图像通常又十分复杂，包括但不限于光照不均匀，字体/颜色/尺度/方向的变化，部分重叠与遮挡，噪音与模糊，复杂背景与相似物体（如砖块/窗户/栅栏等）等影响因素，这些因素给自然文字的检测定位与识别造成了很大的干扰。由于场景文字信息提取的第一步是文字的检测定位，这一环节的效果直接决定了后续应用，如识别/提取/索引的效果，它的重要性便显得尤为突出。近年来深度学习方法的发展与进步，也催生了许多高性能的文字检测定位方法，这篇文章将介绍分析并对比这些最近的方法，同时结合自然语言描述，提出一种最新的无歧义文字精确检测定位方法，为该领域的后续发展开辟新的方向。

## 2 常用的文字检测定位方法与数据集

### 2.1 近期文字检测定位方法

近年来常见的文字检测定位方法大体可分为两类：1) 基于传统特征的方法；2) 基于深度神经网络特征的方法。前者又可进一步分为基于滑动窗口（sliding window）的方法与基于连通域（connected components）的方法，本节将介绍这些方法的大致特点与发展脉络。

**滑动窗口方法** 基于滑动窗口的方法主要是将多尺度窗口在图像中所有可能的位置进行滑动，随之从对应的窗口位置提取

文本特征并进行分类。因此，所提取特征的效率与优劣对文字检测的结果起到了决定性的作用。最近流行的手工设计文本检测特征包括：局部二值特征（local binary pattern, LBP<sup>[8]</sup>），尺度不变特征描述变换描述符（scale-invariant feature transform, SIFT<sup>[9]</sup>），快速文本检测特征（FASText<sup>[6]</sup>）等。虽然滑动窗口方法直观简单，易于实现，但多尺度下的检测效率通常不高，且手工设计特征的性能与普适性也相对较低。

**连通域方法** 基于连通域的方法主要是通过颜色，纹理，笔画宽度，或者区域极值等属性进行聚类并得到相应的连通域，之后从连通域中提取特征并进行文本 / 非文本的区域二分类。目前最为流行的方法包括笔画宽度变换（stroke width transform, SWT<sup>[6, 31]</sup>）与最大稳定极值区域（maximally stable extremal regions, MSER<sup>[7, 8, 32]</sup>）两种。但 SWT 与 MSER 方法均对光照/颜色变化及噪声/模糊等影响因素较为敏感，文字检测性能容易受到影响。

**深度学习方法** 基于深度学习与神经网络特征的方法<sup>[5, 12, 13, 14, 15]</sup>，主要采用卷积神经网络（convolutional neural network）的优秀特性，在学习大量训练数据的基础上，直接提取文本特征并进行分类。这类方法在多个最新的文本检测数据集上取得了很好的结果。尽管如此，这些方法依然极端依赖于训练数据的规模与质量，并且无法输出除文字本身外，更多有用的语义与环境场景信息。

## 2.2 公开数据集

自 2003 年的国际文字检测定位竞赛（ICDAR Competition）以来，越来越多的文本检测数据集被公布出来，供研究人员检验/评估/对比所提出方法的性能。近年来常用的数据集包括 ICDAR (International Conference on Document Analysis and Recognition) 2013/2015 dataset<sup>[27]</sup>，SVT (Street View Text) dataset<sup>[28]</sup>，COCO-Text dataset<sup>[35]</sup>，与 SynthText (Synthetic Text) dataset<sup>[14]</sup>等，这些数据集的统计数据如表 1

表 1 近期常用文本检测数据集的统计数据（图像数量）

数据集	训练用数量 (图片/文字)	测试用数量 (图片/文字)
ICDAR 2013 <sup>[27]</sup>	229/848	255/1095
ICDAR 2015 <sup>[27]</sup>	1,000/4,500	500/2,000
SVT <sup>[28]</sup>	100/211	249/514
COCO-Text <sup>[35]</sup>	43,686/118,309	10,000/55,280
SynthText <sup>[14]</sup>	800,000/~8,000,000	0

所示，其中 SynthText 数据集仅包含训练数据。

### 3 文字的精确检测定位与无歧义检索

#### 3.1 背景与研究动机

文字呈现形式，即文本对象（text instance，如场景图像中的字符，单词和字符串），可以提供最简洁准确的自然语言表达，以便理解和解释自然场景。从基于图像的自然场景读取文本信息，在场景理解及其相关应用中起着重要作用，例如导航与定位，上下文检索，端到端机器翻译，及自动驾驶系统等。然而，大多数现有的

自然场景文本检测定位方法通常将文本对象视为自然物体的一种，并尝试将文本对象表示为区别于其他类别物体的特殊特征，然后将场景中存在的所有文本对象分配给预定义的种类标签（class label）。这意味着自然场景文字不能比其他自然物体对象贡献更多的对场景的语义描述，即使文本对象自带的文字信息通常与所处的环境更为相关，并且能够在语义上达成自我描述（self-described）。

准确地说，对于自然场景图像中的文本对象，当前主流的文本提取方法可以生成其位置和顺序字符编码，我们可以将其称之为位置（spatial）与字面（literal）信



图 1 结合无歧义自然语言描述的文字精确检测定位示例，各种颜色的候选框表示所有检测出的文本对象的位置，而其中的彩色候选框标明了对应自然语言描述所指代的特定文本对象。

息。然而，为了全面地描述和解释高度复杂的自然场景，我们需要更高级的线索，例如语义（semantic）与上下文（context）信息。特别是针对盲人或视障人士，当前虽已存在众多场景文字检测定位的实际应用，例如商店中基于图像信息的购物助理，但只有当用户能够准确感知文本对象的位置时，文本信息才能真正有效地帮助用户理解实际场景。

例如，当盲人或视障人士在杂货店使用文本检测应用来帮助找到特定产品及其价格时，可想而知，他/她将会更偏好购物助理应用程序生成自然的语言描述，如“在一篮红苹果上面的红色标志上写着{无与伦比的优惠价格}”，而不是从图像中提取出的一堆离散且无序的单词列表，如图 1 所示。

为了更好地利用自然场景中的文本信息，本文探讨了文本对象与其背景之间的关系。我们提出了一个基于文本的自然场景理解的新框架，它将来自场景的文本对象的位置信息和明确的自然语言描述相结合，而这种自然语言描述通常被称为“引用表达”（referring expression）<sup>[3, 4, 21, 25]</sup>。文本对象的上下文描述对于整个场景的理解和描述通常是十分有效的，而能够从自然图像中无歧义地精确检索出文本对象在使用自然语言界面的许多应用程序中非常关键。

### 3.2 文字检测与定位

我们提出一个新的文字检测与定位的密集文本定位网络模型（Dense Text Localization Network, DTLN）。这个模型采用 VGG-16 <sup>[17]</sup> 卷积神经网络

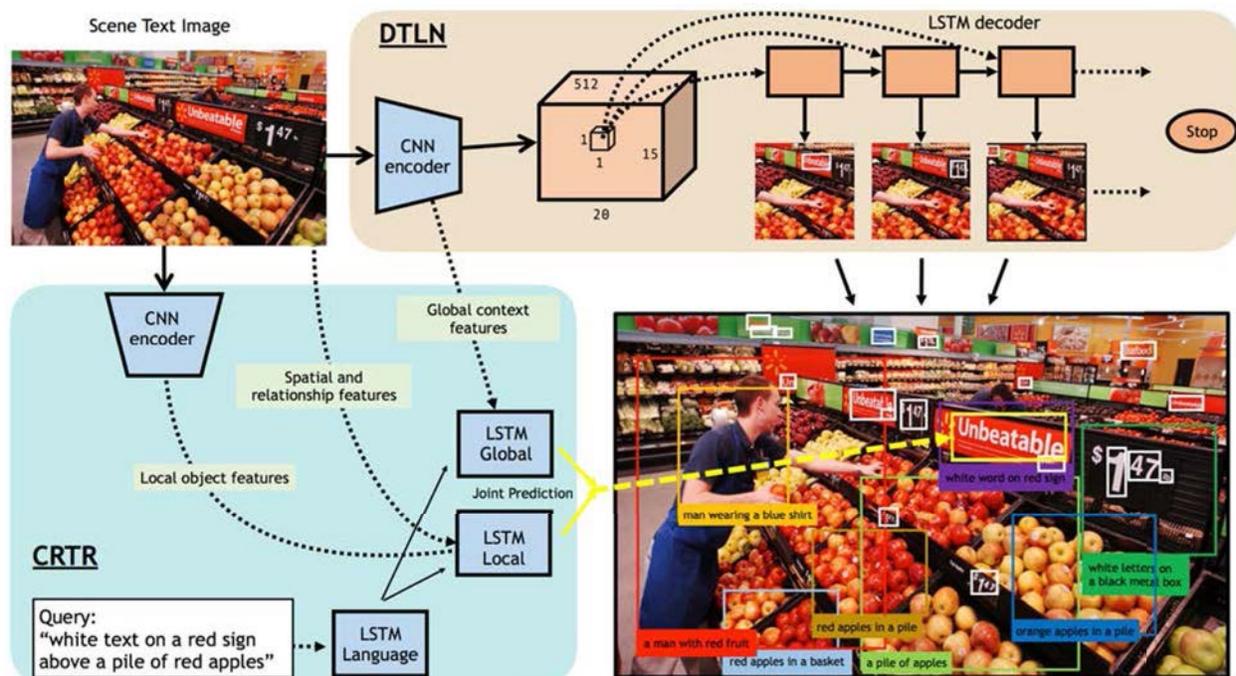


图 2 本文所提出的文本精确检测定位及无歧义检索方法的流程与模型架构图，包含密集文本定位网络模型(Dense Text Localization Network, DTLN)与上下文推理文本检索模型(Context Reasoning Text Retrieval, CRTR)。

(Convolutional Neural Network, CNN) 架构, 将场景图像编码为  $20 \times 15 = 300$  个 512 维特征描述符的特征图。虽然场景文本对象在检测阶段经常被视为自然物体的一个特殊类别, 但它们的高度变化的外观/尺度和自我描述属性将其与自然物体显著区分开来。根据长短期记忆网络 (Long-Short Term Memory, LSTM<sup>[18]</sup>) 模型, 我们构建了一个循环解码器 (recurrent decoder), 对所有潜在的目标文本对象进行联合预测, 基于 CNN 的编码器与基于 LSTM 的解码器的组合在我们的框架中起着至关重要的作用, 因为这种组合可以生成可变长度的预测集, 而这种优势已被成功地用于图像描述<sup>[19, 22, 23]</sup>, 机器翻译<sup>[20]</sup>和行人检测<sup>[24]</sup>等任务。值得注意的是, 虽然此种方法在行人检测中表现良好, 但对于高度不规则的对象表现不佳; 此外, 这种方法主要意图解决的遮挡问题在自然场景文本检测定位中几乎很少发生, 而本文提出的模型有效地解决了这个问题。另外, 在我们的任务中,

生成连贯预测集的能力至关重要, 因为在图像的各个目标区域中不会存在对潜在文本对象个数的先验知识 (prior knowledge), 而我们的系统可以有效记忆之前生成的文本对象, 并避免对同一目标重复进行预测。

先前提取的 512 维特征描述符总结了局部区域的内容, 包括文本对象的大小, 位置和类别的信息, 随后基于 LSTM 的解码器将从这些 CNN 编码的特征描述符中自动提取出目标场景文本对象, 并顺序输出新的候选边界框及其相应的置信度分数, 而该分数表示在边界框的位置可以找到以前未检测到的文本对象的概率。边界框是按照下降的置信度分数产生的, 当基于 LSTM 的解码器不能在这一局部区域中找到具有更高置信度得分的更多边界框时, 将产生停止符号以结束整个解码过程。最终, 我们将收集来自自然场景图像的所有 300 个局部区域的所有输出边界框和置信度得分作为文本对象的检测定位结果。



图 3 自然场景文字检测定位的实验结果示例。绿色框包含了正确的文本检测结果, 红色实体框包含了错误的检测结果, 而红色虚线框包含了遗漏的文本对象。

### 3.3 文本对象的无歧义检索

接着我们提出了通过自然语言来检索场景文本对象的上下文推理文本检索 (Context Reasoning Text Retrieval, CRTR) 模型。在测试阶段, 给定图像以及自然语言查询和一组候选文本边界框, 基于三个 LSTM 解码器 (分别对应自然语言特征, 全局图像特征, 和局部图像特征) 的 CRTR 模型将从 DTLN 的检测输出结果中选择最相关的子集。

自然场景图像中的文本对象通常嵌入在各种复杂背景中, 因此如果不考虑它们

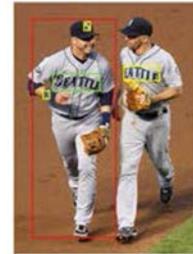
与上下文中的自然对象的关系, 就难以对文本对象进行无歧义的描述。直观地说来, 自然场景中的文本对象通常由出现在某些对象的表面上, 并由印刷或手写字符组成, 而它们的视觉关系通常主导着一张自然场景图像的整体解释。由于文本对象和自然对象 (例如物品或人物) 之间的关系集是巨大并且呈指数增长的, 我们将专注于与文本对象直接相关联和交互的上下文概念。然而, 获得足够的训练示例来涵盖所有这种关系对仍然十分困难。为了简化这个问题并制定出最可行的解决方案, 我们将语义空间减少为仅包含单个文本对象和单个



query = "largest text on the closest object"



query = "white text around a bench"



query = "largest text left to the right human"



query = "text on a motorcycle"



query = "text on the right cup"



query = "text on top of a red boat"



query = "blue text on the largest plane"



query = "most salient text on a bus"

图 4 无歧义文字检索的实验结果示例, 绿色框标明了成功检索出的文本对象, 红色框标明了最相关的场景物体, 而黄色框标明了其余检测出的文本对象。

上下文自然对象之间的关系，因为所有可能的关系对的语义空间比单个文本对象的语义空间大得多。为了避免对文本对象描述的模糊性，这里我们只专注于它们的空间关系和文本属性的预测。

#### 4 实验对比与评估

在实验中，我们采用 SynthText 数据集进行文本检测器的训练，并基于 COCO-Text<sup>[35]</sup>数据集采集了新的无歧义文本检索数据集，命名为 COCO-TextRef<sup>[36]</sup>。上文所提出的密集文本定位网络 DTLN 是通过 SynthText in the Wild 数据集中的 80 万个图像进行训练的，每个图像的大小被预先调整为 480\*640，而 VGG-16 网络的权重则通过 ImageNet<sup>[16]</sup>上预先训练的权重进行初始化，并进行微调，以满足解码过程的新需求。解码器中的所有权重都从均匀分布初始化，并于训练过程中在每次迭代的同时在一个图像的所有网格单元上并行进行，而所有权重都在区域和 LSTM 的解码步骤之间进行绑定。整个基于 SynthText in the Wild 数据集的训练在 NVIDIA Titan X (Maxwell) GPU 上需要大约 15 个小时。文本对象检测定位的实验示例可参见图 3。

我们进一步在新收集的 COCO-TextRef<sup>[36]</sup>数据集上评估了上下文推理文本对象检索 (CRTR) 模型，并将其与 LRCN<sup>[21]</sup>，DenseCap<sup>[33]</sup>，及 SCRC<sup>[26]</sup>模型进行比较，这些也相应地在 COCO-TextRef 数据集上进行了微调，以便检索文本对象。对于 CRTR 模型，给定图像和自然语言描述语句，模型应该从场景文本定位方法生成的一组候选文本区域中检索出最相关的文本对象，图 4 展示了在 COCO-TextRef 数据

集中成功检索到文本对象的实验示例，成功检索的标准为高分候选区域与正确标注重叠面积不小于 50%<sup>[29,30]</sup>，实验结果表明所提出的模型基于输入的自然语言描述有效地定位和检索到了目标文本区域。

#### 5 研究趋势与总结

综上所述，文本对象通常具有高级语义信息，对众多应用具有重要意义。自然场景下的文字检测定位是当前重要的研究领域，而本文结合最新深度学习技术与自然语言描述提出的无歧义精确检测定位方法是未来的重要发展方向之一。未来的扩展方向包括基于像素层级的精确文本分割任务与结合图像恢复技术（图像高分辨率处理与去模糊<sup>[34]</sup>）的模型优化。

#### 6 致谢

本研究获美国自然科学基金项目（编号：EFRI-1137172, IIP-1343402, IIS-1400802）资助。

#### 参考文献

- [1] B. Xiong and K. Grauman. Text detection in stores using a repetition prior. WACV, 2016.
- [2] C. Yi, Y. Tian, and A. Arditì. Portable camera-based assistive text and product label reading from hand-held objects for blind persons. IEEE Trans. on Mechatronics, 2014.
- [3] L. Yu, P. Poirson, S. Yang, A. C. Berg, and T. L. Berg. Modeling context in referring expressions. ECCV, 2016.
- [4] S. Kazemzadeh, V. Ordonez, M. Matten, and T. L. Berg. ReferIt game: Referring to objects in photographs of natural scenes. EMNLP, 2014.
- [5] M. Jaderberg, K. Simonyan, A. Vedaldi, and A. Zisserman. Reading text in the wild with

- convolutional neural networks. *International Journal of Computer Vision*, 2015.
- [6] M. Busta, L. Neumann, and J. Matas. FASText: Efficient Unconstrained Scene Text Detector. *ICCV*, 2015.
- [7] B. Epshtein, E. Ofek, and Y. Wexler. Detecting text in natural scenes with stroke width transform. *CVPR*, 2010.
- [8] L. Neumann and J. Matas. Real-time scene text localization and recognition. *CVPR*, 2012.
- [9] X. Yin, K. Huang, and H. Hao. Robust text detection in natural scene images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2014.
- [10] X. Rong, C. Yi, X. Yang, and Y. Tian. Scene text recognition in multiple frames based on text tracking. *ICME*, 2014.
- [11] X. Yin, W. Pei, J. Zhang, and H. Hao. Multi-orientation scene text detection with adaptive clustering. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2015.
- [12] M. Jaderberg, A. Vedaldi, and A. Zisserman. Deep features for text spotting. *ECCV*, 2014.
- [13] Z. Zhang, C. Zhang, W. Shen, C. Yao, W. Liu, and X. Bai. Multi-oriented text detection with fully convolutional networks. *CVPR*, 2016.
- [14] A. Gupta, A. Vedaldi, and A. Zisserman. Synthetic data for text localization in natural images. *CVPR*, 2016.
- [15] Z. Tian, W. Huang, T. He, Pan. He, and Y. Qiao. Detecting text in natural image with connectionist text proposal network. *ECCV*, 2016.
- [16] A. Krizhevsky, I. Sutskever, and G. E. Hinton. ImageNet classification with deep convolutional neural networks. *NIPS*, 2012.
- [17] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *ICLR*, 2015.
- [18] S. Hochreiter and J. Schmidhuber. Long short term memory. *Neural Computation*, 1997.
- [19] A. Karpathy and Fei-Fei Li. Deep visual semantic alignments for generating image descriptions. *CVPR*, 2015.
- [20] I. Sutskever, O. Vinyals, and Q. V. Le. Sequence to sequence learning with neural networks. *NIPS*, 2014.
- [21] E. Krahmer and K. van Deemter. Computational generation of referring expressions. *Comp. Linguistics*, 2012.
- [22] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. Berg, and F. Li. Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 2015.
- [23] J. Donahue, L. Anne Hendricks, S. Guadarrama, M. Rohrbach, S. Venugopalan, K. Saenko, and T. Darrell. Long-term recurrent convolutional networks for visual recognition and description. *CVPR*, 2015.
- [24] R. Stewart and M. Andriluka. End-to-end people detection in crowded scenes. *CVPR*, 2016.
- [25] C. Lu, R. Krishna, M. Bernstein, and F. Li. Visual relationship detection. *ECCV*, 2016.
- [26] R. Hu, H. Xu, M. Rohrbach, J. Feng, K. Saenko, and T. Darrell. Natural language object retrieval. *CVPR*, 2016.
- [27] D. Karatzas. ICDAR 2013 robust reading competition. *ICDAR*, 2013.
- [28] K. Wang and S. Belongie. Word spotting in the wild. *ECCV*, 2010.
- [29] M. Everingham, S. M. Ali Eslami, L. V. Gool, C. Williams, J. Winn, and A. Zisserman. The pascal visual object classes challenge: A retrospective. *International Journal of Computer Vision*, 2015.
- [30] C. Wolf and J. M. Jolion. Object count/area graphs for the evaluation of object detection and segmentation algorithms. *International Journal on Document Analysis and Recognition*, 2006.
- [31] S. Lu, T. Chen, S. Tian, J. Lim, and C. Tan. Scene text extraction based on edges and support vector regression. *IJDAR*, 2015.
- [32] Z. Zhang, W. Shen, C. Yao, and X. Bai. Symmetry-based text line detection in natural scenes. *CVPR*, 2015.
- [33] J. Johnson, A. Karpathy, and Li Fei-Fei. DenseCap: Fully convolutional localization networks for dense captioning. *CVPR*, 2016.
- [34] R. Dahl, M. Norouzi, and J. Shlens. Pixel recursive super resolution.

arXiv:1702.00783, 2017.

- [35] A. Veit, T. Matera, L. Neuman, J. Matas, and S. Belongie. COCO-Text: Dataset and Benchmark for Text Detection and Recognition in Natural Images. arXiv preprint arXiv:1601.07140, 2016.
- [36] X. Rong, C. Yi, and Y. Tian. Unambiguous Text Localization and Retrieval for Cluttered Scenes. CVPR, 2017.

## 作者介绍

**戎雪健**，纽约城市大学城市学院电子工程系在读博士生，本科毕业于南京航空航天大学航天学院航天控制系。当前主要研究领域为自然场景下的文本检测与识别，自然场景识别与解释，图像去模糊等领域。



**田英利**，纽约城市大学研究生院计算机系与城市学院电子工程系终身教授。创办并主持多媒体实验室。加盟纽约城市大学前，曾任 IBM 沃森研究中心（IBM T.J. Watson Research Center）研究员和美国哥伦比亚大学（Columbia University）客座教授。主要研究领域为计算机视觉和人工智能。

# Lower Extremity Exoskeletons and Active Orthoses: Challenges

## 下肢外骨骼机器人：研究进展与关键问题分析

Hao Su\*, Chunzhi Yi

衣淳植, 苏浩\*

The City College, The City University of New York, New York, NY 10031, USA;

纽约城市大学城市学院, 纽约市, 纽约州 10031, 美国;

\*联系人, E-mail: haosu.robotix@gmail.com

### Abstract

In the nearly six decades since researchers began to explore methods of creating them, exoskeletons have progressed from the stuff of science fiction to nearly commercialized products. While there are still many challenges associated with exoskeleton development that have yet to be perfected, the advances in the field have been enormous. In this paper, we review the history and discuss the state-of-the-art of lower limb exoskeletons and active orthosis. We provide a design overview of hardware, actuation, sensory, and control systems for most of the devices that have been described in the literature, and end with a discussion of the major advances that have been made and hurdles yet to be overcome.

### 摘要

在近六十年中，外骨骼机器人在传感器，运动执行器和控制器等领域都发生了重大的进展。尽管还有许多问题需要解决，

但是该领域的发展非常迅速。本文将回顾下肢外骨骼及主动矫正器的发展历程，并对其发展现状进行探讨。并对前有论文中大多数该类设备的硬件、驱动方式、传感方式以及控制系统的设计进行回顾，并在结尾处探讨在该领域中已经取得的主要进展和仍需解决的困难与挑战。

**Keywords:** Exoskeleton, Lower extremity, Orthosis, Orthotics, Rehabilitation, Robotics, Walking, Wearable

**关键词:** 外骨骼 下肢 矫正器 矫正器械 学 康复 机器人学 行走 可穿戴

### 1 引言

带有动力装置的人类外骨骼设备的研究始于 60 年代末期<sup>[1]</sup>，由美国和前南斯拉夫的研究团队几乎同时开展。美国的研究主要集中于增强健全人的能力，通常用于军事目的。而南斯拉夫的研究团队则把辅助残疾人当做主要研究方向。尽管在应用上存在分歧，这两个领域面临着许多共同的挑战和阻碍，尤其是在可移植性和与操

作者方面。外骨骼（exoskeleton）是一种拟人的、被操作者“穿戴”的、跟他/她的身体紧密贴合并能够跟随操作者移动的主动机械设备。总的来说，“外骨骼（exoskeleton）”用来描述能增强健全人能力的设备，而“主动矫正器”用来描述用于增强腿部残疾的人的行走能力的设备。但是，“外骨骼（exoskeleton）”有时也可以用来描述覆盖大部分下肢的矫正器。

本文讲述的下肢外骨骼和矫正器，只包括那些与人的腿部并联连接并操作的设备，而非像 Spring Walker<sup>[2]</sup>那种与操作者串联连接的设备。对于主动矫正器，仅仅讨论能增强下肢关节力量的机器：包括在关节处增加和耗散功率，或者是在步态的各个相位中，有控制的释放存储于弹簧中的能量。也就是说，本文主动矫正器不包括其主动部件仅仅能对关节进行锁定和解锁

的设备，也不包括只是由被动的矫正支架和功能性电刺激（Functional Electrical Stimulation, FES）设备结合的系统。最后，不可移动、不能实现物理上独立的用于治疗的外骨骼设备（比如像 Locomat<sup>[3]</sup>这种基于跑步机的设备），不做讨论。

本文将关注不同设想的最初发展状况，较少关注以研究为目的制造的相似设备。并且将描述关于外骨骼和主动矫正器在完成预期任务时使用效果的定量评估结果。

为了描述后文将用到的一些术语和概念，先简要介绍关于人类行走的生物力学知识。然后，介绍增强型外骨骼和辅助残疾人的外骨骼和矫正器。最后，总结现状，探讨发展前景。

## 2 下肢外骨骼设计的生理学依据

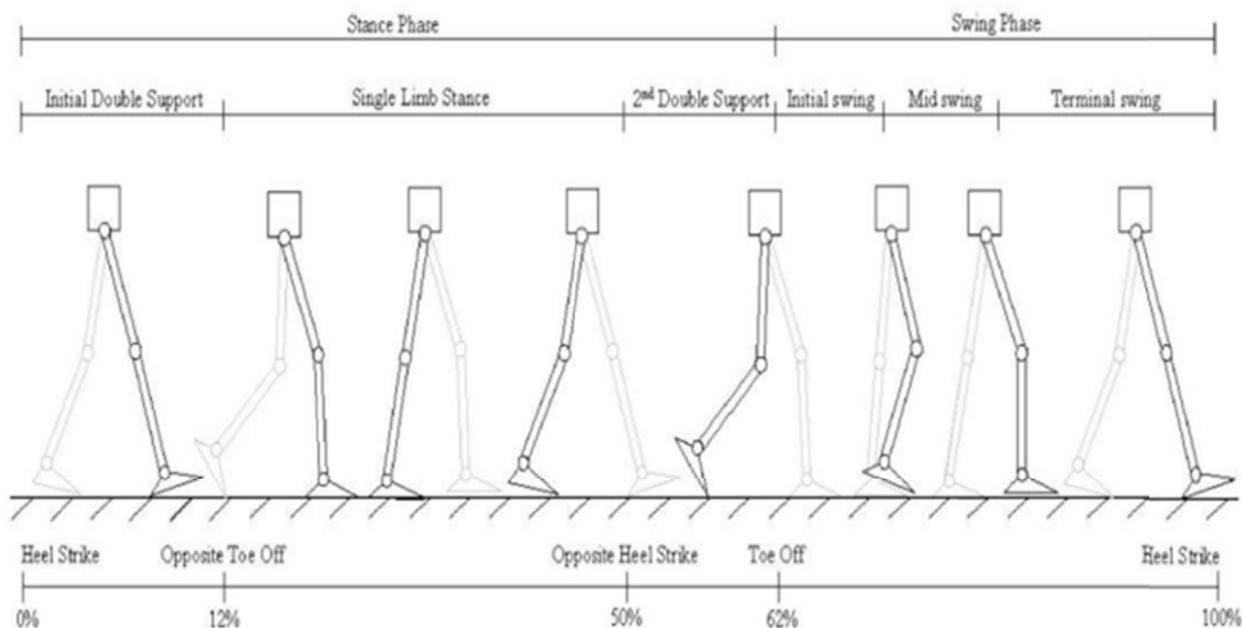


图 1 人行走步态图<sup>[4]</sup>

图 1 (见文献<sup>[4]</sup>) 展示了一个人类行走步态的简图和一些本文需要用到的一些概念。步态周期中被标记出的状态在时间上的分布是极其相似的, 又因个体和情况的不同而异。典型的人类步态周期以相同足跟触底的时间点为开始和结束, 其中另一只足跟触地发生在步态周期的将近 62% 处。

总而言之, 可以认为人类的每条腿是一个 7 自由度的结构, 三个转动自由度在髋关节。一个在膝关节, 三个在踝关节。图 2 展示了人体解剖学平面和人腿部在矢状面内的运动学模型, 其中人腿部的主要运动都在矢状面内 (图 2B)。关节在矢状面内的运动全都叫做弯曲 (正方向) 和伸展 (负方向)。髋关节在额状面内的运动叫做外展 (离开身体中心) 和内收。踝关节在额状面内的运动称为外翻 (远离身体

中心) 和内翻。髋、踝关节的其余自由度的运动都叫做旋转运动。

图 3 展示了一个正常、健康个体 (82 kg, 腿长 0.99 m, 28 岁的男性) 以 1.27 m/s 的速度行走时的生物力学实验。该实验测量了平地行走时三个关节在做弯曲伸展运动时的角度、力矩和功率。用来获取这些数据的实验方法可以在<sup>[5]</sup>中找到。在不同个体和不同条件下行走的数据不同, 但是这些数据具有相似的定性特征<sup>[6]-[9]</sup>。下面, 分析数据, (尤其在低速条件下) 髋关节的功率是正的或者接近零, 膝关节的功率主要是负的 (耗散功率), 踝关节的功率在横纵标上下均匀分布。值得注意的是, 在以不变的速度在地面行走时, 总体上人体所做的净机械功接近于零, 这是因为

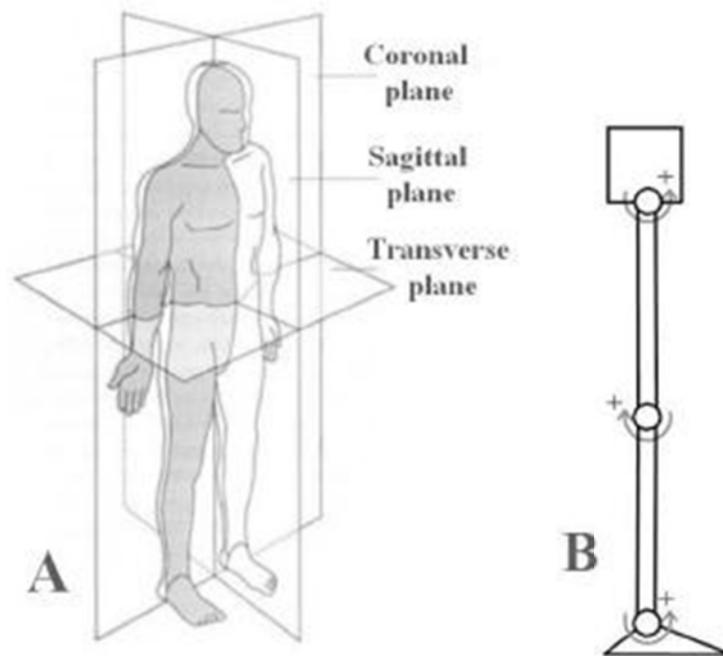


Fig. 2. (A) Description of the anatomical planes. (B) Diagram of the leg shown in the rest position (0 deg at all joints) with the positive direction indicated.

图 2 (A) 人体解剖学平面<sup>[4]</sup> (B) 矢状面内运动方向的规定<sup>[4]</sup>。

人所做总功为零并且运动阻力很小。

考虑到髋膝踝关节在低速行走时的能量转换，动力外骨骼和动力矫正器通常包含一些装置可以在髋关节处做正功，在膝关节处做负功（比如利用一些制动装置或者阻尼器），而且在踝关节处还设置一些从

动的弹性元件。但是，如果操作者的行走速度由慢到快，或者在爬坡或爬楼梯，此人的关节上的功率就会发生较大变化。因此许多设备还可以在膝关节（有时也在踝关节）处做正功。

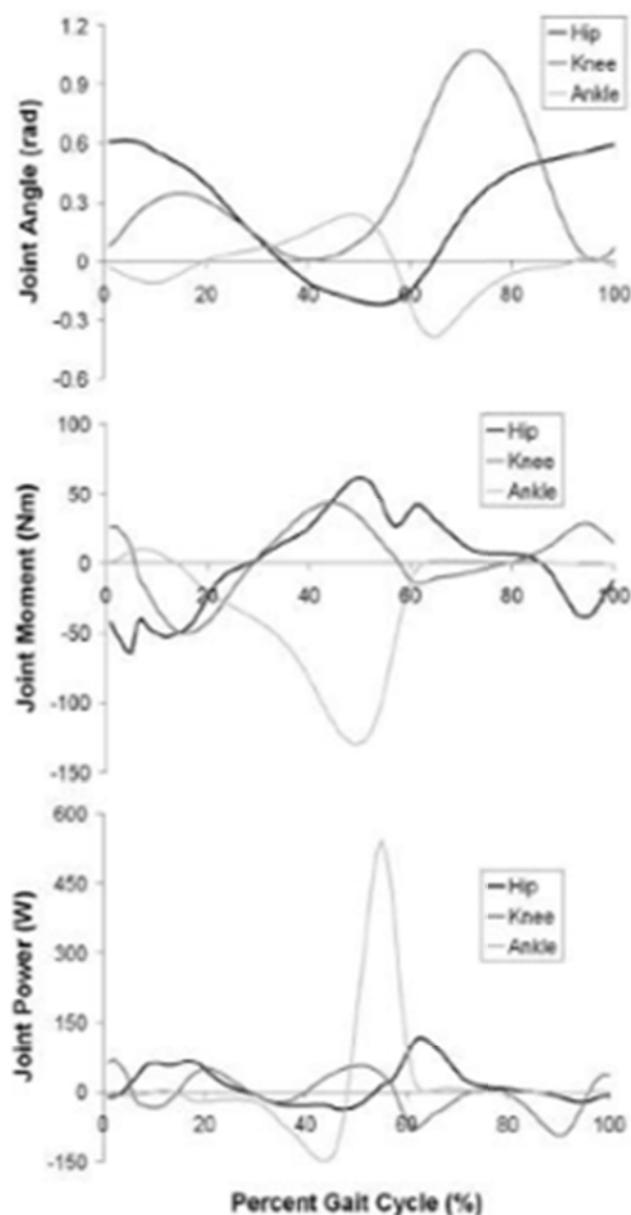


图3 下肢各关节角度、扭矩、功率随步态周期的变化曲线<sup>[5]</sup>。

## 2.1 代谢值

衡量一个增强型外骨骼的性能的一项关键指标就是行走或跑步时的代谢值。通过检测操作者在完成一项特定任务时氧气消耗量与二氧化碳产生量之比，显示这项任务对于操作者来说的费力程度。有很多便宜又简单的设备可以测量这些代谢参数（比如<sup>[11]</sup>中提到的 K4 遥测系统），因此这个指标在经济上是很好测量的。将是否使用外骨骼完成一项工作的代谢值进行对比，可以很好地判断使用一台设备能否省力。

## 3 增强型外骨骼

增强型下肢外骨骼的使用对象是健全人，其目的是使健全人更容易地完成高难度工作或使其完成人类自身无法完成的工作。

### 3.1 完整的下肢外骨骼

1) 伯克利 (Berkeley) 外骨骼 (BLEEX)：BLEEX 与其他外骨骼相比最明显的差别就是它能带着能源移动。实际上，它是第一台“能量自给而且可移动的承重外骨骼”<sup>[13]</sup>。BLEEX 在髋关节上有 3 个自由度，膝关节上有 1 个自由度，踝关节上有 3 个自由度。其中，四个自由度是主动的：髋关节的弯曲/伸展、外展/内收，膝关节的弯曲/伸展，踝关节的弯曲/伸展。被动自由度中，踝关节的外翻/内翻和髋关节的旋转自由度由弹簧承载，而踝关节的旋转自由度不受限制<sup>[14]</sup>。用一个 75 kg 的成年人和行走时的临床步态分析的数据来对这台外骨骼进行运动学和动力需求的分析<sup>[14][15]</sup>。

BLEEX 将人机交互的传感信息的作用降到最小，主要利用外骨骼的传感信息进行控制。和双足机器人相似，这台外骨骼可以自我平衡，但是其使用者必须在行走时为其提供一个向前的引导力。该控制系统利用 8 个编码器和 16 个线性加速度计来测量八个主动关节的角度、角速度和角加速度；每只脚用脚踏开关和负载分布传感器来测量在站立相中地面反馈的力和两脚间的力的分布；利用八个单轴力传感器来控制每个驱动器输出的力；并且考虑到重力的影响，用一个倾角仪来测量背包的方位<sup>[14],[17]</sup>。

根据相关实验，BLEEX 的使用者可以在搬运上限为 75kg 的重物时以 0.9 m/s 的速度行走，而在没有负载的情况下能以 1.3 m/s 的速度行走。第二代伯克利外骨骼正在测试中。改进后的设备大概是原设备重量的一半 (~14kg<sup>[18]</sup>)，部分原因是实现了电驱动和液压变速系统。

2) 萨克斯 (Sarcos) 外骨骼：萨克斯研究公司（位于美国犹他州的盐湖城）在 DARPA EHPA 项目下完成了一个覆盖全身的“可穿戴能源自给机器人 (XoR)”<sup>[18]</sup>。顾名思义，萨克斯外骨骼也是能源自给的，也能携带自己的能源。和伯克利外骨骼类似，萨克斯外骨骼发展了液压驱动外骨骼的技术。然而，萨克斯外骨骼没有使用线性液压驱动器，而是将旋转液压驱动器直接安装在设备的主动关节上。萨克斯公司没有公布这台外骨骼的动力需求，但是他们为了实现有效的液压驱动，在发展能源供应和伺服阀方面做了大量的工作<sup>[18][19]</sup>。

萨克斯外骨骼利用人机间的力传感信息来实现“让路”控制策略。使用者的脚通过一个包含力传感元件的硬金属板来与机器连接，因此，使用者的脚不允许弯曲。

萨克斯外骨骼的结构总承重 84kg，使用者可以在单腿站立时背一个人，在以 1.6m/s 的速度行走时背负 64kg、在胳膊上放 23kg 的重物，可以穿过 23cm 深的泥浆，允许使用者转身、蹲下和跪下<sup>[18], [20]</sup>。

3) HAL (Hybrid Assistive Leg)：由日本筑波大学的山海教授和他的团队设计了一种增强人体机能和康复目的的外骨骼<sup>[29][30]</sup>。这种覆盖全身的外骨骼 HAL-5 的腿部对应人体髌关节和膝关节的弯曲/伸展方向的关节由安装在关节上的谐波激励的直

流电机驱动（见图 4）。踝关节的弯曲/伸展方向的自由度是被动的。其与使用者的连接如下：套在小腿和大腿上带有地面反作用力传感器的特制的鞋子，和一条腰带。值得注意的是，HAL 与 BLEEX, Sarcos, MIT 外骨骼的区别是 HAL 不能将重量传递给地面，而仅仅是增强髌膝踝关节上的力矩。

HAL-5 系统利用了一些传感方式来进行控制：放置在髌关节下方，膝关节上方，位于人体前后两侧的皮肤表面肌电信号（skin-surface electromyographic, EMG）电极；测量关节角度的电位器，地面反作用力传感器，测量躯干姿势的安装在背包中的陀螺仪和加速度计。

尽管有很多健全人使用 HAL 的实例，



图 4 a) BLEEX<sup>[13]</sup>, b) XoR<sup>[18]</sup>, c) 护士辅助外骨骼<sup>[32]</sup>, d) BE (Body Extender)<sup>[34]-[37]</sup>, e) HAL<sup>[29]</sup>, f) 哈佛 soft exosuit<sup>[39]</sup>。

但是并没有找到残疾人使用 HAL 的例子。在这些实例中，HAL 增强人用手臂抬起重物的能力已经得到证明。然而，其各个部件的作用还不清楚。Cyberdyne 公司的网站上声称 HAL 的使用者可以用手臂抬起最大 40kg 的重物，并且可以使使用者的承受“腿部压力”的能力从 100kg 增加到 180kg。

4) 护士辅助外骨骼 (Nuuse-Assistive Exoskeleton)：日本神奈川理工学院的研究人员用十多年的时间发明了一种能够协助护士转移病人的外骨骼<sup>[32][33]</sup>。这台设备的组成部件包括：直接安装在膝关节弯曲/伸展方向的直驱气动旋转驱动器。气体的压力由直接安装在每个驱动器旁的多个小气泵提供，这就使该外骨骼变成完全可移动的。但是该外骨骼非动力的关节和其与操作者的连接方式不清楚。

神奈川理工学院的这台外骨骼的机械机构上特点之一是操作者身前没有任何的机械零件，这就使护理人员能与他/她抱起的病人在物理上有直接的接触。这对确保病人的舒适性和安全性方面很重要。

5) 意大利“BE” (Body Extender)<sup>[34]-[37]</sup>：一款全身型的外骨骼，由意大利国防部支持研发。该设备总重 160kg，具有 22 个自由度，每个自由度均采用直流有刷电机驱动，每条腿拥有 6 个自由度，每只手臂有 4 个自由度，每只手臂前段还有 1 个自由度的手抓，用于抓取物体。在手臂平举状态下，每只手臂可以承载 50 kg。在控制方式上，外骨骼和人的 5 个接触点（双手、双脚、躯干）都装有六维力传感器，并在关节部位安装角度传感器配合电

机编码器，躯干上的加速度计等，使外骨骼能够顺利的跟随人体的运动。

6) 法国 RB3D 公司的“HERCULE”系列下肢助力外骨骼<sup>[38]</sup>：2012 年推出了一款能够帮助穿戴者运送重物的全身外骨骼服“HERCULE V2”，髋关节和膝关节各有一个主动自由度，采用电机驱动，能够辅助穿戴者上下楼梯、斜坡，或进行蹲下/站起动作，在搬运 100 kg 重物的条件下行走自如。在 2013 年，RB3D 公司推出了针对民用的下肢助力外骨骼“HERCULE V3”，重量为 30kg，采用锂电池供电，续航时间 4h。胸前的平台可以用来放置重物。

### 3.2 模块化下肢外骨骼

哈佛大学开发的软性外骨骼机械服“Soft Exosuit”，覆盖使用者的从腰部到脚后跟的整个下肢。初代样机采用气动肌肉驱动，以提高柔性并降低质量<sup>[39]</sup>。总重约 9.1 kg；能源及控制部分都安装到后背。之后的改进版本在外骨骼髋关节和踝关节上采用伺服电机和钢丝绳传动的方式主动助力，如图 4<sup>[40]</sup>。实验表明，改进后的样机可以在负重 24kg 的条件下，以 1.25m/s 的速度行走，并能成功降低人体正常行走所消耗新陈代谢量的 14.6%。

在控制方面，exosuit 采用基于能量的位置控制策略，当检测到人体踝关节的能量过零时，绳索拉伸，通过具有弹性的软性材料，提供辅助力。由于无法检测人体关节的能量，因此根据人体踝关节速度与能量的对应关系，通过踝关节过零点来判断辅助力输出的时机。

### 3.3 单关节下肢外骨骼

1) Panasonic suit : 由日本松下在 AWN-03 之后推出, 重量只有约 13 磅<sup>[41]</sup>。采用轻量的碳纤维马达进行驱动, 同样采用碳纤维材料支撑以减轻机器人的重量。在穿戴方式上同人体的背部和大腿捆绑连接, 并当穿戴者举起或携带一个物体时, 感应器会激活马达。但是有效负载仅 33 磅, 以帮助普通工人进行负重工作。

2) 卡耐基梅隆大学的机械工程学院开发了一款完全被动的踝关节助力外骨骼 (如图 5), 通过与踝关节和小腿并联的弹簧来储存人体踝关节在行走过程耗散的能量, 并在适当的时候释放。通过仿真得到弹簧的最优刚度, 从而达到优化人类行走过程中代谢率的目的<sup>[42, 43]</sup>。该设备利用一个离合器, 在摆动相时松开弹簧, 在步态周期的其余相位中绷紧弹簧。经试验, 使用该设备, 人行走消耗的代谢率减少了约 7.2%。

3) 机械膝关节 (Roboknee): 美国 Yobotics 股份有限公司 (美国俄亥俄州辛辛

那提市, [www.yobotics.com](http://www.yobotics.com)) 发明了一种在完成负载搬运任务时在膝关节处做正功来辅助操作者爬楼梯和蹲下的单一外骨骼<sup>[44, 45]</sup>。这台设备由连接在膝关节上方 (就在髌关节下方)、下方 (在小腿处) 的线性串联弹性执行器 (series elastic actuator, SEA)。这台设备的设计理念是要在机器与操作者间的相互作用呈物理上的低阻抗时, 在膝关节处做正功。这就允许设备在保证操作者安全的前提下, 获得更高的控制增益。

### 4 主动矫正器 (Active Orthoses)

主动矫正器指传统被动支架结合在设备的关节处做正功或负功结合把在步态周期的某相位中储存的能量有控制的释放的元件的设备。

#### 4.1 完整的下肢矫正器

1) 德拉华大学 (University of Delaware) 的机械工程与物理治疗系发明了一种能减少患者行走过程中承受的重力和减轻移动所需的力量的被动腿部矫正器<sup>[46]</sup>。这台设备为了在几何上定位腿部矫正器系统的重



图 5 a) Panasonic suit<sup>[41]</sup>, b) 卡耐基梅隆大学被动踝关节助力装置<sup>[42]</sup>, c) 机械膝关节<sup>[44]</sup>。

心从而抵消重力的影响，有趣地将弹簧和连杆组合应用。

实验结果展示了这台设备的效果，在没有影响膝关节所需力矩的条件下，操作者髋关节所需的力矩平均减少 61%。这组研究人员仍然在这一领域积极的研究，期待他们由更多的成果。

2) 韩国首尔西江大学的研究人员提出了一个想要减轻开发可移动主动矫正器难度的有趣的设计理念<sup>[47]</sup>。该设备由两部分组成：一台完整的下肢外骨骼和一个装有电池、直流电机、控制用电脑的特制的步行器，这大大减轻了矫正器的重量。步行器中的驱动器通过绳索将机械能传递给操作者的关节。正是由于这种动力传递模式，操作者需要和步行器保持一个固定的距离。这台矫正器在髋关节和膝关节的弯曲/伸展方向做正功，并且除了踝关节的旋转运动是固定的之外，允许腿部其他自由度方向的运动。通过关节角度传感器结合传递四头肌处的力传感信息的压力传感器来判断操作者的运动意图。

3) ReWalk: 以色列开发了一款名为

ReWalk 的助力装置<sup>[48]</sup>，用于帮助因脊髓损伤等原因导致下肢截瘫的残疾人行走。它是通过拐杖来维持身体平衡，它主要由电动腿部支架、身体感应器和背包组成。采用机器主动控制模式，人体跟随机械运动。系统提供的助力由控制系统确定，不随人的动作意愿而随时改变。在行走模式下，背包上的陀螺仪传感器将对运动意图进行检测，当使用者向前倾斜，伸出某侧拐杖时，背包中的计算机将采集的传感器信息分析，产生步态指令指导髋关节以及膝关节的电机带动这条腿进入摆动相，另一条腿与之相反。这是第一款通过美国食品药品监督管理局（FDA）认证的可在家庭和社区中使用的个人外骨骼设备。通过对该款产品的临床研究，表明截瘫患者经过训练可以自如地使用它站立行走，并且生活质量得到了很大的改善。

3) 埃克索矫正器 (Ekso exoskeleton): 美国里士满市 (Richmond, USA) 的埃克索仿真学科技公司 (Eksobionics Ltd) 生产的矫正器 Ekso 是一个可穿戴的下肢矫正器，它的两条腿与一个装有电脑和电池得到躯干结构相连<sup>[49]</sup>。躯干结构与使用者腰部相



图 6 a) 韩国首尔西江大学设计的外骨骼<sup>[47]</sup>, b) Rewalk<sup>[48]</sup>, c) Ekso<sup>[49]</sup>, d) Indego<sup>[51]</sup>。

连，该设备的腿部结构通过钩环扣皮带与使用者腰部和关节相连。使用者肩膀上还有两根背带用于辅助支撑该设备的躯干结构。该设备的髋关节和膝关节在矢状面内是动力驱动的（由电池提供动力），其踝关节在矢状面内的运动是有弹簧承载的，其他方向的运动时受限制的。现在，Ekso 拥有四中行走模式：在前两种模式中，物理治疗师或者使用者都可以通过按钮来激活坐-立转换模式和上楼梯模式；而在后两种模式中，通过探测使用者髋关节向前和横向的移动（来完成重心的转换）来判断步态意图，从而指挥该设备运动。Ekso 需要用拐杖来保证使用者的稳定和安全；拐杖的底部装有力传感器，以保证拐杖稳固地放在地面上，并且至少支撑了身体部分的重量。除非拐杖稳固地放在地上，否则该设备将不会运动。另外该设备用分层有限状态机来作为控制器来控制不同步态间的过渡，以保证使用者的安全<sup>[50]</sup>。制造商声称，Ekso 的更新的版本能为使用者提供多

种模式的辅助以及按需辅助的力量。因此，治疗师就能根据患者的情况为其选择一个特定数值的力，或者让矫正器实时地根据患者需要而进行动力上的调整。

## 4.2 模块化主动矫正器

**Indego:** 美国派克哈尼芬公司的 Indego 模块化矫正器由双边的动力髋关节、膝关节组成，还包含了为使用者增强踝关节稳定性和将矫正器重量传递给地面的内置 AFO<sup>[51]</sup>。该设备的髋关节、大腿、小腿的模块可以通过快速链接装置装配，可以轻易地穿上、脱下、运输和储存。它紧凑的轮廓可以与标准器具兼容，可以在坐轮椅的时候使用。该设备能够完成坐下、行走、站立、坐-立过渡、立-行过渡、行-立过渡、立-坐过渡的动作。运动和非运动任务的意图是通过上身与步行器结合，基于使用者影响其压力中心（Center of Pressure, CoP）的能力进行判断的。控制器用 CoP 与前侧踝关节的位置的距离作为指挥设备动



图 7 a) 密歇根大学踝关节矫正器<sup>[52]-[54]</sup>, b) Kinesis<sup>[55]</sup>。

作的主要输入。这需要使用者向前或向后倾斜身体,使 CoP 前/后移动,从而控制设备过渡到一个不同的运动模式中。出于安全考虑,在膝关节处有制动器来防止膝关节在停电时弯曲<sup>[50]</sup>。

### 4.3 单关节主动矫正器

1) 密歇根大学 (University of Michigan): 密歇根大学的人类神经机械学实验室 (the Human Neuromechanics Laboratory) 制造了一批主动外骨骼,尤其关注用于医疗的康复设备<sup>[52]-[54]</sup>。因此,这些设备并不是完全可移动的,并且大部分是由气压驱动的,通过一条拴带与一个固定式空气压缩机相连。该设备用到的气压驱动器是安装在碳纤维和聚丙烯材质的外壳上的人造气动肌肉,这种材质的外壳导致了设备在保持高输出的同时,质量很轻。另外,驱动器的低阻抗特性增强了设备的安全性。

2) Kinesis:由西班牙马德里的 Technaid 公司研制的 Kinesis 是一台下肢康复机器人,为了提供混合行走治疗给不完全脊髓损伤的患者。其目标人群是可以进行功能性恢复的患者,比如可以短距离行走但是需要依赖轮椅才能达到社区性步行能力(通常保留了髌关节屈肌的部分能力)。Kinesis 是一台膝关节-踝关节-足部矫正器,在膝关节处有一个主动驱动器,而在踝关节跖屈背屈方向有一个被动的弹性执行机构,用力敏电阻来探测足-地接触,用电位器来监视关节角位置,用基于应变片的嵌入式惠斯通电桥来探测人-机作用力矩<sup>[55]</sup>。Kinesis 能够通过表面电极向膝关节的伸肌和屈肌提供 FES。

Kinesis 有一些内置安全措施:机器人关节在人移动的生理极限处由机械制动装置;通过在膝关节的导纳控制器里的软件限制最低和最高位置,最大输出力矩,以及 FES 脉冲信号的持续时间和幅值;还有用来切断所有硬件能源的机械安全按钮<sup>[56]</sup>。

3) 其他踝关节-足部外骨骼:亚利桑那州立大学的研究者们提出了一个新奇的设计:用在拇脚趾下方,连接足部左侧和右侧(和足跟形成一个三脚架)的两个“肌肉上方的弹簧(spring over muscle)”驱动器驱动的主动 AFO<sup>[57]</sup>。这些驱动本质上是内部有一个倾向于拉伸肌肉的、能在跖屈背屈方向都使力的弹簧气动肌肉。这个三脚架结构能够让踝关节在弯曲/伸展方向上同时激励,在外展/内收方向上单一方向激励。另外,该研究小组也对用 SEAs 驱动矫正器关节进行了探索<sup>[58]</sup>。

特拉华大学的机械工程与物理治疗系 (Departments of Mechanical Engineering and Physical Therapy at the University of Delaware) 的研究者们也提出了一个关于在使用者弯曲/伸展方向和外展/内收方向上增加力量的主动踝关节矫正器的设计<sup>[59]</sup>。

4) 主动膝关节矫正器:柏林工业大学的研究者发明了一种通过直流电机和滚珠螺杆传动的在膝关节处增加力量的矫正器<sup>[60][61]</sup>。但是到目前为止的工作主要是基于 EMG 的控制系统的研究。

## 5 研究趋势与总结

在写这篇综述的过程中,一些关于功能建立、能源自给、主动矫正设备的挑战不断出现。在这些争论点当中,能量供应、

轻质量驱动器以及有效的动力传输是所有的研究人员都不得不面对的问题。很显然，尤其对于那些已经处于外骨骼发展的较高阶段的研究机构而言，为了达到设计目标，现成的工业化产品不能满足能源、驱动和其他子系统的低重量、高效率以及其它条件的需求<sup>[13], [18], [40], [41]</sup>。实际上，这是一个很多可移动机器人领域需要面对的问题，尤其是那些有拟人机构的机器人。

尽管这些问题还有待解决，在最近五年当中，做出了有很多在外骨骼和主动矫正器方面的巨大进展。近十年来，人体生物力学领域逐渐成熟，为最大限度模仿操作者运动的设备的设计提供了科学背景。电池与直流电机技术在近十年来取得大幅进展，但它们仍然不能满足许多外骨骼的需求。计算、检测、控制方面的技术发展水平十分领先，以至于这些领域不再是机器人硬件的实现上的主要障碍。

### 5.1 增强型外骨骼

在这方面，完整的、能源自给的系统的优点大多是不可靠的。实际上，缺乏关于外骨骼设备能提高人类运动能力的已发表的定量试验结果。仅有几个文献<sup>[24]-[26]</sup>给出的结果并没有证实现有设计的优点。人们想知道这些复杂、昂贵的设备的真正优势在哪里。当然，外骨骼在使佩戴者完成一项他或她自身不能完成的任务时是有价值的<sup>[13], [18]</sup>。但是，如果外骨骼帮助佩戴者完成其本身也能完成的任务而不减少其代谢值和/或疲劳感的话，那么它是几乎没有价值的<sup>[71]-[73]</sup>。除了用代谢值衡量外骨骼运动上的性能外，适合衡量这些类型的系统的其他性能参数包括：肌肉-骨骼系统负担

力量的减少量、肌肉疲劳感的减少、双足稳定性的提高。

外骨骼定量实验结果的缺乏突出了还有大量的挑战。当然，还有很多设计上的问题可能会导致性能下降：操作者的关节与硬件的关节不重合、人机连接方式带来的运动学上的限制、没有对搬运负载时的步态进行优化设计<sup>[70]-[76]</sup>、施加到操作者身上抗拒运动的力、没有以最佳的方式在操作者身上施加力（比如：时机错误、力太小、力太大）等等。所有这些问题都非常难以解决，但是基础研究很有可能解决这些问题。

### 5.2 主动矫正器

主动矫正器由于患者患病的情况有极大的不同，使得发明一个实用性强的设备变得很困难。事实上，这对于很多辅助设备而言都是一个挑战。市面上还没有能源自给的主动矫正器。现在单纯处于治疗目的的外骨骼是那种独立的、基于跑步机的设备，比如 Locomat<sup>[3]</sup>；但是发明一个能在医院、诊所外使用的可移动的设备是有极大价值的。理想情况下，人们会喜欢一个在使用者日常生活中既能提供辅助又能提供治疗的、结构紧凑的、能源自给的矫正器。

可移动性的问题是限制主动矫正器在医院外应用的主要因素之一。本篇综述覆盖的矫正器中的大多都是无法能源自给的，它们都被一些外部的能源供应设备限制住——空气压缩机、液压泵、或者电源。

主动矫正器方面也缺乏证明其有效性的定量实验。与已有的辅助设备进行比较

是一个有效的方法。例如，一台主动矫正器，想要辅助一个人运动，而这个人还可以在 RGO 的辅助下运动，那么其实验结果就应该与使用 RGO 的实验结果进行对比。合适的性能参数包括：运动过程中的代谢值<sup>[70]-[72]</sup>、行走速度、运动的可重复性和柔顺性、肌肉疲劳度、以及稳定性，等等。

### 5.3 展望

外骨骼与主动矫正器未来的发展方向有可能会围绕着“使能”技术，比如能量供应、驱动器、和轻便高效的动力传输。外骨骼方面未来发展所必需的技术（比较普遍的应用，比如：计算、检测、以及控制）的研究是由外骨骼研究机构自身推动的。

还有一些关于外骨骼机械结构的设计方面的领域被忽略了。随着对在行走和其他动作中肌肉和肌腱发挥的作用得理解的加深，会有更有效的外骨骼腿部结构被设计出来。基于能捕获人运动主要特点的实际机械元件的步态模型将会提高对人类腿部形态与控制的理解，并导致高效率低质量的外骨骼方面类似的提高。

非拟人结构的研究将会提供在外骨骼结构方面的一些问题的解决方法，比如：机器人和佩戴者关节轴线共线的要求。而且，还很少有人做“娱乐”外骨骼方面的工作，比如增强跑步和跳高能力的外骨骼，这个领域有可能是将来发展的重点。

除了使能技术机械结构设计之外，有很多关于实现外骨骼和主动矫正器的问题被大大地忽略了。关于被困在动力外骨骼内操作者的安全问题的研究还有待研究。另外，有效的人机交流方法，不管在机械

上还是在神经上，都是未来的研究领域。

### 参考文献

- [1] R. A. Heinlein, *Starship Troopers*[M]. New York: Putnam, 1959.
- [2] J. Dick and E. Edwards. Human bipedal locomotion device: U.S.Patent 5 016 869[P], 1991.
- [3] Lokomat Rehabilitation Device. (2007). Hocoma AG Medical Engineering, Switzerland [Online]. Available: [www.hocoma.ch](http://www.hocoma.ch)
- [4] J. Rose and J. G. Gamble. *Human Walking*[M]. 2nd ed. Baltimore. MD:Williams and Wilkins, 1994.
- [5] M. Popovic, A. Goswami, and H. Herr. Ground reference points in legged locomotion: Definitions, biological trajectories and control implications[J]. *Int. J. Robot. Res.*, vol. 24, no. 12, pp. 1013–1032, 2005.
- [6] A. J. Van Den Bogert,. Exotendons for assistance of human locomotion[J].*Biomed. Eng. Online*, vol. 2, p. 17, 2003.
- [7] C. Kirtley. (2005). CGA Normative Gait Database, HongKong Polytechnic University. [Online]. Available: <http://guardian.curtin.edu.au/cga/data>
- [8] J. Linskill. (2005). CGA Normative Gait Database, Limb Fitting Centre, Dundee, Scotland [Online]. Available: <http://guardian.curtin.edu.au/cga/data>
- [9] A. Winter. (2005). Gait Data, International Society of Biomechanics, Biomechanical Data Resources [Online]. Available: <http://guardian.curtin.edu.au/org/data>
- [10] J. M. Brockway. Derivation of formulae used to calculate energy expenditure in man[J]. *Hum. Nutrition: Clin. Nutrition*, vol. 41, pp. 463–471,1987.
- [11] J. M. Donelan, R. Kram, and A. D. Kuo. Mechanical work for step-to-step transitions is a major determinant of the metabolic cost of human walking[J]. *J. Exp. Biol.*, vol. 205, pp. 3717–3727, 2002.
- [12] E. Garcia, J. M. Sater, and J. Main.

- Exoskeletons for human performance augmentation (EHPA): A program summary [J]. *J. Robot. Soc. Japan*, vol. 20, no. 8, pp. 44–48, 2002.
- [13] H. Kazerooni and R. Steger. The Berkeley Lower Extremity Exoskeleton [J]. *Trans. ASME, J. Dyn. Syst., Meas., Control*, vol. 128, pp. 14–25, Mar. 2006.
- [14] A. B. Zoss, H. Kazerooni, and A. Chu. Biomechanical design of the Berkeley Lower Extremity Exoskeleton (BLEEX) [J]. *IEEE/ASME Trans. Mechatronics*, vol. 11, no. 2, pp. 128–138, Apr. 2006.
- [15] A. Chu, H. Kazerooni, and A. Zoss. On the biomimetic design of the Berkeley Lower Extremity Exoskeleton (BLEEX) [J]. in *Proc. IEEE Int. Conf. Robot. Autom.*, Barcelona, Spain, 2005, pp. 4345–4352.
- [16] A. Zoss and H. Kazerooni. Design of an electrically actuated lower extremity exoskeleton [J]. *Adv. Robot.*, vol. 20, no. 9, pp. 967–988, 2006.
- [17] K. Amundson, J. Raade, N. Harding, and H. Kazerooni. Hybrid hydraulic-electric power unit for field and service robots [J]. in *Proc. 2005 Int. Conf. Intell. Robots Syst.*, pp. 3453–3458.
- [18] E. Guizzo and H. Goldstein. The rise of the body bots [J]. *IEEE Spectr.*, vol. 42, no. 10, pp. 50–56, Oct. 2005.
- [19] G. T. Huang. Wearable robots [R]. *Technol. Rev.*, pp. 70–73, Jul./Aug. 2004.
- [20] U.S. Army Research Laboratory. 2006 ARO in review [z]. U.S. Army Research Office, Adelphi, MD, 2006.
- [21] C. J. Walsh, K. Pasch, and H. Herr. An autonomous, underactuated exoskeleton for load-carrying augmentation [J]. in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Beijing, China, 2006, pp. 1410–1415.
- [22] C. J. Walsh, D. Paluska, K. Pasch, W. Grand, A. Valiente, and H. Herr. Development of a lightweight, underactuated exoskeleton for load-carrying augmentation [J]. in *Proc. IEEE Int. Conf. Robot. Autom.*, Orlando, FL, 2006, pp. 3485–3491.
- [23] A. Valiente. Design of a quasi-passive parallel leg exoskeleton to augment load carrying for walking [D]. Master's thesis, Dept. Mech. Eng., Massachusetts Inst. Technol., Cambridge, Aug. 2005.
- [24] C. J. Walsh. Biomimetic design of an underactuated leg exoskeleton for load-carrying augmentation [D]. Master's thesis, Dept. Mech. Eng., Massachusetts Inst. Technol., Cambridge, Feb. 2006.
- [25] C. J. Walsh, K. Endo, and H. Herr. Quasi-passive leg exoskeleton for load-carrying augmentation [J]. *Int. J. Hum. Robot.* vol. 4, no. 3, pp. 487–506.
- [26] K. N. Gregorczyk, J. P. Obusek, L. Hasselquist, J. M. Schiffman, C.K. Bensele, D. Gutekunst, and P. Frykman. The effects of a lower body exoskeleton load carriage assistive device on oxygen consumption and kinematics during walking with loads [C]. presented at the 25th Army Sci. Conf., Orlando, FL, Nov. 27–30, 2006.
- [27] J. F. Jansen, J. F. Birdwell, A. C. Boynton, H. P. Crowell, III, W. K. Durfee, J. D. Gongola, S. M. Killough, D. J. Leo, R. F. Lind, L. J. Love, M. Mungiole, F. G. Pin, B. S. Richardson, J. C. Rowe, O. A. Velev, and T. Zambrano. Phase I report: DARPA exoskeleton program [P]. Oak Ridge Nat. Lab., Oak Ridge, TN, Rep. ORNL/TM-2003/216, 2003.
- [28] P. Marks. Power dressing [J]. *New Scientist*, vol. 2316, pp. 33–35, Nov. 2001.
- [29] M. Gogola, E. J. Barth, and M. Goldfarb. Monopropellant powered actuators for use in autonomous human scale robotics [J]. in *Proc. IEEE Int. Conf. Robot. Autom.*, Washington, DC, 2002, pp. 2357–2362.
- [30] H. Kawamoto and Y. Sankai. Power assist system HAL-3 for gait disorder person [C]. in *Proc. Int. Conf. Comput. Helping People Special Needs (ICCHP) (Lecture Notes on Computer Science)*, vol. 2398, Berlin, Germany: Springer-Verlag, 2002.
- [31] H. Kawamoto, S. Lee, S. Kanbe, and Y. Sankai. Power assist method for HAL-3 using EMG-based feedback controller [C] in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, 2003, pp. 1648–1653.
- [32] K. Yamamoto, K. Hyodo, M. Ishii, and T.

- Matsuo. Development of power assisting suit for assisting nurse labor [J]. *JSME Int. J., Ser. C*, vol. 45, no. 3, pp. 703–711, 2002.
- [33] K. Yamamoto, M. Ishii, K. Hyodo, T. Yoshimitsu, and T. Matsuo. Development of power assisting suit (miniaturization of supply system to realize wearable suit) [J]. *JSME Int. J., Ser. C*, vol. 46, no. 3, pp. 923–930, 2003.
- [34] Marco Fontana, Rocco Vertechy, Simone Marcheschi, Fabio Salsedo, and Massimo Bergamasco. —The Body Extender: A Full-Body Exoskeleton for the Transport and Handling of Heavy Loads[J], *IEEE Robotics & Automation Magazine*, 2014, 21(4): 34-44.
- [35] Massimo Bergamasco, Fabio Salsedo, Simone Marcheschi, Nicola Lucchesi and Marco Fontana. —A Novel Compact and Lightweight Actuator for Wearable Robots[C], 2010 IEEE International Conference on Robotics and Automation, Anchorage, USA, 2010:4197-4203.
- [36] N. Lucchesi, S. Marcheschi, L. Borelli, F. Salsedo, M. Fontana, M. Bergamasco. An Approach to the Design of Fully Actuated Body Extenders for Material Handling[C]. 19th IEEE International Symposium on Robot and Human Interactive Communication. Viareggio, Italy, 2010:482-487.
- [37] Simone Marcheschi, Fabio Salsedo, Marco Fontana and Massimo Bergamasco. Body Extender: whole body exoskeleton for human power augmentation[C]. 2011 IEEE International Conference on Robotics and Automation. Shanghai, China, 2011:611-616.
- [38] X. Liu, K. H. Low, and H. Y. Yu. Development of a lower extremity exoskeleton for human performance enhancement [C]. in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sendai, Japan, 2004, pp. 3889–3894.
- [39] K. H. Low, X. Liu, and H. Y. Yu. Development of NTU wearable exoskeleton system for assistive technologies [C]. in *Proc. IEEE/RSJ Int. Conf. Mechatronics Autom.*, Niagara Falls, ON, Canada, 2005, pp. 1099–1106.
- [40] T. Onishi, T. Arai, K. Inoue, and Y. Mae. Development of the basic structure for an exoskeleton cyborg system [J]. *Artif. Life Robot.*, vol. 7, pp. 95–101, 2003.
- [41] 法国 RB3D 公司. <http://www.rb3d.com/en/>
- [42] E. V. Kosso. A minimum energy exoskeleton [C]. in *Proc. Carnahan Conf. Electron. Prosthetics*, 1973, pp. 86–89.
- [43] G. L. Cobb. Walking motion [P]. U.S. Patent 2 010 482, 1935.
- [44] J. E. Pratt, B. T. Krupp, C. J. Morse, and S. H. Collins. The Robo Knee: An exoskeleton for enhancing strength and endurance during walking [C]. in *Proc. IEEE Int. Conf. Robot. Autom.*, New Orleans, LA, 2004, pp. 2430–2435.
- [45] P. Filippi. Device for the automatic control of the articulation of the knee applicable to a prosthesis of the thigh [P]. U.S. Patent 2 305 291, 1942.
- [46] Collins S H, Wiggin M B, Sawicki G S. Reducing the energy cost of human walking using an unpowered exoskeleton[J]. *Nature*, 2015, 522(7555):212-215.
- [47] Grimmer M, Eslamy M, Glied S, et al. A comparison of parallel- and series elastic elements in an actuator for mimicking human ankle joint in walking and running[C]// IEEE International Conference on Robotics and Automation. IEEE, 2012:2463-2470.
- [48] M. Vukobratovic, B. Borovac, D. Surla, and D. Stokic. *Scientific Fundamentals of Robotics 7, Biped Locomotion: Dynamics Stability, Control, and Application*[M]. New York: Springer-Verlag, 1990.
- [49] <http://eksobionics.com/>
- [50] Strausser K A and Kazerooni. The development and testing of a human machine interface for a mobile medical exoskeleton [C]. *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)* (Piscataway, NJ: IEEE)., 2011.
- [51] Strausser K A and Kazerooni. The development and testing of a human machine interface for a mobile medical exoskeleton [C]. *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)* (Piscataway, NJ: IEEE)., 2011.

- [52] S. K. Banala, S. K. Agrawal, A. Fattah, V. Krishnamoorthy, W. Hsu, J. Scholz, and K. Rudolph. Gravity-balancing leg orthosis and its performance evaluation [J]. *IEEE Trans. Robot.*, vol. 22, no. 6, pp. 1228–1239, Dec. 2006.
- [53] K. Kong and D. Jeon. Design and control of an exoskeleton for the elderly and patients [J]. *IEEE/ASME Trans. Mechatronics*, vol. 11, no. 4, pp. 428–432, Aug. 2006.
- [54] Y. Mori, J. Okada, and K. Takayama. Development of straight style transfer equipment for lower limbs disabled “ABLE” [C]. in *Proc. IEEE/ASME Int. Conf. Adv. Intell. Mechatronics*, 2005, pp. 1176–1181.
- [55] Hartigan C, Kandilakis C, Dalley S, Clausen M, Wilson E, Morrison S, Etheridge S and Farris R. Mobility outcomes following five training sessions with a powered exoskeleton *Top. Spinal Cord Inj. Rehabil*[Z]. 21 93–9, 2015.
- [56] Quintero H, Farris R, Hartigan C, Clesson I and Goldfarb M. A powered lower limb orthosis for providing legged mobility in paraplegic individuals *Top. Spinal Cord Injury Rehabil*[Z]. 17 25–33, 2011
- [57] D. P. Ferris, J. M. Czerniecki, and B. Hannaford, “An ankle-foot orthosis powered by artificial muscles,” *J. Appl. Biomech.*, vol. 21, pp. 189–197, 2005.
- [58] D. P. Ferris, K. E. Gordon, G. S. Sawicki, and A. Peethambaran. An improved powered ankle-foot orthosis using proportional myoelectric control [J]. *Gait Posture*, vol. 23, pp. 425–428, 2006.
- [59] G. S. Sawicki, K. E. Gordon, and D. P. Ferris. Powered lower limb orthoses: Applications in motor adaptation and rehabilitation [C]. in *Proc. 2005 IEEE Int. Conf. Rehabil. Robot. (ICORR)*, pp. 206–211.
- [60] del-Ama A J, Moreno J C, Gil-Agudo A, de-Los-Reyes A and Pons J L. Online assessment of human-robot interaction for hybrid control of walking *Sensors*[Z] 12 215–25, 2012.
- [61] del-Ama A J, Gil-Agudo Á, Pons J L and Moreno J C. a pilot study *Front. Hum. Neurosci*[J]. Hybrid gait training with an overground robot for people with incomplete spinal cord injury. pp: 8 298, 2014
- [62] K. W. Hollander, R. Hg, T. G. Sugar, and D. Herring. An efficient robotic tendon for gait assistance [J]. *Trans. ASME, J. Biomech. Eng.*, vol. 128, pp. 788–791, 2006.
- [63] K. Bharadwaj, T. G. Sugar, J. B. Koeneman, and E. J. Koeneman. Design of a robotic gait trainer using spring over muscle actuators for ankle stroke rehabilitation [J]. *Trans. ASME, J. Biomech. Eng.*, vol. 127, pp. 1009–1013, 2005.
- [64] A. Agrawal, S. K. Banala, S. K. Agrawal, and S. A. Binder-Macleod. Design of a two degree-of-freedom ankle-foot orthosis for robotic rehabilitation [C]. in *Proc. 2005 IEEE Int. Conf. Rehabil. Robot. (ICORR)*, pp. 41–44.
- [65] J. Nikitczuk, B. Weinberg, and C. Mavroidis. Rehabilitative knee orthosis driven by electro-rheological fluid based actuators [C]. in *Proc. 2005 IEEE Int. Conf. Robot. Autom.*, pp. 2283–2289.
- [66] C. Fleischer and G. Hummel. Embedded control system for a powered leg exoskeleton [J]. in *Embedded Systems—Modeling, Technology, and Applications*. New York: Springer-Verlag, 2006, pp. 177–185.
- [67] N. Kawashima, Y. Sone, K. Nakazawa, M. Akai, and H. Yano. Energy expenditure during walking with weight-bearing control (WBC) orthosis in thoracic level of paraplegic patients [J]. *Spinal Cord*, vol. 41, pp. 506–510, 2003.
- [68] G. Belforte, L. Gastaldi, and M. Sorli. Pneumatic active gait orthosis [J]. *Mechatronics*, vol. 11, pp. 301–323, 2001.
- [69] T. M. Griffin, T. J. Roberts, and R. Kram. Metabolic cost of generating muscular force in human walking: Insights from load carrying and speed experiments [J]. *J. Appl. Physiol.*, vol. 95, pp. 172–183, 2003.
- [70] H. Crowell, III, A. Boynton, and M. Mungiole. Exoskeleton power and torque requirements based on human biomechanics [R]. *Army Res. Lab., Aberdeen Proving Ground, MD, Tech. Rep. ARL-TR-*

2764, 2002.

- [71] A. Boynton and H. Crowell, III. A human factors evaluation of exoskeleton boot interface sole thickness[R]. Army Res. Lab., Aberdeen Proving Ground, MD, Tech. Rep. ARL-TR-2764, 2002.
- [72] E. Harman, H. K. Hoon, P. Frykman, and C. Pandorf. The effects of backpack weight on the biomechanics of load carriage[R]. U.S. Army Res.Inst. Environ. Med., Natick MA, Tech. Rep. USARIEM TR-00/17, 2000.
- [73] R. Atwells, S. Birrell, R. Hooper, and N. Mansfield. Influence of carry- ing heavy loads on soldiers' posture, movements, and gait[J]. Ergonomics, vol. 49, no. 14, pp. 1527–1537, 2006.
- [74] K. Endo, D. Paluska, and H. Herr. A quasi-passive model of human leg function in level-ground walking[C]. in Proc. IEEE/RSJ Int. Conf. Intell.Robots Syst. (IROS), Beijing, China, 2006, pp. 4935–4939.
- [75] M Wehner, B Quinlivan, PM Aubin, E Martinez-Villalpando. A lightweight soft exosuit for gait assistance[C]. 2013 IEEE International Conference on Robotics and Automation (ICRA), Karlsruhe, Germany, 2013:3362-3369.
- [76] 39AT Asbeck, SMMD Rossi, KG Holt, CJ Walsh. A biologically inspired soft exosuit for walking assistance[J]. International Journal of Robotics Research, 2015, 34(6):744-762.
- [77] 40D Ye, I Galiana, A Asbeck, B Quinlivan. Multi-joint actuation platform for lower extremity soft exosuits[C]. 2014 IEEE International Conference on Robotics & Automation (ICRA), Hong Kong, China, 2014:1327-1334.
- [78] MIT Technology Review. <https://www.technologyreview.com/s/539251/the-exoskeletons-are-coming/>. 2015.

## 作者介绍

**衣淳植**，哈尔滨工业大学机械制造及其自动化系本科生。

**苏浩**，纽约城市大学城市学院机械工程系助理教授。曾在飞利浦北美研发中心任职科学家。苏教授是 IEEE 国际机器人和自动化大会的副主编，IEEE 生物医学工程大会的副主编，IEEE Robotics and Automation Society 的 Design and Mechanism 委员会的 Junior Chair，Frontier in Robotics and AI 的编委。主要研究领域为可穿戴机器人，外骨骼，机器人假肢，手术机器人，软体机器人。



**中國旅美科技協會**  
Chinese Association for Science and Technology, USA

**《海外學人》**

**慶祝旅美科协成立二十五周年特辑**

# 风华正茂，继往开来

潘星华，蒋为民，万乐，盛晓明

二十五年的成长，从蹒跚学步、跌跌撞撞，到如今风华正茂、豪情满怀。从纽约几十个人到全美近二十个分会近万人，分布于美东美南美西美北；并从美国走向中国，建立了十几个联络处或办事处，我们不改初衷，我们始终致力于促进中美之间文化、科技、教育、经贸等领域的合作与发展；弘扬中国传统文化，促进中美两国人民的相互了解；促进旅美学人、华人专业人士之间的团结、合作与交流，以便更好地服务华人社区。

虽然我们或早或晚，我们或从大陆、从台湾、从香港，或者经过欧洲、新加坡、日本、澳大利亚、新西兰、加拿大，但是我们最后都来到这片神奇的美利坚，找到了共同的精神家园，中国旅美科技协会！近万会员里，99%拥有硕士以上学位，86%为博士。虽然还只是旅美华人的一小部分，我们中间产生了邓中翰、李彦宏、张亚勤、张朝阳、詹启敏、马启元、周孟初、王飞跃、李百炼、曾大军、方彤、徐昌东、刘志翔、戴一凡、陈友斌、周怀北、陈忠苏、屈志军、毕奇、贾鹏程、张良杰、张以哲、杨崧等一系列高层次人才。我们的志向是凝集全体旅美华人学者的精神和力量，并代表他们去服务社会。我们努力、坚守、奉献，崇尚科技，坚信一切科技分支如生物、信息、电子、能源、材料、工程、金融、数学、物理、建筑、管理、教育等领

域的创新、合作和应用，都能加速经济发展，改善人类生活，同时带来更好的环境保护和人类和谐，促进世界和平。

一代一代的协会前辈特别是前会长们的无私奉献、坚守和宽容，才换来协会今天的兴旺和团结。我们念念不忘早期留美的中国一代大师精英们及国际学术界顶尖人物的关爱，如旅美科协成立之初就担任名誉顾问的陈省身教授、杨振宁教授、朱光亚教授、田长霖教授，及近期受聘的名誉顾问诺贝尔奖获得者达尼埃尔·谢赫特曼教授、马里奥·卡佩奇教授等学术知名人士；我们也为新时期旅美华人科学家的卓越成果、为他们对协会的认同而自豪：庄小威（哈佛）、李凯（普林斯顿）、陆奇（百度）、陈列平（耶鲁）、张翔（伯克利）、王中林（佐治亚理工）等旅美科协卓越科技大奖获奖者；我们也为美国的科技领袖和政府官员人士的支持而衷心感谢：美国工程院院长小丹尼尔·莫特教授，2006年诺贝尔物理学奖获得者约翰·马瑟博士，2013年诺贝尔医学与生理学奖得主托马斯·祖德霍夫教授，美国工程院院士、犹他大学杰拉尔德·斯特林费洛教授，犹他州前州长、驻中国大使洪博培，现州长盖瑞·赫伯特等。

这是一个风云际会、英雄辈出的时代，我们希望旅美科协作为一个平台能更好的实现我们的宗旨，为协会会员和全体华人的成功、为弘扬中华文化、为中美两国人

民的友谊与幸福添砖加瓦。为此，我们提出组织的“正规化、专业化、常态化、可持续化”这新时期的“四化”建设。这里一个重要的标志就是总会章程的修改和完善，其中包含重建理事会的计划，明文规定了理事会的立法监督功能、董事会的组织协调功能以及执委会的计划执行功能，并详细规范会长选举这一有关协会发展导向的重大决策过程。

从五年前开始，我们就开始建立并不断完善成立新分会的指导纲要，明确指出新分会成立之初应致力于建立平稳发展的领导结构，逐步完善（与总会章程不抵触的）分会章程，于所在地区正式注册为非赢利性机构，组织、参与所在地区的各种文化交流、科技、商贸交流活动，有五个新分会加入旅美科协的大家庭，还有好几个新分会正在筹备之中。我们也提出在中国设立办事处和联络处的指导纲要，要求联络处的设立应在慎重的前提下，支持认同旅美科协章程、愿意为旅美科协服务、有一定实力的单位和个人提出申请，现已设立北上广三个办事处以及遍布全国的十几个联络处。

我们逐步建立了协会的学术协调委员会及不同科技领域内的专业协调委员会，目的是为了能够更好地整合和利用旅美科协的学术资源，加强和提升旅美科协的学术活动水平。旅美科协还建立了技术转化协调委员会，以期更好地整合和利用旅美科协的科技资源优势，加强和提升旅美科协对企业在科技创新和科技成果转化应用中的辅助作用。这两个委员会的成立进一步提高

了旅美科协在美国、中国和国际科技界的地位和影响。

近几年来我们加强了学术顾问委员会的建设，得到许多著名科学家的大力支持。我们设立了旅美科协卓越科技大奖，并正在把它打造成科技界值得骄傲的荣誉和协会闪亮的品牌。我们每年一度的年会已经成为中美科技界的一大盛会，美中科技峰会、美中发展与创新峰会、旅美科协杯创新创业大赛、美中高层次人才交流大会、美中合作峰会等精彩纷呈的议程吸引了越来越多的中美两国政府、企业及民间的关注和支持。我们还陆续编辑出版了一系列《海外学人》专刊，如今年的信息技术专刊、生物医药专刊等。旅美科协总会和分会还常年开展各种科技交流活动，以及职业发展论坛、旅美杰出青少年科技奖、教育技术研讨会、优秀高中生奖学金等。

世界潮流，浩浩荡荡。我们张开双臂，欢迎各界合作。我们一定不辜负时代和重任，为中美和全球的科技、经济的合作及发展贡献力量。

## 作者介绍



### 潘星华, Xinghua (Victor)

**Pan**, 博士, 美国耶鲁大学医学院遗传系顾问研究员, 广东省单细胞技术与应用重点实验室主任及南方医科大学生化系主任、教授。他在第一军医大学获得医学学士和硕士学位(1988), 复旦大学

遗传学和遗传工程系获得理学博士学位

(1993), 北京协和医学院肿瘤研究所(1994)及耶鲁大学遗传学系(1999)获得博士后培训, 先后师从于谈家桢、吴旻和 Sherman Weissman 等 3 位中美科学院院士, 并曾在北京大学、墨尔本大学等短期培训。2004 年加盟耶鲁大学医学院遗传系先后任副研究员、研究员, 主持单细胞基因组学研究组, 在美国 NIH 的包括基因组百科全书计划(ENCODE)、癌症和干细胞等重点研究项目子课题和 R01、R21 等课题支持下, 主要从事单细胞测序技术研究, 在单细胞基因组、转录组、表观组和端粒长度分析等方面独立建立了一系列原创性技术; 发表 PNAS、Cell、Nature、Cancer Res 等学术论文和专著 90 多篇, 申请专利 10 余项其中 6 项获得授权, 多次参加组织或主持国际专业大会、发表大会主题报告或专题报告。在耶鲁大学工作之前他在美国分子平台(MSI)生物技术公司任研究员(2000—2004); 出国之前曾服务于第二军医大学任副教授(1994—1997)。

近年, 潘博士担任 Nature Protocols 等 20 多种杂志外审专家或编辑, 美国 *Single Cell Biology* 创刊主编, 及中国首家精准医学杂志 *Precision Clinical Medicine* (由牛津大学出版社 OUP 出版) 创刊副主编, 也是瑞士出版社杂志 *Frontiers* 的研究专题 *Single Cell Genomics: Technology and Application* 的特邀主编, 并应邀撰写国际癌症百科全书章节。他还担任英国医学研究理事会(MRC)、比利时癌症研究基金会和中国科技部 973 基金等的特邀国际评审专家, 为印度 *Advances in NGS* 国际顾问委员会委员、杭州市政府 115 计划特邀外国专家(2014)、江西省赣鄱英才 555 柔性特聘专家兼江西省系统生物医学重点实验室学术委员会主任(2011—2015, 现为首席顾问), 是华西医学院精准医学中心、广东药科大学等的客座教授。

他在美国合伙创建了 ImStem 生物技术公司研发干细胞治疗产品, 获得 USA Today 等主流媒体报道。他也是中国旅美科技协会(CAST-USA)康州分会的首任理事长和两任会长, 多

次担任 CAST-USA 全美总会学术副会长(2013—2016), 并当选为 2017 年候任会长即 2018 年会长。



**蒋为民**, 中国旅美科技协会总会会长、犹他州州长亚裔事务前顾问、重庆大学及青海大学客座教授、四川省海外顾问、US New Century Group 董事长。曾在重庆大学、犹他大学学习工作十多年, 在采矿工程资源及环境

工程领域出版专著两部、发表学术论文 10 余篇, 并获得多个国家、省市级奖项。主要从事矿业/环境工程、教育/文化交流、投资开发/管理。在过去二十几年里, 积极致力于中美文化、教育交流和经贸往来, 致力于促进中美两国间科技、文化、教育、贸易等领域的国际交流与合作, 并促进中美之间各种合作项目的广泛宣传和成功举办。曾受邀到美国白宫参加欢迎胡锦涛主席、习近平主席的国事活动。曾陪同犹他州州长加里·赫伯特中国商贸代表团成功地访问中国, 曾参与组织接待第一届中美省州长论坛中方代表团及后续的各种中美国际合作。

积极协助中国多个省市地区开展中美双边关系各种形式的交流合作, 积极推动犹他与多个省市开展中美双边友好城市、民间交流交往。积极协助促进中美省州、部委州、企业与企业、大学与大学国际交流合作。近年一直积极协助并成功地组织中国中西部地区、欠发达地区与旅美科协的各种形式的交流合作, 积极参与并推动各种形式的国际创新创业活动、高层次人才间的学术交流和国际合作。



**万乐**, 博士, 1982 年毕业于武汉地质学院(原北京地质学院, 1987 年更名为中国地质大学)获工学学士学位。并留校任教, 历任助教、讲师、副教授等。担任了大量教学和科研工作。获得多项部级优秀科

研成果奖。1992年获工学硕士学位。

1993年至1995年，任地球物理系负责教学管理的系主任（地球物理系2003年5月更名为地球物理与空间信息学院）。任职期间，在中国地质大学每年评选出的三个“教学管理先进系”中，地球物理系在1993、1994、1995年分获第三名、第二名、第一名。曾创建三个不同专业：“计算机专业”、“电子无线电专业”和“测控技术与仪器专业”，并任首任测控技术与仪器教研室主任。学校在“电子无线电专业”基础上成立了“机械与电子信息学院”，在“测控技术与仪器专业”基础上成立了“自动化学院”。

1995年评为地质矿产部优秀教师。1997年评为中国地质大学跨世纪中青年学科带头人；1998年评为教育部全国高校优秀青年教师，获教育部霍英东教育基金会奖。湖北省地球物理学会社会推广委员会副主任委员；湖北省地质学会物化探、数学地质、电算专业委员会副主任委员。中国地球物理学会终身会员。

2000年获博士学位。在美国犹他大学地质地球物理系做博士后。现为犹他大学地质地球物理系研究员。Technoimaging公司地球物理师。在犹他大学指导和协助指导硕士和博士十多人，指导和协助指导国内知名高校和研究所的访问学者二十多人。SEG（国际勘探地球物理学家协会）会员。曾任SEG年会分会场主席。是多个国际专业刊物的审稿员。主持和参加过十多项研究项目，获得近十项省部级科研成果奖，发表过七十多篇学术论文。在电磁场模拟中计算机应用技术、海洋大地电磁探测方法反演成像、位场张量偏移成像等研究领域成果卓著，所做的研究工作和成果处于世界领先水平，在行业内具有较高的知名度。国内大学特聘教授。

2015年和2016年任犹他州华人春节晚会组委会主席。组织了本地十八个华人社团（包括来自大陆、台湾、东南亚的华裔侨社组织）共同举办庆祝中国新年的大型联欢晚会，取得空前成功。现任犹他州华人春节晚会组委会顾问和犹他州东方中文学校校长。

中国旅美科技协会终身会员，现任旅美科协总会副会长兼秘书长。在旅美科协犹他分会曾任秘书长、副会长、会长。现任旅美科协犹他分会理事。

在旅美科协总会曾担任负责组织发展的负责人，为硅谷分会、波士顿分会、南佛州分会的成立，做出了贡献。在总会担任秘书长期间，组织协调总会和各分会的关系，组织和协调总会年会的各项工作，做出了贡献。2015年和2016年连续获得旅美科协总会颁发的卓越领导奖。



**盛晓明博士**，1987年进入中国科技大学少年班学习，获工学学士、硕士学位。2002年在加拿大阿尔伯塔大学获得统计学博士学位。先后在美国犹他大学家庭与预防医学系及儿童医学系任助理教授、副教授、教授，并兼任

犹他大学临床和转化科学中心实验设计及生物统计中心共同主任，2013—2016年任犹他大学统计委员会主席、统计学硕士课程主任。主要研究方向为统计方法及其在生物学、医学、社会学、经济学等领域的应用、实验设计与数据分析、临床试验、临床与转化科学、大数据、公共健康与公共卫生及管理。先后参与加拿大四省肺癌研究、新生儿严重性病毒感染成因探讨、美国内华达州核试验放射性灰尘回落健康评估项目等NIH、CDC重大科研课题，在国际杂志上发表几十篇学术论文。在本专业内曾任美国统计协会继续教育委员会主席、数理统计学会新研究员委员会主席，为多家医学、统计学科期刊和杂志审稿。迄今指导了十几名硕士毕业生和三名博士毕业生。先后担任中国旅美科技协会犹他分会以及总会会长；现为犹他州华人联合会主席，《海外学人》杂志编辑。2013年作为海外特邀代表列席中国人民政治协商会议第十二届一次全体会议。

# 回顾科协成立初，团聚纽约庆发展

周华康

1992年夏天，3位来自北京的留美学人（周华康、龚小兵、高克家）聚在一起商讨。我们原本计划获得博士或硕士学位后，尽快归国、报效国家，不期由于种种原因留在美国工作多年。在此期间国内继续改革开放，经济连续多年高速发展，世界各国为之惊叹，旅美学人也都希望能够为国家的经济建设和发展做一些力所能及的事情。我们留美学习和工作多年，亲身感受到美国作为世界最发达的工业化国家，在科技水平和金融资本领域领先中国几十年。我们旅美学人可以利用自身的优势，将美国的先进技术、信息和资源交流到国内。我们认为，应该组建一个团体作为平台，促进旅美学人之间的专业信息交流和事业发展、促进中美两国人民的相互了解，推动及协助中国与美国之间的科技及工商合作，促进旅美学人之间的团结、友谊和弘扬中华文化。

中国驻纽约总领事馆科技组得知此信息后，表示了极大的支持。通过朋友之间的联系推荐，旅美科协筹备组从3个人开始，很快扩大到十几人，几十人到上百人参与筹办。1992年8月25日，在纽约希尔顿大饭店成立了中国旅美科技协会，数百名美东地区的旅美学人参加了成立大会。中国国际人才交流协会马俊如副主席，中国驻纽约总领事馆张伟超大使分别发来了贺电。大会选举了中国旅美科技协会的第一届理事会。

1992年旅美科协成立时，得到了海内外专家学者们的大力支持，杨振宁、李远哲、田长霖、谈家桢和唐敖庆等诺贝尔奖获得者及国际著名学者应邀担任科协名誉顾问。

1995年，旅美科协部分会员组团《旅美学人九五规划考察团》，作为改革开放后留学的第一个博士代表团回国参观考察，受到了国务委员国家科委主任宋健、中国科学院院长周光召、中国科协主席朱光亚和时任统战部副部长刘延东等领导的亲切接见，勉励广大海外学人“为国服务”。

回顾旅美科协刚刚成立的情景，我心情激动，感慨万千！二十五年过去了，旅美科协为中美科技和文化的民间交流做出了有目共睹的贡献。在习近平主席伟大复兴中国梦的指引下，中国取得了翻天覆地变化。作为旅美华人，我为中华民族的伟大复兴而刚到骄傲和自豪。祝愿旅美科协的年青一代，在各自的事业取得更加辉煌的成就，为中美科技文化的交流做出更大贡献！

## 作者介绍



周华康博士，毕业于纽约理工大学。在AT&T和SCHERING等世界500强企业从事软件开发10余年。1993年创办美国华纳国际投资公司，从事国际贸易，国

际培训，风险投资和资本运作。现任四川阿波罗太阳能科技有限责任公司董事长。2005年起投资多家民营高科技企业。2010年3月，资助成立 NJIT-阿波罗太阳能研究中心（后更名为 NJIT-中建材太阳能研究中心）——“这是中国企业第一次正式和美国公立大学合作建立的创新技术研究中心，对中国企业的技术创新发展方向和模式影响重大”。周华康博士是中国旅美科协创建成立的第一任会长。

# 纽带、渠道、桥梁

——记旅美科协成立廿五周年

陆重庆

最近，旅美科协大纽约分会的肖继忠理事长来电话要我写篇文章，讲讲旅美科协的历史。他说，今年十月在纽约举办年会，纽约是旅美科协的诞生地，这里的许多年轻人、包括加入科协已有十几年的他自己对旅美科协的早期都不太了解，很想知道当年为什么成立了这个社团、那时是如何成立的、怎么走到了今天以及其它种种有影响的往事。

都说流年似水，没想到不经意间，旅美科协成立居然已有二十五年了。最初在一起商讨创建筹备的那些同仁，如今有的退休在家含饴弄孙，有的仍在社区服务民众，有的继续在专注学问潜心研究，有的还在奔波四方驰骋业场，也有个别的已撒手人间……健在者多星散各地，鲜有联络。常说往事如烟云，平时没将这话挂在心上，暮然回首，这讲法还有点道理。创建之初那些夜以继日忙碌不休的岁月，电话里会议上共处携手奋战的熟人，曾经为之困惑纠缠的棘手难题，随着时间的流逝，虽然不时在眼前飘现，但亦慢慢依稀淡出。本想借助保存的那些通讯、刊物、纪要、照片，从中发掘史料准确辅证以完成这一写作，不巧前不久整理旧物，书柜纸箱方位大变一时难觅去处。为免出错而误导后人，这里暂且围绕上述节点讲些故事聊个天。至于完整的历史回顾，尤其是事件人名的罗列记载，姑且留待日后。

时针先拨回到九十年代初。那时，中国改革开放后最早一批于八十年代赴美攻读博士学位的留学生先后完成学业步入院校、企业、政府等部门工作。有了稳定的职业生活，有了更多的自由支配时间，精力充沛正当年华，渴望充分借助兼具中西文化的独特优势，发挥所学专业特长，为自身发展、为中美交流做些事。这需要平台、需要团队。然而，当时华裔社团多由早期经商为主的老侨组成，鲜见有以来自中国大陆专业人士为主体的。于是“中国旅美科技协会”、简称“旅美科协”应运而生。这是一九九二年在美国纽约注册成立的一个非政治、非赢利的民间社团，英文名为“The Chinese Association for Science and Technology, USA”，简称“CAST-USA”。说来有趣，成立不久旅美科协去美国机械工程师学会（ASME）总部访问交流，见面伊始对方就问，你们来纽约开设分部了吗？原来他们将“中国旅美科协（CAST-USA）”误认为是“中国科协（CAST）”了。还有，旅美科协犹他分会成立时，被聘为分会顾问的一位老前辈谈起，四十年代就有个旅美科协，他是会员。不过，那个科协在五十年代初随着中美关系的变化消逝。成员大部回国，其中不少是后来众所周知的著名学者科学家。几十年后，留在美国的他还接到一些国内来信，请他作证他们是当年的会员。

旅美科协的出现在北美带来了很大影响，受到中美有关方面的高度关注。这不仅是由于成立得早，还因为其规模、运作及专业人士社团的特色。这与当时的外界大环境，以及那些创建筹办人不无关系。他/她们不少在国内高校就读时就有组织学生会活动的经历，来美就学时又在当地的学生学者联谊会承担过不同的工作，丰富的社团活动经验和广泛的社会联系为协会的创建和运作奠定了重要的基础。一九九二年年年初伊始，他/她们利用周末和下班之余，频繁地交换想法探讨途径，对比参考借鉴周围社团的经验，根据这一群体的特点与需求，从宏观定向到操作细节如协会的名称、宗旨、会徽、组织、结构、运作、外联、地区等方方面面反复商讨，拟写一页页文稿方案。几个月后的初夏，召开了有逾半百人参加的预备大会。八月，数百人云集纽约中城，举办了成立大会，通过章程、选出理事会、聘请了有中美双方包括诺贝尔奖得主在内的科技文化教育各界著名人士的顾问。作为联谊的纽带、沟通的渠道、合作的桥梁，旅美科协顺应时势诞生了，这在中美两地引起了极大反响，有关方面纷纷派员或电文祝贺，一时间许多媒体大幅报道，海外学人趋之若鹜。成立后为迎来第一个春节组织的聚会，有一千二百多人蜂拥而至。当时不少与会者都惊讶于作为华人生活在纽约多年，却一直不知道自己身边竟然有这么多同族专业人士。

旅美科协很快在全美各地组建了十几个地区分会以及工程、电子商务、化学化工、金融商务、网络信息等专业学会。组织了各种有关职场、创业、求职、求学、

理财、治家、育子的系列专业讲座及郊游野餐等社交联谊活动，创办中文学校，还与其它社团联合举办各种专业研讨会，以及如抗战胜利五十周年音乐会等大型社会活动。在诸多活动中，颇具影响的是从未间断过的年会。年会上有来自中美两地科技文化教育经贸等各界著名人士的主题演讲以及各专业领域代表人士的专题研讨，年会又是每年一度大家团聚的场所。在纽约举办第二届年会时，华盛顿、匹兹堡、康州分会开了好几辆车前来参加。晚上联欢，远道而来的会员们高兴得跑到桌子上跳起了舞蹈。旅美科协搭平台助大家绘出各自人生画卷，平台上走出了各领域各行业各色人才。年年岁岁，新知旧雨欢聚一堂，叙旧论新交流切磋。第十届年会上，有位会员手持一张塑封卡片过来说：还记得我吗？我已经去香港工作了，这次专门回来参加年会。你看，这是我的会员卡，会员序号在前一百以内，我是老会员。“老会员”，三个字，自豪之情，溢于言表。

旅美科协刊发多种出版物。有报道协会日常活动的《通讯》和介绍协会昨天今日和明朝的网站，有方便会员相互联络的《通讯录》。值得一提的是《海风》季刊，每年出版四期，十年不断，登载会员在一线原创的长文短篇，展示各领域最新成果发展动向，人文科技兼收并蓄，深入浅出高屋建瓴，外行看得懂、内行得门道。出版的《异邦之龙》一书，记录了两代旅美学人求知、探索、搏击、创业的历程，在中科院首发，被纽约图书馆收藏。众多著述，围绕中国人才引进交流政策、非典防治引出的国家紧急应对机制、金融危机对

策、网络技术电子商务、全球化前景、能源政策、人文社会等专题，以及组织百余专家编撰数册“建言集”中针对中国基础科学、生物技术、农业科技及科技发展所提的看法、建议、措施，现今许多都已得到验证、被有关部门采纳。预见的前瞻、分析的深刻、建言的实在，首开先河、令人侧目。

一九九六年，旅美科协首次组织了旅美优秀学者回国访问团，在北京受到国家领导人的专门会见（图 1），在人民大会堂发言讲话被媒体采访、当晚在全国新闻联播节目报道。这是一支由获得美国总统学者奖的华裔教授所组成的自费交流讲学团，其中不少成员如今已是各自领域的领军人物。当时经广州、杭州、上海到北京，为海外人才高层次大规模多方位有组织的交

流合作率先进行了有益的探索，引起了多方重视反响，有关部门专门派人全程同行。有意思的是，那时有些地方还并不感到有如此人才需求，以为水平高了对口对接不易。没过几年，情况大变，九十年代末在纽约当时的留学生活动中心举办一场招聘活动，千余专业人士竞相前往，人满为患。发展到后来每月都有团组来访，应接不暇。各地对高层次人才需求如饥似渴，人才引进活动如火如荼。如今更是年年邀请组团前往，创新创业大赛络绎不绝，补助机票、提供行宿、全方位支持。二十多年前的涓涓细流已成今日钱江大潮。

二十五载弹指一挥间。如今旅美科协一万多会员遍布北美各地，地区分会稳步扩展，大平台多资源优势彰显，创新创业交流合作活动横跨大洋两岸有口皆碑。回



图 1 1996 年 7 月 20 日，陆重庆会长组织的旅美优秀学者访问团受到时任中共中央书记处书记的温家宝接见（照片转载多维新闻网《温家宝的 81 个瞬间》）

首以往，所走路途亦非一帆风顺。在内部，单以前十年为例，就经历了三次不小的波动。在外部，新老社团生生息息资源力量频频重组变化。网络技术信息世界改变着思维行为方式，面临的挑战层出不穷从未稍止。凡此种种都是旅美科协发展的助推动力。不断交织紧密的美中关系、飞速发展的中国经济、日益增多的赴美中国学人，是旅美科协成长的土壤和基础。传承发展、创新引领，围绕会员需求，牢记服务奉献，吸纳更多新人，搭建更大平台，做好纽带、渠道、桥梁，加强制度组织结构的不断完善，充分发挥如预备会长、现任会长、理事长(上届会长)、董事长(上上届会长)“四驾马车”，执委会、理事会、董事会“三足鼎立”等运作模式的特点。历经锻炼步入成年的旅美科协，当能顺应历史潮流谱写绽放更大辉煌。

登高望自远、海阔纳百川。旅美科协成立十周年时在大纽约区举办了两天年会，年会会刊中汇编了前十届会长就任时的就职演说。旅美科协二十年华诞之时，年会专门摄制视频短片中定格了前二十位会长的任职感言。他们的那些字句、那些声音，随着时光终将湮佚。但他们以及许许多多台前幕后各方友人对旅美科协的那份情感和期望，他们为此所付出的心血、智慧和奉献将永伴旅美科协成为史册上的一笔重彩。

#### 作者简介



**陆重庆**，博士，能源专家。旅美科协终身会员，1992年入会，曾任科协理事、会长，《海风》主编等。

# 新的时代 新的方向

## ——庆祝旅美科协成立二十五周年

谢家叶

近年来旅美科协在协会领导层和广大会员的努力下，呈现出蓬勃发展，欣欣向荣的局面。作为一个从科协初期就参加协会工作的老会员倍感欢欣鼓舞。在科协成立二十五周年之际，我在此简单地回顾科协的发展，和广大会员一起庆祝旅美科协成立二十五周年。科协二十五年的发展是当代中国发展的一个缩影。科协二十五年的发展既反映了负笈海外的莘莘学子的心路历程，也反映了中国当代科技和经济以及社会的飞速发展。

在二十五年中，科协的发展方向基本上是经历了两个主要的变化。科协最初的创立宗旨是为了团结当时在美国发展事业的来自大陆各地的学生（主要是博士生）和学者，建立一个可以互相交流和互相帮助的平台。同在异国他乡奋斗，中国改革开放初期来美国学习和创业的莘莘学子与现在相比人数不多，在经济上的力量有限，但大家工作和学习都很努力，克服了很多困难。在美国初步安定之后，都希望能有机会能够互相交流和支持，以利于在美事业和生活的发展。在这样的形势下，旅美科协就应运而生了。创始会长是周华康博士。最初的协会宗旨是联谊、交流、互相帮助，丰富生活。所以协会最初的活动也是以联谊为主。经常组织的活动主要是旅游、野餐、晚会（包括舞会）和各种聚会。

其中以晚会最受欢迎，参加者众。有一次春节晚会（94年春节）有近二千人参加，盛况空前。那时的晚会颇有质量，很多明星都在科协的晚会上表演过。例如，著名主持人杨澜，著名歌唱家刘欢，当时百老汇音乐剧“西贡小姐”的男主角王洛勇，纽约大都会歌剧院男中音歌唱家田浩江等等。纽约当时有一个留学生中心，地点很好，位于九大道和十八街交界处，里面有一个颇为正规的剧场。旅美科协的晚会经常在那里举行，协会的联谊活动吸引了很多专业人士参加旅美科协。大陆专业人士社团崭露头角，在纽约产生了较大的影响，大陆学者群体开始在美国显示力量。

随着大陆赴美留学生的增多，帮助大陆专业人士进入主流社会的呼声渐高。当初协会理事会成员大多是在美国研究机构和大学工作，于是帮助大陆专业人士更好地进入主流社会发展成为了协会的宗旨和工作方向。为此协会举办了各种讲座，例如，新技术发展方向，如何更有效地寻找工作等。新泽西理工学院的周孟初教授提供场地和组织了许多这方面的讲座。笔者也曾应邀做过“减少口音”（accent reduction）的讲座，每次都有一百多人参加。同时，协会的杂志《海风》上也发表了一些有关进入主流社会的文章。这些工作在当时对广大旅美学人有一定的帮助。协会

内部的组织结构也开始扩大了。华盛顿分会和犹他分会是最早成立的分会。旅美科协成立以后与国内科技界的往来逐年加强。在旅美科协成立的初期就接待了很多国内来访团。也举办过一些培训班，这种培训班接待是全程的，有偿的。通过这种接待，虽然加强了和国内一些单位的联系，但同时也产生了一些利益纠纷。后来这种类似的商业活动就和旅美科协剥离了。旅美协和国内更多的交流还是组织和国内来访科技人员和团体会谈。当时的科委主任宋健，科学院院长路甬祥，现在的科学院院长白春礼访美时也都和旅美科协进行过交流。旅美科协也组织了学者回国访问交流。其中影响比较大的有陆重庆会长组织的获得总统奖的“旅美优秀青年学者代表团”。这些交流对旅美科协的自身发展和促进国内科技的发展都有一定的意义。除此以外，科协还曾多次组织会员撰写文章，对中国的科技发展提出建议。这些建议书都受到了国内的广泛重视。其他令人印象深刻的活动有和中央乐团合作在纽约林肯中心举行的纪念抗日战争胜利五十周年音乐会，演出“黄河大合唱”，在美国社会影响较大。

在协会早期，经费匮乏，绝大部分工作都由会员利用业余时间完成。为了各项活动的组织和开展，会员们不但贡献了时间和精力，还在财务上帮助协会。尤其是会长和理事们每次活动都是出钱出力，默默付出。早期《海风》杂志的出版更是凝聚了广大会员的心血。从组织稿件到印刷出版，都是编辑人员和作者不求回报的无偿奉献。协会只是负担了印刷费，其余的费用都是编辑部工作人员共同负担的。《海

风》一年四期准时出版。有不少质量较高的文章，受到会员的欢迎，也得到国内有关单位的高度重视。

九十年代末，随着高科技的迅速发展，国内也开始认识到高科技的发展在经济发展中的重要作用，对世界高科技的应用和发展开始重视。与此一趋势相适应，旅美科协开始组织每两周一次的科技沙龙，讨论一些新技术的发展以及商业转化前景。同时，对国内的科技发展和可持续发展提出了许多良好的建议。这时候国内开始相继出台了一些吸引国外学者回国创业和回国服务的政策，旅美科协的很多会员也开始萌动了回国创业和回国服务的心思。自此旅美科协的发展方向开始转向协助会员回国创业和服务。直到今天，协会的这个发展大方向基本未变。当时，来美国举办创新创业和招才引智活动的基本上是北京、上海和广州这三个大城市。除了北上广之外，还有江苏、浙江也是早期积极争取国外人才回国服务的省份。旅美科协与这些城市和省份都建立了合作关系。进入新世纪之后，大批旅美科协会员先后加入了回国创业和服务的行列。有不少会员在一些领域做出了亮丽的成绩，取得了很大的成功。

时至今日，回国创业已经洪流滚滚，蔚为壮观。旅美科协会员积极地投入到这一历史发展的机遇之中，千帆竞先，万木争春。在当前的形势下，旅美科协也呈现出崭新的发展局面。科协在国内各城市的联络站如雨后春笋，纷纷成立。旅美科协作为一个具有二十五年历史的重要专业人士组织在协助个人创业方面已经很有成绩。

今后，可以进一步发展一些具有旅美科协特色的工作方向为会员服务。例如，旅美科协编写建议书的传统可以继续开展。建议可以不仅仅局限于科技范畴，可以从科技角度辐射到其他方面，如金融、管理、宏观和微观经济等领域。在创业方面，除了帮助个人创业之外，利用协会优势整合资源，帮助组织团队创业也可以同时开展。从目前商业和技术发展的趋势来看，团队创业应该比个人创业的成功机率高。旅美科协本身是一个非盈利组织，不能经营赢利性的商业活动。但可以帮助组织和整合有合作意愿的会员进行创业或投资合作。这在旅美科协历史上是有先例的，但是并不成功。然而，当时的不成功不等于现在不能成功。在新的形势下，只要组织得法，

合乎法规，发展方向正确，有合作意愿和凝聚力，成功的机会还是很大的。

山重水复，任重道远。团结合作，前程似锦。在全体会员的努力下，在广大朋友的支持下，旅美科协的下一个二十五年将会更加辉煌。

#### 作者简介



**谢家叶**，博士，儿童语言病理专家（pediatric speech pathologist）纽约自由执业。1992年成为旅美科协会员，终身会员。历任《海风》编辑，总编，旅美科协理事会理事，副会长，会长。

# 中国旅美科技协会介绍

中国旅美科技协会（Chinese Association for Science and Technology, USA，简称旅美科协，CAST-USA）是1992年在纽约成立的非政治性、非营利性的民间团体。旅美科协的三大宗旨是：促进中美之间科学和技术、教育和文化、商业与贸易等领域的交流与合作；弘扬中国传统文化、促进中美两国人民的相互了解；加强旅美学人、华人专业人士之间的团结、合作与交流。

旅美科协是一个跨地区（在美国）、跨行业的综合性科技团体，会员主要来自科技、文化、教育、法律、金融、人文等领域的中国旅美专业人士组成，现有会员近万人。许多会员在世界五百强企业或知名大公司、高等院校或研究机构从事科技开发和研究工作，许多会员已经成为中、高层管理人员。旅美科协目前在全美有十六个分会及专业学会，会员分布在美国几十个州，并在中国国内十几个省市设立了联络处。

总会设执行委员会负责日常工作，还设有董事会、理事会、学术协调委员会和顾问委员会。旅美科协成立以来的知名名誉顾问包括陈省身教授、宋健教授、杨振宁教授、朱光亚教授、陈香梅女士、田长霖教授、周光召教授、朱丽兰教授、路甬祥教授、邓文中先生，及诺贝尔奖获得者达尼埃尔·谢赫特曼教授、马里奥·卡佩奇教授等学术及社会知名人士。

旅美科协总会及各分会定期举行学术研讨活动，为会员提供学术交流的平台，并定期组织其他系列活动，为会员们提供学习、工作及生活方面的交流平台。旅美科协总会定期出版《海外学人》杂志及实时通讯，内容包括协会学术活动的介绍与中美科技界、工商界的最新动态，以及各种工作与投资机会等许多会员们切身关心的内容。

每年总会及各分会举办包括全国年会及分会年会、学术讲座等在内的几十次大中型学术研讨活动，活动中旅美科协邀请中美各界知名人士对所关心的学术及社会问题进行探讨。旅美科协注重与其他专业协会的交流与合作，加强不同学科华人的交流，同时促进中美之间科技人才的沟通和科技的发展。旅美科协各分会也注重参加当地的华人社区活动，与所在地的其它侨团建立良好的关系。

2017年总会年会将于10月7-9日在纽约曼哈顿召开。

旅美科协总会现任会长蒋为民，董事会主席陈志雄，理事会主席宋云明，候任会长潘星华。旅美科协之前的历任会长为周华康、章球、徐震春、陆重庆、马启元、周孟初、谢家叶、肖水根、石宏、邹有所、林民跃、王飞跃、李百炼、左力、沈陆、陆强、曾大军、方形、盛晓明、蔡逸强、于浩、宋云明和陈志雄。

总会网站：[www.cast-usa.org](http://www.cast-usa.org)

## 旅美科协分会分布



1. 大纽约分会 - Greater New York Chapter (New York, New Jersey)
2. 康州分会 - Connecticut Chapter
3. 匹兹堡分会 - Pennsylvania Chapter (Pittsburgh)
4. 华盛顿特区分会及网络信息学会 - Washington DC Chapter/Network Society
5. 亚利桑那分会 - Arizona Chapter
6. 加州洛杉矶分会 - California Chapters at Los Angeles
7. 加州圣地亚哥分会 - California Chapters at San Diego
8. 加州硅谷分会 - California Chapters at Silicon Valley
9. 佛罗里达分会 - Florida Chapter
10. 北卡分会 - North Carolina Chapter
11. 德州分会 - Texas Chapter (Dallas)
12. 犹他分会 - Utah Chapter
13. 西弗吉尼亚分会 - West Virginia Chapter
14. 大波士顿分会 - Great Boston Chapter
15. 南佛州分会 - South Florida Chapter

## 16. 内布拉斯加分会 - Nebraska Chapter

筹备分会（及拟落地城市）：佐治亚州（亚特兰大）、伊利诺伊州（芝加哥）、科罗拉多州（丹佛）、德州（休斯顿）、宾夕法尼亚州（费城）、华盛顿州（西雅图）

The Chinese Association for Science and Technology, USA (CAST-USA: [www.cast-usa.org](http://www.cast-usa.org)) is a non-profit and non-political organization. It was founded in August 1992 in New York City. The missions of CAST-USA include: 1) to build relationships and exchange programs between China and United States in the areas of science, technology, education, culture, business, trade and other fields; 2) to promote traditional Chinese culture and provide local Chinese community activities to the general public in the United States; and 3) to provide networking and collaboration opportunities among Chinese scholars and all professionals in the United States.

CAST-USA is a cross-regional, cross-industry, integrated technology group. The majority of CAST-USA members are Chinese-American professionals, who hold advanced degrees in science and technology, education, business, law, medicine, art and other fields of endeavor. CAST-USA is growing at a rapid pace with more than thousands of members in more than 30 states across the United States, working in universities, industries, government agencies and other sectors, many in the world's top 500 multinational corporations, or well-known large companies, universities or

research institutions. Currently CAST-USA consists of 16 regional chapters & 6 to-be-built regional chapters in the United States and 14 liaison offices in China.

CAST-USA has an Executive Board responsible for daily operations and management, a Board of Governors to propose a revision to CAST-USA Bylaws, an Advisory Board, and a Board of Directors to review and approve amendments to CAST-USA Bylaws and vote on the impeachment motion initiated by the Board of Governors. CAST-USA has the following famous Honorary Advisors: Prof. Jian Song, Prof. Chen-Ning Franklin Yang, Mrs. Anna Chan Chennault, Prof. Guangzhao Zhou, Prof. Lilan Zhu, Prof. Yongxiang Lu, Mr. Man-Chung Tang, and Prof. Dan Shechtman, Prof. Mario Capecchi and other academic and social celebrities.

CAST-USA and its chapters hold regular academic seminars and provide academic exchange platform for members. CAST-USA and its chapters also hold other series activities and seminars to provide a platform for members to support each other to study, work and live well. CAST-USA regularly publishes the “Overseas Scholars” magazine and newsletter. The contents include a presentation of academic activities, the latest news of the science, technology, and business sectors, job and investment opportunities, and immediate concerns from many members.

CAST-USA and its chapters hold nationwide annual conferences and regional

conferences, and dozens of academic lectures and academic seminars, which invite many well-known celebrities to address and discuss academic and social issues. CAST-USA is now one of the most-recognized organizations among Chinese professionals in the United States. It is the leading institution dedicated to the promotion of understanding between Americans and ethnic Chinese in the United States. Since its foundation, CAST-USA has established strong ties with the scientific and technological communities both in the United States and in China.

The 2017 annual national conference will be held in New York City from Oct. 6-8.

Current President of CAST-USA is Bill Jiang, Chairman of the Board of Governors Zhixiong Chen, Chairman of the Board of Directors Yunming Song, President-Elect Xinghua Pan. Former presidents of CAST-USA are Huakang Zhou, Qiu Zhang, Zhenchun Xu, Chongqing Lu, Qiyuan Ma, Mengchu Zhou, Jiaye Xie, Shuigen Xiao, Hong Shi, Yousuo Zou, Minyue Lin, Feiyue Wang, Bailian Li, Li Zuo, Lu Shen, Qiang Lu, Daniel Zeng, Tony Fang, Xiaoming Sheng, Yiqiang Cai, Howie Yu, Yunming Song, Zhixiong Chen.

## 1. 大纽约分会（创始分会）

### 城市介绍

纽约是美国最大城市及第一大港。联合国总部位於该市。其政治、经贸、文化影响力在世界上首屈一指，是公认的“世

界之都”。因为该市广大的地下铁路系统，和繁华的建筑物、交通及人们，所以纽约也常被称为不夜城。纽约市有非常多的地标、社区是世界知名的。如自由女神像欢迎数百万移民在 19 世纪到 20 世纪初的到来。



### 当地经贸

大纽约地区是全球商业经贸的重心，是全世界最重要的金融中心，位于华尔街的纽约证券交易所是全球交易量最大的交易所。财富 500 强中就有 45 家公司总部位于该市。大纽约也是全美的保险、房地产、媒体和艺术的重镇。2010 年的统会区有大约 1.28 万亿美元的 GDP，是世界第 13 大经济体。由于纽约市是跨大西洋光纤缆线的主要支点，促进了该市生物科技、软体开发、游戏设计和网路服务等高科技产业的成长，其他重要经济领域还有医疗研究与科技、非营利组织和大学。

### 分会介绍

中国旅美科技协会是于 1992 年 8 月在美国纽约注册成立的一个非政治和非盈利

的华人专业人士社团组织，是旅美科协创始分会。



目前科协大纽约分会拥有来自美国纽约州和新泽西州 2000 多名会员，会员大多数拥有科技，教育，商业，法律，医学，艺术和其他领域的高级学位。科协大纽约分会是大纽约地区华人专业人士最知名和最活跃的华人科技社团组织之一，一直致力于促进中美之间的学术和人才交流，以及会员的职业发展。科协大纽约分会还通过举办多种形式的研讨会、学术会议、IT 高峰论坛，海外人才交流论坛等，努力打造科技，经济，教育和贸易等领域中美合作与交流的平台。此外大纽约分会还组织很多高层次代表团参加在中国的各种学术和人才交流会议。通过和其他专业协会的合作，旅美大纽约分会在促进当地社区发展、加强会员间交流和友谊等方面起到日益重要的作用。旅美科协大纽约分会也将继续致力于帮助会员们在美国或中国的进一步的职业发展。

作为旅美科协创始分会，旅美科协大纽约分会将与其他分会一起，为打造中美专业人士沟通的桥梁做出更大的贡献。

## 联络方式

网址：<http://castusa-gny.org/>

联系人：会长 沈琦博士  
([gnycast@gmail.com](mailto:gnycast@gmail.com))；理事长 肖继忠；  
候任会长 颜为民

## 2. 康州分会

### 康州介绍 (Connecticut)

康涅狄格州是新英格兰六州中最偏南的一州，依山傍海，经济发达，距离纽约



曼哈顿，麻省波士顿仅 2 小时车程，北接马萨诸塞州，东接罗德岛州，西毗纽约州，南隔海峡与长岛相望。康州在全美首先创立基本法，故有“宪法之州”自称，州府哈特福特是美国的保险业都市。康州有个号称是北美最大的金神 (Mohegan Sun) 赌场，为首屈一指的娱乐场所，备有顶级的品牌商店，此外，康州快活豪华大赌场是美国快活大赌场度假村中的最大赌场之一。马克吐温在哈特福特的故居也是观光者必到之处。州内著名大学有耶鲁大学、康涅狄格大学。康涅狄格州是美国最富有的州之一，根据 2005 年的数据，康州是美国人均收入最高的州，拥有众多的亿万富翁和

千万富翁在此州居住，美国亿万豪宅区格林威治就坐落在康涅狄格州，是美国最富有的小镇之一。也是《时代广场的蟋蟀》发源地。康涅狄格州在美国独立时期，是 13 州联盟之一。

### 当地经贸

康州是新英格兰六州中工业最盛的一州。主要是制造核潜艇、喷射飞机引擎、直升机等。康州生物医药产业较为发达，并已形成了企业、科研机构和政府三位互动的集群式发展格局。康州现有生物医药企业约 800 家，一些国际知名的医药企业，如辉瑞、拜耳、勃林格殷格翰、百时美施贵宝等，也在此设有地区总部、生产设施或研发机构。位于康州的耶鲁大学和康涅狄格大学等高水平研究机构，为该州生物医药产业发展提供了宝贵的智力和人才资源。为吸引医药企业在康州投资，帮助企业在康州快速成长，康州经济与社区发展局 (DECD) 出台了一系列优惠政策，包括提供相匹配的拨款、贷款，贷款和税收优惠以及为企业量身定制相应的金融协助方案等。

### 分会介绍

中国旅美康州科技协会 (CAST-CT) 于 2001 年创办，既作为中国旅美科技协会 (CAST-USA, 1992 年成立) 的一个分会，也于 2006 年在康州独立注册为非赢利机构，并于 2008 年初独立取得非赢利机构免税资格。康州科协致力于建立新代华人学者和

专业人士包括康州从北到南的科学家、工程师、企业家、经贸和金融界专业人士等的交流、合作、创业和发展的平台，并成为美中文化、科技、教育及经贸等领域友谊和合作桥梁的康州“桥头堡”。



### 联络方式

网站: [www.castct.org](http://www.castct.org) 联络人: 会长  
余国营 [guoyingyu1@gmail.com](mailto:guoyingyu1@gmail.com)

## 3. 匹兹堡分会

### 分会介绍

中国旅美科技协会-匹兹堡分会 (CAST-P) 创立于一九九四年, 经过二十多年的发展壮大, 至今拥有约千名会员, 一直以来是中国旅美科技协会最大, 最有



影响的分会之一。匹兹堡分会是由高学历旅美华人为主组成的在宾州正式注册的一个非盈利, 非政治性的组织; 它目前也是大匹兹堡地区历史最悠久, 最具影响力的华人社团。

匹兹堡分会的成员主要来自于生活在美国宾州西南地区并在各行各业取得卓越成就的华人科学家, 工程师, 医学专家, 商界人士, 以及艺术家。为了适应新形势的需要, 服务于广大留学生, 分会特设了学生会会员, 旨在吸引更多的留学生参与, 为学生们提供学习, 就业, 创业, 职场等





方面的咨询，指导和交流的平台。从成立的第一天起，分会就以推动中美科技文化领域的广泛交流，促进两国商业的友好往来，帮助旅美华人专业人士改善工作和生存环境，和弘扬中华文化作为自身的工作宗旨。

为了加强会员之间的交流和联络，我们会经常组织一些活动，包括一年一度的大型春节晚会，夏季野餐会等；除此之外，我们还定期举办各类知识性，实用性论坛和讲座，邀请有关方面专家，介绍科技领域的最新进展，以及为会员们提供健康，金融，税务，法律，职场等方面的知识和解惑。同时，我们也积极支持，参与和协助组织本地华人社区的各种活动。这些丰富多彩的活动，加强了会员之间的交流，方便了会员拓展人脉，也丰富了大家的业余生活。

匹兹堡分会一直致力于为会员谋取更多的福利。为了实现一卡（会员卡）在手，华人社区轻松走，分会已同越来越多的本地华人服务公司，商铺和餐馆等建立战略合作伙伴关系，为会员提供商品折扣或优惠服务。

## 联络方式

网站：[www.castp.org](http://www.castp.org) 联络人：会长 程鹏 [pengcheng119@yahoo.com](mailto:pengcheng119@yahoo.com)

## 4. 华盛顿特区分会及网络信息学会 分会介绍

旅美科协华盛顿分会（CAST-DC）是中国旅美科技协会中最活跃的分会之一，



会员主要分布在美国首都华盛顿、马里兰州和北弗吉尼亚。中国旅美科技协会（旅美科协）DC分会（The Chinese Association for Science and Technology - Washington DC Chapter (CAST-DC)），是1993年12月11日成立于马里兰大学的非政治性，非盈利性的民间团体，历任会长为吴少海、薛

澜、李宏诚、肖水根、林民跃、叶玉彬、杨崧、胡增建、孔超翔、汤斌、宋云明、薄智泉。旅美科协信息学会（CAST-NIS）是中国旅美科技协会的一个专业学会，成立于 2000 年。

CAST-DC 和 CAST-NIS 的会员主要分布在美国首都华盛顿、马里州和北弗吉尼亚，遍及各个行业，包括 IT，生物，金融，

网络信息学会牵头主办的中美高级网络研讨会，已在中美成功举办 10 次。从 2016 年 8 月开始，CAST-DC/NIS 成功举办了六人次“大师名人学者学术系列讲座”，其中包括：纳米技术在个人健康领域中的未来巨大发展，商业环境中的大数据分析，光控基因治疗技术及光学网络，光纤和量子通信，等等。这些讲座不仅为会员提供



经济，房地产，管理和法律等专业人士。旅美科协的宗旨是促进旅美学人的信息交流，事业发展；加强会员友谊；发扬中华文化传统；增进中美两国人民之间的了解；推动中美在科技，经济，文化方面的交流与合作，组织筹办各种教育、技术和商业的学术研讨会来促进中美科研人员以及企业家之间的合作与交流。同时也组织中美专业代表团进行互访，以促进双方更密切的合作，通过会议、讲座、互访等为中美科技文化和教育交流起了良好的桥梁作用。从 2013 开始，CAST-DC 与中国科协和江苏等省科协成功举办了“海智大会”。从参加了与中国大陆的人才和科技交流活动。

了纵横向的技术和知识交流，同时讲座也开放给大华府社区，展现了科协立足科技，服务社区，回馈社会的奉献精神。

CAST-DC/NIS 还积极倡导民间的文化和美中友好的草根性活动。注重以分会的形式参加当地的华人社区活动，与所在地的其它侨团建立了良好的关系。作为主要发起社团之一，从 1999 年开始每年组织“华盛顿中国文化节”以弘扬华夏文化、促进族裔团结、增进美中友好。华盛顿市长每次均发布文告，宣布当日为“中国文化节日”。

### 联络方式

网站: <http://www.castdc.org> 华盛顿分会: 会长 石德秀 [dexiu@yahoo.com](mailto:dexiu@yahoo.com) 网络信息学会: 会长 胡嘉南博士 [whu888@yahoo.com](mailto:whu888@yahoo.com)

## 5. 加州洛杉矶分会

### 城市介绍: 洛杉矶 (Los Angeles)

为美国第二大都市地区, 位于美国西岸。风景如画, 海滩迷人。阳光明媚, 气候宜人。从北到南, 有多处自然景观: 圣巴巴拉, 范杜拉海滩, 海峡岛国家公园, 马里布, 圣莫尼卡, 曼哈顿/钻石海滩和长滩,



更有世界闻名的迪斯尼, 影城和 2 小时车程外的拉斯维加斯大赌城。也有现代艺术博物馆, 歌剧院, 科技馆以及尼克松和里根总统图书馆。在喧哗城中区外, 还有清静悠闲且非以旅游业为主体经济的小城如千橡, 卡马里欧, 西米谷等等。千橡去年被评为全美人口十万以上城市中, 第四位最安全的城市。它以保护生态, 讲究生活质量 (包括子女教育) 以及可持续发展作为吸引企业和人才的资本和财富。

### 当地经贸

洛杉矶大都市地区是仅次于纽约和东京外, 世界第三城市经济体。是美国娱乐文化, 影视和航空工业的摇篮。拥有美国 40% 以上的进口货运量, 进住了大部分在美外国汽车厂商的总部 (包括中国比亚迪), 占据了全美一级互联网节点和几千个支持南加州和全美的互联网数据中心 (包括中



国联通和中国移动的美国总部)。它拥有全美著名的生物医学, 制药和工程设备的科研基地 (包括加州理工大学, 加大的洛杉矶, 圣巴巴拉和尔弯分校, 南加州大学, ) 以及相关企业厂家。其中以昂健生物制药 (世界最大生物制药), 贝克曼库尔特 (世界最大血液分离, 检测和分析产品) 和艾尔建 (眼, 神经, 皮肤, 泌尿等专科治疗产品) 著称。世界各大制药, 医疗集团在此都有企业系列产品 (如百特, 美敦力, 圣犹达医疗) 的生产或研发机构。中国企业也利用华人华侨以及天时地利人

和的优势，不断通过洛杉矶开拓美国市场。已有近八十家企业和城市的美洲总部，分支机构，或办事处在此落户。

## 分会介绍

洛杉矶分会 2002 年成立，是在美国注册的非政治性、非赢利性的机构。现有会员六百余人，主要来自高科技企业（特别是生物和信息技术行业中）的资深技术及管理骨干。洛杉矶分会成立的 11 年中，促进中美之间的技术了解与合作，增进了会员之间的专业交流，为会员提供了成长，发展，创业的机会与平台。协会多次组织了会员回国考察，并接待了从国内各个地方代表团；为会员在中国推荐机会，把中国的商机引入大洛杉矶都市地区，洛杉矶分会是一座连接多方的桥梁。

## 联络方式

网站：[www.cast-la.org](http://www.cast-la.org) 联络人：会长何敏琰 [heminyan7@gmail.com](mailto:heminyan7@gmail.com)

## 6. 加州圣地亚哥分会

### 城市介绍：圣地亚哥（San Diego）



美国加州第二大城市，位于美国本土的极端西南角，与墨西哥接壤，与洛杉矶两个小时车程。圣地亚哥气候四季如春，海滩美丽如画。科罗拉多岛，La Jolla 海滩，海洋公园，动物园，Balboa 公园，航空母舰博物馆全国闻名。是美国著名的旅游胜地。

## 当地经贸

圣地亚哥经济发达，是美国无线通信技术中心，美国生物制药行业重镇，新能源与国防工业也非常强大。圣地亚哥是世



界市值最大的无线半导体公司，通讯技术领头羊高通公司总部所在地。并拥有索尼，诺基亚，三星，LG，雅虎，辉瑞，强生，诺华，里来等等许多世界著名企业的重要研发和生产基地。中国企业中兴，华为也在此设有研发中心。圣地亚哥教育科研能力强大，同时拥有世界著名的大学加州大学圣地亚哥分校（UCSD），SCRIPPS，SALK Institute 等研究所。大圣地亚哥地区 GDP 若单独计算在全世界国家中排名第 15。

## 分会介绍



圣地亚哥分会于 2002 年成立，是在美国注册的非政治性、非赢利性的机构，是圣地亚哥最活跃和最具影响力的华人协会之一。现有会员三百余人，主要来自高科技企业的技术骨干。圣地亚哥分会成立的十一年中，促进中美之间的技术合作，增进了会员之间的专业交流，为会员提供了成长、发展，创业的机会与平台。协会多次组织了会员回国考察，并接待了从国内各个地方代表团。为会员在中国推荐机会，把中国的商机引入圣地亚哥。

### 联络方式

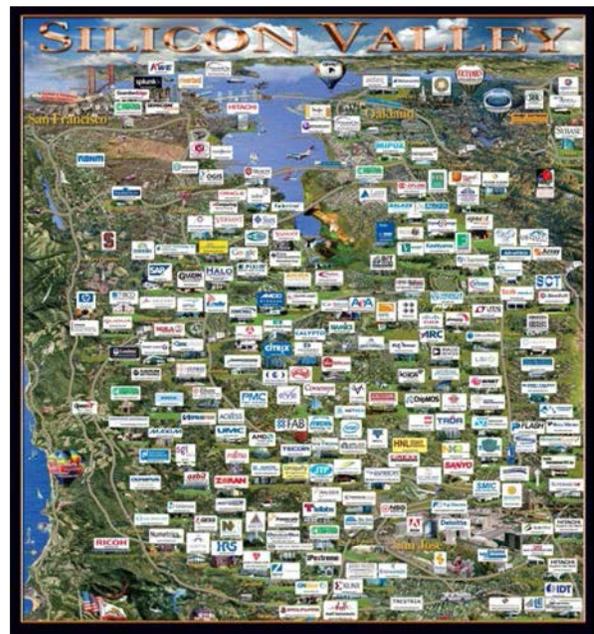
网站：[www.cast-sd.org](http://www.cast-sd.org) 联络人：会长  
陈海刚 [haigang\\_chen78@hotmail.com](mailto:haigang_chen78@hotmail.com)

## 7. 加州硅谷分会

### 地区介绍：硅谷（Silicon Valley）

硅谷位于美国西海岸的北加州，是旧金山湾区的一部分、美国西部最大的金融中心和重要的高新技术研发和制造基地，它是美国最富裕的地区之一以及第 3 大的美国大都市地区。此地背靠太平洋海岸山脉，面对旧金山湾的一片海洋，为山海所环绕。在这块连接旧金山市和圣何塞市之间 40 英里长、10 英里宽的狭长地带云集了近万家大小的高科技公司。80 年代

后，随着半导体、互联网、生物、空间、海洋、通讯、能源材料等新兴技术的研究机构在该地区纷纷出现，硅谷客观上成为美国高新技术的摇篮。现在硅谷已成为世界各国高科技聚集区的代名词。



### 当地经贸

硅谷是美国信息产业人才的集中地，高技术从业人员的密度居美国之首，人均 GDP 全美第一。硅谷集结着美国各地和世界各国的科技人员达 100 万以上，美国科学院院士在硅谷任职的就有近千人，诺贝尔奖的科学家达 30 多人。在硅谷，很多公



司都实行科学研究、技术开发和生产营销三位一体的经营机制，其特点是以附近一些具有雄厚科研力量的美国一流大学如 Stanford University 斯坦福大学、UC Berkeley 伯克利加大和 Cal Tech 加州理工等世界知名大学为依托，以高技术的中小公司群为基础，并拥有 Apple 苹果、Google 谷歌、Facebook 脸书、Intel 英特尔、HP 惠普、Cisco 思科、Lucent 朗讯，Tesla 特斯拉电动车等大公司，融科研、技术、生产为一体。这里高学历的专业科技人员往往占公司员工的 80% 以上。硅谷的科技人员大都是来自世界各地的佼佼者，许多亚裔是拥有高学历的科技新贵。

### 分会介绍

硅谷分会 (CAST-SV) 是中国旅美科协各分会中较新的一个，成立于 2010 年 4 月。硅谷分会秉承旅美科协的发展宗旨，弘扬中国文化，增进中美科技/文化/教育等方面的交流，努力为立志回国创业的同胞提供平台和信息，同时为在美华人融入美

国主流社会创造机会。硅谷有业已成熟的科技领先公司，也有正在蓬勃发展的新兴公司，更汇聚了来自世界各地的技术精英。这样自由、宽松又具竞争性的环境将会给旅美科协硅谷分会的发展提供一个广阔的空间。

### 联络方式

联络人：会长 丁博  
ericdingbo@gmail.com；理事长 于娜，  
annayu@sinolanguage.org

### 8. 佛罗里达分会

#### 城市介绍：戴托纳比奇 (Daytona Beach)



中国旅美科技协会佛州分会位于佛罗里达中部的戴托纳比奇 (Daytona Beach)。此地拥有世界最著名的海滩，四季阳光充裕，气候宜人。毗邻圣奥古斯丁历史遗址、奥兰多迪斯尼主题公园、海洋大世界、环球影城、乐高乐园和美国国家航空航天局 (NASA) 肯尼迪太空中心等享誉全球的景点，是著名的旅游胜地。戴托纳国际赛道是美国赛车协会 (NASCAR) 多项重大赛事如 Daytona 500, Coke-Zero 400 的举办地，吸引着全美、乃至世界各地的赛车爱好者。

## 当地经贸

佛罗里达州是美国第四大经济体。旅游业、农业、工业等都很发达。以 NASA 肯尼迪太空中心为中心的航空航天工业基础雄厚。佛州中部的高新技术走廊东起戴托纳比奇，穿越奥兰多都市区，延伸至西岸的坦帕，聚集了大量的新技术公司，



在农业科技、商业与军用飞行、国际金融、信息技术、微电子与纳米技术、光晶材料、生命科学与医疗技术、仿真训练技术、数字交互游戏等领域处于领先的地位。

## 分会介绍

佛罗里达分会于 2012 年成立。它的成员主要来自于生活在美国中佛罗里达地区，并在各行各业取得卓越成就的华人教授、高科技企业的技术骨干、及医学界和金融界专业人士。协会致力于为佛州地区旅美华人提供专业交流、合作和发展的平台，并积极搭架美国佛罗里达地区与中国国内的文化、科技、教育及经贸等领域交流与合作的桥梁。

## 联络方式

网站：[www.castfl.org](http://www.castfl.org) 联络人：会长 廖国斌：[globaltraderleo@gmail.com](mailto:globaltraderleo@gmail.com)

## 9. 北卡分会

**城市介绍：北卡研究三角园（RTP，NC）**

北卡罗莱纳州研究三角园（Research Triangle Park）位于美国北卡罗莱纳州的 Raleigh、Durham 和 Chapel Hill 三个主要城市之间的交接地带，并被北卡罗莱纳大学、北卡罗莱纳州立大学和杜克大学三所名校环绕，形式类似一个三角形。研究三角园

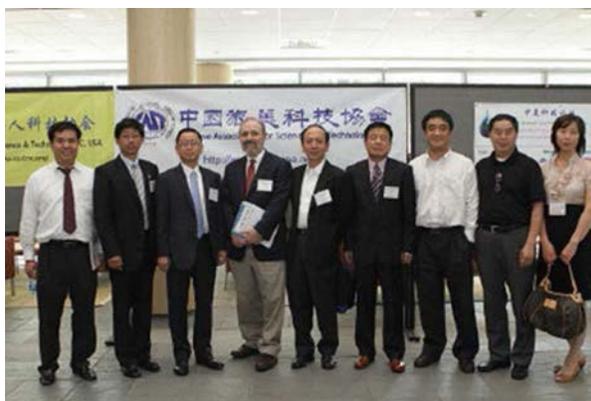


于 1959 年开始建设，规划面积约 7,000 英亩，经过 40 多年的建设，已发展成美国最大也是最成功的高科技园区之一。园区聚集了约 120 家研发机构，呈点状隐没于茂密的原始森林之中。研究领域包括生物医药、微电子学、计算机软硬件、信息技术、环境科学、材料科学等。主要的研究机构有：国家环境卫生科学研究院、美国国家环保署、北卡生物技术中心、微电子研究中心以及 GE、IBM、思科、北电网络、杜邦、葛兰素使克和联想等大型跨国公司。

鉴于三所著名大学、世界顶级公司和美国联邦卫生研究机构实验室间的相互合作，使北卡三角园区成为美国东部技术上的领跑者。园内拥有上万名科学家和工程师，曾有 3 位诺贝尔奖得主，有近百位美国联邦一级科研机构国家科学院、国家医学科学院和国家工程科学院院士，具有博士头衔的人口密度居全美第一。目前美国最重要的科学研究项目有 24%是在三角园研究成功的，每年世界上最重要的学术刊物发表的科研论文有 21%是三角园科学家撰写的，三角园科学家每年获得的技术专利约占全美 29%。北卡科技研究三角园也发展成与斯坦福工业园的“硅谷”及得克萨斯研究园齐名的美国三大科研中心，誉为美国“东部硅谷”。

### 分会介绍

旅美科技协会北卡分会位于享有盛名的美国北卡州三角区科技园，是在北卡州注册的非政治性、非赢利性的机构。自



2000 年成立以来，协会迅速发展壮大，拥有会员近两百名。会员绝大部分拥有硕士或博士学位，专业领域包括科学、技术、工程、教育、商业、法律、医学等，任职于本地的知名高等学府以及举世闻名的商

业巨擎如 GSK、IBM、思科、北方电讯、通用电气以及 SAS 等。北卡科协每年都组织部分有志于回国创业的会员回国考察访问，同时大力引进国内的招聘、科技交流访问团来北卡组织招聘活动，介绍国内的发展情况和机遇，为有意回国发展的朋友创造机会，立志将会员的科技能力转化成社会财富。2013 年 5 月 11 日，北卡科技分会与位于北卡州夏洛特市的南北卡华人工商联合，扩展成南北卡华人工商科技联合总会，着重于科学技术的产业化。

### 联络方式

网站：[www.castnc.org](http://www.castnc.org)； 联络人：会长 陆俊杰（[lujunjie2000@hotmail.com](mailto:lujunjie2000@hotmail.com)）； 理事长 娄玉（[yulou999@gmail.com](mailto:yulou999@gmail.com)）

### 10. 德州分会

#### 城市介绍：达拉斯（Dallas）

中国旅美科技协会德州分会位于德克



萨斯东北部的达拉斯市（Dallas）。达拉斯气候温和，少雨干燥，气候宜人。达拉斯有享誉全美的“达拉斯牛仔”橄榄球队、“小牛”篮球队、“伙伴”足球队、“得州轻骑兵”棒球队“达拉斯之星”曲棍球

队等。著名的旅游景点有勃勒杨小屋（Bryan Cabin），菲尔公园（Fair Park），公园中还有为了纪念德州独立一百周年而于 1936 年兴建的州会厅(Hall of State)。还有著名的达拉斯剧场中心（Dallas Theater Center）、肯尼迪总统纪念碑（Memorial Plaqueto John F.Kennedy）、白岩湖（White Rock Lake）等，吸引着世界各地的游客。

### 当地经贸

美国城市达拉斯（Dallas）是德克萨斯州（Texas）继休斯顿和圣安东尼奥之后的第三大城市，全美第八大城市。达拉斯是美国西南部金融和工商业中心，号称西南部的纽约：达拉斯联邦储备银行、100 多家商业银行、200 多家保险公司都设在此地；同时达拉斯也有雄厚的科技实力，位居美



“硅原”和“电讯走廊”的心脏地带，芯片制造业从业人数占全美 30% 以上。德州仪器、AT&T 总部，世界前十位电讯设备生产商等 600 多家高科技公司均落户于此。位于达拉斯市和福沃斯市之间的达福国际机场，占地 1.7 万英亩，是全球最繁忙的机场之一。

### 分会介绍

德州分会 1999 年成立。其会员分布在达拉斯、奥斯汀、休斯敦和圣安东尼奥等地，主要为并在各行各业取得卓越成就的华人教授、高科技企业的技术骨干、及医学界和金融界专业人士。德州分会致力于促进本地区学术交流与专业发展，支持年轻一代科技教育，及弘扬中国传统文化。

### 联络方式

网站：[cn.cast-tx.org](http://cn.cast-tx.org) 联络人：会长 胡文闯 [walterhu@gmail.com](mailto:walterhu@gmail.com)

### 11. 犹他分会

#### 城市介绍：盐湖城（Salt Lake City）



盐湖城是犹他州的政治、经济、文化中心，也是摩尔门教的发源地及总部。盐湖城地区曾被多家政府机构及媒体评选为

“最适合居住城市”、“最优投资环境城市”。盐湖城及周边城市拥有多个世界顶级滑雪场。此外，犹他州还有五个著名国家公园（Arches, Bryce Canyon, Canyonlands, Capitol Reef, and Zion），两个国家级休闲中心（Flaming Gorge and Glen Canyon），是常年休闲度假的好去处。

### 当地经贸

犹他州是极具经济活力的州。当今经济主要以高新技术、矿冶、贸易、运输、公用事业、旅游、州及联邦政府机构、专业及商业服务行业等为主。盐湖城还是美国的工业银行中心。犹他是信息技术的领导者。包括世界 500 强在内的十多家跨国公司总部设在盐湖城。

犹他州与中国有着长期的政治、经济、文化交流与合作。第十六任州长 Jon

的机构。犹他科协有注册会员二百余人，大多数会员是活跃于文化、经贸、科研和教育等领域，任职于美国知名公司、地方企业、高等院校的中高层管理人员或教学、科研工作者。犹他科协是犹他州最大、最活跃的华人社团之一。

犹他科协最近几年与中国在生物制药、医疗卫生、信息技术、能源材料、环境保护等方面开展了有效的交流活动，组织会员参加了国内省市的招商引资和人才引进的活动。

犹他科协成功组织每年组织或参与组织的主要社区活动包括犹他科协年会、面向华人同胞的各项讲座、盐湖城中国文化节、犹他州华人社区春节联欢晚会、面向美国当地民众开放的庆祝春节活动，以及夏天的露营活动、会员活动等。



Huntsman 于 2009~2011 年出任美国驻中国大使。“美中州长省长论坛”于 2011 年 7 月在盐湖城举办，为中美区域合作奠定了坚实基础。

### 分会介绍

犹他分会于 1996 年在盐湖城成立，是在美国犹他州注册的非政治性、非赢利性

### 联系方式

网站：<http://www.castut.org> 联系人：会长 冷剑伟 [jwleng@gmail.com](mailto:jwleng@gmail.com)

## 12. 西弗吉尼亚分会

### 城市介绍：摩根城（Morgantown）



摩根城地处西弗吉尼亚州北部，是一座宁静而美丽的大学城。西弗州最大的州立大学——西弗吉尼亚大学（West Virginia University）便坐落在这里。西弗吉尼亚大学成立于 1867 年，目前学生约 3 万人，西弗吉尼亚大学在能源、医学、生物识别等领域具有很强的研究实力。另外，西弗吉尼亚大学的橄榄球队、篮球队、足球队、射击队也是闻名全美。除了大学，摩根城还有两个国家实验室：能源部能源技术实验室和卫生部职业病预防中心。摩根城四季分明、安全舒适，连续多年被评为全美最宜居的城市之一。摩根城的交通也很便利，距华盛顿近 4 小时车程，距纽约 6 小时车程。

### 当地经贸

西弗吉尼亚州位于美国东部，是唯一一个全境处于阿巴拉契亚山脉的州，人口约 180 万。西弗州自然资源极为丰富：该州是美国第二大产煤州，近几年的页岩气革命也起源于此。另外，西弗州还有丰

富的石油、森林等自然资源。西弗州因其战略位置、高效的劳动力以及低经营成本吸引了来自全球各地的企业，如丰田、亚马逊、杜邦、洛克希德·马丁、普惠等世界知名公司。西弗州的经济也正由依赖传统能源向更为多元化的经济转变，如新能源、航空、汽车、生物信息、化工、生物技术、木材加工、旅游等。西弗州近年来也与中国保持了紧密的合作，目前西弗州与山西省是友好州省，摩根城与徐州市是友好城



市。西弗吉尼亚大学也与中国 20 余所大学建立了友好合作关系，该校也设有中美清洁煤研究中心、中国经济中心，并和天津财经大学共同建立了孔子学院。

### 分会介绍

CAST 西弗吉尼亚分会（筹委会）于 2015 年 7 月正式成立。会员主要来自西弗吉尼亚大学、卫生部职业病预防中心、能源部能源技术实验室和迈兰制药公司。

### 联络方式

联络人：会长 张汉霆 phone: 304-293-1488; email: hantingzhang@yahoo.com

### 13. 大波士顿分会

#### 地区介绍：大波士顿（Greater Boston）

波士顿是麻州首府，美国东部重镇，也是新英格兰最大的城市，始建于 1630 年，是美国最古老、最有文化价值的城市之一。大波士顿包括 Boston, Cambridge, Quincy 和

地区设立了研发基地，包括 Novartis, Pfizer, GSK, Merck, AstraZeneca, BMS, Abbott, Amgen 等。因此大波士顿地区聚集了数万名从事生物医药研究的高科技人才，科研资金雄厚。以全美国排名第一的哈佛医学院麻省总医院(MGH)为例，它拥有世界上所有医院中规模最大的科研项目，



其他周边城市群，人口 470 多万，在美国大都市中排名第 10。它是美国精神（Spirit of America）的发源地。美国独立战争的第一枪在这里打响。大波士顿地区历史悠久，是美国文化教育，科研创新和医疗保健中心。该地区拥有 Harvard, MIT, Tufts, Brandies, Boston College, BU, Northeastern and U. Mass 等 122 所知名高等院校。这在美国乃至全世界都没有可以相比的。

#### 当地经贸

大波士顿的经济基础是科研、金融与技术，特别是生物工程。它是美国的生物技术中心，也是重要经济区之一。在著名大学 Harvard 和 MIT 周围，聚集着 Biogen-Idec, Genzyme, Millennium, Vertex, Genetic Institute 等数百家以创新药物研发为目标的生物医药公司。除此之外，世界上销售额排名前十位的制药公司基本上都在该

地区设立了研发基地，包括 Novartis, Pfizer, GSK, Merck, AstraZeneca, BMS, Abbott, Amgen 等。因此大波士顿地区聚集了数万名从事生物医药研究的高科技人才，科研资金雄厚。以全美国排名第一的哈佛医学院麻省总医院(MGH)为例，它拥有世界上所有医院中规模最大的科研项目，仅此一家医院就拥有年度科研基金 7.5 亿美元。同时，大波士顿地区吸引了众多科技和风险投资，企业资金充沛，融资渠道广泛，是名符其实的美国生物医学研究及产业化的重镇。

#### 分会介绍

旅美科协大波士顿分会是 2016 年 1 月 30 日成立的，是旅美科协大家庭中的一个新成员。波士顿分会传承了旅美科协总会促进中美科技和学术交流的宗旨，专注于推动学术成果向产业化转化，探讨科技创新的不同模式，立志打造大波士顿地区对接



优秀科技项目，企业界人才和资本的平台。自成立以来，积极开展各项活动，与波士顿地区其他华人专业协会合作，为促进中美之间的文化，教育，科技和人才交流做出了应有的贡献。目前会员有 100 多人，绝大部分拥有博士或硕士学位，专业领域包括生物，医学，工程、教育、商业、法律等，任职于本地的知名高校或各生物和医疗技术公司，律师事务所，金融机构等，也有成功的创业者和正在创业或准备创业的勇士。

### 联络方式

网站：[www.cast-boston.org](http://www.cast-boston.org); 联络人：  
会长 雷明 [mlei\\_05@yahoo.com](mailto:mlei_05@yahoo.com)

## 14. 南佛州分会

### 地区介绍

南佛罗里达地处于美国的最东南部，这里气候宜人，有美丽的阳光和海滩，是世界著名的旅游胜地，集科技经贸、旅游航天、主题公园、及世界 500 强跨国公司为一体，堪称是企业家和实业家的乐园。自 2015 年 11 月 8 号筹备工作开始，中国旅美科协南佛州分会举办了多次大型活动，吸引了众多

的州府、县郡官员、企业家以及华人华侨科技人员和企业，活跃了南佛州华人华侨的生活，为中、美两国的各界人士提供了一个分享信息资源和互相交流的平台。

### 分会介绍

中国旅美科协南佛州分会（CAST-SFL）在总会的支持下，于 2016 年 9 月 17 日在南佛州的博卡正式成立。中国旅美科协南佛州分会会长为 凌肖；副会长为 王庆，邓松，李光义，乐晓钟，钱锋，王军，周业；中国旅美科协南佛州分会理事会理事长为：李忠刚，副理事长为：林松，殷钟，杨莅；秘书长为：周薇。

在中国旅美科协总会的领导下，南佛州分会将努力为华人华裔服务，充分利用并分享旅美科协在各州的科技资源与信息，



成为一个引领华人科技企业协会并服务于南佛州华人社区及科技企业界的优秀社团。

### 联络方式

联络人：会长凌肖  
alex\_xiao\_ling@yahoo.com

## 15. 内布拉斯加州分会

### 内州介绍：（Nebraska）

内布拉斯加州位于美国中西部的大平原地区。州府位于林肯市。最大的城市是位于密苏里河畔的奥马哈。奥马哈主办每年的大学棒球世界联赛和每四年一次的美国奥运游泳队选拔赛。旅游网站 Trip Advisor 在 2014 评选位于奥马哈的 Henry Doorly 动物园为世界最好的动物园。内布拉斯加大学医学中心初级卫生保健项目 2015 年在全美排名第四。

### 当地经贸

内州的经济以农业为主。盛产牛肉，猪肉，玉米，大豆和高粱等。其它重要的经济形式包括货物运输，制造，通讯，信息技术和保险。内州拥有许多著名的公司，包括世界首富 Warren Buffett 担任 CEO 的 Berkshire Hathaway。将总部设于内州的世界 500 强公司有 Mutual of Omaha, Kiewit Corporation 和 Union Pacific Railroad。另外其它大公司还有 InfoUSA, TD Ameritrade, West Corporation, Valmont Industries, Woodmen of the World 和 Gallup 等。

### 分会介绍

内州分会于 2016 年 11 月成立。会员主要由来自科技、文化、教育、法律、金融、人文等领域的居住于内州的中国旅美专业人士组成。协会致力于为内州地区旅美华人提供专业交流、合作和发展的平台，并积极搭架美国内州与中国国内的文化、科技、教育及经贸等领域交流与合作的桥梁。





### 联络方式

网址：<http://www.cast-ne.org/> 联系人：  
会长 刘宇彤 (Yutong.liu@gmail.com)

## 旅美科协驻中国办事处/联络处 负责人和联系方式

国内办事处、联络处：北京、上海、广州；大连、济南、青岛、重庆、成都、安徽、长沙、杭州、南京、东莞、沈阳

- 北京办事处：  
沈陆 lushen818@yahoo.com  
冯杏建 gracef2001@hotmail.com
- 上海办事处：  
闵洪忠 stephen@sunfory.com  
温红媚 char98@126.com
- 广州办事处：  
贾鹏程 jiapc@yahoo.com  
王喜英 wangxy@sunmnet.com  
张涛 zhangt@sunmnet.com
- 大连联络处：  
李拓 tuoli88@yahoo.com
- 济南联络处：  
刘新国 liuxinguo@hbio.cn
- 青岛联络处：  
季文海 wenhaiji@hotmail.com
- 重庆联络处：  
冯毅 harvey.feng@gmail.com
- 成都联络处：  
唐亦凡 19585907@qq.com
- 安徽联络处：  
廖结楼  
孔德怡 kongdeyi@tsinghua.org.cn
- 长沙联络处：  
宋晓兰 Xlsong8822@126.com  
骆军 Luo.john@hotmail.com
- 杭州联络处：  
谢长春 cxiem@yahoo.com
- 南京联络处：  
邹云峰 yunfengzou@hotmail.com
- 东莞联络处：陈友斌
- 沈阳联络处：  
张兴 xingucsd@163.com





中国科学院宁波材料技术与工程研究所

**CNITECH**

**先进制造技术研究所**

**招贤纳士**

### 先进制造技术研究所简介

先进制造技术研究所是中国科学院宁波材料技术与工程研究所下属的二级所之一。先进制造所以中科院“三个面向”、“四个率先”办院方针为指导，秉承材料所“料要成材、材要成器、器要好用”建所理念，坚持多学科交叉融合，贯通“材料-设计-制造”全产业链技术创新链条。先进制造所已规划部署复合材料制造与装备、功能器件制造与系统、精密制造工艺与系统、机器人与智能制造装备技术、制造信息技术等五大科研领域；已引进和招募科研人员近200名，其中“旗舰”研究员3人、研究员及正高级工程师16人、副研究员及高级工程师19人、助理研究员及工程师等科研人员50余人，高层次人才中有4名中组部“千人计划”专家、5名浙江省“千人计划”专家、3名中科院“百人计划”专家。先进制造所已成功获批中科院“机器人与智能制造”创新研究院宁波分部、浙江省级重点企业设计院，拥有浙江省机器人与智能制造装备技术重点实验室、宁波市激光精密制造重点实验室；初步建成材料制造、器件制造、精密制造、智能制造和数字化设计仿真等装备平台，以及一个占地3000平方米的大型工程实验中心，各类实验设备总值1亿余元；建所7年来，已承担国家、科学院、省市地方及重点企业委托项目等百余项，累计获得竞争性经费近2.5亿元；共发表文章450余篇，申请专利411项，获授权专利223项。

(详见中科院宁波工研院先进制造技术研究所网址 <http://manufacture.nimte.cas.cn/>)

根据研究所发展的需要，为进一步加强研究力量，拓展研究方向，先进制造所将继续面向全球公开招聘优秀人才。

### 招聘方向

制造信息技术	复合材料制造与装备	
<ul style="list-style-type: none"> <li>● 工业传感器网络与监控技术</li> <li>● 设备预后诊断与健康管理技术</li> </ul>	<ul style="list-style-type: none"> <li>● 复合材料集成设计</li> <li>● 复合材料连接与装备技术</li> </ul>	<ul style="list-style-type: none"> <li>● 复合材料成型工艺与装备</li> <li>● 复合材料部件测试与评价</li> </ul>
功能器件制造与系统	精密制造工艺与系统	机器人与智能制造装备技术
<ul style="list-style-type: none"> <li>● 光电探测材料与器件</li> <li>● 热电转换材料与器件</li> </ul>	<ul style="list-style-type: none"> <li>● 精密与极端减材制造</li> <li>● 精密与极端表面处理</li> <li>● 精密与极端增材制造</li> </ul>	<ul style="list-style-type: none"> <li>● 精密驱动控制技术</li> <li>● 先进机器人及自动化技术</li> <li>● 智能感知技术</li> </ul>

### 招聘基本条件

1. 应聘人员具有博士学位或有海外科研工作经历者优先考虑；
2. 在国际学术界或企业界有一定的影响或做出过具有国际水平的研究成果，或拥有重大发明（专利），掌握该学科领域能影响高新技术产业化的关键技术的优先考虑；
3. 胸怀宽广，具有立足全国、放眼全球，为我国科技事业发展和国民经济转型升级而艰苦奋斗的奉献精神；
4. 具备项目管理经验和团队人员管理能力者优先考虑。

### 招聘岗位及待遇

岗位名称	科研启动经费	安家补贴
“旗舰人才”高级研究员	300-1000万元	享受135-160平米优惠购房与30万安家费
“团队人才”研究员/项目研究员	300万元/260万元	享受115-125平米优惠购房与20万安家费
“关键人才”正高级工程师/高级工程师	100万元/50万元	享受125/100平米优惠购房与20/10万安家费
“春蕾人才”副研究员	100万元	享受100平米优惠购房与10万安家费
特聘杰出青年研究员	100万元	出站留所可享受20万安家费
优秀博士后	在站期间20万元/年（最长可享两年） 出站留单位工作，经考核可再获20万元	出站留所可享受20万安家费

◆ 聘为研究员/项目研究员或副研究员，符合条件者可推荐申请中组部“千人计划”或中科院“百人计划”及省、市人才计划；其中青年千人计划可获得中组部提供的200-300万元，中科院提供的200万元与宁波材料所提供的300万元科研经费资助，同时额外享受270万元的个人补贴。

◆ 按中科院三元结构工资，提供有竞争力的薪酬（如特聘杰出青年研究员年薪25万、优秀博士后年薪不低于20万）；本人和子女纳入宁波市社保体系；高层次人才享受带薪休假及研究所各类补贴。同时，研究所安心工程政策可协助引进人才的配偶工作安置、安排义务教育段子女的入学等。

联系人：毛佳佳

联系电话：+86 ( 574 ) 87605861

联系邮箱：maojijia@nimte.ac.cn



中國旅美科技協會

Chinese Association for Science and Technology, USA (CAST-USA)  
[www.cast-usa.org](http://www.cast-usa.org)