

CCF

计算机视觉 专委会简报

COMPUTER VISION NEWSLETTER

2017/03 期
总第 07 期

专委会动态

走进高校系列活动

科技前沿

顶会参会感悟

专题综述



主 办：CCF 计算机视觉专业委员会

主 编：王 亮

执行主编：李实英

网 址：<http://ccfcv.ccf.org.cn>

E m a i l：ccfcvn@gmail.com

COMPUTER VISION NEWSLETTER

计算机视觉

专委简报编委会

主 编 王 亮 中国科学院自动化研究所
执行主编 李实英 湖南大学

专委动态

主编 毋立芳 北京工业大学
编委 马占宇 北京邮电大学
王瑞平 中国科学院计算技术研究所
虞晶怡 上海科学技术大学

科技前沿

主编 申抒含 中国科学院自动化研究所
编委 邓 成 西安电子科技大学
卢国梁 山东大学
任传贤 中山大学
苏 航 清华大学
王金甲 燕山大学
杨巨峰 南开大学

委员风采

主编 余 焯 合肥工业大学
编委 韩爱丽 山东大学
刘海波 哈尔滨工程大学
余志文 华南理工大学

资源平台

主编 苗启广 西安电子科技大学
编委 樊 鑫 大连理工大学
贾 同 东北大学
蹇木伟 中国海洋大学
李 策 兰州理工大学
刘 丽 国防科学技术大学
沈沛意 西安电子科技大学

COMPUTER VISION NEWSLETTER

CONTENTS

目录

专委动态

| | | |
|------|-------------------|----|
| 走进高校 | CCF-CV 走进高校系列报告会 | 04 |
| 走进企业 | CCF-CV 走进企业系列交流会 | 05 |
| 信息速递 | 走进企业系列交流会活动细则 | 06 |
| 会议推介 | CCCV 2017 讲习班讲者征集 | 07 |
| | CCCV2017 技术挑战赛 | 08 |

科技前沿

| | | |
|------|--------------------|----|
| 专题综述 | 张量稀疏性度量 | 11 |
| 热点追击 | 人物动作中的语义挖掘与分类 | 16 |
| | 基于对应关系传递一致性的多图匹配方法 | 17 |
| | 数据深度聚类分析 | 18 |

委员风采

| | | |
|-------|-------------|----|
| 委员访谈 | 同济大学王瀚漓教授访谈 | 19 |
| 委员好消息 | | 22 |

资源平台

| | | |
|------|----------------------|----|
| 数据集 | 场景理解数据集介绍 | 23 |
| 仿真平台 | TensorFlow：开源机器学习软件库 | 25 |
| 招聘信息 | | 27 |
| 征文通知 | | 30 |

CCF-CV 走进高校系列报告会

第 29 期 上海交通大学

时间：2017 年 3 月 20 日

本期报告会邀请了北京大学**田永鸿**教授报告《大规模视觉对象的精准搜索与识别》、中国科学院自动化研究所**王亮**研究员报告《媒体大数据分析 with 理解》、浙江大学**吴飞**教授报告《序列数据深度学习及其思考》和上海交通大学电子信息与电气工程学院副院长**杨小康**教授做《未来人工智能视觉感知的若干思考》专题报告。来自上海多所高校及研究院的约 300 名师生及业界精英参加了本次报告会。



第 30 期 中央民族大学

时间：2017 年 3 月 27 日

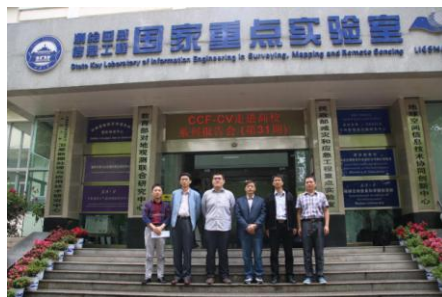
本期报告会由中央民族大学信息工程学院媒体计算实验室承办，分别邀请了中国科学院自动化所**王亮**研究员报告《深度学习及其在视觉模式分析中的应用》、上海大学**张新鹏**教授报告《密文域信号处理》、中国科学院信息工程研究所**操晓春**研究员报告《网络空间视觉内容理解》和厦门大学**纪荣嵘**教授报告《视觉搜索与识别系统中的紧凑性问题》。



第 31 期 武汉大学

时间：2017 年 4 月 16 日

本期报告会由武汉大学测绘遥感信息工程国家重点实验室承办，分别邀请了华南理工大学**金连文**教授报告《Drop to Gain——浅谈深度学习中的“舍弃”法训练技术及其应用》、浙江大学**李玺**教授报告《基于深度学习的光流估计和 Co-Saliency 检测》、清华大学自动化系**鲁继文**副教授报告《深度学习与视觉分析》、西北工业大学光学影像分析与学习中心**王琦**副教授报告《复杂场景中人群行为的研究及其应用》、中国科学院自动化研究所模式识别国家重点实验室**申抒含**副研究员报告《基于图像的大规模场景三维建模》。



第 32 期 中山大学

时间：2017 年 4 月 28 日

本期报告会以“计算机视觉前沿技术及应用”为主题，邀请南京信息工程大学**刘青山**教授报告《从稀疏到深度的视觉特征表示》、清华大学**马惠敏**副教授报告《三维场景图像认知与多模态学习》、华南理工大学**谭明奎**教授报告《黎曼优化及其在机器视觉中的应用》、华中科技大学**白翔**教授报告《自然场景文字检测与识别》四位学者做主题报告。



CCF-CV 走进高校系列活动预告

| 时间 | 走进高校 |
|-------------|--------|
| 2017 年 6 月 | 湘潭大学 |
| | 上海大学 |
| 2017 年 7 月 | 重庆大学 |
| | 陕西科技大学 |
| 2017 年 9 月 | 公安部三所 |
| | 南昌航空大学 |
| 2017 年 10 月 | 西安交通大学 |
| 2017 年 11 月 | 同济大学 |
| | 武汉大学 |
| 2017 年 12 月 | 安徽工业大学 |
| | 华北电力大学 |

(责任编辑：马占宇)

CCF-CV

走进企业系列交流会



2017年4月27日第十期 CCF-CV 走进企业系列交流会走进商汤科技圆满结束

2017年4月27日,中国计算机学会计算机视觉专委会(CCF-CV)走进企业系列交流会第十期活动——走进商汤科技,在北京市海淀区清华科技园科技大厦B座5层商汤科技会议室成功举办。

本期活动由 CCF-CV 秘书处召集,商汤科技**黄璇**女士与**葛瑜**女士负责现场组织协调。活动通知发出后,参会申请邮件纷至沓来,特别是众多在读研究生对这家致力于引领人工智能核心“深度学习”技术突破的科技创新公司表现出了很大的热情与兴趣。由于主办方场地席位所限,最终选择了来自企业、高校、科研院所等 30 余名代表参加了此次活动。商汤科技对此项专委交流活动给予了高度重视和大力支持,交流会由商汤科技高级研究员**彭义刚**博士主持,商汤科技执行研发总监**林惊**教授领衔,高级研究员**石建萍**博士、产品总

监**车英慧**博士联袂参加了本次活动,并分别就商汤科技从学术创新到产业落地、深度学习像素级预测应用于智能手机及智能驾驶项目、公司典型的落地应用产品等作了专题报告。

首先,商汤科技执行研发总监**林惊**教授以“深度驱动的人工智能:从学术创新到产业落地”为题全面介绍了商汤科技围绕人工智能产业发展所进行的方向布局和技术研发。他的报告从人工智能概览、深度学习之创新与颠覆、产业化应用、未来发展趋势这四个方向逐一展开,带领听众系统全面地了解了商汤科技在深度学习方向持续取得的一系列学术创新与产业应用成果。

接下来,由商汤科技高级研究员**石建萍**博士为大家带来深度学习像素级预测在智能手机及智能驾驶中的研究及应用的技术分享报告。石博士的报告技术内涵丰富,深入分析了像素级预测问题的若干核心技

术难点,包括:数据准备、模型压缩、局部结构保持、序列模型、任务相关的损失函数设计等。报告展示了相关前沿技术在智能手机及智能驾驶中的实际应用效果,引导听众真切感受到了这一领域的快速发展。

随后,商汤科技产品总监**车英慧**博士为听众带来了论黑科技如何接地气的生动报告。报告深入介绍了商汤科技具有代表性的几项落地应用成果,比较有趣的环节是,针对今年 3.15 晚会上某人脸识别系统被假脸攻破的案例,给出了商汤科技基于活体检测技术的有效解决方案。通过这些实际案例展示,大家对商汤科技的技术和产品纷纷给予了高度认可与点赞。报告之后的问答环节,大家踊跃提问交流,几位主讲嘉宾均进行了耐心细致的回答,现场气氛非常活跃。

活动间歇,商汤科技工作人员为参会人员精心准备了点心、水果等,还赠送了参会人员精美小礼品,大家边讨论交流边观摩商汤科技的各项精选 demo 演示,充分体会商汤科技产品的独特魅力!



(责任编辑:王瑞平)

信息速递

走进企业系列交流会活动细则（新版）

1、宗旨

为了促进国内计算机视觉学术界与工业界之间的高效交流，搭建企业需求与高校院所供给之间的便捷沟通平台，加快研究机构自主知识产权的技术转移和产品转化，促成计算机视觉领域产学研链条的有效对接和通畅运转，中国计算机学会计算机视觉专委会（CCF-CV）在全国范围的高校科研院所与企业间开展 CCF-CV 走进企业系列交流会活动（以下简称“活动”）。

2、活动形式

活动由 CCF-CV 主办，CCF-CV 秘书处负责协调和指导工作；

每次活动由一家企业承办，活动发起方式有两种：1）由意向企业向秘书处提出开放展示与宣传交流的需求，秘书处向专委会委员广播企业需求以征集相关方向的委员访问意向；2）由专委会委员联名向秘书处提出访问/了解某一感兴趣企业的需求，秘书处在汇总各类需求后与相关企业沟通联系以确定企业的受访意向。上述两种方式均由秘书处与企业方协商后，最终确定受邀参加交流的访问人员。

企业负责安排交流会的场地，并与秘书处商定交流会日程；秘书处负责后续新闻稿撰写、网站更新等事宜；

每次活动根据企业意向邀请适当数量、相关方向的高校与科研院所研究人员，活动内容包括企业技术产品展示、访问人员产品体验、学界专家学术报告、潜在合作方向的互动讨论等；

每月原则上组织 1 次活动。

3、活动申请

活动采取自愿申请的方式，按照上述两种发起方式，有意承办活动的企业、有意访问感兴趣企业的委员均可向秘书处提交申请，专委会鼓励中小企业、初创企业积极承办，申请表格见后附件；如果发起方为专委会委员，则要求联名申请人数不少于 5 人，联名人可以是非专委会委员，意向访问企业不多于 3 家；

活动发起方需提前至少一个月向秘书处提出申请（企业发起方/委员发起方分别填写附件一、二申请表，并发至王瑞平 wangrui ping@ict.ac.cn，并抄送李实英 lishiying66@gmail.com）。

4、活动权益

专委会鼓励委员们积极参加 CCF-CV 走进企业交流活动，并特别倡导专委会委员作为活动发起方（见附件二）主动发起活动申请及联系意向企业，从而为委员们提供更为精准需求的企业交流与合作机会；同时，委员参加/组织活动的活跃度也将作为委员年度贡献的参考之一；

对于承办活动的企业方，专委会将在后续相关活动与宣传中给予优先考虑，包括：

- 1）在专委会委员年度增选中，优先发展企业方组织者作为专委会企业委员；
- 2）在专委会双月刊简报中开设“企业采风”栏目，优先宣传推广承办活动企业方（每期两家企业）；
- 3）在专委会网站和简报中，优先发布相关企业方的招聘信息、技术动态等资讯；
- 4）在专委会年度学术会议如中国计算机视觉大会 CCCV、计算机视觉研究与应用创新论坛

RACV 等重要活动中，优先为承办活动企业方提供宣传展示及展区布置，并为其参加展览等提供优惠价格。

5、注意事项

企业承办方需承担活动场地等相关费用。对于有条件的企业方，可在上述费用基础上，基于企业自愿原则，额外承担访问人员活动期间的餐饮费用。原则上访问人员自己承担交通和住宿费用。

活动开始至少半个月之前，秘书处与企业承办方确定参加交流会人员、交流日程、场地等事宜。秘书处制作活动宣传材料，并在专委会网站上发通知。

活动结束后一周内，秘书处将交流会的活动整理成新闻稿，上传至专委会网站。

该活动以技术交流、合作洽谈为主要目的，严禁以活动名义进行非法集会，传播非技术内容；如果有赞助等商业活动，须与秘书处提前商定。

CCF-CV 走进企业系列交流会申请表（企业发起方）和 CCF-CV 走进企业系列交流会申请表（委员发起方）详见：

<http://ccfcv.ccf.org.cn/ccfcv%E8%B5%B0%E8%BF%9B%E4%BC%81%E4%B8%9A/>

（责任编辑：毋立芳）

CCCV2017 讲习班讲者征集

讲习班主席：**林宙辰** (北京大学)，
耿新 (东南大学)

征集截止日期：2017 年 6 月 30 日
结果通知日期：2017 年 7 月 15 日

中国计算机视觉大会 (The Chinese Conference on Computer Vision) 是由中国计算机学会 (CCF) 主办，计算机视觉专委会承办的计算机视觉领域盛会，旨在为计算机视觉研究的学生、老师和工业界人员提供一个学科互动交流平台，促进学术交流以及学术界与工业界之间的交流，提高国内视觉领域的研究水平。该会议是国内计算机视觉领域最主要的学术活动之一，大会将邀请国内外知名的专家、学者做主题报告，介绍计算机视觉领域最新、最热的理论和方法，使与会者了解学科最前沿的学术动态，分享计算机视觉领域的最新研究成果，创新思想和科学方法。

本届中国计算机视觉大会 (CCCV2017) 将于 2017 年 10 月 12-14 日在天津举行，由中国民航大学承办，天津大学和南开大学联合协办。

现向广大科技工作者公开征集讲习班建议书。CCCV 讲习班的目标是就**某个特定的计算机视觉相关领域进行比较全面和系统的介绍**。我们期望一个理想的讲习班能够具备以下特征：

- 1、类似一门短小精悍的课程，而不是一个浮光掠影的技术综述；
- 2、涉及主题应对计算机视觉领域的科技工作者有足够的覆盖面和吸引力；
- 3、深浅难度适合讲习班的大部分潜在听众，即计算机视觉领域的研究生以及来自学术界和工业界的专家学者；
- 4、总时长约为 3 个小时，如多人组合讲演可延长至 4 小时，中间每小时可安排

10-15 分钟休息时间。

请将讲习班建议书于**2017 年 6 月 30 日前发至 zclin2000@hotmail.com 和 xgeng@seu.edu.cn**。建议书应包括以下内容：

- 1、讲习班题目；
- 2、所有讲者的简介和近照；
- 3、讲习班内容简介；
- 4、讲习班提纲及每部分的大致时间安排；
- 5、讲习班内容涉及的参考文献或其他外部资料；
- 6、如该讲习班已在其他会议讲过，请列出会议名称 (含年份)。

如有任何进一步的问题，欢迎通过 email 咨询：zclin2000@hotmail.com，xgeng@seu.edu.cn。

CCCV2017 赞助商信息列表 (截止目前)

钻石赞助商：天津艾思科尔科技有限公司、北京旷视科技有限公司、北京市商汤科技开发有限公司

铂金赞助商：亮风台 (上海) 信息科技有限公司、广州图普网络科技有限公司

金牌赞助商：英伟达公司、合肥寰景信息技术有限公司、平安科技 (深圳) 有限公司、赛灵思公司

银牌赞助商：北京数字绿土科技有限公司、北京智慧眼科技发展有限公司

铜牌赞助商：银河水滴科技 (北京) 有限公司、上海图漾信息科技有限公司、亚马逊公司、北京云识图信息技术有限公司、上海中科智谷人工智能工业研究院

CCCV2017 特邀讲者

Katsushi Ikeuchi (IJCV 主编, Principal Researcher, Microsoft Research Asia; Emeritus Professor, University of Tokyo)

Marc Pollefeys (Full Professor,

Computer Vision and Geometry lab (CVG), Institute of Visual Computing Department of Computer Science, ETH Zurich)

Sven Dickinson (TPAMI 主编,

Professor, University of Toronto Department of Computer Science)

(责任编辑：毋立芳)

CCCV2017 技术挑战赛

基于 FPGA 的视觉计算

本次赛事由中国计算机学会计算机视觉专委会 (CCF CV) 主办, 赛灵思 (Xilinx) 公司和银河水滴科技 (北京) 有限公司承办。

1、竞赛题目与规则：

本次竞赛要求所有参赛队伍选用基于 Xilinx 的任何一款 FPGA 完成以下两项竞赛中的任何一项或两项, 并采用现场比赛的方式决出最后名次, 具体题目和规则如下：

(1) 面向工业视觉数据的多分类任务。参赛队伍事先可获取标注好的工业表面材料图像训练数据, 通过设计基于 FPGA 的学习算法, 判断该工业表面材料的缺陷种类。计分方式为正确分类一张图片得 1 分。在比赛现场, 各参赛队伍从处理第一张测试图片开始计算时间, 根据先达到指定得分数 (比如 100 分) 的用时来判断成绩。

(2) 面向监控场景的图像中行人检测任务。参赛队伍自行收集训练数据, 通过设计基于 FPGA 的学习算法, 在比赛现场检测出给定测试图像中所有的行人。大赛组织方将公布测试样例。计分方式为正确检测一张图片里所有目标 (与真实目标框重叠率超过 50%) 并且没有误检得 1 分。在比赛现场, 各参赛队伍从处理第一个测试图像开始计算时间, 根据先达到指定得分数 (比如 100 分) 的用时来判断成绩。

参赛者可向赛灵思 (Xilinx) 公司申请租借最新的 Ultrascale + MPSOC (A53 + GPU + FPGA) 开发套件和 Zynq7020 含摄

| 时间 | 任务 |
|----------------------|---|
| 4 月中旬 | 组委会在 CCCV 网站公布赛题、参赛办法和评比方式 |
| 4 月中旬至 7 月中旬 | 大赛宣传阶段, 参赛队伍报名及资格确认。 |
| 7 月下旬 | 组委会公布各赛题的详细介绍、训练数据 (如有) 以及测试样例 (如有) 等。 |
| 7 月下旬至 9 月下旬 | 参赛队伍根据实际情况参加组委会组织的线上线下培训 (赛灵思 FPGA 的使用等)。 |
| 10 月 12 日至 10 月 14 日 | 各队伍现场参赛: 在现场对给出的测试数据进行判断, 一决高下。组委会汇总各参赛队的竞赛结果及排名。在中国计算机视觉大会召开研讨会介绍比赛方法, 评选最具潜力奖, 为所有获奖队伍颁奖。 |

像头套件。如果参赛队伍较多借用平台不够, 可用大赛折扣专价给参赛队伍提供板卡。

2、时间安排

比赛组织的时间安排如表

3、奖项设置：

每个比赛任务分别设

(1) 一等奖 1 名 (1 万元奖金以及赛灵思 Ultrascale + MPSOC (A53 + GPU + FPGA) 开发套件)

(2) 二等奖 2 名 (赛灵思 Ultrascale+ MPSOC(A53+GPU+FPGA) 开发套件)

(3) 最具潜力奖 1 名 (奖品待定, 根据竞赛测试结果和研讨会上的技术报告由评委综合评定, 用于奖励算法具有创新性和前沿性的团队。)

(4) 获奖团队成员优先获得赛灵思 Xilinx 及其合作伙伴的实习及就业机会 银河水滴科技的实习和就业机会

参赛队伍请发送报名表至邮箱：
dong.wang@watrix.cc, 具体信息可以参考

CCCV2017 竞赛单元网站！

<http://ccf-cccv.org/2017/contest1/>

心血管 OCT 易损斑块识别

心血管 OCT 易损斑块识别技术挑战赛竞赛简介

冠心病已成为全人类死亡的第一大原因。在美国, 冠心病的死亡率已经超过所有癌症死亡率的总和, 成为第一位致死的原因; 在我国, 心血管病死亡不仅占总死亡原因的首位, 而且发病人数在逐年增加, 严重威胁着我国居民的健康。因此, 关于研究如何诊断和治疗冠心病的相关技术和方法具有重大意义和价值。

易损斑块是冠状动脉粥样硬化病变当中最危险的斑块, 易损斑块是造成血栓、引发急性冠脉综合症、甚至导致猝死的主要原因, 因此, 检测识别易损斑块具有非常高的临床价值。心血管光学相干断层成像 (IVOCT) 技术可以清晰的观察到易损斑块, 使得基于 IVOCT 影像的易损斑块识别已成为可能, 但是, 由于 IVOCT 采集到的图像序列数据量大, 给医生人工分析图

像带来巨大困难。

本次“心血管 OCT 易损斑块识别技术挑战赛”的意义就在于研究如何利用计算机图像自动识别技术，自动识别 IVOCT 序列中易损斑块的位置区域，从而推动计算机视觉学科与医疗事业的发展，促进智能图像识别技术在心血管 OCT 的临床应用。

本次竞赛是由“中国人民解放军总医院”与“中国科学院”联合组织发起，面向的参赛者包括：国内外各研究团体、企事业单位，各培养单位正式注册教师、在读研究生以及博士生。

1、参赛方式：

以个人或团队方式均可通过邮件方式报名参赛，每个参赛队伍人员不超过 5 人，每名参赛选手只能参加 1 个参赛队。参赛队发送报名信息至 yihui.cao@opt.ac.cn 邮件标题格式“心血管 OCT 易损斑块识别技术挑战赛 + (参赛队名称)”，邮件内容应包括个人或团队基本信息、负责人基本信息、参赛人数、联系方式(手机，邮箱)。组织方收到邮件并与报名者确认后，报名成功。报名截止日期为：2017 年 05 月 30 日

2、赛程如表：

| 时间 | 任务 |
|----------------------|--|
| 3 月 15 日至 3 月 31 日 | 组委会在网站公布比赛任务、参赛办法和评比方式。 |
| 4 月 1 日至 6 月 15 日 | 参赛队伍报名及邮件资格确认。 |
| 6 月 16 日至 9 月 15 日 | 组委会在虚拟机上完成训练样本和验证样本的准备，并对参赛队伍公布,各参赛队通过互联网及参赛账号登录虚拟机，可以在虚拟机上传各自的算法进行调试，也可在参赛队自身的软硬件环境中进行算法调试。 |
| 9 月 15 日至 9 月 30 日 | 比赛开始。各参赛队在规定时间内完成比赛任务，将结果按照规定的格式上传至指定位置。每个参赛队只能提交一次结果参与评测（如果需要修改，请在规定时间内联系组委会进行修改）。 |
| 10 月 1 日至 10 月 15 日 | 组委会公布比赛的测试数据集，评委按照规定的评测方法，验证各队程序执行结果，对结果进行评测，评比各队名次。 |
| 10 月 12 日至 10 月 14 日 | 在中国计算机视觉大会期间,召开 workshop 介绍比赛方法，颁奖。 |

3、奖项设置：

本次竞赛设一等奖 1 名(奖金 3 万元人民币)，二等奖 1 名(奖金 2 万元人民币)，三等奖 2 名(奖金各 1 万元人民币)。

本次比赛奖品由深圳市中科微光医疗器械技术有限公司赞助。

各参赛队伍根据比赛成绩均可获得由计算机视觉专委会、中科院西安光学精密机械研究所、中国人民解放军总医院共同颁发的获奖证书和奖金。

具体信息可参考 CCCV2017 竞赛单元网站 <http://ccf-cccv.org/2017/competition3/>

程序，得到并验证评价图像识别结果，判定有效的图像识别精度，评出名次。

遥感图像目标提取

1、比赛目的与意义

提高智能图像识别分析技术的研究水平，推动计算机视觉学科及相关产业的发展，促进智能图像识别技术在遥感中的应用。

2、竞赛参与者要求

(1) 各培养单位正式注册教师、在读研究生以及博士生。

(2) 国内各研究团体、企事业单位。

3、任务设置

遥感图像目标提取：检测遥感图像内的所有符合检测要求的对象(飞机、船舶)，标出每类对象的类型和中心点 XY 位置。

主办方利用测试数据集，在以下测试环境中，通过运行参赛队伍提交的可执行

测试环境：

CPU：4 核 Intel E5-2609

内存：16G

硬盘：200G

操作系统：Windows 7 X64

4、奖项设置和奖励方法

每项任务设一等奖 1 名(奖金 3 万元人民币)，二等奖 1 名(奖金 2 万元人民币)，三等奖 2 名(奖金各 1 万元人民币)。

各参赛队伍根据比赛成绩均可获得由计算机视觉专委会和北京航天宏图信息技术股份有限公司颁发的获奖证书和奖金。

5、报名方式

以个人或团队方式均可通过邮件方式报名参赛，每个参赛队伍人员不超过 5 人，每名参赛选手只能参加 1 个参赛队。参赛队发送报名信息至 zhuliya@picsat.cn。

邮件标题格式“遥感图像目标提取技术挑战赛+(参赛队名称)”

邮件内容应包括个人或团队基本信息、负责人基本信息、参赛人数、联系方式(手机, 邮箱)。

组织方收到邮件并与报名者确认后, 报名成功。

6、主要时间节点如表

更多挑战赛信息请参考 CCCV2017 竞赛单元网站 <http://ccf-cccw.org/2017/competition2/>

(责任编辑: 毋立芳)

| 时间 | 任务 |
|---------------|--|
| 4月3日-4月5日 | 组委会在网站公布比赛任务、参赛办法和评比方式。 |
| 4月6日-6月30日 | 参赛队伍报名及邮件资格确认。 |
| 4月24日-4月30日 | 组委会完成训练样本和验证样本的准备, 并对参赛队伍公布, 各参赛队通过互联网下载样本数据, 然后在参赛队自身的软硬件环境中进行算法调试。 |
| 5月1日-9月10日 | 各参赛队在规定时间内(截止时间9月4日23点59分)完成比赛任务, 将结果按照规定的格式上传至指定位置。每个参赛队只能提交一次结果参与评测(如果需要修改, 请在规定时间内联系组委会进行修改)。 |
| 9月11日-9月22日 | 组委会公布比赛的测试数据集, 评委按照规定的评测方法, 验证各队程序执行结果, 对结果进行评测, 评比各队名次。 |
| 10月12日-10月14日 | 在中国计算机视觉大会期间, 召开workshop介绍比赛方法, 颁奖。 |

CCCV2017 专题论坛征集

专题主席: **张兆翔**(中科院自动化所)
纪荣嵘(厦门大学)

征集截止日期: 2017年7月30日

结果通知日期: 2017年8月15日

中国计算机视觉大会(The Chinese Conference on Computer Vision)是由中国计算机学会(CCF)主办, 计算机视觉专委会承办的计算机视觉领域盛会, 旨在为从事计算机视觉研究的学生、老师和工业界人员提供一个学科互动交流的平台, 促进领域内学术交流以及学术界与工业界之间的交流, 提高国内视觉领域的研究水平。该会议是国内计算机视觉领域最主要的学术活动之一, 大会将邀请国内外知名的专家、学者做主题报告, 介绍计算机视觉领域最新、最热的理论和方法,

使与会者了解学科最前沿的学术动态, 分享计算机视觉领域的最新研究成果, 创新思想和科学方法。

本届中国计算机视觉大会(CCCV2017)将于2017年10月12-14日在天津举行, 由中国民航大学承办, 天津大学和南开大学联合协办。本届会议将汇聚国内从事计算机视觉理论与应用研究的广大科研工作者, 共同分享我国在计算机视觉领域的最新理论和技术成果, 为大家提供精彩的学术盛宴。

现向广大科技工作者公开征集专题讲座, 其目的是围绕计算机视觉理论、方法、应用乃至产业的特定话题, 召集多个专家, 从多个角度进行报告、研讨, 通过互动强化相关话题的关注度, 促进相关技术的发展。

一个比较理想的专题应该具有如下特点:

- 1、主题明确, 对计算机视觉领域的科技工作者有吸引力
- 2、有3-5名该主题方向的从事者负责组织
- 3、建议采用报告+Panel讨论的形式(如4个报告+1个Panel)

总时长3个小时。建议书应包括以下内容:

- 1、专题题目;
- 2、所有讲者的简介和近照;
- 3、专题日程介绍;

请将主题申请建议书于**2017年7月30日前**发给 zhaoxiang.zhang@ia.ac.cn 和 rrji@xmu.edu.cn 和 zs zhang@nlpr.ac.cn。

(责任编辑: 毋立芳)

张量稀疏性度量

西安交通大学 孟德宇 谢琦

一、引言

稀疏性是现实数据所共有的一般信息表达特性，其含义为数据可由其本质蕴含的少量基元素进行充分表达。目前对于向量(1阶数组)与矩阵(2阶数组)数据均存在较为成熟的稀疏性表达度量。然而，对于张量(高阶数组)数据的合理稀疏性度量的构造，仍然尚未形成统一的解决方案。本文将对张量稀疏性研究的现状进行综合介绍，回顾目前在此方向的研究进展以及所取得的典型应用，并着重介绍本研究小组所构造的一种新型张量稀疏性度量。相比传统度量，所提出的张量稀疏性度量不仅充分编码了对于张量两种经典分解方式(Tucker分解与CP分解，图1)的张量稀疏性内涵，而且具有与传统的向量/矩阵稀疏性度量的统一性，并且蕴含显著的物理含义(张量表达的1秩Kronecker基个数)。我们也将介绍此张量稀疏性度量在多光谱图像去噪、填充等问题中获得的state-of-the-art应用表现，从而说明其合理性与有效性。

二、数据的稀疏先验

从通常意义上理解，先验是一个与经验相对的概念。它先于经验，却对经验的形成有着很大的影响，甚至有时它是经验形成不可或缺的部分。在机器学习问题中，如果观测数据中能提供的经验信息不是非常充分，或是观测数据中蕴含的经验信息难以利用时，先验知识的指导作用将变得尤为关键。数学上看，先验知识的作用可

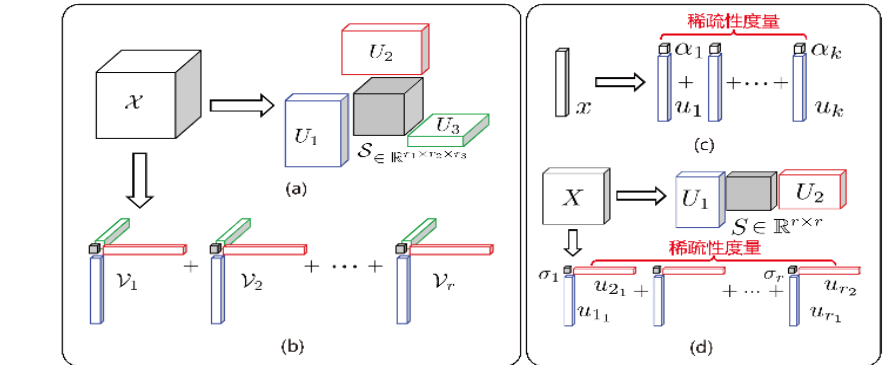


图1 几种分解的展示:(a) Tucker分解;(b) CP分解;(c) 一维情形的分解:向量基表达;(d) 二维情形分解:SVD的两种形式。

以通过模型中的正则函数体现。当损失函数对应的模型是欠定问题或近似欠定问题时，它可以引导模型筛选得到合理的解。举如下的数据恢复模型为例：

$$\min_{x \in \Omega} L(x, y) + \lambda S(x) \quad (1)$$

该模型中， x 代表了待恢复的数据，是模型要求解的优化变量； y 代表了观测数据，是模型的输入信息。相比隐含的真实数据，观测数据往往由于下采样，缺失，噪声污染等原因受到破坏。 Ω 是 x 的可行域，由实际数据的约束条件决定； $L(x, y)$ 代表损失函数，意在表达观测数据与真实数据的差异； $S(x)$ 为正则函数，代表对数据的先验认识表达； λ 是损失函数与正则函数之间的权重参数。对于模型(1)是否能够实现有效的数据恢复， $S(x)$ 往往具有极为重要的作用。当 $\min_{x \in \Omega} L(x, y)$ 没有唯一解或没有绝对优势解时(除了唯一解以外，可行域中还有很多点使其对应的目标函数值虽然比最优值大，但是大得不多)， $\min_{x \in \Omega} L(x, y)$ 的最优解作为数据的恢复结果就不会有太高的可靠性，这种情况下先验知识(正则函数)的充分利用则非常关

键，它可以帮助我们筛选出最符合数据先验的有效解。在数据恢复问题中，常用的先验往往根据待恢复数据的一般化属性制定。例如，对于图像去噪问题，“自然图像是比较平滑的”这一属性常被表达为TV范数形式的正则(全变分正则)形式，对问题的有效求解一般起到非常重要的作用。

稀疏性是一种很多现实数据都共有的一般化属性，因此它作为一种常见的数据先验在大量应用问题中得到了广泛的关注。

简单来说，稀疏性是指现实数据可被少数基元素进行充分表达的特性。很多类型的应用数据本身就具有稀疏性特点，而也有不少数据通过某种变换，也具备较为显著的稀疏性特征。依旧举图像去噪问题为例，当我们用TV范数刻画“自然图像是比较平滑的”这一属性时，其实了可以等价理解为把这一属性近似地用“自然图像数据在差商变换后将具有较强稀疏性”这一易于表达的属性进替换。这意味着经过了差商变换的自然图像，具备了稀疏性的特点。

由于现实数据广泛具有的稀疏性特征, 如何使用量化方式有效刻画数据稀疏性度量, 即如何构造有效刻画数据稀疏性的正则函数形式, 是机器学习、数据挖掘、模式识别等诸多领域关注的问题。

对于具有向量(1阶数组)与矩阵(2阶数组)形式的数据, 已存在较为成熟的稀疏性表达度量。本文将向量意义下的稀疏性也称为1阶稀疏性, 对应的稀疏性度量称为1阶稀疏性度量; 在矩阵意义下, 低秩性可以理解为表达矩阵所需的最少基底的数量, 这正满足我们对稀疏性的定义。出于一致性的考虑, 本文将矩阵的低秩性也称为2阶稀疏性, 对应的稀疏性度量称为2阶稀疏性度量。目前, 1阶稀疏性度量与2阶稀疏性度量均存在受到广泛认可的数学表达形式, 我们一般用向量形式数据¹中的非零元素个数(即 l_0 范数)作为1阶稀疏性度量, 用矩阵形式数据的数据矩阵秩作为2阶稀疏性度量。这两种稀疏性度量及其松弛形式(如, l_1 范数与核范数)都展示出对稀疏性良好的刻画能力, 延伸出一系列基于稀疏或低秩的数学模型[1-8]并已获得广泛应用。

三、高阶稀疏性度量的研究现状

随着数据获取技术的飞速发展, 用以描述数据的特征结构越来越丰富与复杂, 更多问题中数据往往呈现高阶数组, 即张量的表达格式。此时, 机器学习更适宜在张量上直接操作, 而非将其简化为向量或矩阵的格式, 从而避免有用结构信息的损失。因此, 一个自然的问题就是如何将1阶与2阶稀疏性度量向高阶张量做有效推广。相比1阶与2阶情形, 高阶张量具有更加复杂的代数性态与物理特性, 对其构造有效稀疏性度量也具有更大的难度,

对该问题目前仍未形成学界公认的统一解决方案。

CP秩是一种典型的高阶稀疏度量[9], 其定义为组成目标张量的最少秩1张量²的个数。CP秩作为高阶稀疏性度量的优点是其与传统的1阶与2阶稀疏性度量一致, 并且定义十分简单直观。然而, CP秩的计算是一个NP难的问题, 且很难构造方便计算的松弛形式, 其计算的困难性导致其很难有效应用于真实的稀疏建模问题中。

另一种典型的高阶稀疏性度量是Tucker秩[10], 其定义为一个由张量各个方向的展开矩阵的秩组成的数组(其长度为张量的维度数目)。因为Tucker秩是基于矩阵秩定义的, 所以其计算较为方便, 往往可直接基于矩阵相关的计算技术直接移植改造。但Tucker秩是一个数组而非实数, 所以并不合适直接用来作为模型的优化目标。可采用如下简单直观的方法将其量化[11]:

$$S(X) = \sum_{i=1}^N w_i \text{rank}(X_{(i)}) \quad (2)$$

其中, X 是一个 N 阶张量, $X_{(i)}$ 是 X 沿第 i 个维度方向的展开矩阵, 其中 w_i 是一个需要提前确定的权重参数, 一般设为 $1/N$ 。这种形式的稀疏性度量因为容易构造便于计算的松弛形式(例如, $\sum_{i=1}^N \|X_{(i)}\|$), 因而被尝试应用到不少实际问题中[11-16]。然而, 该张量稀疏性度量仍存在一定程度的可改进之处: 1) 需要事先给定权重 w_i , 而合理的 w_i 往往并不容易确定。最简单的作法是使用相同的权重, 但一般张量的各方向展开矩阵的秩是不同的, 因此用相同的权重将它们组合在一起往往不是十分合理; 2) 这种稀疏性度量是基于张量沿各个方向的展开矩阵的性质定义的, 但张量沿任意方向的展开都会本质破坏其沿其它方向的结构信息, 这导致一些重要的张量的

结构信息很难通过各个方向的展开矩阵的简单组合得以充分体现(例如CP秩信息)。

张量奇异值分解(t -SVD)是近期提出的一种新型张量分解形式[17]。它的数学形式简洁, 具备优良的数学性质, 且与传统的1阶与2阶稀疏性度量具有很好的一致性。从形式上看, 其具有与矩阵奇异值分解较为类似的优美形式。[17]同时提出一种基于 t -SVD的高阶稀疏性度量(t -核范数³), 并将其应用到张量恢复中, 验证了 t -核范数的有效性。[18]进一步提出了 t -核范数在张量的稳健主成分分析中的可恢复性理论。然而, 尽管 t -核范数有一些优良的理论性质, 但它不具备CP秩和Tucker秩那样清晰的物理意义。例如, 对于 $n_1 \times n_2 \times n_3$ 的三阶张量 X , 基于 t -SVD的高阶稀疏性度量定义为

$$S(X) = \frac{1}{n_3} \sum_{i=1}^{n_3} \|\bar{x}^{(i)}\|_* \quad (3)$$

其中 $\bar{x}^{(i)}$ 是 X 沿第3个维度的一维的傅里叶变换 \bar{x} 的第 i 层 $n_1 \times n_2$ 矩阵, 其实质是在刻画原数据各层经过傅里叶变换后的低秩加和。因此, 简单来说, t -核范数的实质可理解为对原张量沿某一维度进行傅里叶变换后, 其他维度所呈现的低秩性组合。相比(2), 其优点为可通过傅里叶变换增强某一维度的稀疏性表现, 但与(2)类似, 其仍无法充分刻画细致的张量结构信息。此外, (3)中的傅里叶变换要求对模型的推广性有一定的影响, 例如, 当傅里叶变换并不能增强 X 某一维度的稀疏性, 而该 t -核范数却采用了这一维度进行变换时, 该张量稀疏性可能不仅不能有效的辅助获得更好的计算表现, 反而对模型性能具有一定的负面作用。

¹ 为了方便 这里将向量化后不会引起信息损失的数据的形式都称为向量形式

² 由 N 个向量外积得到的张量, 称为秩1的 N 阶量

³ 这里只以被应用到实际实验的一种松弛形式为例, 实际上基于 t -SVD的高阶稀疏度量还有更为丰富的理论内涵

因此，现有的高阶张量稀疏性度量仍有一定的缺陷，存在提高的空间。针对这一问题，我们将对高阶张量稀疏性进行进一步的探讨，并尝试给出一种新的高阶稀疏性度量。

四、一种新的高阶稀疏性度量

在给出新的稀疏性度量之前，首先通过 CP 分解与 Tucker 分解对高阶稀疏性的本质进行简单的探讨。

如图 1 (b) 所示，CP 分解[9]将张量 $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ 分解为 r 个秩 1 张量的线性组合：

$$\mathcal{X} = \sum_{i=1}^r c_i \mathcal{V}_i = \sum_{i=1}^r c_i u_{i1} \circ \dots \circ u_{iN}$$

其中， \mathcal{V}_i 是 u_{ij} ($j = 1, 2, \dots, N$) 外积得到的秩 1 张量。秩 1 张量亦称为 Kronecker 基。CP 秩 r 即为以这种形式对 \mathcal{X} 进行分解所需要的最少 Kronecker 基个数。CP 秩作为高阶稀疏性度量的定义的优点在于它与传统 1 阶和 2 阶稀疏性度量的定义在形式上有很好的 consistency，这是因为 CP 分解在数据为 1 阶/2 阶形式时可以自然地退化为传统的向量基表达/SVD 分解⁴ (如图 1 (c) 与 (d) 所示)。然而，除了计算较为困难外，CP 秩还存在一个较大的问题，即不能有效刻画 \mathcal{X} 沿各个维度展开矩阵的低秩性。例如，在极端的情况下，当 CP 秩对应 Kronecker 基每个 u_{ij} 在每个维度都呈现不相关的情形中 (即核张量呈现近似对角形态)，尽管 Kronecker 基非常少，但其表达的张量在每个维度展开获得的矩阵均为几乎满秩 (或高秩) 矩阵。此时，CP 秩表达的稀疏性并不能对应于沿张量每个维度展开形成子空间的稀疏性，往往与实际问题不符。

⁴ SVD 分解也可以写成一系列秩一矩阵加和的形式。

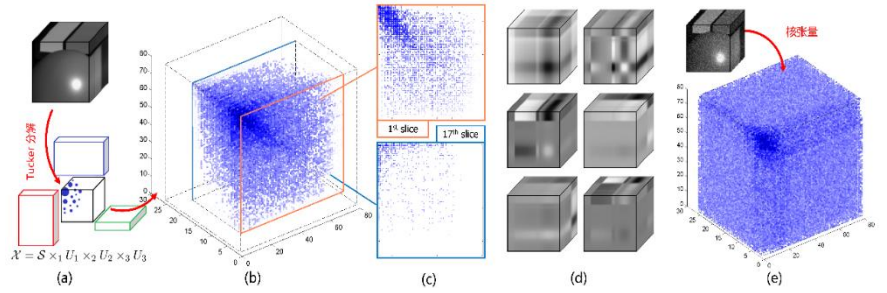


图 2 (a) 一个表示 3 阶张量的多光谱图数据及其“满秩”Tucker 分解；(b) 核张量的可视化显示，颜色越深代表数值越大，非零元素集中在左上角 $69 \times 71 \times 17$ 的方块区域内，且 78.4% 的元素都是零元素；(c) 核张量的第一和第 17 层展示；(d) 与核张量中元素绝对值最大的 6 个元素相对应的 Kronecker 基；(e) 轻微带噪图像及其“满秩”Tucker 分解的核张量，大部分元素都是非零元素。

另一方面如图 1 (a) 所示，Tucker 分解[10]可将张量 $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ 分解 N 个正交矩阵与一个核张量的乘积：

$$\mathcal{X} = \mathcal{S} \times_1 U_1 \times_2 \dots \times_N U_N$$

其中， $\mathcal{S} \in \mathbb{R}^{R_1 \times R_2 \times \dots \times R_N}$ ， $r_i \leq R_i \leq I_i$ 。

当我们对基矩阵 U_i 中的列合理排序时， \mathcal{S} 中的非零元将只分布在左上角 $r_1 \times r_2 \times \dots \times r_N$ 大小的方块区域内。此时，可以看到，Tucker 秩 (r_1, r_2, \dots, r_N) 即核张量中方块非零区域的各边大小。然而，如图 2 所示，实际的 3 阶数据的核张量中，非零元元素沿各层呈现较为复杂的变化形态，而 Tucker 秩并不能有效刻画这样的非零分布差异性，而仅提供了能够包含所有非零分布的一个张量块的大小，提供了一个较为粗糙的核张量非零元个数上界。

从上面的分析中我们可以观察到，无论 CP 秩还是 Tucker 秩对张量的稀疏性的刻画都不全面，而他们刻画的内容一定程度上是互补的。事实上，Tucker 分解也可以写成秩 1 矩阵加和的形式，

$$\mathcal{X} = \mathcal{S} \times_1 U_1 \times_2 \dots \times_N U_N = \sum_{i_1, \dots, i_N} \mathcal{S}_{i_1, \dots, i_N} u_{i_1, \dots, i_N} \quad (4)$$

其中 u_{i_1, \dots, i_N} 为所有 U_j 的第 i_j 列的积 ($j = 1, \dots, N$) 得到的秩 1 张量。 \mathcal{S} 中非零元素的个数即代表表达张量需要的 Kronecker 基个数 (与 CP 分解中的 Kronecker 基的不同在于这里的 Kronecker 基是由各基矩阵 U_j

的列外积而来的)，其提供了 CP 秩的一个上界，由此，我们可以看出 Tucker 分解的核张量提供了一个 CP 秩的约束。

基于以上的认识，我们认为核张量中与高阶稀疏度相关的信息主要可以归纳为以下两点

- 1) 细描述: 核张量的非零元素的数量 (细致刻画核张量中非零元分布)
- 2) 粗约束: 核张量中方块非零区域的大小 (提供核张量非零元个数的上界)

其中第 2 点也是 Tucker 秩以及秩和形式(2)稀疏性度量的刻画内容。

基于上述对张量稀疏性的认识，我们提出了如下的高阶稀疏性度量[19]:

$$S(\mathcal{X}) = t \|\mathcal{S}\|_0 + (1 - t) \prod_{i=1}^N \text{rank}(X_{(i)}) \quad (5)$$

其中， $\mathcal{S} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ 是 \mathcal{X} 的 Tucker 分解的核张量。如图 3 所示，第一项刻画了核张量中非零元素的数量，第二项刻画了核张量中非零方块的大小。相比传统张量稀疏性，该度量能够一些独特的优点：

- 1) 所提稀疏性度量具有明确物理意义。其第一项度量了表达张量所需的 Kronecker 基数量，而其第二项刻画了表达张量所需 Kronecker 基数量的一个约束上界。两者的意义是统一的，第一项是对 Kronecker 基数

量的细致刻画，第二项是对该数量的一个正则约束。

2) 所提稀疏性度量与传统 1 阶与 2 阶稀疏性度量内涵具有一致性。传统的 1 阶向量稀疏性可理解为其在特定字典下本质表达原数据的向量基，即 1 阶 Kroneker 基个数；2 阶矩阵稀疏性即为表达原数据所需秩 1 矩阵，即 2 阶 Kroneker 基个数。因此，其物理意义完全统一。而从表达形式上来看，对于 1 阶稀疏性度量 $\|x\|_0$ ，所提稀疏度量为 $S(x) = t\|x\|_0 + (1-t)\|x\|_1$ ，显然，其关于 $\|x\|_0$ 是单调的；对于传统 2 阶稀疏性度量 $\text{rank}(X)$ ，所提稀疏度量为 $S(x) = t\text{rank}(X) + (1-t)\text{rank}(X)^2$ ，显然，它关于 $\text{rank}(X)$ 也是单调的。因此，从对稀疏性的量化刻画上看，与传统稀疏性度量具有一致性。

3) 所提稀疏性度量与加权秩和形式的稀疏性度量(3)是一致的，且它提供了一种对权重的新理解：

命 $w_i = \frac{1}{N} \prod_{j \neq i} \text{rank}(X_{(j)})$ ，则

$$\prod_{i=1}^N \text{rank}(X_{(i)}) = \sum_{i=1}^N w_i \text{rank}(X_{(i)}) \quad (6)$$

这说明，只要合理地先取 w_i ，传统的加权秩和形式的正则作用与所提模型的第二项将十分相似。反过来，(6) 也提供了一

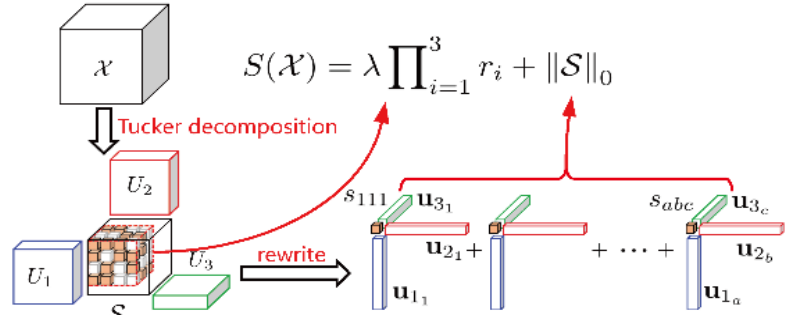


图 3 所提张量稀疏性度量示意图

种 w_i 的选取方式，不难推出这种选取方式在直观是非常合理的：低秩性强的方向要对应更大的权，低秩弱强的方向要对应更小的权。

我们已将所提张量稀疏性度量应用到 3 个经典的张量恢复问题：多光谱图像去噪、填充和背景建模。图 4 展示了所提稀疏性度量应用到多光谱图像去噪[20]时与传统方法结果对比。所提度量在实验中的良好表现也进一步说明了它合理性。

五、总结

本文基于前人的工作，对高阶张量稀疏性度量进行了探讨，认为其本质与张量秩 1 的 Kroneker 基个数紧密相关，并借由张量 Tucker 分解与 CP 分解中对 Kroneker 基相应的内涵理解构造了新型的张量稀疏性度量。相比已有度量，其不仅具有显著

的物理含义，且与传统的 1 阶与 2 阶稀疏性度量具有内在一致性。特别地，其可以提供传统张量稀疏性度量的加权格式理解。在多光谱数据的去噪、填充与视频数据的背景提取问题中，所提张量稀疏性度量均体现了显著的应用有效性。在未来的工作中，我们将研究如何使用更简单的形式对所提的两个性质进行刻画，并进而对所提的稀疏性度量的理论性质进行必要的研究。

参考文献：

- [1] E. J. Candes, X. Li, Y. Ma, and J. Wright. Robust principal component analysis? Journal of the ACM (JACM), 58(3):11, 2011.
- [2] E. J. Candes, M. B. Wakin, and S. P. Boyd. Enhancing sparsity by reweighted l1 minimization. Journal of Fourier analysis and applications, 14(5-6):877–905, 2008.
- [3] F. De La Torre and M. J. Black. A framework for robust subspace learning. IJCV, 54(1-3):117–142, 2003.
- [4] S. Gu, L. Zhang, W. Zuo, and X. Feng. Weighted nuclear norm minimization with application to image denoising. In CVPR, 2014.
- [5] D. Meng and F. Torre. Robust matrix factorization with unknown noise. In ICCV, pages 1337–1344, 2013.
- [6] F. Parvaresh, H. Vikalo, S. Misra, and B. Hassibi. Recovering sparse signals using sparse measurement matrices in compressed dna microarrays. IEEE

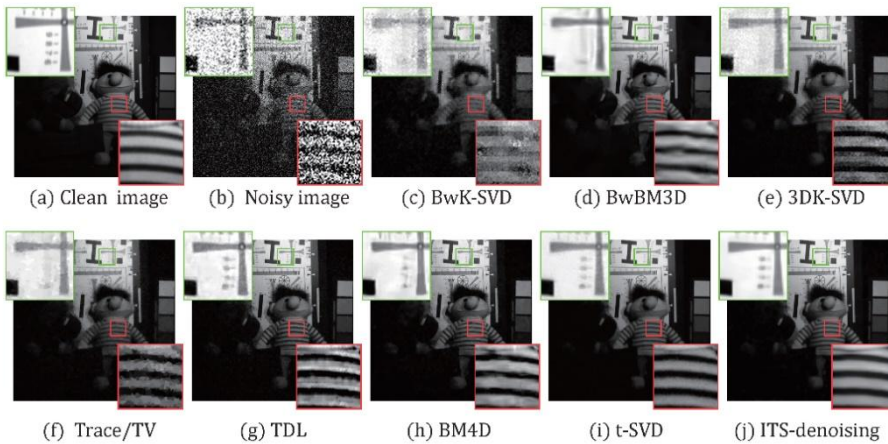
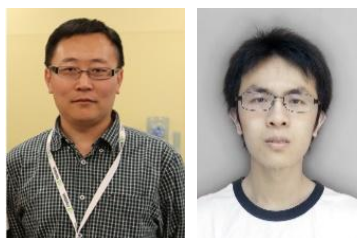


图 4 (a)多光谱图 chart and staffed toy 的两帧；(b) 带方差为 0.2 噪声图片展示；(c)-(f) K-SVD [20]，BM3D[21]，3DK-SVD [22]，LRTV [25]，TDL [24]，BM4D [23]，t-SVD [17]及所提模型应用于去噪 (ITS-denoising [19]) 的结果对比。

- Journal of Selected Topics in Signal Processing, 2(3):275–285, 2008.
- [7] J. Wright, A. Ganesh, S. Rao, Y. Peng, and Y. Ma. Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization. In NIPS, 2009.
- [8] Q. Zhao, D. Meng, Z. Xu, W. Zuo, and L. Zhang. Robust principal component analysis with complex noise. In ICML, pages 55–63, 2014.
- [9] T. G. Kolda and B. W. Bader. Tensor decompositions and applications. SIAM Review, 51(3):455–500, 2009.
- [10] L. R. Tucker. Some mathematical notes on three-mode factor analysis. Psychometrika, 31(3):279–311, 1966.
- [11] J. Liu, P. Musialski, P. Wonka, and J. Ye. Tensor completion for estimating missing values in visual data. TPAMI, 35(1):208–220, 2013.
- [12] S. Gandy, B. Recht, and I. Yamada. Tensor completion and low-n-rank tensor recovery via convex optimization. Inverse Problems, 27(2):025010, 2011.
- [13] R. Tomioka, K. Hayashi, and H. Kashima. Estimation of low-rank tensors via convex optimization. arXiv preprint arXiv:1010.0789, 2010.
- [14] M. Signoretto, Q. T. Dinh, L. De Lathauwer, and J. A. Suykens. Learning with tensors: a framework based on convex optimization and spectral regularization. Machine Learning, 94(3):303–351, 2014.
- [15] B. Huang, C. Mu, D. Goldfarb, and J. Wright. Provable lowrank tensor recovery. Optimization-Online, 4252, 2014.
- [16] W. Cao, Y. Wang, C. Yang, X. Chang, Z. Han, and Z. Xu. Foldedconcave penalization approaches to tensor completion. Neurocomputing, 152:261–273, 2015.
- [17] Z. Zhang, G. Ely, S. Aeron, N. Hao, and M. Kilmer. Novel methods for multilinear data completion and de-noising based on tensorsvd. In CVPR, 2014.
- [18] C. Lu, J. Feng, Y. Chen, W. Liu, Z. Lin, and S. Yan. Tensor robust principal component analysis: Exact recovery of corrupted lowrank tensors via convex optimization. In CVPR 2016.
- [19] Q. Xie, Q. Zhao, D. Meng, Z. Xu, S. Gu, W. Zuo, and L. Zhang. Multispectral images denoising by intrinsic tensor sparsity regularization. In CVPR, 2016.
- [20] M. Aharon. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. IEEE Trans. Signal Processing, 54(11):4311–4322, 2006.
- [21] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. IEEE Trans. Image Processing, 16(8):2080–2095, 2007.
- [22] M. Elad and M. Aharon. Image denoising via sparse and redundant representations over learned dictionaries. IEEE Trans. Image Processing, 15(12):3736–45, 2006.
- [23] M. Maggioni and A. Foi. Nonlocal transform-domain denoising of volumetric data with groupwise adaptive variance estimation. In SPIE, 2012.
- [24] Y. Peng, D. Meng, Z. Xu, C. Gao, Y. Yang, and B. Zhang. Decomposable nonlocal tensor dictionary learning for multispectral image denoising. In CVPR, 2014.
- [25] M. Golbabaee and P. Vandergheynst. Joint trace/tv norm minimization: A new efficient approach for spectral compressive imaging. pages 933–936, 2012.

(责任编辑：苏航)



孟德宇，西安交通大学教授，博导。分别于 2001、2004、2008 年在西安交通大学获学士、硕士和博士学位。共接收/发表 IEEE 汇刊论文 16 篇，CCF A 类会议论文 24 篇。曾获陕西省青年科技奖，陕西省优秀博士学位论文奖，入选首批西安交通大学青年拔尖人才计划。主要研究方向包括自步学习、误差建模、张量稀疏性等。

Email: dymeng@mail.xjtu.edu.cn

谢琦，西安交通大学博士生。主要研究方向包括矩阵分解、张量稀疏性建模、噪声建模、稀疏机器学习。

Email: xq.liwu@stu.xjtu.edu.cn

人物动作中的语义挖掘与分类

清华大学 马惠敏 赵致辰 陈晓智

静态图像中人物的动作分类是指预测单幅图片中人的动作，由于涉及整体观测与局部细节、动作发出与动作受体等，细节的差异即可产生动作类别的差异，因此有很高的难度。

人物动作分类与通用物体分类所不同的是，人与人之间比较类似，人物内部的一些部件之间也比较类似，因此样本具有较小的类间方差。又由于各种姿态的变化，导致样本具有有较大的类内方差，从而给识别任务带来了许多困难。与细粒度物体分类不同，人的部件在不同动作下服从典型的分布，考虑空间信息可以显著地提升算法的性能。

由于动作分类的困难性，传统的解决思路有两种：一种是寻找和人交互的物体，比如“骑马”这个动作都会有一匹马。这样的方式借助了发展更加成熟的物体分类方



图 1 基于语义部件的上下金字塔算法流程

法。另一种方法是针对动作中典型的某些部件的外观进行建模，训练一些小的检测子，作为某种动作存在的判据。

这两种方法都可以从一定程度上解决动作分类的问题，但是也存在一些问题。寻找与人交互的物体或者寻找典型的动作块并不能保证这些线索一定是有帮助的，如果找到的物体与该动作没有紧密的联系可能会影响算法的性能。另外可能有一些其他的有效线索还没有被这两种所涵盖。

基于此，我们提出了语义部件的概念(图1)，

即挖掘图形中能够提供任何有效线索的块，不断地检测，细化，整理。一方面，针对语义部件所训练的检测子来挖掘每种不同动作中能够提供线索的块；另一方面，考虑到部件有特殊的空间分布特性，按照空间顺序排布相关的特征来增大类间方差，如“跳跃”的排布是手-头-腿，而“走路”则是头-手-腿，这样的排布可以帮助我们进一步区分这两种动作，如图2所示。该方法在PASCAL VOC 2012以及Stanford-40两个公开数据集上都取得了目前最好的分类结果。相关成果相继发表于Pattern Recognition、Pattern Recognition Letters、ICIP等国际期刊和国际会议。

(责任编辑：王金甲)

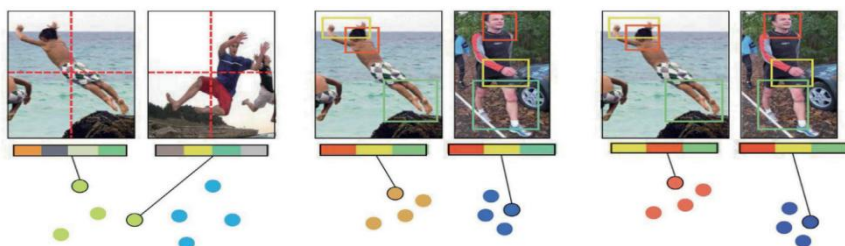


图 2 语义部件形成的空间结构



马惠敏, CCF 计算机视觉专委会委员、清华大学电子工程系副教授、中国图象图形学会秘书长。主要

研究方向为三维目标识别与视觉导航。

Email: mhmpub@mail.tsinghua.edu.cn



赵致辰, 清华大学硕士生, 主要研究方向为物体及动作识别研究。



陈晓智, 清华大学博士生, 主要研究方向为三维物体检测研究。

数据深度聚类分析

新加坡资讯通信研究院 彭玺

人以群分，物以类聚——聚类分析的目的在于根据数据点间的相似性，通过最大化同簇数据之间的相似性和异簇数据之间的相异性，将相似的数据归为同一簇，相异的数据归为不同簇。聚类是无监督学习的一个重要内容，在数据压缩、图像分割、关键点提取、动作分割等领域具有重要意义和价值。

聚类分析的一个核心挑战是高维非线性数据聚类。由于高维数据往往位于非欧空间，这使得 k -means 等基于欧氏距离的方法不能有效地描述数据间的相似性，从而难以揭示数据潜在的真实结构。为解决这一问题，近年来的研究主要集中在如何通过学习的方法将输入数据投影到另外一个线性可分的空间。换言之，聚类分析的关键在于表示学习。

过去几年，子空间聚类结合稀疏编码等浅层表示学习方法，成为高维非线性数据聚类的一个重要研究方向。大多子空间聚类可等同于流形学习+ k -means。简言之，

通过流形学习将数据投影到某一低维欧氏空间，再使用 k -means 得到聚类结果。相关工作的主要区别在于采用了不同的线性/非线性重构系数来度量流形上的不变性，其局限性主要包括：需要求解一个或多个时间复杂度正比于数据规模三次方的优化问题、优化问题的计算可能涉及整个数据集，因此难以实现在线或增量聚类、对数据分布仍然需要较强假设等。

深度学习作为当前最为有效的表示学习方法，在图像识别、场景理解、语音处理等各种有监督应用中取得了巨大突破，受到了国内外学者的广泛关注。深度学习的成功可较大程度归功于其基于数据驱动的端到端特性，即根据数据自身分布建模输入与输出之间关系。在无监督场景下，由于缺少类似标记信息的有效目标输出，使得无监督的深度学习研究成为一个公开的难题。

针对这一难题，以聚类为目标，我们对结合深度学习的聚类分析进行了一些尝

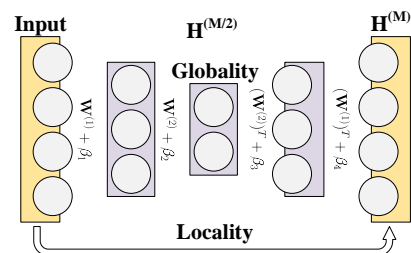


图1 面向聚类的结构化自编码器

试。自编码器是深度学习中最常见的无监督模型，通过最小化单个样本的重构误差（局部性）从而获得数据的紧致表示。为进一步提高自编码器的性能，我们提出一种结构化自编码器（图1），其不仅可以保证编码层的输出是紧致的，同时保留了数据在输入空间直接的结构化信息（全局性）。此外，我们提出的数据-聚类中心不变性假设：给定一个数据 x ，其在流形上与一组簇中心 Ω 之间的条件概率对不同的度量测度是不变的。换言之，在流形上，给定 Ω 的情况下， x 的聚类隶属关系不随所采用的测度变化。基于该假设，我们提出了一个端到端的聚类方法（见图2）。

上述研究的相关成果自2015年以来已相继发表在人工智能和神经网络领域的多个主流国际期刊与会议上（包括IJCAI, AAAI, TNNLS等）。

（责任编辑：邓成）

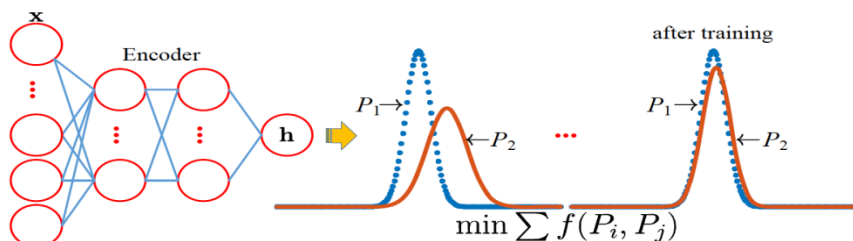
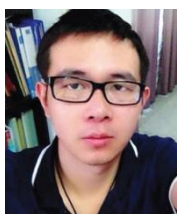


图2 端到端(End-to-End)的聚类



彭玺，新加坡资讯通信研究院研究员，主要研究方向为表示学习及其在计算机视觉和模式识别中的应用。

Email: pengx.gm@gmail.com

同济大学王瀚漓教授访谈

2017年4月27日,委员风采版块编委采访了同济大学电子与信息工程学院计算机科学与技术系教授、国家自然科学基金优秀青年基金获得者、博士生导师王瀚漓教授。采访内容摘录如下:

问:王教授,您在图像处理与计算机视觉领域进行了很多高水平研究工作,获得了国家自然科学基金优秀青年基金,入选了教育部新世纪优秀人才支持计划,是青年科技工作者的楷模。您能否向大家分享一下您的科研经历、主要科研成果和成功经验?

答:首先感谢CCF计算机视觉专委会的采访。作为学术界的晚辈,我非常荣幸能在此与大家进行交流。在研究工作方面,我还需要向国内外的各位前辈和同行多多学习和请教。

关于自己的学习和科研经历,我的感觉是一直都很幸运。我在浙江大学攻读学士学位,后来获得保送资格继续攻读硕士学位。研二时,我的硕士生导师推荐我尝试申请香港城市大学的一份研究助理工作。经过一些准备,我比较顺利地获得了这一职位。当时合作指导的导师是Sam Kwong教授和金耀初教授,主要从事的研究课题是模糊系统的可解释性规则及其应用。这样,我攻读硕士学位的最后一年是在香港城市大学度过的,我很感谢Sam和金老师的指导,带我进入高水平的科研学术殿堂。在两位导师的指导下,我比较顺利地完成了这个研究课题的既定任务。硕士毕业后,我选择了继续在香港城市大学攻读博士学位,研究方向选择了视频编码和多媒体信

号处理。选择这一研究方向一方面是由于课题组的实际需求,另一方面是我个人也希望做一些和实际应用更紧密相关的课题。在三年的博士求学阶段,我很感激导师Sam的悉心指导,从导师Sam那里我不仅获取了开展科研的能力,而且也培养了自己的心性。导师Sam的那种达观、宽容和平和的心态值得我永远学习。临近博士毕业前夕,我申请了德国的洪堡学者奖学金。在等待期间, Sam为我提供了香港城市大学研究员的职位,使我可以继续安心从事科研工作。我作为研究员在香港城市大学工作1个月后,一位好友向我推荐了斯坦福大学访学和硅谷初创公司工作的机会。在征求Sam的同意后,我前往美国加州开始了长达约1年半的访学工作,师从Charles K. Chui教授,主要从事视频编码和传输的研发及应用。美国的访学工作结束后,我来到了德国,开始了德国的洪堡学者研究工作,主要从事视频编码优化理论等方面的研究,师从Wolfgang A. Halang教授。2010年5月,我回到了上海,加入同济大学计算机科学与技术系,现为教授、博士生导师。

自2004年至今,我的主要研究方向之一是视频编码,主要包括:视频的快速编码理论与方法、码率控制等。回国之后,由于团队的扩大,我也开始尝试计算机视觉和机器学习方面的研究工作,包括:图片/视频检索、多媒体事件检测等。回首这十多年来的科研工作,我庆幸自己非常幸运。一路走来,得到了诸位恩师、前辈和朋友的帮助,自己也比较自律和努力,取得了一定的成绩。



问:您在IEEE Transactions on Image Processing等国际顶级期刊和会议上发表学术论文100余篇,并入选爱思唯尔(Elsevier)“中国高被引学者”。您在取得这些科研成果的过程中有哪些不为人知的艰辛过程或有趣经历?您认为优秀科研人员最重要的品质是什么?

答:现在回想起来,很多事情和经历都非常难忘。在这当中,我特别珍惜两个时期的经历。

第一时期是2003-2007年,我在香港城市大学求学和工作。初次来到香港是2003年3月,当时正值SARS肆虐,记得那时每天都带着口罩,即便在办公室里也带着口罩。当时,经济不景气,房租也便宜,初来香港在深水埗找了间唐楼,10楼的一个小隔间,大约7-8个平方,没有电梯,一个上下铺,一个独立洗手间。我一住就是4年,室友换了很多,都是香港城市大学计算机科学与技术系的学生或访问学者。那段时间,我基本都是每天早上6点起床,步行到办公室,途中买个菠萝包或红豆沙包,就着白开水,就是早餐。晚上大约11点和室友及办公室的小伙伴们一起搭伴返回住处。每天过着三点一线的生活:“住处

办公室-食堂”。生活貌似枯燥，但却非常充实。那段时光现在回想起来，是我最珍视的经历之一，我不仅可以全身心地投入科研，而且也收获了诸多珍贵的友谊、结识了很多好朋友。

第二段经历是 2007-2009 年，我在美国访学、工作。刚到美国时，正值美国次贷危机爆发之时。导师 Charles 和同事们在生活方面给予我无微不至的关怀和照顾，至今回想起来亦倍感温暖。当时一起合作的有三名研发工程师，我主要负责算法方面的设计和核心模块的原型编程。我们四个人在一间办公室里，经常讨论可能的技术方案，偶尔还有休息时段的最新当地新闻和趣事，繁忙中夹杂着欢乐。从他们身上我看到和学到了很多，包括技术、积极乐观、沟通和协作。当时的办公室靠近 Menlo Park，我住在旁边的一个城市 Redwood City，很多清晨我都步行前往办公室，行走于两个城市之间，路过稀疏错落的墨西哥小店，感觉很惬意。

关于科研人员的重要品质，我觉得首先对科研要有激情，对所从事的研究领域要热爱和怀有浓厚的兴趣。其次，要有持之以恒的毅力和定力，做科研要耐得住寂寞。科研过程中，遇到困难和挫折时不要气馁，要善于总结经验教训，及时调整思路和方法。

问：高效的科研团队对于深入开展课题研究有至关重要的作用。您作为同济大学多媒体与智能计算实验室主任，您能否简要介绍一下您的科研团队？您认为怎样才能组建和管理一个高效的科研团队？

答：我于 2010 年 7 月加入同济大学，在学院和系里的支持下，分配了一间实验室，于 2011 年开始招收研究生，并成立了有关多媒体与智能计算方面的实验室。迄今为止，已经毕业了 12 名研究生，目前实

验室在读的博士研究生有 9 名，硕士研究生有 10 名。我们课题组规模还比较小，称不上是一个高效的科研团队。在这里，我也打个小广告：从事多媒体、计算机视觉和机器学习等方面研究工作的朋友和同仁，如果您或者您的朋友希望加入同济大学计算机科学与技术系，欢迎联系我。

问：作为博士生导师，您在招收研究生时最注重学生的哪些方面？能否分享一下您在指导研究生方面的心得体会？

答：在招收研究生时，我主要注重学生以下几个方面。首先是学生的入学动机，我比较偏爱对科研真正感兴趣的学生。兴趣是最好的老师，一个真正对科研感兴趣的学生，才能在学术研究的道路上走得更远并享受科研工作带来的乐趣。我还看重学生的品格，我欣赏踏实、勤奋、做事认真细致的学生。此外，我比较重视学生的相关科研基础、数学功底和英语水平。每年招收学生前，我都会和候选学生聊一聊，说一下我的培养方式和对学生的期待，也听听候选学生的想法，考察他们在科研兴趣、品性和相关基础方面的情况。虽然一两次谈话并不能真正了解一个人，招收前的沟通还是或多或少可以起到一定作用。

正如前面所讲，我们课题组目前有约 20 名研究生。我每周都会花 1 到 2 天的时间和每个学生进行单独讨论(或小范围研讨)。通过和每个学生进行单独交流，了解他们的科研进展和近期计划，并给予相应的指导。在与学生的谈话过程中，除了交流科研，时常也会关心他们的生活，希望可以和学生做到亦师亦友。

问：科研基金是开展科研工作的前提。您申请科研基金通常从哪些方面进行准备？您认为书写基金申请书时应注意哪些问题？能否分享一下您个人申请科研基金的成功经验或失败原因？

答：是的，对于科研工作者来说，科研基金是非常重要的，可以看成是开展科研工作的前提。在同济大学计算机系，招收研究生的指标直接和科研经费到款挂钩。科研经费到款少，意味着招收研究生的指标少甚至停招。关于科研基金的申请，我个人觉得一定要做好调研和先期准备工作，要根据自己的实际情况看适合申请哪类科研基金，申请科研基金时要注意是否有相关的科研基础和积累。

关于基金申请书的准备，首先要有一个立意鲜明的题目，做到主题突出。其次要做好国内外相关研究现状的调研工作，分析应该全面、客观，突出申请工作与已有工作的不同之处。申请书要重点介绍申报工作的研究目标、研究内容和具体研究方案，目标、内容和方案之间要相辅相成、环环相扣。此外，要凝练好申请工作的创新点，有理有据地阐述申请工作的可行性以及研究团队已有的相关研究基础。申请书写好后，建议给领域内有经验的专家看看，请专家提提意见，再有针对性地修改。

我第一次写基金申请书时，是申请香港的 RGC 项目，一开始没什么经验，写好后同导师和其他合作者讨论，反复修改多次。后来再写就好些了。我的相关经验也不多，倒是有幸看过一些非常优秀的申请书，感觉好的科研基金申请书是一定要经过千锤百炼的。

问：您能否结合自己的审稿经历，谈一下学术论文写作需要注意哪些方面的问题？您对于研究生撰写学术论文和投稿有哪些建议？

答：在学术论文写作方面，我个人觉得可以考虑以下几个方面：(1)主题明确，立意鲜明。一篇好的学术论文应该具有明确的主题和相关背景介绍，其提出的方法或解决问题的途径与本领域其他相似或相

方法相比应该具有良好的创新特色。(2) 层次清楚, 脉络分明。在论文的谋篇布局方面应该具有较好的逻辑性。例如, 在介绍一个新的算法时, 可以描述出: 他人方法及其局限性(或可以提高的方面), 设计本文方法的动机、目标及具体方案, 本文方法所提方案的验证(包括所提出的每个创新点的验证)。(3) 语言精炼, 表达顺畅。文章的遣词造句要仔细斟酌和推敲。一篇好的论文一般要修改多次才能完成。

对于研究生朋友们, 我建议要多读、多写和多悟。多读一些本领域顶级期刊和会议论文, 培养自己的语感。要擅于将自己的阶段性成果和心得进行记录, 落到字面, 平时注意培养自己的写作能力。读论文时要多思考、多领悟, 有意识地提高自己的逻辑思维能力。关于投稿, 建议要多听取导师的建议, 做好规划, 并勇于大胆尝试。即使被拒, 也不要气馁, 可以从审稿人的意见中学到很多有益的东西并从中得到启发。

问: 您认为计算机视觉领域未来比较有潜力的研究方向主要有哪些? 您能否对计算机视觉领域入门者给出一些建议?

答: 实际上, 我真正开始计算机视觉方面的研究工作起步较晚, 大约是从 2012 年初开始的, 我就结合我目前研究团队在计算机视觉方面拟开展的研究工作说一下吧。目前, 我们团队的研究方向之一是可感知视频编码, 希望更好地从人类视觉的感知角度对视频数据进行压缩, 同时我们也希望借助于计算机视觉技术对大规模视频数据进行协同编码。另一个研究小组在从事图片/视频描述方面的相关研究工作, 通过对图片/视频的内容理解, 生成相应的描述。此外, 我们也尝试在视觉问答方面做一些探索工作。我个人认为通过借助于视觉信息的理解进行预测学习, 应该是一

个很有潜力的研究方向。

问: 您既是一位成功的科研人员, 又是两个孩子的慈父, 拥有幸福的家庭。请问您是如何平衡科研工作和日常生活的?

答: 在家庭方面我觉得自己很幸运, 大女儿今年升初中, 小儿子刚刚学会走路。面对一双儿女有时觉得很愧疚, 我住在浦东张江, 工作在同济大学嘉定校区, 一个在上海的东南角, 一个在上海的西北角, 路上单程耗时 2 小时左右, 我很多时候是周一到嘉定校区, 在学校工作到周五才回家, 家人也表示理解和支持。周末在家, 我会努力陪伴孩子们, 或者陪着去上兴趣辅导班, 或者一起做游戏。

问: 大家都知道, 科研工作是非常辛苦、需要全身心投入的。健康的体魄是科研工作的必要保证。请问您是怎样平衡工作与休息的? 您平时主要进行哪些体育锻炼? 能否与大家分享一下您的业余生活?

答: 我的作息比较有规律, 大约从大学时代就养成了规律作息的习惯。工作日基本上每天晚上 12 点睡觉, 早上 6 点起床, 中午午休 30~60 分钟。太忙时可能会稍微再晚睡一些, 但是我会及时进行调整, 避免自己过于疲劳。

平时我很喜欢走路。说来惭愧, 我到今天还没有去学开车, 每个学期开始时总是和自己说, 下个学期一定要去学, 等到了下个学期又有其他事情和安排, 就又拖到后面了。因此, 我经常走路。此外, 我非常喜欢游泳, 周末或放假有空时会去游泳。平常空闲时间不太多时, 也会督促自己做些小的体育锻炼, 比如伏地挺身和高抬腿等不太花时间的运动, 每次 10 分钟左右, 一天做 2-3 次。

我的业余生活挺单调的。在学校工作时, 基本上就是三点一线的工作和生活模式; 周末回到家, 除了工作外, 就是陪孩

子们、帮家人做家务。有时候, 晚上吃得比较多、较难消化时, 也会唱唱歌, 一来锻炼自己的肺活量, 二来健身效果还是不错的, 各位朋友有兴趣可以试一试☺。

王瀚漓, 男, 博士、教授、博士生导师。分别于 2001 年和 2004 年于浙江大学电气工程学院获工学学士和工学硕士学位;



2007 年于香港城市大学计算机系获哲学博士学位, 并荣获香港城市大学杰出毕业论文奖。2007 年至 2008 年于香港城市大学从事研究员工作, 同期访问美国斯坦福大学从事访学工作。2008 年至 2009 年于美国加州硅谷公司 Precoad Inc. 担任软件工程师。2009 年至 2010 年于德国哈根综合大学从事德国洪堡学者研究员工作。2010 年回国, 现任同济大学计算机系教授、博士生导师。主要研究领域包括视频编码、计算机视觉和机器学习等。发表学术论文 100 余篇, 2014-2016 年连续三年入选爱思唯尔 (Elsevier) “中国高被引学者榜单”, 并获得了包括教育部新世纪优秀人才支持计划、上海市浦江人才计划、上海市曙光学者计划、上海高校特聘教授(东方学者)计划及其跟踪计划、国家自然科学基金优秀青年科学基金等人才类计划项目。主要学术兼职包括 IEEE 高级会员、德国洪堡基金会会员、中国计算机学会高级会员、中国计算机学会多媒体专委会委员、中国计算机学会计算机视觉专委会委员、中国电子学会会员、上海市计算机学会多媒体专委会副主任、上海市计算机学会人工智能专委会副主任等, 担任多个国际学术会议的分会主席和程序委员会委员。

(责任编辑: 韩爱丽 余烨)

委员好消息

- 2016年8月21日,由第23届国际模式识别大会组织的2016 ChaLearn LAP 大规模手势识别竞赛中,CCF-CV 专委会委员、中科院计算所陈熙霖研究员团队获得在线动态组冠军,西安电子科技大学苗启广教授课题组荣获静态组冠军,公安部三所梅林主任与西安电子科技大学软件学院沈沛意教授联合团队荣获季军。ChaLearn LAP 国际手势识别竞赛从2011年至今已经连续举办五届,被誉为模式识别和机器视觉领域最富挑战性的战场,本届大赛吸引了全球各大高校和科研机构的53支参赛队伍。
- 2017年3月,IEEE ICME 2017 评出 World's FIRST 10K Best Paper Award, CCF-CV 专委会委员、中山大学人机物智能融合实验室林惊教授的论文“*Knowledge-Guided Recurrent Neural Network Learning for Task-oriented Action Prediction*”荣登榜首,获钻石奖。IEEE ICME 是国际多媒体领域旗舰会议之一,World's FIRST 10K Best Paper Award 是由 IEEE ICME 组织评议的奖项,以促进通用多媒体相关领域的研究进展,由相关领域的世界顶尖研究人员组成的评奖委员会对决赛作出独立评议。
- 2017年4月1日,据教育部新闻办公室官方微博消息,经学校推荐、通讯评审、会议答辩、人选公示、评审委员会审定、聘任合同签订等程序,教育部最终确定440人为2016年度长江学者特聘教授、讲座教授、青年学者。CCF-CV 专委会委员、上海交通大学熊红凯教授入选长江学者特聘教授,CCF-CV 专委会委员、西安电子科技大学董伟生教授入选青年学者。
- 2017年4月23日,VALSE 2017 在厦门落下帷幕。VALSE 2017 新增了年度贡献奖(评出8人)、年度优秀 WEBINAR 讲者(评出10人)和最受关注 Poster 奖(评出一、二、三等奖各1、2、3篇)。CCF-CV 专委会委员、中科院自动化所何晖光研究员、合肥工业大学贾伟副研究员、电子科技大学姬艳丽副教授、国防科技大学郭裕兰老师获 VALSE 2016-2017 年度贡献奖,CCF-CV 专委会委员、中科院深圳先进技术研究院乔宇研究员获 VALSE 2016-2017 年度优秀 WEBINAR 讲者,CCF-CV 专委会常务委员、华中科技大学白翔教授的 CVPR 2017 论文“*Detecting Oriented Text in Natural Images by Linking Segments*”获 VALSE 2017 最受关注 Poster 奖一等奖,另有多位 CCF-CV 专委会委员论文分别获得二、三等奖。VALSE 发起于2011年,是国内计算机视觉、图像处理、模式识别与机器学习领域青年学者的学术交流舞台。
- 2017年4月26日,2016年度北京市科学技术奖励大会上,由 CCF-CV 专委会委员、北京大学彭宇新教授等完成的“互联网多模态内容分析与识别关键技术及应用”项目荣获北京市科学技术奖技术发明类一等奖。该项目发表论文100多篇(其中 ACM/IEEE Trans.期刊和 CCF A 类论文26篇),获发明专利授权50项,软件著作权16项,6次参加国际权威评测 TRECVID 和 TREC,每次均获得第一名。该项目形成了互联网多模态内容分析与监管系统等系列产品,直接经济效益10.3亿元,并在维护我国互联网内容安全、促进网络文化健康发展等方面产生了显著的社会效益。
- 2017年5月5日,中国图象图形学会视觉大数据专委会(CSIG-BVD)成立,CCF-CV 专委会秘书长、中科院自动化所王亮研究员当选为 CSIG-BVD 专委会主任,CCF-CV 专委会委员、西安交大薛建儒教授,上海交大熊红凯教授和东南大学耿新教授当选副主任,中科院自动化所赫然研究员当选秘书长,复旦大学姜育刚教授、中山大学郑伟诗教授、中科院信工所张晓宇副研究员当选副秘书长。
- 2017年5月6日,CCF-CV 专委会委员、中科院自动化所何晖光研究员团队的工作“*Sharing Deep Generative Representation for Perceived Image Reconstruction from Human Brain Activity*”被 MIT Technology Review 头条报道评述,该工作可以通过大脑扫描数据重建图像,“新提出的技术有效克服了这个领域的一系列难点,能够更加精确地读懂人类大脑信号,还原出人眼所看到的视觉场景。该技术不仅对计算神经科学意义重大,也是研发新一代脑-机接口技术的基石。”

(责任编辑:刘海波 余志文)

数据集：场景理解数据集介绍

东北大学 贾同 兰州理工大学 李策

近年来,场景理解成为计算机视觉领域的热门研究课题之一,广泛地应用于意义地图构建和增强现实。而场景理解的解决方法则以深度学习、条件随机场为代表的机器学习算法为主。为了给深度学习、条件随机场等算法提供大数据训练样本,并评价各个场景理解算法的优劣,需要大量的有标注的场景理解数据集。以下对几种常用的场景理解数据集进行介绍。

1 二维彩色图像数据集

场景理解又具体分为场景分类、目标识别、语义分割等具体研究内容。其中场景分类数据集标注每张图像的场景类别;目标识别数据集标注单目标图像的类别和多目标图像的物体位置和类别;而语义分割数据集为图像中每个像素标注语义标签。下面介绍二维彩色图像数据集。

1.1 Places2 数据集

Places2 数据集是由麻省理工学院(Massachusetts Institute of Technology)的CSAIL(Computer Science and Artificial Intelligence Laboratory)实验室采集,主要用于场景分类的彩色图像数据集。其中包含 1000 万张图像, 400 多个场景类别。每个类别包含 5000 到 30000 张图片。



图 1 Places2 数据集样本

- 数据集网站：

<http://places2.csail.mit.edu/explore.html>

1.2 SUN 数据集

SUN 数据集是由麻省理工学院(Massachusetts Institute of Technology)和布朗大学(Brown University)采集,主要用于场景分类、语义分割的彩色图像数据集。其中包含 899 个类别和 130519 张图像。每种类别包含数量不等的图片。



图 2 SUN 数据集样本

数据集中包含以下内容：

- Color Images：彩色图像。
- Annotated Images：标签图像，为图像中每个像素分配语义标签。
- 数据集网站：

<http://groups.csail.mit.edu/vision/SUN/>

1.3 ImageNet 数据集

ImageNet 数据集是由斯坦福大学(Stanford University)和普林斯顿大学(Princeton University)共同采集,主要用于场景分类、目标识别等。该数据库是根据 WordNet 层次结构组织的图像数据集,每个节点有超过 500 张图像,共包含 14000000+张图片。



图 3 ImageNet 数据集样本

- 数据集网站：

<http://image-net.org/explore>

1.4 Pascal2 数据集

Pascal2 (Pascal VOC)数据集是由 Mark Everingham (University of Leeds), Luc van Gool (ETHZ, Zurich)等来自多所大学和研究机构的组织者共同收集的彩色图像数据库。其中包含 20 个物体类别,共 17125 张图像。



图 4 Pascal2 数据集样本

数据集中包含以下内容：

- Classification/Detection Image：识别图像,标注每张图像类别。
- Segmentation Image：分割图像,标注图像中分割区域的类别。
- Action Classification Image 动作分类图像,标注每张图像类别。
- Person Layout Image：人体图像,标注图像中人体的头、手、脚的位置和类别。
- 数据集网站：

<http://host.robots.ox.ac.uk/pascal/VOC/in dex.html>

2 三维 RGB-D 数据集

相比于二维彩色图像数据,RGB-D 提供额外的深度图像。数据集采集设备以彩色像机和红外测距仪为传感器,同时采

集目标彩色图像和深度图像，并且彩色与深度图像中的像素可以逐点匹配。

2.1 NYU Depth V2 数据集

NYU Depth V2 数据集为纽约大学 (New York University)采集，其中共包含 1449 个已标注场景样本，407,024 个未标注场景样本。

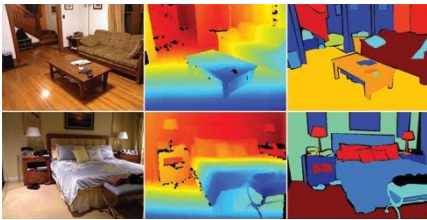


图 5 NYU Depth V2 数据集样本

采集设备：Microsoft Kinect Xbox 360。
数据集中包含以下内容：

- RangeImage：场景的深度图像，分辨率为 640×480 pixels。
- Color Image：场景的彩色图像，分辨率为 640×480 pixels。
- Label Image：场景的标签图像，分辨率为 640×480 pixels，标注每个像素的语义。
- 场景类别和拍摄时的加速计计数。
- 数据集网站：
http://cs.nyu.edu/~silberman/datasets/nyu_depth_v2.html。

2.2 SUN3D 数据集

SUN3D 数据集由普林斯顿大学 (Princeton University)，麻省理工学院 (Massachusetts Institute of Technology)采集，提供全场景视频和对应的点云语义标签。共包含上万帧 RGB-D 图像，录制总计约 668m² 室内场景的视频。

采集设备：ASUS Xtion。
数据集中包含以下内容：



图 6 SUN3D 数据集样本

- RangeImage：场景的深度图像，分辨率为 640×480 pixels。
- Color Image：场景的彩色图像，分辨率为 640×480 pixels。
- Label Image：场景的标签图像，分辨率为 640×480 pixels，不仅提供 RGB-D 图像的语义标签，还通过 SFM 算法提供三维重建点云的语义标签。
- Bounding Box：三维检测框，场景中每个物体的点云被一个外接长方体包围。
- 数据集网站：
<http://sun3d.cs.princeton.edu/>。

2.3 ShapeNet 数据集

ShapeNet 数据集由普林斯顿大学 (Princeton University)、斯坦福大学 (Stanford University) 和 TTIC (Toyota Technological Institute at Chicago)共同采集，包含多个子数据库，依据 WordNet 形式构建，目前拥有 55 大类，270 个子类，共 63,300 个 3D 模型。



图 7 ShapeNet 数据集样本

- 数据集网站：
<https://www.shapenet.org/>。

2.4 ObjectNet3D 数据集

ObjectNet3D 数据集是斯坦福大学 (Stanford University)整理的三维物体识别的数据库，包含 100 种类别，90,127 张图像，201,888 个图像中的物体以及 44,174 个三维模型。数据库提供其图像中物体的三维模型及其精确的位置和姿态。



图 8 ObjectNet3D 数据集样本

数据集中包含以下内容：

- Pose Annotation Image：位姿图像，标注图像中物体的精确位置和姿态。
- Shape Annotation Image：模型图像，匹配彩色图像中的物体与三维模型。
- 数据集网站：
<http://cvgl.stanford.edu/projects/objectnet3d/>。

作者介绍：



贾同，博士，副教授，东北大学信息科学与工程学院。主要研究领域为智能机器人三维视觉感知、视觉测量、机器学习等。



李策，博士，教授，兰州理工大学电气工程与信息工程学院。主要研究领域为计算视觉、模式识别、智能机器人系统。

(责任编辑：蹇木伟)

TensorFlow: 开源机器学习软件库

(本文内容摘录于 <http://www.tensorfly.cn/>)

1 关于 TensorFlow

1.1 什么是 TensorFlow

TensorFlow 是一个由 Google Brain 团队开发的采用数据流图 (data flow graphs) 用于数值计算的开源软件库。其架构灵活, 可以在多种平台上展开计算, 例如台式计算机中的一个或多个 CPU(或 GPU)、服务器、移动设备等等。目前, TensorFlow 已被广泛应用于人工智能领域的研究和生产中, 例如语音识别、计算机视觉和自然语言处理等。



图 1 Tensorflow logo

使用 TensorFlow 必须明白：

- 使用图 (graph) 来表示计算任务；
- 在被称之为 会话 (Session) 的上下文 (context) 中执行图；
- 使用张量 (tensor) 表示数据；
- 通过变量 (Variable) 维护状态。

使用 feed 和 fetch 可以为任意的操作 (arbitrary operation) 赋值或者从其中获取数据。

1.2 TensorFlow 的特征

TensorFlow 的主要特征有：

- 高度的灵活性

用户负责构建图, 描写驱动计算的内部循环, TensorFlow 提供工具来帮助组装“子图”(常用于神经网络)。用户也可以自己在 TensorFlow 基础上写自己的“上层库”, 定义顺手好用的新复合操作和写一个 python 函数一样容易, 而且也不用

担心性能损耗。

- 自动求微分

基于梯度的机器学习算法会受益于 TensorFlow 自动求微分的能力, 其将自动为用户计算相关的微分导数。Tensorflow 计算某个变量相对于其他变量的导数仅仅是通过扩展图来完成的, 如图 2 所示。

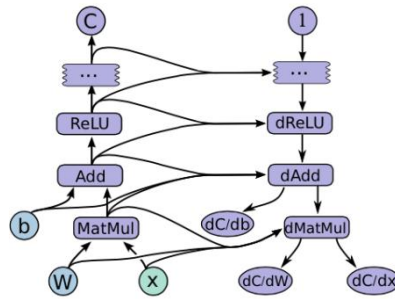


图 2 通过扩展图的方式实现自动求导

- 多语言支持

Tensorflow 有一个合理的 c++ 使用界面, 也有一个易用的 python 使用界面来构建和执行用户的图。用户可以直接写 python/c++ 程序, 也可以用交互式的 ipython 界面来用 TensorFlow 尝试想法。同时, 它可以帮用户将笔记、代码、可视化等有条理地归置好。

- 将科研和产品联系在一起

使用 TensorFlow 可以让应用型研究者将想法迅速运用到产品中, 也可以让学术性研究者更直接地彼此分享代码, 从而提高科研产出率。

- 性能最优化

由于 TensorFlow 给予了线程、队列、异步操作等以最佳的支持, 因此 TensorFlow 可以将用户手边硬件的计算潜能全部发挥出来。用户可以自由地将 TensorFlow 图中的计算元素分配到不同设备上, TensorFlow 可以帮用户管理好不同

的副本。

- 真正的可移植性

Tensorflow 可以在 CPU 和 GPU 上运行, 比如台式机、服务器、手机移动设备等等。Tensorflow 可以让你在没有特殊硬件的前提下, 在自己笔记本上运行机器学习程序; TensorFlow 可以让你无需修改代码便将训练模型在多个 CPU 上规模化运算; TensorFlow 可以将你的训练好的模型作为产品的一部分用到手机 app 里; TensorFlow 也可以将你的模型作为云端服务运行在自己的服务器上, 或者运行在 Docker 容器里。

1.3 TensorFlow 的使用技巧

TensorFlow 的核心使用技巧有：

- 准备训练数据

一般 TensorFlow 应用代码包含 Graph 的定义和 Session 的运行, 代码量不大可以封装到一个文件中。训练前需要准备样本数据和测试数据, 一般数据文件是空格或者逗号分隔的 CSV 文件, 但 TensorFlow 建议使用二进制的 TFRecords 格式, 这样可以支持 QueueRunner 和 Coordinator 进行多线程数据读取, 并且可以通过 batch size 和 epoch 参数来控制训练时单次 batch 的大小和对样本文件迭代训练多少轮。

- 接受命令行参数

众所周知, 深度学习有过多的 Hyperparameter 需要调优, 优化算法、模型层数和不同模型都需要不断进行调整, 这时候使用命令行参数是非常方便的。TensorFlow 底层使用了 python-gflags 项目, 然后封装成 tf.app.flags 接口, 使用起来非常简单和直观, 在实际项目中一般会提前定义命令行参数, 通过参数来简化

Hyperparameter 的调优。

- 定义神经网络模型

定义好的神经网络模型是机器学习中尤为重要的一步。定义模型参数可以很简单，例如创建多个 Variable 即可，也可以做得比较复杂，例如使用 `tf.variable_scope()` 和 `tf.get_variables()` 接口。在定义 Variables 时，TensorFlow 默认使用 GPU，这可能导致参数更新过慢，所以在定义 Variables 时注意要绑定 CPU，同时也要注意保证每个 Variable 都有独特的名字，而且能都轻易地修改隐层节点数和网络层数。

- 使用不同的优化算法

不同优化算法在训练结果、训练速度上都有很大差异，过度优化网络参数可能效果没有使用其他优化算法来得有效，因此选用正确的优化算法也是 Hyperparameter 调优中很重要的一步。不同优化算法对于不同场景和数据集的效果并不相同，因此选择优化算法的最好方式就是实践，通过前面定义的命令行参数可以很方便地使用不同优化算法来训练模型。

- 使用 TensorBoard 优化参数

TensorFlow 还集成了一个功能强大的图形化工具，即 TensorBoard，一般只需要在代码中加入我们关心的训练指标，TensorBoard 就会自动根据这些参数绘图，通过可视化的方式来了解模型训练的情况。

- 分布式 TensorFlow 应用

TensorFlow 有强大的分布式计算功能，传统的计算框架如 Caffe，原生不支持分布式训练，在数据量巨大的情况下往往无法通过增加机器 scale out。TensorFlow 承载了 Google 各个业务 PB 级的数据，在设计之初就考虑到分布式计算的需求，通过 gRPC、Protobuf 等高性能库实现了神经网络模型的分布式计算。

- Online learning 与 Continuous learning

很多机器学习厂商都会宣称自己的产品支持 Online learning，其实这只是 TensorFlow 的一个基本的功能，即支持在线数据不断优化模型。TensorFlow 可以通过 `tf.train.Saver()` 来保存模型和恢复模型参数，使用 Python 加载模型文件后，可不断接受在线请求的数据，更新模型参数后通过 Saver 保存成 checkpoint，用于下一次优化或者线上服务。

而 Continuous training 是指训练即使被中断，也能继续上一次的训练结果继续优化模型，在 TensorFlow 中也是通过 Saver 和 checkpoint 文件来实现。在 `deep_recommend_system` 项目默认能从上一次训练中继续优化模型，也可以在命令行中指定 `train_from_scratch`，不仅不用担心训练进程被中断，也可以一边训练一边做 inference 提供线上服务。

2 什么是数据流图

数据流图用“结点”（nodes）和“线”（edges）的有向图来描述数学计算。“节点”一般用来表示施加的数学操作，但也可以表示数据输入（feed in）的起点/输出（push out）的终点，或者是读取/写入持久变量（persistent variable）的终点。“线”表示“节点”之间的输入/输出关系。这些数据“线”可以运输“大小可动态调整”的多维数据数组，即“张量”（tensor）。张量从图中流过的直观图像是这个工具取名为“Tensorflow”的原因。一旦输入端的所有张量准备好，节点将被分配到各种计算设备完成异步并行地执行运算。

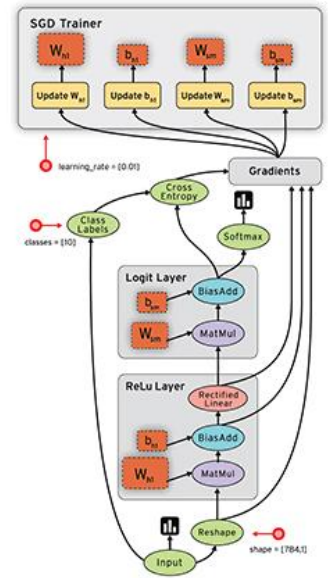


图 3 Tensorflow 数据流图

3 谁可以用 TensorFlow?

任何人都可以用 TensorFlow，学生、研究员、爱好者、极客、工程师、开发者、发明家、创业者等等都可以在 Apache 2.0 开源协议下使用 TensorFlow。开源的 TensorFlow 具有众多的学习资源：

<https://www.tensorflow.org/>

<http://www.tensorfly.cn/>

<https://github.com/tensorflow>

TensorFlow 是很好深度学习框架，对于个人开发者、科研人员已经企业都是值得投资的技术方向，而 Cloud Machine Learning 可以解决用户在环境初始化、训练任务管理以及神经网络模型的在线服务上的管理和调度问题。

（责任编辑：蹇木伟 刘丽 贾同）

招聘信息

一、中国科学院信息工程研究所

多媒体安全研究组招聘信息：

1、岗位名称：多媒体分析算法研究职位
--助理研究员/工程师

岗位部门：第二研究室多媒体安全研究组

需求人数：1~2 人

面向对象：应届毕业生、社会在职人员

工作职责：多媒体内容分析算法研究与实现

岗位要求：

(1) 博士/硕士学位获得者，计算机应用技术、模式识别与智能系统或电子工程等相关专业；

(2) 熟悉模式识别、机器学习和图像处理的常用方法，具备优秀的算法实现能力；

(3) 具有扎实的数学基础、熟练的计算机编程能力和英语交流能力；

(4) 研究能力突出，在本领域高水平国际学术期刊或顶级国际会议上发表论文(硕士可放宽要求)。

2、联系方式

请应聘者以 E-mail 形式发送简历, 请在邮件主题上注明: 姓名+应聘岗位名称。

简历投递截止时间: 招满即止。

联系人: 周宇老师。

简历接收邮箱: zhouyu@iie.ac.cn

二、中科视拓(北京)科技有限公

司职位招聘信息：

1、岗位名称：人脸识别算法研究员

岗位部门：研发部

需求人数：2 人

面向对象：社会在职人员

工作职责：负责人脸检测、对齐、识别、属性分析核心算法研究

岗位要求：

(1) 硕士及以上学位获得者；

(2) 拥有三年以上人脸识别相关领域工作经验, 应熟悉图像处理的常用方法以及深度学习；

(3) 在本领域高水平国际学术期刊或顶级国际会议上发表论文者优先；

(4) 有带团队经验。

2、岗位名称：深度学习算法工程师

岗位部门：研发部

需求人数：4 人

面向对象：社会在职人员

工作职责：负责深度学习核心算法研发, 精通深度目标检测、物体识别、图像分割、人脸识别、物体跟踪技术之一

岗位要求：

(1) 硕士及以上学位获得者, 熟练掌握计算机视觉和深度学习基础理论和方法；

(2) 拥有两年以上人脸识别, 物体识别, 物体检测, 图像检索等相关领域工作经验；

(3) 熟练使用至少一种开源学习框架, 例如 Caffe, Tensorflow, Torch 等；

(4) 扎实的 C++ 编程能力, 能熟练使用 OpenCV 算法库进行 CUDA 编程；

(5) 熟悉至少一种脚本语言, 例如 Python 或者 Lua。

3、岗位名称：目标检测算法研究员

岗位部门：研发部

需求人数：2 人

面向对象：社会在职人员

工作职责：负责目标检测、跟踪、识别核心算法研究

岗位要求：

(1) 硕士及以上学位获得者, 具备相关领域三年及以上研究经验；

(2) 熟悉机器学习和图像处理领域的常用方法, 熟悉目标检测、目标跟踪、目标识别、属性分析、行为分析领域的最新技术；

(3) 在国际顶级期刊及会议中有论文发表的优先；

(4) 有带团队经验的优先。

4、联系方式

请应聘者以 E-mail 形式发送简历, 请在邮件主题上注明: 姓名+应聘岗位名称。

简历投递截止时间: 招满即止。

联系人: 中科视拓人事部。

简历接收邮箱: hr@seetatech.com

三、旷视科技 (FACE++) 研究院

1、岗位名称：研究员

工作职责：将最好的算法在有趣有用的商业场景中落地

岗位要求：

(1) 扎实的计算机基础、数学基础；

(2) 过硬的编程能力, 熟悉至少一种语言, C++, Python 优先。

2、岗位名称：算法开发工程师

工作职责：作为研究组和产品组的纽带, 算法开发工程师负责将算法转化到实际产品中。对算法进行改造、优化, 运用创造性思维迎接来自业务需求的挑战; 整合公司核心算法, 利用软件工程的思想来维护其扩展性、灵活性。工作中需要的开发和研究比例约为 80:20。

岗位要求：

(1) 熟练使用 C/C++, Java, C# 等编程语

言；

(2) 对数据结构算法有一定掌握和运用能力,掌握语言级别的程序性能优化技巧者优先；

(3) 掌握一种脚本语言 (Python 优先)；

(4) 有跨平台开发 (Linux/Windows, x86/ARM) 经验者优先。

3、岗位名称：算法优化工程师

工作职责：与产品趋势紧密结合,挖掘算法的潜力和未来。该职位以产品需求为导向,通过对问题的理解和分析,在不同的解决方案中找到最合适的配置方式并进行整合和调优。与此同时,用工程化的方式将算法进行定制化并封装。

岗位要求：

(1) 熟练使用 C/C++, Java, 或 C# 等编程语言；

(2) 能够充分理解产品的需求,并针对需求调整开发策略；

(3) 有客户端开发经验者 (Android, iOS) 优先。

4、岗位名称：深度学习系统架构师

工作职责：

(1) 维护并优化公司内部的深度学习训练系统

(2) 异构平台上的神经网络性能优化。

岗位要求：

(1) 熟练掌握 C++；

(2) 掌握算法与数据结构；

(3) 了解计算机体系结构；

(4) 有过 CUDA/arm/x86 相关代码优化经验者优先；

(5) 了解深度学习相关知识者优先；

(6) 了解 Caffe/TensorFlow/MXNet/Torch/Theano 等训练框架者优先。

5、岗位名称：活体安全工程师

工作职责：

(1) 图像(人脸、证件等)识别应用进行安全分析；

(2) 理解图像识别算法原理,挖掘各类图像防欺诈技术漏洞；

(3) 针对各类漏洞,研究利用方法,研究攻防技术；

(4) 负责部分项目文档编写与管理。

岗位要求：

(1) 一年以上程序开发或测试经验,熟悉一门脚本语言；

(2) 动手能力强,熟练使用 photoshop 等图像处理软件；

(3) 有好奇心、有探索精神；

(4) 了解图像处理基本原理；

(5) 勤于学习、善于沟通；

(6) 敢于面对挑战,对新鲜事物勇于尝试。

6、联系方式

请应聘者以 E-mail 形式发送简历,并在邮件主题上注明:姓名+应聘岗位名称。

简历投递截止时间:招满即止。

联系人:吴可茵。

简历接收邮箱:wky@megvii.com

四、ULSee Inc.招聘信息：

1、岗位名称：计算机视觉算法专家

岗位部门：ULSee Inc. 杭州研发中心

需求人数：多名

面向对象：应届毕业生、社会在职人员

工作职责：计算机视觉算法开发

岗位要求：

(1) 硕士及以上学位获得者,在计算机视觉、模式识别、机器学习和图像处理方面有扎实的理论基础和实践能力,有较强的计算机编程能力和英语交流能力；

(2) 具有很强的团队合作精神和一定的项目组织管理能力,可以带领小团队开展算法研究和产品研发。

待遇：

薪资 40 万以上,条件优异者薪资不设上限,附加丰厚的公司期权鼓励,可以申报

杭州江干区“百人计划”和杭州“521”计划。

五、大连理工大学-立命馆大学国际信息与软件学院招聘信息：

实际信息与软件学院招聘信息：

1、岗位名称：计算智能/人工智能/可视化/软件工程教学科研职位---教授/副教授

需求人数：2-3 人

面向对象：在职人员

工作职责：承担教学任务和研究生培养工作；承担国家级或省部级科研项目；在国内外发表高水平的学术论文；参与学科建设和实验室建设。

岗位要求：

(1) 具有博士学位；

(2) 五年以上海外留学或三年以上海外工作背景者优先。

2、薪酬待遇

根据大连市和大连理工大学引进人才的有关规定,提供具有竞争力的国际信息与软件学院特色薪酬与发展空间,对于优秀青年人才有“青年才俊”计划支持。

3、联系方式

请应聘者登录大连理工大学人才招聘网上信息平台,注册并按要求填报信息

(<http://202.118.68.7/CMSForward/>)。同时以 E-mail 形式发送简历至

wangzhuo@dlut.edu.cn,并在邮件主题上注明:姓名+应聘岗位名称。

简历投递截止时间:招满即止。

联系人:王卓老师。

六、北京格灵深瞳信息技术有限公司招聘信息：

公司招聘信息：

1、岗位名称：导航算法工程师 (全职)

需求人数：1

面向对象：应届毕业生、社会在职人员

岗位职责：

研发高质量的机器人定位,导航和交互算法。

任职要求:

- (1) 熟悉移动机器人的导航流程;
- (2) 理解主流的路径规划、运动规划算法思想,具备一定的实践经验;
- (3) 了解主流的视觉、雷达定位算法,有雷达、超声等传感器数据处理经验;
- (4) 有良好的编程基础,熟练掌握 C++ 语言,有 ROS 平台开发经验。

联系方式:

请应聘者以 E-mail 形式发送简历,请在邮件主题上注明:姓名+应聘岗位名称。

联系人:任光阔

简历投递到 guanguoren@deepglint.com

2、岗位名称:深度学习算法工程师(全职)

需求人数:2

面向对象:应届毕业生、社会在职人员

岗位职责:研发准确、高效的物体检测、识别、检索等算法

任职要求:

- (1) 编程能力强,熟悉 C/C++, Python, 有 linux 开发经验;
- (2) 了解机器学习和深度学习的基本内容;
- (3) 熟悉至少一个深度学习框架,例如: Caffe, TensorFlow, Mxnet 等;
- (4) 对目标检测、物体识别、图像分割、物体跟踪、图像检索、增强学习、GAN、姿态估计、文字序列识别、模型优化、人脸检测、人脸识别等方向中的至少一个有深入了解;
- (5) 自我驱动能力强,工作细致、认真;
- (6) 【加分】在顶级会议或期刊中发表过相关论文或在相关比赛中获得过较好的名次。

联系方式:

请应聘者以 E-mail 形式发送简历,请在

邮件主题上注明:姓名+应聘岗位名称。

联系人:张德兵

简历投递到 debinzhang@deepglint.com

七、招聘单位简介

1、中科视拓

中科视拓(SeetaTech)是一家由中科院计算所创办的人工智能行业成长期公司,为企业客户提供人脸识别与深度学习开源产品和行业解决方案服务。中科视拓以“开源赋能共发展”的商业理念,帮助各种行业客户在人工智能时代获得可自主迭代和自我学习的人工智能研发能力。

中科视拓的核心创始团队来自中国科学院计算技术研究所,其首席科学家由陈熙霖研究员担任,CTO 由山世光研究员担任,CEO 由刘昕担任。创始团队一直从事计算机视觉特别是人脸识别领域的研究工作,研究成果分别获 2015 年度国家自然科学二等奖和 2005 年度国家科技进步二等奖。

公司成立之初即获得千万量级的天使投资,近期又顺利获得了数千万元的 Pre-A 轮的融资,加上公司已经获得的近千万合同金额,公司具备了快速发展的条件。诚邀计算机视觉和机器学习领域的算法高手和技术大拿加入公司,共同搏击 AI 特别是 CV 领域的星辰大海!

中科视拓(北京)科技有限公司官方网址:
www.seetatech.com

2、旷视科技(FACE++)研究院

旷视(Face++)成立于 2011 年,目前 C 轮,总部设立在北京,同时在北京、杭州、西雅图有分部。旷视(Face++)是一家为行业型和个人型用户提供全方位的智能解决方案和智能数据服务的人工智能企业。旷视的产品以智能云和智能互联为核心,基于行业领先的人脸识别、图像识别技术和深度学习方法,旷视已开发出适用于多

个行业的智能解决方案、智能硬件以及智能云服务并覆盖移动互联网、金融、安防、地产、零售、办公等多个重要领域,与阿里巴巴集团、中信银行、国家公安部第一研究所、东软集团、滴滴出行、秀水街、清华大学、万科集团、神州云海等 800 余家企业机构和政府部门实现了合作。

如果你对人工智能、互联网、高科技充满兴趣,如果你追求极致、简单可靠,那么,欢迎加入我们!

公司网址: www.megvii.com

3、ULSee Inc.

ULSee Inc. 是一家专注计算机视觉的人工智能公司,成立于 2014 年,目前有杭州和台北两个研发中心。公司面向国际化市场,踏实将视觉算法产品化落地,目前已为如 VW, Chanel, LINE, Tencent, Tobii 等国际知名大公司提供了计算机视觉的产品方案。

公司网址: <http://www.ulsee.com/>

4、大连理工大学-立命馆大学国际信息与软件学院

大连理工大学-立命馆大学国际信息与软件学院 2013 年经教育部批准设立,开创了中日两国在高水平学历学位教育层次上合作办学的先例,着力成为国际化办学的桥头堡与工程教育的创新区。学院有全日制本科生、硕士研究生、博士研究生等培养层次,并设有博士后流动站。本科生面向全国部分省份由大连理工大学统一招生,设有软件工程和数字媒体技术两个专业。

学院网址: <http://drise.dlut.edu.cn/>

(责任编辑:樊鑫)

征文通知

一、会议征文

计算机视觉领域相关国内外会议的征文通知如表 1 所示。本期,我们重点推荐在国内举办的若干重要的国际会议,请大家关注。

二、期刊征文

计算机视觉领域近期相关期刊专刊的征文通知如表 2 所示,包括 Pattern Recognition (PR)、IEEE Access、Pattern Recognition Letters (PRL)和 Journal of Visual Communication and Image Representation (JVCI)。请大家和专刊担任 CCF-CV 委员的编委联系。

三、会议简介

1、中国计算机视觉大会 CCCV 是由中国计算机学会 CCF 主办,计算机视觉

专委会承办的计算机视觉领域盛会,旨在为从事计算机视觉研究的学生,老师和工业界人员提供一个学科互动交流平台,促进领域内学术交流以及学术界与工业界之间的交流。中国计算机视觉大会 CCCV2017 将于 2017 年 10 月 12-14 日在天津举行,由中国民航大学承办,天津大学和南开大学联合协办。

大会录用的稿件将由斯普林格出版社出版,并被 EI 和 ISTP 检索,对遴选的优秀论文经扩展后将特别推荐到 Neurocomputing (SCI)、Pattern Recognition Letters (SCI)、IEEE Access (SCI)、Int. Journal of Advanced Robotic Systems (SCI)、Optoelectronics Letters (EI) 等 SCI/EI 期刊上发表。

2、国际图形图像处理会议 ICGIP 在过去的 8 年中已经先后在亚庇、马尼拉、开罗、新加坡、香港、北京、东京成功举办,旨在汇集图形与图像处理领域的研究人员,科学家,工程师,学者和学生,分

享他们的经验,进行思想交流,分型研究成果,并讨论所遇到的实际挑战和采取的解决方案。2017 年,第九届国际图形与图像处理国际会议将由中国海洋大学作为协办单位在青岛举行。

3、中国生物特征识别大会 CCB R 是由中国人工智能学会 CAAI 主办的国内生物特征识别领域的学术盛会。自 2000 年始,已经在北京、杭州、西安、广州、济南、沈阳、天津和成都等地成功举办了 11 届,有力促进了国内本领域的学术和技术发展。第十二届中国生物特征识别大会 CCB R2017 将于 2017 年 10 月 28-29 日在深圳举行,由深圳大学计算机与软件学院和哈尔滨工业大学(深圳)计算机科学与技术学院联合承办。本届会议将汇聚国内从事生物特征识别理论与应用研究的广大科研工作者,并邀请国际同行,共同分享我国生物特征识别研究的最新理论和技术成果,为大家提供精彩的学术盛宴。

(责任编辑:沈沛意)

表 1 计算机视觉相关领域国内外会议(特别推荐)

| 会议名称 | 会议时间 | 会议地点 | 截稿日期 | 会议网站 |
|---|---------------|-------|-----------|---|
| CCCV (Chinese Conf. Computer Vision) | 2017.10.13-15 | 中国 天津 | 2017.5.31 | http://ccf-cccv.org/2017/ |
| ICGIP (International Conf. on Graphic and Image Processing) | 2017.10.14-16 | 中国 青岛 | 2017.6.10 | http://www.icgip.org/ |
| CCBR (Chinese Conf. on Biometric Recognition) | 2017.10.28-29 | 中国 深圳 | 2017.6.15 | http://cv.szu.edu.cn/ccbr2017/ |

表 2 计算机视觉相关领域国内外期刊专刊(特别推荐)

| 期刊名称 | 专刊题目 | 截稿日期 |
|-------------|---|------------|
| PR | Deep Learning for Computer Aided Cancer Detection and Diagnosis with Medical Imaging | 2017.08.15 |
| JVCI | Special Issue on Visual Information Processing for Virtual Reality | 2017.09.30 |
| PRL | Robustness, Security and Regulation Aspects in Current Biometric Systems | 2017.09.30 |
| IEEE ACCESS | Special Issue on Visual Surveillance and Biometrics: Practices, Challenges, and Possibilities | 2017.10.31 |

COMPUTER VISION NEWSLETTER

计算机视觉

专委简报

诚征封底广告

对象：计算机视觉和人工智能领域高科技企业、学术团体或个人等

内容：产品技术、服务理念、科研成果、学术贡献及其他推介

费用：双方协商（联系 Email：ccfcvn@gmail.com）

【简介】专委简报是中国计算机学会计算机视觉专委会（CCF-CV）与委员之间的双向信息通道，每年6期，单月出刊。目前已出四期，每期阅读量几千人次。读者包括活跃于高校、科研院所、高科技企业的计算机视觉和人工智能相关领域的教师、学生和研发人员。CCF-CV于2013年10月成立，在CCF专委评估中连续三年分别获得2014年度“特色活动奖”、2015年度“综合进步奖”、2016年度“优秀专委奖”。
