



# 中国科学院自动化研究所 模式识别国家重点实验室

## 简 报

2015 年第 4 期

总第 36 期

2015 年 10-12 月

### 目 录

最新动态 .....	1
第三届中国图像视频大数据产业创新论坛在自动化所召开 .....	1
第十届中国生物识别学术会议在滨海新区成功召开 .....	1
实验室团队在中国模糊图像处理竞赛中获得佳绩 .....	1
国家自然科学基金委重大研究计划“情感和记忆的神经环路基础”2015 年度总结会在自动化所召开 .....	2
科研进展 .....	2
基于灰度序的特征描述理论 .....	2
基于超图上下文核的行为识别 .....	3
基于联合特征选择和子空间学习的跨模态检索 .....	4
基于深度学习的跨表达域合作的语音分离方法 .....	4
语音驱动发音可视化获得更精确的发音器官轮廓驱动效果 .....	5
阿尔茨海默病影像数据库及其脑功能网络分析 .....	7
学术交流 .....	7
微软亚洲研究院华刚博士来实验室访问交流 .....	7
脑网络组中心邀请 10 位教授来实验室交流访问并进行系列学术讲座 .....	8
项目立项 .....	8
实验室第四季度新建立课题 30 项 .....	8

## 《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 杨明浩 张家俊 殷 飞  
吴 偶 廖胜才 孟高峰 张一帆 戴 佳 王爱华

## 最新动态

### 第三届中国图像视频大数据产业创新论坛在自动化所召开

2015 年 10 月 20 日，由图像视频大数据产业技术创新战略联盟主办，中科院自动化所模式识别国家重点实验室智能感知与计算研究中心承办，亮风台（上海）信息科技有限公司独家赞助的“第三届中国图像视频大数据产业创新论坛”在自动化所召开。论坛得到了社会各界的广泛关注，吸引了来自政府部门、行业企业、科研机构、高等院校等各界近 300 位与会者参会。

随着科技智能化的不断发展，类已经起步迈进技术智能化时代，智能产业将成为经济发展的新引擎。视觉技术是智能产业的一项重要内容，在很多领域，机器视觉不断改变着我们的生活，成为构建智慧城市过程中不可或缺的“第三只眼”。本次论坛以“智能+视觉”为主题，邀请了清华大学周杰教授、中科院计算所张勇东研究员、亮风台共同创始人、首席科学家凌海滨、Base FX 科技副总裁邓宇、阿里巴巴资深总监华先胜研究员作主题报告。他们分别介绍了自己所从事的智能视觉技术，并就在“互联网+”时代，如何加快我国智能科技产业的发展，强化视觉技术对智能产业的引领和支持，推动视觉技术不断发展等方面提出了许多独到见解。为了促进学术界与产业界的交流互动，本次论坛还特别邀请了英伟达、三星研究院、爱奇艺、富士通研发中心、北京视博云科技、浙江捷尚视觉科技、腾讯视频的专家围绕“视觉技术在智能中的作用和挑战”的话题进行了交流和探讨。

### 第十届中国生物识别学术会议在滨海新区成功召开

2015 年 11 月 13-15 日，第十届“中国生物识别学术大会”在天津滨海新区召开。此次大会由中科院自动化所、中国人工智能学会、中国民航大学和天津科技大学联合主办，天津中科智能识别产业技术研究院、中国民航大学、天津科技大学、中国生物识别产业技术创新战略联盟联合承办。大会邀请到中国科学院副院长谭铁牛院士、天津科技大学夏静波书记、中国民航大学吴仁彪副校长，以及滨海新区科委、开发区科技局、科技金融集团、国内著名高校及研究所 280 余名专家出席。生物识别是模式识别、图像处理、人工智能等学科领域的前沿方向，是保障国家和社会公共安全的战略高新技术，同时也是电子信息产业新的很有潜力的增长点。中国生物识别学术会议于 2000 年创办，先后在北京、杭州、西安、广州、山东、沈阳等地成功举办过九届。大会自创办以来有力推动了我国生物识别的学科发展和应用推广，同时为国内生物识别学术界和产业界同行提供了一个交流与合作的平台。

### 实验室团队在中国模糊图像处理竞赛中获得佳绩

2015 年 11 月 8-9 日，由国家自然科学基金委员会信息科学部、重大研究计划“视听觉信息的认知计算”指导专家组和中国自动化学会主办的 2015 年中国模糊图像处理竞赛决赛暨学术交流会在西安举行。由中科院自动化所模式识别国家重点实验室机器视觉课题组吴福朝研究员领队，罗亨亮、杨怡、童贝三位研究生组建的“we-rv”队在“交通标志图像识别挑战赛”中获得了第一名的成绩。

本次比赛针对无人驾驶智能车辆在实际应用中面临的因雾、霾、沙尘、雨雪等气候因素以及车辆运动所导致的图像质量劣化问题,寻求适于该问题的图像恢复、增强与内容分析新方法与新技术。该竞赛分为“交通场景图像增强挑战赛”、“交通标志图像识别挑战赛”和“双目立体视觉环境深度感知挑战赛”等三个项目。由来自国内 30 家高校、科研院所和公司的 52 支队伍参加了网络的初赛,经过角逐,15 支参赛队参加了三项挑战赛的现场决赛。其中有 6 支队伍参与了“交通标志图像识别挑战赛”的决赛。模式识别国家重点实验室“we-rv”队参与的“交通标志图像识别挑战赛”要求对车载相机图像序列数据进行处理与分析,完成交通场景中警告、禁令、指路等 7 个大类、72 种交通标识的检测与识别。该队使用了基于最大稳定极值区域的交通标志候选区域提取和基于卷积神经网络的交通标志识别的解决框架,在真实交通场景的比赛数据中获得了较为精确、鲁棒的检测与识别结果。

## 国家自然科学基金委重大研究计划“情感和记忆的神经环路基础”2015 年度总结会在自动化所召开

国家自然科学基金委重大研究计划“情感和记忆的神经环路基础”2015 年度总结会、国际研讨会暨“脑网络组图谱及其应用研究培训班”于 2015 年 11 月 10-14 日在北京召开。本次会议由国家自然科学基金委主办,中国神经科学学会、浙江大学神经科学研究所协办,自动化所模式识别国家重点实验室脑网络组研究中心和脑网络组北京市重点实验室承办。本次大会主席为浙江大学段树民院士,大会执行主席为模式识别国家重点实验室蒋田仔研究员和浙江大学李晓明教授。

项目总结会上,参与“情感和记忆的神经环路基础”重大研究计划的 25 名课题负责人总结了在过去一年在情感和记忆的分子机制、动物模型及神经环路研究的新技术等方面的进展。部分成果获得了国际同行的认可,论文发表在 Nature 及系列子刊等学术期刊,并且相关的技术产品在同行中得到了应用和技术转让。另外,在项目总结会议前后分别进行了特邀报告和国际研讨会,10 位国际知名的科学家介绍了各自在脑网络研究及相关神经精神疾病的机制研究方面的最新进展。

在“重大研究计划年度总结会”和“国际研讨会”召开前,中国科学院自动化研究所组织了“脑网络组图谱及其应用研究培训班”,由国内具有相关研究经验的专家分别针对脑网络组图谱绘制和神经环路解析需要的技术和方法,包括基于多模态成像的脑区划分及遗传基础、脑电与近红外融合技术、电生理技术、微观成像和三维重建平台、透明脑方法等方面提供系统全面的培训。

## 科研进展

---

### 基于灰度序的特征描述理论

局部图像特征描述一直以来都是计算机视觉领域研究人员非常关注的问题,为了更好地描述局部图像的本质信息、抑制不必要的噪声干扰,同时增强特征描

述子对于不同的图像几何与光度学变换的鲁棒性, 实验室樊彬副研究员和吴福朝研究员等人提出了基于灰度序的特征描述理论, 通过充分挖掘每个像素的邻居点之间的灰度序关系以及局部图像区域内所有像素点的灰度序关系, 提出了基于灰度序特征汇聚的特征描述框架、局部灰度置换群特征以及整体灰度序特征等, 产生了一系列创新性成果, 相关论文相继发表于 ICCV、ECCV、CVPR、IEEE Transactions on Pattern Analysis and Machine Intelligence 等。相关算法的源代码已被 vlfeat 集成收录, 在计算机视觉领域获得了较为广泛的应用。同时, 为了进一步帮助国内外研究者了解图像特征描述领域的相关理论和方法, 樊彬副研究员和吴福朝研究员在他们取得的一系列学术成果的基础上, 对图像特征描述领域进行了系统性总结, 出版了英文学术专著《Local Image Descriptor: Modern Approaches》, 该书涵盖了经典的图像特征描述方法以及该领域的最新技术, 包括基于灰度序的特征描述方法和二进制特征描述方法, 瑞士洛桑联邦理工大学 Pascal Fua 教授和芬兰奥卢大学的 Matti Pietikainen 教授为此书作了序。

## 基于超图上下文核的行为识别

视频序列中人体行为识别的性能主要依赖于两点: 行为的表示和这些表示之间的相似性度量。视频内容安全课题组原春锋博士提出基于一种上下文核来融合两种互补的时空兴趣点特征进行行为识别。对于行为表示, 基于时空兴趣点的词包模型是一种非常有效的方法。然而, 这种方法忽略了兴趣点的时空分布信息。该课题组提出对时空兴趣点形成的 3D 空间进行 R 变换提取一种新的全局特征来描述兴趣点的时空分布。对于相似度度量, 课题组提出将一个视频集建模成一个优化的概率超图, 并构建了一种上下文核来度量视频间的高价关系。克服了以往算法中计算相似度核矩阵时都是基于两两视频之间的相似度(或距离)建立的, 对于噪声比较敏感的问题。并且, 课题组还提出了一种新的概率超图模型, 采用度量学习方法建立超边, 使得上下文数据同类, 避免引入其他类的干扰, 提高上下文核的判别性。多个网络视频库上的测试验证了该算法的有效性, 相关工作已被 International Journal of Computer Vision 接收。

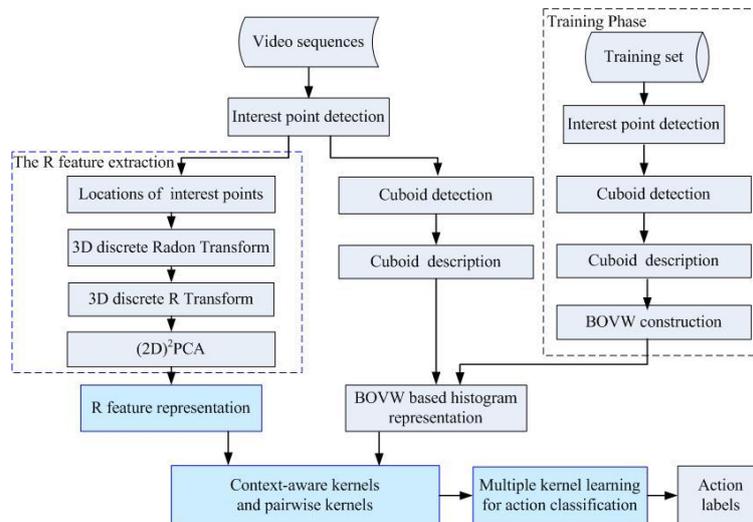


图 1 基于超图上下文核的行为识别算法的流程图

## 基于联合特征选择和子空间学习的跨模态检索

随着网络中不同类型数据的爆炸式增长（如文本、图片、视频等），如何对这些不同类型的数据进行“跨模态检索”以便更加有效地利用这些数据成为亟待解决的问题。跨模态检索的难点是如何度量不同模态数据之间的相似度，同时由于数据的底层特征一般都存在冗余和不相关的特征，所以如何在不同模态数据上同时进行特征选择（耦合特征选择）也是一个非常重要的问题。为了解决这些问题，智能感知与计算研究中心王开业博士、赫然研究员、王亮研究员等提出了一种联合学习方法，同时进行子空间学习和耦合特征选择。该方法的目标函数有三项构成，第一项是耦合线性回归项，目的是学习从不同模态数据的特征空间到共同空间的映射；第二项是 L21 范数项，目的是选择出不同模态数据中那些相关的和具有区分力的特征；第三项是多模态图规则项，用于保持不同模态数据之间的两种相似性关系，即模态间数据的相似性关系和模态内数据的相似性关系。为了求解这一目标函数，我们提出了一种迭代求解算法，同时证明了算法的收敛性。三个跨模态数据库上的实验结果表明了该算法的有效性。相关工作被已被 *IEEE Transactions on Pattern Analysis and Machine Intelligence* 接收。

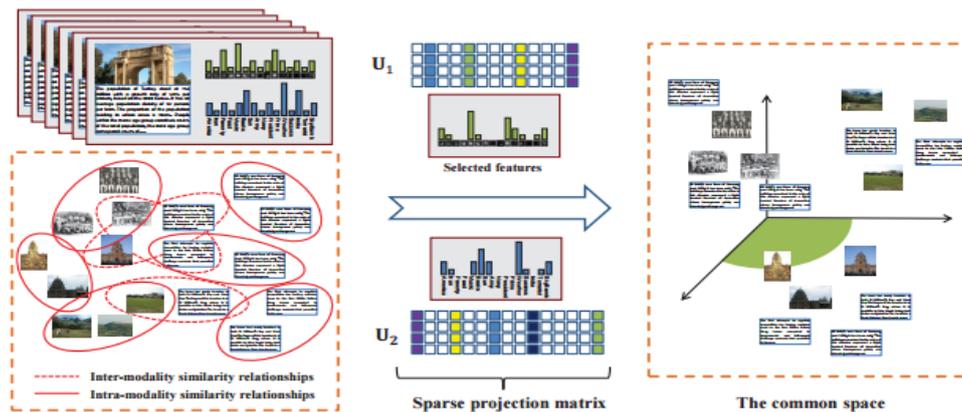


图 2 联合特征选择和子空间学习框架

## 基于深度学习的跨表达域合作的语音分离方法

近年来，基于深度学习的听觉场景分析的语音分离消噪方法受到该领域的特别关注，其核心问题是能够利用深度神经网络学习语音频谱的结构细节，对此问题刘文举研究员等人提出了基于深度学习的跨表达域合作和多分辨率合作的语音分离方法的框架，首先将同一语音信号在不同时频表达域的特征进行融合，其原理在于不同的时频表达方法可以反映同一语音信号的不同侧面，并提供互补信息，从而使得语音分离算法的性能得以提升；更进一步还可表明，即使在同一时频表达域，如果将不同分辨率尺度下的特征进行融合，也能提高语音分离算法的性能。该项成果的论文“Multiscale collaborative speech denoising based on deep stacking network”获 2015 年 IJCNN 国际学术会议墙报比赛荣誉提名奖“Honorable Mention Poster Competition.”

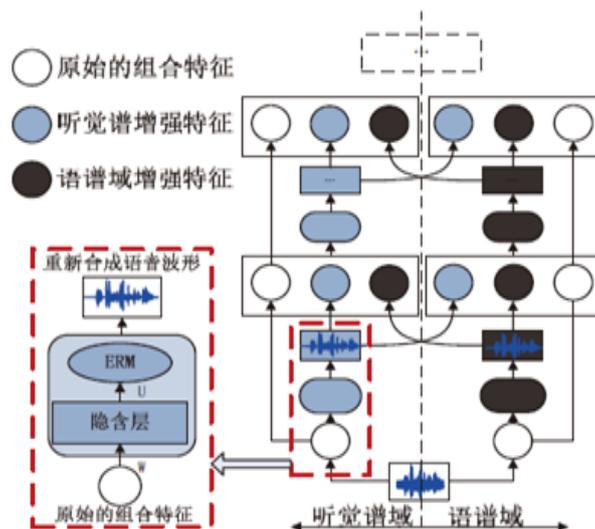


图 3 跨表达域的合作式深度堆叠网络框架

### 语音驱动发音可视化获得更精确的发音器官轮廓驱动效果

从语音信号直接估计发音器官的 MRI 图像对于进一步了解人类发音机理以及病理语音分析与评估都有重要的意义。尽管 MRI 是很好的发音器官观测工具，但 MRI 图像存在大量噪点，当舌位与其他发音器官（如上颚、软腭、喉部等）发生接触时，其轮廓边缘变得十分模糊甚至消失，这给舌位轮廓跟踪带来了很大困难。根据 MRI 成像特点，人机语音交互课题组杨明浩副研究员和陶建华研究员等人提出一种较为鲁棒的舌位轮廓自动提取方法。该方法包括：对于核磁图像序列，在舌位运动区域内，利用多方向 Sobel 算子得到舌位轮廓初始边缘点；建立舌位边缘点映射矩阵，并结合前一帧舌位轮廓位置，对该映射矩阵进行调整；在调整后的映射矩阵中寻找最优边缘点序列，借助过控制点的二次样条曲线拟合技术得到舌位轮廓。该方法能够自动从核磁图像序列中较准确地提取出舌位轮廓，其优势在于当舌位与其他发音器官发生接触时，该方法也具有很好的鲁棒性。此外，课题组还提出了一种由语音信号对 MRI 图像序列进行直接估计的发音器官运动图映射方法，采用了基于多层神经网络的线性回归模型从语音信号中估计发音器官运动图像序列方法，并比较了不同深度结构来对深度神经网络进行预训练的影响。研究的系列成果连续多年发表在 ICASSP 会议上，并在 2015 年全国人机语音学术会议的语音多模态信息获取与处理专题上做了特邀报告

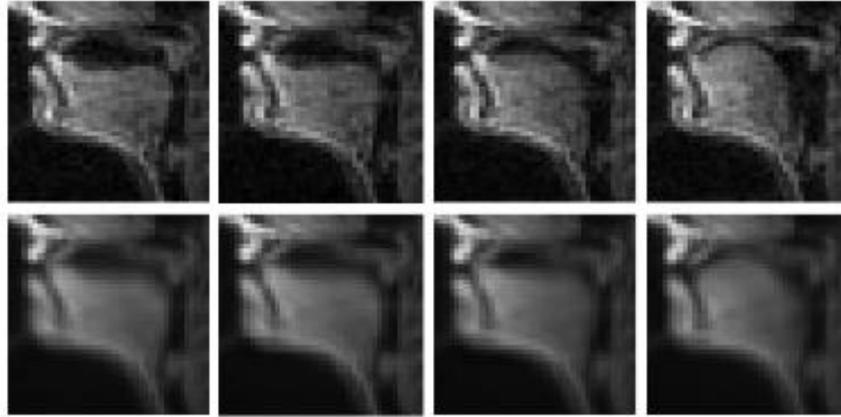


图 6 基于多层神经网络的语音信号到舌位运动轮廓映射模型（上面一行是原图像，下面一行从语音信号预测得到的图像）

### 基于显著性目标候选集的多通道相关滤波跟踪

有效的特征和模型漂移问题是目标跟踪问题的两个重要方面。对于特征表达来说，梯度和颜色特征被广泛应用，但是如何融合它们并有效地应用到跟踪问题仍然是一个难题。图像视频分析与处理课题组王金桥副研究员提出了一个新的特征描述子：**Multiple Channel HOG (MC-HOG)** 特征。该描述子融合多颜色通道和空间梯度信息。课题组将其嵌入到相关滤波跟踪框架中来预测目标的状态。为了处理由于遮挡或干扰项所带来的漂移问题，课题组提出将显著性信息作为先验知识选择目标候选样本，从而降低了背景的干扰。另外，课题组还提出了一种重排序策略来判断候选项的重要程度，该策略通过学习到的外观滤波器、历史保留的目标样本和干扰提案项作为模板进行判断从而可以有效地缓解模型漂移问题。相关研究成果发表在 **BMVC 2015** 和 **AAAI 2016**。



图 7 基于显著性目标候选集的多通道相关滤波跟踪系统框图

## 阿尔茨海默病影像数据库及其脑功能网络分析

模式识别国家重点实验室脑网络组研究中心建立了阿尔茨海默病及轻度认知障碍的多中心、多模态影像学数据库：该数据库包包括阿尔茨海默病（AD）、轻度认知障碍（aMCI）和正常对照数据在内 700 余人次的多模态磁共振影像数据。主要合作单位包括：中国人民解放军总医院，宣武医院，东南大学附属中大医院，天津环湖医院，山东大学附属齐鲁医院。

基于静息态功能磁共振影像，脑网络组研究中心建立了一个脑功能连接分析的方法体系(包括单个脑区的功能连接分析、局部脑网络分析和全脑网络分析)，并推广到了阿尔茨海默病和轻度认知障碍的研究中，获得了对其功能连接和功能网络的异常模式的新发现。近期的主要发现有：1) 皮层下核团海马旁回功能连接在阿尔茨海默病和轻度认知障碍的异常连接模式提示提示阿尔茨海默病是一种“功能失连接综合征”。2) 通过分析阿尔茨海默病网络内及网络间功能连接损害特点，发现了阿尔茨海默病患者认知相关网络内及网络间功能连接均减低，这种减低与患者认知功能下降直接相关；从网络的角度提示阿尔茨海默病的脑功能网络的全局组织结构紊乱和信息传递效率下降，且异常的功能连接和异常的网络表征和患者的认知能力显著地相关，此项研究进一步加深了人们对阿尔茨海默病病理机制的理解和认识。研究结果已发表在领域 *Scientific Reports*, *Journal of Alzheimer's Disease* 等专业期刊上。

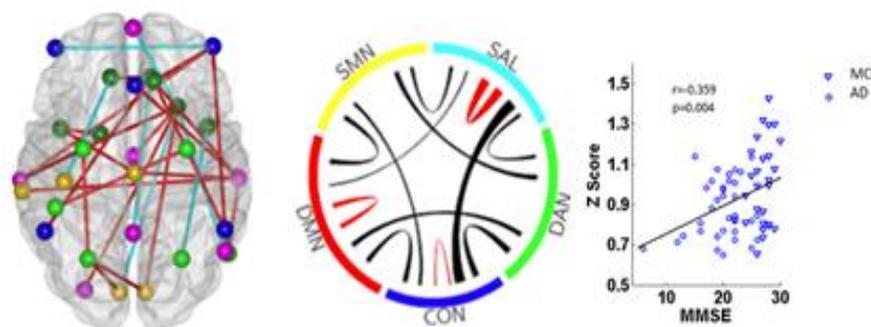


图 8 阿尔茨海默病及轻度认知障碍对大脑不同神经环路的影响

## 学术交流

### 微软亚洲研究院华刚博士来实验室访问交流

2015 年 11 月 12 日，微软亚洲研究院的高级主管研究员华刚博士来实验室访问，并做客自动化所“模式识别学术大讲堂”，为研究所师生作题为 *Probabilistic Elastic Part Model: A Pose-Invariant Representation for Face Recognition* 学术报告。

报告中，华刚博士首先介绍了在非可控情形下，头部姿态变化对人脸识别精度影响有很大的影响。而一个普遍认可的观点是，基于部件的人脸的表达对头部姿态的变化具有很好的鲁棒性。在以往的做法中，是首先进行人脸关键点定位，

然后手工提取人脸中各个部件。他的工作是将这种手工的方式转为一种概率化配准的部件模型，称为概率弹性部件模型（probabilistic elastic part (PEP) model）。这种模型能从姿态各异的人脸的局部表达中，拟合出一个具有空间表现的高斯混合模型。不论是对一张人脸还是一系列的人脸，该模型的每一个分量能选取具有最高概率被选中的局部特征。将这些所有被选中的局部特征组合在一起，就构成了具有头部姿态不变性的人脸表达，称为 PEP 表达。他将这种表达应用到人脸验证和人脸识别，效果较大幅度的优于基于手工选取部件的方法。

华刚博士在美国西北大学获得电气与计算机工程博士学位，先后在 IBM、诺基亚、微软等公司担任研究员，并在 2011 年加入了史蒂文斯理工学院，任计算机专业副教授。2015 年，华刚博士加入微软亚洲研究院视觉计算组，现任微软亚洲研究院(MSRA)的高级主管研究员，曾获得 2015 IAPR Young Biometrics Investigator Award。

## 脑网络组中心邀请 10 位教授来实验室交流访问并进行系列学术讲座

2015 年第四季度，脑网络组中心邀请了 10 位国际学者来实验室交流访问并进行系列学术讲座。他们分别是芬兰埃因霍温理工大学教授 Bart ter Haar Romeny 教授，台湾国立清华大学江安世教授，中国科学院武汉病毒研究所罗敏华教授，中国科学院上海神经所许晓鸿研究员，英国伦敦国王学院 Gunter Schumann 教授，德国马普生物控制论研究所 Xin Yu 教授，澳大利亚神经科学研究所教授 George Paxinos 教授，澳大利亚莫纳什大学教授 James Bourne 教授，中国科学院生物物理研究所赫荣乔研究员，南京航空航天大学的张道强教授。来访的国内外学者从微观到宏观层面系统的介绍了他们在各自研究领域的最新成果，呈现了脑网络组研究广阔的前景和充满希望的一面，给大家带来了一系列有启发性的观点。各位国内外学者同时和脑网组研究中心就数据共享，交流合作和项目申请等各个方面进行了深入交流，探讨如何共同推进脑网络组研究和脑网络组图谱的进一步应用等问题。

## 项目立项

### 实验室第四季度新建立课题 30 项

实验室 2015 年第四季度新建立课题 30 项，总科研经费 1319 万元：

项目名称	项目负责人	项目类型	经费 (万元)	执行期
类脑神经网络自主学习	刘成林	中科院战略性科技先导专项 B 类	50	2015-7-01 至 2015-12-31
视觉计算神经网络模型与方法	王亮	中科院战略性科技先导专项 B 类	150	2015-7-01 至 2015-12-31
语言认知与跨语言理解	宗成庆	中科院战略性科技先导专项 B 类	32	2015-7-01 至 2015-12-31

多模态人机协同与智能生长计算模型	陶建华	中科院战略性科技先导专项 B 类	32	2015-7-01 至 2015-12-31
基于神经生理的立体视觉模型与方法	胡占义	中科院战略性科技先导专项 B 类	29	2015-7-01 至 2015-12-31
视频时空信息的类人感知与理解	胡卫明	中科院战略性科技先导专项 B 类	29	2015-7-01 至 2015-12-31
感知协同的人机交互模式	孙哲南	中科院战略性科技先导专项 B 类	29	2015-7-01 至 2015-12-31
介观脑网络动态活动记录与分析	余山	中科院战略性科技先导专项 B 类	33	2015-7-01 至 2015-12-31
脑功能神经影像大数据挖掘	隋婧	中科院战略性科技先导专项 B 类	29	2015-7-01 至 2015-12-31
人工智能发展战略研究	谭铁牛	院学部咨询评议项目	98	2015-10-01 至 2016-10-31
面向基础教育的知识关联与推理类问题求解关键技术及系统	刘康	863 子课题	135	2015-01-01 至 2017-12-31
脑神经功能重塑及临床应用关键技术研究	吴毅红	863 子课题	75	2015-01-01 至 2017-12-31
图像视频大数据产业技术创新战略联盟	王亮	北京市产业技术创新战略联盟促进科技专项	15	2015-07-13 至 2016-07-13
多媒体大数据分析 & 展示	徐常胜	自然科学基金联合基金项目	150	2015-01-01 至 2018-12-31
基于视觉注意机制的迷彩伪装评价模型研究	张天柱	基金面上子项	20	2015-01-01 至 2018-12-31
多模态体感行为分析、理解与评估	杨明浩	基金重点子项	20	2014-01-01 至 2018-12-31
体感互动中的视觉理解关键技术研究	张晓鹏	基金重点子项	45	2014-01-01 至 2018-12-31
虹膜及眼周识别技术开发和咨询	李海青	北京中科虹霸科技有限公司	20	2015-10-28 至 2017-10-28
移动互联网中的敏感彩信识别	李兵	北京交通大学	20	2014-12-01 至 2017-12-31
在线健康问诊大数据聚类算法研究	罗冠	北京好啦科技有限公司	10	2015-11-30 至 2018-12-31
智能视频图像分析	张俊格	中科唯实科技(湖南)有限公司	30	2015-09-05 至 2017-03-05
面向大规模开放域知识库的深度问答关键技术研究	刘康	中国计算机学会	10	2015-10-01 至 2016-12-31
环场拍摄影像深度自动生成系统二期	高伟	北京文博远大数字技术有限公司	27	2015-11-20 至 2016-06-20
特效模拟与互动漫游	张晓鹏	中国科学院文献情报中心	25	2015-10-30 至 2016-10-29

基于人脸识别的关键应用程序保护	董未名	中国科学院软件研究所	25	2015-04-01 至 2015-10-31
体育赛事中运动员位置跟踪	孟维亮	泽普互动(天津)科技有限公司	25	2015-11-03 至 2016-10-31
三维模型计算分析合作合同	吴怀宇	中国科学院大学	6	2015-11-17 至 2016-12-31
大范围航拍图像分析研究及其系统	赵鑫	湖南中航天目测控技术有限公司	80	2015-09-18 至 2016-09-18
手机部件生产线上工业缺陷智能视觉检测	陈晓棠	中科金睛视觉科技(北京)有限公司	20	2015-10-19 至 2016-10-19
ISMP 算子集开发项目	赵军	长安通信科技有限责任公司	50	2015-02-01 至 2018-02-01