

CCF

计算机视觉 专委会简报

COMPUTER VISION NEWSLETTER

2018/02 期
总第 12 期

专委动态

走进高校系列活动

科技前沿

研究热点追踪

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 上海科技大学

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 上海科技大学

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 焜 合肥工业大学
编 委 黄 岩 中国科学院自动化研究所
刘海波 哈尔滨工程大学
张汗灵 湖南大学

资源平台

主 编 沈沛意 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 山东财经大学
金 鑫 北京电子科技学院
李 策 兰州理工大学
刘 丽 国防科学技术大学

COMPUTER VISION NEWSLETTER

CONTENTS

目录

专委动态

走进高校	走进高校系列活动	04
会议推介	PRCV 2018 诚邀申请组织竞赛	05

科技前沿

专题综述	生物视觉中的物体识别通道	06
热点追击	视觉神经信息编解码	14
	基于多层次或多尺度 FCN 的显著性检测方法	15

委员风采

委员访谈	西北工业大学韩军伟教授访谈	16
委员好消息		19

资源平台

数据集	计算机视觉中的行人检测数据集	20
开源代码	除雾方法及源码整理	22
招聘信息		27
征文通知		29

CCF-CV 走进高校系列报告会

第 46 期 南京邮电大学

时间：2018 年 1 月 7 日

第 46 期走进高校系列活动在南京邮电大学成功举办，圆满的完成了 2018 年至关重要的第一棒。本期报告邀请南昌大学**闵卫东**教授、南京信息工程大学**刘青山**教授以及中山大学**林惊**教授三位专家做特邀报告。报告会由南京邮电大学通信与信息工程学院承办，通信与信息工程学院**周全**副教授为执行主席。来自东南大学、南京航空航天大学、南京理工大学等高校约 140 名师生参加了本次报告。

南京邮电大学通信与信息工程学院副院长**邵曦**教授致欢迎辞。此次组织委员会邀请了人工智能、深度学习最前沿的一些专家，目的是让该领域的同僚们能进行良好的沟通交流，共同促进中国的计算机事业、人工智能的发展。



此次 CCF-CV 走进高校系列报告持续了三个多小时。专家与听众亲切互动和交流，最后活动在雷鸣般的掌声中圆满结束。

第 47 期 中国矿业大学

时间：2018 年 1 月 12 日

第 47 期走进高校系列活动在中国矿业大学成功举办。本期报告邀请北京大学**查红彬**教授、南京信息工程大学**刘青山**教授以及中兴通讯云计算及 IT 研究院**董振江**研究员三位专家做特邀报告。报告会由中国矿业大学计算机科学与技术学院承办，计算机科学与技术学院**周勇**教授、**姚睿**副教授为执行主席。来自中国矿业大学、江苏师范大学等高校约 150 名师生参加了本次报告。



报告会由中国矿业大学计算机科学与技术学院副院长**牛强**教授主持，表示此次中国计算机学会计算机视觉专委会走进中国矿业大学，一定会激发计算机科学与技术学院老师和同学们的科研热情，提升学院的科研水平。

此次 CCF-CV 走进高校系列报告持续了三个多小时。与会师生获得了与专家近距离交流和学习的机会。

(责任编辑：马占宇)

PRCV2018

诚邀申请组织竞赛

中国模式识别与计算机视觉学术会议 (Chinese Conference on Pattern Recognition and Computer Vision, PRCV) 是由中国模式识别学术会议 (CCPR) 和中国计算机视觉大会 (CCCV) 合并而来, 由中国人工智能学会 (CAAI)、中国计算机学会 (CCF)、中国自动化学会 (CAA) 和中国图象图形学学会 (CSIG) 联合主办, 定位国内顶级的模式识别和计算机视觉领域学术盛会。首届中国模式识别与计算机视觉学术会议 (PRCV2018) 将于 2018 年 11 月 23-26 日在广州举行, 由中山大学承办。

现诚挚邀请公司或单位在 PRCV 2018 组织与承办计算机视觉、模式识别与机器学习方面的各种技术竞赛。在 PRCV 的前身 RACV 2016 和 CCCV 2017 的竞赛单元中, 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、Xilinx (赛灵思) 公司、银河水滴科技 (北京) 有限公司、北京航天宏图信息技术股份有限公司、中国科学院西安光学精密机械研究所、北京健康促进会心血管影像中心等机构分别承办了视频图像分析、模糊车牌图像清晰化、视频标注、FPGA 视觉计算、遥感图像目标提取、心血管 OCT 易损斑块识别等相关技术竞赛, 共吸引了 260 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, 有力推动了计算机视觉与模式识别学科及相关产业的发展。具体信息可参考 RACV 2016 和 CCCV 2017 竞赛单元网站。

有意愿在 PRCV 2018 组织或承办竞赛的公司或单位需要于 2018 年 4 月 15 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 PRCV 2018 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、报名方式和竞赛结果等信息。

竞赛申请书需包括以下信息:

竞赛名称	竞赛目的与意义
竞赛参与者要求	参赛团队注册方式
报名方式	主要时间节点
任务设置	知识产权归属
结果的评价方式	结果提交方式
奖项设置和奖励方法	专题研讨会组织方式 (包括场地与时间需求)
竞赛数据、硬件资源的准备与使用方式	竞赛组织方 (包括联系方式)

申请书需要以 PDF 或 word 文档的形式提交给竞赛单元主席。PRCV2018 大会还将为竞赛承办方分别配专门的研讨会时间, 并协助竞赛组织者为参会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, 请及时联系竞赛单元主席。

姓名	邮箱
谢晓华 (中山大学)	xiexiaoh6@mail.sysu.edu.cn
鲁继文 (清华大学)	lujiwen@tsinghua.edu.cn

(责任编辑: 毋立芳)

生物视觉中的物体识别通道

中国科学院自动化研究所 胡占义

一、引言

从计算机视觉当前的研究状况看，以下二方面最有可能取得突破性进展：1) 新的数学方法，如 Boykov 等于 1999 年提出的关于图割 (Graph Cuts) 的优化近似算法[1]，由于显著提高了优化速度，使得这些算法已成为目前图优化的基本算法；2) 基于生物视觉(或生物视觉启发)的方法。近年来，深度学习在图像物体识别方面取得了“变革性”进展，其层次化的网络结构，神经元“感受野”的逐层增加机制，“简单细胞”，“复杂细胞”等概念，都来自于视神经科学。所以，了解一些生物视觉的知识，特别是群体神经元对图像物体的表达机理(或编码机制)，将有助于更好地理解深度网络对图像物体的内在表达机制，并进一步提高深度网络和深度学习在物体识别和场景解释方面的性能。

不同生物种类的视觉系统千差万别，如猴子和老鼠。即使同一种类，如猴子，不同实验手段下得到的一些定量结果也会相差很大。如“猴子的视觉皮层”的总面积到底有多大，不同的文献出处，差距也会很大。所以，对从事信息科学的人员来说，由于习惯于“数字是上帝”，往往会因为神经领域的“数字混乱”感到不知所措。所以，当信息领域人员遇到统计数据差异很大时，应该认为这是“神经领域”的正常现象。神经领域是一个典型的小样本下的外推(extrapolation)领域，统计数据出现大的差异也是不可避免的。

另外，生物视觉既“能力超然”又“愚蠢不堪”。人们可以瞥一眼一张开会的照片，就能给出合理的解释和描述，但人们也经常出现幻觉(illusion)。所以，从学术的观点探究生物视觉的奥秘是必要的，并在可能的情况下加以借鉴。但如果一定认为生物视觉是计算机视觉“学习的榜样”，也许是完全错误的。人工智能先驱 Marvin

Minsky 曾指出智能来源于“多样性”，不是一些特有的“原理”[2]。智能如此，视觉亦如此。当将来有一天人们对生物视觉系统对信息的加工机理基本了解清楚后(可能永远无法完全了解清楚)，也许人们会发现，生物视觉系统“没有什么特殊性”，仅仅是“大量神经元连接的一个超复杂网络系统”，这个“复杂系统”仅仅对“生物日常生存需要的一些功能”体现出优良的性能而已，对生物其它不太需求的功能，其“视觉能力”并没有什么优势。所以，人们绝不能过高估计生物视觉的能力，也不宜借鉴生物视觉不太擅长的能力。

二、视觉系统的基本组成

视觉系统主要由皮层(cortex)和皮下组织(subcortex)构成。认知功能主要靠皮层完成。视觉皮层是指主要对视觉刺激产生响应的皮层。严格地说，很少有某一皮层区域(cortical area)仅仅对一种刺激产生响应，文献中说的“视觉皮层”，“听觉皮层”和“触觉皮层”等，仅仅表示该区域对某种刺激主要产生响应而已。另外，“视觉功能”必然有视觉皮层参与，但绝不是完全靠视觉皮层完成。视觉认知(visual cognition)的大多数任务都几乎需要大脑的各个皮层区域及皮下组织的协作来共同完成。

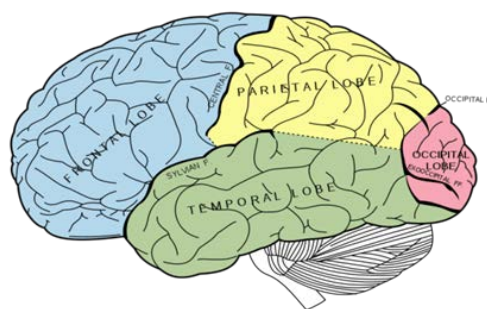


图 1 大脑皮层分区

如图 1 所示，大脑皮层主要分为 4 个不同的

区域：枕叶 (occipital lobe)，颞叶 (temporal lobe)，顶叶 (parietal lobe) 和额叶 (frontal lobe)。

视觉信息加工机制，主体上说，是一个层次化加工过程。目前很多文献中认为深度网络(DNN)借鉴了生物视觉加工机制，也主要指这种“层次化结构的相似性”。物体在视网膜成像后，首先传到后面的枕叶，进行底层特征提取。然后传到颞叶，进行物体识别；从枕叶另一路传到顶叶，进行运动检测和空间位置确定等。额叶主要进行高级认知功能。额叶在人类快到 30 岁才能成熟，尔后体积慢慢缩小，老年痴呆病人的额叶体积会明显减小。

2.1 视网膜及其信息加工机制

不少计算机视觉研究人员，往往认为视网膜 (retina) 与照相机一样，是一种光电转换机构。事实上视网膜有 3 层信息加工过程。如图 2 所示，感光细胞 (photoreceptor) 进行光电转换后，信号通过双极细胞层 (bipolar cells) 中继后，通过神经节细胞层 (ganglion cells) 进一步处理，最后输出到丘脑的外漆体 (LGN)。可以看出，视网膜的信号为“从后向前”的处理机制。

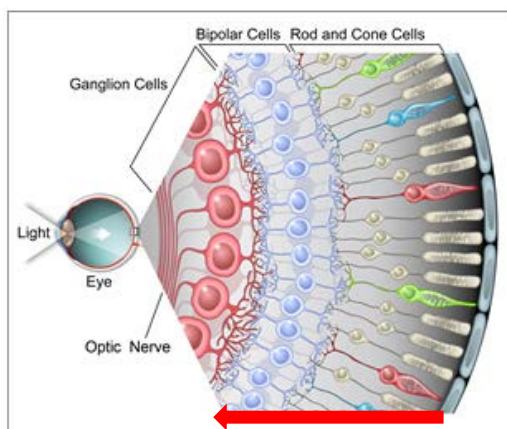


图 2 视网膜的三层处理结构

感光细胞

感光细胞进行光电转换，分为二类：锥状细胞 (cones) 和杆状细胞 (rods)。人类视网膜大约有 5-6 百万个锥状细胞，1 亿 2 千多万杆状

细胞。杆状细胞主要用作夜视。杆状细胞可以捕获单个光子，多个杆状细胞共同作用，在微暗的光度下形成夜视能力。锥状细胞主要分红、绿、蓝三种类型，它们之间的比例约为：红：62%；绿：32%，蓝：2%。这种比例关系与目前的彩色相机中不同颜色的 CCD 单元个数相同的结构也有大的区别。锥状细胞主要分布在视网膜的中央凹 (fovea) 周围，是产生正常清晰视觉的神经基础。由于中央凹非常小，所以正常情况下人们要不断地调整“注视点” (Focus of attention) 进行大范围观察。

双极细胞和神经节细胞

双极细胞 (bipolar cell) 主要起信号中继作用。神经节细胞 (RGC: Retina Ganglion Cell) 大约有 1 百多万个。分为大细胞 (Paraso RGC) 和小细胞 (Midget RGC)。大细胞对运动敏感，对空间频率不敏感，反应速度快；小细胞对颜色敏感，对空间频率敏感，反应速度慢。大细胞与小细胞的比例个数约为：1：9。从上面的介绍可知，从感光细胞到神经节细胞，细胞个数得到大量压缩，视觉信息已得到很复杂的处理，所以，视网膜不能等价为一台照相机。我们日常感受到的五彩缤纷的世界，都是从一百多万个神经节细胞输出的信息进一步加工得到的。从这个意义上说，提高照相机的分辨率未必是提高“计算机视觉”能力的有效途径。

细胞的感受野

细胞的感受野 (receptive field) 是指该细胞对空间视张角内 (或视网膜对应区域内) 刺激产生响应的区域。当一个神经元接收多个神经元的输入时，该神经元对应的感受野为多个输入神经元感受野的组合。从 $V1 \rightarrow V2 \rightarrow V4 \rightarrow PIT \rightarrow IT$ 信息的逐层处理过程中，对应的神经元的感受野越来越大，每层之间感受野增大的系数大体为 2.5。V1: 1.3 度；V2: 3.2 度；V4: 8.0 度；PIT: 20 度；IT: 50 度。感受野越大，说明对应神经元编码的图像特征越复杂 [3]。

2.2 物体视觉和空间视觉通道

如图 3 所示，视觉信息从视网膜成像并加工处理后，先通过丘脑(thalamus)的外漆体(LGN: Lateral geniculate nucleus of thalamus)中继到枕叶的初级视皮层(V1 或 striate cortex)进行初步加工,然后主要分成两个通道(pathway)进行处理:背部通道(dorsal pathway)进行空间视觉(spatial vision)、腹部通道(ventral pathway)进行物体视觉(object vision)。

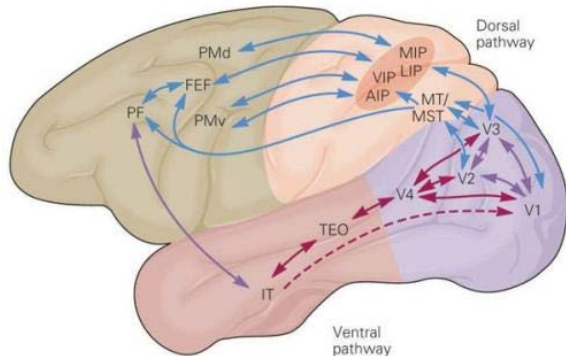


图 3 视觉信息的层次化处理

背部通道：空间视觉的层次化信息加工

空间视觉指为“行为和动作”服务的视觉，如视差计算，运动估计，坐标系变化等。空间视觉主要在枕叶和顶叶(parietal lobe)完成。从 V4 到顶叶后，处理后的信息从后顶叶皮层(Posterior parietal: PP)又投射到三个不同的脑皮层区。具体情况如下：

- 1) 投射到前额皮层(Prefrontal)，参与空间工作记忆(Spatial working memory)；
- 2) 投射到运动前皮层(Premotor)，执行视觉引导的动作(Visually guided action)；
- 3) 投射到内侧颞叶(Medial temporal lobe)，用于导航(Navigation)。

腹部通道：物体视觉的层次化信息加工

腹部通道的主要功能是物体视觉，即对图像物体进行精细识别。图像物体识别是一个层次化处理过程，从 V1→V2→V4→PIT→AIT 区，神经元对图像刺激形状的选择性越来越复杂。IT 区

(inferior temporal cortex) 被认为是视觉物体表达和识别的最高区。

腹部通道的前端(AIT: Anterior IT)又进一步投射到三个皮层区域和三个皮下区域，进行进一步的“高层语义”加工。

投射到的三个皮下区域及功能为：

1) 到新纹状体(neostriatum)，支撑依据视觉的习惯和技能形成；

2) 到腹侧纹状体(ventral striatum)，支撑刺激“效价”的赋值。效价(Valence)是一个心理学术语，指根据“吸引力”(attractiveness)赋予某人或某事的评价值；

3) 到杏仁体(amygdala)，支撑情感刺激处理。

投射到的三个皮层区域及功能为：

1) 到内侧颞叶(medial temporal)，支撑物体长时和上下文记忆；

2) 到眼窝前额皮质(orbitofrontal)，参与奖励处理(mediate reward processing)；

3) 到腹外侧前额叶(ventrolateral prefrontal)，参与物体工作记忆(mediate object working memory)。

腹部通道从 IT 区(背部通道从 PP 区)后，处理的信息不再为单纯的视觉信息，视觉信息参与后续的多模态信息融合、记忆和推理等高级认知功能。所示，视觉能力绝不是仅仅靠“视皮层”完成的。视觉能力是大脑各功能模块协同处理的结果。所以，“理解视觉系统”与“理解整个大脑”本质上没有多少区别。

三、腹部通道和物体识别

腹部通道是物体表达和识别的通道。深度网络所谓的“模拟脑结构”也主要指腹部通道的这种分层处理机制。首先，生物视觉系统的层数(解剖区)一般也就是 10 多层，不像 DNN 动辄几十层，甚至几百层，上千层；其次，生物视觉同一层内部的神经元连接更加普遍(约占 80%)，层与层之间的连接约占 20%；另外，生物视觉的跨层连

接非常普遍，基本上很少有不同区域没有神经连接的现象，仅仅体现为连接强度的差别而已。这些均与目前的 DNN 结构有区别。

图 4 给出了皮层不同区域之间的连接关系及连接强度。图中对应的方块大小表示皮层的面积大小，连线粗细表示连接强度。

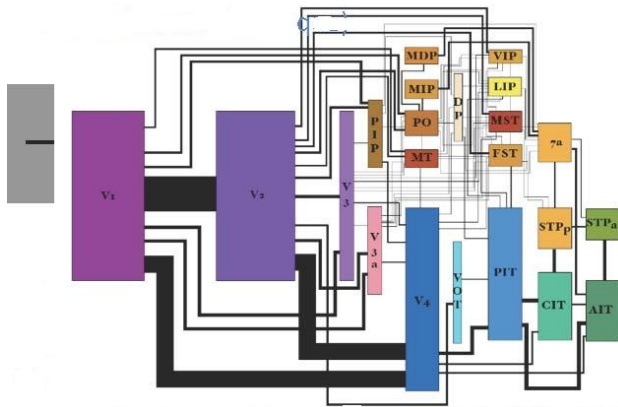


图 4 视皮层不同区域的面积大小及连接强度

每层神经元的个数、输入输出神经元的个数和从刺激出现到该层神经元出现响应的潜伏期 (latency) 均不相同[4]。例如，从视网膜到达 V1 区的潜伏期约为 50ms, 达到 AIT 的时间约为 100 ms。V1 区接收 LGN 的输入个数约为 1 百万左右, V1 区的神经元个数约为 1 亿 9 千万, V1 区的输出约为 7 千 7 百万。

目前除了 V1 区细胞对刺激的加工机制比较清楚外, V2、V4、IT 区的功能尚不是很清楚。大体上说, 从 V2 区开始神经元对错觉轮廓 (illusory contour) 有反应, 已形成色调敏感区; V4 区为颜色感知的主要区域, 参与曲率计算、运动方向选择和背景分离; IT 区为物体表达和识别区。这里需要指出的是, IT 区并没有形成完全的与视角无关的表达。研究表明[5], IT 区细胞的感受野的大小分布不均匀。既有大感受野细胞, 也有中小感受野细胞。IT 区存在小感受野细胞, 说明这些细胞仅仅对物体的局部敏感。

3.1 简单细胞和复杂细胞

简单细胞和复杂细胞 (simple cell and

complex cell) 是 Hubel 等人研究猫的 V1 区细胞对图像边缘 (和运动) 的响应时提出的概念[6]。简单细胞指其响应对边缘 (运动) 的具体位置密切相关的细胞, 复杂细胞指对边缘 (运动) 在一个小范围内变动其响应不明显变化的细胞。Hubel 等认为复杂细胞是由简单细胞组合得到的, 则复杂细胞的感受野就是简单细胞的感受野的空间组合。Hubel 等的复杂细胞, 并不是指“对复杂刺激有响应的细胞”。目前在深度学习领域, “复杂细胞”的概念有点“过度泛化”。在视觉通道的高层区域, 如 IT 区, IT 神经元对输入刺激的形状选择已很复杂, 但似乎没有人称 IT 细胞为复杂细胞。神经领域有“老祖母细胞”(grandmother cell) 的提法, 是指该细胞对特定面孔有响应的细胞。目前看来, 是否存在老祖母细胞也尚无定论。因为实验时发现, 在上百张面孔图像中, 有的细胞确实仅对其中的一幅图像有响应, 但当测试图像多时, 特别是当图像之间的相似性增大时, 那些老祖母细胞是否仍仅对一幅特定面孔图像有响应、对相似性面孔图像无响应仍是一个未定的问题。

3.2 神经元的响应表达和物体表达

单个神经元的响应就是指神经元对视觉刺激发放的脉冲序列。目前的研究表明, 将一段时间 (如 IT 神经元在 50ms -180ms) 神经元发放脉冲个数的平均数作为该神经元对刺激的响应, 是一种有效的响应表达方式。所以, 文献中关于神经元的响应, 一般就是指“给定时间段内发放脉冲个数的平均值”。

单个神经元的作用很小, 群体神经元才能对视觉物体进行表达 (或编码)。群体神经元对某个视觉物体的表达, 就是将同一皮层区域多个神经元对该物体的响应表示为一个向量, $\mathbf{x} = (x_1 \ x_2 \ \dots \ x_n)$, 其中 x_i 为第 i 个神经元对输入图像刺激的响应, 用该向量来作为对该图像物体的表达。

这里需要强调一下, 神经文献中在比较不同表达模型对同一物体的表达之间的关系时, 如果二种表达之间满足一个线性变换, 则认为是相同

表达。例如，假定猴子 IT 区对第 i 个输入图像物体的表达为 X_i^{Monkey} ，该物体在某种深度网络下的表达为 X_i^{DNN} ，如果这二种表达之间满足线性变换关系，即： $X_i^{DNN} = AX_i^{Monkey} + B \quad \forall i$ ，则认为 X_i^{DNN} 和 X_i^{Monkey} 为相同表达，或深度网络给出了一种优秀的对猴子 IT 区的物体表达的定量预测模型。神经元个数固定后，由于输入图像的个数可以非常大，所以对二种不同的表达，上面的线性变换关系不可能对所有输入图像都满足。当然具体应用中，任何时候不可能严格满足这种线性关系，需要利用某些统计方法来进行显著性验证。

3.3 群体神经元对图像物体的表达方式

对视差 (disparity) 和运动 (motion) 等刺激，由于这些刺激可以用参数描述，如运动可以用运动方向和大小来描述，所以神经元对这些刺激的响应曲线可以比较容易测定和参数化表示。但对图像物体，很难给出参数化的响应曲线。文献中分析猴子对图像物体的表达时，先将某个神经元对所有刺激的响应表示为一个概率密度函数，称为该神经元的选择性函数，或将所有神经元对同一刺激的响应（即群体编码）表示为一个概率密度函数，称为群体神经元的稀疏性函数，然后对所有选择性函数（或稀疏性函数）进行统计分析。如利用分布函数的 4-阶矩来刻画其非高斯性，用尾巴指数 (Tail index) 来刻画分布函数的大尾巴性等。分布函数的“尾巴”越大，说明个体神经元的选择性越高，或群体神经元的编码稀疏性越高。

Lehky 等 [7][8] 对二只猴子 674 个 IT 神经元对 806 幅图像的响应（选择性函数和稀疏性概率密度函数）分析表明，IT 神经元的编码具有如下特性：

1) 单个神经元的选择性都不高，但 IT 区存在大量这样的神经元。这种 IT 区神经元的编码方式与传统模式识别理论的物体表达方式不同，主要体现在神经元的表达是一种“过表达” (over-complete representation)，而传统模式识别领

域的物体表达方式则希望不同的表达尽量相互独立。IT 神经元的这种过表达方式与目前 DNN 神经元的表达方式具有相似性；

2) 尽管 IT 区有大量神经元 (AIT 区有约 6 千 6 百多万万个神经元)，但这些神经元响应之间独立的个数不多，一般不超过 100 个。

目前对猴子的物体视觉系统研究表明，如果一种数学表达方法可以与猴子 IT 区神经元对物体的表达相媲美，则该物体表达方法对输入图像物体一定同时具有高的物体分类 (categorization) 和鉴别 (identification) 能力，即这种表达是一种优秀的表达 [9][10]。

Dong 等 [11] 的近期工作表明，DNN (以 VGG19 作为 DNN 研究模型) 神经元的编码特性与猴子 IT 神经元的编码特性有本质的差别。他们的结果同时表明，一种优秀的表达不一定非要与猴子 IT 神经元的表达相似。

3.4 猴子视觉系统对脸孔图像的表达

鉴于面孔在生物生存和社交中的重要性，面孔识别一直是神经领域研究猴子物体识别机理的重要内容。2010 年，Freiwald 等 [12] 发现，猴子对面孔的识别逐渐从视角相关到视角无关。猴子后外侧皮层 (PL: posterior lateral cortex) 的神经元的响应与视角相关，但前内侧皮层 (AM: Anterior medial cortex) 的神经元的响应已与视角无关，达到了对面孔图像的不变识别 (invariance)。

Landi 等 [13] 进一步研究发现，猴子除了上面面孔识别的核心系统外，另有二个区域：一个位于颞极区 (temporal pole)，另一个位于鼻周皮层区 (perirhinal cortex)，主要负责对“熟悉面孔” (familiar face) 的识别。而且，对熟悉面孔的响应呈非线性性，可能体现了人们对曾经见过的人“突然想起来了”的现象。

Chang 等 [14] 研究发现，当面孔图像用一个包含纹理和几何信息的参数向量表示时，则每个 IT 神经元均有自己的一个特定向量 (STA: Spike-Triggered Average)，该神经元对任何一幅新输

入的面孔图像的响应就是该神经元的 STA 与输入图像参数向量的内积。说明猴子 IT 神经元在面孔图像的参数空间具有“简单而优美”的线性编码机制。这一结果与目前的 DeepFace 等深度网络对面孔图像的表达机制有很大区别。

3.5 神经元响应定量建模：目标驱动的框架

DiCarlo 等 2016 年在 Nature Neuroscience 发表了一篇“Perspective”文章，提出了所谓的“目标驱动的感知信息建模方法”[15]。Dicarlo 等认为，只要将物体分类性能作为训练一个深度卷积网络 (HCNN) 的优化目标函数，则训练好的 HCNN 就能很好地预测猴子 IT 神经元对物体的编码。在这种目标驱动的框架下，他们通过在 ImageNet 上训练了一个 6 层的卷积神经网络[15]，训练时代价函数仅仅为物体分类精度，则训练好的网络的输出不仅可以“定量预测”IT 神经元的响应，而且可以预测物体的位置、大小等几何信息。由于训练时仅仅“控制”物体的分类能力，没有使用任何 IT 神经元响应的信息，所以他们称这类建模方法为“目标驱动”的建模。

目标驱动的框架在一些情况下可以对 IT 神经元的响应进行很好预测，但是否如 DiCarlo 等认为的是一种“通用性建模”方法值得磋商。文献中已有报道[16]，相同的 DNN 深度网 (Alex 网络) [17]，当使用 4 种不同的初值训练时，所得到的 4 个 Alex 网络的图像正确分类率基本相同，但 4 个不同 Alex 网最后一层神经元之间的表达无法用一个线性变换表示。作者将这种现象称之为“收敛学习”与“发散学习”现象。这些结果对 DiCarlo 等提出的“目标驱动的感知信息建模方法”提出了挑战[18]。因为网络结构相同、仅仅不同初值下训练得到的不同 DNN，就出现“发散学习”现象，当网络结构不同时，这种“发散学习”现象就会更普遍。另外，从数学的观点看，HCNN 的神经元表达是一种“过表达”，而同一问题的多个等价过表达之间一般不存在线性变换关系。

四、大脑皮层之间的反馈

视觉皮层之间从高层到底层的反馈机制到底是什么？是从高到低逐层反馈？还是从高层先反馈到 V1 区，然后从 V1 区再逐层上传到 V2 和 V4 等后续区域呢？这个问题长期以来一直是一个有争议的问题，即使目前文献中也仍然存在不同的看法。Buffalo 等[19]通过研究注视效果与“潜伏期”长短的关系，发现注视效果到达 V4 区的时间最短，然后是 V2 区，最后是 V1 区，说明“注视”的反馈效果是从高层到底层的，如图 5 所示。

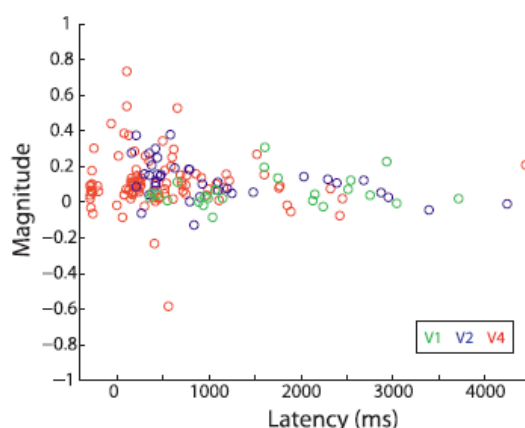


图 5 视觉注视从高层到底层的逐层反馈机制

对计算机视觉的研究人员来说，每当提及反馈，首先问的问题是：反馈什么？这似乎是一个“不太科学”的问题。事实上，包括 VGG 等前馈深度神经网络，底层对高层到底上传了什么，人们似乎也并不清楚。文献中给出的一些有特定意义的“感受野”，如对应眼睛、耳朵等的神经元感受野，也仅仅是挑选出来的极少个例，并不具有一般性。所以，对一个复杂的网络研究其反馈机制时，似乎仅仅研究“反馈环路构成”和“对应的反馈强度”就足够了。“反馈强度”由于受“视觉任务”影响，也是一个难以定义和测定的量，所以，计算机视觉界似乎可以借鉴“脑连接组学” (Connectomics) 的成果，特别是不同皮层区域之间的神经元的投射比例 (neuron projection ratio) 关系，来探索视觉反馈机制。

Friston[20]和 Markov 等[21][22]的研究表明，反馈具有如下的一些基本属性：

1) 高层到底层的反馈起调制作用。“调制”

表示“反馈无法单独激活相应神经元”，仅仅对从底层输入的信号“增强”或“抑制”；调制同时表示“高层对底层作用的区域比较大”。前馈“底层对高层作用的区域（感受野）相对要小”；

2) 皮层之间的反馈连接现象比正向连接更普遍；

3) 跨层反馈现象更普遍；

4) 反馈强度与距离呈指数规律下降；

5) 反馈既可以通过双向神经回路（底层到高层和高层到底层），也可以通过单向回路（仅仅高层到底层）进行；

6) 反馈的信号较前馈要慢（50 多毫秒），前馈层与层之间约为 10 毫秒左右。

Markov 等 [21][22] 通过测定不同脑皮层区之间的神经元投射比例关系 FLN (Fraction of Labeled neurons) 发现：如果 A 区投射到 B 区的距离为 d ，则投射比例（强度）FLN 的概率服从下面的指数分布，其中 $\tau = 0.188 \text{ mm}^{-1}$ ：

$$P(d) = e^{-\tau d}$$

A 区投射到 B 区的神经元比例 $FLN_{A \rightarrow B}$ 的定义为：

$$FLN_{A \rightarrow B} = \frac{\text{区域 A 投射到区域 B 的神经元个数}}{\text{区域 B 的总神经元个数}}$$

Markov 等发现，给定目标区域，其它区域投射到目标区域的 FLN 服从负对数正态分布 (lognormal)。也就是说，其它投射区域到目标区域之间的距离服从正态分布。换句话说，大多数投射区域与目标区域的距离在正态分布的均值

附近 ($\mu = 26.57 \text{ mm}$, $\sigma = 10.11 \text{ mm}$)。

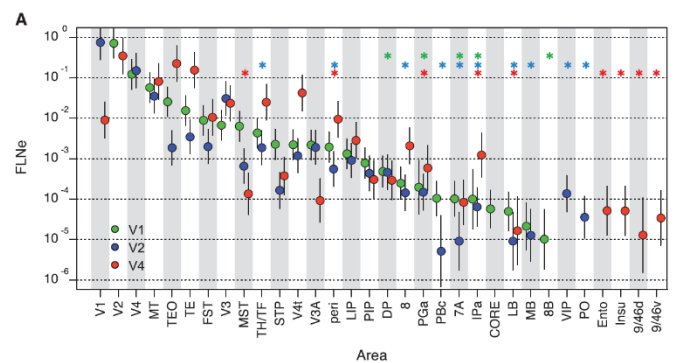


图 6 其它脑皮层对 V1, V2, V4 区投射的神经元比例

从图 6 可以看出，即使对用于初级加工的视皮层 V1, V2, V4，也有大量其它区域的神经元的投射关系。一般来说，邻域之间的投射要多，但也有些例外。有些区域尽管不是邻域关系，但它们之间的神经投射也接近邻域投射的比例强度。如果把大脑皮层之间的不同区域看作脑网络的节点的话，那么对应的脑皮层网络是一个“密集连接”的网络，而不是过去人们认为的仅仅存在局部连接的“小世界”(Small world) 的网络结构。另外，从图 6 也可以看出，其它皮层投射到 V1, V2, V4 的 FLN 的数值有 5 个数量级的差别。如何借鉴这些信息，来构建具有反馈机制的深度网络，是一个值得探索的未来研究方向。

(责任编辑：申抒含)

作者声明：由于有些图是从网上查阅得到的，很难给出具体出处，所以文中图并没有注明具体文献出处，对此笔者表示抱歉并对原作者一并给予致谢。

参考文献

- [1] Boykov Y. et al. Fast approximate energy minimization via graph cuts, ICCV, 1990.
- [2] Minsky M. The Society of Mind, New York: Simon & Schuster. 1986.
- [3] Rolls. T. W & Deco G. Computational Neuroscience of Vision, Oxford University Press, 2004.
- [4] DiCarlo J. J. et al. How Does the Brain Solve Visual Object Recognition? Perspective, Neuron, 73: 415-434, 2012.
- [5] Lafwe-Sousa R et al. Parallel, multi-stage processing of colors, faces, and shapes in macaque IT cortex, Nature Neuroscience, 16:1870-1878, 2013.
- [6] Hubel D H & Wiesel T N. Receptive fields of single neurons in the cat's striate cortex. The Journal of Physiology, 148(3): 574-

- 591, 1959.
- [7] Lehy R. S. et al. Statistics of visual responses in primate inferotemporal cortex to object stimuli, *J. Neurophysiology*, 106:1097-1117, 2011.
 - [8] Lehy R. S. et al. Dimensionality of object representations in monkey inferotemporal cortex, *Neural Computation*, 26: 2135-2162, 2014.
 - [9] Yamins D. L.K. et al. Performance-optimized hierarchical models predict neural responses in higher visual cortex, *Proc. Natl. Acad. Sci.* 111(23): 8619-8624, 2014.
 - [10] Yamins D. L. K & DiCarlo J. J. Using goal-driven deep learning models to understand sensory cortex, *Nature Neuroscience*, 19(3):356-365, 2016.
 - [11] Dong Q. L. et al. Statistics of Visual Responses to Object Stimuli from Primate AIT Neurons to DNN Neurons, *Neural Computation*, 30(2): 447-476, 2018.
 - [12] Freiwald W A & Tsao D Y. Functional compartmentalization and viewpoint generalization within the macaque face-processing system. *Science*, 330(6005): 845, 2010.
 - [13] Landi S.M & Freiwald W. A. Two areas for familiar face recognition in the primate brain, *Science*, 357: 591–595, 2017.
 - [14] Chang L & Tsao D. Y. The Code for Facial Identity in the Primate Brain, *Cell*, 169: 1013–1028, 2017.
 - [15] Yamins D. L. K & DiCarlo J. J. Explicit Information for category-orthogonal object properties increases along the ventral stream, *Nature Neuroscience*, 19(4): 613-622, 2016.
 - [16] Li Y et al. Convergent learning: Do different neural networks learn the same representations? In *International Conference on Learning Representations*, 2016.
 - [17] Krizhevsky A et al. ImageNet classification with deep convolutional neural networks. In: *Advances in Neural Information Processing Systems*, 25: 1106–1114, 2012.
 - [18] Dong Q. L. et al. Commentary: Using goal-driven deep learning models to understand sensory cortex, *Frontiers in Computational Neuroscience*, 12(4), 2018.
 - [19] Buffalo et al. Desimone A backward progression of attentional effects in the ventral stream, *Proc. Natl. Acad. Sci. USA*, 107: 361-365, 2010.
 - [20] Friston K. Learning and inference in the brain. *Neural Networks*, 16: 1325:1352, 2003.
 - [21] Markov T. N. et al. Weight consistency specifies regularities of macaque cortical networks, *Cerebral Cortex*, 21: 1255-1272, 2011.
 - [22] Markov T. N. et al. Anatomy of hierarchy: Feedforward and feedback pathways in Macaque visual cortex, *The Journal of Comparative Neurology: Research in Systems Neuroscience*, 522: 225-259, 2014.



胡占义

中国科学院自动化研究所模式识别国家重点实验室研究员。主要研究方向包括基于生物视觉的视觉建模、三维计算机视觉等，2004年“视觉计算理论与算法研究”获国家自然科学基金二等奖（3/5）。曾任国家自然科学基金委员会信息学部第三、第四届专家咨询委员会委员、国家绕月探测科学应用专家委员会专家，国家863计划“十五”信息获取与处理技术主题专家、ICCV 2005 Local Co-Chair, ACCV 2012 Program Co-Chair。

Email: huzy@nlpr.ia.ac.cn

视觉神经信息编解码

中国科学院自动化研究所 何晖光 杜长德

视觉神经信息编码指人脑将外部视觉刺激转换成神经活动信号的过程。研究人脑视觉神经信息的编码机制，对于提高机器的智能感知能力具有重要意义。另一方面，利用非侵入式脑神经影像采集手段，如功能磁共振成像（fMRI），可以在一定程度上解读大脑视觉神经信息，此即视觉神经信息解码。视觉神经信息的编解码研究（图 1）促进了认知神经科学和类脑智能的发展。

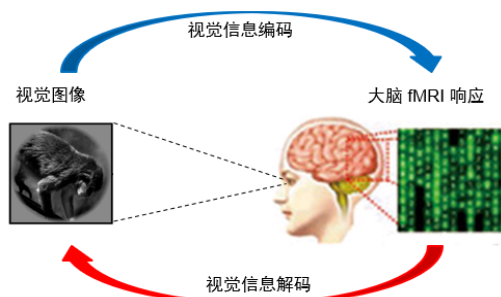


图 1 视觉神经信息编解码

目前的视觉神经信息编解码研究涵盖了对初级、中级以及高级视觉特征的分类、辨识或重建。但大多数方法在根据大脑响应重建视觉场景时效果不佳，且它们通常基于线性变换，与人类视觉系统的信息处理机制不符。近期，以深度学习为基础的人工智能系统已经在许多任务上达到甚至超越了人类水平。虽然深度学习方法比起人脑的学习能力还有很多局限，但多项研究表明深度神经网络在视觉信息处理方面与人脑具有类似之处。因此，基于深度学习的视觉神经信息编解码研究逐渐成为计算神经科学和类脑智能领域的研究热点。

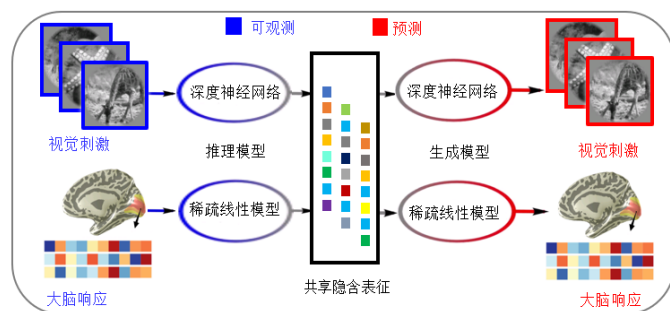


图 2 基于深度生成式多视图模型的视觉信息编解码

基于上述讨论，我们提出了一种基于深度神经网络的视觉神经信息编解码方法“深度生成式多视图模型”（DGMM，图 2）。对于视觉刺激，我们利用深度神经网络从原始图像中逐层提取可解释性强的特征，这和人类视觉系统具有一定的相似性，对建立视觉编码模型非常重要。对于大脑响应，我们采用稀疏贝叶斯学习从大量的 fMRI 体素中自动筛选出对视觉信息解码较重要的体素。这有效降低了数据维数和样本个数的比值，有利于增强模型的稳定性和泛化能力，进而提升视觉信息解码的准确率。实验结果表明，DGMM 成功地根据大脑响应重建出了视觉刺激内容，并且重建结果显著优于已有的同类型研究。该研究成果受邀在 Human Brain Mapping 2017 年会上作了口头报告，并被 MIT Technology Review 头条报道。未来我们将进行自然场景和动态视觉刺激的解码，相信该研究会进一步促进脑-机接口的发展。

（责任编辑：杨巨峰）



何晖光

CCF 计算机视觉专委会委员，中国科学院自动化研究所研究员。主要研究方向为模式识别和神经影像计算与分析。

Email: huiguang.he@ia.ac.cn



杜长德

中国科学院自动化研究所博士生。主要研究方向为机器学习和类脑智能。

Email: duchangde@gmail.com

基于多层次或多尺度卷积神经网络的显著性检测方法

哈尔滨工业大学 邬向前 唐有宝

对于人眼视觉系统而言,显著性检测是指人眼视觉注意机制对场景中具有视觉刺激的目标进行选择注意的能力。这种机制能够过滤掉场景中我们不感兴趣的信息,使得整个视觉系统集中处理感兴趣信息,从而减少能量的消耗和相应资源的浪费。对于计算机视觉而言,显著性检测则指设计相关算法并利用计算机模拟人眼视觉注意机制,定位图像中感兴趣目标并计算其对应的显著性值。

目前已有的显著性检测方法可以大致分为两类:第一类是基于手工设计的特征,该类方法首先从图像中的像素或区域中提取局部或全局的手工特征(比如颜色、纹理等),然后利用机器学习算法进行显著性计算。第二类是基于卷积神经网络学习的特征,该类方法主要通过设计更好的网络结构来学习更具区分能力的特征进行显著性预测。近年来基于卷积神经网络的显著性检测方法取得了较好的检测结果。

为了同时利用图像中像素和区域的信息,我们提出了一种基于多层次(像素级和区域级)卷

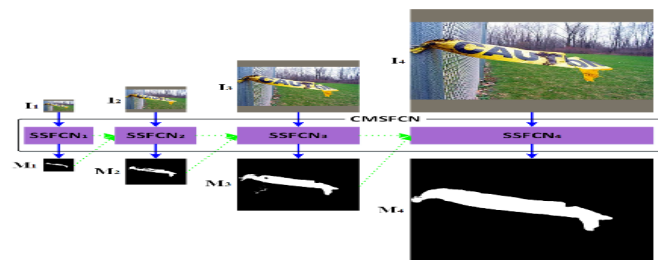


图2 基于多尺度卷积神经网络的显著性检测

积神经网络的显著性检测方法,如图1所示。该方法首先使用全卷积神经网络得到像素级显著性检测结果,其次使用自适应区域生成技术和另一个卷积神经网络得到区域级显著性检测结果,为了充分利用这两个层级显著性图的互补信息,最后使用一个卷积神经网络对其进行融合得到最终的检测结果。该方法虽然很有效,但由于其涉及多个网络结构且不能进行端到端的学习,从而在一定程度上影响了该方法的效率和性能。为此,我们提出了另一种基于链式多尺度卷积神经网络(CMSFCN)的显著性检测方法来进一步提高检测的效率和性能,如图2所示。该网络结构包含了多个单尺度的全卷积神经网络(SSFCN),当前SSFCN将利用前一SSFCN中的检测结果和中间层特征来学习更具分区能力的特征用于显著性预测,从而随着SSFCN的增加,所检测的显著性结果越来越精确。与此同时,由于为全卷积神经网络结构,该方法具有较快的检测速度。该研究的相关成果于发表于国际会议ECCV 2016和ACM MM 2017。

(责任编辑:王金甲)

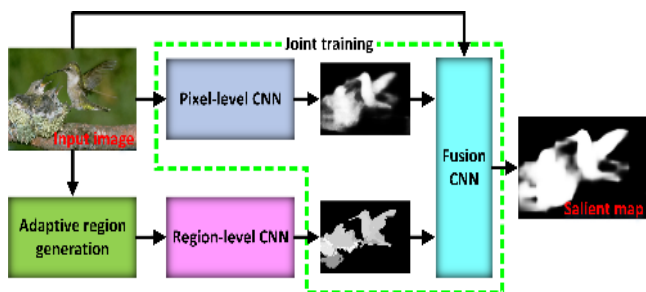


图1 基于多层次卷积神经网络的显著性检测



邬向前

CCF 计算机视觉专委会委员, 哈尔滨工业大学计算机科学与技术学院教授, 博士

生导师。曾获得黑龙江省自然科学一等奖等, 连续三年(2014-2016)入选爱思维尔“中国高被引学者榜单”。

Email: xqwu@hit.edu.cn

西北工业大学韩军伟教授访谈

2018年3月1日,专委秘书处采访了西北工业大学自动化学院副院长、优秀青年学者韩军伟教授。下面是采访实录。

您在计算机视觉与模式识别等领域做出了很多高水平研究工作,获得了国家自然科学基金优秀青年基金支持,入选了教育部新世纪优秀人才支持计划和陕西省青年科技新星计划,是青年科技工作者学习的楷模。您能否向我们介绍一下您现在主要的研究方向是什么?是如何选择现在的研究方向的?以及怎么沿着当前的研究方向慢慢做大的?

我目前的主要研究方向是仿脑视觉信息智能处理,通过磁共振功能成像手段观测大脑的功能活动,从而理解大脑处理视觉信息的一些工作规律,接下来设计算法来模拟大脑处理视觉信息的机制,最后应用于视觉信息理解的一些实际任务中。我这个研究方向需要把脑科学和信息科学进行交叉研究,也需要同时具备两个学科的一些基本知识和实验技能。

其实在2010年前,我主要还是做传统的计算机视觉,然而传统计算机视觉的方法大多并不考虑人脑的处理机制,因此并不是特别智能,存在语义鸿沟等难以解决的难点问题。2010年回国后,我有机会跟我一位师兄合作,他在美国大学做脑成像分析方面的研究,我们在一次学术讨论中,觉得将脑成像和计算机视觉结合起来也许是一个好的研究方向。大脑在观看视频时候的功能活动蕴含了对于视频内容语义层的理解,如果能加以利用,应该能在一定程度上缓解语义鸿沟问题。

讨论出来这个想法后,我们两人都很兴奋,就立即组织一支研究小组进行这方面的研究,很快我们采集了数据,进行了实验,实验结果也符合预期,我们把实验结果整理成论文,投向了多

媒体领域的顶级会议ACM Multimedia,三个审稿人一致赞同我们工作的创新性,论文以长文的形式发表了,并且还获得了最佳学生论文奖提名。

这篇顶级论文的发表,给我们很大的肯定和鼓励,我们研究组就沿着这个方向持续进行研究,并且还扩展了合作伙伴,例如跟美国耶鲁大学,澳大利亚昆士兰医学院进行合作研究。

您在国际顶级期刊和会议上发表了多篇论文并被多次引用,在您取得的这些成果中,有哪些工作让您印象特别深刻?其中有哪些不为人知的艰辛过程或者有趣经历?

让我印象特别深刻的工作就是我前面提到的2010年我回国后打算在脑科学与计算机视觉交叉这个新的研究方向上开展研究。我们当时都觉得这个工作如果能做出来,将会对多媒体领域产生较大的影响,所以我们当时把目标定在了把这个工作发表在多媒体领域的顶级会议ACM Multimedia上。

当时从我们有这个想法到投稿截止只有两个多月时间,时间紧,任务重,我们立即组织了20多人的研究队伍,一部分人在美国设计实验,采集脑成像数据,另一部分人设计算法,根据实验结果不断对算法进行调整和提升。终于在论文提交前两天,我们完成了所有实验,并且取得了很好的实验结果,我们立即投入到论文写作中,同样是中美两边的团队分工协作,最后的48小时,主要的写作人员基本上没有休息,最终赶在deadline之前2个小时完成了论文。这个工作得到了审稿人一致的高度评价,以长文的形式发表在ACM Multimedia上,并最终获得了最佳学生论文奖提名。

高效的科研团队对于深入开展课题研究有着至关重要的作用。您作为西北工业大学信息融

合技术教育部重点实验室副主任，您能否简要介绍一下您的科研团队？以及您是如何高效管理科研团队的？

我回国后就着手组建自己的科研团队，经过几年的努力，目前团队已经有教授 1 人，副教授 3 人，助理教授 4 人，研究生 20 余人。

我们团队比较重视学术交流，我们经常会邀请国内外研究者到团队来进行学术报告，团队的研究成员每年都会参加领域里的一些顶级国际会议如 CVPR、AAAI 等。团队内部日常的学术交流也比较规范，每两周举办一次 reading group，团队成员会轮流负责为大家讲述最近发表在顶级会议或者期刊上的工作。

科研团队根据研究方向又细分成三个小组，每个小组每个月都会至少开一次学术讨论会议，小组成员轮流介绍研究进展及遇到的问题，大家讨论解决方案。我对于博士生，要求每十天要书面介绍研究进展。除了会议，我们也经常通过高效的通讯方式如：QQ 和微信进行交流讨论。

作为博士生导师，您在招收研究生时最注重学生的哪些方面？能否跟大家分享一下您在指导学生方面的心得体会？在您的实验室是否有某种奖罚制度用以激励学生？

我比较看重学生两个方面的情况，第一是是否勤奋刻苦，其实我发现能到我们学校和团队来的研究生智商和基础都没有太大差别，勤奋的学生一定会收获更多的成果。第二是是否自身有很高的目标和追求，自身的目标和追求高，他努力的动力就足，而有些学生抱着混个学历的态度工作，就不会很努力，导师督促也不一定会有作用，所以这一类的学生我就不愿意招收。

在指导学生方面，我个人的一个心得是还需要因材施教，每个学生的性格、追求、目标、优缺点都还是会有不同，导师还是要多和学生交流，为不同类型的学生定制出来适合他们的培养方案。在我的实验室，对于发表高水平论文的学生，我会发些奖金作为激励。

科研基金是开展科研工作的前提，您认为书写基金申请书时应注意哪些问题？能否分享一下您个人申请科研基金的成功经验或者失败原因？

我觉得基金申请跟写论文有很多相似的地方，往往能够发表高水平论文的学者基金也会写的很好。我个人理解，书写基金最重要的一点还是要突出创新性，能够把你的研究和其他人研究的差别和优势准确的表述出来。基金申请书跟论文写作的区别在于基金往往是要对一个研究方向持续研究，研究内容和创新点会更多，往往要包含多个研究工作。而通常情况下，一篇论文基本上对应一个研究工作，所以基金书写还要注意描述清楚这些研究工作之间的关系。

您是多个期刊和会议的编委或审稿人，您能否结合自己的审稿经历，分享一下学术论文写作投稿过程中需要注意哪些方面的问题？或者有哪些建议？

我对于论文写作方面的建议主要是针对初学者。要写好论文，首先要多看论文，尤其是多看领域顶级会议或者期刊的论文，现在的研究生，英语都很好，语法上面的问题不会太多，但是对于科技论文来说，语言只是一个方面，论文的结构，描述问题的逻辑这些更加重要。

我通常建议我的研究生，在动手写第一篇论文的时候，首先要先看 100 篇顶级的论文，要仔细揣摩高水平论文怎么来组织逻辑，怎么把要解决的问题和自己的创新地方表述清楚。写论文的时候要更多从读者的角度出发，不能光想着自己明白就好了，还要换个角度多考虑这样写，读者是否能够看明白。

您认为计算机视觉与模式识别领域未来比较有潜力的研究方向主要有哪些？您认为未来哪些技术的在社会上的应用需求会比较大？

我觉得未来几年里，深度学习肯定还是计算机视觉和模式识别领域的主流研究方向，未来深

度学习的发展方向可能集中在深度学习的理论解释、脑工作机理启发的深度学习方法等这些方面。

另外，我对于弱监督 / 无监督学习比较看好。目前很多具体的任务都缺少训练样本，更缺少标注很好的训练样本，而且很多情况下训练样本标注是高噪声的。对于这些实际应用，弱监督 / 无监督学习是很好的解决方案，但是这个方向的研究理论上存在很多难点和挑战，如果有所突破，会有很广泛的应用场景。

您对现在深度学习技术发展的看法是什么？您认为优势和劣势都有哪些？未来的发展趋势是什么？

毋庸置疑，深度学习技术在过去 5 年是整个计算机视觉和模式识别最主流的技术，应用深度学习技术的论文数量也是这几年在领域各大顶级会议和期刊上最多的。它的优势劣势领域里面都争论多年了，我相信大家都比较清楚了。

我个人认为未来几年，深度学习还将是领域研究的主要方向，未来发展方向可能集中在深度学习的理论解释、脑工作机理启发的深度学习方法、提高深度学习训练效率、深度学习与其他学习方法结合、深度学习应用于其他领域如医学、机械等这些方面。

您可否可以大家分享一下您的业余生活？是否经常进行体育锻炼？以及您是如何协调工作和家庭的呢？

我的兴趣爱好比较少，业余时间会打打篮球，如果不遇到出差，基本上每周会打一次球。这些年来，很感谢家人对我工作的理解和支持，让我能够把更多的时间用于工作上。



韩军伟

西北工业大学教授，自动化学院副院长，信息融合技术教育部重点实验室副主任，陕西省科技创新团队负责人，欧盟玛丽居里国际人才引进基金获得者。入选了国家优青、教育部新世纪优秀人才、陕西省中青年科技领军人才、陕西省青年科技新星等人才计划。主要研究方向是多媒体信息处理、人工智能和生物医学图像处理。已在领域顶级期刊如：Proceedings of the IEEE, IEEE TPAMI, IJCV, NeuroImage, Cerebral Cortex 等发表学术论文 50 余篇，在领域顶级的国际会议如：CVPR, ICCV, MICCAI, IPMI, IJCAI 等发表学术论文 30 余篇。论文被引用 4000 余次，ESI 高被引论文 14 篇，ESI 热点论文 7 篇。获得国际会议 ACM Multimedia 2010, MICCAI 2011 和 ICME 2016 最佳学生论文奖提名。获得教育部自然科学二等奖等 3 个省部级科技奖励，担任国际期刊 IEEE Trans. on Neural Networks and Learning Systems、IEEE Trans. on Circuits and Systems for Video Technology 等 7 个国际期刊编委，担任领域权威国际会议如：ICPR, ACCV 等的区域主席。

（责任编辑：黄岩 余烨）

委员好消息

◇ 2018年1月3日,64位CCF专业会员晋升为高级会员,其中CCF-CV专委会委员5人:中科院信工所葛仕明副研究员、江南大学李朝锋教授、兰州工业学院李向伟教授、中科院心理所王甦菁助理研究员、西安理工大学魏嵬副教授。39位CCF连续会龄在5年以上、在学科教学、学术研究、人才培养、企业管理、项目开发等方面做出突出成就,并为CCF发展做出重要贡献的CCF高级会员当选,其中CCF-CV专委会委员4人:合肥寰景信息技术有限公司董事长陈拥权、中兴通讯云计算及IT研究院副院长董振江、中国科学院自动化研究所黄凯奇研究员、中南大学邹北骥教授。

◇ 2018年1月8日,国务院发布了《关于2017年度国家科学技术奖励的决定》,授予王泽山院士、侯云德院士国家最高科学技术奖,授国家自然科学基金颁奖项目35项,国家技术发明奖颁奖项目66项,国家科学技术进步奖颁奖项目170项,授予7名外籍专家中华人民共和国国际科学技术合作奖。CCF-CV专委会委员、西安电子科技大学董伟生教授参加的《编码混叠成像与计算重建理论方法研究》项目获国家自然科学基金二等奖,CCF-CV专委会委员、北京理工大学刘越教授参加的《交互式显示关键技术及应用》项目获国家技术发明二等奖。

◇ 2018年1月10日,2017年度CCF会士评选会在学会总部举行,共10人当选,其中CCF-CV专委会委员3人:中国科学院大学黄庆明教授、北京航空航天大学王蕴红教授和北京交通大学于剑教授。

◇ 2018年1月19日,Elsevier发布了2017年中国高被引学者榜单,本次国内共有1793位学者入选。CCF-CV专委会有14位委员上榜:华中科技大学白翔教授、浙江工业大学陈胜勇教授、南京航空航天大学陈松灿教授和谭晓阳教授、南开

大学程明明教授、西安电子科技大学高新波教授、清华大学郭振华副研究员、中科院自动化所谭铁牛研究员、刘成林研究员和王亮研究员、深圳大学沈琳琳教授、同济大学王瀚漓教授、哈尔滨工业大学邬向前教授、哈尔滨工业大学深圳研究生院徐勇教授。

◇ 2018年1月30日,由DeepTech深科技主办的2017年《麻省理工科技评论》中国区“35岁以下科技创新35人”颁奖典礼于北京举办,这是“MIT Technology Review Innovators Under 35”榜单首次落地中国。CCF-CV专委会委员、中国科学技术大学查正军教授因在解决多媒体分析检索中的“语义鸿沟”和“意图鸿沟”难题方面取得了一系列创新成果而获此殊荣。

◇ 2018年2月1日,CCF公布了2017年CCF杰出演讲者名单,共52人。CCF-CV专委会副主任、中科院计算所陈熙霖研究员、CCF-CV专委会委员、哈尔滨工业大学邬向前教授位列其中。

◇ 2018年2月3日,CCF颁奖大会在京举行,计算机视觉专委会荣获CCF 2017年度特别奖。计算机视觉专委会2013年10月成立,荣获CCF 2014年度特色活动奖、2015年度综合进步奖、2016年度优秀专委会奖。

◇ 2018年3月1日获悉,CCF-CV专委会常委、华中科技大学白翔教授的团队近日提出了一种基于角定位与区域分割多向文本检测方法。在计算机视觉的应用场景里,对图像中的文本进行准确识别是重要而相对困难的任務,该方法大幅提高了机器学习的识别准确度。该研究论文 *Multi-Oriented Scene Text Detection via Corner Localization and Region Segmentation* 已被即将于6月18日在美国盐湖城举行的CVPR 2018大会接收。

(责任编辑:刘海波)

计算机视觉中的行人检测数据集

东北大学 贾同 袁希

行人检测是计算机视觉领域内应用比较广泛和比较热门的算法，常与行人跟踪，行人重识别等技术进行结合，实现对区域内的行人进行检测识别，在安全防护等领域体现出极其重要的作用。行人检测的挑战性体现在多个方面，行人的穿着、姿态，道路上的遮挡物，摄像头的视角等，都影响着检测的准确性。

本文针对行人检测的常用数据集进行总结。

1 MIT-CBCL Pedestrian Database

MIT 行人数据集为较早公开的行人数据集，共 924 张行人图片，64x28 PPM 格式的图像，肩到脚的距离为 80 像素。数据集中只包含行人正面和背面两个视角，无负样本，未区分训练集和测试集。Dala1 等采用“HOG+SVM”，在该数据集上的检测准确率接近 100%。



图 1 MIT 行人检测数据集样例

2 USC Pedestrian Detection Test Set

USC 行人数据集包含三组数据集，分别为 USC-A、USC-B 和 USC-C，均以 XML 格式提供标注

信息。USC-A 取自互联网，以评估人体无遮挡的正面、背面、行走或站立的人体检测算法，该组包含 205 张图片，313 人；USC-B 取自 CAVIAR 视频语料库，以评估人体存在部分遮挡时的正视、后视、行走或站立的人体检测算法，该组包含 54 张图片，271 人；USC-C 取自互联网，以评估行人间无遮挡的多视图行走或站立的人体检测算法，该组包含 100 张图片，232 人。



图 2 USC 行人检测数据集样例

3 Caltech Pedestrian Detection Benchmark

Caltech 行人数据集是目前规模较大的行人数据集，取自城市环境中正常交通行驶的车辆拍摄的视频，包含约 10 小时的 640x480 30Hz 视频，在 137 分钟左右的时长中，约 250000 帧，共计 350000 个边界框和 2300 个独特的行人被注释，注释包括边界框和详细的遮挡标签之间的时间对应关系。

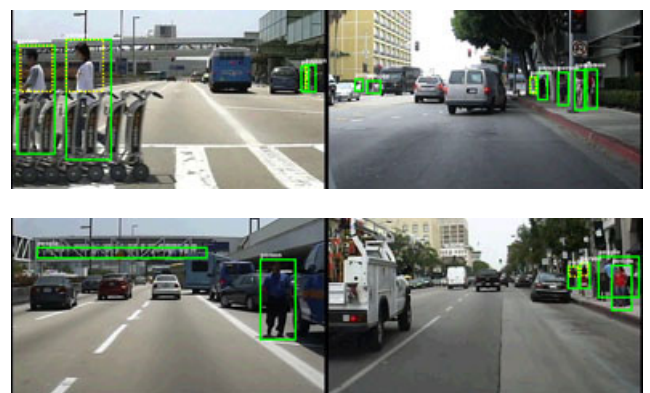


图 3 Caltech 行人检测数据集样例

4 Daimler Pedestrian Detection Benchmark

Daimler 行人数据集包括从城市环境中的车载校准立体摄像机捕获的手动轮廓标记的行人

人图像，分为检测和分类两个数据集。检测数据集包含 18x36 和 48x96 大小的正样本图片各 15560 张，行人最小高度为 72 像素，640x480 或 360x288 大小的负样本 6744 张；测试集为一段 640x480 分辨率时长约 27 分钟的视频，共 21790 张图片，56492 个行人。分类数据集有三个训练集和两个测试集，每个数据集有 4800 张行人图片，5000 张非行人图片，大小均为 18×36，另外还有 3 个辅助的非行人图片集，各 1200 张图片。

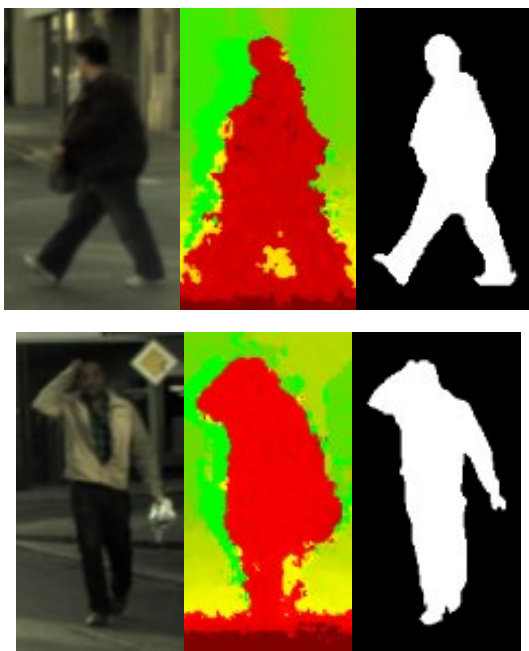


图 4 Daimler 行人检测数据集样例

5 INRIA Person Dataset

INRIA 数据集是目前使用最多的静态行人检测数据集，提供原始图片及相应的标注文件。训练集有正样本 614 张，2416 个行人，负样本 1218 张；测试集有正样本 288 张，1126 个行人，负样本 453 张。

图片中人体大部分为站立姿势且高度大于 100 个像素，部分标注可能不正确。图片主要来源于 GRAZ-01、个人照片及 google，图片的清晰度较高。



图 5 INRIA 行人检测数据集样例

（责任编辑：沈沛意）



贾同

教授，博士，东北大学信息科学与工程学院机器学习与统计研究所。主要研究领域为三维视觉感知、视觉测量、机器学习，医学影像计算等。



袁希

东北大学硕士研究生，主要研究方向为基于深度学习的目标识别、视觉测量、图像处理和分析等。

除雾方法及源码整理

大连理工大学 樊鑫 侯岷君

1 雾图成像模型

在计算机视觉中,下述关于雾图成像的大气散射模型被广泛使用: $I(x) = J(x)t(x) + A[1 - t(x)]$ 。

通过这个公式,除雾的目标是从已知雾图 $I(x)$ 中恢复原始无雾图像 $J(x)$ 。其中,全球大气光参数 A 和透射率 $t(x)$ 这两个重要的参数均为未知。由于在实际情况中很难获得拍摄图像时的具体相关参数,因此需要对透射率 $t(x)$ 和大气光参数 A 进行估计以求出定解。这也意味着参数估计越准确,除雾性能越好,而其核心是如何更精确地估计透射率 $t(x)$ 。

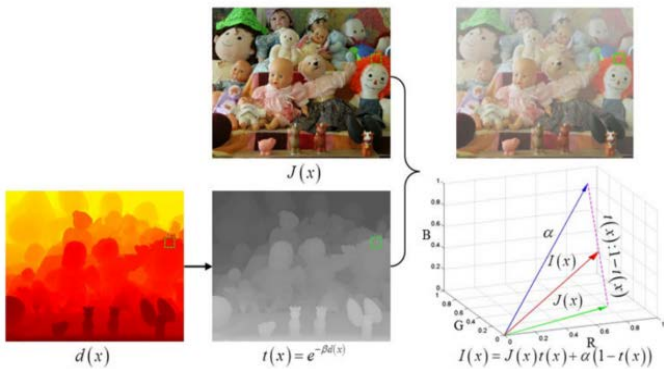


图1 大气散射模型,雾图 $I(x)$ 由真实场景 $J(x)$, 透射率 $t(x)$ 和全局大气光共同生成,且 $t(x)$ 随景深 $d(x)$ 增大而减小。

2 基于先验的除雾方法

基于先验的方法一般是在一些先验信息的基础上对透射率 t 和大气光 A 进行估计,进而求出定解,其关键在于先验的设计。

2.1 暗通道先验(DCP)除雾

Single Image Haze Removal Using Dark Channel Prior (CVPR 2009)

作者在对大量清晰的室外图像进行研究统计后发现,在大多数非天空的局部区域中,某些像素总会有至少一个颜色通道具有很低的值,并称其为暗通道。而在雾图中,暗通道可较好地反映雾度的分布,据此作者首先提出了暗通道先验(DCP)理论,通过用滑动窗口对雾图进行最小值

滤波求得图像的暗通道,进而估计透射率。

由于在计算过程中使用了最小值滤波,暗通道图像会具有块状伪影并影响到最终的恢复图像,为此DCP算法使用Soft Matting操作来优化透射率。通过DCP理论也可获得大气光参数 A : 首先在暗通道中选择具有前0.1%最亮像素的局部区域,然后在原始雾图的该区域中选择具有最高像素值的点,并以此像素值作为大气光 A 。估计出相应参数 $t(x)$ 和 A 后,将其代入到成像模型即可求出恢复图像。



图2 清晰图像和有雾图像的暗通道。

- Paper: <http://www.robots.ox.ac.uk/~vgg/rg/papers/hazeremoval.pdf>
- Code: <https://github.com/sjtrny/Dark-Channel-Haze-Removal>

2.2 改进的DCP除雾算法

Efficient Image Dehazing with Boundary Constraint and Contextual Regularization (ICCV 2013)

作者提出一种改进的DCP除雾算法,通过对透射率公式施加固有的边界约束求得包含雾度分布信息的图像特征,之后使用加权的基于L1范数的上下文正则化来优化透射率,并提出了一种变分算法解决优化问题。而大气光参数 A 可通过用移动窗口对原始雾图的每个通道执行最小滤波和最大值滤波后得到。

在这个算法中,恢复图像的像素值总是有界

的，以此改进原始 DCP 算法在天空区域的失效问题，然而迭代次数和两个边界约束不能自适应调整。

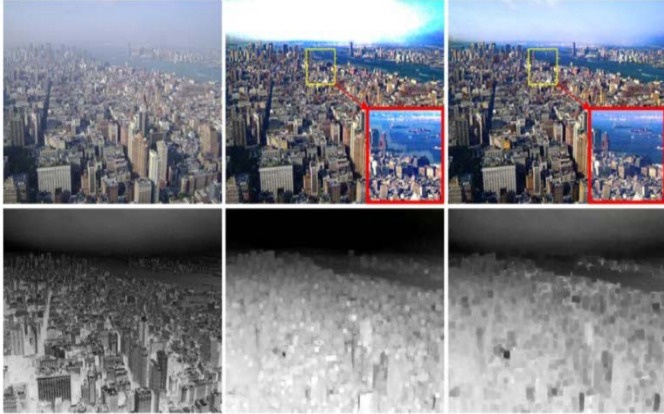


图 3 从左到右：（顶部）有雾图像，原始 DCP 去雾结果，改进后的 DCP 去雾结果；（底部）边界约束映射，原始 DCP 透射率结果和改进后的 DCP 透射率结果。

- Paper: <http://www.escience.cn/system/file?fileId=65063>
- Code: <http://www.escience.cn/system/file?fileId=65069>

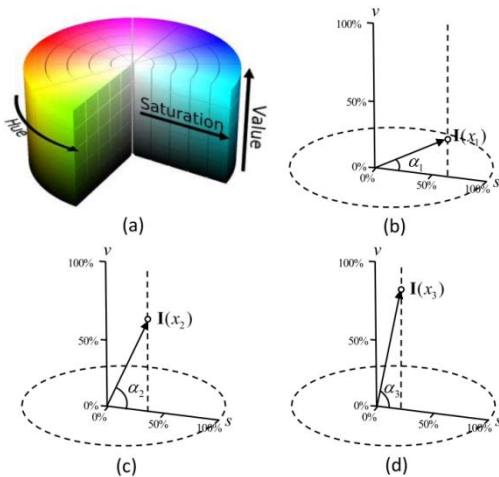


图 4 CAP 几何描述：(a)HSV 颜色模型 (b)近距景深 (c)中距景深 (d) 远距景深。

2.3 颜色衰减先验(CAP)除雾

A Fast Single Image Haze Removal Algorithm Using Color Attenuation Prior (TIP 2015)

这篇文章提出了一个颜色衰减先验(Color Attenuation Prior)，基于此先验，作者创建一个线性模型来对雾图的景深建模，并通过监督学

习方法学习模型的参数，以此恢复图像的深度信息，进而恢复透射率及清晰图像，并使用导向滤波对透射率图进行优化。训练样本由合成雾图及其相应的真实深度图组成。但此算法在一些特定情况会失效（如天空和白色物体区域）。

- Paper: <http://www.wineyard.in/Abstract/mtech/DIP/2015/bp/15D008.pdf>
- Code: <https://github.com/JiamingMai/Color-Attenuation-Prior-Dehazing>

2.4 Haze-line 先验除雾

Non-Local Image Dehazing (CVPR 2016)

一幅无雾图像可以用几百个不同的颜色近似表示，近似 RGB 值在 RGB 空间分布为一个聚类，但是在有雾图像中这些 RGB 值的对应分布为一条线，作者称之为 haze-line，如下图所示：

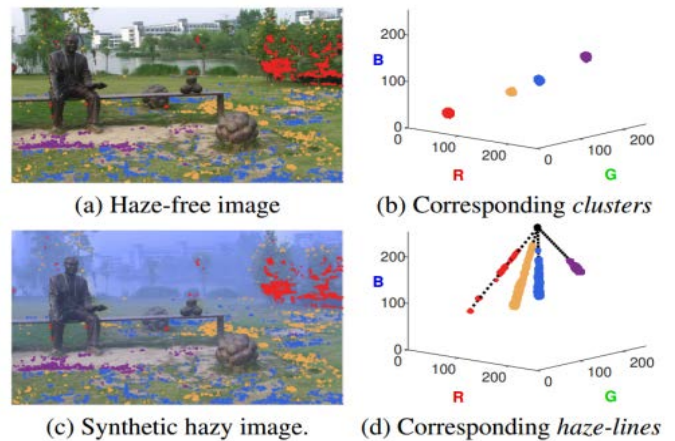


图 5Haze-line 模型 (a) 使用 K-means 对无雾彩色图像的像素进行聚类，属于四个群集的像素被标记。(b) RGB 空间中的相应四个颜色集合。(c) 合成雾图中相同的聚类像素被标记，但他们观测到的颜色受到不同雾度的影响。(d) RGB 空间中雾图像素，haze-line 标记为黑色。

每一个聚类所包含的具有相近 RGB 值的像素点分布在整个图像区域内(即 Non-Local 思想)，但是与摄像机间的距离远近不同(即景深不同)，因此同一聚类里不同的像素点具有不同透射率值，其对应到 haze-line 上即可反映深度信息。haze-line 的一端坐标值为图像中无雾区域(即距相机较远处)的聚类点的坐标，而所有 haze-line 相交于一点，该点即为雾图中距相机最远处，因此该点坐标即为环境光 A。作者通过用球坐标系表示雾图，将所有像素与 haze-line 建立联系

来估计初始透射率，在经过正则化优化过程后即可得到最终透射率并进行除雾操作。

- Paper: http://www.eng.tau.ac.il/~berman/NonLocalDehazing/NonLocalDehazing_CVPR2016.pdf
- Code: <https://github.com/danaberman/non-local-dehazing>

2.5 雾度特征的探究及学习框架

Investigating Haze-relevant Features in A Learning Framework for Image Dehazing (CVPR 2014)

该文系统地研究了学习框架中不同的雾度相关特征（多尺度暗通道、多尺度局部最大对比度、色偏、局部最大饱和度等），以确定图像去雾的最佳特征组合。基于上述雾度特征，作者使用随机森林学习一个回归模型来估计透射率。随机森林的输入是从训练数据（合成的雾图图像块）中提取的与相关特征，输出是对应的透射率。为打破相关特征与图像内容之间的相关性，作者在训练之前对特征值进行排序以确保学习到的模型不会针对特定图像内容（避免图像纹理的干扰）。该学习框架可以自动发现不同特征之间的高阶关系，以更精确地估计雾度，同时实验证实了暗通道特征为可用于雾度估计中的最有效的特征。同时，由于在一些图像中“均匀大气光”假设不成立，作者根据输入图像的亮度对大气光 A 进行了自适应调整。

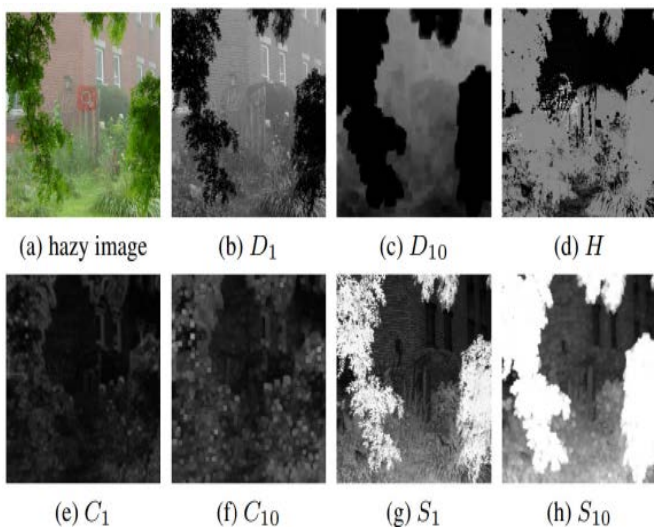


图 6 雾度特征 (a)雾图 (b,c)不同尺度的暗通道 (d)色差特征 (e,f)不同尺度局部最大对比度 (g,h)不同尺度

局部最大饱和度

- Paper: https://www.cv-foundation.org/openaccess/content_cvpr_2014/papers/Tan_g_Investigating_Haze-relevant_Features_2014_CVPR_paper.pdf
- Code: https://github.com/zlinker/haze_2014

2.6 最小化梯度残差消除伪影

Robust Image and Video Dehazing with Visual Artifact Suppression via Gradient Residual Minimization (ECCV 2016)

作者提出了一种新方法以抑制在图像和视频去雾中不同类型的伪影，该方法在雾度估计和图像恢复中均做出了贡献。首先，为改进局部先验生成的初始透射率，作者提出了一种图像引导的深度边缘感知平滑算法，将透射率的细化看作一个针对数据保真项和正则化项的全局优化问题，并采用带导向图的二阶广义变分 (TGV) 用于正则化以使透射率随着深度的变化平稳地变化。

在图像恢复过程中，作者提出梯度残差最小化 (GRM)，引入梯度残差和误差层，使该方法能够在不对每种情况特别建模的情况下最小化各种伪影，同时联合地恢复无雾图像。然而由于需要较多的迭代，此算法需要耗费较多的运行时间，并且在图像恢复的过程中，在消除伪影的同时可能会由于过度平滑而使恢复结果丢失一些图像细节。

- Paper: <http://www.juew.org/publication/eccv16-dehaze.pdf>
- Code: <http://web.engr.illinois.edu/~cchen156/code/robustdehaze.zip>

2.7 双层高斯过程回归去雾

Two-layer Gaussian Process Regression with Example Selection for Image Dehazing (TCSVT 2016)

近年来，基于实例的方法通常缺乏物理成像模型且需要人工合成雾图，而从观测图像中得到或由合成图像训练得的先验信息不能总是反映自然图像的真实结构信息。为解决这些问题，作者提出了基于双层高斯过程回归 (GPR) 的学习模型，该模型形成成对的马尔科夫随机场 (MRF)，

并且能够在保留局部细节的同时实现更准确的透射率估计，从而减少明显的伪影。通过使用训练实例，双层 GPR 建立了从输入的图像特征到透射率的非线性映射，并且学习局部的图像先验以进一步改进透射率估计。在 Gaussian 回归过程中作者使用超像素来组合区域间纹理相似性、区域间亮度相似度和曲线连续性等七种特征向量。使用这些中级特征而不是图像块有利于较小的训练实例和较简单的回归器，同时可提供结构信息。

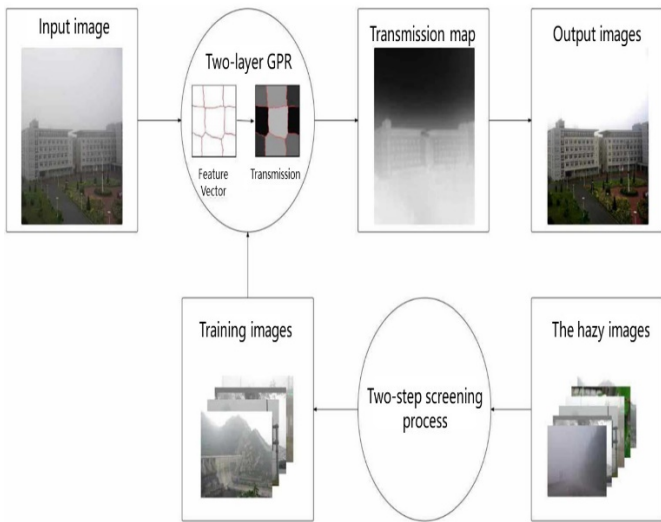


图 7 算法流程

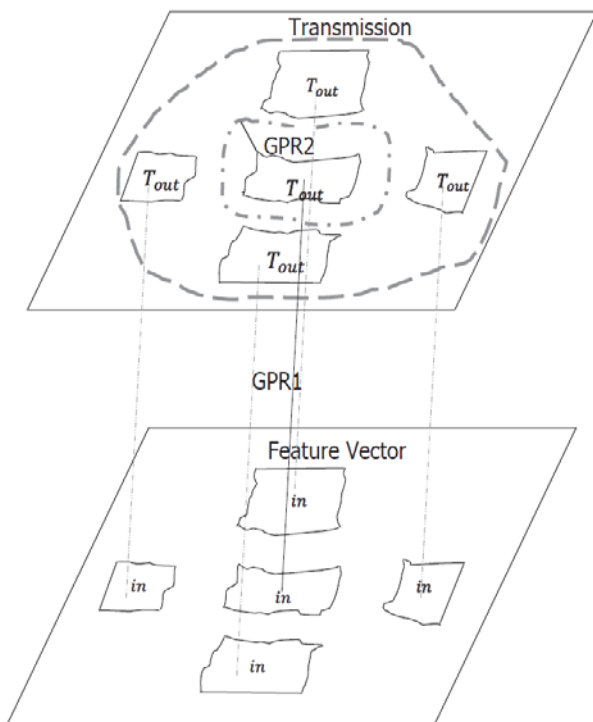


图 8 双层 GPR 模型

此外，作者还提供了一个系统的方案以合成或真实世界模糊图像中为双层模型选择合适的训练实例。在传统的基于先验的方法可能失效的白色或明亮物体区域，该方法体现出良好的性能。

- Paper: https://www.researchgate.net/publication/305417643_Two-Layer_Gaussian_Process_Regression_with_Example_Selection_for_Image_Deblurring
- Code: <https://github.com/dlut-dimt/Two-Layer-GPR-Deblurring>

3 卷积神经网络除雾

基于先验的图像去雾方法往往会在某些特定情况下失效，为避免手工设计的先验的局限性，鉴于深度卷积神经网络 (CNN) 在计算机视觉领域的迅猛发展及其良好的学习性能，越来越多的研究者将 CNN 应用到图像除雾中。

3.1 多尺度卷积神经网络除雾

Single Image Dehazing via Multi-Scale Convolutional Neural Networks (ECCV 2016)

这篇文章通过学习雾图与其相应的透射率之间的映射，提出了一种用于单幅图像去雾的多尺度深度神经网络，该网络由基于整幅图像预测。

大致透射率结构的粗尺度网络和局部细化结果的细尺度网络组成，网络结构如下图所示：

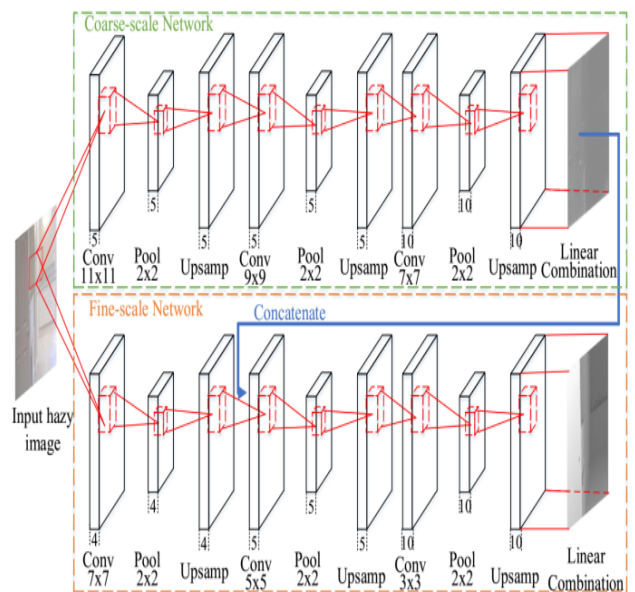


图 9 Multi-Scale CNN

其中两种尺度的网络都可应用于原始输入雾图，但粗尺度网络的输出作为附加信息传递给细尺度网络以便其可以细化粗略的预测。网络的训练数据由合成雾图和相应准确透射率组成。

- Paper: <http://visionlab.hanyang.ac.kr/wordpress/wp-content/uploads/2017/09/dehaze-net.pdf>
- Code: https://sites.google.com/site/renwenqi888/research/dehazing/mscnndehazing/MSCNN_dehazing.zip?attredirects=0&d=1

3.2 端到端除雾网络

DehazeNet: An End-to-End System for Single Image Haze Removal (TIP 2016)

DehazeNet 是一个特殊设计的深度卷积网络，利用深度学习去智能地学习雾霾特征，解决手工特征设计的难点，网络结构如图 10 所示。

其中 DehazeNet 的特征提取采用“卷积

+Maxout”的结构作为网络第一层，其等价于传统的手工去雾特征。并采用多尺度特征提取方法，使用 3 组不同尺度的滤波器以提高不同分辨率下特征提取的鲁棒性。通过 Max Pooling 层约束透射率估计中的噪声，同时提出双边纠正线性单元 (BReLU)，在双边约束的同时保证局部的线性。网络的训练数据为 100000 对清晰图像块及相应的合成雾图图块，因此训练耗时较大，而网络的除雾性能会受到训练数据的影响，在处理较浓雾图时稍显不足。

- Paper: <http://caibolun.github.io/papers/DehazeNet.pdf>
- Code: <https://github.com/caibolun/DehazeNet>。

(责任编辑: 蹇木伟)

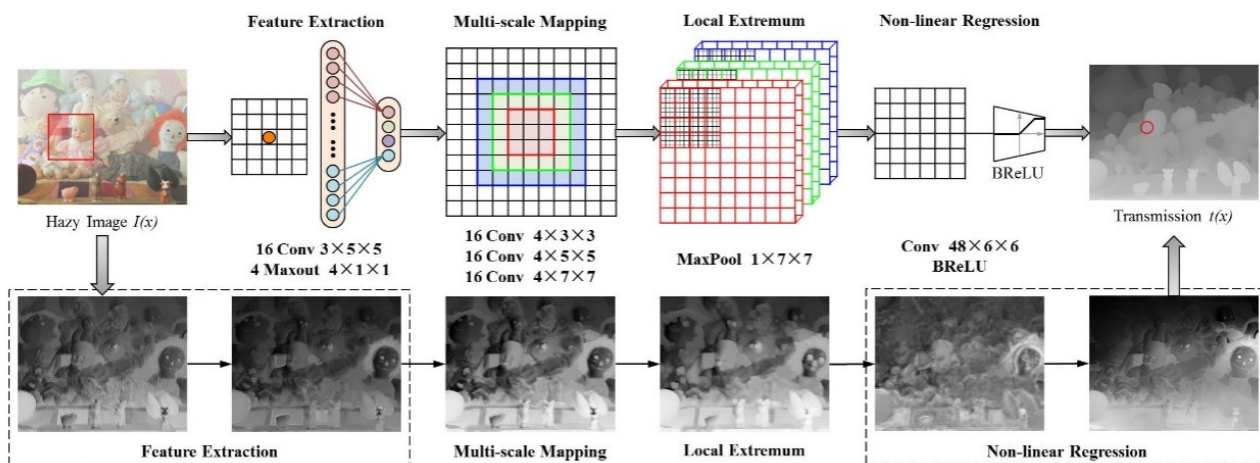


图 10 DehazeNet 网络结构

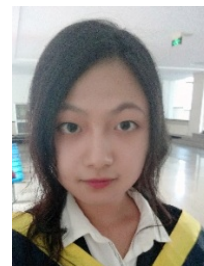


樊鑫

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

侯岷君

硕士研究生，主要研究方向为计算机视觉、图像增强等技术。



招聘信息

一、自动化所模式分析与学习团队招聘

助理员/博士后招聘信息:

1、岗位名称: 模式识别与机器学习助理研究员/博士后; 文档图像分析与识别助理研究员/博士后

岗位部门: 自动化所

需求人数: 2~4 人

面向对象: 应届毕业生、社会在职人员

工作职责: 模式识别基础理论、机器学习模型与方法研究; 文档图像分析与识别(包括: 版面分析、文本检测与分割、文本行识别、上下文融合、文档和字符结构分析等)方法与技术研究。

岗位要求:

(1) 计算机科学与技术、控制科学与工程、信息处理等专业博士毕业;

(2) 在模式识别、机器学习、图像处理、计算机视觉或人工智能其他方向取得高水平研究成果(在高档次期刊/会议上发表论文);

(3) 对基础研究具有强烈的兴趣, 具有坚实的理论基础和较强的学习能力、分析思考能力和总结表达能力;

(4) 具有较强的科研主动性、分析问题解决问题的能力、文献阅读和总结表达能力

(5) 具有良好的团队协作精神和学术交流能力, 以及一定的组织协调能力;

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。

简历投递截止时间: 招满即止。

联系人: 刘成林老师。

简历接收邮箱:

liucl@nlpr.ia.ac.cn

二、上海交通大学人工智能研究院

Track 师资招聘信息:

1、岗位名称: 长聘教轨副教授

岗位部门: 人工智能研究院

需求人数: 5 人

面向对象: 副教授

工作职责: 承担相关领域教育教学工作; 从事科学研究, 积极组织承担国家级、省部级科技项目, 取得一批具有国际先进水平的标志性科技成果, 发表高水平学术论文, 同时积极推进实验室研究成果转化; 参与相关领域学科建设, 促进多学科的交叉融合, 加强研究院的国际交流与合作; 参加研究院建设和其他社会服务性工作。

岗位要求:

(1) 研究生学历, 博士学位。具有世界一流大学博士学位并有连续 3 年以上海外科研工作经历, 年龄不超过 40 周岁;

(2) 在学术前沿领域开展创新性研究工作, 取得重要的研究成果, 表现出很强的学术潜力, 达到世界一流大学助理教授水平;

(3) 对特别优秀的海外博士, 可以突破任职年限限制。

2、岗位名称: 助理教授

岗位部门: 人工智能研究院

需求人数: 5 人

面向对象: 副教授及以上

工作职责: 承担相关领域教育教学工作; 从事科学研究, 申请各类课题, 发表高水平论文, 积极推进实验室研究成果转化; 参与相关领域学科建设, 加强研究院的国际交流与合作, 协助教授、副教授指导研究生; 参加研究院建设和其他社会服务性工作。

岗位要求:

(1) 研究生学历, 博士学位。参与相关领域学科建设, 加强研究院的国际交流与合作, 协助教授、副教授指导研究生;

(2) 在学术前沿领域开展创新性研究工作, 取得重要的研究成果, 表现出很强的学术潜力;

(3)对特别优秀的海外博士,可以突破任职年限限制。

3、联系方式

请应聘者以 E-mail 形式发送简历,并请在邮件主题上注明:姓名+应聘岗位名称。简历内容发表论文清单和代表作 5 篇全文;推荐信 2-3 封。

简历投递截止时间:招满即止。

联系人:杨老师、何老师。

简历接收邮箱:

xkyang@sjtu.edu.cn

heliming@sjtu.edu.cn

三、中科院信工所网络空间技术实验室

员工/博士后招聘信息:

1、**岗位名称:**2018 年度员工、博士后

岗位部门:网络空间技术实验室

需求人数:若干

面向对象:硕士或博士毕业

工作职责:在信息安全科技领域,展开基础理论与前沿技术研究;开发应用性技术与系统;实验室面向国家战略需求,基于云计算和大数据、数据挖掘和机器学习、知识图谱等技术解决网络与信息安全问题。

岗位要求:

(1)多媒体信息处理、机器学习、信息检索、社会计算、云计算、网络存储、网络与信息安全相关专业硕士或者博士毕业;

(2)从事知识图谱、多媒体处理、云平台相关的科研或者开发工作;

(3)博士发表过高质量学术论文者优先;

(4)硕士具有较强开发经验者优先。

2、联系方式

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

简历投递截止时间:招满即止。

联系人:工程部。

简历接收邮箱:

gcb5@iie.ac.cn

四、北卡罗来纳大学教堂山分校沈定刚教授课题组招聘博士后

博士后招聘信息:

1、**岗位名称:**2018 年博士后

岗位部门:沈定刚教授课题组

需求人数:若干

面向对象:博士

工作职责:研究方向是基于脑影像和机器学习的脑疾病诊断(包括但不限于基于多模态数据的老年痴呆症的早期诊断,婴儿正常脑发育轨线刻画和脑发育问题的早期诊断)。

课题组网页: <https://www.med.unc.edu/bric/ideagroup>

岗位要求:

(1)认知神经科学,生物医学工程,计算机,自动化,应用数学,信电等专业的博士;

(2)机器学习和深度学习方法及其临床应用研究感兴趣的科研人员;

(3)博士发表过高质量学术论文者优先;

2、联系方式

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

简历投递截止时间:招满即止。

联系人:沈定刚、张寒。

简历接收邮箱:

dgshen@med.unc.edu

hanzhang@med.unc.edu

(责任编辑:李策)

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，可继续关注每个会议举办的 workshop 或 special session。

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 IEEE Journal on Selected Topics in Signal Processing (JSTSP)、IEEE Transactions on Neural Networks and Learning Systems (TNNLS)、ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)、

3 会议简介

第 24 届国际模式识别大会 ICPR 2018 (24th International Conference on Pattern Recognition) 是模式识别领域的权威会议，今年

将首次在中国举办。ICPR 是一个国际性的学术会议，将为全球模式识别领域的研究人员及参与者提供一个交流平台，其内容涉及近期在模式识别、机器学习、计算机视觉方面的进展，以及这些技术在多个领域的应用。

本届 ICPR 会议由国际模式识别协会 (IAPR: International Association for Pattern Recognition) 主办，由中科院自动化研究所，模式识别国家重点实验室承办，将于 2018 年 8 月 20-24 日在北京市中国国家会议中心举行。涉及模式识别与机器学习、计算机视觉、语音、图像、视频、多媒体、生物信息、人机交互、文本分析与识别、生物成像等方向的研究进展。同时将举办相关的研讨会(Workshop)、教程(Tutorials)、竞赛(Contests)。

(责任编辑：金鑫)

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
CCFICAI 2018	2018.08.09-10	中国 济南	2018.03.31	http://ccf-icai2018.sdufe.edu.cn/
S+SSPR 2018	2018.08.17-19	中国 北京	2018.04.15	http://ssspr2018.buaa.edu.cn
CCBR 2018	2018.08.11-12	中国 乌鲁木齐	2018.04.15	http://ccbr2018.xju.edu.cn/
BMVC 2018	2018.09.03-06	Newcastle, UK	2018.04.30	http://bmvc2018.org/
PCM 2018	2018.09.21-22	中国 合肥	2018.05.05	http://www.pcm2018.org/
PRCV 2018	2018.11.22-25	中国 广州	2018.05.08	https://prcv-conf.org/2018/
PRICAI 2018	2018.08.27-31	中国 南京	2018.05.31	http://cse.seu.edu.cn/pricai18/

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
JSTSP	Robust Subspace Learning and Tracking: Theory, Algorithms, and Applications	2018.04.15
TNNLS	Recent Advances in Theory, Methodology and Applications of Imbalanced Learning	2018.04.30
TOMM	Cross-Media Analysis for Visual Question Answering	2018.06.30

银河水滴： 步态识别引领身份识别新未来

银河水滴是中国领先的专注于计算机视觉和视频大数据分析的人工智能企业。2016年6月由中科院自动化所孵化，联想等知名机构投资。银河水滴团队在计算机视觉领域，已有20多年的技术积累，拥有全球最先进的步态识别技术和领先的工业视觉技术，为安防、交通、制造、家居、零售、机器人等行业提供领先的人工智能解决方案。

银河水滴是全球唯一将跨视角步态识别做到精确度高达94%、并推进商业化的企业；拥有全球最大的步态数据库和全球最先进的步态识别技术。

在中央电视台《机智过人》节目中，银河水滴的步态识别技术，战胜人类“最强大脑”，成功靠步态识别“嫌疑人”、金毛犬，被图灵奖得主姚期智称赞“机智过人”，银河水滴获得中央电视台“机智先锋团队”年度大奖。

银河水滴团队，包括创始人兼CEO黄永祯博士、首席科学家谭铁牛院士、联合创始人王亮博士和孙哲南博士等，出版论著11部，发表论文超过500篇，申请专利超过100项，创办和参与了国内外多个高影响力的计算机视觉组织和会议、并担任要职，获得国家自然科学奖、国家科技进步奖、国家技术发明三项中国最高科技大奖，

多次获得国内外计算机视觉竞赛冠军。中科院自动化所100多人的顶尖人工智能团队，作为银河水滴的研发储备力量，为公司发展提供最具前沿性、创新性的科研探索。

银河水滴志在以人工智能释放人类的体力和脑力，拓展人类生存时空！



（责任编辑：王瑞平）

地平线： 嵌入式人工智能领导者

地平线作为嵌入式人工智能全球领导者，致力于提供高性能、低功耗、低成本、完整开放的嵌入式人工智能解决方案。面向智能驾驶、智慧城市和智能商业等应用场景，为多种终端设备装上人工智能“大脑”，让它们具有从感知、交互、理解到决策的智能，让人们的生活更安全、更便捷、更美好。



地平线凝聚了一支研发能力强大、工业界经验丰富的团队，共同打造万物智能。目前地平线已得包括晨兴资本、高瓴资本、红杉资本、金沙江创投、创新工场、真格基金、线性资本、DST、双湖资本、祥峰基金、青云创投、英特尔投资、嘉实基金、建投华科、理成资产、混沌投资和云晖资本的投资。2017年10月，完成了由英特尔投资领投，嘉实投资等新老投资人共同参与的总额超过一亿美元的A+轮融资。

经过两年多的研发，地平线基于BPU自主研发的中国首款全球领先的嵌入式人工智能视觉芯片已成功面世，推出了面向智能驾驶的征程（Journey）系列处理器和面向智能摄像头的旭日（Sunrise）系列处理器，并向行业客户提供“算法+芯片+云”的完整解决方案。这是完全由地平线自主研发的人工智能芯片，采用地平线的第一代BPU架构，具有全球领先的性能：可实时处理1080p@30视频，每帧中可同时对200个目标进行检测、跟踪、识别，典型功耗1.5W，每帧延时小于30ms。

征程1.0面向智能驾驶，能够同时对行人、机动车、非机动车、车道线、交通标志牌、红绿灯等多类目标进行精准的实时监测与识别，同时满足车载严苛的环境要求以及不同环境下的视觉感知需求。旭日1.0处理器则面向智能摄像头，能够在本地进行大规模人脸抓拍与识别、视频结构化处理等，可广泛用于商业、安防等多个实际应用场景。

地平线始终坚信技术和创新的价值，坚信人工智能可以让世界更美好。在砥砺前行的道路上，我们希望越来越多有梦想、敢拼搏的人加入，一起打造万物智能的世界。简历投递至dream@hobot.cc。（责任编辑：王瑞平）



地平线始终坚信技术和创新的价值，坚信人工智能可以让世界更美好。在砥砺前行的道路上，我们希望越来越多有梦想、敢拼搏的人加入，一起打造万物智能的世界。简历投递至dream@hobot.cc。（责任编辑：王瑞平）