



中国科学院自动化研究所 模式识别国家重点实验室

简 报

2016 年第 4 期

总第 40 期

2016 年 10-12 月

目 录

最新动态	1
国家重点研发计划“面向大范围场景透彻感知的视觉大数据智能分析关键技术与验证系统”项目启动会顺利召开	1
国家重点研发计划“云端融合多模态数据交互意图理解”课题实施方案研讨会暨启动会顺利召开	1
2016 年中欧信息科学与技术研讨会成功举办	1
第 15 届国际手写识别前沿会议成功举办	2
第一届多模态情感识别竞赛成功举办	3
第三届脑网络组和基因组汇聚国际学术研讨会召开	3
多媒体计算团队在社会事件分析方面获得两项最佳论文奖	4
桑基韬和申抒含副研究员分获 2016 年 ACM 中国和北京新星奖	4
刘康副研究员获 Google Focused Research Award.....	5
黄凯奇研究员获 2016 年 CCF-IEEE CS 青年科学家奖.....	5
董晶副研究员入选 IEEE 亚太区执行委员会.....	5
孙哲南研究员当选国际模式识别学会会士	5
实验室博士生刘强获“微软学者”奖学金.....	6
实验室团队在 ECCV 2016 描述子竞赛中获得佳绩.....	6
实验室团队在 2016 年上海 BOT 视觉大数据竞赛中获得佳绩	6
科研进展	7
基于精确光照环境估计的人脸图像拼接检测	7

基于 RGB-D 数据的 One-shot Learning 手势识别	7
基于无监督多层描述子学习的行人再辨识	8
基于重要事件时序-权值分析图的体育新闻自动生成方法	8
基于温和正样本挖掘和深度度量学习的行人再辨识	9
面向序列预测的神经网络深层结构设计	10
基于句子级关注机制和实体描述的弱监督关系抽取	10
基于层次化上下文多示例多标签学习的图像标注方法	11
基于主动结构光线的文档图像矫正	12
动态并行及分布式图割算法	13
基于概率图模型的视频人物标识	13
学术交流	14
德国人工智能研究中心 Hans Uszkoreit 教授等来实验室访问交流	14
南加利福尼亚大学的 C.-C. Jay Kuo 教授来实验室访问交流	14
脑网络组研究中心邀请 8 位国内外专家来实验室访问交流	15
项目立项	15
实验室第四季度新建立课题 25 项	15

《模式识别国家重点实验室简报》编委会

内容审核

刘成林 陶建华

编辑小组

组长：申抒含	左年明				
成员：桑基韬	叶军涛	张 曼	刘 斌	张家俊	殷 飞
原春锋	廖胜才	孟高峰	张一帆	戴 佳	王爱华

最新动态

国家重点研发计划“面向大范围场景透彻感知的视觉大数据智能分析关键技术与验证系统”项目启动会顺利召开

2016 年 10 月 31 日, 由自动化所模式识别国家重点实验室王亮研究员任首席科学家的 2016 年国家重点研发计划云计算和大数据重点专项“面向大范围场景透彻感知的视觉大数据智能分析关键技术与验证系统”项目启动会在自动化所召开, 标志着该项目正式进入全面实施阶段。

由郑南宁院士、谭铁牛院士、李波教授、赵耀教授、陈熙霖研究员、查红彬教授等组成的专家组成员、项目牵头单位自动化所刘成林副所长, 以及科技部高技术研究发展中心、中科院科技促进发展局的相关领导出席了此次项目启动会。项目首席科学家王亮研究员对本次重点研发计划项目做了整体汇报, 从立项背景、关键科学问题、主要研究内容与预期目标、总体研究方案与技术路线、课题设置与研究团队、年度计划以及项目组织管理等方面加以介绍。专家组成员针对每个课题分别提出了意见。此次项目启动会的顺利召开, 使本项目进一步凝练了总体目标, 明确了研究任务, 梳理了研究思路, 促进了项目组成员间的学术交流与合作, 同时加深了对项目和课题研究方案的理解, 为项目的顺利实施奠定了良好基础。

国家重点研发计划“云端融合多模态数据交互意图理解”课题实施方案研讨会暨启动会顺利召开

2016 年 11 月 15 日, 2016 年国家重点研发计划课题“云端融合多模态数据交互意图理解”在第三军医大学第一附属医院(西南医院)召开启动会。本课题由中科院自动化所牵头, 由北京科技大学、复旦大学、山东大学、西南医院共同承担。

本次会议主要包括课题内容讨论交流、人机交互计算机智能辅助无菌手术室应用需求及后续工作部署等内容。课题负责人陶建华研究员和各参与单位就工作目标、研究内容等进行了详细汇报, 项目负责人、各课题负责人员及参与单位人员共 30 余人就项目的总体实施方案、工作目标、研究内容、进度安排等具体方案进行了详细汇报和讨论。随后, 课题组对西南医院手术室无菌环境及手术现场进行了实地观摩。参观完后, 课题组继续在医教部会议室进行了具体的技术路线讨论, 根据临床专家提出的需求及实地观摩情况, 进一步细化了任务分工, 明确了时间节点。本次启动会的成功举办为课题研究的有序推进奠定了坚实的基础。

2016 年中欧信息科学与技术研讨会成功举办

10 月 10-11 日, 由中科院国际合作局、中科院自动化所、法国国家信息与自动化研究学院(INRIA)、法国国家科研中心、瑞士洛桑联邦理工学院等单位共同主办, 中欧信息、自动化与应用数学联合实验室(LIAMA)承办的 2016 年中欧信息科学与技术研讨会(The 1st Sino-European Workshop on Information Science and Technology, IST2016)在自动化所顺利举行。中科院自动化所、法国国家信息与自动化研究学院、法国国家科研中心以及瑞士洛桑联邦理工学院等

参与了 IST2016 的会议组织工作, LIAMA 中方主任、模式识别国家重点实验室副主任陶建华研究员和 LIAMA 法方主任 Jean-Francois Monin 共同担任大会主席,会议邀请了中科院副院长谭铁牛院士、中科院自动化研究所所长徐波研究员、中科院自动化所副所长刘成林研究员以及 INRIA 首席执行官 Antoine Petit 教授作为特邀嘉宾。

本次研讨会邀请了多位知名专家学者做了 20 余场主题报告,报告涉及脑科学、人机交互以及人工智能等多个领域。与会来宾分享了各自领域中最新的研究工作,介绍了相关领域中最新最热的理论和方法,使与会者接触到了本学科最前沿的研究工作和团队。会议受到了广泛的关注,共有 80 余人参加了会议,与会人员与各个领域的教授进行了热烈的讨论,会议圆满成功。

会议为中方和欧方的科研工作者以及研究机构提供了一个分享科研成果和科研想法的平台,并为双方在信息科技领域共同申请新的合作项目提供良好机会,有助于推动中欧双方学术的交流与进步。

第 15 届国际手写识别前沿会议成功举办

2016 年 10 月 23-26 日,第 15 届国际手写识别前沿会议(International Conference on Frontiers in Handwriting Recognition, ICFHR 2016)在深圳召开。这是该会议第一次在中国召开,由中科院自动化所承办。会议围绕手写文档处理与分割、深度学习、联机 and 脱机手写识别、手写文本识别、文档检索、笔迹鉴别等内容开展了深入的交流和探讨,吸引了众多学术界和工业界的研究人员前来参会,到会代表 150 多人。本次大会的主席是中科院自动化所刘成林研究员和清华大学深圳研究生院陈友斌教授。加拿大著名学者,Pattern Recognition 期刊主编、ICFHR 创始人 Ching Y. Suen 教授担任本次会议的名誉主席。

本次大会邀请多位国际知名学者专家做大会特邀报告。荷兰格罗宁根大学的 Lambert Schomaker 教授做了题为 "How deep is deep and what's next in computational intelligence?" 的报告,深刻阐述了深度学习以及计算智能的关系,以及手写识别和文档分析中很多深度学习不能解决的问题。德国亚琛工业大学的 Hermann Ney 教授做了题为 "Handwriting and Speech Recognition: From Bayes Decision Rule to Deep Neural Networks" 的报告,详细介绍了从贝叶斯决策到深度神经网络的转变过程,以及语音识别和手写识别中的隐马尔可夫模型和深度神经网络建模最新进展。日本东京农工大学的 Masaki Nakagawa 教授做了题为 "Online handwriting recognition: past, present and future" 的报告,生动刻画了联机手写识别的过去、现在与未来,指出手写识别研究和应用的广阔前景。会议还组织了一个由学术界和工业界研究者参加的 Panel Discussion,期间显示了很多手写识别领域的应用需求、挑战性技术问题和人才需求。

本次大会包含 7 个 Oral Session 以及 2 个 Poster Session,与本次大会同时举办的还包括 8 个围绕手写识别多方面技术的国际竞赛,吸引了来自全球的众多参赛团队,会议期间对竞赛获胜团队进行了表彰。同时大会还选出了最佳论文奖、最佳学生论文奖、最佳 Poster 奖。会议期间还投票选出了 ICFHR 2020 的举办地为德国的多特蒙德。

第一届多模态情感识别竞赛成功举办

第一届多模态情感识别竞赛 Multimodal Emotion Recognition Challenge (MEC 2016) 于 2016 年 11 月 6 日, 作为全国模式识别学术会议 (Chinese Conference on Pattern Recognition, CCPR) 的特殊议题在电子科技大学成功举办。

近年来, 情感计算研究已受到学术界的广泛重视, 然而目前国内外尚缺乏大规模的中文情感数据库。为此, 模式识别国家重点实验室人机交互课题组李雅副研究员带领学生构建了离散情感标注数据库 (CASIA Chinese Emotional Audio-Visual Database, CHEAVD)。该数据库的数据来源于中文影视剧中所截取的音视频片段, 每一个音视频片段分别标注为一些常见情感 (高兴、悲伤、生气、惊讶、厌恶、担心、焦虑) 及中性情感中的一种。相关研究成果发表在人机交互领域国际期刊 Journal of Ambient Intelligence and Humanized Computing 上。

在 CHEAVD 数据库基础上, 陶建华研究员同情感计算领域 Björn Schuller 教授作为共同主席组织了第一届多模态情感识别竞赛, 期望为面向情感识别领域的研究人员提供统一的数据集及验证平台, 加快情感计算研究的发展。这届竞赛包括音频情感识别、表情识别及音视频融合的情感识别三项子任务。经过严格的两轮评比以及论文评审, 最终选取了 9 篇论文在全国模式识别会议 CCPR 2016 上进行交流。

第三届脑网络组和基因组汇聚国际学术研讨会召开

2016 年 11 月 4-6 日, 第三届脑网络组和基因组汇聚国际学术研讨会 (The Second Symposium on Brainnetome Meets Genome, SBMG 2016) 在海口成功召开。SBMG 2016 由模式识别国家重点实验室脑网络组研究中心、脑网络组北京市重点实验室、澳大利亚昆士兰大学脑研究所、海南金港生物科技有限公司和昆明理工大学共同主办, 得到了国家自然科学基金委、中国科学院、中国科学院自动化研究所、海南金港生物科技有限公司和昆明理工大学的大力支持, 吸引了来自澳大利亚、美国、德国、英国、法国等国外本领域的著名学者参加, 中科院生物物理所、北京大学、第四军医大学、电子科技大学等高校及研究机构的相关团队也派出代表参加。这个系列会议旨在汇集本领域从事脑网络科学、神经影像学、基因组学、临床神经科学等领域的专家学者共同探讨脑网络组和基因组汇聚的最新进展、面临的挑战及未来发展机遇、推动相关领域的国际合作。

会议期间, 与会专家对脑网络组和基因组及其汇聚的前沿科学问题, 特别是微观和宏观脑网络组研究、脑网络组计算理论和方法、全基因组关联研究、基因组和脑网络组如何汇聚及其在重大脑疾病中的应用等方面的问题进行了深入的探讨。研讨会最后一个议题邀请部分来自不同国家的专家介绍各自机构和实验室的研究方向、拥有的研究条件、期望的合作领域等, 帮助与会者在共同感兴趣的研究问题上开展切实合作。

脑网络组和基因组及其汇聚研讨会促进了脑网络组研究中心与国际上本领域学者之间的学术交流合作与联系, 也为今后的进一步深入研究和相互合作奠定了良好基础。

多媒体计算团队在社会事件分析方面获得两项最佳论文奖

2016 年 10 月 19 日,在荷兰阿姆斯特丹举办的 ACM 多媒体国际会议(ACM Multimedia 2016)上,模式识别国家重点实验室多媒体计算团队博士生钱胜胜、张天柱副研究员和徐常胜研究员共同署名的文章“Multi-modal Multi-view Topic-opinion Mining for Social Event Analysis”荣获大会唯一的最佳论文奖。鲍秉坤副研究员和徐常胜研究员共同署名的文章“Cross-Platform Emerging Topic Detection and Elaboration from Multimedia Streams”获得 ACM Transactions on Multimedia Computing, Communications, and Applications 2016 年度唯一的最佳论文奖。

ACM Multimedia 创办于 1993 年,是多媒体领域最具影响力和最重要的国际会议。ACM Multimedia 2016 是第 24 届会议,共收到投稿 650 篇,录用长文 52 篇,经由多媒体领域资深学者和专家组成的评奖委员会评定和推荐,在四篇最佳论文提名文章中,一篇被确定为最佳论文。

获奖论文针对社会事件数据中呈现的多视角多模态等属性,提出一种多模态多视角的主题观点挖掘方法。该方法利用不同新闻媒体对同一个社会热点事件观点的差异性,学习不同视角的主题及对应的观点。通过本方法,既能够细致地分析和研究不同媒体对社会热点事件的主题和观点,又能够对社会热点事件进行全方位展示,便于探究热点事件的全局态势与舆情走势。

ACM Transactions on Multimedia Computing, Communications, and Applications 由国际计算机协会(ACM)承办并出版,是多媒体领域的顶级期刊。2016 年度最佳论文评选是对 2015 年度发表的 52 篇论文经过由国际多媒体领域的资深专家组成的评奖委员会两轮评审确定的,获奖论文的作者受邀请出席 ACM 国际多媒体会议,并在颁奖大会上领奖。

获奖论文从多个社交平台上检测社会热点事件,并将其全面的展现出来,不仅切合政府和民众的需求,也是国际学术的研究热点。获奖论文敏锐地利用不同社交平台的在传播速度和传播内容上的特点,提出了一个有效的跨平台跨模态的社会热点事件检测和展现方法。选题的新颖性、解决思路的创新性、以及技术的有效性,均受到评审专家的一致认可。

桑基韬和申抒含副研究员分获 2016 年 ACM 中国 and 北京新星奖

模式识别国家重点实验室桑基韬和申抒含副研究员分别获得国际计算机学会(ACM)颁发的 2016 年中国新星奖(ACM China Rising Star Award)和北京新星奖(ACM Beijing Rising Star Award)。ACM 中国理事会于 2014 年设立“ACM 中国新星奖”,每年评选出 1-2 名在计算机领域有突出科研潜力的青年学者,以表彰中国计算机领域的青年学者所取得的研究成绩,推动青年学者在科研领域的发展与进步。

2016 年 ACM 新星奖的评选首先由 ACM 全国各分会分别评选 1-2 名分会新星奖,本年度 16 个 ACM 分会共评选出 24 位分会新星奖获奖者,ACM 中国奖励委员会从中选出 4 位候选人参加中国新星奖答辩。10 月 19 日,ACM 主席 Vicki Hanson、ACM 首席执行官 Robert Schnabel、ACM 中国理事会主席刘云浩教授,以及 ACM 中国理事、各分会主席等共计 30 多位专家评委通过现场答辩最终投票选出 2 位中国新星奖获奖者。

刘康副研究员获 Google Focused Research Award

模式识别国家重点实验室刘康副研究员经过 Google 公司多名专家匿名评审, 获得 2017 年度 Google Focused Research Award, 资助项目为:《Cross-lingual Text and Knowledge Graph Jointly Embeddings》。Google Focused Research Awards 是由 Google 公司每年面向全球大学、研究所等学术研究机构发布, 致力于针对 Google 公司所关注的计算机科学领域关键前沿问题展开研究。

刘康副研究员申请的专项研究计划是“Google Language Understanding and Knowledge Discovery Focused Research Awards Program”, 该课题主要致力于跨语言知识和文本的联合表示学习问题, 力图将多种语言的知识、文本放入统一的学习框架中, 学习它们的语义表示, 从而解决多语言知识图谱的构建、小语种文本内容理解与知识推理问题。该研究对于促进多语言知识图谱构建具有重要的科学意义。

黄凯奇研究员获 2016 年 CCF-IEEE CS 青年科学家奖

在 2016 年 10 月 20-22 日召开的 2016 年中国计算机大会上, 模式识别国家重点实验室黄凯奇研究员获得 2016 年 CCF-IEEE CS 青年科学家奖, 以表彰他在计算机视觉学术及应用领域做出的贡献。

CCF-IEEE CS 青年科学家奖授予在科学研究方面有突出成就和重要贡献、年龄不超过 40 岁的青年学者。从 2016 年起由 CCF 和 IEEE Computer Society 共同评选和颁发。本年度的颁奖仪式于 10 月 22 日在山西国际会议中心举行, 由评奖委员会主席张晓东教授主持颁奖仪式, CCF 理事长高文教授, IEEE-CS 理事长 Roger Fujii, 百度公司代表赵世奇博士为获奖者颁奖。

董晶副研究员入选 IEEE 亚太区执行委员会

2016 年 12 月 2 日, 国际电气电子工程师协会 (IEEE) 公布了亚太区新执行委员会名单, 模式识别国家重点实验室智能感知与计算研究中心董晶副研究员再次入选 IEEE 亚太区 (Region 10) 执行委员会, 任期为 2017 至 2018 年, 同时担任该地区 WIE (Women in Engineering Coordinator) 协调员。在本届 IEEE 亚太区执行委员会委员中, 董晶副研究员是中国大陆的唯一执委, 也是历年担任此职位的最年轻的中国女性。

IEEE 是全世界最大的国际性专业技术学会, 也是在电子电气与信息技术领域最具影响力的国际学术组织。IEEE 目前拥有超过 40 万会员, 分布世界 160 多个国家。IEEE 按地域共分为 10 个区, 亚洲太平洋区属于第 10 区 (R10), 拥有近 8 万会员。

孙哲南研究员当选国际模式识别学会会士

2016 年 12 月 5-8 日召开的第 23 届国际模式识别大会发布了新一届的会士 (Fellow) 名单, 模式识别国家重点实验室智能感知与计算研究中心的孙哲南研究员因其在自动虹膜识别及其应用领域的贡献当选国际模式识别学会 (The International Association for Pattern Recognition) 会士。

国际模式识别学会是由世界各国模式识别方面的专家和团体组成的国际性权威学术组织，IAPR Fellow 每两年评选一次，当选率少于会员总数的 0.25%。2016 年全球共 28 人当选 IAPR Fellow。

实验室博士生刘强获“微软学者”奖学金

模式识别国家重点实验室智能感知与计算研究中心博士生刘强同学荣获 2016 年度“微软学者”奖学金。今年的“微软学者”奖学金吸引了来自中国大陆、香港、台湾、日本、韩国和新加坡等地的四十余所著名高校和研究机构的近百名博士生申请。经过评委委员会的严格考核和评审，共有 10 名计算机领域的博士生成为 2016 年“微软学者”，获得微软研究院提供的 10000 美元奖学金资助，以及为期 3 个月的学术访问机会。刘强同学于 2013 年进入中科院自动化所学习，研究方向为数据挖掘与用户建模，其研究成果相继发表于 AACL, SIGIR, CIKM, ICDM, IEEE TKDE 等国际会议和国际期刊。

实验室团队在 ECCV 2016 描述子竞赛中获得佳绩

在 ECCV 2016 Workshop“Local Features: State of the art, open problems and performance evaluation”组织的局部图像描述子竞赛中，自动化所模式识别国家重点实验室吴福朝研究员、樊斌副研究员和博士研究生田雨润所组成的队伍在比赛中获得了第一名的成绩。

局部图像特征描述一直是计算机视觉领域的一个重要研究方向，以 SIFT 为代表的一系列经典的手工设计描述子在各个领域都有广泛的应用。ECCV 2016 Workshop 组织此次竞赛的目的在于发掘更多优秀的学习型描述子。此次竞赛分为图像匹配，图像检索以及图像分类三个项目，比赛数据包括了实际应用中图像的视角变换、光照变化等问题。共有来自布拉格捷克理工大学，汉诺威大学，萨格勒布大学和中科院自动化研究所的五支队伍参加了此次竞赛。来自模式识别国家重点实验室的队伍提出了一种基于 CNN 的方法来进行描述子的学习，该新型描述子的效果超过现有的主流学习型与手工描述子，并在三个项目的比赛中均获得了第一名。

实验室团队在 2016 年上海 BOT 视觉大数据竞赛中获得佳绩

2016 年 11 月 11 日，“2016 上海大数据应用大赛”总决赛在上海市大数据产业基地落下帷幕，来自模式识别国家重点实验室图像视频分析课题组的王金桥研究员带领 ITdog 团队获得计算机视觉竞赛的第一名和最佳算法奖。

本届大赛由上海大数据产业基地（市北高新）、上海大数据联盟、英特尔（中国）有限公司和华院数据技术（上海）有限公司联合主办，奖金总数 80 万元。今年是上海 BOT 大数据应用大赛首届举办，大赛聚焦计算机视觉与人工智能聊天机器人商业应用两大热门赛题。大赛历时 4 个月，报名参赛近 800 人，组建了约 400 支团队，选手来自 13 个国家和地区，提交作品 500 多次。计算机视觉赛题有“隐藏任务”、“双赛题制”等创新性环节和设置，为参赛选手不断提供新的挑战。11 月 10 日开始的计算机视觉赛题决赛，采用了世界流行的开发者模式——黑客松的形式，并现场公布赛题——机器视觉视觉智能问答(VQA)。十强团队既

有来自中科院、清华大学、北京大学、上海交大等高校和科研机构的团队，也有来自百度、360 等企业的团队。最终实验室 ITdog 团队获得冠军，包揽了第一名和最佳算法奖，获得奖金 21 万元。

科研进展

基于精确光照环境估计的人脸图像拼接检测

由于人脸图像的高度敏感性和重要性，涉及人脸的图像拼接和篡改具有非常严重、恶劣的影响，这也是数字图像取证和信息可信性研究中的一个重要课题。基于光照环境估计的拼接检测算法假设真实图片中的人脸所处的光照环境在方向上是一致的，而拼接图片中拼接人脸和原人脸所处的光照环境不同。这种方法具有对低质量图片鲁棒的优点，但更加精确地从单张图片估计人脸所处的光照环境则是提升拼接检测算法有效性的关键途径。实验室智能感知与计算研究中心王伟博士等人提出了更加准确和一般化的三维光照反射模型，在重建三维人脸模型的基础上，考虑了真实人脸所具有的几何非凸性（自我遮挡、阴影）以及具有纹理的特性，并使用球形谐波函数表示光照环境以及一般化的光照转移函数。实验表明这一方法在多个数据集上均有较大的性能提升，证明了所提出的一般化模型具有更好的光照估计准确度。相关研究成果发表在信息取证与安全领域国际期刊 IEEE Transactions on Information Forensics and Security。

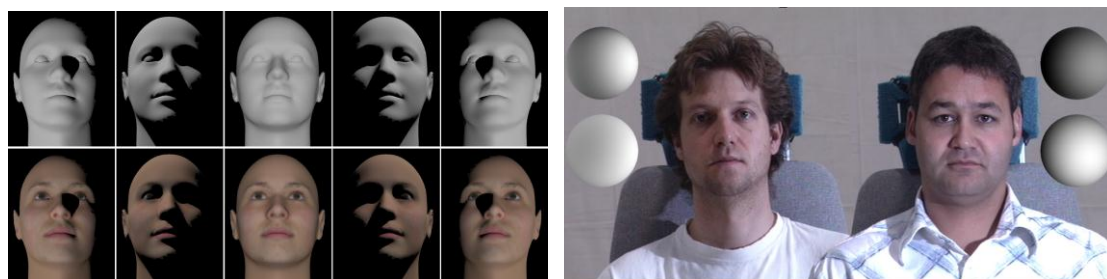


图 1 左图为拟合光照转移系数所使用的渲染图片，右图为人脸拼接检测示例图

基于 RGB-D 数据的 One-shot Learning 手势识别

生物特征识别与安全技术课题组万军博士等人在前期系列工作(JMLR 2013, TIP 2014, JEI 2014)的基础上,经过系统回顾基于 RGB-D 数据的 one-shot learning 手势识别研究方法以及对该领域的一些展望,提出了一种如何在有限的 RGB-D 数据构成的训练样本(每类手势只有一个训练样本)中挖掘更有效、更丰富的时空特征算子-mixed features around sparse keypoints (MFSK)。该算子具有旋转和尺度不变性。在词袋模型框架下测试,这一算法比当前主流的特征性能更好,在手势识别中取得了当前最好的效果。相关研究成果发表发表在国际期刊 IEEE Transactions on Pattern Analysis and Machine Intelligence。

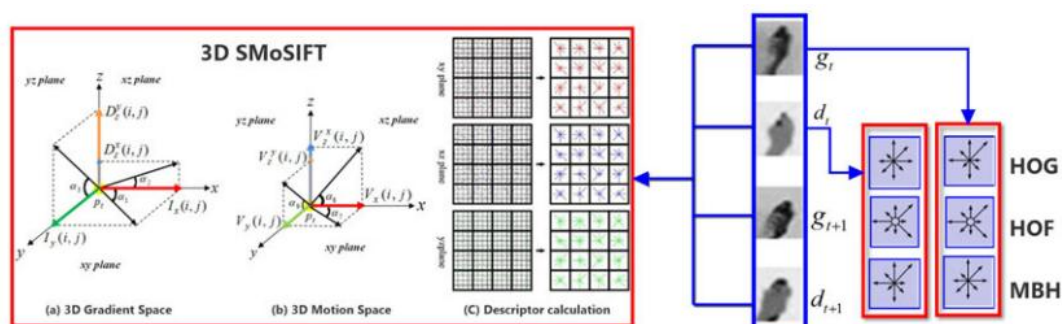


图 2 MFSK 算子: 在构建的金字塔中, 对检测到的关键点附近区域构建三维梯度和运动空间, 分别计算 3D SMoSIFT, HOG, HOF, MBH 算子, 这些算子归一化后组成 MFSK 算子

基于无监督多层描述子学习的行人再辨识

行人再辨识是智能视频分析领域的一个重要研究方向, 为目标持续性跟踪、跨摄像机行人关联以及行人检索提供了重要技术手段。其中, 如何描述视频中的行人成为解决行人再辨识的根本性任务。现有工作通常基于颜色或者纹理特征等手工特征来描述不同的行人。但是, 由于监控场景中图像分辨率较低, 不同摄像机所拍摄的行人存在姿态、视角变化, 已有的人工设计特征缺乏足够的类内(相同行人)鲁棒性和类间(不同行人)判别性。为此, 生物特征识别与安全技术课题组杨阳博士等人提出了基于多层(像素层、块层和图像层)描述子的分数层融合策略以及一种新的加权线性编码方法, 能够有效且快速地从原始数据中学习多层描述子。该编码可以保持空间一致性, 提升了描述子的判别性。该工作基于一种简单的相似度量作为基准, 在多个公开数据库上取得了国际领先的结果。相关研究成果发表在人工智能领域国际会议 AAI 2017 (Oral)。

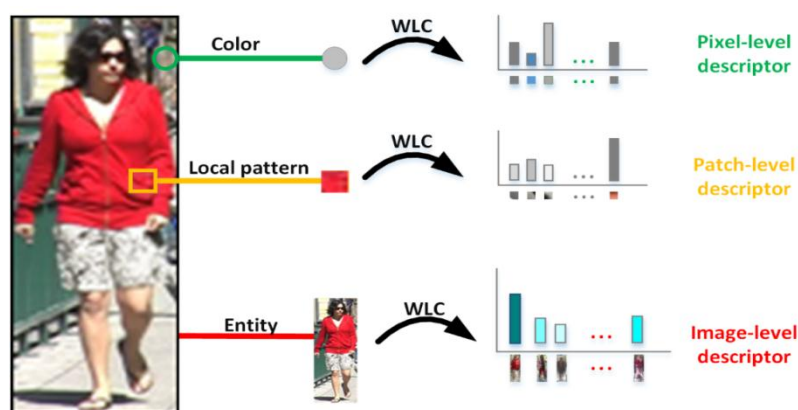


图 3 基于加权线性编码 (WLC) 的多层特征描述子学习

基于重要事件时序-权值分析图的体育新闻自动生成方法

每年的体育新闻超过千万篇, 其中大部分电视或者文字转播的体育比赛均会采用实时记录的方式保存下来, 但这些实时记录的文字非常琐碎, 并且含有很多错误。从这些文字中重新整理一篇体育新闻, 会耗费体育编辑大量精力。针对此

问题，人机交互课题组杨明浩副研究员等人提出了一种基于重要事件时序-权值分析图的体育新闻自动生成方法，通过构建关键词库，采用循环神经网络筛选出重要语句后，以时序为主轴排列重要事件时序及重要性图，结合关键句权值与事件热点之间的关系，最后，针对体育新闻报道的特点，结合语义分析，判断动作主体、动作意图，基于模板生成描述性关联语句，完成体育新闻的自动编写。该方法参加了 2016 年 12 月 2-6 日在昆明举行的国际自然语言处理与中文计算会议语义理解竞赛，在"NLPC2016 Competition Task 5: Sports News Generation from Live Webcast Scripts"竞赛单元综合排名第一，并在九项评价指标中获得六项第一。

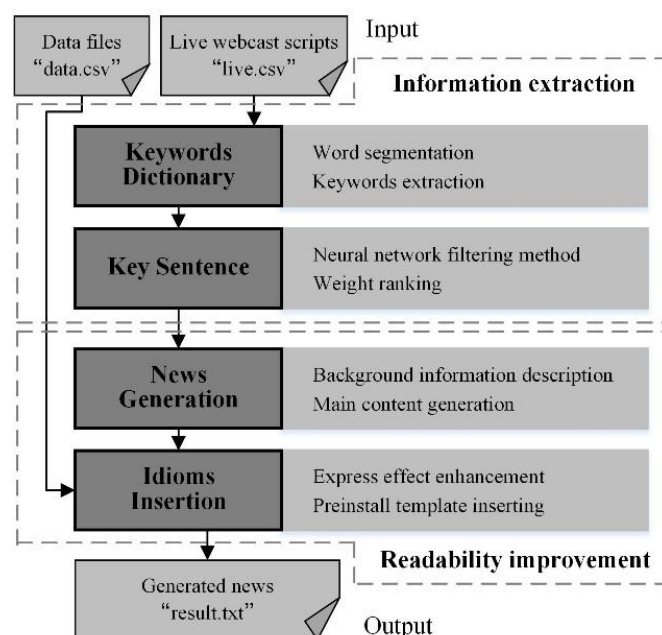


图 4 重要事件时序-权值分析图的体育新闻自动生成方法

基于温和正样本挖掘和深度度量学习的行人再辨识

监控视频中行人数据的一大问题是存在较大的类内差异，这主要是由于不同摄像头采集到的视频图像受到光照、人体姿态、遮挡等因素的影响较大。这些类内差异导致样本在特征空间中的分布呈现高度扭曲的流形。此时常规的欧式距离不适合作为度量来训练和测试算法模型，而且由于特征流形的分布是未知的，测地距离在此情形中也无法直接使用。

在流形学习一般方法的基础上，生物特征识别与安全技术课题组博士生石海林等人提出了温和正样本挖掘算法作为神经网络的训练策略。该方法采用局部的欧式距离作为度量，并且配合样本间的近邻关系，从而近似特征的测地距离，在训练神经网络时不仅减小类内差异而且保持数据的固有近邻关系结构。此外，该工作还提出将传统的马氏距离通过人工神经网络的方式建模与学习，并且采用特殊的正则项，增强了模型的推广性能。该研究有效地解决了较大类内差异在训练人工神经网络时带来的问题，显著提高了收敛效率和算法精度，在目前主流行人再辨识公开库上达到领先性能。相关研究成果发表在计算机视觉领域国际会议 ECCV 2016。

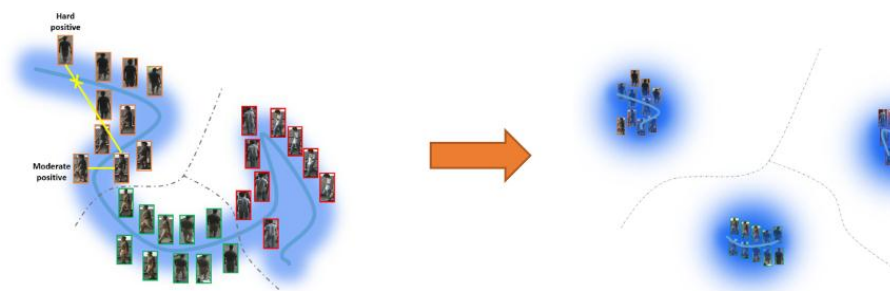


图 5 通过温和正样本挖掘，在减小类内差异的同时保持数据固有的内在结构

面向序列预测的神经网络深层结构设计

序列预测是自然语言处理中的核心任务之一。例如，汉语分词、词性标注、语义分析以及机器翻译等都属于序列标注问题。由于深度学习的发展，长短期循环神经网络（LSTM）已经成为解决序列标注的主流方法之一。但是网络结构的深度（Stacked LSTM）直接影响模型效果与训练效率。对此，学术界还没有给出令人满意的解决方案。

自然语言处理课题组宗成庆研究员等近年来一直探索如何针对自然语言处理设计更好的深层网络结构，近期从跨层的连接方式与连接函数两个角度分别提出了一系列深层 LSTM 的结构设计方案。在连接方式上不仅探索了跨层的模型设计，而且提出了三种自底往上的信息传递方式。这一方法在词性标注和语义标签（CCG Supertagging）预测任务中超越了其他主流方法，取得了目前最好的性能。相关研究成果发表在自然语言处理领域国际会议 COLING 2016 以及人工智能领域国际会议 AAAI 2017。

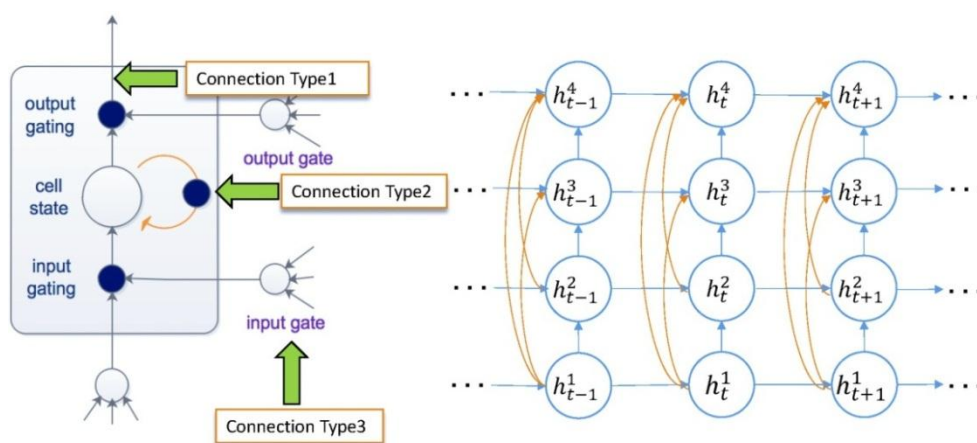


图 6 Stacked LSTM 跨层设计与信息传递方法

基于句子级关注机制和实体描述的弱监督关系抽取

关系抽取是构建结构化知识的重要途径，在自然语言处理的许多任务中具有重要意义，如知识图谱、智能问答系统等。相对于传统的关系抽取任务，弱监督

关系抽取不需要人工标注数据，它以知识图谱为监督源，其核心问题是剔除知识图谱回标错误产生的噪声以及准确判断给定实体对之间的关系类别。针对第一个问题，自然语言处理课题组博士生纪国良等人提出一种嵌入句子级关注机制的卷积神经网络模型，该模型首先使用分段卷积神经网络抽取句子特征，然后通过隐藏层网络计算特征与类别的相似度，选取有效的句子实例。该方法避免了对人工设计特征的过分依赖，并且能够将噪声数据的剔除融入到学习过程当中，在减弱噪声影响的同时，能够充分利用知识图谱的监督信息。针对第二个问题，课题组从 Freebase 中抽取实体的描述信息，补充实体的背景知识，为关系类别的判断提供更加丰富的依据。在公开的数据集上实验表明，该方法能够有效地提升弱监督关系抽取的性能，获得了较基线系统最佳的效果。相关研究成果发表在人工智能领域国际会议 AAAI 2017。

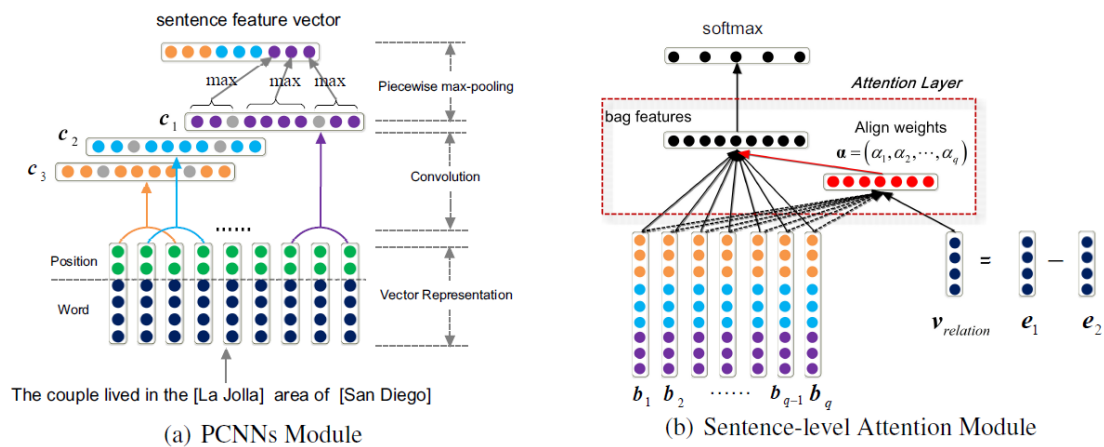


图 7 嵌入句子级关注机制的卷积神经网络结构图

基于层次化上下文多示例多标签学习的图像标注方法

基于多示例多标签学习的图像标注方法可以很好的建模图像区域与标签的相互联系，使标注结果更全面合理。研究证明，标签之间和图像区块之间均存在着上下文的关联关系，如果能够将这些上下文关系融入图像标注算法可以很好的改善标注的效果。为此，视频内容安全课题组李兵副研究员等人将这些上下文关联融入到学习过程中，提出了基于层次化上下文多示例多标签学习的图像标注算法。在该算法中，一个图像被看作一个包，图像内的区域看作包内的示例。首先，通过图结构来建模包内示例间的上下文关系；然后，通过多标签公用的隐藏标签层实现标签的上下文关系建模；最后，通过图核以及表示定理将包的图结构融合到基于标签上下文的多标签标注模型中，实现了层次化上下文结构的多示例多标签图像标注模型。在多个公开数据集上的测试实验验证了算法的有效性。相关研究成果发表在三多媒领域国际期刊 IEEE Transaction on Multimedia。

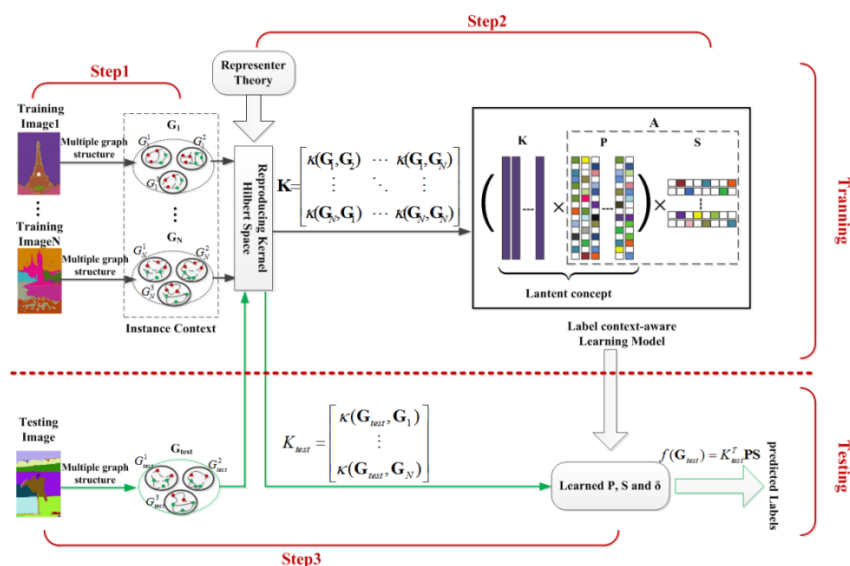


图 8 基于多层上下文结构的多示例多标签图像标注算法框图

基于主动结构光线的文档图像矫正

相机因其成像速度快、成像角度灵活、非接触性以及不受所拍摄文档尺寸限制等优点,近年来被广泛的应用于文档资料的数字化,并呈逐渐取代传统式平板扫描仪的趋势。然而,由于相机的透视效应以及文档页面的弯曲,拍摄的文档图像常常出现严重的畸变以及阴影不均等非线性失真现象。这些成像过程中引入的非线性质量退化,为后续的文档图像处理带来很多挑战。因此,如何利用相机获取高质量的文档图像成为研究者近年来关注的热点和难点问题。

针对这一问题,遥感图像处理课题组孟高峰副研究员等人提出了基于结构光线的主动式文档畸变矫正方法,并构造了文档畸变矫正原型系统。该系统由两个线激光器和一个高分辨率相机组成。线激光器将两条直线光线投射到需要拍摄的文档页面上。通过事先标定相机和激光线,页面的 3D 弯曲形状可以根据激光线的弯曲程度估计出来,进而可构造图像的矫正网格完成图像几何畸变的矫正。估计得到页面的 3D 曲面后,结合一个简单的郎伯光照模型,并利用凸包重构技术,可以提取图像的阴影曲面实现图像非均匀阴影的矫正。相关研究成果发表在计算机视觉领域国际会议 CVPR 2014 和国际期刊 International Journal of Computer Vision 2016。

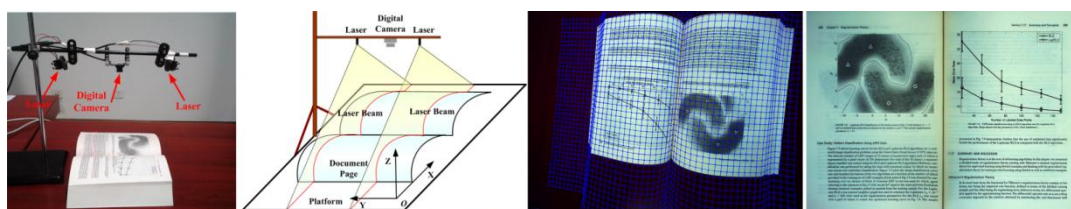


图 9 文档畸变矫正原型系统与文档图像矫正效果

动态并行及分布式图割算法

作为求解马尔科夫随机场能量函数优化问题重要的工具之一，图割算法（Graph cuts）的计算效率和大规模数据处理能力对于很多视觉问题的求解具有重要的理论和现实意义。图割算法并行化是提高计算效率的有效途径，但并行图割计算过程中可能出现不收敛的情况。为了解决这一问题，机器视觉课题组博士生余淼等人提出了一种并行及分布式图割算法。该方法可以在迭代过程中对相邻子图进行动态合并，并且能够有效地复用子图上的流，因而具有很高的计算效率。同时，这一并行及分布式图割算法是以基于对偶分解的子图分解法为基础，从理论上保证了算法的严格收敛性。这项研究工作构建的是一种通用的动态并行及分布式图割算法框架，在该框架的每一次迭代过程中，可进行任意的子图分解及合并，能够在并行图割的计算过程中动态地调整并行的粒度和负载的均衡。相关研究成果发表在国际期刊 IEEE Transactions on Image Processing。

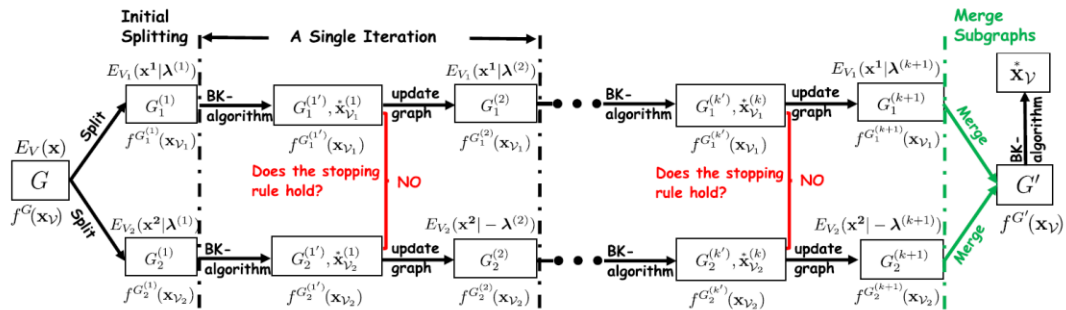


图 10 两个子图的并行 BK 合并流程

基于概率图模型的视频人物标识

对人物的识别和标注是对影视剧和新闻视频内容分析、检索与再编辑的关键环节。基于海量视频数据，在无人工标记样本的条件下，采用无监督的方式实现人脸的聚类以及人脸与人物名称的映射对齐是视频人物标识的基本任务。由于在无约束的真实场景中，人脸图像会受到姿态、穿戴、光照、运动等诸多因素的影响而降低其不同类之间的可分性，因此将人脸与从文本中自动抽取的人物名称标签相匹配也会存在较大的歧义性。针对这些问题，图像与视频分析课题组张一帆副研究员带领团队就该问题开展了深入研究，在概率图模型理论上提出了一整套高效的视频人物标识方法，主要包括以下几部分：1）人脸聚类：通过条件随机场来建立视频中人脸之间的近邻系统，利用视频中特有的互斥性和连续性生成 cannot-link 和 must-link 约束，并将约束条件在近邻系统中传播，提高约束矩阵的稠密性，然后通过 simulated field 算法实现人脸的聚类。2）视频文本对齐：从网络上我们可以方便的获取视频的脚本文本如剧本和新闻稿等，从文本中提取人物名称并使之与人脸在时间上对齐，可以实现对视频人物的标注。在标注过程中，课题组主要解决了在没有时间戳情况下的文本视频对齐问题。基于人脸聚类的结果，利用人脸人名在视频中的出现顺序和全局分布特性，在条件随机场中构造一阶和高阶匹配代价势函数，采用 Needleman-Wunsch 算法实现文本与视频的

对齐。3) 人脸聚类 and 文本视频对齐的联合优化：由于这两个问题的求解是相互依赖的，课题组将两个问题的随机场模型耦合在一起，通过迭代优化的策略交替学习和优化模型参数，达到协同学习，共同提升的效果。相关研究成果发表在多个计算机视觉和人工智能领域国际期刊和国际会议上，包括 IEEE T-IP, CVPR, IJCAI 等。

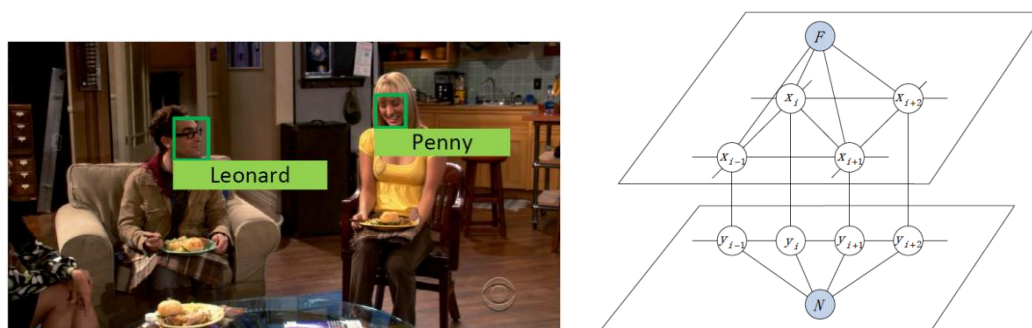


图 11 视频人物标识示意图和耦合条件随机场模型

学术交流

德国人工智能研究中心 Hans Uszkoreit 教授等来实验室访问交流

2016 年 11 月 23 日，德国人工智能研究中心语言技术实验室 Hans Uszkoreit 教授、Feiyu Xu 博士访问自动化所模式识别国家重点实验室，并做题为“Language Understanding cum Knowledge Yield”的学术报告。报告中，Uszkoreit 教授介绍了他们在关系提取领域的研究成果，特别是在他们的 Google 聚焦研究奖项目——LUcKY 中取得的成果。

Hans Uszkoreit 教授现任德国人工智能研究中心（German Research Center for Artificial Intelligence, DFKI）语言技术实验室科技总监、DFKI 柏林站的站点协调员以及 DFKI 的部门主任。Uszkoreit 教授现为欧洲科学院成员，同时担任计算语言学国际委员会的永久会员、柏林技术大学的荣誉教授。Feiyu Xu 博士现任 DFKI 语言技术实验室首席研究员和文本分析研究组主任。

南加利福尼亚大学的 C.-C. Jay Kuo 教授来实验室访问交流

2016 年 12 月 20 日，南加利福尼亚大学 C.-C. Jay Kuo 教授来实验室访问交流，并作了题为“Deep Learning Networks – Architectural Evolution and Theoretical Foundation”的学术报告。本次讲座聚焦于多层感知机、多层前馈网络、卷积神经网络，并且深入讨论了网络结构的演化和理论基础，洞察了 CNN 研究和应用的过去、现在、未来。

C.-C. Jay Kuo 教授现任南加州大学媒体通讯实验室主任。他的主要研究领域是数字媒体处理、压缩、通信和网络技术。Kuo 教授曾担任 IEEE Transactions on

Information Forensics and Security 主编、Journal of Visual Communication and Image Representation 主编。Kuo 教授是 AAAS、IEEE 和 SPIE 会士，曾获得 1993 年美国自然科学基金青年学者奖、2010 年电子成像科学家奖和 2011 年潘文渊杰出研究奖。

脑网络组研究中心邀请 8 位国内外专家来实验室访问交流

2016 年第四季度实验室脑网络组研究中心邀请了 8 人次的脑网络组系列报告，来自德国 Heinrich-Heine 大学的 Simon B. Eickhoff 教授、香港大学的 Ed X. Wu 教授、美国 Montana 州立大学的 Charles M. Gray 教授、荷兰 TU/e 大学的 Bart ter Haar Romeny 教授、英国 UCL 的 Stewart Shipp 教授、法国 CNRS 的 Henry Kennedy 教授、中科院心理所的王亮研究员先后访问了脑网络组研究中心并做专题报告，就各自的研究领域与相关研究人员进行了深入研讨，探讨了进一步合作方向。

其中，法国 CNRS 的 Henry Kennedy 教授于 11 月对脑网络组研究中心进行了为期 4 周的访问。在访问期间，他两次给脑网络组中心的研究人员进行专题讲座，介绍他基于逆向神经束路的跟踪技术在脑网络组研究中的应用，协助脑网络组研究中心搭建该平台，并结合中心高场磁共振技术共同开展脑网络组图谱研究。

项目立项

实验室第四季度新建立课题 25 项

实验室 2016 年第四季度新建立课题 25 项，总科研经费 1949.96 万元，部分项目如下：

项目名称	项目负责人	项目类型	经费 (万元)	执行期
精神分裂症系统水平表型组学研究	刘冰	科技部国家重点研发计划项目	90	2016/07/01 至 2019/06/01
大规模视频目标的跨时空多维度关联	廖胜才	北京航空航天大学	125	2016/07/01 至 2019/06/01
博士后国际交流计划资助项目	郑强	博士后基金会	30	2016/11/14 至 2018/11/01
基于场景感知的视频目标的结构化描述与识别	李子青	科技部重点研发计划项目	268	2016/07/01 至 2019/06/01
术中人脑功能区实时成像仪研发	张鑫	科技部数字诊疗装备研发专项	80	2016/07/01 至 2018/12/01
多语言处理管理平台及主流语言与民族语言翻译系统开发	宗成庆	晋城市国家安全局	130	2016/10/24 至 2019/10/01
深度学习与生物识别联合实验室	李子青	北京中科奥森数据科技有限公司	600	2016/09/30 至 2019/09/01

中科联想身份认证云服务联合实验室	孙哲南	联想（北京）有限公司	100	2015/12/01 至 2017/12/01
法庭视频图像分析识别算法研究（二期）	叶军涛	北京华夏电通科技有限公司	83	2016/11/21 至 2017/06/01
VR 头部位置跟踪	吴毅红	华为技术有限公司	150	2016/10/13 至 2017/12/01
虚拟试衣系统中的服装仿真算法与实现	叶军涛	深圳市云之梦科技有限公司	30	2016/09/01 至 2017/08/01
面向公安应用场景下的行人检测	张俊格	中科唯实科技（北京）有限公司	50	2016/10/10 至 2018/07/01
“优图研究”联合项目四期	董未名	腾讯科技（上海）有限公司	55	2016/08/21 至 2017/08/01