



中国科学院自动化研究所 模式识别国家重点实验室

简 报

2017 年第 1 期

总第 41 期

2017 年 1-3 月

目 录

最新动态	1
模式识别国家重点实验室 2016 年度工作会议顺利召开	1
宗成庆研究员当选中国中文信息学会副理事长和亚洲自然语言处理联合会副主席 ...	2
科研进展	2
基于长短时记忆递归神经网络的鲁棒基频提取	2
差异一致性的多视角子空间聚类	3
基于整体和自适应部件特征学习的行人再辨识	4
感知语义实例的图像文本匹配	4
面向机器翻译的神经融合模型	5
面向冷启动问题的垃圾评论检测	6
基于有监督关注机制融合角色信息的事件识别方法	6
学术交流	7
乔治亚理工学院 B.H. Juang 教授来实验室访问交流	7
清华大学王小勤教授来实验室访问交流	7
项目立项	8
实验室第一季度新建立课题 38 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 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

模式识别国家重点实验室 2016 年度工作会议顺利召开

2017 年 1 月 13 至 14 日，模式识别国家重点实验室 2016 年度工作会议在中科院自动化所智能化大厦顺利召开。中科院自动化所徐波所长、实验室学术委员会委员、技术委员会委员及实验室全体成员参加了本次会议。

1 月 13 日上午，学术委员会委员，北京大学查红彬教授和中科院声学所颜永红研究员，分别做了题为“基于三维数据流融合的场景重建与传感器定位技术”和“以语音为核心的人机交互机遇和挑战”的学术报告。1 月 13 日下午，实验室召开年度工作会议正式开始，徐波所长首先致辞，所长首先肯定了模式识别国家重点实验室成立 30 年来取得了很多成果，然后强调了人工智能发展新高潮带来很多机遇和新的挑战，最后针对国家重点实验室新一轮评估更加重视学术水平和国际影响的情况，希望各位学术委员和技术委员给实验室提出更多的宝贵意见。随后，实验室主任刘成林研究员做 2016 年度实验室工作报告。报告主要介绍了实验室在人才队伍建设、承担科研任务情况、科研产出与代表性成果、运行管理与开放交流、实验室文化建设等方面的主要进展。之后，实验室副主任陶建华研究员陪同技术委员会委员们参观了实验室的代表性工作演示。同时，学术委员会主任谭铁牛院士主持召开了学术委员会会议。委员们听取了 2017 年开放课题的申请情况、2016 年在研开放课题的进展情况以及 2015 年开放课题的结题情况汇报。会议对新申请课题的评审意见进行了综合评定，最终评选出 2017 年拟资助的 19 个课题。



1 月 14 日上午，实验室多个工作小组分别对本年度在基础研究、学术交流、宣传与成果演示、研究生教育、网站与平台建设等方面的工作做了详细汇报。随后，全体实验室员工分三个小组展开讨论，对实验室评估工作的准备、学科建设、人才队伍建设、资源优化使用等问题进行了分组讨论。在工作年会闭幕式上，刘成林主任再次感谢实验室全体成员一年来的辛勤努力，感谢大家对实验室发展提出的宝贵建议，并对获得 2016 年实验室优秀员工的向世明、王金桥、李兵、董未名、张煦尧五位同志进行了表彰。

宗成庆研究员当选中国中文信息学会副理事长和亚洲自然语言处理联合会副主席

2016年12月23-24日，中国中文信息学会第八次全国会员代表大会暨学会成立35周年学术会议在北京中国科技馆举行，并进行了学会的换届选举工作。中科院自动化所模式识别国家重点实验室的宗成庆研究员当选为第八届理事会副理事长。另外，在刚刚结束的亚洲自然语言处理联合会（Asian Federation of Natural Language Processing, AFNLP）换届选举中，宗成庆研究员当选新一届AFNLP副主席。

中国中文信息学会成立于1981年6月，由钱伟长、甄健民、安其春等学者发起，是中国科协直接领导的一级学会。学会的主要学术研究内容是利用计算机对中文（包括汉语和少数民族语言）的音、形、义等语言文字信息进行的加工和操作，包括对字、词、短语、句、篇章的输入、输出、识别、转换、压缩、存储、检索、分析、理解和生成等各方面的处理技术。近年来，中文信息处理技术在语言文字信息处理、网络内容安全和个性化信息服务等各方面都发挥了重要作用，中文信息学会在国内外同行的共同努力下逐渐壮大，学会影响力不断攀升，在本学术领域享有崇高的声誉。

亚洲自然语言处理联合会（AFNLP）于2000年发起成立，是亚洲自然语言处理领域权威的民间学术组织，致力于加强亚洲地区自然语言处理和计算语言学相关领域的学术研究与合作，促进信息和资源共享，推动亚太地区本领域与国际学术界的交流与发展。AFNLP主席、副主席和秘书长等领导任期为两年，主席任期到届后副主席自动升任主席。在2015至2016年期间，宗成庆研究员担任该联合会的秘书长，他的副主席任期为2017至2019年。

科研进展

基于长短时记忆递归神经网络的鲁棒基频提取

基频是语音信号中最为重要的语音参数特征参数，基频提取的准确性直接影响到语音分离和语音合成的性能；当前噪声环境下的基频提取仍然存在着挑战，尤其是在非平稳噪声和低信噪比环境下。针对这一问题，人机语音交互课题组刘斌博士等人提出了一种联合训练特征增强模型和基频估计模型的方法提高噪声环境下基频估计的精度。这一方法的优点为：采用特征增强神经网络，可以抑制不同环境下的噪声信息，从而缩小了基频估计神经网络输入特征的动态范围；将特征增强神经网络和基频估计神经网络联合训练可以使得特征增强模块以基频重构误差为目标函数，从而可以获得更为精准的基频值的预测；通过长短时记忆递归神经网络可以捕获到整个序列的信息检测基频。实验结果表明这一方法在不同噪声环境下优于主流的语音基频提取方法，可以适用于不同说话人、不同语音、不同噪声环境下的语音基频提取，为基于语音的情感分析、韵律分析等技术研发奠定了非常好的基础。相关研究成果发表在 ICASSP 2017。

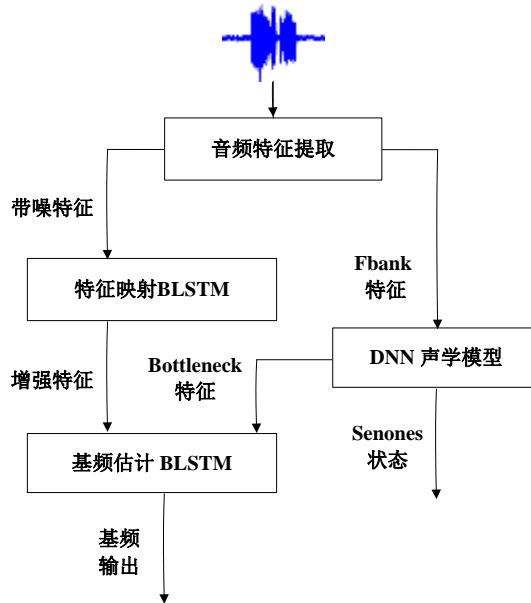


图 1 基于长短时记忆递归神经网络的鲁棒基频提取方法

差异一致性的多视角子空间聚类

评价视角间的多样性是多视角子空间聚类任务中重要的组成部分。目前已有不少的评价准则，在实际多视角聚类问题中也取得了合理的性能。但是现有的方法大多是数值感知的，也就是视角间的多样性与每个视角相似图的数值是相关的。而实际问题中，相似图的块对角性质虽然比较类似，但是相似图的数值可能相差显著，导致现有的多样性评价标准达不到预期的结果。为此，生物特征识别与安全技术课题组博士生王晓波与李子青研究员等人提出了一种新颖的多样性评价准则。相较于之前数值感知的多样性准则，他们所提出的新准则是位置感知的，与相似图之间的数值无关，较好地解决了相似图数值尺度不一的问题，取得了较优的结果。该方法与后续的谱聚类处理相结合，融入到一个优化框架下，进一步提升了多视角子空间聚类的性能。相关研究成果发表在 CVPR 2017 (Spotlight)。

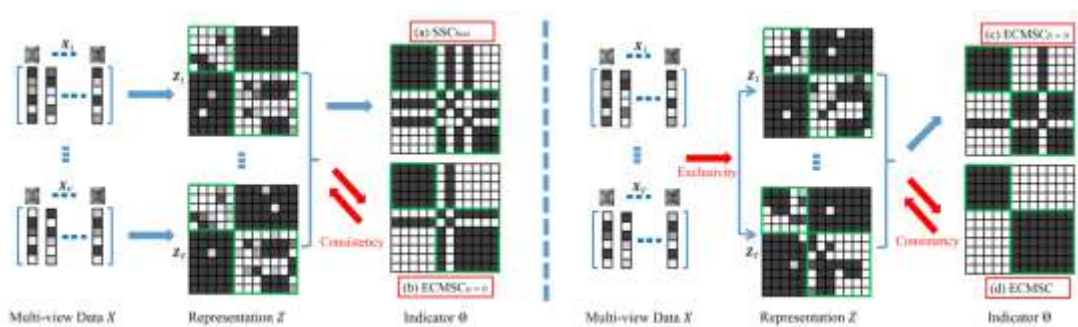


图 2 差异一致性的多视角子空间聚类算法框架

基于整体和自适应部件特征学习的行人再辨识

大规模视频监控中的行人搜索是智能视觉监控中的一个基础问题，行人再辨识技术是解决这个问题的重要途径。目前行人再辨识面临许多挑战，如图像分辨率低、光照变化、姿态和视角变化、遮挡、背景噪声等，这些都给行人的特征学习带来了巨大的困难。为了更好的学习行人特征表达，智能感知与计算研究中心博士生李党伟等人提出了一种多尺度的上下文感知网络，来更好地学习局部上下文知识，捕捉不同尺度的视觉信息，如眼镜，上衣等。考虑到行人自身的姿态变化和由于不完备的检测算法产生的大量背景，该工作还提出了一种自适应的行人部件定位算法来获取行人部件，以更好的捕捉行人的前景信息，克服姿态变化和背景噪声带来的影响。该方法通过融合行人整体和自适应部件，同时结合多尺度上下文感知的深度网络进行特征学习，有效的解决了行人姿态变化和大量背景给行人特征学习带来的困难，在国际主流的行人再辨识数据集中取得了领先的结果。相关研究成果发表在 CVPR 2017。

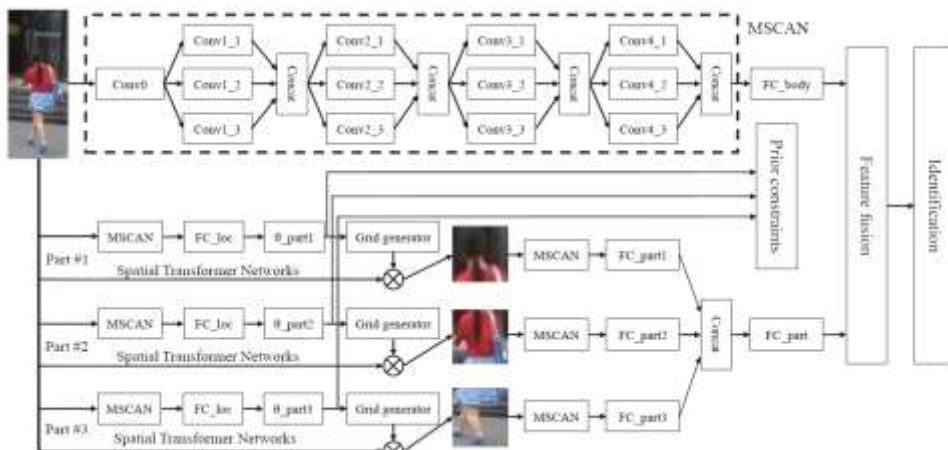


图3 基于整体和自适应部件的行人再辨识框架

感知语义实例的图像文本匹配

图像文本匹配是许多模式识别任务(如图像文本跨模态检索、图像描述生成、视觉问答系统等)的一个关键科学问题。当前用于图像文本匹配的算法通常对图像文本提取全局特征，然后利用典型关联分析或者子空间学习策略来拉近匹配的图像文本对。考虑到图像文本都包含很多冗余信息，即只有图像文本中的部分显著的、表征同一语义概念的内容才能被关联起来。为此，智能感知与计算研究中心提出了一种基于选择式多模态循环网络的图像文本匹配方法，可以选择性关注和匹配图像文本中的语义实例。所提出的选择式多模态循环网络是一个动态模型，在每一时间步，利用基于上下文的多模态注意机制选择图像文本中语义上相同的目标和词语，并计算其相似性作为图像文本的局部相似性。然后，该模型利用一个包含多个时间步的循环网络将多个局部相似性进行序列化融合，并最终得到图像文本的整体相似性。该模型的多模态注意机制和局部相似性融合模块可以通过随机梯度下降算法进行联合训练。该方法在当前两个公开数据集 Flickr30k 和 MSCOCO 上进行了验证，取得了较好结果。相关研究成果发表在 CVPR 2017。

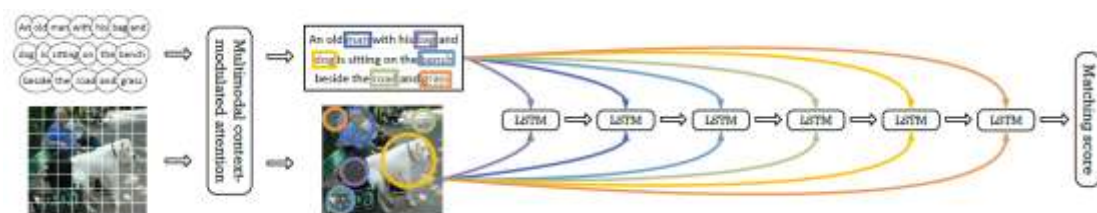


图4 基于选择式多模态循环网络的图像文本匹配框架

面向机器翻译的神经融合模型

近两年来，机器翻译模型经历了研究范式的转变，神经机器翻译逐渐取代统计机器翻译，取得了译文质量的突破性提升。神经机器翻译虽然能够生成自然流畅的译文，但是仍存在很多不足，特别是容易产生词不达意的情况。基本可以总结为：神经机器翻译的译文更加流畅通顺，而统计机器翻译更能忠实地传递原文语义。因此，如何博取众长生成既忠实原文又通顺流畅的译文成为一个具有重要意义和挑战的研究问题。针对这一问题，自然语言处理课题组博士生周龙与宗成庆研究员等人提出了一种基于多源输入的神经融合模型。该模型拓展了典型的端到端的神经机器翻译模型，在源语言表示端允许多个编码器，每个编码器对统计机器翻译和神经机器翻译的译文进行编码，然后通过层次化的注意机制模型逐词生成目标语言译文。这种方式不仅能够保持端到端模型的译文流畅性，同时能够将统计机器翻译的忠实原文的优点继承下来，显著提升了译文质量。相关研究成果发表在 ACL 2017。

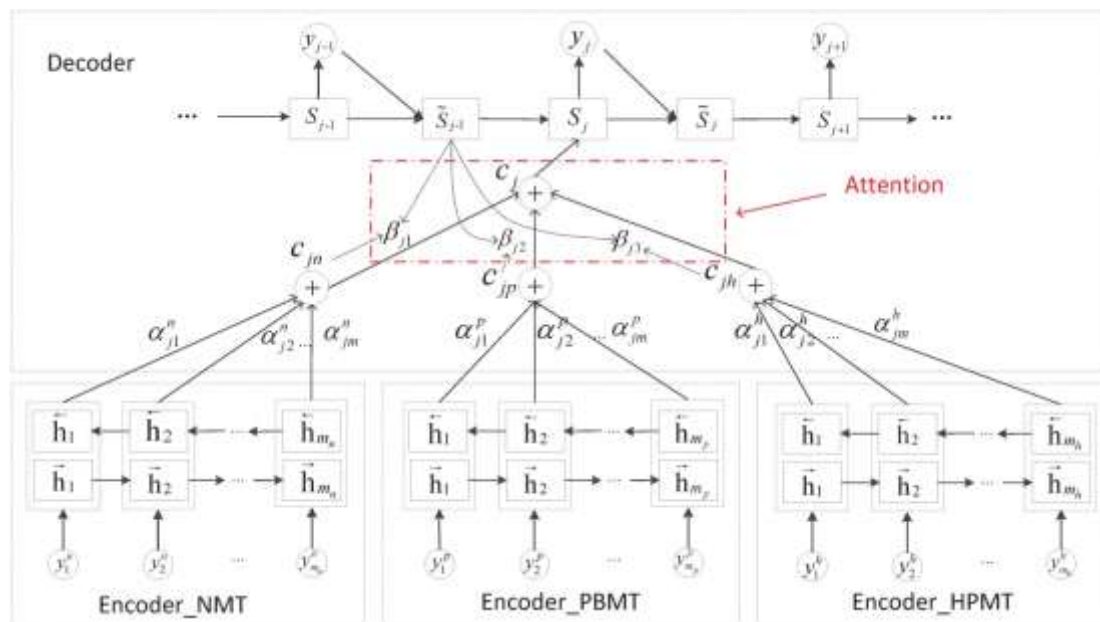


图5 多种翻译模型的神经融合系统示意图

面向冷启动问题的垃圾评论检测

垃圾评论检测是观点挖掘领域的一个重要任务，是工业界与学术界亟待解决的问题。针对该任务，传统的统计特征提取方法存在依赖丰富行为信息而导致无法解决冷启动的问题。为此，自然语言处理课题组提出了一种基于图结构与卷积神经网络的模型，在该领域内第一次尝试量化分析并处理冷启动问题。该项工作通过实验量化分析证明，传统统计特征需要建立在用户丰富的行为记录的基础上，而对于那些只发布了一条评论的新用户而言，基于传统统计特征的系统无法对其及时做出检测判定，因此传统特征方法无法处理冷启动状态下的垃圾评论检测任务。该工作所提出的模型将评论系统中的评论文本信息与用户行为信息联合编码，寻找与新用户评论文本信息相似的老用户，进而将其行为信息用来补充增强新用户不足的行为信息，以此检测垃圾评论。实验结果表明，相比于传统统计特征，该方法能够有效地及时地检测冷启动状态下的垃圾评论。相关研究成果发表在 ACL 2017。

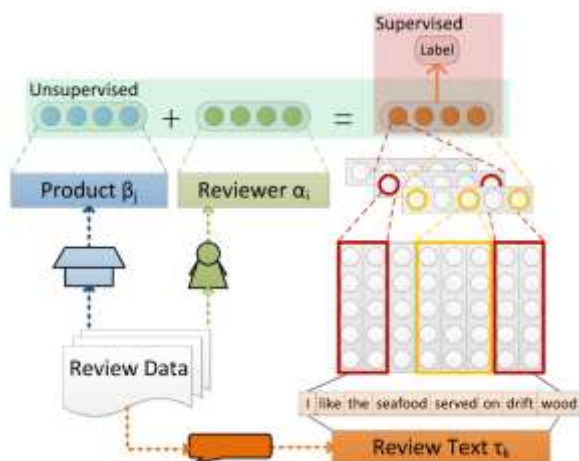


图 6 基于图结构与卷积神经网络的冷启动状态下的垃圾评论检测系统框图

基于有监督关注机制融合角色信息的事件识别方法

事件是由触发词和事件角色组成的复杂结构，事件的角色信息能够为事件识别提供重要的线索。然而，几乎所有的相关工作在进行事件识别时，要么完全忽略了训练集中标注的事件角色信息，要么只是通过间接方式使用该信息，不能有效地对该信息进行建模。针对该问题，自然语言处理课题组提出了一种直接应用角色信息进行事件识别的方法，其基本思想是在事件识别的过程中重点关注角色信息。为了实现该目标，首先提出了一种利用有监督关注机制融合角色信息进行事件识别的方法，并探索了不同关注策略对模型的影响。在训练阶段，首先利用训练集中标注的事件角色信息构建标准关注向量，然后将其作为监督信息训练模型的关注机制，因此事件识别模型和关注机制都在有监督的条件下进行训练。在测试阶段，利用学习得到的关注机制和事件识别模型在测试文本中进行事件识别。公开数据集上的一系列实验的结果证明了本方法的有效性。相关研究成果发表在 ACL 2017。

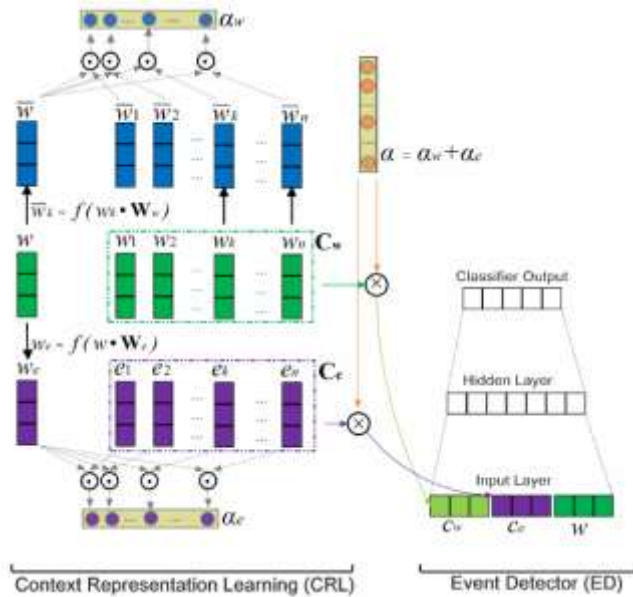


图7 基于有监督关注机制的事件识别模型框架图

学术交流

乔治亚理工学院 B.H. Juang 教授来实验室访问交流

2017年3月23日，美国乔治亚理工学院讲座教授、美国工程院院士、美国发明家学院院士、IEEE会士、中央研究院院士 Biing Hwang (Fred) Juang（庄炳煌）教授来访中科院自动化所，并做了题为“Deep Neural Networks from a Developmental Perspective”的学术报告。庄教授在报告中指出，近年来深度神经网络的研究和应用异常火热，相关的文章非常多。然而一个令人沮丧的现象是：许多学者和学生都把深度神经网络视为一个黑盒子，并没有真正理解其本质。庄教授利用贝叶斯决策以及马尔科夫随机场等基础理论去解释深度神经网络，建立深度神经网络与模式识别经典理论与方法的联系，从而使大家对深度神经网络有更为深入的理解。同时，庄教授回顾了深度神经网络的发展历程，解释了随着Hinton对受限玻尔兹曼机（RBM）理论的贡献，神经网络从早期的多层感知器（MLP）演变成现在颇为强大复杂的深度神经网络和深度循环网络。

庄教授是语音处理及信号处理领域国际知名学者，曾任美国贝尔实验室信息研究部主任，曾任IEEE Fellow Committee主席。庄教授学术贡献甚丰，包括经典的矢量量化、隐马尔科夫模型、语音编码与识别、多频回声控制等技术，其所著《语音识别基础》(Fundamentals of Speech Recognition, co-authored with L.R. Rabiner)一书被公认是经典之作。庄教授所获奖项也极多，包括IEEE信号处理学会技术成就奖、IEEE技术领域J.L. Flanagan奖章等。

清华大学王小勤教授来实验室访问交流

2017年3月23日，美国约翰霍普金斯大学、清华大学生物医学工程系王小

勤教授访问实验室脑网络组研究中心，并做题为“Marmoset as a model system for studying neural basis of audition and social communication”的学术报告。王小勤教授的研究方向为灵长类听觉大脑皮层处理语音信号和音乐信号的机制、听觉系统在复杂声场中提取信号的原理、灵长类大脑控制发声和听觉反馈的机理，以及注意力和行为对听觉大脑皮层的调制。在过去 20 多年中，他所领导的实验室创立了使用绒猴（Marmoset）研究神经机理的独特实验模型，开创了大量绒猴神经电生理和行为学研究的核心实验技术。因他对“听觉神经生物学做出的杰出贡献以及开创了在语音通讯生理机制研究中的实验和计算方法”，王小勤教授于 1999 年荣获“美国总统青年科学家奖”（Presidential Early Career Awards for Scientists and Engineers）。该项奖是美国政府授予在美国工作的杰出青年科学家的最高荣誉。

项目立项

实验室第一季度新建立课题 38 项

实验室 2017 年第一季度新建立课题 38 项，总科研经费 3453 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
图像分析与模式识别	赫然	优秀青年基金	130	2017/01/01 至 2019/12/31
模拟人脑多重储存记忆模型的在线视觉跟踪研究	高晋	青年科学基金	21	2017/01/01 至 2019/12/31
四维光场人脸图像的预处理和识别	孙运莲	青年科学基金	21	2017/01/01 至 2019/12/31
面向虚拟听觉的个性化头相关传输函数快速自适应研究	戚肖克	青年科学基金	20	2017/01/01 至 2019/12/31
谷氨酸信号通路影响精神分裂症患病风险的脑机制研究	郑凡凡	青年科学基金	17	2017/01/01 至 2019/12/31
基于扩散磁共振成像的概率纤维追踪算法研究	叶初阳,	青年科学基金	16	2017/01/01 至 2019/12/31
基于社交多媒体行为的物理世界用户属性研究	桑基韬	基金面上项目	63	2017/01/01 至 2020/12/31
无约束人脸识别问题研究	兴军亮	基金面上项目	63	2017/01/01 至 2020/12/31
数据驱动的图片合成	董未名	基金面上项目	62	2017/01/01 至 2020/12/31
跨视域行人检索方法研究及刑侦应用	廖胜才	基金面上项目	60	2017/01/01 至 2020/12/31
基于层次化表达的大规模图	黄凯奇	基金面上项目	62	2017/01/01 至

像物体识别				2020/12/31
排序学习中的对象集结构分析和等序关系输出研究	原春锋	基金面上项目	62	2017/01/01 至 2020/12/31
基于弱监督的神经网络翻译模型研究	张家俊	基金面上项目	62	2017/01/01 至 2020/12/31
绘制绒猴脑网络组图谱	杨正宜	基金面上项目	56	2017/01/01 至 2020/12/31
基于多视角航拍图像的大规模城市场景快速重建	徐士彪	基金面上项目	60	2017/01/01 至 2020/12/31
深度认知神经网络理论与方法研究	王亮	基金重点项目	245	2017/01/01 至 2021/12/31
城市道路交通大数据融合方法研究	向世明	基金重点项目	240	2017/01/01 至 2020/12/31
多模态人类脑图谱绘制	蒋田仔	重点国际(地区)合作研究项目	242	2017/01/01 至 2021/12/31
联合的图像大数据语义分析与城市场景重建	张晓鹏	重点国际(地区)合作研究项目	245	2017/01/01 至 2021/12/31
网络敏感多媒体识别与比对关键技术研究	胡卫明	联合基金项目	253	2017/01/01 至 2020/12/31
基于多级生理特征信号的心理应激状态判别研究	张鑫	联合基金项目	67	2017/01/01 至 2019/12/31
应对大数据分析的情景融合个性化隐写理论与方法	董晶	联合基金项目子课题	74.4	2017/01/01 至 2020/12/31
基于语义分析的大规模视觉场景重建关键技术研究	申抒含	基金重点项目子课题	68.85	2017/01/01 至 2021/12/31
危险因素靶向干预与神经环路再激活	刘勇	国家重点研发计划	25	2016/09/01 至 2020/12/31
多尺度猴脑网络组图谱的构建及转录组研究	蒋田仔	中科院重点研究项目	300	2016/01/01 至 2020/12/31
类人视觉语义建模与理解	黄凯奇	中科院前沿科学重点研究项目	250	2017/01/01 至 2017/12/31
院青年促进会项目--张家俊	张家俊	院青年促进会	19.5	2017/01/01 至 2020/12/31
院青年促进会项目--李兵	李兵	院青年促进会	19.5	2017/01/01 至 2020/12/31
院青年促进会项目--樊令仲	樊令仲	院青年促进会	19.5	2017/01/01 至 2020/12/31
院青年促进会项目--董晶	董晶	院青年促进会	19.5	2017/01/01 至 2020/12/31
个性化可穿戴产品的3D模型设计与社会化制造核心技术研究	叶军涛	广东省科技计划项目	50	2015/07/01 至 2017/06/30

基于家系队列的阿尔茨海默病痴呆前诊断技术研究	刘勇	北京市科技计划项目	20	2016/01/01 至 2018/12/31
视觉定位导航联合实验室	李和平	北京布科思科技有限公司	300	2016/12/30 至 2019/12/31
基于手机图像的桌面场景三维重建软件开发合作协议书	申抒含	亮风台信息科技有限公司	60	2017/01/19 至 2017/04/30
Cross-lingual Text and Knowledge Graph Jointly Embedding	刘康	Google 公司	34.28	2016/12/01 至 2020/06/01
基于深度学习的人体视觉特性研究	兴军亮	华为技术有限公司	21.2	2016/11/22 至 2017/05/22
稀疏分析新方法及其在视觉运动分析中的应用	胡卫明	英国伦敦大学 Birkbeck 学院	100	2017/01/01 至 2019/12/31