

CCF

计算机视觉 专委会简报

COMPUTER VISION NEWSLETTER

2017/02 期
总第 06 期

专委动态

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
信息速递	CCF-CV 2017 年委员增选申请通知	05
	CCF-CV 秘书处 2017 年度工作会议召开	05
会议推介	中国计算机视觉大会 (CCCV 2017)	06
走进企业	CCF-CV 走进企业系列交流会	07

科技前沿

专题综述	三维形状数据的深度特征表示	08
热点追击	运用人工智能技术对医学影像进行自动判读	12
	视觉显著性检测：计算机视觉领域的基础性研究	13
	由点到面：基于运动特征点的人群行为分析及其应用	14
	AAAI 2017 会议总结	15

委员风采

委员访谈	西安电子科技大学董伟生教授访谈	16
委员好消息		19

资源平台

数据集	3D 人脸数据集介绍	20
仿真平台	Webots 仿真机器人平台简介	23
招聘信息		24
征文通知		25

CCF-CV 走进高校系列报告会



**CCF计算机视觉专委会
走进高校系列报告会**

CCF - CV Lecture Series
深圳大学-深圳 (第 28 期)

- 明仲教授 CCF理事, 深圳大学计算机学院院长
致欢迎词
- 沈琳琳教授 计算机视觉研究所所长
介绍研究所科研情况
- 贾奎博士 华南理工大学教授
深度学习模型中卷积核及网络连接优化探究
- 乔宇博士 深圳先进技术研究院研究员
视频行为的深度建模与识别方法
- 郑伟诗博士 中山大学教授
视频监控下的行为识别分析研究

【活动背景】自 2015 年 11 月起, CCF 计算机视觉专委会(CCF-CV)在全国范围内率先开展走进高校系列报告会、走进企业系列交流会等特色活动, 在学术界、工业界产生了热烈反响, 受众遍及祖国大江南北。CCF 计算机视觉专委会欢迎各兄弟学会、专委会借鉴, 共同推动我国相关领域的学术繁荣和产业发展!

第二十八期 CCF-CV 走进高校系列报告会于深圳大学圆满结束

2017 年 1 月 14 日, 中国计算机学会计算机视觉专委会 (CCF-CV) 走进高校系列报告会第 28 期活动在深圳大学计算机与软件学院一楼学术报告厅成功举行。

本期报告会特别邀请了来自华南理工大学的贾奎教授、中科院深圳先进技术研究院的乔宇研究员以及中山大学的郑伟

诗教授做精彩演讲。执行主席沈琳琳教授和于仕琪副教授联合主持了本次会议。会议共吸引了来自深圳大学、深圳先进技术研究院、哈尔滨工业大学深圳研究生院等相关科研院所与高校, 以及腾讯、华为等企业的工程技术人员 100 余人参会, 为珠三角地区特别是深圳地区计算机视觉与深度学习领域的科研工作者、相关企业研发人员等提供一个高层次的学术交流平台。

来自华南理工大学的贾奎教授长期从事计算机视觉、机器学习、模式识别及图像处理领域的教学、科研和产业化工作, 本次报告贾教授围绕题为“深度学习模型中卷积核及网络连接优化探究”展开, 分享了其团队在深度学习模型中卷积核(及网络连接)优化方面的最新研究进展。

中科院深圳先进技术研究院的乔宇

研究员的报告主题是“视频行为的深度建模与识别算法”, 报告首先介绍视频行为分析与理解的最新进展, 然后重点介绍课题组近年来开展的一些工作(CVPR 13, ICCV 13, CVPR 14, ECCV 14, CVPR15, IJCV 15, CVPR 16, ECCV 16)。内容包括视频中层表示、多视角特征编码、轨迹卷积特征池化、运动向量 CNN、时空注意模型等。

中山大学的郑伟诗教授以“视频监控下的行为预测分析研究”为题展开, 重点介绍了其团队最近在 RGB-D 视频上的行为预测分析的研究成果。同时也简要介绍了本组与之相关的交互行为分析工作, 包括基于模板学习的人物交互行为分析、基于 RGB-D 的异质特征数据融合的人物交互行为分析、基于广义相似性函数的人与人交互建模的多人交互行为分析学习模型。

本次报告会持续了三个多小时, 每位讲者的报告内容都非常精彩, 令人印象深刻。每位讲者都在报告结束后与现场的师生进行互动, 师生们积极发言, 讨论热烈, 现场气氛非常活跃。

CCF-CV 走进高校系列活动近期预告

时间	走进学校 (执行主席)
.4 月	武汉大学 (夏桂松/白翔)
	中山大学 (郑伟诗/任传贤)
.5 月	西南科技大学 (吴亚东)
	东南大学 (耿新)
.6 月	湘潭大学 (欧阳建权)
	上海大学 (方昱春)

CCF-CV

2017 年委员增选申请现已开放

自 2013 年 10 月成立以来,中国计算机学会 (CCF) 计算机视觉专业委员会 (ccfcv.ccf.org.cn) 发展迅速,举办了很有影响力的活动,搭建了全方位、高水平、大规模的计算机视觉领域交流平台。专委会成立三年以来,已经发展委员 247 人,连续三年在 CCF 专委评估中分别获得“特色活动奖”、“综合进步奖”、“优秀专委奖”。为了保持专委会的活力、促进国内外视觉领域人员的交流和合作,专委会现开放 2017 年计算机视觉专委会的委员增选工作。

申请时间: 2017 年 1 月 16 日至 2017

年 9 月 16 日

申请流程: 填写申请表 (专委会网站下载) 后, 发送至专委秘书处邮箱 (ccfcv@139.com), 主题写明“2017 新委员申请-姓名-单位”。注: 推荐人必须是现任专委委员, 名单可以从专委网站查询。电子版申请表中需填写推荐人姓名和意见, 委员增选成功后可以补签名。

申请资格: 任职国内外学术界或企业界副教授或等同级别以上的人员, 拥有计算机视觉相关领域的高水平研究成果, 是 CCF 会员, 且积极参加计算机学会计算机

视觉专委会的各项活动。特别优秀的讲师、企业人士亦可考虑。

申请需知: 每个现任专委委员可以推荐最多 3 名候选人, 主任委员、副主任委员、秘书长、常务委员可以推荐最多 5 名候选人。本次申请结果将在“2017 年中国计算机视觉大会 (CCCV2017)” (<http://www.ccf-cccv.org/2017>) 期间举行的专委工作年会上投票确定 (申请者届时必须注册参会)。

秘书处 2017 年度工作会议召开



2017 年 2 月 22 日, 中国计算机学会计算机视觉专委会 (CCF-CV) 秘书处 2017 年度工作会议于北邮科技酒店召开。专委王涛副主任与秘书处全体成员参会, 会议由秘书长王亮研究员主持。

本次会议主要议题是专委 2017 年度工作计划的制定。过去的 2016 年, 专委各项学术活动和组织建设稳步推进, 并逐渐形成了若干具有引领性的特色活动。新一

年度, 各项活动如何更有实效地开展, 如何更加广泛地激发委员们参与专委建设的积极性, 如何吸引更多优秀的产学研人才加入专委, 如何办好专委年度学术大会 CCCV2017, 大家围绕这些议题开展了热烈的讨论, 并形成了具体可行的执行方案。

走进高校活动将延续 2016 年度的火爆, 在活动申请、讲者征集、活动形式等方面将更为开放, 惠及更多高校学子; 走进企业活动将注重挖掘新型创业企业的需求, 将深化产学研交流的目标进一步落地; 专委简报全新改版, 编委阵容得到强力扩充, 由季刊改为双月刊之后的新一期将极大丰富其内涵; 专委网站将强化其窗口宣传与交流功能, 在为广大委员提供更为丰富便

捷的专委信息、专业资源的同时, 积极发挥其国内 CV 专业门户网站作用, 为更多海外留学生提供信息服务; 专委人才资源丰富、青年才俊辈出, CCCV2017 会议期间将有望举行优秀青年学术论坛, 为参会者带来更多精彩分享。

最后, 王涛副主任为大家做了题为《以规范促进发展, 让 CV 专委更具影响力》的报告, 与大家分享了 2016 年 CCF 总结交流大会上来自 CCF 总部的建议以及其它优秀专委的成功经验, 为秘书处在新一年度的工作开展提供了很多具有启发性的新思路。

CCCV 2017

诚邀申请组织竞赛



CCCV 2017诚邀申请组织竞赛

CCF Chinese Conference on
Computer Vision 2017, Tianjin

诚挚邀请公司或单位在 CCCV 2017 组织与承办视觉与学习方面的各种技术竞赛。本次大会竞赛主办方为中国计算机学会计算机视觉专委会，承办方为图像视频大数据产业技术创新战略联盟。在去年专委会组织的 RACV2016 竞赛单元中，公安部第三研究所，浙江捷尚视觉科技股份有限公司，北京爱奇艺科技有限公司分别承办了视频图像分析、模糊车牌图像清晰化、视频标注三个竞赛，共吸引了 80 余支队伍报名参加，成功架起了学术界与工业界之间的沟通桥梁，并有力推动了计算机视觉学科及相关产业的发展。具体信息可参考 RACV2016 竞赛单元网站（<http://sist.shanghaitech.edu.cn/racv2016/competition.html>）。

有意愿组织或承办竞赛的公司或单位需要于 2017 年 3 月 25 日前将竞赛申请书以电子邮件的形式将竞赛组织申请书提交给竞赛单元主席。申请结果将会在收到申请书的 2 周内通知组织者申请是否被大会采纳。所有被采纳的竞赛将会在 CCCV 2017 网站上公布。大会将会为每个竞赛分配各自的网页链接，用于公布竞赛细则、报名方式 and 竞赛结果等信息。

竞赛申请书需包括以下信息：

1. 竞赛名称、申请人邮箱
2. 竞赛目的与意义
3. 竞赛组织方
4. 竞赛参与者要求
5. 报名方式
6. 主要时间节点

7. 竞赛数据使用方式
8. 任务设置
9. 结果的评价方式
10. 结果提交方式
11. 奖项设置和奖励方法
12. 专题研讨会组织
13. 知识产权归属
14. 参赛团队注册方式

具体的申请书内容和形式可参考附件的 word 文档。申请书需要以 PDF 或 word 文档的形式提交给竞赛单元主席。

CCCV 2017 大会还将为竞赛承办方分别配专门的研讨会时间，并协助竞赛组织者对参赛的竞赛获胜者颁发奖励证书。

如果竞赛组织申请人对上述信息有任何疑问，请联系竞赛单元主席：鲁继文（清华大学，lujiwen@tsinghua.edu.cn）；左旺孟（哈尔滨工业大学，wmzuo@hit.edu.cn）

主办方：中国计算机学会计算机视觉专委会

承办方：图像视频大数据产业技术创新战略联盟

优秀论文推荐期刊

CCCV2017 论文集将由 Springer 出版，并提交 EI 检索；优秀的论文经扩展

后将推荐到 SCI 或者 EI 期刊发表：

- (1) Neurocomputing (SCI)
- (2) Pattern Recognition Letters (SCI)

- (3) International Journal of Advanced Robotic Systems (SCI)

- (4) Optoelectronics Letters (EI)

国际顾问委员会

CCCV 作为中国计算机视觉最顶级的会议，也受到国际计算机视觉协会 (Computer Vision Foundation) 和 PAMI-TC 的关注。这两个协会由往年 CVPR 和 ICCV 的程序主席组成，是国际视觉领域声誉最高的组织。今年我们有幸邀请到了 CVF 和 PAMI-TC 里 9 位国际最顶尖的学者担任国际顾问委员会。顾问委员会将组织国际顶尖学者来 CCCV 做主旨演讲，并推荐 CCCV 给国际同行，以推进中外计算机视觉学者的交流合作。

国际顾问委员会成员：

Ramin Zabih (Professor of Computer Science at Cornell University)

David Forsyth (American Computer Scientist and Full Professor at the University of Illinois at Urbana-Champaign)

Bill Freeman (Professor in the Electrical Engineering and Computer Science Department at MIT)

Jitendra Malik (Professor of EECS University of California at Berkeley)

Fei-Fei Li (Associate Professor at Stanford University)

Gerard Medioni (Professor in the Computer Science Department at University of Southern California)

Ikeuchi Katsushi (Principal Researcher, Microsoft Research Asia)

Song-chun Zhu (Professor of University of California, Los Angeles)

Long Quan (Professor of The Hong Kong University of Science and Technology)

CCF-CV

走进企业系列交流会



2017 年 3 月 3 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九期活动——走进驭势科技，在北京市海淀区成府路 45 号驭势科技会议室成功举办。

本期活动由 CCF-CV 秘书处召集，驭势科技 CMO 李雨嘉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，由于主办方场地所限，最终选择了来自企业、高校、科研院所等近 30 名代表参加了此次活动。驭势科技对此次专

委交流活动给予了高度重视和大力支持，公司创始人、CEO 吴甘沙先生亲自主讲，以“智能驾驶：值得赌上 20 年的职业”为题，围绕智能驾驶当前的发展现状、存在的挑战、未来的趋势等话题，与参会人员做了分享。

众多周知，本轮人工智能热潮不仅在学术界产生了强烈的反响，同时在产业界也催生了多个方向的技术落地开花，智能驾驶毫无疑问是其中的热宠。作为英特尔中国研究院的前任院长，吴总在其职业生涯巅峰期转身投入这一朝阳产业方向，他的所感所悟为听众们揭开了智能驾驶发展道路上的众多谜团。吴总的报告主要从四个方面切入主题，着重回答了四个“为什么”的问题，包括：为什么智能驾驶是值得赌上 20 年的职业、为什么成为智能驾驶工程师并不难、为什么这个领域

吸引 hardcore 的你、为什么进入这个领域并不晚。报告内容干货满满，精彩纷呈，以严谨独特的视角生动诠释了 AI 工程师心中“技术是为了改变未来生活”的情怀。报告结束之后，吴总与听众进行了热情的讨论互动，与大家就智能驾驶发展的前景、面临的技术挑战、交叉行业的互动等一些相关问题进行了深入讨论。

在整场活动进行过程中，驭势科技工作人员为参会人员精心准备了点心、水果等，同时为感兴趣的小伙伴们提供了众多实习的机会，很多同学已经跃跃欲试要投入智能驾驶的大潮中，跟随驭势科技一起“成为时代弄潮儿”！

活动现场气氛热烈，大家的交流讨论热度不减。活动结束后，大家一同合影留念。

三维形状数据的深度特征表示

国防科学技术大学 郭裕兰

一、引言

随着三维成像技术的快速发展，低成本小型化三维传感器近年来大量涌现并逐步配备到移动设备中，典型代表如 Kinect, Realsense 和 Google Tango。三维传感器能很好地捕获场景三维信息，使智能设备更好地感知和理解周围环境，在机器人、AR/VR、人机交互、遥感测绘等多个领域具有广阔的应用前景(图 1)。三维传感器获得的数据是物体三维形状的直观反映，其可以表示为深度图、点云或者网格等多种形式。其中，深度图只能表示物体在单个视点下的距离信息，因此常称为 2.5 维数据，而点云和网格能表示物体在多个视点下的信息，因此常称为三维形状数据。有效的三维形状特征表示是实现点云配准、三维模型重建、三维形状检索、三维目标识别、三维生物特征识别、场景语义分割、以及定位制图等视觉任务的基础。

一个好的特征应该具有良好的鉴别力、稳健性、不变性以及计算效率。自上世纪 90 年代开始，三维形状特征提取算法经历了 20 余年的发展，逐步从手工特征过渡到基于深度学习的特征。手工特征通常通过提取三维形状几何属性的空间分布或直方图统计等方法得到，典型代表如 Spin Image、FPFH、Heat Kernel Signature (HKS)、MeshHOG、RoPS 等[1]。但这类方法依赖于研究者的领域知识，无法获得适用于某一特定任务的最优三维形状特征表示。近年来，随着三维形状数据集的不断完善，深度学习模型开始应用于三维形状特征表

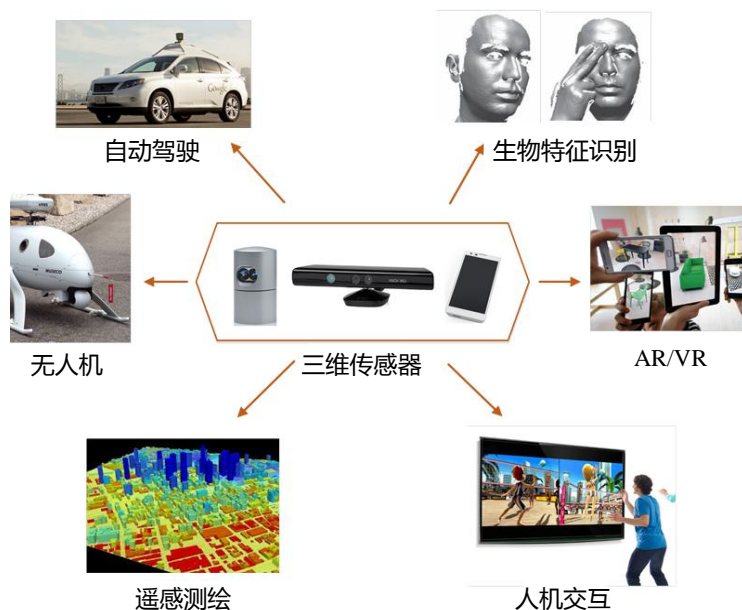


图 1 三维传感器的应用场景

示，并产生了大量的研究成果。本文旨在简要综述该领域面临的主要挑战，研究进展以及潜在研究方向。

二、面临的挑战

相对于二维图像领域，深度学习模型在三维形状上的研究起步较晚，仅在近三年取得了较大的进展。该领域面临的主要问题如下。

1. 三维形状的结构化表示

图像是结构化的，可以表示为二维平面上一个矩阵，但三维点云和网格都是非结构化的，不能直接输入到神经网络中。因此，采用深度学习模型首先要解决三维形状的结构化表示问题。主要思路包括：1) 在三维形状上手工提取低级特征，

再采用深度学习模型提取高级特征；2) 将三维点云或网格转化为系列二维图像，再采用深度学习模型提取特征；3) 将三维形状体素化得到结构化表示，再构建三维深度学习模型提取特征；4) 设计能适应原始三维数据特点的深度学习模型。

2. 公共数据集较小

相比于 ImageNet 等千万量级的二维图像数据集，传统的三维形状数据集很小。近年来发布的较大的数据集包括用于形状分类与检索的 ModelNet 和 ShapeNet。ModelNet 包含了来自 662 类的 127915 个三维形状，其子集 Model10 包含了来自 10 类的 4899 个三维形状，ModelNet40 包含了来自 40 类的 12311 个三维形状。ShapeNet 包含了约 300 万个形状，其子集 ShapeNetCore 包含了来自 55 类的 51300 个形状。较小的

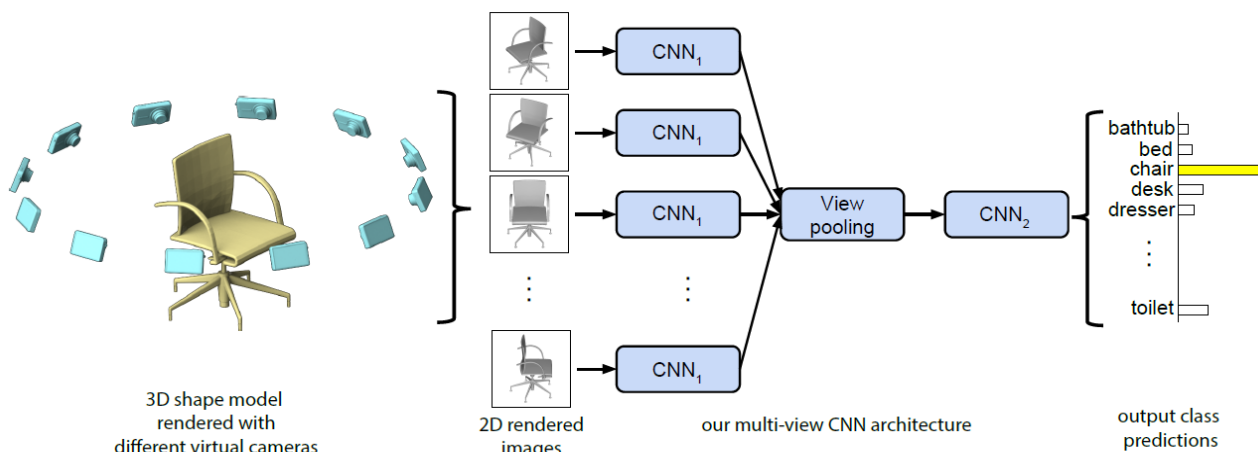


图2 用于三维形状识别的3D CNN网络结构图[4]

数据集对深度学习模型的设计与训练提出了更高要求。

3. 神经网络运算量大

由于三维形状比二维图像在空间中多了一维信息，因此在保持相同空间分辨率的前提下，神经网络的运算量比图像大很多。比如，即使将三维形状采样为 $30 \times 30 \times 30$ 的体素，其运算量已相当于 165×165 的图像，而实际上分辨率为 $30 \times 30 \times 30$ 的体素对形状表示是非常粗糙的。因此，如何既能获得精细的三维形状表示，又能将运算量控制在可接受的范围内，是一个富有挑战的问题。

4. 物体姿态变化

由于三维物体的姿态是任意的，因此如何使得学习到的特征具有对姿态的不变性也是设计深度学习模型时需要考虑的问题。典型思路包括在预处理中对输入数据进行姿态归一化，或者在神经网络中加入Max-Pooling等操作以消除姿态的影响。

三、现有方法分类

依据不同的数据表示方式，现有的三维形状深度特征表示方法可以分为：基于

手工特征的方法、基于投影图像的方法、基于三维体素的方法以及基于原始数据的方法。

1. 基于手工特征的方法

这类方法首先在三维形状上提取手工特征，进而将这些特征作为深度神经网络的输入，用以学习高层特征表示。其优势在于可以充分利用现有的低层特征描述子及深度学习模型。

比如，Bu等人[2]首先将热核特征和平均测地距离等构成的低层特征通过Bag-of-Feature模型转化为中层特征，接着采用深度置信网络(DBN)从中层特征中学习高层特征表示，并成功应用于三维形状检索与识别。Xie等人[3]首先提取三维形状Heat Kernel Signature特征的多尺度直方图分布作为自编码器的输入，然后在每个尺度上训练一个自编码器并将多个尺度隐含层的输出连接得到特征描述子，并在多个数据集上测试了该方法用于形状检索的有效性。

这类方法的缺陷在于，其仍然依赖手工特征的选择与参数优化，因此在某种程度上损失了深度学习的优势，无法从根本上克服手工特征存在的问题。

2. 基于投影图像的方法

这类方法首先将三维形状投影到二维图像空间，进而在二维图像上采用深度学习模型进行特征学习，其优势在于：1) 可以充分利用二维图像领域性能优越的网络架构；2) 存在海量图像数据供深度学习模型进行预训练。

比如，Su等人[4]首先获得三维形状在12个不同视点下的投影图，进而采用VGG-M卷积神经网络学习各个视点下投影图的特征，最后将多视点下的特征进行池化并送入下一个CNN网络中得到最终的形状特征，其网络结构如图2所示。其第一级CNN网络采用ImageNet进行预训练并利用三维形状的二维投影图进行精调，实验结果表明多视点图像能获得比单视点图像更好的性能。Shi等人[5]通过沿主轴方向进行圆柱投影将三维形状转化为多个全景图，进而利用CNN从全景图中学习特征表示。由于在卷积层和全连接层之间加入了Max-Pooling操作，因此特征具有对绕主轴旋转的不变性。Sinha等人[6]将三维形状参数化到球形表面，进而将球形表面投影到八面体后展开成二维平面，并采用主曲率或HKS在平面的分布获得二维图像，最后采用CNN网络从二维图像中学习特征表示。Kalogerakis等人[7]在不同视点

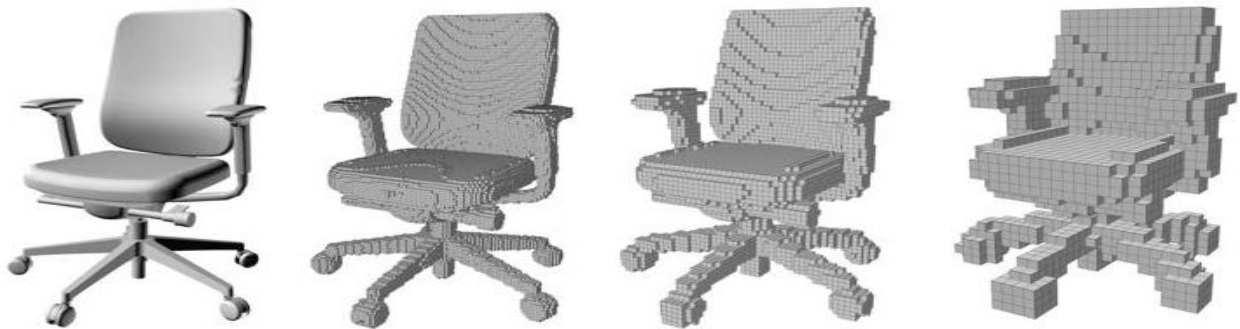


图3 椅子模型在不同分辨率下的三维体素表示[11]

及尺度下获得三维形状的系列阴影图 (Shaded Image) 和深度图, 进而采用全卷积网络 (FCN) 进行特征学习。

虽然这类方法通过变换能一定程度上保留三维物体的形状信息, 但变换过程本身改变了三维形状的局部和全局结构, 使得特征鉴别力下降。此外, 将三维形状进行投影损失了大量的结构信息, 且这类方法通常要求三维形状在垂直方向上进行了对齐。

3. 基于三维体素的方法

这类方法将三维形状视为三维体素网格中的概率分布, 从而将其表示为二值或实值的三维张量。图 3 给出了椅子模型在不同分辨率下的三维体素表示。这类方法的优势在于三维体素完整保留了三维形状信息, 有利于提高特征的鉴别力。

比如, Wu 等人[8]将三维形状表示为二值概率分布, 若体素在三维表面内则其值为 1, 否则为 0, 并采用卷积深度置信网络 (CDBN) 学习三维体素与标签之间的联合分布。Xu 等人[9]采用定向搜索对 3D ShapeNets 进行优化以减少 CNN 网络的参数数量并提高分类精度。Xu 等人[10]则把三维体素的每一层抽取出来组合成一个二值图像输入到 CNN 网络中进行特征学习。为进一步降低计算量, Li 等人[11]将三维形状表示成体素场 (Volumetric Fields) 以克

服三维体素表示的稀疏性问题, 并采用一个场探索滤波器 (Field Probing Filter) 取代 CNN 中的卷积层来学习特征。Qi 等人[12]系统比较了基于多视点投影和基于三维体素的方法, 并通过增加辅助学习任务、多姿态数据增广与池化融合等来提高三维形状分类的性能。此外, Wu 等人 [13]采用体素卷积网络和生成对抗式网络从概率空间中生成三维形状, 其通过非监督学习得到的特征能获得很好的三维目标识别性能。

但是, 这类方法也面临一些挑战, 如: 1) 为使得网络训练不过于复杂, 三维体素的分辨率不能太高 (通常为 $30 \times 30 \times 30$), 而较低的分辨率限制了所学特征的鉴别力; 2) 三维形状表面所占的体素比例不高, 使得体素化结果较稀疏, 因此需要设计合理的网络以避免大量乘 0 或者为空的运算。

4. 基于原始数据的方法

这类方法针对三维形状数据的特点设计特定的神经网络输入层, 使网络能很好地应对三维形状数据非规则化的问题。

对于网格表示, Han 等人[14]提出了一种 Mesh 卷积受限玻尔兹曼机 (MCRBM) 以实现三维形状的非监督特征学习。该方法首先在三维形状上均匀放置固定点数的节点, 并在这些节点上用局部函数能量分布 (LFED) 来表达三维形状局部区域的几何和结构信息, 进而采用卷积深度置信网

络来学习高层特征。Han 等人[15]还提出了一种圆形卷积受限玻尔兹曼机 (CCRBM) 来学习三维局部区域的几何和结构信息。该方法首先将局部区域上的点投影到该区域中心的切平面上并获得投影距离分布 (PDD) 特征, 进而采用傅里叶变换系数获得旋转不变的特征, 最后采用卷积受限玻尔兹曼机学习高层特征。

对于点云表示, Qi 等人[16]认为解决输入点云的无序问题有三种解决思路: 1) 将输入点云进行排序, 但在高维空间中难以确保排序结果的稳定性; 2) 将输入点云看作一个序列去训练 RNN 网络, 并采用不同排列组合得到的点云作为增广数据训练 RNN 网络, 但对于包含大量点的点云来说这一方法并不能获得很好的不变性; 3) 采用一个对称函数融合每个点的信息, 该对称函数以所有点作为输入但输出一个对点序不变的向量。Qi 等人[16]采用第三种思路, 以多层感知网络及 Max-Pooling 来近似该对称函数, 通过训练获得对点序不敏感的特征表示。该方法被成功应用于三维形状分类、物体部件分割以及语义场景分割。

四、未来发展方向

相对于二维图像而言, 对三维形状的深度特征表示研究才刚刚起步。随着三维成像传感器的进一步普及, 三维形状特征学习及相关应用的研究将得到更多关注。

在这一领域，依然有很多方向值得挖掘，比如：非刚性三维形状的特征学习；大规模点云的特征学习，特别是测绘制图和自动驾驶场景下的点云实时特征学习；以及遮挡和背景干扰下的三维目标检测与识别。

参考文献：

- [1] Y. Guo, M. Bennamoun, F. Sohel, M. Lu, J. Wan. 3D object recognition in cluttered scenes with local surface features: a survey. *IEEE TPAMI*, 36(11): 2270-2287, 2014
- [2] S. Bu, Z. Liu, J. Han, J. Wu, R. Ji. Learning High-Level Feature by Deep Belief Networks for 3-D Model Retrieval and Recognition. *IEEE TMM*, 16(8): 2154-2167, 2014
- [3] J. Xie, G. Dai, F. Zhu, E. Wong, Y. Fang. DeepShape: Deep-Learned Shape Descriptor for 3D Shape Retrieval. *IEEE TPAMI*, 2016
- [4] H. Su, S. Maji, E. Kalogerakis, E. Learned-Miller. Multi-view Convolutional Neural Networks for 3D Shape Recognition, *ICCV*, 2015
- [5] B. Shi, S. Bai, Z. Zhou, X. Bai. DeepPano: Deep Panoramic Representation for 3-D Shape Recognition. *IEEE Signal Processing Letters*, 22: 2339-2344, 2015
- [6] A. Sinha, J. Bai, K. Ramani. Deep Learning 3D Shape Surfaces Using Geometry Images. *ECCV*, 223-240, 2016
- [7] E. Kalogerakis, M. Averkiou, S. Maji, S. Chaudhuri. 3D Shape Segmentation with Projective Convolutional Networks. *arXiv preprint arXiv:1612.02808*, 2016
- [8] Z. Wu, S. Song, A. Khosla, F. Yu, L. Zhang. 3D ShapeNets: A Deep Representation for Volumetric Shapes, *CVPR*, 1912-1920, 2015
- [9] X. Xu, S. Todorovic. Beam Search for Learning a Deep Convolutional Neural Network of 3D Shapes. *arXiv preprint arXiv:1612.04774*, 2016
- [10] X. Xu, D. Corrigan, A. Dehghani, S. Caulfield, D. Moloney. 3D Object Recognition Based on Volumetric Representation Using Convolutional Neural Networks. *International Conference on Articulated Motion and Deformable Objects*. 147-156, 2016
- [11] Y. Li, S. Pirk, H. Su, C. R. Qi, L. J. Guibas. FPNN: Field Probing Neural Networks for 3D Data. *NIPS*, 307-315, 2016
- [12] C. R. Qi, H. Su, M. Niessner, A. Dai, M. Yan, L. J. Guibas. Volumetric and Multi-View CNNs for Object Classification on 3D Data, *arXiv*, 2016
- [13] J. Wu, C. Zhang, T. Xue, W. T. Freeman and, J. B. Tenenbaum. Learning a Probabilistic Latent Space of Object Shapes via 3D Generative-Adversarial Modeling. *NIPS*, 2016
- [14] Z. Han, Z. Liu, J. Han, C. M. Vong, S. Bu, C. L. P. Chen. Mesh Convolutional Restricted Boltzmann Machines for Unsupervised Learning of Features With Structure Preservation on 3-D Meshes. *IEEE TNNLS*, 2016
- [15] Z. Han, Z. Liu, J. Han, C. M. Vong, S. Bu, X. Li. Unsupervised 3D Local Feature Learning by Circle Convolutional Restricted Boltzmann Machine. *IEEE TIP* 25(11): 5331-5344, 2016.
- [16] C. R. Qi, H. Su, K. Mo, L. J. Guibas. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. *arXiv preprint arXiv:1612.00593*, 2016.

(责任编辑：申抒含)



郭裕兰，国防科技大学电子科学与工程学院讲师。2011年至2014年于西澳大学从事访问研究，2016年11月起于中科院计算所从事博士后研究。已在 *IEEE TPAMI*、*IJCV*、*IEEE TIP* 等国际期刊及会议上发表论文 40 余篇。曾获“中国人工智能学会优秀博士学位论文奖”和“军队优秀博士学位论文奖”，指导学生获第十一届中国研究生电子设计

竞赛第一名暨特等奖。现为 CCF 计算机视觉专委会委员，中国自动化学会模式识别与机器学习专委会委员，中国人工智能学会模式识别专委会委员。主要研究方向包括点云特征学习、三维目标识别、三维人脸识别和三维场景重建等。

Email: yulan.guo@nudt.edu.cn

运用人工智能技术对医学影像进行自动判读

深圳大学 沈琳琳

医疗数据中有超过90%的数据来自于医学影像，但是这些数据大多需要进行人工分析。人工分析的缺点很明显，第一是不精确，只能凭借经验去判断，很容易误判。中国医学会的一份误诊数据资料显示，中国临床医疗每年的误诊人数约为5700万人，总误诊率为27.8%，器官异位误诊率为60%，恶性肿瘤平均误诊率为40%。第二是缺口大，目前我国医学影像数据的年增长率约为30%，而放射科医师数量的年增长率约为4.1%，放射科医师数量增长远不及影像数据增长。这意味着我国放射科医师在未来处理医学影像数据的压力会越来越大。

深度学习自2006年提出以来，先后

在ImageNet ILSVRC大规模图像识别评测、人脸识别、以及图像理解等领域取得了突破性地进展。目前，神经网络深度学习也以极快的速度跃入医疗领域。由于拥有大量疾病诊断数据，IBM公司开发的Watson系统可以在十分钟内诊断出罕见的白血病，从而在诊断能力和速度上大大超过普通医生。Enlitic公司开发的恶性肿瘤检测系统检出精度比一名放射技师检查肺癌的精度高50%以上。美国休斯敦癌症研究中心开发的AI乳腺癌诊断软件比专业医生快30倍。斯坦福基于卷积神经网络的皮肤癌诊断算法也取得了和专业医生相当的准确率。深圳大学计算机视觉研究所开发的细胞分割与分类系统，连续在ICIP 2013



图1 医疗影像智能分析机器人

和ICPR 2016组织的荧光细胞图像分类算法大赛上获得准确率冠军。该系统基于深度学习开发，在HEp-2(Human epithelial type 2)细胞图像判读准确率比训练过的专业人员高10%，可广泛应用于免疫系统疾病诊断。

人工智能不仅能帮助患者更快速地完成健康检查，同时也可以帮助影像医生削减读片时间、提升效率、降低误诊概率。随着人工智能和医学影像大数据的逐渐普及和应用，医学影像所面临的准确度和大缺口问题可以迎刃而解。两者的融合将成为未来医学影像发展的重要方向。

(责任编辑：卢国梁)

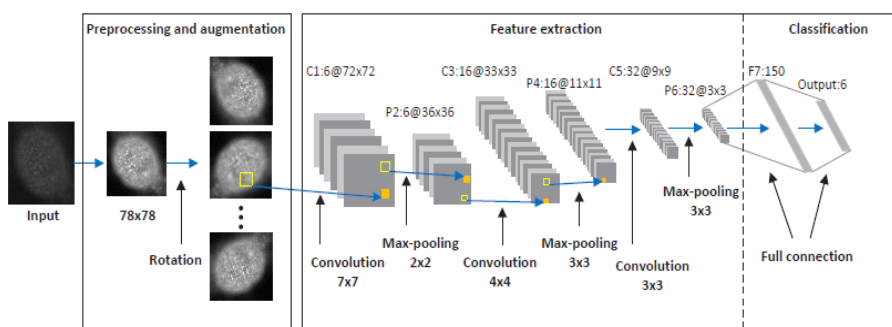


图2 基于深度学习的HEp-2细胞图像判读系统



沈琳琳，CCF 计算机视觉专委会委员。深圳大学计算机学院教授、计算机视觉研究所所长，英国诺丁汉大学荣誉教授，广东省中英合作视觉信息处理实验室、深大-大华智能系统工程联合实验室主任。主要研

究方向为人脸识别分析，医学图像处理。

Email: llshen@szu.cn

视觉显著性检测：计算机视觉领域的基础性研究

大连理工大学 卢湖川

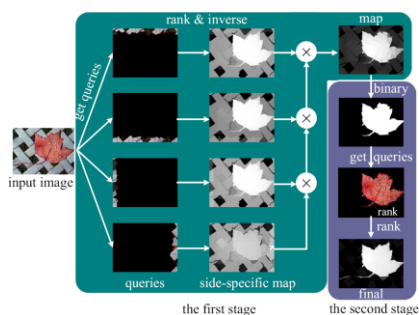


图 1 传统模型

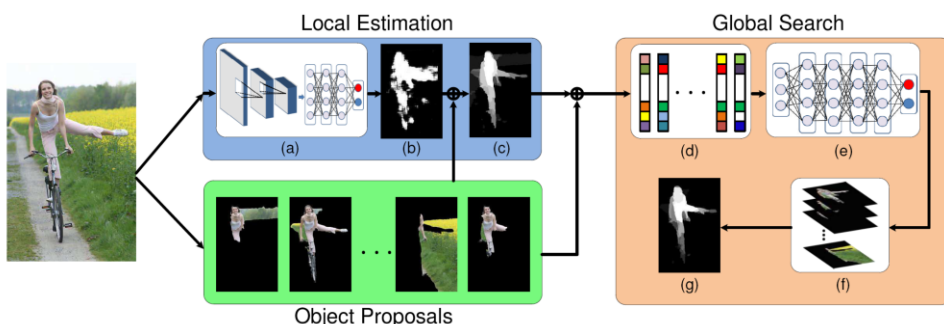


图 2 深度模型

视觉显著性检测，顾名思义就是将场景中比较突出的、容易引起人眼注意的区域检测出来。视觉显著性检测可以分为眼动点检测和显著性目标检测两大类，这里我们主要讨论显著性目标检测。显著性目标检测通常作为一种图像预处理手段，在许多领域得到了广泛应用，如目标分割、图像压缩等。传统的显著性目标检测算法大多是基于图像的颜色、纹理、对比度等特征，对于背景简单、对比度强的图像库取得了较好的效果（图 1）。而近几年兴起的深度学习方法，在传统方法的基础上，对自然图像的语义信息有了更高层次的表达，能够处理相对复杂的场景，从而得到了更加广泛的应用。

显著性检测的数据大多来源于自然

图像，基于一幅图像中至少有一个显著目标的假设，通过统计的方法进行人工标注，从而获取图像的显著性区域，并在此基础上构造有效的模型进行预测。传统方法在处理简单场景，即前背景对比度明显、前景目标特征具有一致性等理想情况时具有精度高和速度快的优点，但在处理实际问题时往往得不到理想的结果。为了克服传统方法的缺点，大连理工大学信息与通信工程学院 IIAU (Intelligent Image Analysis and Understanding) 实验室致力于将传统模型与深度学习方法相融合，提出了多个新的预测模型，以进行更加精确有效的分析和检测。该课题组自 2013 年开始陆续有 9 篇论文发表在国际计算机视觉三大会议 (CVPR、ICCV、ECCV) 上。其中，2015

年在 CVPR 发表的“Deep Networks for Saliency Detection via Local Estimation and Global Search”（图 2）和 2016 年在 ECCV 发表的“Saliency Detection with Recurrent Fully Convolutional Networks”结合了传统方法中的有效信息和深度学习方法的高适应性，在许多简单和复杂的场景下都取得了理想的检测结果。

从 CVPR 2016 和 ECCV 2016 发表的论文来看，最近该领域的发展趋势是在 FCN 的基础上，围绕 RNN 等热点网络，进一步提升显著性检测的精度。

（责任编辑：杨巨峰）



卢湖川，CCF 计算机视觉专委会常委，大连理工大学教授，大连理工大学电子信息与电气工程学部副部长。研究方向包括计算机视觉、模式识别等。目前担任 IEEE

Transaction On Systems, Man, and Cybernetics Part B 编委。

Email : lhchuan@dlut.edu.cn

由点到面：基于运动特征点的人群行为分析及其应用

西北工业大学 王琦 陈穆林

随着公共场所安全事故的不断发生，人群行为分析成为计算机视觉领域的重要研究内容。人群行为分析的主要内容是对监控视频中的人群行为进行识别与理解，该项研究对公共场所设计、社会安全保障等工作具有重要意义，对平安城市、智慧城市愿景的实现也有推进作用。

人群场景中通常存在严重遮挡，这给监测个体的准确提取带来一定困难。另外，由于个体之间往往存在动态的交互与作用，使得群体行为研究的难度也相对较大。为了减少上述问题带来的负面影响，西北工业大学光学影像分析与学习中心计算机视觉课题组提出了一种基于行为意图的动态模型对人群中的运动特征点进行建模，并利用流形学习算法研究特征点间的拓扑

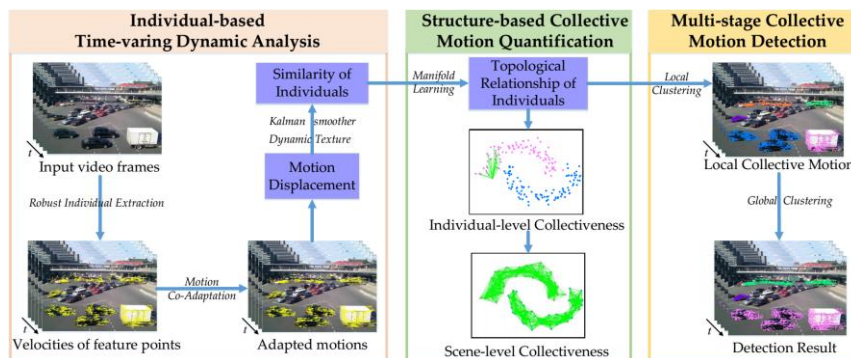


图 1 基于流形学习的群组检测与度量方法

关系，进而根据其运动模式进一步对高密度人群的运动状态进行分析，如图 1 所示。目前该方法已应用于异常检测、场景语义分割等问题中。此外，针对特征点信息过于局部化的问题，课题组提出了一种基于多特征融合的无参数聚类方法，有效结合了速度信息与结构上下文信息，并将其应

用于群组检测问题中，如图 2 所示。

该研究的相关成果自 2014 年以来已相继发表在多媒体和人工智能领域的多个主流国际期刊与会议上（包括 ACM TOMM、IEEE T-ITS、PR、AAAI）。

近年来，深度学习技术在物体检测、识别等方面取得了很大进展，受到了国内外学者的广泛关注。而利用深度特征对人群行为进行理解与分析，是有效提升群体行为表达能力的有效途径，将会成为未来一个新的研究热点。

（责任编辑：任传贤）

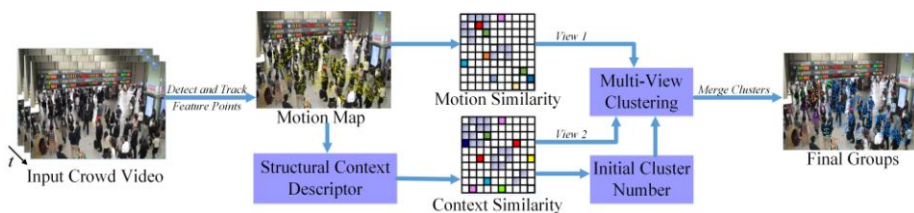


图 2 基于多视角学习的无参数群组检测方法



王琦，CCF 计算机视觉专委会委员，西北工业大学光学影像分析与学习中心（人才特区）副教授。主要研究方向为模式识别、计算机视觉、视频分析理解、遥感图像解析等。

Email: crabwq@nwpu.edu.cn

陈穆林，西北工业大学博士生，主要研究方向为机器学习与计算机视觉。

Email: chenmulin001@gmail.com

AAAI 2017 会议总结

人工智能领域的顶级会议 AAAI 2017 于 2017 年 2 月 4 日在美国旧金山召开。AAAI (Association for the Advance of Artificial Intelligence) 成立于 1979 年, 是人工智能领域的主要学术组织之一, 其主办的年会也是人工智能顶级学术会议之一。近年来, 人工智能受到了前所未有的关注, AAAI 投稿量也连续 4 年论文创新高。根据大会公布的结果, 今年 AAAI 收到投稿论文 2571 篇、录取论文 639 篇, 录取率为 24.9%、注册参会人数 1692 人, 这些数字均创近年来新高。

“中国人工智能力量的崛起”成为本届大会上被提及最多的一个话题, 在现场看到了大量来自中国的面孔。除了香港科技大学杨强教授、清华大学朱小燕教授等学术界的领军人物外, 今日头条实验室总监李磊、iPIN 创始人杨洋以及来自百度、腾讯等诸多中国产业界的行业领袖也参加了本次大会, 充分显示出华人学者在更加国际化的 AI 领域做出了斐然的贡献, 获得了更多的话语权。另外, 2016 年新增的 6 名 AAAI Fellow 中, 华人学者有 3 位; 南京大学周志华教授当选为 AAAI 2019 程序主席, 成为 AAAI 大会创办以来欧美之外担任大会程序主席的第一人。

本届 AAAI 大会论文投稿数量上中国已经多于美国, 尽管在最终录取的论文数量上还低于美国, 但是仍然远远领先于其它国家。另一个标志性的事件是 AAAI 2017 考虑到了中国春节的因素, 将原定的时间延后一个星期举行, 充分显示了中国学者在 AAAI 上的影响力。

从议程上来说, 本届大会包含有研讨会 (workshop)、教程 (tutorial)、博士生论坛 (doctoral consortium)、技术项目 (technical program)、专题讨论 (panel) 等。其中最受关注的 Panel 是关于德州扑克进展的讨论, 阿尔伯塔大学 DeepStack 的指导教师 Michael Bowling 和卡内基梅隆大学 Libratus 的指导教师 Tuomas Sandholm 教授参与了讨论。这两个程序代表了世界上最强的德州扑克计算机程序, 对不完全信息博弈有着深远的影响。另外, 强化学习领域的知名研究人员 Peter Dayan 从神经学和人工智能两个方面讨论了强化学习, 指出强化学习和神经科学可以互相激发、为彼此提供新方法和新想法。英国剑桥大学 Steve Young 教授从统计口语对话系统及机器学习的角度讨论了语音对话系统, 以及应用机器学习建立稳健的、高性能开源语音交互系统的机遇。同时, AAAI 还邀请了原谷歌自动驾驶团队对自动驾驶汽车的未来进行了讨论。推理同样是本次大会的热点之一, 大会邀请了美国 Naval 研究室的 David W. Aha 对目前深度学习的薄弱环节进行了讨论, 对目标推理的新兴应用、基础和前景进行了展望。

AAAI 2017 还颁发了多项重要奖项。其中 AAAI 杰出论文奖颁给在人工智能领域实现了广度和独特性的研究者, 比较青睐于体现学科交叉的论文。本次大会的杰出论文奖颁给了融合物理学和领域知识的神经网络非监督学习论文《Label-Free Supervision of Neural Networks with Physics and Domain Knowledge》, 其思路来自于人类的学习过程, 利用了先前的领域知识将

输出空间约束到一个特定的学习结构。这种做法使神经网络训练不再依赖大量有标签数据, 从而在不降低系统性能的前提下, 显著减少了标注工作量。

AAAI 另一个重要奖项 Blue Sky Idea 奖则为表彰那些论文中所提出的想法和愿景能够促进新的研究方向的论文。今年, 将该奖项授予了论文《Moral Decision Making Frameworks for Artificial Intelligence》。该论文探讨当我们赋予人工智能系统更多自主性时, 关于人工智能系统的伦理视角的反思, 希望通过一个框架支持人工智能系统在面临道德困境时自主做出合理的道德决策。

值得一提的是, 在本届大会特设的计算可持续发展环节中, 最佳学生论文授予了清华大学本科生尤佳轩, 以表彰其所提出的结合高斯过程的深度学习方法在谷物产量预测中的成功应用。

总之, AAAI 2017 体现了人工智能获得了各领域前所未有的关注。由学术展开深度纵向交流, 从应用出发广泛横向探讨, 以点映面, 由面概点, 人工智能的思想在产学研界交织共鸣。各国学者在人工智能各研究方向上付出了自己的勤劳和智慧, 推动着人工智能稳步向前发展, 而华人学者也已成为国际人工智能舞台上的重要角色, 用他们独特的思想共同将人工智能发展壮大。

(责任编辑 苏航)

西安电子科技大学董伟生教授访谈

2017年3月10日晚 20:00, 专委秘书处通过网络在线方式采访了西安电子科技大学电子工程系董伟生教授。

董老师, 感谢您百忙之中安排时间接受我们的访谈, 分享您研究过程中的经验!

您在图像处理方面做出了令人瞩目的工作, 能否和大家分享一下您在科研道路中获得成功的经验, 以及您取得这些成就的动力呢?

其实我没有做出什么令人瞩目的工作, 也没有什么可以分享的成功经验。我还是先跟大家分享一下我的科研经历吧。

我本科是在华中科技大学读的, 读本科时, 也没有经历过什么科研训练, 对偏理论一点的研究没有兴趣, 只是比较喜欢编程, 参加了一些电子设计竞赛方面的活动。后来到西安电子科技大学师从石光明教授读研究生时, 也主要是接触一些工程性的工作。研一有段时间导师不在, 我自己申请了个课题进行研究, 后来也捣鼓出一些成果发了篇 SCI 和一些会议论文。但当时其实还不太懂什么叫科研。真正改变我的是博士一年级导师推荐我去微软亚洲研究院吴枫老师研究组实习的经历, 这应该是引导我慢慢走向学术研究道路的一个转折点。这段实习经历虽然也有痛苦挣扎, 但开拓了我的视野, 让我深切地体会到了高水平科学研究的魅力, 立志要做高水平的科研。回到学校后, 经过导师的推荐, 我还认识了后来对我影响也比较大的加拿大麦克马斯特大学的武筱林教授, 体会到了武老师那种“一门心思都在科研上”

的专注, 感觉他时时刻刻都在想科研问题, 这给我的冲击非常大。我从他们身上学会了勤于思考, 做研究时要专注于一个问题, 不能轻言放弃, 三心二意, 我想这应该是科研道路上一个值得借鉴的经验。

后来经导师和武老师推荐至香港理工大学跟着张磊教授做项目和科研, 接触到了一些稀疏编码的内容, 这在当时看来是一个较新且热门的方向。自从开始了这方面的研究后, 我的研究工作就变得相对顺利一些。所以这也是我想和大家分享的一条经验——想要把科研做好, 选择一个好的方向特别重要。一个相对主流、前沿的方向, 相对也会更容易取得成功。

我后来取得的成果离不开前期工作的积累, 我正是把以前在西电做图像处理时候的想法应用在稀疏表达上, 才开始比较顺利。所以说科研应该是一个积累的过程, 在这个过程中我们可能会体验到很多失败和挫折, 但不能因为这些失败和挫折就轻言放弃, 而是要一直坚持下去。坚持, 也是我想和大家分享的一条经验。

我做的方向相对比较窄, 这对一个成功的研究者来说可能是一个不好的地方, 但好处是能保持对一个方向深入的思考, 做出比较有质量的成果。如果研究方向广而不精, 只在别人的工作上简单做些修改而发表成果, 这样的论文含金量并不大, 我想这是不值得学习的。

我们想请教一下董老师, 当您处于科研瓶颈期的时候, 是怎样克服这些挫折的?

我读博士的前期其实也是挺不顺利

的, 做了大量的实验, 但是效果并不太理想, 博士前两年也并没有发表论文。但是正是这一段时间的研究积累, 为我后来在做稀疏编码方法的工作奠定了基础。我也听过很多在各领域取得优异成果的研究者介绍他们的经历, 在科研道路上至始至终都非常顺利的人比较少。那些一开始经历很多挫折, 做不出成果的人, 一旦当知识和经验积累到一定程度时, 对这个问题有深入的理解, 成果可能也会随之而来。所以我觉得要养成一种积极良好的心态, 把挫折当作是一种积累。

有没有哪段研究工作让您印象特别深刻?

我印象最深的是在香港交流的那一段经历。虽然只有一年半的时间, 但是因为没有什么杂事的干扰, 能安心的做科研, 加上找到了一个比较好的方向, 所以也取得了一点成果, 发表了一些论文。其实在那边开始的时候也是不顺利, 一篇论文被拒, 另一篇论文有一个审稿人给的意见很差, 后来我突然想到博士初期做的一些小波降噪方面的东西可以应用在我当时做的稀疏编码方面的工作上, 最终打动了审稿人。这是我发表的第一篇 TIP 论文, 也是我目前为止引用次数最高的论文, 加上这过程的波折, 印象还是比较深刻的, 后来就相对顺利多了。

您曾经在香港理工大学担任 Research Assistant, 有机会感受香港的学术氛围。

您认为大陆和香港的科研区别最大的地方是什么?

我认为最大的区别可能是国内氛围

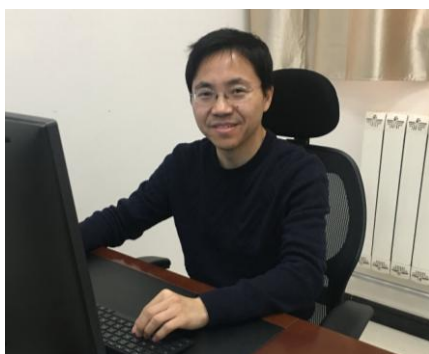
相对浮躁一点，干扰比较多。好像很多学生在国内表现并不太出色，但到了香港，或者是国外其他地方，他能做得很好。可能是因为国外的这些地方有一个比较好的科研氛围，有环境的熏陶，学生就会不自觉地静下心来学习，向更高水平的研究者靠拢。国内的学生可能由于毕业要求的限制，会更势利一些，比如说“我一定要发一篇 SCI 论文”这种想法，但是在国外他们就不在乎这个问题，他们在乎的是我能做出什么好的、有影响力的工作，而不是我什么时候能发表论文。还有一点可能是国外的学校老师们研究的课题相对来说更前沿一些，参加这些课题的学生成果产出可能比国内的学生机会更高一些。

作为 IEEE TIP、CSSP、SIIMS、FCS 等高水平期刊的编委，能否结合一些印象比较深的审稿经历，谈一下您认为论文写作中需要注意什么问题呢？

这是一个很好的问题。我认为写学术论文对于研究生，特别是刚开始进行科研工作的博士来说，是一个比较具有挑战性的问题。我一开始在论文写作方面做得也不太好，随意写出一句话而不琢磨琢磨，句子中充满着语法和逻辑上的错误。后来被武老师多次批评后，我才变得更加谨慎，写每一句话都力争意思表达清晰，数学符号等定义清楚。另外我觉得在论文写作之前，要把论文的思路理清楚，就像讲故事一样把工作的 motivation 讲清楚，突出自己工作的创新性，然后把自己的工作按各部分逻辑关系介绍清楚，实验验证充分并且结果不错，这样的话审稿人也就没有理由拒你的稿子了。

您一般会从哪几个方面来评价一篇稿件呢？您认为稿件中最重要的是什么？

我觉得最重要的还是创新性吧，这篇文章所述的工作有没有新意，应该是审稿



人最看重的。特别是现在很多来稿的论文，只不过是简单地将几个部分进行结合，或者是在前人工作的基础上做一些小的改进，给人的感觉并没有什么新意，这样的工作其实我是不太欣赏的。从我个人的科研经历来说，我觉得如果完全随波逐流地去做科研工作，能够获得能产生影响力的成果的几率比较小。好的工作需要有自己的思考，体现自己的特色，才能让审稿人有眼前一亮的感觉。但是如果你想做的和别人不一样，首先要对这个问题认识得足够深刻。

您在图像处理方面研究成果卓著，能为刚入门的青年人提供一点建议吗？

其实也只是在我做的方面上取得了一点成果，远远谈不上成果卓著。要展开研究工作，我觉得首先要选定一个合适的并且自己感兴趣的方向。刚开始入门主要还是要多读一些相关的经典论文，了解经典的算法来加深对问题的理解，但同时也要了解这个学科方向的发展前沿工作。如果只是对前人的工作做一些简单的修改，这样也许可以发表文章，但是这样的论文对业界不会产生重要的影响，含金量不高。只有当你一个问题持续地思考，那么你对这个问题就会有比一般人更深刻的认识，才有可能做出比较有创新性的工作。

如果您的学生想要涉足一些您不是很擅长的领域，您会支持他们吗？

我肯定是支持他们的。我觉得这种学

生是好学生，他能有自己的想法，可以去尝试其他的东西。对于这样的学生，我可以和他們一起去开拓这个方向，不需要局限于我原来的方向。也许我平时不会有太多时间在这个方向上做研究，但是如果他有这个兴趣，而且这个方向是可以做的话，那么我觉得这是一件好事情，我是百分之百支持的。

能否简单介绍一下您的科研团队？

目前我这边没有年轻教师，以研究生为主。每个年级我大概招 4 个硕士研究生，博士生方面目前我和石光明老师一起培养的大约有 3 到 4 人，现在我的团队规模还是相对比较小。我主要是和学生，以及其他学校的一些老师进行合作。以后有机会的话，希望可以适当扩充一下这个团队。我们团队现在主要偏向于算法性的项目上，在工程性项目上没有多少投入。

研究团队的组建和管理是研究工作顺利开展的基础，能不能分享下您在学生指导方面的心得？

我觉得目前我在这一方面做得一般。当了老师以后，感觉自己在管理方面确实存在一些问题，应该说现在还处于摸索阶段。在管理学生的问题上我也和其他老师大同小异，主要是定期检查学生工作，和他们讨论一些科研上的想法和心得，试图和他们一起解决遇到的问题。对那些对科研比较感兴趣的学生，抓得会相对紧一点，对他们的要求会相对高一些。我觉得自己做的不太好的是，对学生的指导比较细，这样我的想法可能会过多地干涉了学生的想法，甚至直接帮他们解决了他们的问题，这是不利于学生科研发展的。如果他们自己去解决问题，虽然要经过一个比较痛苦的过程，但同时他们也能学到更多、积累更多，能力会得到比较大的提升。

您对学生在论文方面是如何要求的？

对于硕士我没有什么要求，毕竟他们也没那么大的 motivation 要发很高水平的论文。我倒是希望他们能自由发展，经历一个完整的科研过程，对他们的科研能力有所提升。对于博士生，我的要求会高一点，通常要求他们发一些高水平的论文，不求在量上做突破，但是一定要保证质量。有时候，有的学生会告诉我，某某导师的学生一年发了多少篇文章，我不知道怎么回答他们。我对学生的要求比较高，我希望他们做的工作是有创新性的，不是随大流的、简单的改进、应用和结合之类的工作。

在科研上您的未来发展目标是什么呢？

我觉得还是踏踏实实做科研吧，这也是我的兴趣。我比较享受每天在实验室看论文、做实验，和学生讨论问题的生活。我对后面几年的规划主要还是希望能把科研做好，在原来的工作基础上找到一些更加重要的应用，并适当拓宽一下自己的研究方向。最近我也得到了一些资助，未来一段时间可以好好利用这笔资助发展团队，做一些感兴趣的、有意义的科研。

在未来几年里，您有想过开辟新的研究方向吗？

那肯定是会有这样的想法的，特别是以后在扩充研究团队时，招收的学生多了，他们也会有自己的想法，自然也会开辟一些新的方向。其实这几年我也有关心深度学习，但是我一直没有真正进入到这个领域，刚开始我也没有真正地理解这个东西。因为我对自己的要求是，不要跟风地去做，不要因为深度学习最近非常热门，就直接把别人的东西拿来用了，而是要把这个东西理解得足够深入和透彻，能够提出我自己的、新的问题，所以我要进入一个新

的领域，相对要慢一点。

董老师您是如何看待深度学习的呢？

其实我很早就知道这项技术，但当时对神经网络理解不深。一开始我试图去理解深度学习这些模型，思考它们这么做有什么道理，也试图去设计一些网络。可是后来由于事务比较多，也没花多少时间在上面，所以对这领域涉足还不是很深。现在对这个领域理解更深刻了一点，我觉得深度学习肯定是有道理的，是一个很好的研究方向。前几年里，大家用深度学习解决各种问题，可能偏应用的多一点。深度学习的风格跟我原来的研究风格不太一致，我原来喜欢对问题进行深入理解，结合自己的理解，设计一个数学模型，然后再通过数学优化来求解这些实际问题，希望达到比较好的效果。然而深度学习并不需要很深厚的数学模型，主要是调整网络结构，优化参数，通过数据驱动来获得比较好的结果。这跟我原来的风格差距比较大。在后来的研究工作中，我发现神经网络跟我之前研究的那些优化问题其实也比较相关，深度学习可以把我们原来那套迭代求解用网络来实现，还是比较有意思的。我也有安排一些学生从事深度学习方面的研究，不过我希望他们做出来东西能跟别人不一样，除了结果好之外还要有自己的创新。

您在时间管理上有没有什么心得体会，您是如何平衡科研和日常生活的？

在时间管理上我也没有什么心得，就是一有空余时间就工作，我比较享受工作的过程。一年中主要是过年和暑假会给自己放放假，平时也没有周末和节假日的概念，我想这是大多数科研工作者的常态。在时间管理方面，我觉得提高效率是很重要的，工作的时候就专心完成该完成的事情，这对学生来说也是一样的，在实验室

要专注。效率提高了，就能从时间里挤出更多时间来。现在我每天来回家里和学校的路上要花费比较多的时间，也需要腾出一些时间给家庭，想要在较短的工作时间里产出成果，就要依靠高度的专注和效率了。

董老师，在生活中您是什么样的人呢？能否跟大家分享一下您生活的其他方面么？

在生活方面我是一个比较随和的人，平时比较忙，喜欢听音乐，体育方面喜欢打乒乓球、看 NBA。现在有时间就跑步锻炼身体。每天下班回来陪孩子玩会，周末带孩子出去转转，呼吸一下新鲜空气。平时工作也比较忙，所以尽可能多花时间陪陪孩子。

董伟生，西安电子科技大学电子工程学院教授，博导，教育部“青年长江学者”特聘教授，国家优秀青年科学基金



获得者。2004 年本科毕业于华中科技大学，2010 年博士毕业于西安电子科技大学，曾在香港理工大学、微软亚洲研究院进行访问研究。主要研究方向为图像稀疏表示、图像处理逆问题、图像降噪和超分辨率等。发表论文 40 余篇，其中以第一作者/通信作者在 CCF-A 类期刊 IJCV、IEEE-TIP 以及 A 类学术会议 CVPR、ICCV、NIPS 上发表论文 10 余篇，4 篇论文入选 ESI 热点/高被引论文，论文被引用 2600 余次。曾获 IEEE VCIP 国际会议最佳论文奖，陕西省科学技术一等奖。担任包括国际权威学术期刊 IEEE Transactions on Image Processing、SIAM Journal on Imaging Science 在内的 3 个国际期刊的 AE。

(责任编辑：余辉 余志文 韩爱丽)

委员好消息

● 2016年12月30日,CCF组织评估专家成立专委评估工作组,对下属34个专委会(专业组)进行了评估工作,计算机视觉专委被评为优秀专委。计算机视觉专委学术会议影响力大且吸引众多知名企业赞助,积极开展走进高校活动和走进企业活动,专委网站和简报宣传科研热点,在扩大专委和领域的影响力方面做了很多工作和努力。同时获评优秀的另外4个专委是虚拟现实与可视化技术专委、大数据专委、中文信息技术专委和数据库专委。

● CCF-CV专委会委员、公安部三所物联网中心梅林主任带领的“搜神”代表队,继ImageNet2016挑战赛获得多个单项冠军后,今年首次参加COCO图像目标检测竞赛(总共有检测、分割和关键点检测三个任务),取得令人瞩目的成绩,力压CMU、Universite Paris Est等世界强队,以仅次于谷歌的成绩获得全球第二名。COCO(Common Objects in Context)是由微软发起,包括微软、FaceBook、卡耐基梅隆大学、康奈尔大学、布朗大学等多家高校和研究机构参与构建的一个图像理解数据集。

● 2017年1月11日,中国计算机学会举行CCF“中国科协青年人才托举工程”答辩会,经过专家评议,CCF-CV专委委员、南京理工大学李泽超入选2016-2018年度“中国科协青年人才托举工程”,此次共有6人入选。李泽超毕业于中科院自动化所模式识别国家重点实验室,曾获得中国科学院优秀博士论文奖、中国计算机学会优秀博士论文奖、中科院院长奖等。“中国科协青年人才托举工程”旨在鼓励32岁

以下青年科技工作者开展原创性研究,支持这些青年学者在创造力黄金期脱颖而出,做出突出业绩,努力成长为品德优秀、专业能力出类拔萃、社会责任感强、综合素质全面、具有国际视野的学术技术带头人,成为国家主要科技领域高层次领军人才和高水平创新团队的重要后备力量。

● 2017年2月22日,教育部科技发展中心网站公布了2016年度高等学校科学研究优秀成果奖(科学技术)授奖项目,共有312个项目或个人获奖,其中一等奖116项、二等奖187项、青年奖9人。获奖名单中共有CCF-CV专委委员8人(大连理工大学卢湖川,南京航空航天大学陈松灿、谭晓阳,南京信息工程大学刘青山、袁晓彤,中科院自动化所程健,厦门大学纪荣嵘,复旦大学张文强)。卢湖川教授主持的“复杂场景下显著性检测与目标跟踪”、刘青山教授主持的“视觉特征低维表达理论与方法”获得自然科学二等奖,张文强教授主持的“服务机器人关键技术及应用”获技术发明二等奖。

● 2017年2月27日,爱思唯尔(Elsevier)于爱思唯尔科技部中国区网站发布了2016年中国高被引学者(Most Cited Chinese Researchers)榜单,共收录1776名最具世界影响力的中国学者。计算机科学领域入选者最多,达152人。CCF-CV专委13位委员榜上有名,他们是:南京航空航天大学陈松灿,中国科学院自动化所刘成林、王亮、谭铁牛,南京航空航天大学谭晓阳,浙江工业大学陈胜勇,华中科技大学白翔,深圳大学沈琳琳,清华大学郭振华,哈尔滨工业大学邬向前、徐勇,

同济大学王瀚漓和南开大学程明明。

● 2017年3月7日获悉,CCF-CV专委委员、杭州电子科技大学俞俊教授(合作者芮勇、陶大程)的论文“Click Prediction for Web Image Reranking Using Multimodal Sparse Coding”获得IEEE信号处理学会2016年度最佳论文奖。IEEE信号处理学会(IEEE Signal Processing Society)是IEEE所属的最主要的学会之一,该学会颁发的最佳论文奖也是信号处理(含图像处理)领域最有影响力的国际学术奖项之一。该奖项每年从过去5年的IEEE信号与图像相关的多个著名期刊中,经权威专家评选出不超过6篇最佳论文。此次俞俊教授获奖的论文于2014年发表在《IEEE Transaction on Image Processing》第23卷第5期上。

● 2017年3月获悉,CCF-CV专委委员、上海科技大学教授、虚拟现实中心主任虞晶怡正式担任2017国际计算机视觉大会(ICCV)、国际计算机视觉与模式识别会议(CVPR)、神经信息处理系统大会(NIPS)三大人工智能会议领域主席。ICCV、CVPR、NIPS作为计算机视觉和人工智能顶级大会,是国际人工智能领域评价学术水平和研究成果的主要方式。领域主席担负大会论文筛选和评价重责,一般是由领域内颇有建树的专家学者担任。虞晶怡教授担任这三大顶会领域主席,彰显了华人在计算机视觉和人工智能领域内的国际影响力。

(责任编辑:刘海波)

数据集：3D 人脸数据集介绍

樊鑫 呼延康 大连理工大学

近年来, 3D 人脸广泛的应用于计算机视觉以及虚拟现实。针对二维人脸难以克服光照、表情、姿态等复杂问题, 基于 3D 人脸的人脸识别, 分析技术取得了很大进展。为了评价各个 3D 人脸分析算法的优劣, 需要适用大量精确采集的 3D 人脸数据集。以下对几种常用的 3D 人脸数据集进行介绍。

1 RGB-D 数据集

RGB-D 数据集采集设备通常由摄像机和测距仪组成, 这些设备同时采集目标深度图像和彩色图像。如图 1 所示, 为 RGB-D 相机采集得到的图像及对应的深度图像。



图 1 左图为彩色图像, 右图为其对应的深度图像

1.1 FRGC v2.0 数据集

FRGC v2.0 数据集为圣母大学 (University of Notre Dame) 采集, 其中共包含来自 466 个目标的 4,007 组样本, 每个目标对应 1~22 组样本, 所有的样本均为正脸。

采集设备: Minolta Vivid 900/910 series sensor(3D), Canon PowerShot G2(2D)。数据集中包含以下内容:

- Range Image : 深度图像, 分辨率为 640×480 pixels。

- Texture Image : 纹理图像, 分辨率为 640×480 pixels。
- 2D Image : 4 张控制条件下的静态图像 (2 种光照, 微笑和自然 2 种表情); 2 张非控制条件下的静态图像(采集于各种场景, 包括大厅, 走廊, 户外等)。
- Landmark : 68 个面部特征点, 由 iBug 网站 (<https://ibug.doc.ic.ac.uk/resources/faceia-l-point-annotations/>) 提供。

数据集网站:

<https://sites.google.com/a/nd.edu/public-cvrl/data-sets/>。

1.2 Texas 3DFRD 数据集

Texas 3DFRD (Texas 3D Face Recognition database) 数据集由 ADIR (Advanced Digital Imaging Research), LLC (Friendswood, TX), LIVE (Laboratory for Image and Video Engineering at The University of Texas) 采集, 共包含来自于 118 个成年人的 1149 组样本, 所有样本均为正脸。

采集设备: MU-2 立体成像系统。数据集中包含以下内容:

- Range Image : 深度图像, 分辨率为 751×501 pixels。
- Texture Image : 纹理图像, 分辨率为 751×501 pixels。
- Landmark : 25 个面部特征点, 所有人脸图像均已对齐, 其中鼻尖为图像中心。

数据集网站:

<http://live.ece.utexas.edu/research/texas3dfr/>。

1.3 EURECOM KFD 数据集

EURECOM KFD (The EURECOM Kinect Face Dataset) 组由 Multimedia Image Processing Group – EURECOM Institute 采集, 共包含来自 52 个目标的 468 组样本。每个目标的图像有 9 种不同的状态, 包括: 正常, 微笑, 张嘴, 左侧脸, 右侧脸, 遮挡眼睛, 遮挡嘴, 遮挡半张脸以及强光照。

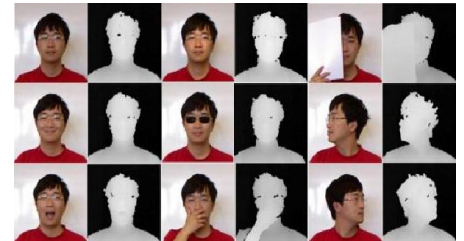


图 2 EURECOM KFD 数据集样本

采集设备: Microsoft Kinect。数据集中包含以下内容:

- Range Image : 深度图像, 分辨率为 256×256 pixels。
- Texture Image : 纹理图像, 分辨率为 256×256 pixels。
- Landmark : 6 个面部特征点, 包括左眼, 右眼, 鼻尖, 左右嘴角以及下巴。

数据集网站: <http://rgb-d.eurecom.fr/>。

1.4 3DMAD 数据集

3DMAD (3D Mask Attack Database) 拥有来自 17 个人的 76,500 帧图像, 所有图像均为正脸、自然表情。

采集设备: Microsoft Kinect。数据集中包含以下内容:

- Range Image : 深度图像, 分辨率为 640×480 pixels。
- Texture Image : 纹理图像, 分辨率为 640×480 pixels。

- Landmark : 标记眼睛位置。

数据集网站 :

<https://www.idiap.ch/dataset/3dmad>。

1.5 UMB-DB 数据集

UMB-DB (University of Milano Bicocca Database) 数据集由米兰比可卡大学采集。数据集共包含来自 143 个目标的 1473 组样本, 目标中包含有 98 名男性, 45 名女性。对于每个目标有自然、微笑、无聊、饥饿四种表情, 另外还包括头发遮挡、眼镜遮挡、手遮挡、衣领遮挡、帽子遮挡等多种遮挡情况。

采集设备 : Minolta Vivid 900 laser scanner。数据集中包含以下内容 :

- Range Image : 深度图像, 分辨率为 640×480 pixels。
- Texture Image : 纹理图像, 分辨率为 640×480 pixels。
- Landmark : 7 个面部特征点, 包括四个眼角, 鼻尖以及两个嘴角。

数据集网站 : <http://www.ivl.disco.unimib.it/minisites/umbdb/description.html>。

1.6 ND-2006 数据集

ND-2006 数据集为圣母大学 (University of Notre Dame) 采集, 其中共包含来自 888 个目标的 13,450 组样本, 每个目标包含多达 63 组样本。数据集中包含自然、快乐、伤心、惊讶、厌恶和其它等六种表情。

采集设备 : Minolta Vivid 900/910 series sensor (3D), Canon PowerShot G2 (2D)。数据集中包含以下内容 :

- Range Image : 深度图像, 分辨率为 640×480 pixels。
- Texture Image : 纹理图像, 分辨率为 640×480 pixels。

数据集网站 : <https://sites.google.com/nd.edu/public-cvrl/data-sets/>。

1.7 3D-TEC 数据集

3D-TEC (3D Twins Expression Challenge) 数据集为圣母大学 (University of Notre Dame) 采集, 其中共包含来自 214 个目标的 428 组样本。这 214 个目标为 107 对双胞胎, 其中每个目标包括自然和微笑两种表情。

采集设备 : Minolta Vivid 900/910 series sensor (3D), Canon PowerShot G2 (2D)。数据集中包含以下内容 :

- Range Image : 深度图像, 分辨率为 640×480 pixels。
- Texture Image : 纹理图像, 分辨率为 640×480 pixels。

数据集网站 : <https://sites.google.com/nd.edu/public-cvrl/data-sets/>。

2 三角网格数据集

3D 网格数据集往往是由激光或者结构光 3D 成像仪在 3D 点云基础上采用三角剖分算法获得。这类数据通常包含一个 3D 数据以及一张纹理图像。其中 3D 数据中包含 3D 点的坐标、三角网格的连接关系、纹理映射等信息。如图 3 为 3D 人脸模型上的三角网格。

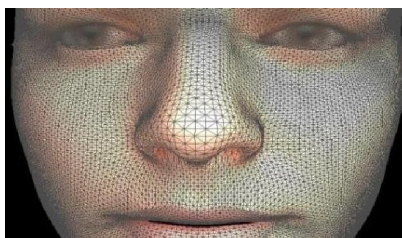


图 3 三角网格数据

2.1 Bosphorus 数据集

Bosphorus 数据集共包含有 105 个目标的 4666 组 3D 人脸数据, 其中 60 个男性目标, 45 个女性目标。每个目标有 34 种表情变化, 13 种姿态变化, 4 种遮挡情况。

采集设备 : Inspeck Mega Capturor II 3D 数字转换设备。数据集中包含以下内容 :

- 3D Data : 3D 数据, 包含 3D 点的坐标以及其对应二维彩色图像上的纹理坐标。
- Texture Image : 纹理图像, 分辨率为 1600×1200。
- Landmark : 24 个面部特征点, 包括 3D Landmark 以及 2D Landmark。

数据集网站 : bosphorus.ce.boun.edu.tr/Content.aspx。

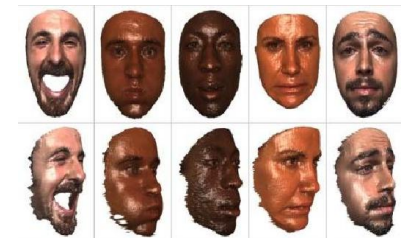


图 4 Bosphorus 数据集样本

2.2 BU-3DFE 数据集

BU-3DFE (Binghamton University 3D Facial Expression) 数据集由宾汉姆顿大学采集。共包含来自 100 个目标的 2500 组样本, 其中 56 名女性, 44 名男性, 年龄为 18~70 岁, 包含白种人, 黑种人, 东亚人, 中东人, 印度人, 葡萄牙人。

每个目标包含自然、高兴、厌恶、害怕、生气、惊讶、伤心七种表情, 除自然外每种表情有 4 种不同的强度, 因此每个目标拥有 25 组样本。每组样本包含一个正脸的 3D 模型以及两张侧脸 (-45°, +45°) 的纹理图像。



图 5 BU-3DFE 数据集样本, 分别为微笑和惊讶两种表情的 4 种不同强度

其中,采集设备:3DMD digitizer。数据集中包含以下内容:

- 3D Data:正脸 3D 模型数据,包含 3D 点的坐标以及其对应二维彩色图像上的纹理坐标。
- Texture Image:纹理图像,分辨率为 1300×900 pixels,分别为目标 -45° 以及 +45° 两张侧脸图像。
- Landmark 为 83 个 3D 面部特征点。数据集网站:http://www.cs.binghamton.edu/~lijun/Research/3DFE/3DFE_Analysis.html。

2.3 BU-4DFE 数据集

BU-4DFE (Binghamton University 4D Facial Expression) 数据集由宾汉姆顿大学采集。这是一个 3D 动态表情数据库。数据库中包含来自 101 个目标的 606 个 3D 人脸序列 (3D Video), 每个目标拥有高兴, 厌恶, 害怕, 生气, 惊讶, 伤心 6 个表情序列, 其中每个表情序列包含有 100 帧 3D 模型。每一帧中的 3D 模型大约有 35,000 个顶点。101 个目标中有 58 名男性, 43 名女性, 包含白种人, 黑种人, 东亚人, 中东人, 印度人以及葡萄牙人。

采集设备:3DMD digitizer。数据集中包含以下内容:

- 3D Video:3D 人脸表情序列。
- Texture Image:纹理图像,分辨率为 1040*1329 pixels。数据集网站:http://www.cs.binghamton.edu/~lijun/Research/3DFE/3DFE_Analysis.html。

2.4 BP4D-Spontaneous 数据集

BP4D-Spontaneous 数据集由宾汉姆顿大学采集。该数据集中共包含来自 41 个目标的 328 个视频序列(3D Video), 其中每个目标完成 8 个不同的表情任务。41 个目标中 23 个女性, 18 个男性, 年龄为

18~29 岁,包括 11 个亚洲人,6 个非裔美国人,4 个葡萄牙人以及 20 个欧裔美国人。其中采集设备:Di3D 动态人脸采集系统。数据集中包含以下内容:

- 3D Video:3D 人脸表情序列。
- 2D Video:二维人脸表情序列。
- Landmark 83 个 3D/2D 面部特征点。

数据集网站:http://www.cs.binghamton.edu/~lijun/Research/3DFE/3DFE_Analysis.html。

2.5 CASIA-3D FaceV1 数据集

CASIA-3D FaceV1 数据集中共包含来自 123 个目标的 4624 组样本。每个目标包括 5 种表情 (微笑, 大笑, 生气, 惊讶, 闭眼), 10 种姿态 (低头, 抬头, 3 种程度的左侧脸, 3 种程度的右侧脸, 左倾斜, 右倾斜, 正脸) 以及 5 种光照情况 (上, 下, 左, 右, 中)。数据集中不仅包含光照, 姿态, 表情三种因素单独作用的样本, 还包含这三种因素结合作用的样本, 对于戴眼镜的目标, 数据集另外采集戴眼镜情况下的样本。这样, 对于数据集中的每个目标, 都有 37 或者 38 个对应的样本。其中, 采集设备:Minolta Vivid 910。数据集中包含以下内容:

- 3D Data:3D 点的坐标以及其对应二维彩色图像上的纹理坐标。
- Texture Image:纹理图像。
- Landmark :包含鼻尖点的 3D 以及双眼的 2D Landmark。

数据集网站:<http://biometrics.idealtest.org/dbDetailForUser.do?id=8>。

2.6 DUT Multi-view 数据集

DUT Multi-view 数据集由大连理工大学国家示范性软件学院几何计算与数字媒体技术团队采集。该数据集中由 2D 多视角人脸图像以及 3D 视频序列(3D Video)两部分组成。其中 2D 多视角人脸

图像共包含来自 547 个中国人的 8205 个样本, 年龄分布为 18 到 25 岁。其中每个目标包含 15 个同步采集的不同视角的样本。15 个不同视角的样本分别由 15 台分布在距离目标相同半径 (1.5m) 的球面上的佳能单反相机(DSLR)采集。15 个视角的图像都在统一的光照和背景下采集。数据集中的所有样本均已手工标注 72 个面部特征点;而 3D 视频序列现包含来自 30 个目标的 30 个样本, 目前 3D 数据集处于完善过程中, 而 2D 数据集已开放申请使用。

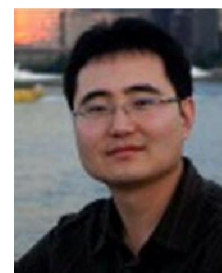
采集设备:Canon digital single-lens reflex (DSLR) cameras(2D), GI Realtime Scanner (3D)。数据集中包含以下内容:

- 2D Image:15 个视角的人脸图像。
- 3D Video:3D 人脸表情序列。
- Landmark :72 个面部特征点, 为 2D Landmark。

数据集网站:<https://github.com/dlut-dimt/Sample-of-DUT-Multi-view>。

作者介绍:

大连理工大学国家示范性软件学院几何计算与数字媒体技术团队, 由大连理工大学软件学院数字媒体技术系相关方向的教师、博士生和硕士生组成, 目前主要从事计算几何、3D 打印、计算机视觉、多媒体、机器学习、深度学习和优化方法等方向的理论与工程研究。



樊鑫, 教授、博士生导师, 主要研究方向为计算几何、图像处理与机器视觉。

呼延康, 硕士研究生, 主要研究方向为 3D 人脸重建、人脸特征点检测等技术。

(责任编辑:沈沛意 李策)

Webots 仿真机器人平台简介

机器人仿真平台因其较低的价格（相较于实体机器人）、便捷性、安全性以及丰富的机器人建模资源，被广泛应用于教学与科研领域。



图 1 Webots 仿真平台标识画面

Webots 是其中一款主要用于地面机器人建模、编程以及模拟移动机器人的仿真开发环境，如图 1 所示。Webots 的用户可以设计复杂的机械装置，并在一个共享的环境中设计一个或多个、相似或完全不同的机器人，如图 2 所示。

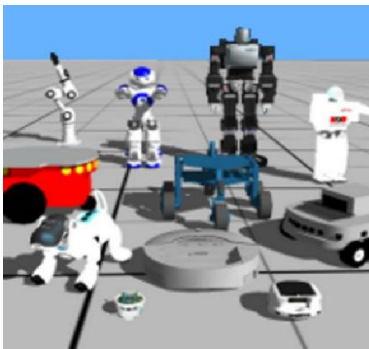


图 2 Webots 提供的部分机器人模型

Webots 提供了丰富的室内、外模拟场景与物体模型（如图 3 所示），例如：墙壁、门、窗户、灯、椅子、沙发、桌子、书架、储物柜、电视、冰箱、烤箱、苏打水、水果、书籍、植物、绘画、车、岩石、桶、足球等。

每个设计对象的性质，如形状、颜色、质地、质量、摩擦系数等等物理属性，都可由用户使用开放动力学引擎 ODE (Open Dynamics Engine) 检测物体碰撞和

李策 兰州理工大学 贾同 东北大学

模拟刚性结构的动力学特性，进行精确的设计。另外，该平台提供了大量的模拟传感器与制动器（如图 4 所示），例如距离传感器、光感传感器、触觉传感器、GPS、加速度计、摄像机、发射器与接收器、伺服电机、夹具、陀螺仪、罗盘等。



图 3 Webots 提供的部分场景模型



图 4 Webots 提供的部分感知元件

Webots 机器人的控制器可通过内部集成化开发环境或者第三方开发环境进行编程，其支持的主流控制器编程语言有：C/C++，Java，Python，Matlab。

另外，Webots 可通过所提供的 roscpp 或 rospy 控制器接口与 ROS 连接，可以使用所有 ROS 堆栈，包括 OpenCV，OpenRAVE，OpenSLAM 等。

如图 5 所示，Webots 界面有四个板块，左边是场景树 (SceneTree)，中间是 3D 窗口，右边是代码编辑区 (Texteditor)，

下边是控制台(Console)。

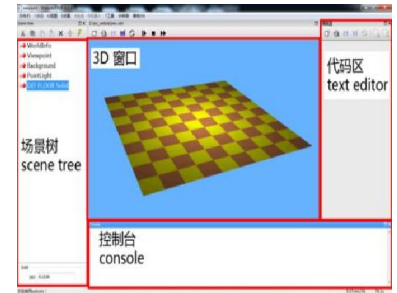


图 5 Webots 仿真设计界面示例

Webots 使用 VRML (虚拟现实建模语言) 进行建模，通过创建节点与子节点，并选择不同的节点属性(形状、物理性质、行为方式等)以及节点的组合来实现场景与机器人的建模设计。全球已有超过 1000 家高校和研究中心使用该仿真软件，支持各主流操作系统包括 Windows、Linux 和 MacOS，是一款深受研究人员喜爱的机器人仿真平台。



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、智能机器人系统等。



贾同, 博士, 副教授, 东北大学信息科学与工程学院机器学习与统计研究所。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, 医学影像计算等。

(责任编辑: 沈沛意 樊鑫)

招聘信息

一、大数据与多模态计算研究组

职位招聘信息：

(一) 研究员/副研究员/助理研究员

研究员岗位优先考虑符合国家“千人计划”、国家“青年千人计划”、中科院“百人计划”等政策条件的人选。

1、岗位名称：文本计算方向

需求人数：1人

岗位要求：

- (1) 具有计算机、统计、自动化、电子信息等相关专业博士学位；
- (2) 在自然语言处理领域具有扎实的研究背景，在国际重要刊物或会议上有与申请职位相匹配的论文发表记录，具备相应的领域影响力；
- (3) 具有严谨的工作作风和良好的沟通能力，具有较强的团队合作精神；
- (4) 具有扎实的数学基础，熟练的编程能力，能独立开展具体的科研任务；
- (5) 具有网络内容理解、网络内容安全、多媒体信息处理、社交媒体分析等专业知识的者优先；

2、岗位名称：数据挖掘方向

需求人数：1人

岗位要求：

- (1) 具有计算机、统计、自动化、电子信息等相关专业博士学位；
- (2) 在数据挖掘、大数据分析、信息检索等领域具有扎实的研究背景，在国际重要刊物或会议上有与申请教职相匹配的论文发表记录，具备相应的领域影响力；
- (3) 具有严谨的工作作风和良好的沟通能力，具有较强的团队合作精神；
- (4) 具有扎实的数学基础，熟练的编程

能力，能独立开展具体的科研任务；

(5) 开展过大数据相关工作的跨学科研发者优先；

3、待遇：

- (1) 协助申请国家、中科院的各类人才，如青年千人、百人计划等，享受国家、中科院相关政策待遇；
- (2) 一流的科研和工作环境，一流的团队支持，充足的科研经费的支持；
- (3) 实行三元结构工资(包括基本工资、岗位津贴和绩效津贴)，提供有竞争力的薪酬福利，并根据实际工作业绩进行绩效考核，实时调整；
- (4) 提供国内知名IT公司的合作机会；

(二) 博士后

博士后岗位优先考虑从事计算机视觉、模式识别、机器学习和数据挖掘方向研究的博士毕业生。

1、岗位要求：

- (1) 具有计算机、统计、自动化、电子信息等相关专业博士学位；
- (2) 具有扎实的数学基础、熟练的计算机编程能力，能独立开展具体的科研任务；
- (3) 具有严谨的工作作风和良好的沟通能力，具有较强的团队合作精神；
- (4) 研究能力突出，在本领域(计算机视觉、模式识别、机器学习、数据挖掘等方向)高水平国际学术期刊或顶级国际会议上有论文发表；

2、待遇：

工资及福利待遇按国家博士后相关规定和中国科学院自动化研究所有关规定执行。

(三) 实习生

实习生岗位招聘来自 211 或者 985

高校的硕士生或博士生。

1、岗位要求：

- (1) 计算机、自动化、电子信息等相关专业，具有良好的计算机编程能力(C/C++/Python/Matlab等)；
- (2) 了解图像视频处理、模式识别或数据挖掘等相关知识，熟悉OpenCV等常用工具；
- (3) 勤奋努力、认真负责，和良好的团队合作精神；
- (4) 实习期半年以上；
- (5) 得到导师和学校的同意；

2、待遇：

- (1) 良好的科研条件和实验环境；
- (2) 具有挑战性和实用性的项目；
- (3) 学习和实践前沿的机器学习方法；
- (4) 提供对实习过程的全面指导；
- (5) 提供一定的实习补助；

(四) 联系方式

有意者请发送个人简历并注明申请职位至：multimodal_comp@126.com

附：招聘单位简介：

大数据与多模态计算研究组隶属于智能感知与计算研究中心、模式识别国家重点实验室和中国科学院自动化研究所。研究组针对文本、图像、视频等大规模的多模态数据，进行模式识别、视觉计算、大规模机器学习、数据挖掘、大数据情境建模等方向的理论及应用研究。现因科研发展需要，诚挚邀请国内外相关领域优秀人才加盟。

(责任编辑：蹇木伟)

二、NVIDIA Deep Learning

Instructor.

Job Description : NVIDIA is searching for a world-class trainer to help create and deliver high-quality hands-on GPU programming content, focused in the area of Deep Learning.

What you'll be doing:

As an integral member of the Deep Learning Institute team, you will be responsible for:

- Developing curriculum and training content for data scientists and software developers.

- Training/certifying other instructors to teach DLI content Presenting on-site.
- Deep Learning workshops, Delivering webcasts and office hours for online courses.

What we need to see:

- Bachelor's degree in computer science/engineering, or equivalent experience.
- 2+ years of experience in Machine Learning or Deep Learning.
- Experience creating training content for developers, Comfortable presenting to technical audiences.

- Strong relationship management skills; ability to handle multiple projects simultaneously.
- Ability to work cross-functionally across many levels of a matrixed organization.
- Willing to travel (45-60%)

Contact : ahou@nvidia.com

(责任编辑：蹇木伟)

征文通知

1 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。同时,可继续关注每个国际会议举办的 workshop。

2 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) 和 Pattern Recognition Letters (PRL)。表 2 的期刊

征文中②的客座编辑有谭铁牛院士。

3 会议简介

中国计算机视觉大会 CCCV 2017 是由中国计算机学会主办,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、教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,提高国内视觉领域的研究水平。

中国计算机视觉大会 CCCV 2017 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优秀学术论文(英文),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,其中优秀论文将被推荐以 SCI 期刊的专刊形式出版。

(责任编辑：刘丽)

表 1 计算机视觉领域相关国内外会议

会议名称	会议时间	会议地点	截稿日期	会议网站
BMVC	2017.9.4-7	英国 伦敦	2017.5.2	https://bmvc2017.london/
ACM Multimedia	2017.10.23-27	美国 加尼福利亚	2017.4.10	http://www.acmmm.org/2017/
CCCV 2017	2017.10.12-14	中国 天津	2017.5.31	http://ccf-cccv.org/2017/

表 2 计算机视觉领域相关国内外期刊专刊

期刊名称	专刊题目	截稿日期
①PR	Deep Learning for Computer Aided Cancer Detection and Diagnosis with Medical Imaging	2017.08.15
②PR	Multimodal Data Analysis and Integration in Smart and Autonomous Systems	2017.05.1
③PRL	Multimodal Fusion for Pattern Recognition	2017.04.30
④PRL	Robustness, Security and Regulation Aspects in Current Biometric Systems	2017.09.30

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”。
