



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 2 期

总第 42 期

2017 年 4-6 月

目 录

最新动态	1
国家重大科学仪器设备开发专项“光电同步脑活动检测仪器开发”通过初步验收	1
陶建华研究员入选科技部“中青年科技创新领军人才”	1
黄凯奇研究员入选第十二届“中国科学院杰出青年”	2
科研进展	2
融合多上下文的多视角多示例学习算法	2
基于深度学习的相对性视觉跟踪方法	3
不同语气的高表现力自然语音生成	3
受人类注意机制启发的句子语义表示方法	4
基于 Multilingual I-vector 音频内容表征的段特征建模方法	5
学术交流	5
德国康斯坦茨大学 Oliver Deussen 教授来实验室访问交流	5
新墨西哥大学 Vince D Calhoun 教授来实验室访问交流	6
项目立项	6
实验室第二季度新建立课题 25 项	6

《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：鲍秉坤 叶军涛 张 曼 刘 斌 张家俊 殷 飞
原春锋 廖胜才 孟高峰 张一帆 张亚萍 王爱华

最新动态

国家重大科学仪器设备开发专项“光电同步脑活动检测仪器开发”通过初步验收

2017 年 6 月 19 日，科技部重大科学仪器设备开发项目“光电同步脑活动检测仪”初步验收会议在自动化所召开。项目由中国科学院自动化研究所牵头，电子科技大学等 10 余家单位合作开发。模式识别国家重点实验室脑网络组研究中心蒋田仔研究员担任项目负责人。经来自科技部、工信部、清华大学、国防科技大学等 12 家单位共同组成的专家评审团认定，项目通过初步验收。中科院条件保障与财务局姜言彬副处长、中科院自动化所刘成林副所长出席验收会议。

验收会上，蒋田仔研究员介绍了项目总体情况，其他 9 个单位汇报了子任务情况。项目总目标是光电同步脑活动检测仪器开发，通过仪器开发、应用开发、工程化开发和产业化，逐步实现目标仪器从工程样机到标准产品的转化，真正将仪器投放于市场，为相关国家重大科技项目和科学研究提供技术支撑。

项目自 2012 年 10 月起开始启动，在四年的项目实施过程中，已成功研制了高性能多通道近红外血氧检测模块，开发了高性能多导脑电采集模块，同时完成了近红外检测模块与脑电刺激模块的硬件以及软件平台的搭建，实现了近红外光谱技术和脑电采集技术的同步采集，真正实现了脑功能信号的多模态数据融合。在硬件研发以及软件研发等各个方面都有许多创新点，申请了国内专利 64 项、国际专利 4 项、软件著作权 21 项。目前已生产 32 通道光电同步脑活动检测系统 5 台，8 通道光电脑活动检测系统 8 台，2 通道血氧检测头戴 5 台。

项目研制的光电同步脑活动检测系统已经在多家应用单位开展研究工作，对于深入研究正常及疾病状态下的脑功能活动具有重要的价值，能够同时获取神经电活动和血氧代谢信息，满足了许多研究机构和临床应用单位的重大需求。该系统具有同步采集神经电活动和脑代谢信息方面的独特优势，在认知以及重大神经精神疾病的临床应用方面具有很好的社会效益。

陶建华研究员入选科技部“中青年科技创新领军人才”

2017 年 6 月 15 日，科技部发布《关于 2016 年创新人才推进计划入选名单的通知》，模式识别国家重点实验室陶建华研究员入选“中青年科技创新领军人才”。根据《创新人才推进计划实施方案》规定，科技部开展了 2016 年创新人才推进计划组织实施工作。经申报推荐、形式审查、专家评议和公示等环节，确定 314 名中青年科技创新领军人才、67 个重点领域创新团队、203 名科技创新创业人才和 33 个创新人才培养示范基地入选 2016 年创新人才推进计划。

创新人才推进计划旨在通过创新体制机制、优化政策环境、强化保障措施，培养和造就一批具有世界水平的科学家、高水平的科技领军人才和工程师、优秀创新团队和创业人才，打造一批创新人才培养示范基地，加强高层次创新型科技人才队伍建设，引领和带动各类科技人才的发展，为提高自主创新能力、建设创新型国家提供有力的人才支撑。

黄凯奇研究员入选第十二届“中国科学院杰出青年”

2017 年 6 月,中国科学院公布了第十二届“中国科学院杰出青年”获奖名单,模式识别国家重点实验室黄凯奇研究员入选。

黄凯奇研究员主要围绕视觉计算理论、关键技术及应用开展研究,对面向安全的图像理解和视频分析中的诸多关键问题进行了系统深入研究。负责了包括国家重点研发计划、国家科技支撑计划、中科院前沿基础重点研究计划等 40 余项国家和地方项目,通过与重大应用相结合,将基于视觉计算理论及关键技术的相关系统成功应用于多个国家安全领域和重要部门。

“中国科学院杰出青年”评选表彰活动是在中科院党建工作领导小组的领导下,由院党建工作领导小组办公室、京区党委、院青联和院团委联合组织,旨在树青年楷模,展青年风采,引导和鼓励广大青年坚定理想信念,进一步调动广大青年的积极性和创造性,培养造就适应科技创新和管理创新需要的优秀青年,为实现“四个率先”作出更大贡献。自 1995 年启动以来,“中国科学院杰出青年”评选已进行了十二届,每届评选出 10 人。

科研进展

融合多上下文的多视角多示例学习算法

多示例学习广泛应用于图像、视频分类等领域。现有的绝大多数多示例学习算法假设示例间是相互独立的。这种特定假设只能够反映示例间的某一种特性。在实际应用中,示例之间存在各种复杂的上下文关系。为了能够有效地对这些复杂的上下文关系进行建模,视频内容安全团队李兵副研究员等提出了一种多视角多示例学习算法,算法框架如图 1 所示。该算法的主要创新点有:(1)提出了一种稀疏图的构图算法,该算法能够根据不同的参数设置,系统地为每一个示例包构建不同的无向图,以表达示例间复杂的上下文关系;(2)提出了一种多视角联合稀疏表示框架,融合多个图结构对示例包进行分类,从而达到考虑多种上下文关系的目的;(3)提出了一种面向多个图结构的字典学习算法,进一步提升了示例包的分类精度。该方法被应用到恐怖视频识别中,有效地提高了网络恐怖视频识别的精度。相关研究成果发表在发表在国际期刊 *IEEE Transactions on Pattern Analysis and Machine Intelligence*。

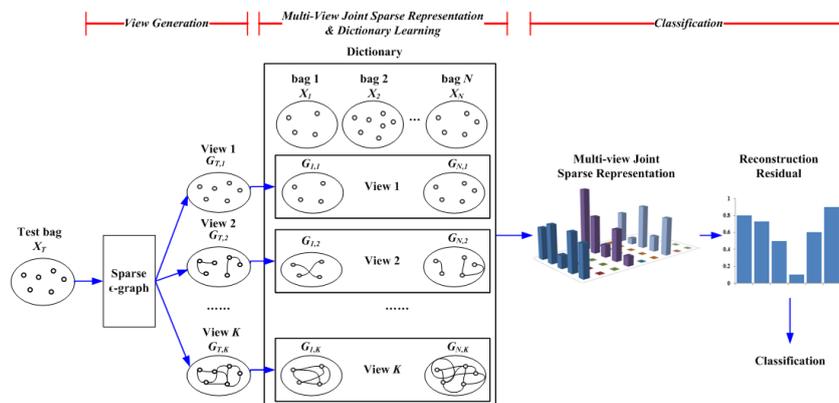


图 1 多视角多示例学习算法框架

基于深度学习的相对性视觉跟踪方法

视觉跟踪是计算机视觉和多媒体分析的研究热点问题之一，已经在视频监控，自动驾驶，人机交互和运动分析等领域得到广泛应用。尽管近年来取得了显著进步，但由于部分遮挡、光照变化、目标形变、运动模糊，背景复杂和尺度变化等因素的存在，对目标进行稳健、高效的跟踪仍然是一个具有挑战性的任务。

大多数传统的跟踪方法没有对跟踪场景中图像块间的潜在关系进行建模。由于未考虑目标物体周围图像块之间的结构化关系，当目标表观因部分遮挡、背景复杂或其他因素而发生显著变化时，这些传统的跟踪方法可能存在漂移问题。为了克服上述问题，多媒体计算课题组徐常胜研究员等提出了一种新颖的相对跟踪方法，可以有效地利用图像块之间的相对关系进行建模。相对性跟踪方法可以找到相对性分数最高的图像块，对目标物体的表观变化有较好的鲁棒性。为了建模大量图片对之间的相对性关系，课题组提出了一种新颖有效的基于卷积神经网络的相对性建模方式。相关研究成果发表在国际期刊 *IEEE Transactions on Image Processing*。

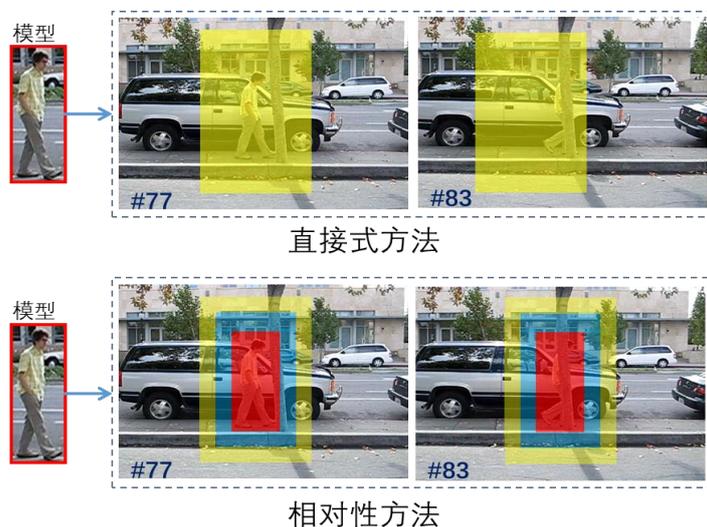


图2 直接式跟踪方法与相对性跟踪方法比较

不同语气的高表现力自然语音生成

对话系统研究与语调、韵律等超音段特征的结合，以及其与语气、情感、信息结构、焦点等因素的互动，为语音研究带来了新鲜而复杂的课题。人机语音交互课题组李雅副研究员等人在大规模的包含陈述和疑问的对话语料基础上，对汉语六种疑问句的韵律、语调进行了统计分析，研究发现：陈述句与各种型疑问句在韵律上的差异主要体现为两点，一是疑问句句首音高高于陈述句，二是疑问句整句音强与句末音强均高于陈述句。同时疑问句的语调并非传统感知上的上升，而是同陈述句一样，具有句末音高下倾模式。同时，该方法借鉴了汉语传统的韵律模型，构建了疑问句的语调数学模型，该语调模型可以用于语音合成系统中的韵律生成。以上研究成果将能够很好促进自然流畅的汉语疑问句的语音生成，同时促进汉语口语对话系统的研究和应用。相关研究成果发表在国际期刊 *Speech Communication*。

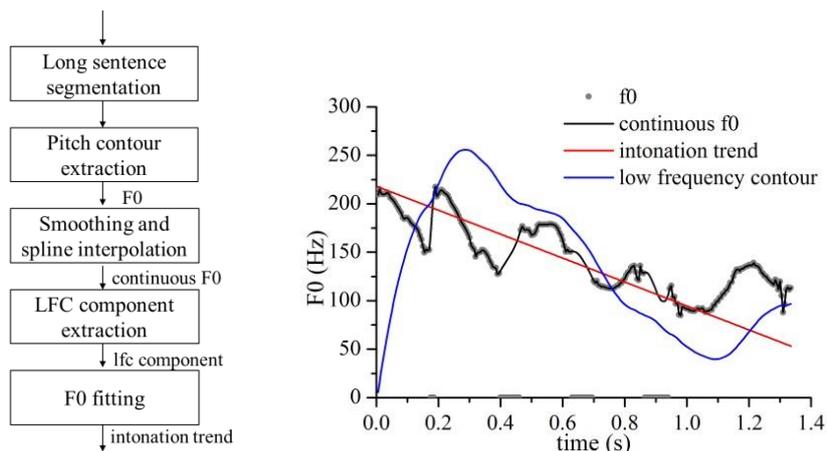


图3 汉语疑问句语调建模（左：语调建模流程图；右：分步语调建模后的基频曲线）

受人类注意机制启发的句子语义表示方法

人类在阅读和理解句子时会选择性的注意句子中的某些词汇，也会选择性的跳读一些词汇，这种注意力机制让人阅读和理解句子变的更加高效。受人类注意力机制的启发，自然语言处理课题组博士生王少楠在宗成庆研究员的指导下发现在构建句子表示时给重要的词汇赋予较高的权重可以得到更好的句子表示。大量的有关人类阅读时间的研究证明了词汇的特性，如词性、词长、词频、词汇惊异度（Surprisal）等，都会影响人类阅读文本时对这个词汇的关注程度。为此，可以选择两种词汇特性来对词汇的重要程度进行建模。一种是词汇惊异度，这是一个连续值的变量。词汇惊异度反映了一个单词在一个句子中传达的信息量，由词汇在句子中的负对数条件概率计算得到，通常这个值越高表示理解这个词越困难，需要更多的阅读时间。另一种是离散的词汇特性，利用了词汇的词性和组合语法范畴(CCG)(词类特征)。上述两种注意力模型可以得到每个词汇的注意力权重，然后用加权平均的方式对词向量进行组合得到这个句子的语义表示。该工作是结合认知心理学的研究建立自然语言处理模型的初步尝试，利用了人类的注意力机制来改进句子表示模型的效果，取得了显著的提升。相关研究成果发表于人工智能国际会议 IJCAI 2017。

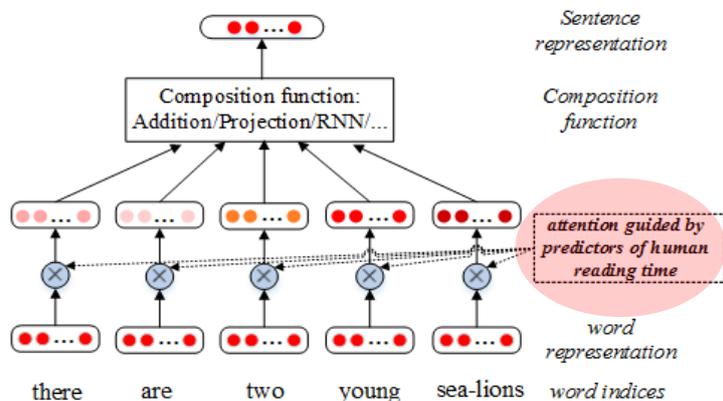


图4 受人类注意机制启发的句子语义表示方法

基于 Multilingual I-vector 音频内容表征的段特征建模方法

音乐特征是对音乐内容的一种描述,如果所提音乐特征包含了各种重要的音乐特性,如节拍、旋律、响度、歌手信息、所用乐器或者其他特征描述等等,那么这样的特征便能很好地表达出音乐的内容。但提取能包含如此多音乐特性的特征是很困难的,而且不同分类任务所需要的特性是不同的,如何选择并提取符合当前分类任务的特性并组合起来更加困难,针对这一问题,戴佳博士提出了一种基于高斯分布来拟合并自动学习生成不同的音频特性的新方法。该方法避免了复杂的特性选取,组合和计算,方法假设每种音乐特性可以用一个高斯混合模型描述,而高斯混合模型中每个高斯分量都可以看作一个微特性(“音频单元”)。在基于 I-vector 的特征建模框架中,采用基于高斯混合模型的全局背景模型(GMM-UBM)学习“音频单元”。在理论上,如果有足够多的数据和高斯个数,高斯混合模型可以拟合任意的数据分布。因此为了使学到的“音频单元”更充分,该方法结合了 Multilingual 模型,通过网上下载的大量的无标签的音乐数据训练 GMM-UBM 模型,最终得到基于 Multilingual I-vector 的“音频单元”表征的特征。实验证明该方法得到的特征能很好地表征音乐信号内容,分类性能显著优于一般的音乐信号表征方法。相关研究成果发表在国际会议 Interspeech 2017。

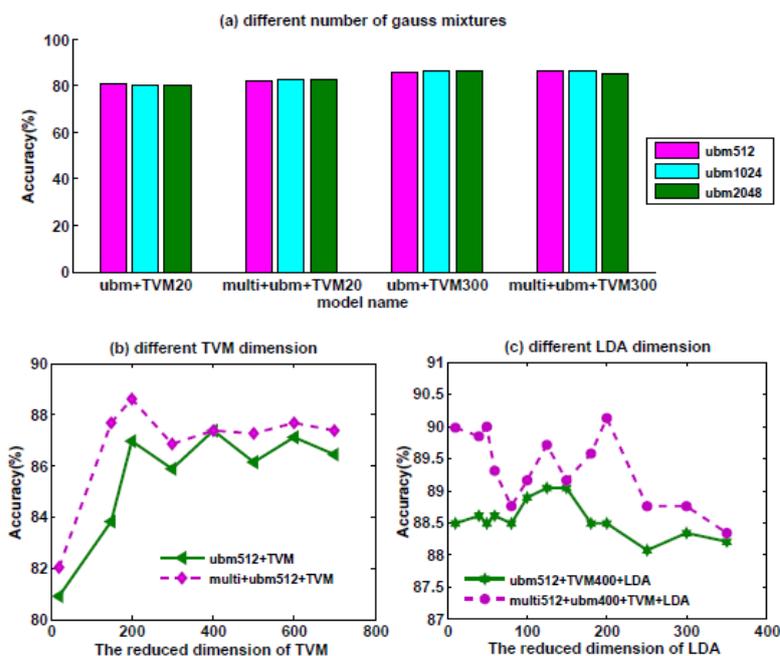


图5 各种方法的音乐信号表征对比

学术交流

德国康斯坦茨大学 Oliver Deussen 教授来实验室访问交流

2017年4月14日,德国康斯坦茨大学 Oliver Deussen 教授来实验室访问交

流，并作了题为“定量信息可视化:处理不确定性”的学术报告。Deussen 教授围绕可视媒体计算中的定量方法研究这一主题，首先针对可视媒体计算这一概念进行了介绍，包含了计算机视觉、图像分析、计算机图形学和信息感知等多个研究领域的内容，并将这些看似独立的领域通过可视分析技术（如可视化、人机交互和机器学习等）成功的联系起来。Deussen 教授指出，在可视计算领域，尤其是自然数据方面，最常用的观察和评价方法是基于人的主观性的定性评价，从而缺少了准确的定量信息。他分别以全球温度变化颜色图和斯坦福大学的朋友社交关系图为例，说明了定量的分析技术可以更准确和有效的得到可视计算所要反映的信息。Deussen 教授在报告中总结了可视计算的一般流程，并重点介绍了可视计算中的定量分析方法，从人类视觉差异预测和不确定信息的可视化两个方面进行了阐述。在报告的结尾，Deussen 教授希望通过数据可视化的定量分析研究，可以从这些有趣和大规模的数据中找到规律，从而启发和促进新的科学知识的发现。

Oliver Deussen 教授毕业于德国卡尔斯鲁厄理工学院，目前是德国康斯坦茨大学终身教授，2010 年受聘为中国科学院外国专家特聘研究员，2012 年入选国家外专局高端外国专家，2014 年入选外国专家千人计划项目。曾担任欧洲图形学会 Eurographics 副主席，欧洲图形学会刊 Computer Graphics Forum 主编。研究领域主要包括复杂场景建模、非真实感渲染、几何处理、信息可视化等。他在国际会议和期刊发表学术论文 140 余篇，其中有 3 篇为会议最佳论文。Deussen 教授是国际著名植物建模技术 Xfrog 的创始人。2014 年，Deussen 教授荣获奥斯卡科学技术成果提名。

新墨西哥大学 Vince D Calhoun 教授来实验室访问交流

2017 年 6 月 2 日，新墨西哥大学 Vince D Calhoun 教授访问实验室脑网络组研究中心，做题为“Maximizing information in brain imaging studies: data capture, prediction, dynamics, and data-fusion”的学术报告，系统地阐述了基于脑影像的脑网络构建、脑网络动态表征、大型数据库建设等相关研究内容。Calhoun 教授是美国科学促进协会 AAAS、国际电气和电子工程师协会 IEEE、美国医学与生物工程协会 AIMBE、国际医学磁共振学会 ISMRM 与美国神经心理药理学联盟 ACNP 5 个学会的会士(Fellow)，国际人脑成像图谱协会（OHBM）新任主席（chair-in-elect），美国新墨西哥大学杰出教授，the Mind Research Network 高级科学官。目前已发表 500 多篇 peer-reviewed 期刊论文和 550 篇会议论文，google scholar 索引大于 30000 次，H-index=87。作为一名拥有电子信息/生物医学交叉技术的科研先锋，Calhoun 教授提出了多种基于数据驱动的脑影像分析方法，在业内具有广泛影响，其中包括 fMRI 的组间独立成分分析、全脑动态网络分析、影像基因组学的独立成分分析等等，引领了神经影像领域动态连接组分析的潮流。他研发的 GIFT 软件被业内广泛用于 fMRI 的分析。

项目立项

实验室第二季度新建立课题 25 项

实验室 2017 年第二季度新建立课题 25 项，总科研经费 1558.54 万元，其中

包括中科院项目 4 项、地方政府和部委项目 3 项、企业委托项目 11 项，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
万人计划—胡卫明	胡卫明	万人计划	90	2017-01-01 至 2019-12-31
推进超高场磁共振神经影像技术	蒋田仔	中科院国际合作局对外合作重点项目	75	2017-01-01 至 2018-12-31
稀疏分析新方法及其在视觉	胡卫明	中科院国际合作局对外合作重点项目	100	2017-01-01 至 2019-12-31
灵长类动物介观脑网络组研究平台	蒋田仔	中科院修购专项	330	2017-02-22 至 2017-12-31
2017 年脑网络组与基因组汇聚国际研讨会	蒋田仔	中科院国际合作局国际会议资助	15	2017-02-13 至 2017-12-31
人脑重要功能重塑精细动态图谱的构建	樊令仲	北京市科委项目	150	2017-01-01 至 2019-12-31
社会热点事件协同分析技术研究	张天柱	北京市自然科学基金面上项目	18	2017-01-01 至 2019-12-31
基于多模态磁共振影像的认知损害脑网络研究	刘勇	中国医学科学院北京协和医院	33	2017-03-09 至 2018-12-31
整机仿真几何特征识别技术合作协议	严冬明	华为终端(东莞)有限公司	50	2016-08-10 至 2018-02-28
视觉描述	徐常胜	微软(中国)有限公司	20	2016-10-01 至 2017-08-31
英伟达 DGX-1 合作研究	王亮	英伟达公司	68.79	2017-02-17 至 2022-02-17
航天领域红外制导中的图像跟踪器技术研究	樊彬	上海航天控制技术研究所	28	2016-09-30 至 2018-09-30
基于高斯过程回归的图像处理技术	高晋	灵动科技(北京)有限公司	20	2017-03-13 至 2019-03-12
综合 AR 识别引擎开发协议	王金桥	联想新视界(北京)科技有限公司	80	2017-03-22 至 2018-03-22
基于深度神经网络的高性能中日互译机器翻译系统的开发与研究	张家俊	东芝(中国)有限公司	30	2017-04-01 至 2018-03-31
自动跟踪技术合作项目	王金桥	华为技术有限公司	37.1	2017-03-29 至 2018-03-31
面向智慧停车的车辆检测与跟踪系统	王金桥	智慧互通科技有限公司	60	2017-05-02 至 2017-11-01

导电游丝自动化焊接机器 视觉技术开发	侯广琦	北京航天控制仪 器研究所	48	2016-11-28 至 2017-11-28
-----------------------	-----	-----------------	----	----------------------------