

# CCF 计算机视觉专委简报

## Computer Vision Newsletter

2016 年第 3 期 (总第 4 期) 2016 年 9 月

主办: CCF 计算机视觉专委会

主编: 王亮

执行主编: 李实英

编委: 王瑞平 马占宇 毋立芳 虞晶怡

网址: <http://ccfcv.ccf.org.cn> (专委简报)

Email: [ccfcvn@gmail.com](mailto:ccfcvn@gmail.com)

(欢迎委员投稿、荐稿)

微信公众号: CCF 计算机视觉专委会



(公众号二维码)

### 目 录

#### 2 特别消息

CCF-CV 第一次常务委员会圆满召开

#### 3 科技前沿

山世光研究员: 深度学习与计算机视觉

#### 7 研究热点

CVPR2016: 深度视觉盛宴

#### 11 委员风采

访谈华中科技大学白翔教授

王亮研究员: 视觉大数据领域“慧眼识珠”

#### 15 专委动态

CCF-CV 走进高校系列活动

CCF-CV 走进企业系列活动

#### 22 资源平台

#### 23 征文通知

#### 26 委员分享

## CCF-CV 第一次常务委员会圆满召开

专委秘书处王瑞平撰稿



2016 年 7 月 27 日，中国计算机学会计算机视觉专委会（CCF-CV）第一次常务委员会于中科院自动化所圆满召开。除个别常务委员出差请假之外，专委会现任常务委员全部出席了本次会议（其中林宙辰、刘青山两名委员在国外通过网络视频全程参会），专委秘书处成员列席本次会议。

本次会议是 2015 年 9 月专委改选并选举出常务委员以来召开的第一次常务委员会。会议首先由专委主任谭铁牛院士做了简短发言，谭院士站在国家创新驱动发展战略的高度，结合计算机视觉领域的国际前沿趋势，围绕专委如何做大做强、如何持续发挥专业影响力、如何为国家科技发展建言献策等方面提出了多项内涵丰富、切实可行的发展建议。

接下来，由专委秘书长王亮研究员代表秘书处向各位委员报告了 2015 年 9 月至今专委建设的各项进展情况。在过去不到一年时间里，秘书处将专委各项工作有序稳步向前推进：本年度专委最重要的学术活动—9 月份即将举办的“2016 计算机视觉研究与应用创新论坛”正按计划紧锣密鼓筹备中；专委网站内容建设持续丰富多样化；走进高校系列活动正在形成品牌效应，足迹已遍及全国 14 座城市的 18 所高校；走进企业系列活动逐渐为企业界

与学术界提供更为多样的交流渠道；已连续出版 3 期的专委简报为委员们呈上了当前领域发展的最新资讯，构建起一个集委员风采、科技前沿、专委动态、资讯分享等功能于一体的形式新颖、内容丰富的专委交流平台。

之后，各位常务委员们针对专委发展的几项议题展开了热烈的讨论，议题涵盖专委承办会议的组织形式、现有几项活动的改进发展、专委组织结构规划、专委与 CCF 相关活动的互动等等。委员们踊跃发言，积极献计献策，最终讨论形成了多项富有成效的建议，为专委下一步发展明确了方向。

会议最后由谭主任作了总结发言。会议结束后，全体与会委员合影留念。



# 深度学习在计算机视觉领域成功的启示与开放问题讨论

山世光 中国科学院计算技术研究所

**摘要：**近几年来，深度学习技术在图像分类、物体检测、人脸识别、语义分割乃至“看图作文（Image Captioning）”等诸多计算机视觉（CV）任务上均取得了令人欢欣鼓舞的性能提升，甚至可以说是革命性的进步。这已是业界共知的事实，故本文不再赘述其进步，而旨在讨论与思考。首先谈谈深度学习成功带给我们的一些启示，然后对相关开放性问题进行讨论。所述仅为个人粗浅看法，不当之处请读者斧正。

## 1、深度学习成功的启示

深度学习的成功不仅仅带来了人工智能相关技术的快速进步，解决了许多过去被认为难以解决的难题，更重要的是它给我们带来了思想观念的变革，个人以为主要体现在以下几个方面。

### 1.1 优化方法的变革是开启深度学习复兴之门的钥匙

回顾自 2006 年（所谓深度学习元年）以来深度学习的十年大发展，我们必须首先注意到优化方法不断进步的重要作用。需要注意的是，深度学习并非全新的技术方法，而是上世纪 80 年代出现的多层神经网络的复兴。CV 领域现在炙手可热的深度模型（如深度卷积神经网络 DCNN）在上个世纪八十年代就已经基本成型，当时未能普及的原因很多，其中之一是长期缺少有效地优化多层网络的高效方法，特别是对多层神经网络进行初始化的有效方法。从这个意义上讲，Hinton 等人 2006 年的主要贡献是开创了无监督的、分层预训练多层神经网络的先河，从而使众多研究者重拾了对多层神经网络的信心。但实际上最近三年来 DCNN 的繁荣与无监督、分层预训练并无多大关系，而更多的与优化方法或者有利于优化的技巧有关，如 Mini-Batch SGD、ReLU 激活函数、Batch Normalization、ResNet 中的 shortcut 等，特别是其中处理梯度消失问题的手段，对 DCNN 网络不断加深、性能不断提升功不可没。

### 1.2 从经验驱动的人造特征范式到数据驱动的学习范式

在深度学习兴起之前，专家知识和经验驱动的 AI 范式主宰了语音处理、计算机视觉和模式识别等众多领域很多年，特别是在信息表示和特征设计方面，过去大量依赖人工，严重影响了智能处理技术的有效性和通用性。深度学习彻底颠覆了这种“人造特征”的范式，开启了数据驱动的“表示学习”范式。具体体现在两点：1) 所谓的经验和知识也在数据中，在数据量足够大时无需显式的经验或知识的嵌入，直接从数据中可以学到；2) 可以直接从原始信号开始学习表示，而无需人为转换到某个所谓“更好”的空间再进行学习。数据驱动的表示学习范式使得研发人员无需根据经验和知识针对不同问题设计不同的处理流程，从而大大提高了 AI 算法的通用性，也大大降低了解决新问题的难度。

### 1.3 从“分步、分治”到“端到端的学习”

分治或分步法，即将复杂问题分解为若干简单子问题或子步骤，曾经是解决复杂问题的常用思路。在 AI 领域，也是被广泛采用的方法论。比如，为了解决图像模式识别问题，过去经常将其分解为预处理、特征提取与选择、分类器设计等若干步骤。再如，为了解决非线性问题，可以采用分段线性方式来逼近全局的非线性。这样做的动机是很清晰的，即：子问题或子步骤变得简单、可控，更易解决。但从深度学习的视角来看，其劣势也是同样明显的：子问题最优未必意味着全局的最优，每个子步骤是

最优的也不意味着全过程来看是最优的。相反，深度学习更强调端到端的学习（end-to-end learning），即：不去人为地分步骤或者划分子问题，而是完全交给神经网络直接学习从原始输入到期望输出的映射。相比分治策略，端到端的学习具有协同增效（synergy）的优势，有更大的可能获得全局上更优的解。当然，如果我们一定要把分层看成是“子步骤或子问题”也是可以的，但这些分层各自完成什么功能并不是我们预先设定好的，而是通过基于数据的全局优化来自动学习的。

#### 1.4 深度学习具备超强的非线性建模能力

众多复杂问题本质上是高度非线性的，而深度学习实现了从输入到输出的非线性变换，这是深度学习在众多复杂问题上取得突破的重要原因之一。在深度学习之前，众多线性模型或近似线性模型曾大行其道。特别是从上个世纪 90 年代开始，以判别式降维为目的的线性子空间方法得到大家的重视，如主成分分析、Fisher 线性判别分析、独立成分分析等。后来，为了处理非线性问题，Kernel 技巧、流形学习等非线性处理方法相继得到重视。其中 Kernel 方法试图实现对原始输入的非线性变换，但却无法定义显式的非线性变换，只能借助有限种类的 kernel 函数，定义目标空间中的点积，间接实现非线性。而 2000 年之后曾一度广受重视的流形学习方法则试图通过对样本点之间测地距离或局部邻域关系的保持来学习非线性映射，遗憾的是这类方法难以真正实现对非训练样本的显式非线性变换。而深度学习则通过作用于大量神经元的非线性激活函数（如 Sigmoid 或 ReLU），获得了可以适配足够复杂的非线性变换的能力。

#### 1.5 大模型未必总是不好的

奥卡姆剃刀原理在诸多领域特别是机器学习领域广为人知，它告诫人们：“如无必要，勿增实体”，换句话说，求解问题的模型能简单最好不要复杂。这一原理在机器学习领域是提高模型推广能力的重要法则，也使得复杂的大模型往往不被看好。

而深度学习恰恰在这一点上是令人费解的，以 AlexNet 为例，其需要学习的参数（权重）多达 6000 万个，如此之巨的参数似乎表明这是一个非常复杂（如果不是过分复杂的话）的模型。当然，模型中需要学习的参数的多少并不直接等于模型的复杂度，但毋庸置疑的是，深度学习乍看起来是“复杂度”非常高的。那么，奥卡姆剃刀原理失效了吗？抑或看似复杂的深度学习模型的复杂度并不高？目前似乎尚无明确的理论支撑。最近的一些工作表明，很多已经训练好的、复杂的深度学习模型可以通过剪枝、压缩等手段进行约简，其性能并不降低甚至可以提高。

这里的关键也许在于“大数据”带来的“红利”。一种可能是：科研人员过去长期面对着“小数据”问题，因而过于偏爱简单模型了。而在数据量陡增的今天，适度复杂的模型变得更加适应科研人员面对的复杂问题，当训练数据大到与测试数据同分布，甚至测试数据基本“跑不出”训练数据所在范围的时候，在训练数据上的“过拟合”就变得不那么可怕了。

#### 1.6 脑神经科学启发的思路值得更多的重视

深度学习作为多层神经网络是受脑神经科学的启发而发展起来的。特别是卷积神经网络，其根源于 Fukushima 在 1980 年代提出的认知机模型，而该模型的提出动机就是模拟哺乳动物视觉神经系统的感受野逐渐变大、逐层提取由简及繁的特征，从而实现语义逐级抽象的视觉神经通路。在诺贝尔奖获得者 Hubel 和 Wiesel 的共同努力下，该通路从上个世纪 60 年代开始逐渐清晰，为 CNN 的诞生提供了良好的参照。但值得注意的是，生物视觉神经通路极其复杂，神经科学家对初级视觉皮层区中简单神经细胞的边缘提取功能是清晰的，对通路后面部分越来越复杂的神经细胞的功能也有一些探索，但对更高层级上的超复杂细胞的功能及其作用机制尚不清晰。这意味着 CNN 等深度模型是否真的能够模拟生物视觉通路还不得而知。但可以确定的

是，生物神经系统的连接极为复杂，不仅仅有自下而上的前馈和同层递归，更有大量的自上而下的反馈，以及来自其他神经子系统的外部连接，这些都是目前的深度模型尚未建模的。

但无论如何，脑神经科学的进步可以为深度模型的发展提供更多的可能性，是非常值得关注的。例如，最近越来越多的神经科学研究表明，曾一度被认为功能极为特异化的神经细胞其实具有良好的可塑性，例如，视觉皮层的大量神经细胞在失去视觉处理需求后不久，即被“重塑”转而处理触觉或其他模态的数据。神经系统的这种可塑性意味着不同的智能处理任务具有良好的通用性，为通用人工智能的发展提供了参照。

## 2、开放问题

大数据支撑的深度学习的复兴固然是 AI 领域的里程碑式进步，但并不意味着深度学习具有解决全部 AI 问题的潜力。下面对深度学习领域的开放性问题的讨论。

### 2.1 举一反三：大数据是否学习之必需？

大数据是深度学习成功的基石，大数据之于深度学习，恰如燃料之于火箭。越来越多的应用领域正持续积累着日趋丰富的应用数据，这对深度学习的进一步发展和应用至关重要。然而，过分倚重有标注的大数据也恰恰是深度学习的局限性之一。数据收集是有成本的，而且标注成本已经开始水涨船高，而且还有一些领域存在着难以收集数据的问题，例如在医疗诊断领域，一些较为罕见的疾病的相关数据收集是困难的。

更重要的，当我们把人的智能作为参照系的时候，自然就会问：人的智能是否是大数据学习的结果呢？其答案并不显然。从人类个体的角度来说，答案很可能是否定的：我们甚至可以只见过一个苹果（甚至只是一张苹果图片）就学会了识别苹果，而无需观察成百上千个不同的苹果。但是，这样批判深度学习看似有理有据，却未必是公平的：人类

作为一个种群，进化过程中已经见过了何止成百上千个苹果？

但无论如何，“小数据”如何驱动深度学习或其他机器学习方法是一个值得探索的新方向。在这个意义上，基于无监督数据的学习、相似领域的迁移学习、通用模型的领域适应、知识与经验的嵌入等方法是非常值得关注的。

### 2.2 无师自通：如何获取无监督学习能力？

获取有标注数据的时间和金钱成本很高，但大量无监督数据的获取成本却是微乎其微的。而目前深度学习对无监督数据的学习能力严重不足，以致大量无监督数据就像富含黄金的沙海，我们却没有高效淘金的利器。有趣的是，回顾深度学习的历史，我们应该记得 2006 年 Hinton 教授等人倡导的却恰恰是利用无监督学习来对深层神经网络进行预训练。但此后，特别是 DCNN 兴起之后，无监督的预训练似乎已经被很多研究者所抛弃（特别是在 CV 领域）。

直接从大量无监督数据中学习模型确实是非常困难的，即便是人这部“机器”，也有“狼孩”的例子警告我们“无师自通”似乎是不现实的。但“少量有导师数据+大量无导师数据”的模式也许是更值得大力研究的。

### 2.3 从参数学习到结构学习？

深度学习以“数据驱动”范式颠覆了“人造特征”范式，这是一个重大的进步。但与此同时，它自己又陷入了一个“人造结构”窠臼中。无论 Hinton 教授组最初设计的 AlexNet，还是后来的 VGG, GoogLeNet, ResNet 等等，都是富有经验的专家人工设计出来的。给定一个新问题，到底什么样的网络结构是最佳的（如多少卷积层）却不得而知，这在一定程度上阻碍了深度学习在更多智能任务上的普及和应用。因此，同时学习网络结构和网络参数是一个值得大力关注的研究方向。从计算的角度来看，全面的学习网络结构是极其复杂的。尽

管近期已经有一些这方面的尝试，如剪枝算法、网络约简等，可以在一定程度上调整网络结构。同时也出现了少量对网络的结构参数（如 DCNN 的 kernel bank 数量）进行学习的探索性工作，但尚处于起步阶段。

#### 2.4 如何在预测阶段进行反馈与网络调制？

我们知道，人类视觉系统在实现“看见”的过程中，视觉通路上神经细胞接受的输入并非仅仅来自低层或同层的神经元，还大量接受高层神经元的反馈信号，并受到来自其他神经子系统（如听觉）的信号的调制。与之相比，我们目前常用的深度神经网络（特别是 DCNN）在训练完成后的特征提取或预测阶段，低层神经元大多无法接受高层神经元的反馈信号，也没有机制接收其他信号（比如先验或其他模态信息）的调制。这意味着先验知识、上下文、猜测和想象（脑补）等“智能”能力难以在现有深度网络上得到应用和体现。如何突破这一困局，赋予深度网络感知阶段的自适应调制能力，是值得大力研究的。

#### 2.5 如何赋予机器演绎推理能力？

基于大数据的深度学习可以认为是一种归纳法，而从一般原理出发进行演绎是人类的另一重要能力，特别是在认知和决策过程中，我们大量依赖演绎推理。而演绎推理在很多时候似乎与数据无关。例如，即使不给任何样例，我们也可以依赖符号（语

言）描述，来学会识别之前从未见过的某种物体。这样的 zero-shot 学习问题看似超出了深度学习的触角范畴，但也未必不可企及。例如，近年来越来越多的基于深度学习的产生式模型正在努力实现从符号（概念）到图像的生成。



**山世光**，博士，中科院计算所研究员、博导，中科院智能信息处理重点实验室常务副主任、计算机视觉专委会常务委员。主要从事计算机视觉、模式识别、机器学习等相关研究

工作，特别是与人脸识别相关的研究工作。

他已在计算机学会推荐的 A 类国际刊物和会议上发表论文 50 余篇，论文被 Google Scholar 引用 9000 余次。曾应邀担任过 ICCV, ACCV, ICPR, FG, ICASSP 等多个国际学术会议的领域主席 (Area Chair)，现任 IEEE Trans. on Image Processing, Neurocomputing 和 Pattern Recognition Letters, Frontier of Computer Science 等国际刊物的编委 (AE)。研究成果获 2005 年度国家科技进步二等奖和 2015 年度国家自然科学基金二等奖，2012 年度基金委“优青”获得者，2015 年度 CCF 青年科学奖获得者。

## CVPR2016: 深度视觉盛宴

李实英摘编自微软亚洲研究院官网

(<http://www.msra.cn/zh-cn/research/academic-conferences/cvpr-2016.aspx>)

CVPR 全称为 IEEE Conference on Computer Vision and Pattern Recognition, 计算机视觉及模式识别大会。涵盖的问题包括但不限于: 物体识别与检测、图像高级语义理解、人脸、优化方法、Correspondences 求解、相机定位及三维地图构建 (SLAM)。CVPR 是计算机视觉最大的年度聚会, 今年收到投稿 2,145 份, 接受论文 643 篇, 接收率 29.9%。与会人数去年是 2800 多人, 今年有 3609 人, 接近 30% 的增长, 这与计算机视觉在学术界和工业界得到越来越多的关注不无关系。在这 643 篇接收文章中, 深度学习再次占据了大半壁江山。

### 1、如日中天的深度学习

张祥雨同学分享的关注点主要在深度学习和图像识别方面。深度学习作为眼下最热门的机器学习框架, 在计算机视觉方面的研究和应用可谓如日中天。据不完全统计, 本次会议 60% 以上的文章均与深度学习有关。在图像识别, 尤其是图像(视频)分类、物体检测、图像语义分割等领域, 深度学习已呈一统天下之势。即使在如 3D 视觉、底层图像处理等传统方法相对主流的领域也有不少学者给出了基于深度学习的解决方案。

### 图像识别

自从 Ross Girshick 等人提出基于深度卷积神经网络的物体检测方法 RCNN 以来, 深度学习凭借良好的精度逐渐成为物体检测的主流方法。之后的重要工作如 Fast RCNN 和 Faster RCNN 等更将物体检查的准确度和速度均提升了一个档次。本次会议有 5 篇物体检测方面的工作进行了口头展示。其中, Abhinav Shrivastava 等人在训练检测网络时通过添加困难样本, 以一种非常简洁的方式取得了可观的性能提升, 令人印象深刻; 而在另一项工作 YOLO

(You Only Look Once: Unified, Real-Time Object Detection) 中, 演讲者 Joseph Redmon 通过幽默的语言和生动的现场展示, 介绍了他们是如何设计算法框架以同时满足高精度和高实时性的要求。值得一提的是, Ross Girshick 本人在这两项工作中均有贡献, 足见其在物体检测领域“超级大牛”的地位。

作为图像识别问题的重要分支, 图像语义分割 (Semantic Segmentation) 同样吸引了一批顶尖科学家投入研究。张祥雨同学所在的微软亚洲研究院视觉计算组通过对图像分割问题中的多个目标函数进行级联式的联合训练, 在 MS COCO 等数据集上取得了明显的性能提升。该项工作是去年 ImageNet & MS COCO 比赛中微软冠军团队技术体系中的重要一环。此外, 视觉计算组的另一篇有关图像语义分割的文章 Scribble Sup: Scribble-Supervised Convolutional Networks for Semantic Segmentation 也在会议上做了口头展示。

### 视觉问答 (VQA)

视觉问答是近年来迅速兴起的研究课题。该问题的一般提法为: 给定一张图片, 以及一句使用自然语言描述的与图像相关的问题, 视觉问答系统需要对该问题做出回答。可见, 该课题同时涉及图像识别和自然语言理解, 更具挑战性。近年深度神经网络在图像识别和自然语言处理的迅猛发展, 为解决视觉问答问题提供了重要的思路。本次会议上视觉问答相关的文章几乎都采用了深度学习的方法。

值得一提的是, 今年年初的视觉问答挑战赛 (VQA Challenge) 的比赛结果也在本次会议的视觉问答专题会议上公布。Yuandong Tian (田渊栋) 等一批著名科学家被邀请在讨论会上发言。本次比赛设有 4 个挑战项目, 共有 20 多支队伍参加角逐。

张祥雨表示, 本次比赛所有项目的前三名均使用了他们发表的 152 层深度残差网络作为图像特征的提取模型, 再次表明了他们的方法在不同的应用场景下均具有较强的推广能力。

## 2、灵感碰撞的专题研讨会

今年的主讲嘉宾请来了 Amnon Shashua 教授。Ammon 是耶路撒冷大学的教授以及 Mobileye 公司的创始人。Ammon 在会上分享了他对自动驾驶、深度学习、汽车市场的一些见解。

Keynote 刚开始, Ammon 说, It's not going to be the kind of talk that I'll tell you how things are done. I'll tell you that something more interesting is what are the things that need to be done. I always tell my students that 80% of the work is knowing what to solve. The remaining 20% percent, if you don't do it, somebody else will do it. Knowing what to solve is really the big thing.

“找到需要去做的东西去做才是最重要的”, 此话让人深有感触。计算机视觉经过磕磕碰碰几十年的发展, 似乎终于找到了靠谱的落地方式: 无人车。主题演讲中 Ammon 将自动驾驶中的技术分为三个支柱, 分别是传感技术 (Sensing), 地图绘制 (mapping) 和驾驶策略 (Driving Policy)。Ammon 用一些例子说明了三个问题的发展状况, 讨论了为何这三个问题必须同时发展并作为整体来考虑, 以免一个问题对另一个问题产生分歧。现在全世界一窝蜂想搞无人车的大环境下, Ammon Shashua 教授的演讲为想了解 and 想进入这个领域的同学提供了一个很好的全局观, 让大家理清了需要解决的问题, 以及对这些问题的感性认识。

CVPR 2015 曾有个专题研讨会: 计算机视觉中的女性研究者 (WiCV2015: Women in Computer Vision)。今年 CVPR 2016 的组委会几乎全是女性, 通过邀请在这个行业中的女性榜样研究员提高女性研究者的影响力和被关注程度。同时每天大会的

简报还特意挑选出一两位女性研究者来分享她们关于研究兴趣、职场规划的宝贵经验, 也给予年轻的女性研究者更多的机会去展示她们的成果。

张婷同学说, 在经过一周参会的思想碰撞后, 感觉计算机视觉在生活中几乎无处不在。引用第一天简报的话语: 计算机视觉将会应用到汽车、家庭、搜索引擎、医院等各种各样的领域。它可以用来帮助有残疾的人士、分析社交关系、拯救垂危的生命, 可以用来增加粮食的产量等。而计算机视觉要实现这样的目标, 需要来自各个不同领域具有各种专业背景的研究员。

## 3、CVPR 论文是如何写成的

今年的最佳学生论文毫不意外也是关于深度学习的文章, 来自康奈尔团队的 Structural-RNN: Deep Learning on Spatio-Temporal Graphs, Ashesh Jain, Amir R, Zamir, Silvio Savarese, Ashutosh Saxena。最佳论文第二名 (Honorable Mention) 则颁给了慕尼黑科大(TUM)团队关于优化理论的文章: Sublabel-Accurate Relaxation of Nonconvex Energies, Thomas Moldenhoff, Emanuel Laude, Michael Moeller, Jan Lellmann, Daniel Cremers。可见组委会在深度学习热潮中依然保持着对传统能量优化问题的重视。其中一个重要原因可能是, 尽管深度神经网络在中高层视觉中已所向披靡, 但在低层视觉的一些问题中仍未超越传统方法, 例如去噪、光流、以及立体视觉匹配。

论文中 TUM Cremers 团队提出的优化方法正是针对低层视觉问题中的经典 formulation: 逐像素的非凸数据项+TV 正则化约束。该方法的中心思想是对每一个像素, 用一个高维向 (Lifted Representation) 来表示原问题中要优化的标量 (或低维) 自变量, 将原问题映射到一个更高维空间中的一个等价问题, 从而可以对非凸能量项进行更好的凸近似, 获得更高质量的解。近似后得到的凸能量可以并行优化, 且易于实现, 可以高效地找到近似问题的全局最优解。论文重点讨论了该方法在立

体视觉匹配中的应用，效果大大超过了传统的 TV+非凸数据项的解法。

张弛同学解释，立体视觉匹配是他的研究重点之一，该问题的任务是从双目相机所获取的左眼图像和右眼图像中恢复出逐像素点的深度。由于双目相机中的平行极线约束，该问题又转化为左右眼图之间的逐像素匹配问题。

受限于训练数据的数量，同时也因为可能未找到该问题下的最佳网络结构，深度学习中的端到端训练方法在立体视觉匹配上的表现依然未超过传统方法。这里的端到端训练方法指的是：直接对输入图像进行一系列卷积以及一系列反卷积来回输出输出图像。由于获得双目图像的 ground truth 深度数据本身就是一个难题，目前立体视觉匹配两个主要 benchmark Middlebury 和 KITTI 上也分别只有数十对和数百对训练数据，对训练端到端的深度估计网络还远远不够。

为了解决这个难题，TUM Cremers 团队在另一篇论文中讨论了如何生成足够且靠谱的训练数据用以训练深度估计网络：A Large Dataset to Train Convolutional Networks for Disparity, Optical Flow, and Scene Flow Estimation, Nikolaus Mayer, Eddy Ilg, Philip Häusser, Philipp Fischer, Daniel Cremers, Alexey Dosovitskiy, Thomas Brox。该论文使用开源的 3D Creation Suite Blender 渲染出一系列带有复杂运动模式的物体的双目图像，由于渲染过程是从 3D 到 2D 且物体及场景的三维模型已知，生成 ground truth 的深度图变得非常容易。生成的训练数据集中包含 35,000 对训练样本，训练得到的深度估计网络在 KITTI 上的排名也比较靠前，但离第一名传统方法 Displet 仍有相当一段距离。值得一提的是，虽然该网络使用人工合成的数据集进行训练，它在真实数据集上似乎具有良好的泛化能力。随着训练数据的不断完善，深度网络在几个重要的低层视觉问题上是否会超越以及如何超越传统方法，又或者能给传统解法带来什么样的启发，我们拭目以待。

张弛同学还分享了他们组关于物体扫描的论文 Joint Multiview Segmentation and Localization of RGB-D Images Using Depth-Induced Silhouette Consistency 的准备历程。CVPR 的主会议程非常紧凑，被接收的 600 多篇论文被安排在短短四天里面完成展示。他们的工作在 29 日下午进行展示。该工作由中山大学智能信息处理实验室以及微软亚洲研究院多媒体搜索与挖掘组共同合作完成。论文的动机为利用物体的剪影信息来提高 RGB-D SLAM 中相机位置估计的精度。由于物体剪影难以获得，一般需要繁琐的用户输入，这大大阻碍了使用剪影信息的系统的实用性。针对这个问题，他们提出同时进行多视角物体分割以及相机定位的解决方案，使得物体剪影可以在优化过程中即时产生，并用于改进相机姿态估计。论文的实验中展示了所提出方法产生的更高质量的重建结果。

#### 4、从 ImageNet 比赛冠军到 CVPR 最佳论文

微软亚洲研究院在深度舞台中继续扮演重要角色，2015 年 12 月在 ImageNet 图像识别挑战赛上开创 152 层神话的 Deep Residual Networks 夺得了最佳论文奖。其实这一结果计算机视觉组的研究员们并没有感到十分意外，去年的 ImageNet& Microsoft COCO 比赛他们正是凭借该方法获得了全部 5 个主要项目的冠军。在方法公布至今半年多的时间里，该方法在学术界和工业界受到了较大的关注，之后的许多成果从不同的侧面印证了该模型的有效性。这次获奖再次印证了学术界对这项工作的肯定。随着神经网络层数加深，神经网络会越来越难训练。获奖论文提出了一种针对深度神经网络的残差训练框架，解决了这一当前具有迫切需求的难题。

张祥雨同学作为这篇文章的作者之一，他的研究方向主要是深度网络模型构建及其在图像分类问题上的应用。他认为，由于该方向在深度视觉领域的基础性，优秀的工作往往会得到学术界的高度重视，当然也意味着相对较大的研究难度。

张祥雨同学分享道，作为该项工作的参与者，从中学到了很多。该项研究始于去年 5 月左右。一次偶然的机会，他发现某些含有跨层结构的网络能够很大程度上提升网络的深度，同时模型的分类精度也有大幅提高。沿着这个思路走下去，他们得到了一系列结构复杂但是性能良好的模型。然而此时指导老师指出这条思路是不正确的，复杂的模型包含了太多的参数，很难在多个任务上公平地衡量模型的优劣；更重要的是，由于模型的复杂性，无法得知究竟是哪一块设计对最终结果起到了决定性的作用，从而难以对以后的工作产生有益的指导。因此老师建议对当前的模型结构进行简化，只保留对最终精度影响最大的结构成分。

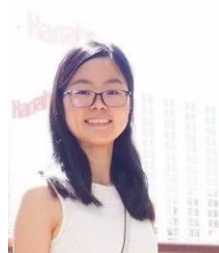
起初他对此十分不解，因为模型结构的简化通常是以损失精度为代价的，尤其是当时随着比赛日期的临近，“追求极致精度”似乎是更为合理的做法。然而事实证明指导老师是正确的，当他们对网络结构进行充分的简化和变形之后，发现模型中的残差结构对网络性能起到了决定性的影响，而跨层的设计则避免了传统深层网络中常见的梯度消失或膨胀的问题。至于简化模型所损失的精度，可以简单地通过增加网络深度来补偿。后来的实验表明，这种极简的网络设计在保证分类精度的同时，极大地方便了模型在物体定位、物体检测、图像分割等问题中的运用。这也使得去年 ImageNet& COCO 比赛中他们在不同的项目上均取得了较大优势。



张弛，是一名微软亚洲研究院实习生，本科毕业于中山大学，目前是中山大学和微软亚洲研究院联合培养博士生。曾在 CVPR/ICCV/ECCV/

TVCG 等会议和期刊上发表论文，研究方向为立体视觉匹配和即时定位与地图构建 (SLAM)。

张弛论文: Joint Multiview Segmentation and Localization of RGB-D Images Using Depth-Induced Silhouette Consistency. Chi Zhang, Zhiwei Li, Rui Cai, Hongyang Chao, Yong Rui.



张婷，本科毕业于中国科大少年班学院学习计算数学专业，曾参与微软亚洲研究院创新人才学院 (Pre-PHD) 培养项目，目前是中国科学技术大学和微软亚洲研究院联合培养博士生，在微软亚洲研究院实习。曾在 ICML 2014/CVPR 2015/CVPR 2016 发表论文，主要研究方向为近似最近邻搜索。

张婷论文: ① Collaborative Quantization for Cross-Modal Similarity Search, Ting Zhang, Jingdong Wang. ② Supervised Quantization for Similarity Search, Xiaojuan Wang, Ting Zhang, Guo-Jun Qi, Jinhui Tang, Jingdong Wang.

张祥雨论文: Deep Residual Learning for Image Recognition, Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun.



张祥雨，本科毕业于西安交通大学。三年前我参加了西安交通大学与微软亚洲研究院的联合培养博士生项目，一直实习至今。研究方向为深度学习与图像识别相关问题。

张祥雨论文: Deep Residual Learning for Image Recognition, Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun.

## 访谈华中科技大学白翔教授

专委秘书处李实英采编

白老师好，感谢您百忙之中接受我们的访谈、给我们机会分享您研究过程中的经验！

### 您在视频图像的文字检测、识别等方面取得丰硕的成果，当初您是如何开始这方面研究的

视频图像文字检测与识别是一个非常重要且极具挑战性的一个课题。选择此方面的研究，主要出于两个方面的考虑：一、自身的研究背景。我的博士阶段主要研究领域为形状表示与匹配。工作之后，希望能对自己的研究方向进行一定的扩展。识别文字的主要特征就是形状，尤其是边缘或区域特征跟文字识别有着密切的关联，让我对此产生了浓厚的兴趣；二、国内此前在视频图像文字检测与识别的研究并不多见（除了自动化所刘成林教授团队及少数几家单位在此取得了一些显著成果），虽然拥有重要的理论研究价值和实际应用需求，当时并没有引起足够的关注。鉴于此，我选择了此研究方向，希望能够做点事。

幸运的是，在此方面的研究能够得到国家自然科学基金优秀青年基金的支持。实话说，我当时心里也没底，毕竟和之前研究基础有一定跨度，而场景文字检测与识别的研究对系统实现有着相当高的要求。研究期间遇到了各种各样此前没经历过的困难，自己没有太多别的想法，和学生一起闷头慢慢做。

在场景文字检测与识别领域的研究过程中，始终遵循的精神是“坚持”，即便已经研究了 4、5 年，仍然有很多问题我们没解决好，因此还会“坚持”下去。

### 这些年您在顶级国际期刊和会议上发表了很多成果，是否可以分享您的研究成果产出经验

这是一个永恒的话题，很多专家在此方面比我

更有发言权。关于模式识别领域的顶级期刊 PAMI 和 IJCV，更加强调工作的完整性和坚实性，一定要准备充分并且经得起考验，其工作量往往要多于普通期刊，甚至是普通期刊的好几倍；而顶级会议 CVPR、ICCV、ECCV 的文章也要十分细致。因为会议论文就是一锤子买卖，reviewer 可以通过一处公式错误或一篇重要文献的缺失直接拒稿。

如果一定要说经验，我觉得就是大胆多投，被拒稿也不要沮丧，应该虚心从 reviewer 指出的问题里去学习去提高。事实上，文章被拒稿是最能帮助人提高的事，reviewer 的批评意见比起我平时指出学生做研究的问题起到的效果要好得多。当 reviewer 找不到强烈理由拒稿的时候，文章自然就中了。当然，对于顶级期刊和会议偶尔也需要一定的运气，以一颗平常心去对待，坚持不懈必有所获。总体来说顶级期刊和顶级会议都有自己的风格和要求，经验的积累需要一个过程。

### 您今后将主要关注哪些方面的研究动向

当然还是场景文字检测与识别，这里面还有很多问题尚未解决，尤其是多语种多方向文字等一些复杂情况。除此之外，我应该会关注 Image Captioning 和无监督或弱监督学习模型。Image Captioning 是从图像到文本结构化描述的主要手段之一，和文字识别在算法模型上有着密切的联系。而关注无监督或弱监督学习是因为场景文字的标注是一项十分繁琐的任务，我希望能使用少量的数据标注来实现海量视频图像文字的语义提取，彻底解决场景文字的检测与识别这一难题。

### 是否可以分享您指导学生过程的一些体会

坦率地说，如何指导学生是我仍然一直在琢磨的问题。大体上，我会遵循三个原则：因材施教、诚恳交流、全力支持。

首先，我会根据每个学生的知识背景和他/她自己的研究兴趣去设计合适的研究题目；其次，在课题研究期间，我会尽可能地找时间与每个学生交流，指出其可能存在的问题，并给出我自己的建议。需要指出的是，谈话的内容并不仅限于工作，这样有助于我更好地了解学生，更便于师生之间的沟通和工作的开展。在我刚开始工作时，曾经遇到一个四年级博士生。该学生十分刻苦努力，研究上却进展不明显。经过一定了解，我发现该学生性格十分内向，不擅长和老师同学交流，研究上思维也不够活跃。针对该学生的情况，我跟他提出了一个要求：每周至少有一天工作日不准来实验室，而且尽量在人多的地方（如商场、公园等）活动。这样坚持了 2-3 个月后，该学生心态上有了明显好转，与同学老师之间的交流也变得频繁起来。当然，他的研究工作同时也取得了重要进展，顺利获得了博士学位。

关于“全力支持”，就是尽可能给予学生研究的环境、条件及精神上的鼓励。举个例子，我的一位学生曾跟我提出了要购买一台 GPU 服务器的建议，当天下午我们研究组的第一台 GPU 服务器就送货上门了。事实证明，确实对我们之后的研究工作起到了极大的帮助。对我来说，指导学生的方法很难用言语去准确总结，但保持着一定的“灵活性”是必要的。

以下附上我最喜欢对初入研究的学生说的几句话：“研究一个问题就好比追女孩子，你首先要成为这个世界上最了解她的人，离追到手就成功了一半。”“Research 的含义就是反复搜寻，10 次实验 9 次失败是完全正常的。当你实在没招的时候，

就学蝙蝠侠放火烧了森林，那小丑不就无处可藏了吗？”“没有取得好的实验结果其实并不可怕，可怕的是你不知道其中原因。”“你對自己要求更严格一点，reviewers 就会对你温柔一点。”

### 您研究和教学工作繁忙，学术服务活动也非常活跃，在时间和精力上您如何协调

没有什么特殊的技巧，肯定是比较辛苦的，但也不缺动力，毕竟是自己的兴趣所在。现在尽量控制自己出差的次数，以便有效地利用时间。



**白翔**，华中科技大学电子信息与通信学院教授，博导，国家防伪工程中心副主任。先后于华中科技大学获得学士、硕士、博士学位。曾先后访问于美国 Temple 大学和加州大学洛杉矶分校，

入选微软铸星计划。

他的主要研究领域为计算机视觉与模式识别，具体包括目标识别、形状分析、图像视频文字检测与识别及智能交通系统。尤其在形状的匹配与检索、场景 OCR 取得了一系列重要研究成果，引起了国际同行的关注，入选 2014、2015 年中国高被引学者。他的研究工作曾获微软学者、国家自然科学基金优秀青年基金的资助。担任中国计算机学会计算机视觉专委会（CCF-CV）常务委员、中国图象图形学学会理事，是视觉与学习青年研讨会（VALSE）在线活动主要发起人之一。

## 王亮：在视觉大数据领域“慧眼识珠”

李实英转自《中国科学报》(2016年8月8日第6版,作者:《中国科学报》记者沈春蕾)

计算机视觉就是给计算机安装上眼睛(摄像机)和大脑(算法),使其能模拟人类视觉系统的结构和功能,实现对周围环境视觉信息(图像/视频)的智能分析与处理,进而感知环境、适应环境。



中国科学院自动化所研究员、模式识别国家重点实验室副主任王亮主要从事计算机视觉和模式识别等相关领域的研究。因出色的科研表现,王亮日前荣获

第十四届“中国青年科技奖”。该奖项旨在造就一批进入世界科技前沿的青年学术和技术带头人,表彰奖励在国家经济发展、社会进步和科技创新中作出突出成就的青年科技人才,每两年评选一次,每届获奖人数不超过100名。

王亮在接受《中国科学报》记者采访时说:

“我的研究工作之一就是通过计算机视觉方法,让人们可以便捷迅速地从大规模、没有关联的视觉大数据中获取所需的重要信息。”

### 求学海外

2004年,王亮在中科院自动化所获得工学博士学位。2004~2010年,他先后在英国帝国理工学院、澳大利亚莫纳什大学、澳大利亚墨尔本大学及英国巴斯大学工作。

他告诉记者:“我在前面三所大学主要从事博士后的研究工作,2009年在墨尔本大学的博士后项目结题后我开始找正式工作,当时收到来自英国和澳大利亚的两份录用通知书,最终选择了巴斯大学的讲师工作。”

6年两个国家、4所高校的工作和学习,王亮收获颇丰。无论是在英国还是在澳大利亚,王亮称自己不仅了解了异国的文化及风土人情,还接受了不同的教育和科研理念。

“要玩就玩得痛快,要学就学得踏实。”这是国外导师对科研工作和生活态度的真实写照。“在国内从事科研工作加班加点是常事,但国外科研工作者将工作和生活通常划分得很清楚。”王亮说,“导师在周末或度假的时候几乎是不会处理工作事情的。”

除了高效的研究小组讨论会外,他还发现,国外导师不管多忙,每周都会认真阅读学生的工作周报并给出意见和建议;同时每周也会留出固定时间与每位学生面对面地交流。通过在巴斯大学的工作经历,王亮也体会到了不同的教育体制和教学经验。

通过比较国内外的科研环境,王亮认为,国内的硬件设施并不比国外差,可能比国外还好;但软实力还有待提升,比如科研基础和创新性等。

### 报效国家

“我是在国内读的博士,当时毕业时很想出去看看和体验一下。”王亮向记者回忆了他当初回国的想法,“在国外呆得越久,越发感觉原来的好奇和新鲜感正在慢慢褪去,从国外可以学习的东西也越来越有限。”一旦动了回国的念头,便一发不可收拾。

2010年,王亮成功入选自动化所“百人计划”招聘,为此他辞退了巴斯大学的终身教职。“我忘不了当时辞职时系主任惊愕的表情,他说海外经济不景气找正式工作是多么不易,一再问我你确定辞职吗?他不能了解,这对我来说,不仅仅是回国,

更是回家。”中科院自动化所就是王亮口中所说的那个“家”。

如果自己所学能为自己的国家作点贡献，哪怕只是一点点，心里的满足感也将是不可言喻的。回国5年来，王亮获得了国家杰出青年科学基金，“百人计划”终期评估优秀。因在模式识别领域的突出贡献，王亮在2014年被授予“国际模式识别学会会士”荣誉称号。目前他是中科院脑科学与智能技术卓越创新中心骨干人员、中国计算机学会计算机视觉专委会秘书长、中国电子学会青年科学家俱乐部副主席、中国图像视频大数据产业技术创新战略联盟秘书长等。

如今，王亮已申请或授权发明专利47项，出版编著、专著7部，发表或接收论文176篇，其中发表在《IEEE 模式分析与机器智能汇刊》《IEEE 图像处理汇刊》等领域高质量国际期刊38篇；发表在IEEE国际计算机视觉大会、IEEE国际计算机视觉与模式识别会议等领域顶级国际会议27篇。

王亮向记者透露，他们团队的远距离步态识别技术已经顺利实现转化，成立了银河水滴（北京）科技有限公司。公司将专注于人工智能和计算机视觉技术，引领作为人工智能核心的深度学习技术的突破与革新，提供世界领先的视觉大数据分析技术。

### 潜心科研

从自动化所出国，再回到自动化所，并取得了一系列的成绩，王亮非常感谢自己的导师——中科院副院长谭铁牛院士。刚回国的時候，王亮在谭老师指导下，主要致力于拓展智能识别与数字安全团队的研究工作，主要研究方向是“视觉模式分析与理解”。

在动态图像序列模式分析与理解方面，“我们针对监控视频，不仅要分析场景中人是谁（身份）的问题，更要解决他或他们在执行什么活动（行为）的问题，同时考虑人的身高、性别、年龄、体态等相关属性的推理问题。”王亮介绍道。

王亮在国际上较早开始步态识别研究，并系统研究了步态识别理论与方法。他发表在国际顶级期刊《IEEE 模式分析与机器智能汇刊》上的首篇步态识别论文成为该领域经典文章，目前已经实现步态识别从单一视角到跨视角的突破，较好解决了步态识别易受视角变化影响这一难题。

他进一步解释道，我们力争通过系统性的研究，实现从“简单、个体”行为到“复杂、群体”行为的深度研究，为行为识别技术的研究与应用提供算法基础。

在图像模式分析与理解方面，王亮主要研究图像分类、检索和分割三块相互关联的重要内容。他还就这三块内容进行了说明：图像分类是判断图像中目标的类别或者场景类别，图像检索是根据图像的表达式来寻找与其相似的图像，图像分割是找到图像中属于某个特定类别目标的区域。

王亮团队提出了图像分类中特征编码的一般理论框架，给出了局部特征编码方法的族谱图和演化图，2014年连续4个月成为《IEEE 模式分析与机器智能汇刊》的热点文章。

此外，王亮团队提出了基于深度语义排序的哈希编码方法，较好解决了多标签图像中如何保留复杂的多级语义结构的难题；提出多尺度上下文深度卷积神经网络的前景分割算法，提升了复杂图像目标分割的精确度和鲁棒性。

除了步态识别、个体/群体行为识别、图像分类与分割等应用研究外，王亮还针对大规模、多模态数据分析和深度学习及应用等方面开展一系列理论和方法研究。

多年的科研工作也让王亮发现，虽然国内很多技术在国际上已经处于领先地位，但仍缺少原创性的发现，一些研究工作不够深入和系统。他希望自己和团队能够潜心科研，做出一些真正有影响力的、原创性的、有价值的工作。

## CCF-CV 走进高校系列报告会

中国计算机学会计算机视觉专业委员会（CCF-CV）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，在全国范围的高校和科研院所等开展 CCF-CV 走进高校系列报告会活动。详情请见 <http://ccfcv.ccf.org.cn>。

<p><b>天津科技大学·天津（第二十一期）</b></p> <p>2016 年 9 月 29 日（星期四）</p> <p>执行主席：<b>杨巨成</b>教授，天津科技大学</p>	<p><b>10 月活动安排</b></p> <p>第 22 期 走进中科院深圳先进技术研究院·深圳，执行主席<b>乔宇</b>研究员</p> <p>（注：各月活动的具体时间请关注秘书处的最新通知）</p>
<p><b>11 月活动安排</b></p> <p>第 23 期 走进兰州理工大学·兰州，执行主席<b>李晓旭</b>副教授</p> <p>第 24 期 走进天津大学·天津，执行主席<b>胡清华</b>教授</p>	<p><b>12 月活动安排</b></p> <p>第 25 期 走进中国石油大学（华东）·青岛，执行主席<b>宫文娟</b>博士</p> <p>第 26 期 走进西安电子科技大学·西安，执行主席<b>苗启广</b>教授</p>

### 华中科技大学·武汉（第十六期）

2016 年 6 月 21 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十六期在华中科技大学国际学术交流中心 8 号楼报告厅举行。

报告会由中国计算机学会计算机视觉专委会委员、华中科技大学**桑农**教授主持。中国科学院自动化所**王亮**研究员、西安电子科技大学**苗启广**教授、南京信息工程大学**刘青山**教授、中山大学**郑伟诗**副教授以及武汉大学**荆晓远**教授受邀出席本会，并就计算机视觉相关专业领域做出精彩的学术演讲。华中科技大学、武汉大学、华中师范大学、武汉理工大学、中国地质大学、华南理工大学、河南师范大学、中南民族大学、常州大学、武汉华目信息技术

有限公司、基恩士（中国）有限公司等多家单位相关人士参加了本次报告会。面对精彩的讲座内容，同学们积极提问，学术氛围浓厚。

**王亮**博士以“Multi-modal Data Analysis with Application to Retrieval and Classification”为题，介绍多模态数据分析的重要性及应用背景，针对多模态数据的一致性与互补性，介绍其研究小组在多模态数据分析方面的研究以及在检索与分类方面的应用，最后，给出多模态数据分析未来的研究方向。

**苗启广**博士以“单幅图像去霾方法”为题，介绍其小组去霾方面的研究工作：根据霾与雾等天气条件的物理特性之差异，提出了一种基于黄霾物理特性的复原方法；基于融合的思想，从单幅输入图像生成两幅增强图像，通过两者的融合使复原图像同时得到近景和远景处的细节信息；提出了一种基

于熵和模糊融合的方法,该方法能够有效解决暗通道方法中白色物体景深估计错误的问题。

刘青山博士以“Feature learning for visual recognition”为题,针对计算机视觉中特征学习问题介绍其研究小组在视觉识别领域的研究工作。首先介绍了基于子空间学习的视觉识别,及其在人脸对齐和遥感图像处理等方面的应用。还介绍了一种基于超图的视觉理解方法,利用超图代替传统的数据中的复杂关系建模。

郑伟诗博士以“行人重识别:发展与挑战”为题,分析了行人重识别面临的困难,介绍了其研究小组在视域敏感的跨视域度量模型、深度先验正则化行人特征表达、以及基于视频的行人重识别模型方面的研究工作:针对行人残缺和开集测试情况,提出一种局部相对比较度量学习模型。最后,给出了行人重识别未来的研究方向。

荆晓远博士以“视图特征学习及行人重识别”为题,介绍了其研究小组在多视图特征学习方面的研究工作,在多视字典学习和多视子空间学习方面提出一系列的方法,实验证明了其有效性。最后,介绍其在基于图像与视频特征学习的行人重识别方面的研究工作。

此次CCF-CV走进高校系列报告会现场学术气氛浓厚,报告结束后师生踊跃发言,参与讨论与交流,讲者对提出的问题进行了详细的讲解,收效良好。至此,第十六期CCF-CV走进高校系列报告会于华中科技大学圆满结束。



## 西北工业大学·西安(第十七期)

2016年7月9日,中国计算机学会计算机视觉专委会走进高校系列报告会第十七期活动“计算机视觉前沿技术及应用”,在西北工业大学成功举行。

执行主席、西北工业大学韩军伟教授邀请了西安交通大学薛建儒教授、美国佛罗里达大学杨林副教授、浙江大学李玺教授、清华大学鲁继文副教授四位专家做专题报告。来自西北工业大学、西安交通大学、中科院自动化所、西安电子科技大学、西北农林科技大学、西安邮电大学、陕西师范大学、西安科技大学、咸阳师范学院等高校教师和研究生等近150人参加了本次报告会。西北工业大学校长助理张艳宁教授代表学校致欢迎词。

薛建儒博士带来一场题为“智能车的视觉场景理解的集成计算框架”的精彩报告。薛博士简要介绍了视觉与学习在智能车上的最新进展,然后介绍课题组近年来基于无人驾驶实验平台所提出的以视觉为主导,融合立体对应、显著性交通要素及其他传感信息的跨模态跨尺度时空对齐与配准计算框架。最后,薛博士为大家展示了课题组近年来参加国家自然科学基金委组织的“中国智能车未来挑战赛”的视频片段。

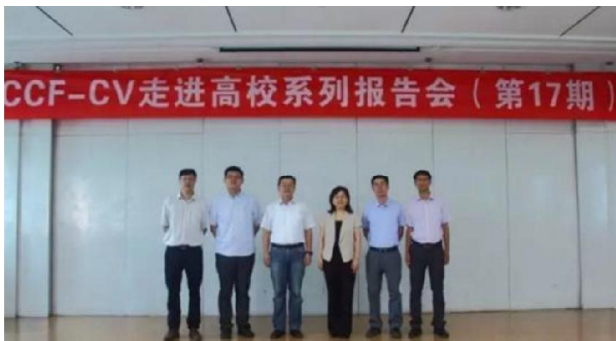
杨林博士的报告题目是“Large Scale Biomedical Image Analysis and Imaging Informatics”。他首先介绍了生物医学图像分析和图像信息学(包括生物学、病理学、放射学)的内容,然后回顾了一些新兴生物医学图像分析应用程序中使用机器学习工具,比如深度学习,最终希望激发不同领域的研究人员合作,应对个性化医疗所带来的挑战。

李玺博士做了“Deep Correspondence Structure Learning for Person Re-identification and Self-paced Boosting Learning for Classification”的报告。他首先从行人视觉感知特性、深度学习器构建机制、自节奏学习、鲁棒学习等多维度视角进行了深入剖析,并引入了行人再识别对应结构学习和自节奏

boosting 学习所涉及的主要研究问题和技术方法；然后系统地展示了行人再识别和自节奏 boosting 学习的实验对比性能；最后和大家探讨了涉及视觉特征学习所面临的一些开放性问题 and 难题。

鲁继文博士的报告题目是“距离度量学习在人脸识别中的应用”。他主要介绍了他们课题组近年来提出的系列距离度量学习方法：包括代价敏感度学习、近邻排斥度量学习、稀疏重构度量学习、局部多核度量学习和深度度量学习等，以及它们如何应用于人脸识别中的特征表示和相似性度量以提升识别性能。

此次 CCF-CV 走进高校系列报告会之西北工业大学站的讲座持续了三个多小时，现场学术气氛浓厚，互动活跃，每位讲者在报告后与提问嘉宾进行了热烈的讨论。通过此次活动，不同高校同行之间推进了交流与合作，提升各高校在计算机视觉研究和应用中的能力。



### 电子科技大学·成都（第十八期）

2016 年 7 月 15 日下午，第十八期中国计算机学会计算机视觉专业委员会（CCF-CV）走进高校系列报告会在电子科技大学图书馆成功举行，来自中国科学院自动化研究所、成都计算机研究所、南京大学、武汉大学、国防科技大学、东南大学、北京师范大学、北京航空航天大学、西安电子科技大学、电子科技大学、四川大学、西华大学、重庆大学、重庆邮电大学、重庆师范大学、东北林业大学、河北工业大学等高校的专家、老师和学生共计 60

余人参与了本次活动，报告会现场座无虚席。

电子科技大学机器人研究中心执行主任、自动化工程学院程洪教授作为此次活动的执行主席主持本次报告会。上海科技大学国家青年千人计划学者高盛华研究员、中国人民大学信息学院金琴副教授、西北工业大学光学影像分析与学习中心王琦副教授分别做社会机器人中的机器学习与视觉计算领域的前沿专题报告。

高盛华博士为大家带来了题为“Crowd Scene Understanding with Multi-column Convolutional Neural Networks”的报告。他从 2015 年发生在我们身边的上海踩踏事件为例解释了人群行为分析在安防监控系统的应用和意义，进一步介绍了存在遮挡、人群密集的图像中如何计算人数的方法，以及在人群密集的图像中侦测异常行为的方法，为面向公共安全的社会机器人提供了典型应用案例。

金琴博士做了题为“Semantic Image Profiling for Historic Events”的报告。她从社会学和人工智能的交叉领域对历史事件下了一个定义“An echo of the past in the future. A reflex from the future on the past”，指出对历史图片进行语义分析遇到的困难，然后分享了如何利用卷积神经网络在记录有历史事件的图像中用一句话概括出事件发生的时间、地点、人物和其他物体信息。并向大家演示了用 Image Profiling 的方法可以合成穿越时空效果的历史性照片。金博士的研究给出了社会机器人学在历史事件检索领域的新应用方向，反映出社会机器人与社会学跨界融合的趋势。

王琦博士在题为“视觉技术在交通数据解析中的应用”的报告中，首先分析了国家 2016 年提出的《机动车运行安全技术条件》对智能驾驶技术发展的推动，接着从检测、跟踪、识别三个方面讲解了基于增量学习的交通标志检测，特别介绍了适用于国内复杂道路的基于序列分类的深度学习交通标志检测方法，最后向大家展示了西北工业大学在智能驾驶车辆视觉感知系统的研究进展。

本次 CCF-CV 走进高校活动的讲座紧凑而高效，内容丰富多彩。整场报告会持续 3 个半小时，现场学术气氛浓厚，互动活跃，透析当下无人驾驶、深度学习和语义理解这 3 个热门机器人与人工智能话题，讲者在报告后和茶歇过程中与同学们进行了热烈的讨论。



### 中国海洋大学·青岛（第十九期）

2016 年 8 月 23 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十九期在中国海洋大学崂山校区图书馆第二会议室举行。报告会由中国计算机学会计算机视觉专委会委员、中国海洋大学董军宇教授和中国海洋大学仲国强副教授共同主持。中国科学院自动化研究所王亮研究员、复旦大学姜育刚教授和浙江大学李玺教授受邀出席本会并做精彩报告。来自中国海洋大学、中国石油大学（华东）、中国科学院海洋研究所、青岛大学、青岛科技大学、青岛理工大学、青岛农业大学、海信集团等单位约 150 人参加了本次报告会。面对精彩的报告内容，老师同学们积极互动，学术氛围浓厚。

王亮博士的报告题目为“深度学习及在视觉模式分析中的应用”。他在报告中首先简单介绍了以大数据时代为契机的深度学习的发展，以及深度学习发展所经历的和正在经历的三个发展阶段，并指出了视觉大数据的兴起及其重要性；然后重点介绍了近年来在深度学习和视觉数据分析方面的一些研究工作，如深度语义检索、深度关系学习、群体

行为分析、相关 RBM、骨架行为识别、图像集的深度表示、多帧超分辨率等；最后对未来的发展方向做了进一步探讨，引发了到场老师和同学们的深刻思考。

姜育刚博士的报告题目是“Video Content Recognition with Deep Learning”，主要介绍了在当前大数据的背景下如何利用深度学习方法来解决海量视频内容识别问题。姜博士以最近建立的一些互联网视频数据集作为引入，详细讲述了几种深度学习方法在视频分析和理解中的有效应用，如改进的稠密轨迹、特征编码、基于图像的深度视频分类等；最后进一步讨论了未来的研究方向。

李玺博士的报告题目是“基于深度结构学习的行人再识别研究以及自节奏 boosting 驱动的鲁棒分类研究”。李博士在报告中主要围绕计算机视觉和机器学习领域行人再识别和 boosting 学习问题，从行人视觉感知特性、深度学习构建机制、自节奏学习、鲁棒学习等多维度视角对问题进行了深入剖析，并引入了行人再识别对应结构学习和自节奏 boosting 学习所涉及的主要研究问题和技术方法，然后系统地展示了行人再识别和自节奏 boosting 学习的实验对比性能。另外，还介绍了近年来利用视觉特征学习进行视觉语义分析和理解所做的一系列代表性的研究工作及其实际应用。最后对涉及视觉特征学习所面临的一些开放性问题 and 难题进行了深入讨论。

三位讲者从自身研究领域出发，就深度学习及其在计算机视觉的前沿应用问题展开深入探讨，为相关领域的研究人员和技术人员提供交流学术成果、进行广泛讨论的平台。报告会内容精彩，高潮迭起。报告过程中，专家与听众们密切互动和交流，对大家关注的问题进行详细解答。通过 CCF-CV 走进高校系列报告会，真正实现了同学们与领域专家的近距离交流。与会的老师和学生们纷纷对 CCF-CV 搭建的这样一场高水平的学术大餐表达深切的谢意。（仲国强、董军宇）



### 厦门大学·厦门 (第二十期)

2016 年 9 月 3 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二十期在厦门大学海韵行政楼 B-505 会议室举行。报告会由中国计算机学会计算机视觉专委会委员、厦门大学纪荣嵘教授主持。来自北京大学的查红彬教授、中山大学的赖剑煌教授和天津大学的胡清华教授受邀出席本期活动并做精彩报告。来自厦门大学、华侨大学、厦门理工大学和芬兰 Oulu 大学的师生参加了本次报告会。本次报告会的主题是“计算机视觉前沿技术及应用”, 三位报告人结合自己的研究内容作了精彩纷呈的报告。

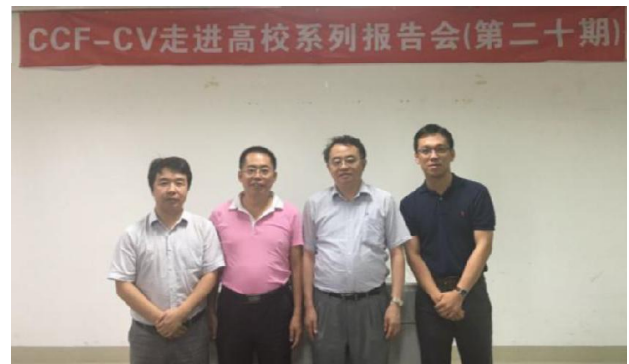
查红彬博士的报告题目为“3D Reconstruction for Object Modeling and Scene Analysis”。他首先介绍了三维物体重建与场景分析的。就相关研究话题分析了物理网络系统 (CPS) 中存在的问题, 介绍其研究组使用 catadioptric 相机系统进行三维建模的研究, 详细阐述了使用两面镜系统简化相机标定过程的方法及原理。查博士分析了当前采集三维数据的各类传感器存在的问题, 引入其团队在深度图像分辨率提高方面的创新性工作。最后向听众展示了该研究在实际问题中的应用价值。

赖剑煌博士以“行人再标识问题的若干研究进展”为题介绍了他们团队在该领域非常深入的工作。首先给出行人再识别问题定义和该问题的挑战性, 然后报告归纳了两类方法: 摄像机之间有重叠视域及摄像机间具有非重叠视域。后者是当前研究的主

流, 该类方法常见的有基于行人特征表达的方法、训练行人匹配模型的方法、上下文线索的挖掘、基于迁移学习的跨场景行人再标识。最后, 赖博士重点讲解了其团队在行人再标识领域陆续提出的 6 个方法: 基于行人特征表达的方法; 基于深度排序和联合表达的行人再标识; 基于不对称激励学习的行人再标识; 基于部分匹配的行人再标识; 基于深度信息的行人再标识; 现实场景中基于去身份混淆的行人再标识。每种方法都图文并茂地展示了计算过程和详细的实验结果对比。

胡清华博士以“大数据时代的不确定性挑战”为题, 报告了多源异构不确定数据的机器学习中存在的挑战与对策。胡博士介绍他们团队与国家天文台空间天气预报、林火概率预报和石油钻井平台检测三个应用需求的研究情况, 展示实际应用的实例数据, 启发性地从属性视角、决策视角、样本视角分析大数据的特点, 令在场观众耳目一新。另外, 基于总结出的数据特性, 以风速分布建模预报为例, 分享了针对大数据学习任务的解决思路。最后总结其丰富应用经验, 深入讨论了未来工作的展望。

三位讲者学术功底深厚、严谨又不失幽默, 为本期 CCF-CV 走进高校系列报告会活动奉献了三场高质量的专业学术报告, 获得了厦门高校师生们的热烈反馈。报告内容从经典的相机标定、行人标识、超分辨率, 到热门的深度学习、大数据三维大数据等, 在提问与回答环节, 专家与听众们密切互动和交流, 对大家关注的问题进行详细解答。活动在师生的热烈掌声中圆满结束。(黄剑波、纪荣嵘)



## CCF-CV 走进企业系列交流会

中国计算机学会计算机视觉专业委员会（CCF-CV）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，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。本活动每月安排一次左右，名额有限，委员优先，报名从速。详细情况请关注 <http://ccfcv.ccf.org.cn>。

### 走进格林深瞳 北京（第七期）

时间：2016 年 9 月 29 日下午（星期四）

地点：北京市海淀区玉带路 6 号

（活动日程等信息请关注专委秘书处的最新通知）

### CCF-CV 走进阿里巴巴·杭州（第六期）

2016 年 6 月 29 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六期活动“走进阿里巴巴”，在杭州市阿里巴巴西溪园区成功举办。

本期活动由 CCF-CV 秘书处召集，浙江大学许威威研究员（CCF YOCSEF 杭州地区 AC 委员）带队。活动通知发出后，参会申请邮件纷至沓来，由于接待方场地限制，最终秘书处选择了来自企业、高校、科研院所等 40 多名代表参加了此次活动。阿里云研究员、IEEE Fellow 华先胜、阿里巴巴 OS 事业群高级专家易东、阿里巴巴无线事业部高级专家袁岳峰参与了本次活动，并分别就视觉大数据智能分析、识别和搜索，YunOS 中的计算机视觉技术，VR 淘宝等三个专题作了报告分享。

首先，华先胜博士以“视觉大数据智能分析、识别和搜索”为题作了精彩的专题报告。他简要回顾了图像识别与搜索研究背景，阐述了图像识别与图像搜索在商业应用上密不可分的关系。华先胜博

士在报告中通过手机“拍立淘”应用向大家生动展示了阿里巴巴在图像分析方面的前沿技术，即图像识别与图像搜索相结合、局部特征与深度学习相结合。与会者通过自己手机上安装的手机淘宝 app 的“拍立淘”应用，直观地感受到科技带来的便利与科技带来的巨大商机。

接下来，由阿里高级专家易东介绍阿里 YunOS——世界第三大手机操作系统中与计算机视觉相关的核心技术，包括基于人脸特征点检测的人脸美颜、人脸视觉增强、虚拟化妆等多个与淘宝关系紧密的应用，也包括 YunOS 中的基于人脸识别技术的人脸解锁功能。在报告中，易东研究员向参会人员展示了他们的方法在各大公开的人脸数据集上的识别准确率，阿里在这方面已经达到了世界领先水平。

之后，活动进入了茶歇环节。在阿里同事的精心安排下，参会人员一边享用阿里巴巴为大家精心准备的水果，一边就彼此感兴趣的话题进行热烈的讨论，并与阿里主讲嘉宾进行了深入探讨。

在茶歇过后，由阿里巴巴高级专家袁岳峰为大家带来最后一个报告—VR 淘宝。这个报告主要介绍了阿里巴巴围绕 VR（虚拟现实）这一当今非常火热的研究热点在软件与硬件方面的投入，着重介绍了虚拟试衣、VR tour 和 Buy+三个产品导向的研究内容。报告最后通过一项虚拟试衣的 Demo 生动地展示了 VR 技术在网购中的巨大应用潜力。

本次活动全程比较轻松和愉快，三位专家和与会人员在报告过程中进行了深入的探讨与交流。听众提问积极踊跃，对阿里专家报告深感兴趣。大企业计算机视觉技术直观的现实应用与所面临的技

术挑战给参会者留下深刻印象。活动结束后，大家一同合影留念。



## Computer Vision Resources

李实英摘自 *The Ultimate List of 300+ Computer Vision Resources*

(<https://hackerlists.com/computer-vision-resources>)。

以下链接已于 2016 年 8 月 26 日确认更新(原网址有些链接无效)。

### BOOKS

- Computer Vision: Models, Learning, and Inference – Simon J. D. Prince 2012 (可下载 PDF 和课件)
- Computer Vision for Visual Effects – Richard J. Radke, 2012 (有 code 下载)
- Computer Vision: A Modern Approach (2nd edition) – David Forsyth and Jean Ponce 2011
- Visual Object Recognition synthesis lecture – Kristen Grauman and Bastian Leibe 2011 (可下载 PDF)
- Computer Vision: Theory and Application – Rick Szeliski 2010 (可下载 PDF)
- High dynamic range imaging: acquisition, display, and image-based lighting – Reinhard, E., Heidrich, W., Debevec, P., Pattanaik, S., Ward, G., Myszkowski, K 2010
- Multiple View Geometry in Computer Vision – Richard Hartley and Andrew Zisserman 2004 (可下载不同格式的插图)
- Computer Vision – Linda G. Shapiro 2001
- Vision Science: Photons to Phenomenology – Stephen E. Palmer 1999

### COURSES

- EENG 512 / CSCI 512 – Computer Vision – William Hoff (Colorado School of Mines)
- Visual Object and Activity Recognition – Alexei A. Efros and Trevor Darrell (UC Berkeley)
- Computer Vision – Steve Seitz (University of Washington)

- Visual Recognition – Kristen Grauman (UT Austin)
- Language and Vision – Tamara Berg (UNC Chapel Hill)
- Convolutional Neural Networks for Visual Recognition – Fei-Fei Li and Andrej Karpathy (Stanford University)
- Computer Vision – Rob Fergus (NYU)
- Computer Vision – Derek Hoiem (UIUC)
- Computer Vision: Foundations and Applications – Fei-Fei Li (Stanford University)
- High-Level Vision: Behaviors, Neurons and Computational Models – Kalanit Grill-Spector and Fei-Fei Li (Stanford University)
- Advances in Computer Vision – Antonio Torralba and Bill Freeman (MIT)
- Computer Vision – Bastian Leibe (RWTH Aachen University)
- Computer Vision 2 – Bastian Leibe (RWTH Aachen University)

### PAPERS

- CVPapers – Computer vision papers on the web
- CVonline – Bob Fisher (University of Edinburgh)
- SIGGRAPH Paper – Graphics papers on the web
- NIPS Proceedings – NIPS papers on the web
- Computer Vision Foundation open access
- Annotated Computer Vision Bibliography – Keith Price (USC)
- Calendar of Computer Image Analysis, Computer Vision Conferences – (USC)

## 计算机视觉相关的主要国际期刊专刊

(李实英搜集整理未来三个月以内的专刊征文通知, 详情见专刊链接。欢迎承担专刊编辑的委员来信。)

---

### Pattern Recognition Special Issues on

<http://www.journals.elsevier.com/pattern-recognition/call-for-papers/>

#### ■ Distance Metric Learning for Pattern Recognition

Important dates:

Paper submission: September 15, 2016

Final decision: June 15, 2017

Tentative publication: September, 2017

Guest editors:

Jiwen Lu, Tsinghua University

Ruiping Wang, ICT, CAS

Ajmal Mian, The University of Western Australia

Ajay Kumar, Hong Kong Polytechnic University

Sudeep Sarkar, The University of South Florida

#### ■ Pattern Recognition for High Performance Imaging

Important dates:

Paper submission: January 5, 2017

Final decision: July 15, 2017

Tentative publication: September 5, 2017

#### ■ Articulated Motion and Deformable Objects

Important dates:

Paper submission: December 15, 2016

Final decision: April 1, 2017

Tentative publication: September 15, 2017

Guest editors:

Josef Kittler, University of Surrey

Francisco Perales, Universitat de les Illes Balears

Sergio Escalera, Computer Vision Center and

University of Bcelona

Guest editors:

Xiao Bai, Beihang University

Jun Zhou, Griffith University

Antonio Robles-Kelly, Data61, CSIRO

---

### IEEE Transactions on Circuits and Systems for Video Technology Special Issues on

[http://tcsvt.polito.it/call\\_for\\_papers.html](http://tcsvt.polito.it/call_for_papers.html)

#### ■ Large Scale and Nonlinear Similarity Learning for Intelligent Video Analysis

Important dates:

Paper submission: December 31, 2016

Acceptance notification: June 30, 2017

Final Manuscripts: December, 2017

Guest editors:

Wangmeng Zuo, Harbin Institute of Technology

Liang Lin, Sun Yat-Sen University

Alan L. Yuille, The Johns Hopkins University

Horst Bischof, Graz University of Technology

Lei Zhang, Microsoft Research

Fatih Porikli, Australian National University

**Neurocomputing** Special Issues on

<http://www.journals.elsevier.com/neurocomputing/call-for-papers/>

■ Emergence in Human-like Intelligence toward Next-Generation Web

Important dates:

Paper submission: September 30, 2016

Final decision: March 31, 2017

Tentative publication: July 31, 2017

Guest editors:

Neil Yen, University of Aizu

Odej Kao, Technische Universität Berlin

Qun Jin, Waseda University

Ching-Hsien Hsu, Chung Hua University

■ Machine Learning for Non-Gaussian Data Processing

Important dates:

Paper submission: October 15, 2016

Final decision: June 15, 2017

Tentative publication: September 15, 2017

Guest editors:

Zhanyu Ma, Beijing University of Posts and Telecommunications

Jen-Tzung Chien, National Chiao Tung University

Zheng-Hua Tan, Aalborg University

Yi-Zhe Song, Queen Mary University of London

Jalil Taghia, Stanford University

Ming Xiao, KTH- Royal Institute of Technology

## 计算机视觉相关的 2016 主要国内国际会议

(李实英搜集整理未来 6 个月以内的相关会议论文截止日期, 依时间排序。尽管已参考 CCF 推荐和网络资源, 并在会议网站核查, 仍难免有疏漏之处, 详情请亲自确认。)

会议名称	重要日期	网站链接和会议地址
<b>FG 2017</b> 12th IEEE International Conference on Automatic Face and Gesture Recognition	Paper submission: Sep 26, 2016 Manuscripts due: Jan 15, 2017 Main conference: May 30-Jun 3, 2017	<a href="http://www.fg2017.org/">http://www.fg2017.org/</a> (Washington, DC, USA)
<b>WACV 2017</b> IEEE Winter Conference on Applications of Computer Vision	Paper submission: Oct 3, 2016 Paper notification: Dec 20, 2016 Main conference: Mar 24-31, 2017	<a href="http://pamitc.org/wacv2017/">http://pamitc.org/wacv2017/</a> (Santa Rosa, CA, USA)

<p><b>CVPR 2017</b> 30th IEEE Conference on Computer Vision and Pattern Recognition</p>	<p>Paper submission: Nov 8, 2016 Paper notification: Mar 3, 2017 Main conference: Jul 21-26, 2017</p>	<p><a href="http://cvpr2017.thecvf.com/">http://cvpr2017.thecvf.com/</a> (Honolulu, Hawaii)</p>
<p><b>MVA 2017</b> 15th IAPR International Conference on Machine Vision Applications</p>	<p>Paper submission: Dec 5, 2016 Paper notification: Feb 13, 2017 Main conference: May 8-10, 2017</p>	<p><a href="http://www.mva-org.jp/mva2017/">http://www.mva-org.jp/mva2017/</a> (Nagoya, Japan)</p>
<p><b>ICIAR</b> 14th International Conference on Image Analysis and Recognition</p>	<p>Paper submission: Jan 9, 2017 Manuscripts due: Mar 1, 2017 Main conference: Jul 5-7, 2017</p>	<p><a href="https://www.aimiconf.org/iciar17/">https://www.aimiconf.org/iciar17/</a> (Montreal, Canada)</p>
<p><b>MIUA 2017</b> 21st Medical Image Understanding and Analysis Conference</p>	<p>Paper submission: Feb 24, 2017 Manuscripts due: Mar 31, 2017 Main conference: Jul 11-13, 2017</p>	<p><a href="https://miua2017.wordpress.com/">https://miua2017.wordpress.com/</a> (Edinburgh, UK)</p>
<p><b>ICCV 2017</b> 19th International Conference on Computer Vision</p>	<p>Paper submission: Mar 1, 2017 Paper notification: Jul 1, 2017 Main conference: Oct 22-29, 2017</p>	<p><a href="http://www.cv-foundation.org/ICCV2013/ICCV2017-Bid-Venice.pdf">http://www.cv-foundation.org/ICCV2013/ICCV2017-Bid-Venice.pdf</a> (Venice, Italy)</p>

## 委员好消息

([李实英](#)搜集整理专委委员的晋级、获奖等各类好消息, 欢迎委员来信或推荐。以下依时间排序。)

- ☆ 7 位专委委员, 西安电子科技大学[董伟生](#)教授、东南大学[耿新](#)教授、中国科学院自动化研究所[赫然](#)研究员、复旦大学[姜育刚](#)教授、中山大学[林惊](#)教授、同济大学[王瀚漓](#)教授、杭州电子科技大学[俞俊](#)教授, 获得 2016 年国家优秀青年科学基金资助项目 (<http://www.cingta.com/>, 2016.08.18)。
- ☆ 专委副主任、北京大学[查红彬](#)教授, 专委秘书长、中国科学院自动化研究所[王亮](#)研究员, 专委委员、北京交通大学[于剑](#)教授等获得 2016 年国家自然科学基金重点项目资助 (<http://www.nsf.gov.cn/>, 2016.08.17)。
- ☆ 专委副主任、中国科学院计算技术研究所[陈熙霖](#)研究员连同专委委员、国防科学技术大学[郭裕兰](#)博士获得博士后创新人才计划项目资助 (<http://www.chinapostdoctor.org.cn/>, 2016.08.04)。
- ☆ 专委常务委员、北京大学[林宙辰](#)教授获得 2016 年国家杰出青年科学基金项目资助 (<http://www.nsf.gov.cn/>, 2016.08.04)。
- ☆ 专委主任、中国科学院[谭铁牛](#)院士应邀参加英国约克大学举行的 2016 年毕业典礼暨学位授予仪式, 并被授予约克大学荣誉博士学位。这是继白春礼、潘云鹤、陈竺等中国学者被授予该校荣誉博士学位后, 我国又一科学家获得此项荣誉 (<http://www.cas.cn/>, 2016.08.01)。
- ☆ 专委秘书长、中国科学院自动化研究所[王亮](#)研究员和专委委员、北京理工大学[黄华](#)教授获得 2016 年国家重点研发计划重点专项资助 (<http://service.most.gov.cn/>, 2016.06.22)。
- ☆ 专委委员、国防科学技术大学[郭裕兰](#)博士指导团队研制的“鹰眼”自动跟踪无人机获“华为杯”第十一届中国研究生电子设计竞赛最高奖“团体特等奖” (<http://digitalpaper.stdaily.com/>, 2016.09.05)。
- ☆ 专委委员、中山大学[林惊](#)教授, 专委副主任、中山大学[赖剑煌](#)教授, 以及中山大学谢晓华研究员各自指导的团队在“维盛杯”第三届全国研究生智慧城市技术与创意设计大赛中分别获得一、二、三等奖。林惊教授另外获得优秀指导奖 (<http://sdc.sysu.edu.cn/>, 2016.08.31)。
- ☆ 专委委员、华南理工大学[金连文](#)教授和张鑫副教授带领的团队在 CVPR 2016 大会 VIVA 国际挑战赛的手部检测及分类两项任务中均获得第一 (<http://www2.scut.edu.cn/>, 2016.07.20)。
- ☆ 专委委员、中科院深圳先进技术研究院[乔宇](#)研究员团队在 CVPR 2016 大会 LSUN 大规模场景分类、ChaLearn 性别表情识别和装饰品识别、ActivityNet 大规模行为识别三项竞赛中均获得第一 (<http://www.tiat.ac.cn/>, 2016.07.14)。
- ☆ 专委委员、东南大学[耿新](#)教授团队在 CVPR 2016 大会 Chalearn 年龄估计竞赛中获得第二名 (<http://cse.seu.edu.cn/>, 2016.07.03)。