

CCF

计算机视觉 专委会简报

COMPUTER VISION NEWSLETTER

2017/05 期
总第 09 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	CCF-CV 专委多项新活动开启	06
会议推介	CCCV 2017 最新进展	08

科技前沿

专题综述	主动视觉及其应用	10
热点追击	二维人脸的三维感知	16
	基于运动向量卷积神经网络的实时动作识别系统	17
	局部图像特征描述	18
顶会总结	CVPR 2017 会议总结	19

委员风采

委员访谈	中山大学林惊教授访谈	20
委员好消息		23

资源平台

数据集	OUC-VISION 水下图像数据集	24
开源代码	手势识别开源代码介绍	26
招聘信息		28
征文通知		29

CCF-CV 走进高校系列报告会

第三十七期 重庆大学

时间：2017年7月1日

由中国计算机学会计算机视觉专委会主办、重庆大学软件学院和重庆大学智能服务与软件工程中心承办的第37期CCF-CV走进高校系列报告会“距离度量学习与视觉分析”，在重庆大学软件学院学术报告厅成功举行。



郑伟诗教授的报告题目为“多模态和在线哈希函数学习”。郑伟诗老师给大家介绍了其课题组最近结合行人重识别的哈希函数学习。**鲁继文**副教授的报告题目为“视觉大数据哈希学习”。鲁继文老师富有体系性的研究思路，引发了参会人员的思考。

王瑞平副研究员的报告题目为“面向视频人脸识别与检索的非线性度量学习”。王瑞平老师介绍了人脸识别主流方法的发展历程。**邓伟洪**副教授的报告题目为“人脸验证和表情识别新问题与数据库”。邓伟洪老师深入浅出的讲解，给现场的听众抛出了新的科研思路和科研方法。

四位讲者为本期CCF-CV走进高校报告会倾情奉献了四场高质量的学术报告。



第三十八期 浙江大学

时间：2017年7月16日

由中国计算机学会计算机视觉专委会主办、浙江大学计算机科学与技术学院承办的第38期CCF-CV走进高校系列报告会“计算机视觉和机器学习前沿技术及应用”在浙江大学计算机科学与技术学院成功举行。

石光明教授的报告题目为“智能催生语义通讯引导通讯技术变革”。石光明老师的报告充分体现了人工智能与传统通讯领域的结合，以及良好的研究前景。**王亮**研究员的报告题目为“人工智能时代的视觉大数据分析”。王亮老师的报告启示了人工智能与计算机视觉的趋势。**薛建儒**教授的报告题目为“动作与事件的序列分析研究”。薛老师的报告介绍了课题组基于记忆网络在视觉动作识别等问题上取得的一些研究进展。

于剑教授的报告题目为“深度学习的能和不能”。于老师的报告从一个新颖的角度深入的分析了深度学习的适用范围。**李玺**教授的报告题目为“人工智能驱动的视觉理解与计算”。报告现场激发了现场师生的兴趣和热情。

讲座现场，五位讲者与听众亲切互动和交流。报告会在热烈的掌声中圆满结束。

第三十九期 上海公安部第三研究所

时间：2017年9月6日

由中国计算机学会计算机视觉专委会主办、公安部第三研究所承办的第39期CCF-CV走进高校系列报告会“计算机视觉前沿技术及应用”在公安部第三研究所成功举行。

熊红凯教授的报告题目为“信号处理的表示学习”。报告分析了传统信号处理的重构框架与基于学习表示的研究趋势。**刘青山**教授的报告从视觉特征表示的角度分享了课题组近年来的一些研究进展。

乔宇研究员的报告题目为“面向复杂行为理解的深度学习方法与应用”。报告介绍了面向复杂行为的深度学习方法的最新进展，并分享了课题组的开源成果。**林巍峭**副教授的报告介绍了多目标跟踪的主要思路 and 现有研究进展。

此次CCF-CV走进高校系列报告圆满完成，场上场下互动热烈，高潮迭起。



（责任编辑：马占宇）

CCF-CV 走进企业系列交流会

第十一期 走进银河水滴科技



时间：2017年7月18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十一期活动——走进银河水滴科技，在北京市海淀区银谷大厦 2008 室银河水滴科技的展示厅成功举办。

本期活动由 CCF-CV 秘书处召集，银河水滴科技杨璠女士和焦琰女士负责现场组织协调。银河水滴科技对此次“CCF-CV 走进水滴科技”的活动高度重视，交流会现场，银河水滴科技 CEO 黄永祯博士和 COO 杨璠女士参加了此次活动，并且就银河水滴科技的技术产品布局、产业落地应用等方面做了专题报告。

首先，银河水滴科技 CEO 黄永祯博士以“水滴科技远距离多模态识别技术介绍”为题，全面介绍了银河水滴科技在计算机视觉领域的技术发展和产品布局。接下来，由银河水滴科技 COO 杨璠女士为大家带来分享报告：远距离多模态识别技术在平安城市及智慧商业等领域的应用。

两场报告之后是中场休息的环节，参会人员表示对远距离步态识别技术和远距离人脸识别技术非常感兴趣，纷纷进行现

场体验，并和工作人员展开了热烈的讨论。

之后银河水滴科技 CEO 黄永祯博士作为主持人，邀请来自北京航空航天大学的王蕴红教授、清华大学的鲁继文副教授、北京邮电大学的邓伟洪副教授、中科院计算所的王瑞平副研究员，在现场展开了 AI 领域的专题 Panel 讨论。

此次活动在热烈的气氛中结束，大家觉得经过分享和讨论，对于深度学习领域有了更深入的了解，希望银河水滴科技能够发展的越来越好，也希望 CCF-CV 能继续坚持举办此系列的活动，提供校-研-企-资交流的平台，助力整个人工智能行业的发展。最后与会人员一起合影留念。

第十二期 走进地平线

时间：2017年8月25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十二期活动——走进地平线，在北京市海淀区中关村海龙大厦地平线会议室成功举办。

本期活动由 CCF-CV 秘书处召集，地平线市场部市场经理赵琦负责现场组织协调。地平线对此次专委交流活动给予了高度重视和大力支持，联合创始人兼软件副总裁杨铭博士领衔，资深 IC 工程师谭洪贺

博士、智能驾驶高级工程师伟强博士联袂参加了本次活动，并分别就地平线科技的发展历程与研发概况、人工智能芯片的高能效设计和实现、地平线自动驾驶解决方案等作了专题报告。

首先，地平线联合创始人兼软件副总裁杨铭博士重温了地平线的成长历程和发展航路，全面介绍了地平线围绕人工智能产业发展所进行的方向布局和技术研发。

接下来，由地平线资深 IC 工程师谭洪贺博士为大家带来人工智能芯片的高能效设计和实现的技术分享报告。

随后，地平线智能驾驶高级工程师伟强博士带来了关于地平线自动驾驶解决方案的专题报告。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，地平线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地平线的各项精选技术演示，充分体会地平线科技产品的独特魅力！

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。

(责任编辑：王瑞平)

CCF-CV 走进高校活动征集

借 AI 东风, 布 CV 之道-- CCF-CV 走进高校系列报告 2018 年度开始申请啦!



为了更好地推动计算机视觉学科专业领域的学术与技术交流, 促进国内外学者间的了解与合作, 全面推动国内计算机视觉的学科发展, 提升我国计算机视觉研究在国际领域的影响力, 中国计算机学会计算机视觉专委会在全国范围的高校和科研院所等开展了 CCF-CV 走进高校系列报告会活动。

自 2015 年 11 月开始以来, 本活动得到了高校、讲者、听众的大力支持。截至 2017 年 9 月底, 本活动已成功举办了 40

届, 遍及祖国大江南北 23 个城市, 邀请讲者 89 人, 分享专题报告 161 场, 活动现场人均听众近 200 人次, 微信公众号平均阅读 1000 余次, 在征得讲者同意的前提下, 于爱奇艺视频网站共享报告会现场盛况 40 多小时, 在国内外计算机视觉领域引起了强烈反响。在这里, 向对活动顺利开展提供支持 and 帮助的承办方以及分享精彩报告的讲者们表示由衷的感谢!

图为 CCF-CV 走进高校系列报告会活动分布图。



值此金秋时节, 我们面向全国的高等

院校、科研院所等机构征集 2018 年度活动申请, 希望能够进一步贯彻落实《新一代人工智能发展规划》纲要中有关计算机视觉领域发展的精神, 继续推动计算机视觉领域的学术交流、思想碰撞, 为广大师生提供更多近距离接触行业领域专家、互动沟通的机会。

请感兴趣的申请人下载活动申请表表: ccfcv.ccf.org.cn/resources/ 填写相关信息后发给专委会秘书处。联系方式:

马占宇 mazhanyu@bupt.edu.cn

毋立芳 lfwu@bjut.edu.cn

CCF-CV 网站: <http://ccfcv.ccf.org.cn/>

视频分享

<http://www.iqiyi.com/u/1235027067/v>

(责任编辑: 马占宇)

CCF-CV 提名与奖励工作组成立

为发挥中国计算机学会计算机视觉专委会的学术评价作用, 激励、调动计算机视觉及相关领域专业工作者的积极性, 促进计算机视觉及相关领域的创新和进步, CCF-CV 设立中国计算机学会计算机视觉专委会奖 (简称 CCF-CV 奖)。

为了尽快落实 2017 年 6 月专委会常务会议确定的成立 CCF-CV 奖工作组的决议, 2017 年 8 月 26 日 CCF-CV 提名与奖励工作组正式成立。工作组商议了有关奖

项的设置、目的、对象、提名办法以及启动时间等议题, 其中明确了今年将要启动的奖项的内容、名称以及评选时间与方式。

组织架构:

工作组组长:

查红彬 北京大学

工作组成员 (按姓氏拼音顺序):

贾云得 北京理工大学

刘青山 南京信息工程大学

林宙辰 北京大学

马占宇 北京邮电大学

王亮 中国科学院自动化研究所

王涛 北京爱奇艺科技有限公司

张艳宁 西北工业大学

(责任编辑: 马占宇)

CCF-CV 青年工作组成立

为了更好地调动计算机视觉领域青年学者的积极性,活跃青年学者间的学术交流,专委会决定成立 CCFCV 青年工作组。工作目标包括如下:1)扩大专委会对青年学者的影响力,吸引优秀青年学者加入专委会;2)宣传青年委员的成果,加强青年委员的交流与合作,促进青年委员快速成长;3)协同专委会组织各项专委会活动。

组织架构:

工作组组长:

白翔 华中科技大学

副组长(按姓氏拼音顺序):

邓成 西安电子科技大学
金城 复旦大学
李策 兰州理工大学
苗启广 西安电子科技大学

工作组成员:

白翔 华中科技大学
丛杨 中科院沈阳自动化所
查正军 中国科技大学
邓成 西安电子科技大学
方显春 上海大学
金城 复旦大学
姬艳丽 电子科技大学
李策 兰州理工大学

刘光灿 南京信息工程大学
苗启广 西安电子科技大学
王亮 中科院自动化所
张姗姗 南京理工大学
郑伟诗 中山大学

(责任编辑:毋立芳)

CCF-CV 网站招聘志愿者

为了更好的活跃专委会网站建设,使网站真正起到宣传专委会、宣传委员、活跃计算机视觉领域研究与交流的平台作用,特征集网站编辑志愿者若干,具体需求如下:

1. 网站英文编辑 2 名,负责网站英文版宣传页面的翻译与校对;
2. 委员好消息编辑 2 名,负责邀请有成绩的委员将自己的好消息整理成稿,或者将已有新闻整理成稿,在网站相应栏目发布;
3. 委员风采编辑 2 名,负责委员采访,

并将采访稿件在网站相应栏目发布;

4. 技术报告编辑 2 名,负责邀请本领域顶会或顶刊发表论文的作者撰写文章介绍相关领域工作进展;

5. 数据资源编辑 2 名,负责收集 cv 领域的的数据,代码,评测资源等,并在专委会网站相应栏目分享;

6. 学术信息编辑 1 名,负责提供视觉领域最新学术会议和专刊征文信息。

7. 技术交流编辑 1 名,负责收集编辑企业的技术需求和成果相关信息和链接。

8. 人才需求编辑 1 名,负责收集相关领域的招聘信息。

为了鼓励大家积极参与,网站编辑志愿者将计入委员年度贡献。

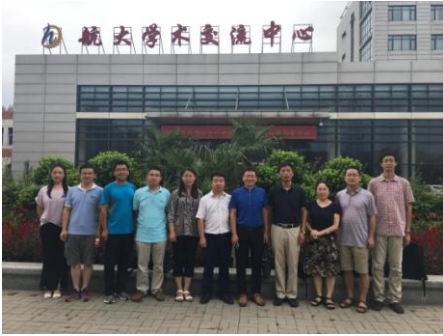
感兴趣的委员请通过邮箱联系。联系方式:

毋立芳 lfwu@bjut.edu.cn

马占宇 mazhanyu@bupt.edu.cn

(责任编辑:毋立芳)

CCCV2017 大会筹备进展交流会



2017年8月23日,中国计算机学会计算机视觉专委会(CCF-CV)秘书处由秘书长王亮研究员带队访问了中国民航大学,此行主要实地交流将于10月份举办的“2017中国计算机视觉大会”(CCCV2017)各项会务筹备情况,与论坛本地组委人员商讨落实会议各项组织事宜。中国民航大学杨金锋教授(CCCV2017大会主席)、贾桂敏教授(CCCV2017财务

主席)作为主办方负责人对全程交流活动做了周密细致的安排,天津大学胡清华教授(CCCV2017程序主席)、南开大学程明明教授(CCCV2017程序主席)、天津大学徐超教授(CCCV2017展示主席)等大会组委共同参加了交流讨论。

交流会议的日程安排紧凑而充实。交流团一行围坐一起就会务组织中的各项事宜进行了充分详细的讨论,包括会场布置、志愿者服务、会议材料印制、参会注册、出版论文提交、会议赞助、参会人员食宿交通等议题均进行了——落实,并最终形成了交流会书面记录。

总的来看,在大会主席与各位程序主席的带领下,目前会议各项筹备工作均在有条不紊进行中,会议注册系统已于近期

对外开放,热忱欢迎国内外计算机视觉同行们注册参会,金秋十月相聚在海河之滨,共享视觉盛宴!



关于会议动态,请随时关注会议官方网站:<http://ccf-cccv.org/2017/>,及CCF-CV专委会微信公众号。

(责任编辑:王瑞平)

CCCV2017 Demo 展示征集

2017年中国计算机视觉大会(CCCV 2017)邀请学术界研究人员和行业内的从业人员提交展示申请,介绍计算机视觉方向的最新研究成果和应用系统。这为计算机视觉方向的工作人员及单位提供一个难得的机会:在数千名CCCV 2017会议参与者面前展示直接展示自己的成果。所有提交的内容将接受审查。

学术研究方面展示的申请不局限于被CCCV 2017和其研讨会接收的文章,我们欢迎任何来自于其他科研机构或个人的最新计算机视觉方法的研究展示。

工业应用方面的展示申请不局限于CCCV 2017相关赞助单位,我们欢迎业内的同行们展示最新的计算机视觉相关产品。

展示活动将会与大会同期举行(10月12日-14日,2017)。

感兴趣的展示申请人需按照相关要求填写展示提案表并附相关辅助材料一并发送电子邮件给展示会: cccv2017_demo@sina.cn。本次展示的申请必须注册到CCCV 2017。组委会将根据申请内容与CCCV 2017的相关性及其本身

价值进行选择。

日期安排

提交截止:2017年9月25日

接收通知:2017年9月28日

演示日期:2017年10月12-14日

更多信息请访问 <http://ccf-cccv.org/2017/>或联系展示: cccv2017_demo@sina.cn

(责任编辑:毋立芳)

CCCV 2017

中国计算机视觉大会开始注册啦

大会注册链接

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

大会注册流程

一、被 CCCV2017 会议录用的论文至少一位论文作者进行全注册或会员注册才能覆盖本篇论文，否则您的论文将不能发表。

二、作者须在 2017 年 9 月 25 日前(含当日)向会议注册并缴纳注册费，每份注册费只能用于一篇论文，

三、特别声明：仅选期刊发表，不入会议论文集的文章作者，也须进行大会注册，并张贴或报告相关论文，否则期刊将撤出相关稿件。

大会注册支付方式

CCCV2017 组委会委托“天津紫苑策划发展有限公司”代收本次大会的会务和赞助等相关费用，支付方式如下：

1、支付宝转账

姓名：买洁

支付宝账号：mai jie0508@126.com

2、银行汇款（或单位转账）

开户名称：天津市紫苑策划发展有限公司

开户银行：中国工商银行天津第一支行

开户账号：0302 0902 0901 0185 314

3、转账和发票问题联系人

买洁：18622540929

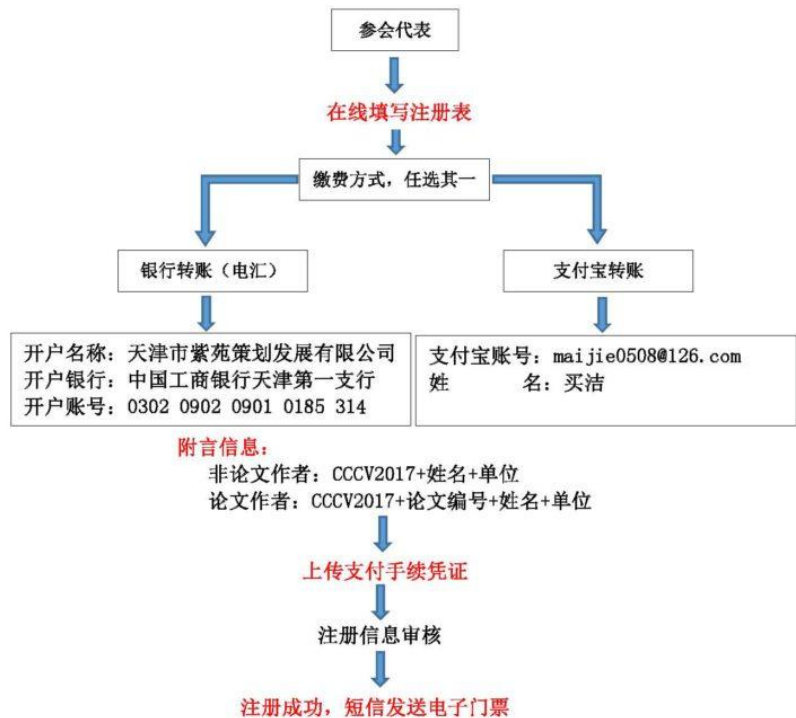
贾桂敏：18920580716

四、CCCV2017 会议（2017 年 10 月 11 日 - 14 日）注册费标准如下表。表中，会员代表是指：中国计算机学会会员。学生代表是指：全日制在校本科生、硕士生和博士生，不含在职研究生和博士后。

CCF CCCV2017 会议会务费 (10 月 12 日-14 日)				
注册时间	全额注册	会员代表	学生代表	备注
9 月 25 前	2000 元/人	1600 元/人	800 元/人	必选
9 月 25 后	2500 元/人	2000 元/人	1000 元/人	必选
现场注册	3000 元/人	2400 元/人	1200 元/人	必选
迎宾会 (10.12 日晚)	一律 400 元/人			任选
超页费 (15 页以上)	800 元/页			超页必选

五、2017 年 10 月 11 日：计算机视觉技术发展前沿讲习班，本讲习班不收取费用，欢迎大家踊跃参加！

六、大会注册分有网页版、手机版种模式。为了使注册过程流畅，请参会人员按如图步骤。



(责任编辑：毋立芳)

主动视觉及其应用

清华大学 徐静

一、引言

近年来,随制造业发展,对工件三维尺寸的测量速度和测量精度要求越来越高,主动视觉,即基于结构光的三维测量技术也因此成为研究热门。相比于传统的接触式测量及被动视觉测量技术,结构光三维测量具有非接触、精度高、密度大、匹配简单等特点,而在工业检测、质量控制、逆向工程和医学工程等领域有广泛应用。结构光三维测量系统由相机投影仪和计算机组成,结构光是指经过计算机特别设计,具有特定结构的编码图案,不同的编码方法对应三维扫描过程中不同的解码方式,适用于不同的使用环境。编码图案经过投影仪投影和相机采集后,需要结合标定结果进行三维重建,而标定精度直接决定了系统的测量精度,因此结构光的标定方法被大量研究。在此基础上,本文对结构光的应用进行了介绍与相关分析。

二、结构光测量原理

基于结构光的三维测量系统一般由相机投影仪和计算机组成,如图1所示。通过将结构光编码图案投影至待测物体表面,结构光图像会因为待测目标的深度信息而发生不同程度的畸变,再利用相机采集图像,计算机对采集的图像上每个像素进行解码,获得每个像素对应的相位,进而建立物体表面三维坐标点与编码图案及采集图像像素的对应关系,然后结合标定

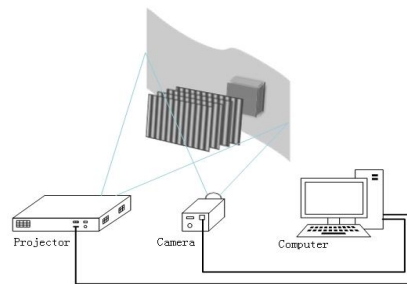


图1 结构光测量系统

结果进行计算,实现待测物体的三维重建。

基于结构光的三维测量系统实现待测目标三维重建过程的核心在于:1)根据编码图案信息获得相机与投影仪像素间的点匹配函数;2)利用标定结果,结合相机投影仪的点匹配函数,解码物体的三维坐标。因此,本文主要对编码原理和标定方法进行介绍。

三、结构光编码方法

结构光设计编码图案的目的是更方便的建立相机投影仪图像像素间的匹配关系。编码策略主要包括空间编码、直接编码和时间编码三种方法。

1. 空间编码

空间编码法根据每个点周围邻近点的信息(如颜色、梯度等)对该像素点进行编码。它只需要投影一幅编码图案[1],因此大部分情况下适用于运动物体表面三维信息的测量。但由于每个像素点的解码过程依赖于邻域点的信息,在不连续的目标表面会因为信息缺失而解码失败,解码困

难且较难获得很高的三维扫描密度。

2. 直接编码

直接编码法利用每一个像素点自身信息进行编码,不需要额外附加信息,一般需要一至两幅编码图案[2]。其原理是根据像素点自身特别的色彩值解码,所以编码图案要求大量的颜色值。但由于图像像素颜色值不仅仅取决于投影的颜色,还与物体表面的颜色直接相关,易受噪音影响,解码分辨率较低,难以直接运用于工件的三维扫描。

3. 时间编码

相比于空间编码和直接编码方法,时间编码法应用更为广泛[3]。时间编码法是通过按顺序投影一系列编码图案,对多张编码图案进行解码获得每个像素的编码信息,具有解码简单、测量精度高、分辨率高等优点。时间编码法代表方法有格雷码和相移编码法等。

3.1. 格雷码

格雷码是二值编码方法的改进[4],通过投影 m 张编码图案,可获得长度为 m 的二值码序列,最多可对 2^m 宽度的图像像素进行唯一表征和解码。格雷码方法编码简单,但为获得高分辨率往往需要投影大量编码图案,实时性较低。

3.2. 相移编码

相移编码法[5,6]则通过在被测物体表面投影一系列正弦条纹图案,再由相机捕捉经物体反射后的变形图案。其中变形条

纹包含了相位变化的信息，而相位信息则包含了物体高度信息。因此只需根据变形图案解码出相位值，即可获得物体高度信息，实现物体的三维重建。由于正弦条纹解码的限制，相移编码法至少需要三张编码图案，且随编码图案的增多，三维扫描的抗干扰性和测量精度获得提高，但测量时间也相应增加。此外，计算机对正弦条纹的解码限制相移编码法的相位在 $(-\pi, \pi)$ 范围，因此需要额外的相位解包络 (phase unwrapping) 过程。一种方法是通过在与相位变化垂直的方向投影一条直线作为参考标记[7]，并通过连续性假设获得相对于参考标记的绝对相位。但相移编码法解码出的相位依然存在二义性问题，虽然可以通过减小相移频率解决，但低频相移易受噪声影响，从而降低测量精度。

3.3. 其他编码方法

格雷码与相移结合的编码算法与多频相移法能够解决相位二义性问题。格雷码与相移结合算法[8]通过格雷码将图像划分为更小的图像，然后利用相移法求解每一小图像中的相对相位，进而获得绝对相位。多频相移法[9,10]通过投影至少两组不同频率的相移图案，分别计算对应相对相位，再利用不同频率编码图案间相对相位关系求解绝对相位。虽然这两种方法具有很强的抗干扰性和很高的测量精度，但需要大量的编码图案，时间较长，主要用于静止物体的测量。

四、结构光标定方法

精度在物体的三维测量系统中往往是评价的标准。而对于基于结构光的三维测量系统，测量精度取决于标定的精度，因此标定方法对于结构光系统有着至关重要的作用。

已有的结构光三维测量系统的标定方法主要有三类：基于矩阵变换的摄影测量法，基于几何关系的三角测量法和多项式拟合法。其中，前两种方法以小孔成像模型作为三维测量系统的数学模型，两者区别在于对投影仪的描述不同，前者直接将投影仪近似为小孔成像模型；而后者则利用几何关系添加了系统约束，简化参数后再对投影仪进行描述。多项式拟合法则抛开成像模型，直接利用拟合方法建立物体三维坐标与解码相位之间的关系。

1. 基于矩阵变换的摄影测量法

基于矩阵变换的摄影测量法建立在小孔成像的基础上，需要对相机内外参数和投影仪进行标定。相机内外参数的标定较易获得，但投影仪的标定方法分为两种，分别是标定投影仪的内外参数，和标定投影仪的光平面方程，前者包括伪相机法和逆向相机法，后者则主要是光平面法。

1.1. 伪相机法

伪相机法将投影仪看作伪相机，投影仪的编码图案即为伪相机拍摄图像，所以投影仪标定被转化为已知标定板的三维坐标 (x,y,z) ，求对应投影仪像素坐标 (i,j) 对应的过程。Zhang 和 Huang[11]通过分别投影横向和纵向的相移编码图案至棋盘格标定板，获得投影仪与相机图像像素的匹配关系，建立一对一映射。其优点在于投影仪相机标定相互独立，误差不会传播，但匹配精度有限。Li 和 Shi 等[12]在此基础上投影 25 张图案，减小了匹配误差，但测量速度有所降低。Gao 和 Wang 等[13]通过红蓝棋盘格求解相机投影仪图像间的单应性矩阵，获得转换关系，但标定精度不高。

1.2. 逆向相机法

逆向相机法将投影仪看作逆向相机，标定过程被转化为已知投影仪像素坐标

(i,j) ，求对应三维坐标 (x,y,z) 的过程。Falcao 和 Hurtos 等[14]利用投影仪向标定板投影编码图案，相机捕捉不同位姿处的标定板图案。先恢复出相机参数，然后求棋盘格角点三维坐标，然后计算投影仪内外参数，因此，投影仪的参数依赖于相机参数的计算，测量精度低于伪相机法。

1.3. 光平面法

光平面法需要求解光平面方程，然后利用线线相交或线面相交的原理求出待测点三维坐标，主要包括多点法和两线法。其区别在于多点法[15]根据光平面上的特征点计算光平面方程，而两线法[16]通过光平面与不同平面交线计算光平面方程。光平面法仅针对于线移编码方法且依赖于相机参数，标定耗时较长。

2. 基于几何关系的三角测量法

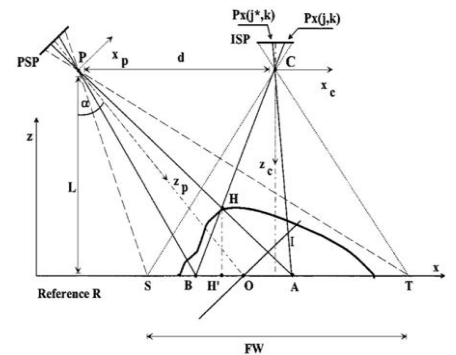


图 2 文献[17]中标定方法

基于几何关系的三角测量法利用了系统的几何关系，建立待测点三维坐标与参数间联系。其代表性方法由 Sansoni 等[17]提出的，如图 2 所示。图中左侧为投影仪，右侧为相机，系统安装要求相机投影仪光心连线 PC 平行于参考平面，光心间距离为 d，参考平面与相机光轴垂直且距离为 L，投影仪相机光轴相差角度为 α 。因此，待测点 H 深度可根据相似三角形计算：

$$Z_H(x,y) = \frac{L \cdot S_R(x,y)}{d + S_R(x,y)}$$

$$S_R(x, y) = |x_A - x_B|$$

A,B 两点关系如图 2, 两点沿 x 轴间隔可通过解码的相位与系统参数确定。其中, d 需要经过精确测量获得, L 与 α 需要进行标定计算得到。从上可以看出, 虽然标定与计算过程得到简化, 但相机投影仪的位置要求严格, 安装十分困难。

Hu 和 Huang 等[18]提出了一种有 9 个参数的标定算法, 但对系统安装同样有很高要求, 较难实现。Jia 和 Zhang 等[19]提出了一种包括相机投影仪焦距参数、基线长度、以及相机投影仪光轴与世界坐标系 x 轴夹角的标定算法, 虽然该系统对安装没有要求, 但标定过程中基线长度和夹角参数难以测量。

综上所述, 基于几何关系的三角测量法虽然避免了投影仪参数的标定, 但存在部分参数难以测量或系统设备安装困难的缺点。

3. 多项式拟合法

多项式拟合法原理较为简单, 根据像素点(u,v)及该点处相位 φ 与对应的待测点坐标(x,y,z)——对应的原理, 假设三维坐标(x,y,z)与像素坐标(u,v)及相位 φ 之间关系可用多项式表示, 建立多项式拟合模型, 进而利用实验确定拟合参数, 建立二维图像坐标与三维空间坐标之间的映射。

Lendray 等[20]假设待测点坐标(x,y,z)与像素坐标(u,v)及相位 φ 之间可直接用三次多项式表示, 需要确定 3 个三次多项式共 57 个系数。若要提高标定精度, 采用四次多项式, 则需要确定一共 102 个系数。显然直接拟合进行标定工作量大, 耗时很长。

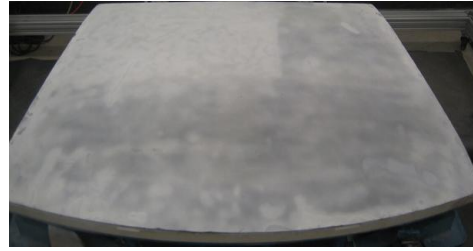


图 3 文献[24]质量检测 (左为实物, 右为检测结构)

Liu 等[21]通过推导简化了多项式模型:

$$\begin{cases} x = a_1 z + a_0 \\ y = b_1 z + b_0 \\ z = \frac{m_1 \varphi + m_0}{n_1 \varphi + 1} + \sum_{0 \leq i \leq 6} \frac{g_i - h_i \varphi}{n_i \varphi + 1} \left(\frac{m_1 \varphi + m_0}{n_1 \varphi + 1} \right)^i \end{cases}$$

其中, 一共需要标定 21 个参数, 为获得精确的标定结果, 需要利用高精度的平移装置对标定板深度方向 21 个不同位置进行测量, 耗时依旧较长。

Tavares 等[22]根据实验经验建立深度信息与相移间的线性模型:

$$z(x, y) = \varepsilon + \lambda \Delta \phi(x, y)$$

通过选取深度为 0 的参考平面 R1 以及旋转一定角度后的斜平面 R2, 分别测量两平面上的位相分布, 经过最小二乘拟合得出系数 ε 和 λ 。此方法虽然测量过程较为简单, 但精度有所降低。

Vargas 等[23]提出先标定相机参数, 然后再移动标定板计算标定板位置, 进一步拟合深度与相位之间多项式, 但标定过程依旧较为复杂。

总而言之, 多项式拟合法优点在于不需要对相机投影仪的具体参数进行标定, 对结构光测量系统的安装没有特殊要求。但需要精确的辅助设备 (如平移装置等),

成本较高; 高精度的标定结果则对应需要大量的工作量, 耗时很长。

五、结构光的应用

基于结构光的三维测量技术作为物体表面三维重建最可靠的技术, 具有非接触、速度快、结构简单、精度高等优点, 在工业检测与质量控制等方面有广泛应用。结构光测量由于非接触性能够无伤成像, 因此能应用于医学工程与建模领域。结构光三维测量实时性强, 可应用于实时交流、手势识别、人机协作等交流合作领域。而选用特殊光源或采用激光线结构光等方式, 抗干扰强, 能够实现复杂地形环境的三维重建。此外, 结构光三维测量技术获得的三维成像信息相比二维更加丰富, 因此在特征识别等领域也有所应用。

以下首先介绍了结构光三维测量系统在制造业中的应用, 然后对结构光的拓展应用进行了相关介绍。

1. 结构光在质量检测中的应用

基于结构光的三维测量技术具有速度快、精度高、分辨率高等特点, 能够满足制造业的产品质量检测要求。Xu 等[24]提出一种实时测量编码算法, 实现大工件目标的快速高精度三维检测, 如图 3 所示。

2. 结构光在医学工程中的应用

二维成像技术在医学中已有大量应用,如胃镜、X光等。而结构光三维测量技术相比二维成像能够额外捕捉到深度信息,有助于医生更好地进行诊断。如 Armbruster 和 Scheffler 等[25]设计了一种投影线结构光的内窥镜,通过分析变形测量内腔三维形貌; Bendall 等[26]设计一个小型化的相移法投影系统,能够替代内窥镜的镜头实现人体内腔的三维重建。此外,结构光技术也可用于器官的三维监测,如 Iyengar 等[27]利用结构光实现对牛心脏器官的外界监控,从而分析心脏的动态特性,如图 4 所示。

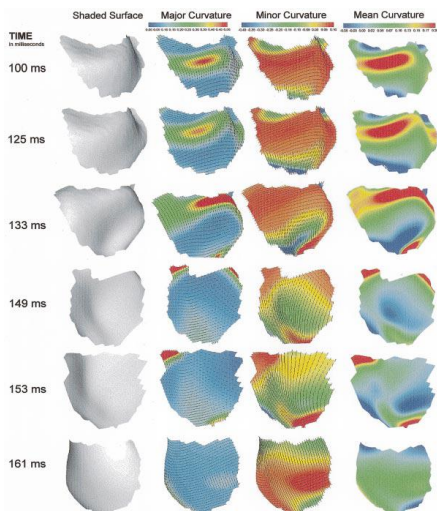


图 4 文献[27]中心脏动态监测

3. 结构光在实时交流中的应用

结构光技术可对人体姿势进行扫描,结合三维人体模型,可获得人体姿势较准确的估计[28]。在此基础上,可实现人体的手势和表情识别,所以可用于人机交互[29]。结合结构光技术的实时性,便可实现人物面部间的三维实时通话[30],如图 5 所示。

4. 结构光在地形测绘中的应用

结构光具有光学传输距离远的特性,能够测量人难以到达的危险区域或一些复杂区域,如水下物体三维形貌重建和地形

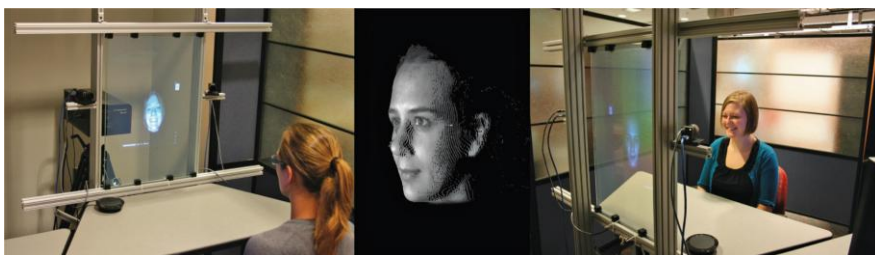


图 5 文献[30]中三维实时通话

检测。Bruno 等[31]利用结构光实现了水下近距离三维重建。而在地形检测方面,Zhan 等[32]通过设计一种多摄像头结构光高速测量系统,安装在轨道车上,在行驶过程中重建铁路隧道,有助于铁路隧道的定期检查,实现铁路隧道的快速三维扫描。

5. 结构光在全景测量中的应用

为拓宽结构光三维测量应用,Wang 等[33]设计了一种特殊的曲面镜,将面结构光反射至全景空间并捕捉,如图 6 所示,通过设计特别的投影图案与标定方法,实现了全景空间的三维扫描。结构光全景测量技术可进一步应用,如智能移动机器人的路径规划与自动驾驶设计等。

6. 结构光在复杂光学物体测量中的应用

基于结构光的三维测量技术应用在复杂光学物体时存在局部过曝的问题,Li 等[34]提出一种自适应编码图案测量方法,通过预先两次测量便可调整编码图案强度,进而实现复杂光学物体的三维重建。这方面的相关研究拓宽了结构光的应用范围,进一步可应用于高亮物体的三维重建,如

焊接过程中焊道,镜面物体的测量等。

六、总结与展望

随着工业现代化智能化的发展,主动视觉,即基于结构光的三维测量技术在工业上会有更广泛的应用。结构光三维测量关键技术在于投影图案的编码与测量系统的标定,两者分别决定解码方式与系统精度。本文总结了现有的编码图案设计的方式与标定方法,并在此基础上介绍了结构光的相关应用及拓展。但结构光测量过程仍存在一些有待解决的问题,如受环境光影响大、大景深及复杂光学物体难以测量、实时性与测量精度难以同时提高、点云密度高导致处理过程缓慢等。此外,随着智能终端的普及,结构光三维测量系统也在朝微型化,便携式方向发展,基于智能终端的三维扫描技术将开拓主动视觉新的市场。

(责任编辑:卢国梁)

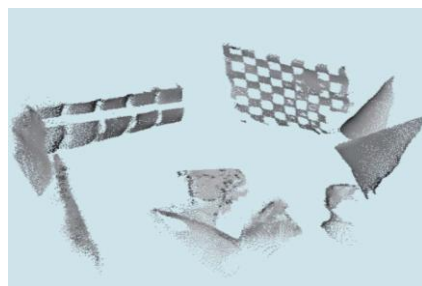
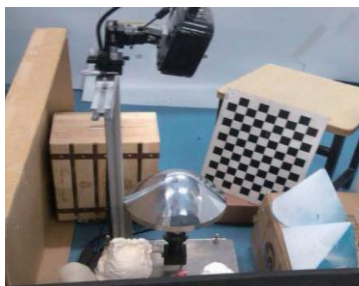


图 6 文献[33]全景测量模型(左为实物模型,右为重建结果)

参考文献：

- [1] Wong A K C, Niu P, He X. Fast acquisition of dense depth data by a new structured light scheme[J]. *Computer Vision & Image Understanding*, 2005, 98(3):398-422.
- [2] 喻擎苍. 基于符号 M 阵列二值结构光的三维检测方法的研究[D]. 浙江大学生物系统工程与食品科学学院 浙江大学, 2007.
- [3] 于晓洋, 单鹂娜, 曹沈楠, 吴海滨. 结构光时间编码技术进展, 哈尔滨理工大学学报, vol.15,no.1, pp98-102, 2010.
- [4] S. Inokuchi, K. Sato, F. Matsuda, Range imaging system for 3-D object recognition, in: *Proceedings of the International Conference on Pattern Recognition*, 1984, pp. 806–808.
- [5] Huntley J M, and Saldner H, Temporal phase-unwrapping algorithm for automated interferogram analysis *Appl. Opt.*, vol.32, pp 3047-3052, 1993
- [6] Zhang S. Recent progresses on real-time 3D shape measurement using digital fringe projection techniques. *Optics and Lasers in Engineering*, vol.48, pp149–157, 2010.
- [7] Ghiglia D C, Pritt M D. Two-Dimensional Phase Unwrapping: Theory, Algorithms, and Software[J]. 1998.
- [8] 尹丽萍, 于晓洋, 吴海滨. 格雷码与相移结合的结构光三维测量技术研究, 哈尔滨理工大学学报, vol.12, no.5, pp5-11, 2007.
- [9] Towers C E, Towers D P, Jones J D C. Absolute fringe order calculation using optimised multi-frequency selection in full-field profilometry[J]. *Optics & Lasers in Engineering*, 2005, 43(7):788-800.
- [10] Reich C, Ritter R, Thesing J. White light heterodyne principle for 3D-measurement[J]. *Proceedings of SPIE - The International Society for Optical Engineering*, 1997, 3100:236-244.
- [11] S. Zhang and P. S. Huang. Novel method for structured light system calibration. *Optical Engineering*, 2006, 45 (8): 083601.
- [12] W. Gao, L. Wang, and Z.-Y. Hu. Flexible calibration of a portable structured light system through surface plane. *Acta Automatica Sinica*, 2008, 34 (11): 1358-1362.
- [13] W. Gao, L. Wang, and Z. Hu. Flexible method for structured light system calibration. *Optical Engineering*, 2008, 47 (8): 083602.
- [14] G. Falcao, N. Hurtos, and J. Massich. Plane-based calibration of a projector-camera system. *VIBOT Master*, 2008, 9
- [15] D. Q. Huynh, R. A. Owens, and P. Hartmann. Calibrating a structured light stripe system: a novel approach. *International Journal of computer vision*, 1999, 33 (1): 73-86.
- [16] X. Zexiao, Z. Weitong, Z. Zhiwei, and J. Ming. A novel approach for the field calibration of line structured-light sensors. *Measurement*, 2010, 43 (2): 190-196.
- [17] G. Sansoni, M. Carocci, and R. Rodella. Calibration and performance evaluation of a 3-D imaging sensor based on the projection of structured light. *Instrumentation and Measurement, IEEE Transactions on*, 2000, 49 (3): 628-636.
- [18] Q. Hu, P. S. Huang, Q. Fu, and F.-P. Chiang. Calibration of a three-dimensional shape measurement system. *Optical Engineering*, 2003, 42 (2): 487.
- [19] X. Jia, Z. Zhang, F. Cao, and D. Zeng. Model and error analysis for coded structured light measurement system. *Optical Engineering*, 2010, 49 (12): 123603.
- [20] I. Léndry, C. Bràque, and V. Valle. Calibration of a structured-light projection system: Development to large dimension objects. *Optics and Lasers in Engineering*, 2011, 50 (3): 373-379.
- [21] H. Liu, W.-H. Su, K. Reichard, and S. Yin. Calibration-based phase-shifting projected fringe profilometry for accurate absolute 3D surface profile measurement. *Optics Communications*, 2003, 216 (1-3): 65-80.
- [22] P. J. Tavares and M. A. Vaz. Linear calibration procedure for the phase-to-height relationship in phase measurement profilometry. *Optics Communications*, 2007, 274 (2): 307-314.
- [23] J. Vargas, J. Antonio Quiroga, and M. José Terrón-López. Flexible calibration procedure for fringe projection profilometry. *Optical Engineering*, 2007, 46 (2): 023601.
- [24] Xu J, Xi N, Zhang C, Shi Quan, and Gregory J. Real time 3D shape inspection system of automotive part based on structured light pattern [J]. *Optics & Laser Technology*. 2011, 43(1), pp1-8.
- [25] K. Armbruster and M. Scheffler. Messendes 3D-endoskop. *Horizonte* 1998, 12, pp 15–16.
- [26] C. A. Bendall, G. Song, L. Tao, K. G. Harding, and T. Karpen. Fringe projection system and method for a probe suitable for phase-shift analysis. U.S. Patent 7,821,649, 2010.
- [27] A. K. Iyengar, H. Sugimoto, D. B. Smith, and M. S. Sacks. Dynamic In Vitro Quantification of Bioprosthetic Heart Valve Leaflet Motion Using Structured Light Projection. *Ann.*

- Biomed. Eng. 2001, 29, pp 963–973.
- [28] Thang N D, Uddin M Z, Lee Y K, et al. Recovering 3-D human body postures from depth maps and its application in human activity recognition[M]//Depth map and 3D imaging applications: Algorithms and technologies. IGI Global, 2012: 540-561.
- [29] C. J. Casey, D. L. Lau, and L. G. Hassebrook. Structured Light Illumination Methods for Continuous Motion Hand and Face–Computer Interaction. InTech, 2008.
- [30] N. Karpinskys. Portal-s: high-resolution, real-time 3D video telepresence. Ph.D. thesis, Iowa State University, 2013.
- [31] Bruno F, Bianco G, Muzzupappa M, et al. Experimentation of structured light and stereo vision for underwater 3D reconstruction[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2011, 66(4): 508-518.
- [32] Zhan D, Yu L, Xiao J, et al. Multi-camera and structured-light vision system (MSVS) for dynamic high-accuracy 3D measurements of railway tunnels[J]. Sensors, 2015, 15(4): 8664-8684.
- [33] Wang P, Xu J, Guan Y, et al. Three-dimensional panoramic sensor based on a variable-frequency fringe pattern.[J]. Applied Optics, 2017, 56(3):620.
- [34] Li D, Kofman J. Adaptive fringe-pattern projection for image saturation avoidance in 3D surface-shape measurement.[J]. Optics Express, 2014, 22(8):9887.



徐静，于 2003 年在哈尔滨工业大学获学士学位，于 2008 年在清华大学获博士学位，2008 年至 2010 年在美国密西根州立大学从事博士后研究。2010 年加入清华大学机械学院，现为副教授。主要研究方向包括主动视觉、智能机器人等。在国际期刊和

国际会议发表论文 70 余篇，授权中国发明专利 20 余项。

Email: jingxu@tsinghua.edu.cn

二维人脸的三维感知：基于三维辅助的人脸识别分析和应用

中国科学院自动化研究所 雷震 朱翔昱

人脸识别通常需要在非接触、非用户配合、甚至隐蔽条件下进行。在这些场合中，由于姿态、表情等干扰因素的影响，人脸表象的类内变化会大于类间变化，从而影响人脸识别系统准确度。

由于姿态变化是人脸在三维空间旋转造成的，而表情变化源于人脸的三维非刚性形变，因此从二维图像生成三维人脸模型并使用三维信息对干扰因素进行校正，可以极大抑制这些变化对人脸识别系统的影响。中国科学院自动化研究所生物识别课题组提出了高保真的人脸正面化算法：采用基于稀疏 SIFT 流的快速拟合方法，利用人脸三维可变模型从单张人脸图像中重构出三维人脸。通过调整模型系数，在保留人脸类别信息的前提下将人脸图像校正至中性表情，使用图像像素渲染三维模型并将模型旋转至正面，生成正面无表情的人脸图像。该校正过程可以有效消除姿态和表情的影响，提高人脸识别系统在实际场景中的准确率。

Before Normalization - Clustered by Pose



After Normalization - Clustered by Identity

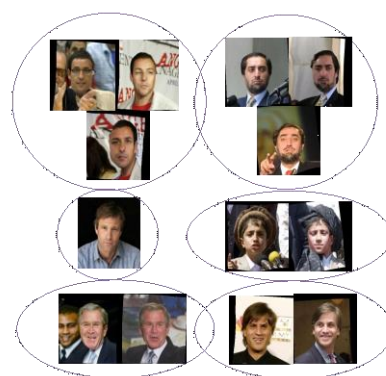


图 1 基于三维辅助的人脸识别

大姿态下的三维人脸模型拟合是一个挑战性问题。首先，自遮挡会导致部分关键点不可见，使得大姿态下无法依赖于二维关键点约束或者三维关键点处提取的图像特征。其次，当存在大姿态时人脸的表象变化更加复杂，甚至会会发生从正脸到侧脸的结构变化，这对拟合方法提出了更高的要求。最后，在大姿态下标定训练样本过于困难导致拟合算法的训练缺乏数据。为此，课题组提出了双路级联卷积神经网络，

利用两种拟合特征学习一个强大的回归器，并提出了一种新的损失函数以描述模型参数的优先级，使拟合过程倾向于更加重要的参数。课题组提出的结合级联回归和卷积神经网络进行三维模型拟合，首次解决了三维人脸拟合在 90 度的大姿态下的自遮挡问题。该研究的相关研究成果自 2015 年来已发表在多个模式识别国际会议，如 IEEE FG、CVPR 等。

(责任编辑：任传贤)



雷震，CCF 计算机视觉专委会委员、中国科学院自动化研究所副研究员。主要研究方向为智能视频分析和人脸识别。

Email: zlei@nlpr.ia.ac.cn

朱翔昱，中国科学院自动化研究所助理研究员。主要研究方向为模式识别和计算机视觉，包括三维人脸模型，人脸关键点定位和人脸识别。

Email: xiangyu.zhu@nlpr.ia.ac.cn

基于运动向量卷积神经网络的实时动作识别系统

同济大学 王瀚漓

基于视频的人类动作识别技术在智能监控、行为分析等实际问题中发挥着重要的作用。随着近年来大规模视频数据集以及全球范围内相关竞赛的出现和发展，动作识别领域受到了越来越广泛的关注。

领先的动作识别算法大多需要使用光流技术对运动信息进行描述，而光流计算速度较慢，即使使用 GPU 进行加速，只有约 16 帧/秒，无法达到实时处理的需求。对此，我们提出了一种基于运动向量卷积神

流一般精确细致的运动信息，因此直接使用运动向量代替光流会降低识别精度。为此，我们设计了知识传递 (Knowledge Transfer) 策略，使用基于光流的卷积神经网络以及动作类别作为监督信息，对基于

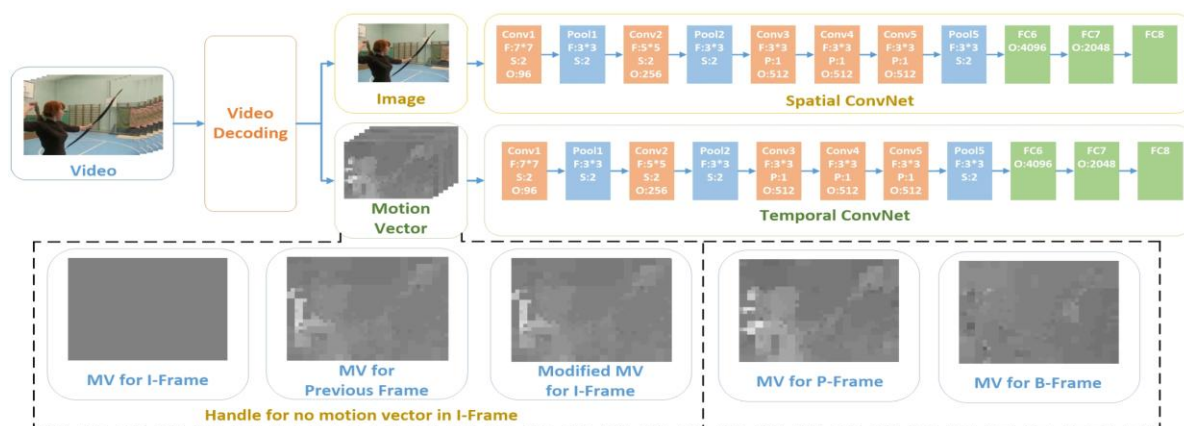


图 1 基于运动向量 CNN 的实时动作识别系统结构框图

动作识别方法可以大致分为两类，第一类基于传统手工设计特征以及“词袋”模型进行动作识别，该类别的代表算法有：改进的稠密轨迹算法、Fisher 向量算法等。第二类是基于卷积神经网络等深度学习技术的动作识别方法，代表算法有双流神经网络、C3D 算法、TDD 算法等。近年来得益于深度学习的发展，基于卷积神经网络的算法在各大数据集和竞赛中获得了较好的识别精度。

神经网络 (Motion Vector CNN) 的实时动作识别系统，如图 1 所示。在该系统中，我们使用运动向量对运动进行描述。运动向量是视频编码中为了减小相邻帧间冗余所提出的一种技术，其优点在于提取速度快 (运动向量在视频编码中已经计算完成，只需对码流进行解码操作就能获取)。基于运动向量，我们所提出的识别系统的计算速度达到了约 400 帧/秒。但是，由于运动向量是基于块进行描述的，无法提供如光

运动向量的神经网络进行监督学习。基于该策略，我们在大规模数据集如 UCF101 和 THUMOS 等中可以取得较快的处理速度和较高的识别精度。该研究的相关成果发表于国际期刊和会议 IEEE Trans. Image Processing、CVPR 2016。

(责任编辑：王金甲)



王瀚漓，CCF 计算机视觉专委会委员，同济大学教授，主要研究方向是视频编码、计算机视觉和机器学习等。

Email: hanliwang@tongji.edu.cn

局部图像特征描述

中国科学院自动化研究所 樊彬 吴福朝

局部图像特征描述指为局部图像信息构造特征描述符的过程或方法，目的是要提取局部图像区域内的本质信息用向量进行表示，使得在特征描述符空间中，三维世界中的同一点在不同图像中提取的特征描述符之间的距离相近，同时使得不同三维点在图像中提取的特征描述符之间的距离差距很大，如图 1 所示。前者对应特征描述符的鲁棒性，而后者对应特征描述符的特异性。作为计算机视觉理论的经典研究内容之一，局部图像特征描述方法广泛用于建立鲁棒的图像点对应关系，为解决三维计算机视觉中的相机标定、图像定位、几何形状与位置计算等问题提供可靠的输入源；此外，它还应用于高层视觉特征提取，在图像分类、识别等任务中发挥作用。

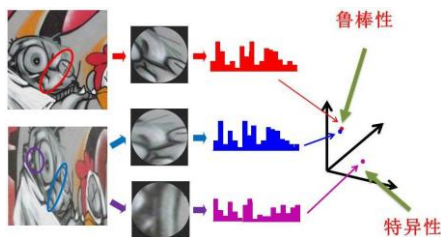


图 1 局部图像特征描述

在实际应用中，由于图像拍摄条件的差异，例如相机相对于物体的角度、位置、光照、气候条件等外部因素以及相机类型、

曝光、光圈设置等内部因素的变化，使得不同图像中的同一三维点及其局部图像邻域往往存在很大的表现以及形状差异，这对局部特征描述方法的鲁棒性提出了巨大的挑战；另一方面，由于特征描述符需要面对三维世界中近乎无穷且未知的点，如何为这些点保持特异性是局部图像特征描述方法需要处理的另一大难点。

为了设计既兼顾鲁棒性、又具有很强的特异性的特征描述符，广泛采用的方法是以 SIFT 为代表的知识驱动的方法。这类特征描述方法可以简单地归结为“从底层特征提取到特征汇聚再到归一化”的过程。研究人员通过对“底层特征提取”方法和“特征汇聚”方法两方面进行集中创新，在继 2004 年 SIFT 风靡之后的 10 多年里提出了许许多多 SIFT 的替代方法，例如 GLOH、SURF、DAISY、CS-LBP 等。这类知识驱动方法的最新成果是我们提出的基于灰度序的特征描述理论，在一个基于灰度序的特征汇聚框架下（图 2），提出了多种基于灰度序的特征描述符，包括 MROGH、MRRID、LIOP 和 OIOP。这些方法的相关成果相继发表在 CVPR、ICCV、IEEE TPAMI 等国际会议与期刊上。

近年来，随着深度学习技术的兴起，数据驱动的局部图像特征描述方法逐渐引

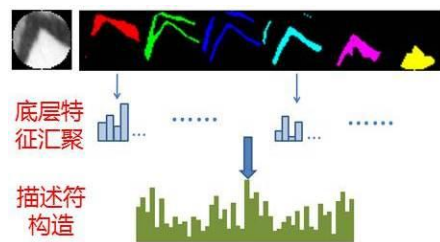


图 2 基于灰度序的特征汇聚方法

起了研究人员的关注，陆续提出了一些基于深度神经网络结构的特征描述符。根据是否耦合特征描述符的距离度量网络，这类方法可以分为带度量网络的方法（例如 MatchNet、DeepCompare）和直接在欧式空间使用的方法（例如 DeepDesc、TFeat、LIFT）。一般而言，带度量网络的方法通常匹配性能更好，但直接在欧式空间使用的方法应用起来更加灵活，在很多应用系统中可以直接替代传统的特征描述符进行使用，因而受关注度更高。在 2017 年的 CVPR 上，我们提出了 L2-Net，利用全卷积结构的深度神经网络，不依赖额外的距离度量网络，基于特征匹配的相对特性来进行网络训练，获得可直接在欧式空间中使用的特征描述符，不仅在性能上大幅度超越了同类方法，而且匹配性能也超越了目前最好的带度量网络的方法。

（责任编辑：申抒含）



樊彬，CCF 计算机视觉专委会委员、中国科学院自动化研究所副研究员，主要研究方向为图像特征提取与匹配。

Email: bfan@nlpr.ia.ac.cn

吴福朝，中国科学院自动化研究所研究员，主要研究方向三维计算机视觉、图像特征提取、视觉定位与测量。

Email: fcwu@nlpr.ia.ac.cn

CVPR 2017 会议总结

中国科学院自动化研究所 崔海楠

计算机视觉与模式识别领域的顶级会议 CVPR 2017 于 2017 年 7 月 21-26 日在美国夏威夷召开。CVPR(IEEE Conference on Computer Vision and Pattern Recognition) 是 IEEE 举办的一年一度的大型学术性会议，主要内容是探讨计算机视觉与模式识别领域的技术发展及最新应用。

随着人工智能领域受到了越来越多的关注，CVPR 投稿量近些年来连续攀升。根据大会公布的统计结果显示，今年 CVPR 总共收到有效投稿论文 2620 篇，录取论文 783 篇，录取率为 29.9%。按照研究方向划分，录取论文中所占比例从高到低排序前三名分别是：机器学习与计算机视觉（23.5%）、物体识别与场景理解（22.0%）以及三维计算机视觉（12.0%）。而在三维计算机视觉的许多工作中，研究者们越来越关注于如何将传统几何计算与深度学习结合起来进行三维视觉分析。

CVPR2017 注册参会人数达 4950 人，赞助商共 127 个，这些数字均创造了历史新高。除了学术界，工业界研究者们也带来了他们的研究成果和 demo 展示。从文章录取情况来看，阿里巴巴入选 4 篇，腾讯 6 篇，谷歌 21 篇，商汤科技 23 篇。最让大家惊喜的是历来保密的苹果公司，今年对外发布的第一篇 AI 领域的论文，也意味着今后我们将会随时看见苹果 AI 技术发展的身影。该论文提出了一种由模拟器 Simulator、精制器 Refiner 以及判别器 Discriminator 组成的 SimGAN 训练方法，斩获了 CVPR 2017 最佳论文。

CVPR2017 共推选两篇最佳论文、两篇最佳论文提名、一篇最佳学生论文。除了苹果推出的 SimGAN 之外，由康奈尔大学、清华大学与 Facebook AI 研究院 (FAIR) 的研究人员共同提出的 Densey Connected Convolutional Networks 也获得了最佳论文的殊荣。多伦多大学的“Annotating Object Instance with a Polygon-RNN”和 华盛顿大学的“YOLO 9000: Better, Faster, Stronger”获得了最佳论文提名奖，其中“YOLO”是“You Only Look Once”的缩写，是两位作者开发的实时目标检测系统。最佳学生论文奖颁发给以色列理工学院的一名博士生“Computational Imaging on the Electric Grid”，该论文横跨了光学、电气工程、计算机视觉、图像处理等多个领域，主要是一套借助肉眼不可见灯光设备的闪烁，来收集城市电网信息的系统。此外，Longuet-Higgins 奖授予 James Philbin 等人在 2007 年发表于 CVPR 的关于物体检索方面的工作，PAMI Young Researcher Award 授予了 Facebook 人工智能研究院研究员 Ross Girshick 以及法国国家信息与自动化研究所 (Inria Grenoble) 研究员 Julien Mairal。

从议程上来看，本届大会包含有研讨会 (workshop)、讲习班 (tutorial)、主会 (main conference)、演示 (demos)、专家报告 (plenary speakers)，以及半日的 AR/VR 展示。由于论文数目庞大，在三天半的主会中同时并行进行了三个议程。录取论文展示形式主要分为三种：大会长报

告 (oral)，大会短报告 (spotlight)，以及海报展示 (poster)。其中 oral 和 spotlight 均有视频录制并免费分享到视频网站上。此次 CVPR 大会三位特邀专家分别是来自 MIT 的专注于大脑神经科学与认知科学的 James J. DiCarlo 教授，来自微软的全球执行副总裁沈向洋讲述计算机视觉的商业化，以及来自斯坦福大学 Dan Jurafsky 教授讲述如何从语言中提取社会意义。

此外，CVPR2017 workshop 中备受关注的莫过于 ImageNet 挑战赛的结束。李飞飞与邓佳做主题演讲，对 ImageNet 挑战赛的 8 年历史进行了总结，并宣布之后的 ImageNet 挑战赛将转由 Kaggle 主办。

总之，无论是从注册参会的人数上，还是从论文工作的研究内容上，CVPR2017 都体现了当前人工智能已经获得了各个领域前所未有的关注。从学术上的深度交流，到视觉应用上的仔细探讨，来自各个国家的研究者们都在为计算机视觉与人工智能领域贡献着自己的智慧与力量。尤其是华人作者在国际舞台上已经成为了主要领军力量，2021 年 CVPR 主办权也将归属于华人，届时中科院谭铁牛院士将担任大会主席 (General Chair)，上海科技大学的虞晶怡教授将担任程序主席 (Program chair)。



崔海楠，中国科学院自动化研究所助理研究员，研究方向为三维计算机视觉理论与应用。

中山大学林惊教授访谈

2017年8月25日,CCF-CV专委会简报委员风采栏目编委采访了中山大学林惊教授。下面是采访实录。

林老师,您在视觉计算与智能感知方面取得了突出的成果,目前,您已获得国家“优秀”等多项学术荣誉,研究成果也在工业界有所应用,能否跟大家分享一下您的科研经历、获得成功的经验,以及您取得这些成就的动力呢?

这其实是个比较大的话题,我相信每个学者都有各自的答案。首先说说动力方面,我觉得自己的动力来自于兴趣和好奇心,同时也相信自己研究的人工智能、计算机视觉、机器学习这些技术能推动各个行业的进步,甚至改变世界,所以也有一些理想主义在里面。

第二点,在求学过程中受导师们影响很多(包括北京理工大学的王涌天教授和美国 UCLA 的朱松纯教授等),他们给我树立了追求科研和学术的榜样,培养了我的兴趣和理念。我也希望自己成为像他们一样优秀的学者。

第三点,我到中山大学任职后,接触到很多优秀的同学,他们有扎实的基础,强烈的进取心。作为老师,我有一种强烈的培养学生的意愿,希望能够把他们引导到科研的道路上,做出比我更好的成就。

第四点,我也希望自己做的学术工作能够进一步与产业应用相结合。随着大数据时代的到来,新一代人工智能技术的兴起,使得学术界和工业界的边界越来越模糊。涌现出不少以 AI 技术作为产品核心竞争力的企业,这类企业除了能提供大量的

数据和计算资源,还有真实的应用环境以及对科研工作的评价标准。在这样的背景下,我选择加入商汤科技,既是机遇,也是新的挑战。

科研经验方面,也是一个很大的话题,从我的角度来做一个简单探讨吧。首先,要学会坚持,不必过于计较眼前得失。这几年,人工智能相关领域的研究很热闹,每隔一段时间都有新的概念或者方法提出来。我觉得盲目追赶潮流和热点不值得鼓励。我时常建议我的研究生们:“与其做得快,不如做得扎实一些、深刻一些”。举例来说,我自己坚持做的一个方向就是统计语法模型和推理计算,对图像视频内容进行层次化语义解析,早期受限于特征表达能力和标注数据不足,进展不算快。到了2013年深度学习和神经网络成为研究热点,我也是更多地去思考这种数据驱动的深度学习方法与统计语法模型的内在关联,并设计了一些互相结合的方法。基于这样的研究思路,也取得了一些成果例如人物服饰精细化解析、场景层次解析、3D 人体姿势解析,都作为大会口头报告论文发表在 ICCV2015、CVPR2016、CVPR2017 这三次会议上。

我的第二个经验就是要深入细节,坚持在第一线做科研。现在大家都很忙,尤其是高校教授,有很多人才竞聘、考核、项目申请的压力,这种情况下很多教授容易浮在表面上,没有时间看经典文献和深入细节讨论分析。在这方面,我自己认为坚持的还算不错,包括与学生或者同事们讨论具体实现、实验的设计等。



第三个经验是保持 open mind,尤其是注意与相近领域甚至不同领域的学者的广泛交流与合作。

此外,我还建议学术界的教授、研究员们要发挥自己的优势,有条件的话,做一些前沿性的新问题。感觉正常的状态应该是学术界在引导工业界前进。

说起科研经历,也有一些小的波折,其实我在读博士之前根本没想过今后要从事学术科研,曾经一心想做产业应用。在本科大四阶段,我还曾参与过创业,与朋友们共同创办过小公司,做一些互联网和手机增值业务的项目。开始还有些收益,但由于缺乏核心技术突破和积累,维持的非常艰难。于是我就思考,是不是该更深入地去钻研一些技术。

当时正好有机会保送研究生,而且有幸遇到了我的导师王涌天教授,他给我提供了很好的科研平台,我从增强现实、三维视觉方向开始做,在老师指导下做了一个户外环境注册定位算法。这个过程极大地激发了我的兴趣,但是也很难继续深入,

主要是我当时缺乏计算机视觉的理论基础，遇到问题无法分析解决。那时我曾想放弃做科研，准备开始找工作或者实习机会，没想到又有了新的转折点。2005 年的一个会议上，我有幸认识了朱松纯老师，被他的渊博知识和学术情怀所吸引（其实见到他之前曾查阅过他的论文，但是完全没看懂）。当朱老师邀请我去他在湖北莲花山新创办的研究院进行访问学习时，我丝毫没有犹豫就答应了。在莲花山研究院的访问以及后来在美国 UCLA 的学习工作经历，对我来说是一个脱胎换骨的变化过程，尤其是系统性地学习和掌握了计算机视觉、模式识别、统计分析等理论知识和方法。从 2006 年开始，我在贝叶斯概率框架下研究基于马尔科夫链蒙特卡罗原理的聚类采样算法，应用于图匹配、多目标轨迹分析等任务。此外，还完成了一个基于概率与或图的物体识别框架，在当时的几个公开数据集上取得领先性能。在 UCLA 做博后期间，又找到了一些很有意思的方向，例如把图像视频解析的思想引入视频非真实感绘制，后来还获得了 NPAR 的最佳论文 Runner-up 奖；与同事合作了第一个从图像解析到文本描述生成的工作，并被 MIT 的 Technology Review 所报道。

2010 年初，综合考虑工作环境和家庭的原因，我选择了到中山大学工作。刚回国时，除了上课外，我还在学校里办了一个兴趣班，吸收大三的学生（我当时正给他们上课）参与，我跟大家讲一些前沿的研究，共同学习提高。在中山大学工作最令我高兴的事就是培养了一批优秀的同学。例如，我的第一个博士生梁小丹（目前在 CMU 的 Eric Xing 教授课题组做博后），在博士期间就取得了不少创新性很强的工作，并且在毕业后还持续有新的突破，非常令我骄傲。

后来几年经历了中山大学的院系调

整，我自己的研究方向也做了一些拓展，主要是与机器学习、大数据技术进行融合交叉，也遇到了一些志同道合的朋友同事，共同建设科研团队。

您当时是怎么决定想要加入商汤的？

因为商汤对我有很强的吸引力。2015 年 1 月我受邀去香港中文大学作学术报告，之后与汤晓鸥老师聊了一个下午，通过深入交流，我对汤老师的理念非常认同。汤老师希望能做出一个坚持原创技术的企业，通过技术突破来创造蓝海市场，驱动产业发展，并且与合作伙伴共建产业生态。这个想法在当时看来非常大胆且前卫，我当时其实也不完全信服。不过我还是同意暂时以技术顾问的身份加入商汤，提供技术支持，共享一些研究成果。从 2015 年开始，随着合作的深入，我也逐步对商汤的理念以及快速发展更加认同。此外，我也渴望将自己的科研能力和应用场景相结合，产生更大的一些突破，不仅是在学术上，更多在产业界和应用上。我也有很强的意愿去做一些具有挑战的尝试。所以 2016 年底开始，我就以执行研发总监的身份在商汤负责技术研发、产品化等一系列工作。商汤在这一年的发展非常快，我很庆幸自己是这个变化和发展的亲历者，而不是一个旁观者。也很高兴，我能够在这个过程中贡献了自己的力量。

您觉得加入商汤之后，对您原本作为教授的工作有什么影响或者冲突么？

肯定是对我原来的计划有一些调整和改变。我首先还是要完成学校的教学和科研任务，也会花很多精力来指导学生。庆幸的是，商汤有很好的科研平台，包括充分的资源，真实并具有挑战性的应用环境，对我的研究工作促进很大，并使我在一些问题和技術上的视野更加开阔。当然了，在个人发展方面肯定会做一些取舍，我希

望将来能够专注于应用性强、具有产业前景的研究工作。我也很愿意推动商汤与学术界同行学者的交流和互动。

林老师，您一向特别注重科研成果的转化，可否跟大家介绍一下您在科研成果应用和转化方面的经历和经验？您觉得高校老师应该怎样把应用成果转化做好？

我觉得这是一个非常好的问题，因为在我们国家快速发展的背景下，急需科技的创新、成果的转化。传统成果的转化以委托课题为主，企业从它的业务系统里把这些课题剥离出来，委托给高校老师去开发，开发完后把源代码或者系统移交给企业，这是一种很有效的方式。但现在更多的是教授、科研人员深入企业去做一些指导、交流互动。这种模式以前更多发生在美国。美国工业界和学术界的边界是很模糊的，很多著名教授在暑期或者学术休假期间都会去大公司进行访问，把自己的技能、知识、观点带给企业的技术团队。我认为这种方式在中国也将逐步成为主流。首先，现在成果的转化方式跟以前不太一样。单独提取出难题给高校教授来做，可能困难较大，很多技术往往需要科学家与工程团队密切协作才能完善。第二，企业在发展中，往往缺乏一些具有丰富研究经验的带头人，在应用环境下攻关技术难题。此外，就是我说过的数据、计算资源的问题。

林老师，加入商汤之后，您觉得对您的研究方向会不会有什么影响？您是否会考虑开展一些新的方向？

从我的角度来说，商汤带给我的东西很多，最重要的一点就是研究方向和研究思路的调整。商汤不是一个保守的公司，它做了很多大胆的尝试，比如说无人车、新一代的深度学习芯片，在这里能够做一些在学校里没有资源，无法开展的一些研

究方向，可以去尝试一些更前沿、更新的研究，这个对我来说是非常重要的。

加入商汤让我有更多机会去接触一些实际的更真实的问题，因为研究很多时候依赖于假设，而这些假设并非真实存在的。举个例子，在学术界，工作的标准是不能在一些公开测试集上取得好的效果。到了产品级别的技术就完全不一样了。比如说做人脸识别，在真实情况下会碰到各种各样的数据，来自监控摄像头的的数据、来自身份证照片的数据和社交网络上的数据，怎么把这些多源、多模态、具有噪声的数据用好，并持续提升模型性能，是非常具有挑战性的。我团队中的研究员告诉我，很多学术界里面做出来看起来很好的方法，直接应用起来效果并不好。因此这样的过程会促进一些新的研究工作。

林老师，您是中山大学人机物智能融合实验室的主任，能否向大家介绍一下这个实验室？您当时成立这个实验室的使命是什么？您是如何管理这个团队的呢？

我和成慧、王青两位老师，以及几位研究员一起，成立了这样的一个实验室，它并不是官方的实验室，更像是研究组。我们希望去研究人工智能下面的一些核心问题，比如说感知计算、智能学习、自主机器人、大数据分析等，并且将这些问题作为整体来考虑。实验室的王青副教授，很早在人机交互、软件工程方面有很重要的突破，2010年获得了顶级会议 SIG CHI 的最佳论文奖提名，成慧副教授在机器人控制和感知、路径规划方面有很多的成果。我们当时凑在一起，纯粹是兴趣驱动，看看能不能形成互补，做一些有意思的研究工作。实验室的名字反映了我们的愿景——将 AI 拓展到包含人类社会（人）、信息空间（机）、物理世界（物）的三元世界。

从管理上说，我们是一个既松散、又

紧密的团队。松散指实验室的教授、研究员都有自己研究的自由度，有各自的想法来进行碰撞。我们会有一些定期的交流来把握一些主要方向，围绕项目需求或者兴趣目标来设立一些共同课题。紧密指我们在培养学生、科研经费、科研设备的使用上是一体化的，大家共享资源，我更像一个协调人和主持者的角色。

请问林老师对自己的学生是怎么要求的？又是怎样来选拔学生的呢？能否跟大家分享一下您在管理学生方面的心得？

我选学生，首先要求学生要有很强的意愿和兴趣去从事科研工作。第二，他要有很端正的科研态度，能够踏实工作，能够坦诚面对自己的问题、直视自己的缺点。我也希望学生能够逐渐形成自己的想法，具有批判精神。中山大学校训“博学、审问、慎思、明辨、笃行”，其实从某种意义上说，就是我对学生的要求。

具体来说，我对学生的编程实践能力和数学基础要求比较高。我的很多学生都是从上课、本科毕业设计开始，再报读我的硕士、博士生。一个基本原则就是老师与学生要互相了解和认同，我的观点是：“一日为师、终生为友”。

您在美国加州大学洛杉矶分校学习工作过，在香港作过访问学者，请问您对国内高校、香港高校、美国高校的研究氛围、研究环境分别如何评价呢？

首先我觉得所谓的氛围还是来自于实验室的负责人，他直接决定了这个实验室的方向、氛围。在美国时，我导师朱松纯教授非常执着于学术，是一个有自己思想体系的学者，也有理想和情怀，我受他的启发非常多。他的实验室能坚持自己的方向，做自己的特色研究，不被所谓的热点、热潮所影响。这是非常值得国内高校的一

些实验室学习的。

香港的学术氛围也很好，我曾在香港理工大学张磊教授研究组访问，他们的一个特色就是能够把工作做得深入浅出，能让其他同行广泛理解接受，他们的很多工作都把源代码公布给大家参考。这一点也很值得大家学习。

国内也有很多优秀的实验室，能够达到国际水平，在某些研究方面比欧美国家做得还要好，这是非常值得高兴和骄傲的。不过还是要清醒认识到，相比于欧美发达国家，我们在技术和人才方面的积累还不够，整体上还有较大的差距。例如，目前各个顶级学术会议，来自国内研究组的论文很多，但是受邀做大会报告的学者大部分还是来自于欧美的研究组或者大公司。

此外，需要长时间投入、不容易马上应用的一些研究在中国难以得到支持。欧美由于资源更加充足，他们的科研体制能够保护学者去做一些不受关注，前瞻性很强的研究，我觉得这点应该向他们学习。

目前大家了解的都是工作状态的林老师，请问可否跟大家分享一下您的业余生活，以及您如何协调工作和家庭的呢？

我非常感谢家人的支持！我平时工作忙也时常出差，每当“稚子牵衣问，归来何太迟”，总觉得亏欠家人太多。我觉得每一个科研人员都应该感谢、珍惜他们的家人。

这两年开始我比较注意锻炼身体，每天都会坚持跑步，或者去健身房锻炼，有条件还会游泳。高强度的工作更需要保持身体机能的状态良好。

此外，业余时间我喜欢看书，尤其是一些人文历史方面的书籍，我出差时都会随身带阅读器 Kindle。我也很喜欢旅游，比较存粹的游览观光，不是出差开会。我

觉得丰富的生活是对科研的一个调剂，更有利于启发创造性的思考。



林惊，中山大学数据科学与计算机学院教授，博士生导师，商汤科技执行研发总监，国家优秀青年基金获得者，教育部超算工程软件工程研究中心副主任，

IET Fellow。先后在美国加州大学洛杉矶分校、香港中文大学等机构工作或访问研究。长期从事面向视觉大数据的语义分析与智能学习相关领域的研究，并且在商汤科技的一系列产品中应用落地。迄今在 CCF-A 类期刊与会议上发表论文 70 余篇，包括在 PAMI/IJCV 期刊发表论文 12 篇，在 CVPR/ICCV/NIPS/Multimedia 发表论文 40 余篇，2 篇论文成为 ESI 高被引论文。获得

NPAP 2010 最佳论文奖，2012 Google Faculty Award, 2017 年度 World' s FIRST 10K Best Paper Diamond Award by IEEE ICME, 2014 Hong Kong Scholars Award，率队获得 2016 英特尔杯全国并行应用挑战赛金奖。目前担任 IEEE Trans. Human-Machine Systems 等多个著名学术期刊的编委 (AE)。

(责任编辑：余焯 韩爱丽)

委员好消息

- 2017 年 7 月 17 日，ImageNet 公布了 ILSVRC2017 (ImageNet Large Scale Visual Recognition Challenge) 比赛结果。该赛设有目标检测、目标定位和视频目标检测三大类任务。CCF-CV 专委会常委、南京信息工程大学**刘青山**教授率领的 BDAT 团队夺得目标检测任务桂冠，并包揽了用提供的训练数据进行目标检测的前 3 名以及用额外训练数据进行目标检测的前 2 名。CCF-CV 专委会委员、公安部三所**梅林**研究员带领的 Trimps-Soushen 团队在目标定位任务中斩获佳绩，在仅使用标准数据的情况下，分别获得按定位错误率和分类错误率两个排行榜的第二名。自 2010 年以来，ILSVRC 每年由 ImageNet 举办一次，2017 年是最后一届，以后将由 Kaggle 主办。
- 2017 年 7 月 18 日，65 位 CCF 专业会员晋升为高级会员，其中包含 CCF-CV 专委会 7 位委员：中国科学院自动化研究所**何晖光**、北京航空航天大学**李甲**、北京邮电大学**马占宇**、西北工业大学**王琦**、同济大学**张林**、中国科学院信息工程研究所**张晓宇**、中国科学院自动化研究所**张兆翔**。
- 2017 年 7 月 19 日，央视 10 套《走近科学》栏目播出了 CCF-CV 委员、清华大

学**马惠敏**副教授历时十年的研究成果“基于图像认知的心理测评与干预技术”。

- 2017 年 7 月 21 日，CCF-CV 专委会委员、复旦大学**姜育刚**教授与腾讯 AI Lab 合作的团队获得 2017 iMaterialist Challenge 亚军。该赛事是 CVPR 2017 中 FGVC⁴ (4th Fine-grained Visual Categorization) 专题研讨会的一部分，由 Google 主办，主要任务是完成产品图像细粒度属性识别。
- 2017 年 7 月 26 日，CCF-CV 专委会常委、南京信息工程大学**刘青山**教授带领的团队在 CVPR 的 Faces "in the wild" 竞赛中获面部特征点定位算法第一名。
- 2017 年 7 月 28 日，CCF-CV 专委会主任、中科院自动化所**谭铁牛**院士当选 CVPR 2021 大会主席，CCF-CV 专委会副秘书长、上海科技大学**虞晶怡**教授当选程序主席。
- 2017 年 8 月 4 日，国家自然科学基金委公布了 2017 年度国家杰出青年科学基金建议资助项目申请人名单，建议资助的国家杰青共有 200 人，CCF-CV 专委会委员、大连理工大学**卢湖川**教授入选。
- 2017 年 8 月 5 日，CCF-CV 委员、上

海交通大学**林巍峒**副教授团队和中兴通讯的合作提出的一种基于超平面匹配的一体化多目标检测与跟踪算法在 MOT (Multiple Object Tracking) Challenge 测评中取得第一名的佳绩。MOT Challenge 是国际多目标跟踪领域最权威的测评平台，由阿德莱德大学、苏黎世联邦理工学院以及达姆施塔特工业大学联合创办，旨在评测多行人对象在复杂场景下进行同时检测及跟踪的算法性能，迄今为止参加测评的机构多达 60 多支。

- 2017 年 8 月 17 日，2017 年国家优秀青年科学基金获得者名单出炉，共 399 个项目获得资助立项，直接费用资助强度为 130 万元/项。CCF-CV 专委会委员、华南理工大学**余志文**教授和中科院沈阳自动化所**丛杨**研究员获得资助。

(责任编辑：刘海波 余志文)

OUC-VISION 水下图像数据集

山东财经大学/中国海洋大学 蹇木伟 国防科技大学 刘丽

"The OUC-Vision Large-Scale Underwater Image Database"是一种用于水下显著目标检测或显著性检测的大型水下图像数据集,其中包含 220 个独立物体的 4400 幅水下图像。对每个物体捕获了四种姿态变化(每个水下物体的正面、反面、左侧和右侧)和五个空间位置变化(水下物体位于左上角、右上角、中心、左下角和右下角),从而每个物体采集了 20 幅图像。同时,这个公开的 OUC-VISION 数据集还为相关工业领域和学术研究人员提供了不同变化来源的水下图像,特别是物体姿态、空间位置、光照、水浊度等。OUC-VISION 数据集不仅可以广泛用于评估现有显著目标检测和显著性检测算法的性能,而且还有利于未来水下视觉技术的进一步发展。

1 水下拍摄装置

为了模拟海底的环境,考虑了拍摄物体的姿态、空间位置、光照和水浊度等变化。我们建立了一个尺寸为 $1.5 \times 1.5 \times 1.5 \text{ m}^3$ 水池。池内充满水,并将沙子铺满在水池底部模拟海底环境。拍摄装置安装在 OUC-VISION 实验室的水下摄影室,包括图像采集系统和光照系统两部分,如图 1 所示。

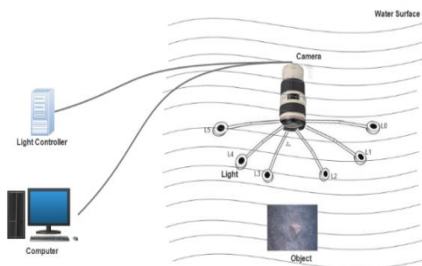


图 1 图像采集和光照系统。

1.1 图像采集系统

图像采集系统由 MV-GE500C 摄像机和电脑组成。相机是具有 2592×1944 像素的 MV-GE500C 水下镜头,用于捕获水下图像,计算机用以收集和存储图像。相机通过 USB 接口连接到计算机。为了控制相机并拍摄图像,我们开发了一个软件系统来实现这一功能。当我们点击捕获按钮时,软件可以控制相机采集物体的水下图像,并将图像存储在存储设备中。

1.2 光照系统

我们使用了具有 6 个 LED (发光二极管)灯的特殊照明系统来模拟水下复杂的光照环境,从而可以获得不同光照条件下的水下图像。六个 LED 灯安装在一个六边形上,距离相同,如图 1 所示。它们按顺时针方向依次标记为 L0-L5。每个灯都有独立的开关,所以我们可以分开控制它们。图 1 显示了 6 个 LED 灯排列成六边形的草图。

我们使用 L0、L2、L4 (每个灯与水平保持 60°) 三个对称灯光来模拟不同的光照条件。通过控制三个 LED 灯的开关,可以模拟四种不同的光照条件。在实验中,光照变化并没有严格的进行控制,但是可以很容易地区分这些照明变化。

2 水下数据集的设计

构建 OUC-VISION 水下数据集时考虑了四种变化,即姿态、空间位置、照明和水浑浊度。图像中的显著性物体被不同工作人员标注出来,作为 Ground-truths 用于实验评估和对比。

2.1 姿态变化

为了使捕捉到的水下图像具有不同

的姿态,我们考虑了每一个物体的四个侧面(正面、反面、左侧和右侧)。图 2 展示了一个水下物体所包含的四个侧面,我们把它们分别叫做正面视图图像、反面视图图像、左侧视图图像和右侧视图图像。



图 2 同一物体的四个不同姿态图像。

2.2 空间位置变化

为了展现同一个物体不同空间位置的变化,我们把每一个物体放置在图像的左上角、右上角、中心、左下角和右下角五个位置。因此,对于每一个物体的每一个侧面,我们可以得到五种不同空间位置的图像。图 3 展示了同一个物体同一个侧面的五种不同空间位置的示例图。



图 3 同一物体(同一侧面)的不同空间位置图像。

2.3 光照变化

光照系统中的 L0、L2 和 L4 三个 LED 灯用来模拟不同的水下光照条件。通过控制三个灯的开关来模拟了四种不同

水下的光照，四种光照分别对应着三个灯全亮、两个灯亮、一个灯亮、全熄灭。图 4 展示了四种灯光条件下的水下物体。



图 4 四种不同灯光下的水下物体图像。

2.4 水浊度变化

水浊度的变化是影响水下显著性检测的一个重要因素。在真实的水下应用中，水的浑浊度是经常变化的。为了评估水浊度对水下图像显著性检测方法的影响，我们考虑了三种不同程度的水浊度。图 5 是三种不同水浊度的示例图。我们是通过添加泥土的方法来改变水的浑浊度。从图 5 可以看出，水下图像的清晰度很大程度上依赖于水浊度，当水比较浑浊的时候，水下物体会变得模糊不清。



(a) Limpidity (b) Medium (c) Turbidity

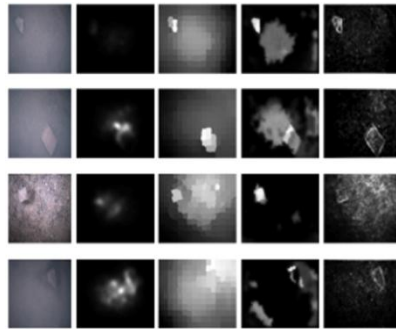
图 5 不同水浊度下的水下图像。

3 OUC-VISION 数据集评估

我们用该水下数据集来进行算法评估的主要目的如下：第一，评估目前主流的显著性检测算法对水下图像的检测难度；第二，为致力于水下目标检测和水下视觉的研究工作人员提供参考性的评估结果；第三，判断目前各个算法对水下图像进行显著性检测的优缺点。

如图 6 所示不同显著性算法的检测结果，从结果中我们可以看到，这些显著性检测算法对水下图像很难取得令人满

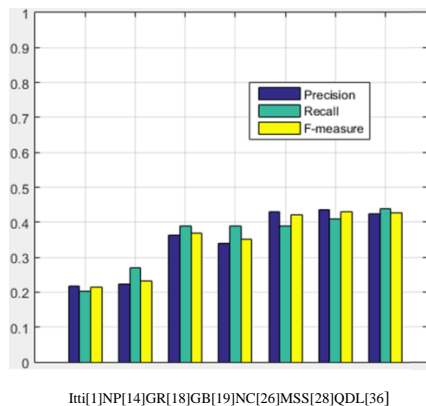
意的结果。



Input NC[26] GR[18] MSS[28] QDL[36]

图 6 不同的显著性算法检测结果对比。

为了定量的评估这几种代表性显著性检测方法，我们使用平均精度、召回率和 $F\beta$ 测量值来进行对比。图 7 显示了不同的方法在平均精度、召回率和 $F\beta$ 测量值的实验结果。根据这三个评估标准，我们可以看到，所有方法的值都低于 50%，这表明水下图像显著性检测是一个非常具有挑战性的问题。在实验中，除了 MSS 和 QDL，大多数方法的召回率都低于 40%。对于准确率，只有 NC、MSS 和 QDL 超过 40%。对于 $F\beta$ 测量值，只有 NC、MSS 和 QDL 达到 40% 以上，但都低于 45%。与普通图像显著性检测相比，水下显著性检测的性能较低。实验结果表明，水下图像的显著性检测难度较大。



Irfi[1]NPI[14]GR[18]GB[19]NC[26]MSS[28]QDL[36]

图 7 不同显著性检测方法在平均精度、召回率和 $F\beta$ 测量值的比较。

近年来，尽管研究人员提出了许多提高显著性检测性能的方法，并在简单背景图像上取得了出色的性能。然而，水下场景显著性检测是一个新的研究方向和

挑战性的任务。大多数主流的图像显著性检测模型几乎不适应复杂的水下环境。现有显著性检测算法性能不佳的主要原因是复杂的水下环境使捕获的水下图像受到以下影响所致：复杂的光照条件变化；低对比度；严重的水浊度；水下图像缺乏物体的颜色信息。

所有这些影响因素相作用，使得水下图像显著性检测成为一个具有挑战性的问题，未来需要研究更先进的显著性检测算法来解决此难题，推动水下视觉的进一步发展。

4 获取 OUC-VISION 数据集

OUC-VISION 数据集包含具有不同类型的 220 个水下物体的 4400 幅水下图像，人工标注的 Ground-truths 信息也包括在公开的数据集中。该数据集考虑了四种不同变化：姿态、空间位置、照明和水浊度。数据集还为相关工业领域和学术研究人员提供了不同变化来源的水下图像，不仅可以广泛用于评估现有显著目标检测和显著性检测算法的性能，而且还有利于未来水下视觉技术的发展。

对 OUC-VISION 水下图像数据集感兴趣的学术研究者或者相关工业领域人员可以联系 jianmuwei@ouc.edu.cn 获取该数据集。

[Ref1] Muwei Jian (蹇木伟), Qiang Qi, Junyu Dong (董军宇), Yilong Yin (尹义龙), Wenyin Zhang (张问银), Kin-Man Lam, "The OUC-Vision Large-Scale Underwater Image Database". 2017 IEEE International Conference on Multimedia and Expo (ICME), pp. 1297 - 1302, Hong Kong.

(责任编辑：樊鑫)

开源代码：手势识别开源代码介绍

西安电子科技大学 李宇楠 朱光明 苗启广

手势是利用人的肢体动作来说明其意图或态度的行为的交互形式。由于在视频监测控制、标志语言理解、虚拟现实和人机交互等领域有着巨大的应用前景,越来越多的研究人员开始研究手势及姿态识别算法,以实现将人类手势解释给机器的目标。以下对几种常用的开源手势/姿态识别源码进行介绍。

1 C3D 三维模型

C3D 是 D. Tran 等人于 2015 年 ICCV 提出的一种三维 CNN 模型,相比传统的 CNN 只能在空域进行卷积、池化操作,C3D 模型将时域信息也纳入操作的范围,可以同时时对域和空域进行卷积和池化,由此提取到具有时空特性的特征。该方法也因此更加适合视频的处理。

项目主页：<http://vlg.cs.dartmouth.edu/c3d/>
源码地址：

<https://github.com/facebook/C3D>

2 LRCN 长期循环卷积网络

长期循环卷积神经网络 (Long-term Recurrent Convolutional Network, LRCN) 是 J.Donahue 等人于 2015 年 CVPR 上提出的一种模型。为了处理具有时序特征的数据 (视频),他们首先利用传统的 CNN 提取每一帧的特征,然后利用神经网络结果的部件——LSTM 单元来描述这种时序特征。作者将该模型应用与动作识别、图像标注和视频标注,均取得了不错的效果。

项目主页：<http://jeffdonahue.com/lrcn/>

源码地址

<https://github.com/LisaAnne/lisa-caffe-publ>

ic/tree/lstm_video_deploy

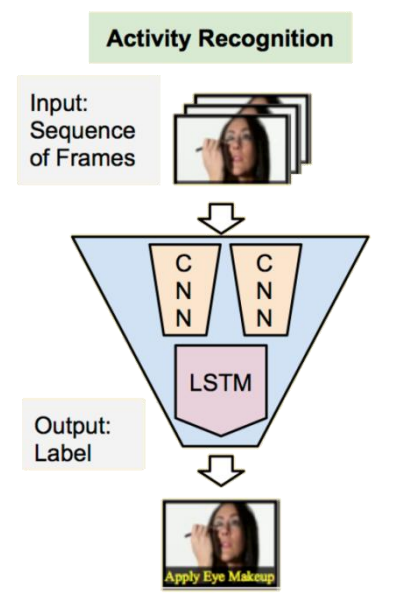


图 1 LRCN 在不同应用当中的网络结构

3 TSN 视频监督模型

Temporal segment network 是另外一种处理基于视频的手势或动作识别的方法。它将稀疏时间采样策略和基于视频的监督相结合,使用整个视频支持有效的学习。

首先,TSN 网络结构能够在一段长的视频序列中通过稀疏采样的方法提取短片断 (short snippets)。

这些样本在时间维度上服从均匀分布,然后作者分别用原始的 RGB 视频和基于其生成的光流视频来描述视频的时域和空域特征,并通过 consensus function 来融合两个网络的结果,作为最终的分类结果输出。

源码地址：

<https://github.com/yixiong/temporal-segment-networks>

4 Chalearn-LAP 大规模手势识别竞赛的相关方法

Chalearn LAP Large-scale Isolated Gesture Recognition Challenge 是一个专门针对大规模手势识别的竞赛。该竞赛已举办两届,要求参赛者在基于 RGB-D (深度) 视频的 Chalearn LAP IsoGD dataset 上进行手势识别。该数据集共有 249 类,包含 47933 个视频 (训练集 35878 个,验证集 5784 个,测试集 6271 个),每个视频中包含 1 个独立的手势。在该比赛中,亦涌现出不少具有很好识别效果的方法。

4.1 ASU

该方法首先针对输入视频中,存在的亮度差异和噪声,分别通过 Retinex 和中值滤波分别加以消除,随后,根据视频中不同阶段对动作识别重要程度的差异,以每一段的光流为指标,进行加权采样,获得统一的帧数。而特征的提取则是基于 ResC3D (一种结合 C3D 和残差网络的模型) 完成。最终的特征由每一种视频所提取的特征经典型相关性分析融合得到。

源码地址：

https://github.com/sevico/ChaLearn_ASU_code

4.2 SYSU_ISEE

该方法考虑了两种模型:动态和静态动作线索来进行手势识别。对于动态线索,从 RGB 视频、深度视频、光流序列和骨骼信息 (skeletons) 中学习具有辨别性的运动信息。对于静态线索,使用 rank

pool 方法,通过一个静态超级帧来表示每一个样本所有光流帧和深度帧。

这些输入(除了骨骼信息)分别被送到 VGG-16 中,而对于骨骼信息,则将它们送入深层次的 LSTM 网络,以学习被观察的骨骼间的依赖关系。

源码地址:

<https://github.com/EthanTaylor2/AMS>

4.3 Lostoy

该方法利用了 cross-modality 训练的思路,交替利用 RGB 和深度视频所训练得到的模型来对另一种数据进行参数微调(finetune),与此同时,利用 bounding box 标记处手的位置。采用这样的方法,可以很好地消除手势无关因素对识别的影响,也避免了过拟合的发生。

源码地址:

<https://github.com/yingweili/isogd>

4.4. AMRL

该方法采用基于异构的多模态手势识别方法。CNN 和 convLSTM 因其在时空维度上的表示能力而被用来构建异构网络。

首先将 RGB 和深度图片序列分成身体动态图片和手的动态图片分别输入到 ConvNet 网络中。然后通过 3D 卷积网络来学习手势的短期的时序特征,并基于提取到的短期时序特征来学习长期的时序特征。

为了学习到较细粒度的时序特征,该方法还用 fast-rcnn 来检测手的部分。这个基于异构网络的方法可以学习到不同级别的时序特征,并通过它们的相互补充提高识别率。

源码地址:

<https://github.com/wanghuogen/IsoGestureChallenge4ICCV>

4.5 XDETVP

该方法结合了三种不同类型的网络:首先利用 3D CNN 获取短期特征,然后利用 convLSTM 来学习它们的长期特征。通过 convLSTM 得到二维的时序图后,最终,基于二维特征图,用二维的卷积神经网络来识别手势。

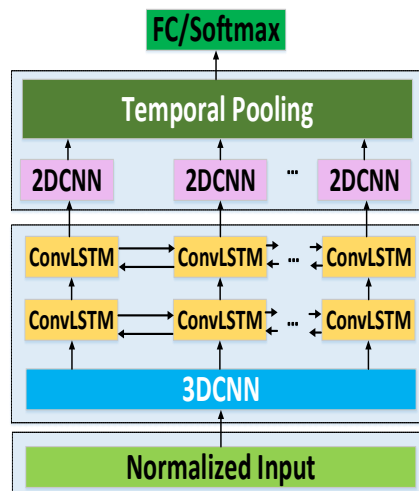


图 2 XDETVP 的网络结构

源码地址:

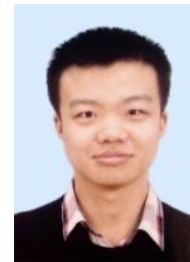
https://github.com/GuangmingZhu/Conv3D_BICLSTMg/

参考文献:

- [1] D. Tran, L. Bourdev, R. Fergus, L. et al., "Learning spatiotemporal features with 3d convolutional networks," ICCV, 2015, pp. 4489–4497.
- [2] J. Donahue, L. A. Hendricks, S. Guadarrama, et al., "Long-term recurrent convolutional networks for visual recognition and description," ICPR, 2015, pp.2625–2634.
- [3] L. Wang, Y. Xiong, Z. Wang, et al., "Temporal segment networks: towards good practices for deep action recognition," ECCV, Springer, 2016, pp. 20–36.
- [4] J. Wan, S. Z. Li, Y. Zhao, et al.,

"Chalearn looking at people rgb-d isolated and continuous datasets for gesture recognition," in CVPR Workshops, 2016, pp. 1–9.

作者简介:



李宇楠, 博士生,就读于西安电子科技大学,主要研究方向为计算机视觉与机器学习。于 2016 年和 2017 年带领团队 FLiXT 和 ASU 两次

参加 Chalearn LAP Large-scale Isolated Gesture Recognition Challenge, 在测试集上正确率分别达 56.90%和 67.71%, 蝉联冠军。



朱光明, 博士,西安电子科技大学讲师。2009 年和 2015 年在浙江大学获得学士、博士学位, 2015 年至今在西安电子科技大学软件学院任

职。主要研究方向包括人体行为识别、手势识别、场景理解、深度学习等。曾获首届 Chalearn LAP Large-scale Isolated Gesture Recognition Challenge 季军。



苗启广, 博士,西安电子科技大学教授,博导。主要研究方向为智能图像处理与机器学习、高性能计算。近年来,在 IEEE TNNLS、TIP、计算

机学报等国内外重要学术期刊、国际会议上发表 SCI/EI 收录论文 100 余篇。指导研究生参加 Chalearn LAP Large-scale Isolated Gesture Recognition Challenge 两度获得冠军。

(责任编辑:沈沛意)

招聘信息

一、清华大学招聘信息：

1、岗位名称：机器视觉博士后

岗位部门：机器人实验室

需求人数：1~2人

岗位要求：

(1) 近两年内博士学位获得者，计算机科学等相关专业；

(2) 熟悉计算机视觉、机器学习和深度学习等相关专业；

(3) 在计算机视觉等相关领域顶级学术会议和期刊上发表论文4篇以上；

(4) 具有熟练的计算机编程能力和英语写作交流能力；

(5) 具有较强的团队管理能力和合作精神。

2、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。简历投递截至时间：招满即止。

联系人：徐静

简历接收邮箱：

jingxu@tsinghua.edu.cn

二、军事医学科学院招聘信息：

1、岗位名称：机器学习算法开发

需求人数：1人

岗位要求：

(1) 相关专业博士学位获得者；

(2) 扎实的最优化或信号处理等知识；

(3) 有丰富的深度神经网络研发或应用经验，熟悉 VS、MATLAB 以及 Python 环境下递归神经网络或卷积神经网络；

(4) 熟悉 Caffe、Theano、TensorFlow 等至少一种深度学习框架；

(5) 独立完成过至少 1 项发明或实用新

型专利；

(6) 发表过国际期刊/会议论文。

2、联系方式

请应聘者以 E-mail 形式发送简历。

简历投递截至时间：招满即止。

联系人：徐老师

简历接收邮箱：

xugesen1989@126.com

三、深圳大学招聘信息：

1、岗位名称：图像及视频分析处理，计算机视觉，机器学习博士后

需求人数：1人

岗位要求：

(1) 已取得相关领域的博士学位或接近完成博士学位；

(2) 具有严谨的科研态度，扎实的训练基础和良好的团队合作精神；

(3) 发表过高水平论文，具有较强的英语写作和口语交流能力；

(4) 具有原创研究兴趣者优先；

(5) 全职从事博士后研究工作。

2、研究方向：

(1) 技术理论及算法：图像/视频内容分析处理，计算机视觉技术，机器学习包括深度学习最新技术的研究和开发；

(2) 应用及产业化：高动态范围数字相机系统，视觉技术在机器人和工业控制中的应用，视频监控和公共安全视频图像分析，智慧城市和智慧交通视频图像分析，数字显微镜及数字病理图像分析，生态及环境的遥感及无人机图像视频分，商业广告、娱乐和消费行业图像和视频分析，人机交互、虚拟现实和混合现实的应用等；

3、待遇：

(1) 聘请 2 年，薪酬为每年不低于 27

万元，主要包括：深圳大学综合年薪 15 万元（税前，含工资、社保、公积金、住房补贴等），深圳市政府博士后津贴 12 万元（税后），课题组科研绩效奖励；

(2) 提供一流的办公设施、场所及研究环境，提供最先进计算机视觉实验室设备和带 GPU 的计算机工作站，提供博士后宿舍，提供国际国内学术交流机会；

(3) 出站时条件优秀者可优先推荐留深圳大学工作，出站后留深圳工作可申请 30 万元科研启动经费，符合深圳市后备级人才和孔雀计划 C 类条件的可以获得 160 万元个人奖励（具体条件和金额以当年深圳市人力资源和社会保障局发布的标准为准）。

4、联系方式

请应聘者以 E-mail 形式发送简历，包括个人简历，反映本人学术水平的代表性成果以及其他相关支持材料。简历投递截至时间：招满即止。

联系人：邱老师

简历接收邮箱：

qiu@szu.edu.cn

四、华中科技大学招聘信息：

1、岗位名称：机器人系统控制

需求人数：1人

岗位要求：

(1) 全职博士后，年龄在 35 岁以下，博士毕业原则上不超过 3 年；

(2) 专业方向：自动化、计算机、物理等；

(3) 对控制理论与算法有深刻理解与应用，有嵌入式开发经验与机器人、无人机设计研究经验者优先考虑；

(4) 具有较强的理论基础、实践能力、

独立的科研能力良好的团队合作精神。

2、待遇：

(1) 应聘人员被应聘后，享受《华中科技大学博士后研究人员管理工作的暂行规定》的相关待遇，课题组另外根据实际工作情况给予每月补助和年终奖励，年薪为 20-30 万+每月 1000 的住房补助，具体面谈，特别优秀者上不封顶；

(2) 课题组提供良好的科研环境，创造良好的个人发展平台，并鼓励结合课题组的发展创立新的研究方向，申报国基金等

研究课题；

(3) 课题组提供与国外著名高校以及企业合作的机会，鼓励参与关键技术的研究，推动创新创业。

3、联系方式

请应聘者以 E-mail 形式发送简历，包括个人简历，主要成绩介绍，以及代表性论文或专利、软件等成果证明。简历投递截至时间：招满即止。

联系人：袁老师

简历接收邮箱：

ye.yuan@outlook.com

附：招聘单位简介：

军事医学科学院基础医学研究所、前沿交叉学科研究室多年来长期从事神经工程与脑科学等方向研究。研究室承担了基金委重点、杰青、优青，以及科技部“863”、“937”等 10 余项重大、重点课题。正在开展信号处理等系列研究工作，诚邀国内外相关领域优秀人才加盟。

(责任编辑：贾同)

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 Letters (PRL)、Computer Vision and Image Understanding

(CVIU)和 Neurocomputing。

3 会议简介

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，提升视觉领域的研究水平。

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，欢迎广大科技工作者参加。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：李策)

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
CCCV 2017	2017.10.12-14	中国 天津	-----	http://ccf-cccv.org/2017/
FG 2018	2018.5.16-18	中国 西安	2017.9.22	https://fg2018.cse.sc.edu/
ICDIP2018	2018.5.11-14	中国 上海	2018.3.10	http://www.icdip.org/

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
PRL	Robustness, Security and Regulation Aspects in Current Biometric Systems	2017.09.30
CVIU	Organizing and exploiting photo/video collections	2017.10.15
Neurocomputing	Advances in Data Representation and Learning for Pattern Analysis	2017.12.01

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”。
