



主办: 中国科学院空天信息创新研究院
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象图形学报

2021
11
VOL.26

ISSN1006-8961
CN11-3758/TB



电力视觉前沿技术



第26卷第11期（总第307期）
2021年11月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊
中文核心期刊

版权声明

凡向《中国图象图形学报》投稿，均视为同意在本刊网站及CNKI等全文数据库出版，所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用，所有内容，未经许可，不得转载或以其他方式使用。

Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院
主办单位 中国科学院空天信息创新研究院
中国图象图形学学会
北京应用物理与计算数学研究所

主 编 吴一戎
编辑出版 《中国图象图形学报》编辑出版委员会
通信地址 北京市海淀区北四环西路19号
邮 编 100190
电子信箱 jig@aircas.ac.cn
电 话 010-58887035
网 址 www.cjig.cn

广告发布登记号 京朝工商广登字20170218号
总 发 行 北京报刊发行局
订 购 全国各地邮局
海外发行 中国国际图书贸易集团有限公司
(邮政信箱: 北京399信箱 邮编: 100048)
印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

Title inscription: Song Jian | Monthly, Started in 1996

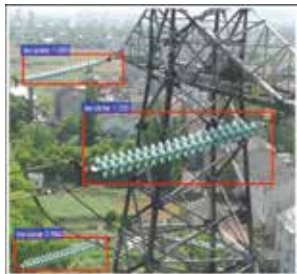
Superintended by Chinese Academy of Sciences
Sponsored by Aerospace Information Research Institute, CAS
China Society of Image and Graphics
Institute of Applied Physics and Computational Mathematics

Editor-in-Chief Wu Yirong
Editor, Publisher Editorial and Publishing Board of Journal of Image and Graphics
Address No. 19, North 4th Ring Road West, Haidian District, Beijing, P. R. China
Zip code 100190
E-mail jig@aircas.ac.cn
Telephone 010-58887035
Website www.cjig.cn

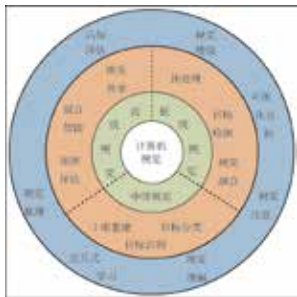
Distributed by Beijing Bureau for Distribution of Newspapers and Journals
Domestic All Local Post Offices in China
Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China)
Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB
ISSN 1006-8961
CODEN ZTTXFZ

国外发行代号 M1406
国内邮发代号 82-831
国内定价 60.00元



面向改进尺度缩放网络的绝缘子识别(第2561页)



混合增强视觉认知架构及其关键技术进展(第2619页)



激光点云的稀疏体素金字塔邻域构建与分类(第2703页)

编者按 I

电力视觉前沿技术

输电线路部件视觉缺陷检测综述

赵振兵, 蒋志钢, 李延旭, 戚银城, 翟永杰, 赵文清, 张珂 2545

面向改进尺度缩放网络的绝缘子识别

赵文清, 张海明, 徐敏夫 2561

最优知识传递宽残差网络输电线路螺栓缺陷图像分类

戚银城, 金超熊, 赵振兵, 丁洁涛, 吕斌 2571

可变形NTS-Net的螺栓属性多标签分类

张珂, 何颖宣, 赵凯, 冯晓晗, 赵振兵, 马占宇 2582

嵌入双注意力机制的Faster R-CNN航拍输电线路螺栓缺陷检测

戚银城, 武学良, 赵振兵, 史博强, 聂礼强 2594

轻量化航拍图像电力线语义分割

许刚, 李果 2605

综述

混合增强视觉认知架构及其关键技术进展

王培元, 关欣 2619

深度学习人脸特征点自动定位综述

徐亚丽, 赵俊莉, 吕智涵, 张志梅, 李劲华, 潘振宽 2630

图像处理和编码

多尺度特征复用混合注意力网络的图像重建

卢正浩, 刘丛 2645

图像分析和识别

自监督深度离散哈希图像检索

万方, 强浩鹏, 雷光波 2659

L-UNet: 轻量化云遮挡道路提取网络

许苗, 李元祥, 钟娟娟, 左宗成, 熊伟 2670

基于多尺度特征多对抗网络的雾天图像识别

陈硕, 钟汇才, 李勇周, 王师峰, 杨建刚 2680

图像理解和计算机视觉

结合动态图卷积和空间注意力的点云分类与分割

宋巍, 蔡万源, 何盛琪, 李文俊 2691

激光点云的稀疏体素金字塔邻域构建与分类

陶帅兵, 梁冲, 蒋腾平, 杨玉娇, 王永君 2703

计算机图形学

单幅植物叶片图像的3D重建

任非儿, 刘通, 杨龙 2713

医学图像处理

深度学习多部位病灶检测与分割

黎生丹, 柏正尧 2723

遥感图像处理

对抗学习遥感图像场景识别

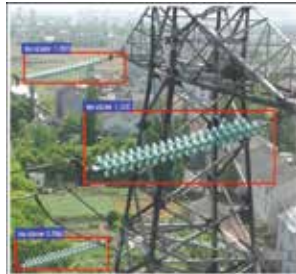
李彤, 张钧萍 2732

融合特征的卫星视频车辆单目标跟踪

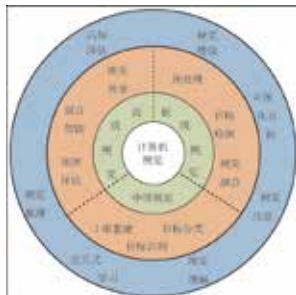
韩鸣飞, 李盛阳, 万雪, 轩诗宇, 赵子飞, 谭洪, 张万峰 2741

CONTENTS

JOURNAL OF IMAGE AND GRAPHICS



Insulator recognition based on an improved scale-transferrable network(P2561)



Hybrid enhanced visual cognition framework and its key technologies(P2619)



Sparse voxel pyramid neighborhood construction and classification of LiDAR point cloud(P2703)

Frontier Technology of Power Computer Vision

Overview of visual defect detection of transmission line components

Zhao Zhenbing, Jiang Zhigang, Li Yanxu, Qi Yincheng, Zhai Yongjie, Zhao Wenqing, Zhang Ke 2545

Insulator recognition based on an improved scale-transferrable network

Zhao Wenqing, Zhang Haiming, Xu Minfu 2561

Image classification method of transmission line bolt defects using the optimal knowledge transfer wide residual network

Qi Yincheng, Jin Chaoxiong, Zhao Zhenbing, Ding Jietao, Lyu Bin 2571

Multi-label classification method of bolt attributes based on deformable NTS-Net

Zhang Ke, He Yingxuan, Zhao Kai, Feng Xiaohan, Zhao Zhenbing, Ma Zhanyu 2582

Bolt defect detection for aerial transmission lines using Faster R-CNN with an embedded dual attention mechanism

Qi Yincheng, Wu Xueliang, Zhao Zhenbing, Shi Boqiang, Nie Liqiang 2594

Research on lightweight neural network of aerial powerline image segmentation

Xu Gang, Li Guo 2605

Review

Hybrid enhanced visual cognition framework and its key technologies

Wang Peiyuan, Guan Xin 2619

Automatic facial feature points location based on deep learning: a review

Xu Yali, Zhao Junli, Lyu Zhihan, Zhang Zhimei, Li Jinhua, Pan Zhenkuan 2630

Image Processing and Coding

Multiscale feature reuse mixed attention network for image reconstruction

Lu Zhenghao, Liu Cong 2645

Image Analysis and Recognition

Self-supervised deep discrete hashing for image retrieval

Wan Fang, Qiang Haopeng, Lei Guangbo 2659

L-UNet: lightweight network for road extraction in cloud occlusion scene

Xu Miao, Li Yuanxiang, Zhong Juanjuan, Zuo Zongcheng, Xiong Wei 2670

Haze image recognition based on multi-scale feature and multi-adversarial networks

Chen Shuo, Zhong Huicai, Li Yongzhou, Wang Shizheng, Yang Jiangang 2680

Image Understanding and Computer Vision

Dynamic graph convolution with spatial attention for point cloud classification and segmentation

Song Wei, Cai Wanyuan, He Shengqi, Li Wenjun 2691

Sparse voxel pyramid neighborhood construction and classification of LiDAR point cloud

Tao Shuaibing, Liang Chong, Jiang Tengping, Yang Yujiao, Wang Yongjun 2703

Computer Graphics

3D reconstruction of a single plant leaf image

Ren Feier, Liu Tong, Yang Long 2713

Medical Image Processing

Multiorgan lesion detection and segmentation based on deep learning

Li Shengdan, Bai Zhengyao 2723

Remote Sensing Image Processing

Remote sensing image scene recognition based on adversarial learning

Li Tong, Zhang Junping 2732

Integrating multiple features for tracking vehicles in satellite videos

Han Mingfei, Li Shengyang, Wan Xue, Xuan Shiyu, Zhao Zifei, Tan Hong, Zhang Wanfeng 2741

中图法分类号: TP3914, BB42.1 文献标识码: A 文章编号: 1006-8961(2021)11-2619-11

论文引用格式: Wang P Y and Guan X. 2021. Hybrid enhanced visual cognition framework and its key technologies. Journal of Image and Graphics, 26(11): 2619-2629 (王培元, 关欣. 2021. 混合增强视觉认知架构及其关键技术进展. 中国图象图形学报, 26(11): 2619-2629) [DOI:10.11834/jig.200446]

混合增强视觉认知架构及其关键技术进展

王培元*, 关欣

海军航空大学, 烟台 264001

摘要: 智能视觉系统虽然在大规模信息的特征检测、提取与匹配等处理上具备一定优势,但是在深层次认知上仍存在不确定性和脆弱性,尤其是针对视觉感知基础上的视觉认知任务,相关数理逻辑和图像处理方法并未实现质的突破,智能算法难以取代人类执行较为复杂的理解、推理、决策和学习等操作。为助力智能视觉感知和认知技术的进一步发展,本文总结了混合增强智能在视觉认知领域的应用现状,给出了混合增强视觉认知的基本架构,并对可纳入该架构下的应用领域及关键技术进行了综述。首先,在分析智能视觉感知内涵和基本范畴的基础上,融合人的视觉感知与心理认知,探讨混合增强视觉认知的定义、范畴及其深化过程,对不同的视觉信息处理阶段进行对比,进而在分析相关认知模型发展现状的基础上,构建混合增强视觉认知的基本框架。该架构不仅可依靠智能算法进行快速地检测、识别、理解等处理,最大限度地挖掘“机”的计算潜能,而且可凭借适时、适当的人工推理、预测和决策有效增强系统认知的准确性和可靠性,最大程度地发挥人的认知优势。其次,分别从混合增强的视觉监测、视觉驾驶、视觉决策以及视觉共享等4个领域探讨可纳入该架构的代表性应用及存在的问题,指出混合增强视觉认知架构是现有技术条件下能够更好地发挥计算机效能、减轻人处理信息压力的方式。最后,基于高、中、低计算机视觉处理技术体系,分析混合增强视觉认知架构中部分中高级视觉处理技术的宏观、微观关系,重点综述可视化分析、视觉增强、视觉注意、视觉理解、视觉推理、交互式学习以及认知评估等关键技术。混合增强视觉认知架构有助于突破当前视觉信息认知“弱人工智能”的瓶颈,将有力促进智能视觉系统向人机深度融合方向发展。下一步,还需在纯粹的基础创新、高效的人机交互、柔性的连接通路等方面开展更加深入的研究。

关键词: 视觉认知;视觉感知;智能视觉感知;混合增强视觉认知;人机融合

Hybrid enhanced visual cognition framework and its key technologies

Wang Peiyuan*, Guan Xin

Naval Aviation University, Yantai 264001, China

Abstract: Although the current intelligent vision system has certain advantages in feature detection, the extraction and matching of large-scale visual information and the cognition of deep-seated visual information remain uncertain and fragile. How to mine and understand the connotation of visual information efficiently, and make cognitive decisions is an engaging research field in computer vision. Especially for the visual cognitive task based on visual perception, the related mathematical logic and image processing methods have not achieved a qualitative breakthrough at present due to limitations by the western philosophy system. It makes the development of computer vision processing intelligent algorithm enter a bottleneck period and completely replacing human to perform more complex operations such as understanding, reasoning, decision

收稿日期: 2020-08-12; 修回日期: 2020-12-18; 预印本日期: 2020-12-25

* 通信作者: 王培元 wangpy629@163.com

making, and learning difficult. The basic framework of hybrid enhanced visual cognition and the application fields and key technologies that can be included in the framework to promote the development of intelligent visual perception and cognitive technology based on the application status of hybrid enhanced intelligence in the field of visual cognition are summarized in this paper. First, on the basis of analyzing the connotation and basic category of intelligent visual perception, human visual perception and psychological cognition are integrated; the definition, category, and deepening of hybrid enhanced visual cognition are discussed; different visual information processing stages are compared and analyzed; and then the basic framework of hybrid enhanced visual cognition on analyzing the development status of relevant cognitive models is constructed. The framework can rely on intelligent algorithms for rapid detection, recognition, understanding, and other processing to maximize the computational potential of “machine”; can effectively enhance the accuracy and reliability of system cognition with timely, appropriate artificial reasoning, prediction, and decision making; and give full play to human cognitive advantages. Second, the representative applications and existing problems of the framework are discussed from four fields, namely, hybrid enhanced visual monitoring, hybrid enhanced visual driving, hybrid enhanced visual decision making, and hybrid enhanced visual sharing, and the hybrid enhanced visual cognitive framework is identified as an expedient measure to enhance computer efficiency and reduce the pressure on people to process information under existing technical conditions. Then, based on high, medium, and low computer vision processing technology systems, the macro and micro relationships of several medium- and high-level visual processing technologies in a hybrid enhanced visual cognition framework are analyzed, focusing on key technologies such as visual analysis, visual enhancement, visual attention, visual understanding, visual reasoning, interactive learning, and cognitive evaluation. This framework will help break through the bottleneck of “weak artificial intelligence” in current visual information cognition and effectively promote the further development of intelligent vision system toward the direction of human-computer deep integration. Next, more indepth research must be carried out on pure basic innovation, efficient human-computer interaction, and flexible connection path.

Key words: visual cognition; visual perception; intelligent visual perception; hybrid enhanced visual cognition; man-machine fusion

0 引言

随着光学数字采集设备的普及和互联网技术的发展,人们可轻易获取海量的图像信息,但要高效地检测、识别和评估其中的异常变化,则存在巨大的感知负担和认知缺陷。虽然现有的图像处理技术能够自动检测场景变化并示警,但是大部分算法处理能力单一、辨别能力有限、决策能力不足,甚至在繁杂环境中会不断增加人的感知和认知负担。可见,在计算机视觉信息处理的智能化进程中(Liu等, 2016),中高级的视觉认知技术在促进人机深度融合中将扮演越来越重要的角色。

部分计算智能系统已远超人类,有的感知智能系统也可与人类媲美,而认知智能的发展远远不够,存在沟通能力较差、理解能力较弱和学习成本较高等短板,最典型的挑战就是对视觉信息的理解和推理。人类视觉系统具备的认知能力高度复杂,计算机视觉认知系统在短时间内无法超越。因此,若要

使之达到自主适应环境、自主决策和主动学习(Chernenkiy等,2018)的目标,不建议抛开人来构建纯粹的计算机视觉认知模型。

在智能视觉感知不断向视觉认知发展的今天,本文归纳出的混合增强视觉认知架构整合机器感知、人类认知和计算机运算,是现有技术条件下一种较为可行的方案,可实现对大规模、非完整和非结构化的视觉信息的认知处理,解决模糊和不确定感知问题,弥补单一系统的认知不足(Pan, 2016),主要涉及3个方面:1)从视觉感知的层面出发,基于各类光学传感器信息和人的主观感知,形成融合信息的输入;2)计算机对重要信息进行筛选、分类、理解、推理、判断和学习等操作,然后以视觉增强、复杂数据可视化等方式,形成人的认知输入,进而将人的主观决策反馈给计算机认知,形成独特的认知闭环;3)将计算机的运算结果与人的决策相匹配,形成概率化与规则化有机协调的优化输出,如图1所示,其中,“ $\otimes$ ”表示对视觉感知信息的深度融合,“ $\oplus$ ”表示计算机视觉认知信息对人眼认知的反馈,二者叠加互补。

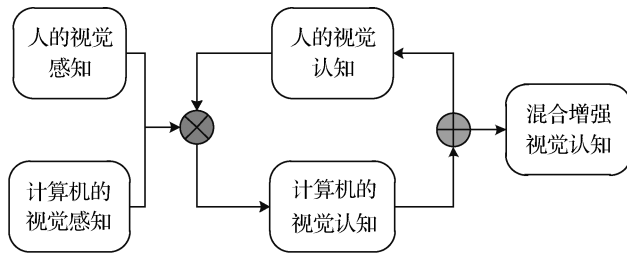


图1 混合增强视觉认知闭环

Fig. 1 The closed loop of hybrid enhanced visual cognition

1 智能视觉感知的进化

感知是指信息从物理域或信息域映射到认知域的一种输入过程,人类80%的感知信息来自视觉。视觉感知是一个从图像中找出物体是什么、理解并预测物体运动变化的过程(Shi,2017)。认知与智力相关但并不相同,可解释为基于感知、推理、学习和先验模型的知识生成,包含认识、理解、决策等心理过程。视觉认知是对人类执行视觉任务(如理解和识别形状)中心理机制的抽象描述,其过程依赖于感知、记忆、想象和逻辑判断(Liu等,2016),主要分为两个阶段:一是对客观世界信息的感知,识别并理解目标形状、属性及场景;二是基于记忆和检索能力的推理、决策过程。

智能视觉感知是指针对光学传感器信息,借鉴生物视觉感知机制,融合计算机视觉处理方法和相关智能算法,执行以计算机运算为核心的图像信息察觉、理解和预测的过程,如图2所示。智能视觉感知是计算机、感知、视觉与人工智能等领域的交叉,这些领域已取得一定成就,然而在计算机视觉认知方面还没有重大进展。

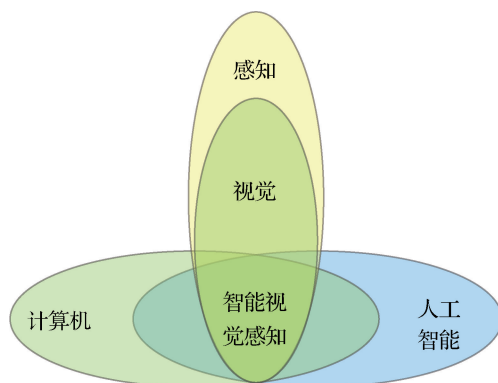


图2 智能计算机视觉感知

Fig. 2 The intelligent computer vision perception

计算机视觉认知是指基于感知系统理解的基本环境信息,进一步通过推理、决策和学习等程序运算表达内在知识,进而为改变环境状态提供支持,是未来态势认知(尤其是战场态势认知)的关键(Ma等,2015)。然而,受限于弱人工智能阶段的技术水平,无法完全以计算机运算为核心完成较为复杂的视觉认知任务。因此,基于外在环境信息和内在系统状态,融合计算机处理大规模数据的感知优势和人类推理、决策的认知优势,提出混合增强视觉认知概念。混合增强视觉认知是以计算机运算为核心的智能视觉感知来增强人对复杂数据的理解认知,以人的推理、辅助决策支持增强计算机的认知学习能力,执行以深度人机交互为核心的图像信息察觉、理解、推理、预测、决策、学习的过程,如图3所示。

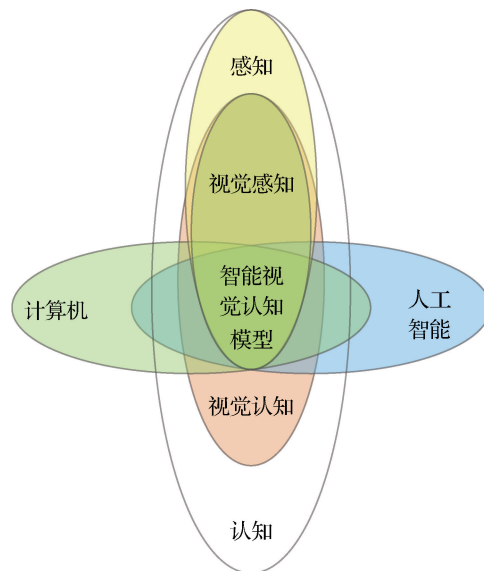


图3 混合增强视觉认知的基本范畴

Fig. 3 The hybrid enhanced visual cognition

图4给出了计算机视觉感知到计算机视觉认知的深化过程。其中, T 表示不同阶段视觉信息处理技术的演变状态。显然,人的视觉系统能够完成视觉感知到视觉认知的自然过渡。在以“人机交互”为核心的计算机视觉感知阶段,主要依靠人完成图像信息的理解、预测等高级操作,计算机作为辅助工具;在以“计算机运算”为核心的智能视觉感知阶段,具有一定智能化程度的算法能够有效察觉、理解大规模视觉信息,但受认知计算水平所限,难以对复杂信息所蕴含的知识进行推理、决策;在以“深度人机融合”为核心的混合增强视觉认知阶段,进一步将人的认知作为辅助元素促进计算机的学习进化;

最终达到完全以“计算机认知计算”为核心的计算机视觉认知水平。

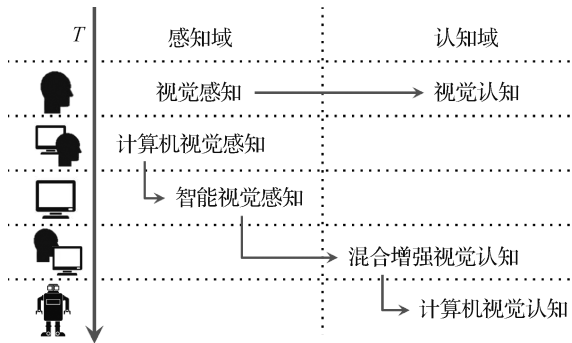


图 4 计算机视觉认知深化过程
Fig. 4 The deepening process of computer vision cognition

目前已有部分视觉认知模型的研究。视觉认知计算模型在人与计算机之间搭建了认知桥梁,旨在使计算机完成人类大脑的部分工作(Li 等,2018);基于海马—前额叶记忆系统的认知计算模型(Aleshinskaya 和 Albatsha,2020)可使系统在单一感知的基础上自主学习更为复杂的感知—动作映射;ABGP-CCSM(awareness, belief, goal, plan-convolutional generative stochastic model)认知模型(Ma 等,2015),对场景噪声具有较强的鲁棒性;此外,还有基于PMJ(perception, memory and judgement)模型的分布式计算认知模型(Forbus 等,2017)及其对象识别模型等(Taffar 和 Miguët,2019)。但是,现有视觉认知模型的存储及处理能力有限,混合增强智能

视觉认知模型一方面需要针对认知对象和事件进行信息管理、功能关联和物理布局,通过与环境的交互减轻大规模视觉信息所施加的认知负荷(Kirschner 等,2018);另一方面要能够利用数据挖掘或机器学习的方法生成抽象或派生知识的表示(Christensen 和 Nagel, 2006),通过注意机制来选择感兴趣的信息。

可见,视觉感知智能应与更深层次的混合增强视觉认知智能融合发展,在逻辑判断基础上不断发展认知推理、情感互动和辅助性决策的视觉认知应用。因此,基于人机协作增强智能,归纳总结了人在回路的混合增强视觉认知方法的基本架构,如图5所示。首先,由视觉传感器采集目标与环境数据;其次,计算机基于图像处理算法完成低层次的视觉要素察觉,基于智能算法完成中层次的视觉理解,并通过视觉共享机制形成联合态势,完成对大规模复杂数据的智能视觉感知;然后,基于可视化技术将特征信息与人脑信息处理通路连接,人脑与计算机协同进行态势预测、辅助决策,完成对复杂数据的高层次混合增强视觉认知;最后,将此次人机协作过程存储为经验数据,分配交互权重,完成对视觉认知策略的学习。该框架表征了混合增强视觉认知模型的基本范畴,通过可视化单元和多周期寻优,可将视觉传感系统、计算机系统、海量存储系统与人类的认知系统进行特征级与决策级的融合,在语义活动、信息共享、态势预测和决策支持等方面有独特的优势。

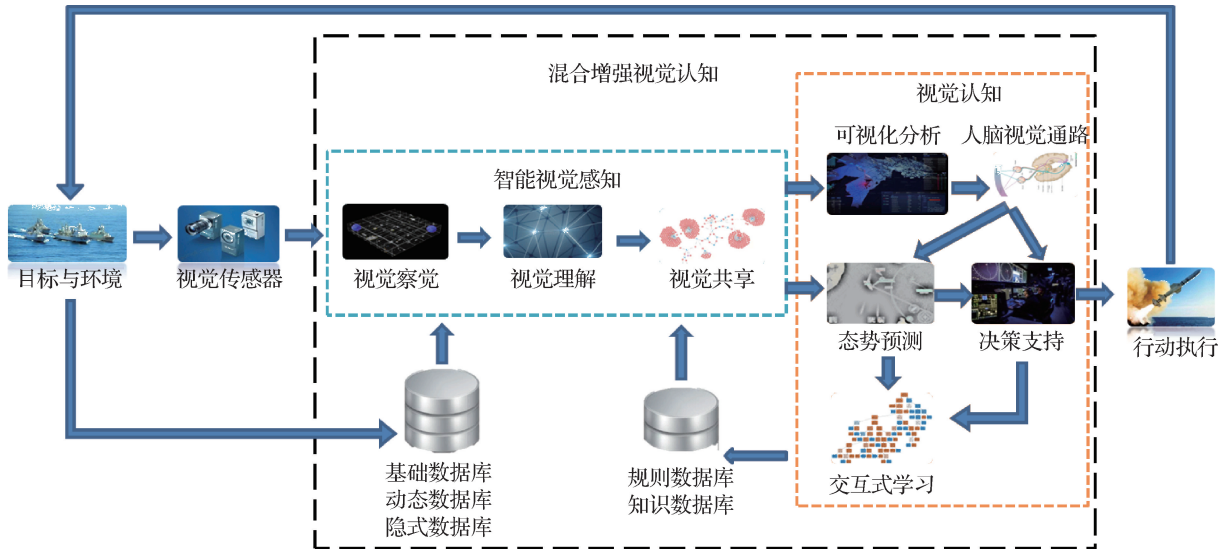


图 5 混合增强视觉认知基本架构
Fig. 5 Basic framework of hybrid enhanced visual cognition

2 混合增强视觉认知架构下的典型应用

视觉认知系统至少需要在语义、解释和学习特征上具备主动性或自主性。目前,语义和解释规则已可编码到视觉感知系统中,赋予其知道和理解的能力,部分视觉认知系统已可利用运动注意机制和视觉数据的长期一致性,自动学习并建立固定场景中实体的语义。例如,欧盟的视觉认知系统可将多个识别模型与不同类型的存储器相结合,建立场景信息与特定目标的联系,进行不确定性条件下的高效推理;国内也有了“人工智能认知+”的视觉认知系统基本架构(张迪等,2018)。可纳入混合增强视觉认知架构的典型应用领域主要包括混合增强视觉监测、混合增强视觉驾驶、混合增强视觉决策和混合增强视觉共享。

1)混合增强视觉监测。针对大范围区域内的多目标活动分析,需要及时发现并处理异常事件。通常,异常事件的处理分为3个部分:预测、检测和后续处理。异常预测的一个典型应用是通过物联网数据分析监测目标的状态,但是在面对海量的非结构化图像数据时,很难预测异常事件。这就要求预测模块能够自动处理大规模数据,并将结果传递给人类,对于置信度较低或影响较大的结果,通过专家经验做出进一步判断。进而,认知模型可以利用专家的反馈,提高对图像信息的理解和分析能力,最终实现更好、更快的异常事件预测、检测和后续处置机制。在这种交互机制下,已有了部分智能监测系统,如基于人群分析的报警与误报控制系统(Xi等,2020),基于知识库的语义环境的监测系统(Braun等,2020)等。

2)混合增强视觉驾驶。当前的自动驾驶系统是高度集成的人工智能系统,但是要实现全自动驾驶仍面临许多技术难题。混合增强视觉驾驶是指驾驶员与视觉认知系统对交通工具的控制共享,协同完成驾驶任务,二者之间需要很强的互补性。一方面,视觉认知系统可实现对驾驶场景的高精度连续监控,拓宽视觉感知领域,弥补人的视觉易受光、天气和其他因素影响的不足,降低人为失误的风险;另一方面,认知系统从人机协同中学习驾驶员的心理

和行为,可以为自动驾驶提供经验数据。例如,自动驾驶车辆的视觉认知系统,已可基于记忆和推理的混合学习网络完成对道路场景的理解。该应用的关键问题是如何基于视觉感知信息,进行低确认偏差判断和人机高效交互(Hubbard,2018)。

3)混合增强视觉决策。人机合作决策可能会带来更有价值的解决方案和丰厚的回报。混合增强视觉认知系统可创建一个以视觉信息感知为主导,支持学习、理解、推理和决策的人机交互环境,将大幅提高态势估计能力,并为大规模认知任务协调提供解决方案,具有巨大潜力。这种人机协作决策模式需要集成多机学习、决策模型和领域经验,允许所有合作伙伴访问,且内部工作流透明。一般情况下,协作应用程序应包括专家系统,通过融合知识库中的显式知识、推理规则和专家隐式知识,提供最优决策建议。同时,该程序还需要不同模块之间的标准接口,供各成员进行信息共享,将不同的解决方案组合到合作决策模型中,以获取推荐的最佳解决方案。

4)混合增强视觉共享。将数百万计的混合增强视觉认知系统连接到移动互联网云,形成云视觉混合增强认知网络,是人们追求的一个长期目标。目前,这种互联已经应用到了云机器人、在线智能学习和医疗保健等领域(Atif等,2015;Chen等,2016)。在这些系统中,通过对不同认知任务的优化,可使各系统独立处理特定任务,并能够以合作的方式传递知识,从而在共享空间和有限时间内实现多视觉系统的协同认知。这不仅可显著提高视觉认知的学习潜力和连接性,而且还可以有效地解决视觉认知指令简单化和视景单一化的问题。

注意,从这几个典型应用领域来看,视觉认知发展的最终目的仍是以高度的自主化彻底摆脱人工参与,但是当前的技术条件无法达到这一目标。混合增强视觉认知架构的应用,是在现有技术条件下,为更好发挥计算机效能、减轻人处理信息压力的权宜之计。虽然暂时无法将人的认知能力彻底迁移到机器上,但是却能最大程度地拓展现有技术应用空间、挖掘现有技术应用潜力。虽然在向强智能视觉认知发展的道路上处于过渡阶段,但对相关方法和技术的集成与融合,势必带来视觉认知范式的创新与突破。

3 混合增强视觉认知架构下的关键技术进展

计算机视觉处理技术虽然种类繁多、派系复杂,但基本是按照从低级视觉感知向高级视觉认知的脉络发展。在对动态目标的处理上,低级视觉主要承担类似于人脑视觉信息处理的 where 通路功能(饶恒毅,2002),属于视觉察觉层次;中级视觉承担类似 what 通路功能,属于视觉理解层次;而高级视觉则承担类似于对两个通路整合基础上的视觉估计功能,属于视觉预测层次。计算机视觉处理关键技术体系如图 6 所示。

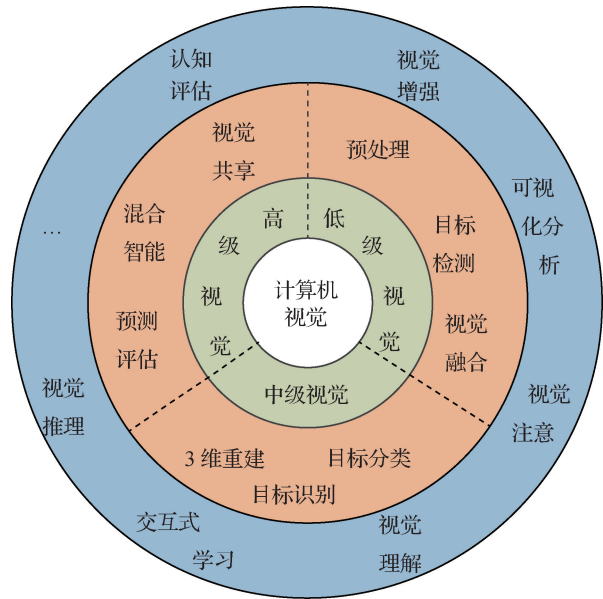


图 6 计算机视觉处理关键技术体系
Fig. 6 Key technology system of computer vision

前文叙述混合增强视觉监测、驾驶、决策和共享等典型应用领域涉及了计算机视觉处理技术的各个层次。例如,预处理、目标检测、目标分类、目标识别、检测跟踪、3 维重建和视觉融合等技术属于视觉感知的范畴(周建军和王培元,2019)。

与混合增强视觉认知架构相关的技术主要包括可视化分析、视觉增强、视觉注意、视觉理解、视觉推理、交互式学习和认知评估等,相关技术与混合增强视觉认知的关系如图 7 所示。

宏观来看,一方面,计算机视觉系统通过可视化分析技术将复杂高维数据直观地向人展示,增强人对数据内部知识的理解;通过视觉增强技术增强人

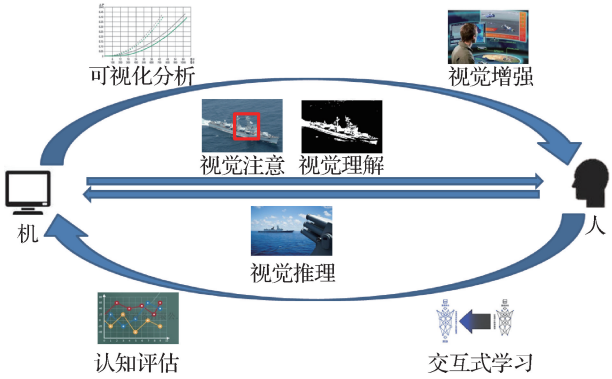


图 7 服务于混合增强视觉认知的部分中高级视觉处理技术
Fig. 7 Some intermediate and advanced visual processing technologies for hybrid enhanced visual cognition

的信息感知维度,提供更多细节信息;通过视觉注意与视觉理解技术,去除冗余信息,增强人的判读效率。另一方面,人利用视觉推理技术,高效完成对事件的决策,增强计算机视觉系统的判决能力;利用交互式学习技术,丰富计算机的视觉系统的认知知识库;利用认知评估技术,不断修正计算机视觉系统的认知偏差。

微观来看,每一类技术都不是单向增强的,均在各自的应用场景下进行混合双向增强,且优化与不足是并存的。

1) 可视化分析。人类无法用肉眼直接从高维度复杂数据中提取有用信息,为了能够直接与数据对话,并帮助人从不同角度分析数据的不确定性,需要进行复杂数据的可视化分析(Kovalerchuk, 2017)。高维数据无损可视化是视觉分析、数据挖掘要解决的主要问题,常见方法如表 1 所示。相关的可视化商业软件较多,如 QlikView、Tibco/SpotFire 和 Tableau 以及 Facebook 和 Google Analytics 等。在混合增强视觉认知过程中实现自适应性和可视化的紧密耦合是人机交互面临的一个重要挑战(Nussbaumer 等,2016)。自适应可视化有助于认知系统具备自学习智能,即通过一系列有效的感知设计,在可视化系统中提供智能辅助(Arbulu 等,2018)来弥补人与复杂数据之间的鸿沟。例如,基于改进的层次视觉适应模型(Nazemi, 2018),在考虑人的感知和认知能力基础上,可用于增强的可视化分析。

2) 视觉增强。虚拟现实(virtual reality, VR)技术增强了人对虚拟空间的感知(Wu 等,2018),增强现实(augmented reality, AR)技术将合成传感器信息

表1 典型高维数据可视化方法
Table 1 Typical visualization methods of high dimensional data

方法	可视化	数据规模特点
主成分分析	有损	规模较大数据
热图、饼图和条形图	无损	小规模数据集
平行坐标和径向坐标	无损	规模较大数据
广义线性坐标及其若干规范	无损	小规模数据集
广义位移成对坐标	无损可逆	小规模数据集
具有半监督框架可视化分析	有损	规模较大数据

转化为人对真实环境的感知,而混合现实技术(mixed reality, MR)则利用实时视频、3维建模和多传感器信息融合等技术,将现实世界与虚拟信息有机融合,使人可以同时感受虚拟世界和真实物体,进行超现实的混合增强视觉认知。从视觉增强系统的发展趋势来看,硬件技术在AR和MR领域发展迅速,出现了许多具有里程碑意义的产品。在民用领域,较为典型的谷歌眼镜实现了视觉增强零的突破,可带来较好的视觉体验;微软的全息透镜可使用户在真实环境中与全息图像交互(Yang等,2019);美国TWC(the weather channel)天气频道也用到了混合现实技术,逼真呈现了恶劣天气(<https://www.gkzhan.com/news/detail/114410.html>);但这些应用还远远不能满足用户对海量数据处理的需求。对此,出现了集成云计算、AR技术和深度学习的视觉增强系统。总体来说,现有的视觉增强系统仍然功能单一,处理能力有限,人机交互能力依然较差。

3) 视觉注意。人类具有异常突出的数据筛选能力,注意是人类一项重要的心理调节机制。对视觉而言,即为视觉注意,其作用在于快速选择最重要、最有用和最相关的视觉信息。视觉注意机制基于对静态或动态图像的快速粗略分析,选择目标行为的基础信息,提取具有代表性的信息标签,如亮度、颜色、动态和几何(区域边缘、水平和垂直线、角度)等(Shi等,2019a)。该机制应是自适应的,并能够根据图像的统计属性进行调整,降低计算复杂度。早期的计算机视觉注意模型(Koch和Ullman,1985)遵循特征集成理论的基本假设,有基于贝叶斯、决策论、信息论、图模型、频域分析和模式分类的视觉注意模型(王文冠等,2019),而神经形态视觉工具包

可能是最有影响力的视觉注意模型(<https://www.aiche.org/academy/conferences/aiche-annual-meeting/2019/proceeding>)。在机器人的应用上,分为过度注意模型和面向应用的注意模型两种,具有生物注意系统的独特特性(Zibafar等,2019)。然而,在进行混合增强视觉认知时,由于大多数计算机视觉注意模型具备显著性运算符、隐性焦点转移、自下而上和自上而下分析、可视化搜索离线训练和基于空间对象分析等共同特征,易导致仅限于视觉搜索的普遍约束,限制了视觉注意机制的发展。

4) 视觉理解。理解是神经科学和认知理论的范畴,其思维的基础是人类的一般经验。若要理解和解释视觉中的复杂活动,需要过滤掉一般活动的图像数据,生成更加深入的信息。如图8所示,基于视觉系统和人的认知知识库,不仅要理解上下文信息,还要提供上下文的自动文本描述,或者回应操作员的上下文查询信息,从而达到认知共振的目的。

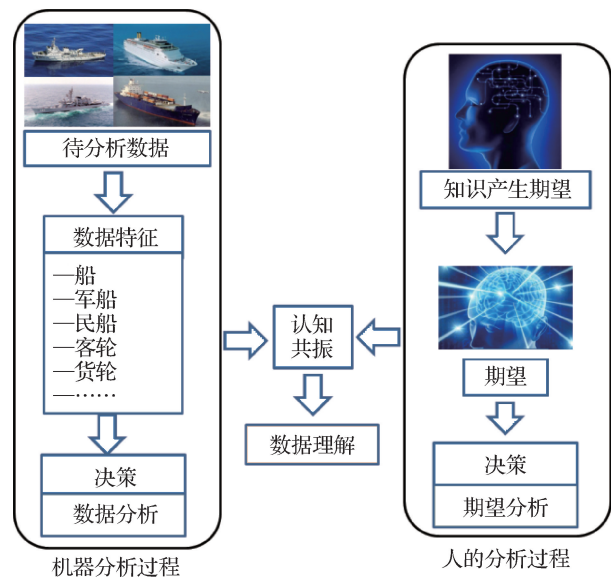


图8 语义分析过程中的视觉理解

Fig. 8 Visual understanding in semantic analysis

当前,视觉理解的研究热点已从低水平视觉任务的算法构建转变为高层次视觉任务的行为分析,尤其是针对大场景的复杂事件识别与分析、社会行为分析、人群分析等研究热度颇高。相关研究大多针对机场、购物中心和火车站等公共场所,采用技术主要包括利用贝叶斯网络(Martinez-Hernandez和Dehghani-Sanij,2018)和矩不变特征描述(Chen等,2017)检测跳跃、撞击、踢腿和坠落等事件(Juang和Wu,2015);基于轨迹信息的最近邻分类器(Puchkin

和 Spokoyny, 2020) 检测冲突、战斗的行为; 基于目标轨迹和相对距离解决社会行为分析的复杂性和模糊性问题 (Zhang 等, 2019; 黄军 等, 2019; Martinez-Villaseñor 和 Ponce, 2020) 等。在混合增强视觉认知中, 人与场景元素交互的方式、顺序和效果是视觉理解必需的, 需要进行更具挑战性的推理。

5) 视觉推理。视觉推理是分析视觉信息并解决其主要矛盾的过程 (Daw, 2020)。人类的推理能力十分强大, 可基于视觉、常识和背景知识, 正确理解歧义信息, 并根据自己的经验推测最可能的结果。以此形成的专家知识系统是视觉推理过程中数据理解的关键。基于训练系统的视觉认知推理过程如图 9 所示。

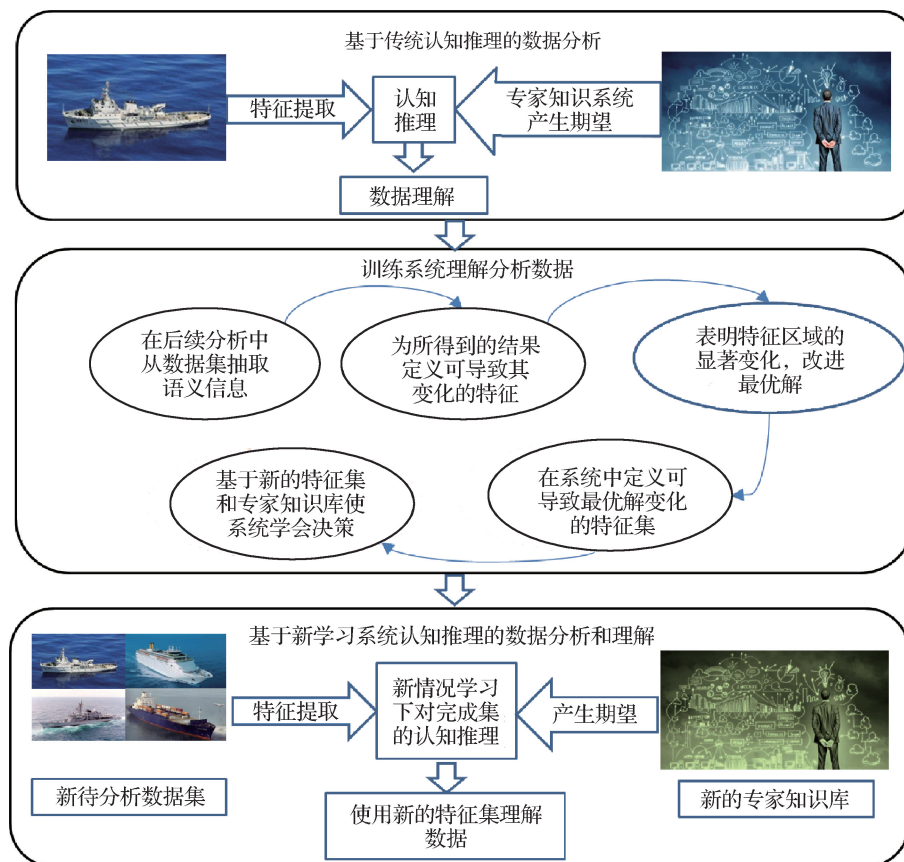


图9 基于训练系统的视觉认知推理

Fig. 9 Visual cognitive reasoning based on training system

在多物体堆叠场景下, 机器人已可基于视觉推理按照正确顺序完成抓取任务 (张翰博 等, 2018); 基于场景图的可解释显式神经模块推理方法可实现低级视觉感知的分离 (Shi 等, 2019b); Chen 等人 (2018) 提出一种新的迭代视觉推理框架, 对缺失的推理区域具有很强的适应性, 在 ADE (auto-deeplab environment) 上实现了 8.4% 的提升; 视觉推理芯片 questcore (求索) 号称全球最智能、性价比最高 (陈玉祥, 2019)。但是, 当前的视觉系统缺乏任意上下文推理功能, 在构建下一代视觉系统时, 需关注将空间推理和语义推理结合起来。

混合增强视觉认知中的视觉推理旨在人机交互条件下, 使认知系统帮助人类处理多种解释, 并且能

够生成与人类基本相同的解释。例如, 通过集成最先进的目标检测网络在对象级别执行此推理, 可在视频中实现语义与空间的详细交互 (Baradel 等, 2018); 基于格式塔心理学原理的可视化推理引擎 (Feest, 2021) 能够模拟经验丰富的人工抉择, 并且在歧义产生时可以在交互中选出最合理的解释, 使视觉认知系统具有与人类推理相似的能力。

6) 交互式学习。在现有视觉系统中增加主动学习能力不仅是对系统认知的挑战, 也是自主适应环境变化、不断进行情报挖掘的实际需求。混合增强视觉认知可基于模仿的交互式学习使计算机获取各种经验数据 (Hou, 2019)。现有的交互式学习模型集成了机器学习、知识库和人类决策, 可从训练数

据或少量样本中学习并预测新的数据,当预测置信度较低时,人类辅助判断。例如,基于视觉信息提供的2维和3维轨迹的交互式学习模块,生成了场景模型和活动模型,场景模型使系统能够理解运动并响应操作员提出的上下文查询,活动模型则用于检测与之不一致的可疑事件。目前,单纯的机器学习过分依赖于规则,系统的移植性和可扩展性较差,只能在有严格约束和有限目标的环境中工作,并且不能处理动态、非完整和非结构化的信息。由人工神经网络、模糊推理、粗糙集、近似推理、进化计算和群体智能构建的学习机制(Zheng等,2017)可整合不同的自适应技术,在一定程度上克服了个体的局限性,但是仍然很难运用数据经验,学习处理开放环境下的时变、复杂、不完备问题。

7) 认知评估。当图像数据具备不确定性、复杂性和时间约束时,会影响视觉认知系统做出最佳决策,出现认知偏差。在这种情况下,混合增强视觉认知系统一般会运用启发式方法进行信息价值的评估,即经验法则。但是这种方法有时会导致严重的系统误判或决策偏差。在情报分析中,这种误判或偏差可能发生在分析周期的每个阶段,从而诱发后续认知的误解、忽视和拒绝等错误。最著名的认知偏差是确认偏差,系统会不成比例地支持初始期望和假设的信息。对此,需要对偏差检测进行新的定义,研究检测和减轻确认偏差新的框架。目前主要通过自动计算预测和主观人类判断(Fan等,2018)两种方式评估视觉认知。总体目标是开发自适应视觉认知分析评估系统,若系统在认知过程中受到确认偏差的影响,则告警并提示用户寻找偏差存在的原因。

4 结 语

随着视觉感知技术不断向认知领域发展,为深化弱人工智能条件下的技术应用,混合增强智能逐渐成为视觉处理技术发展的一个重要方向。在实际应用中,人机相互混合相互补充,可以在现有技术水平下最大限度地挖掘智能视觉系统的潜力,为走向深度人机融合铺路搭桥。为此,还需要在基础创新、人机交互和连接通路等多方面进行努力。1) 更多的基础创新。目前,国内智能视觉领域重技术组合应用,少基础理论创新,受发展环境、专业人才和应

用路径等因素的影响,在算法、芯片和元器件等方面与国际水平仍有差距。2) 高效的人机交互。计算机视觉可自下而上地为人提供视觉增强服务,同时人也要能够自上而下地将主观意图高效地传递给计算机。通过有机的人机互动,使认知模型不断进化。因此,除了研究将高维数据映射到低维、将低维数据重建到高维的视觉增强技术外,还要使计算机能够高效接收并准确理解人的意图及介入行为,同时也要为人确定好合理的介入时机与介入深度。3) 柔性的连接通路。人机间信息的混合交互,需要借助可视化分析或视觉增强等技术建立传感器信息特征与人脑信息处理的连接通路。在计算机端,鉴于算法和数据的复杂度,通路具有宽开性;而人类视觉无法对大型抽象高维数据进行广泛认知,需要在理解其内部规律的基础上,在大脑中形成低维知识的快速实现,即人脑视觉信息处理通路呈现选择性。因此,应进一步加强人机间信息连接通路的柔性研究,使之能够灵活确定连接方式、快速响应系统内外部变化,始终保持稳定、鲁棒的交互。

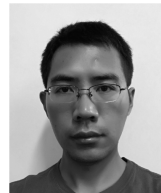
参考文献 (References)

- Aleshinskaya E and Albatsha A. 2020. A cognitive model to enhance professional competence in computer science. *Procedia Computer Science*, 169: 326-329 [DOI: 10.1016/j.procs.2020.02.191]
- Arbulu M, Mateus P, Wagner M, Beltran C and Harada K. 2018. Industry 4.0, intelligent visual assisted picking approach//*Proceedings of International Conference on Mining Intelligence and Knowledge Exploration*. Cluj-Napoca, Romania: Springer: 205-214 [DOI: 10.1007/978-3-030-05918-7_18]
- Atif Y, Mathew S S and Lakas A. 2015. Building a smart campus to support ubiquitous learning. *Journal of Ambient Intelligence and Humanized Computing*, 6(2): 223-238 [DOI: 10.1007/s12652-014-0226-y]
- Baradel F, Neverova N, Wolf C, Mille J and Mori G. 2018. Object level visual reasoning in videos//*Proceedings of Computer Vision-ECCV 2018*. Munich, Germany: Springer: 106-122 [DOI: 10.1007/978-3-030-01261-8_7]
- Braun A, Tuttas S, Borrmann A and Stilla U. 2020. Improving progress monitoring by fusing point clouds, semantic data and computer vision. *Automation in Construction*, 116: #103210 [DOI: 10.1016/j.autcon.2020.103210]
- Chen X J, Ke J, Zhan Y Z, Chen X B, Zhang Q Q, Jiang X M, Song X P, Chen B D, Xu H and Zhang J G. 2017. Improved combined invariant moment for moving targets classification. *Multimedia Tools*

- and Applications, 76(19): 19959-19982 [DOI: 10.1007/s11042-016-4014-x]
- Chen X L, Li L J, Li F F and Gupta A. 2018. Iterative visual reasoning beyond convolutions//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA; IEEE: 7239-7248 [DOI: 10.1109/CVPR.2018.00756]
- Chen Y, Argentinis E and Weber G. 2016. IBM Watson: how cognitive computing can be applied to big data challenges in life sciences research. *Clinical Therapeutics*, 38(4): 688-701 [DOI: 10.1016/j.clinthera.2015.12.001]
- Chen Y X. 2019. Taking advantage of Guizhou public security traffic police big data to move to the whole country [EB/OL]. [2020-07-25]. <http://kpgz.gog.cn/system/2019/05/22/017246184.shtml> (陈玉祥. 2019. 借势贵州公安交警大数据走向全国依图新推视觉推理芯片[EB/OL]. [2020-07-25]. <http://kpgz.gog.cn/system/2019/05/22/017246184.shtml>)
- Chernenkiy V, Gapanyuk Y, Terekhov V, Revunkov G and Kaganov Y. 2018. The hybrid intelligent information system approach as the basis for cognitive architecture. *Procedia Computer Science*, 145: 143-152 [DOI: 10.1016/j.procs.2018.11.022]
- Christensen H I and Nagel H H. 2006. *Cognitive Vision Systems*. Berlin, Heidelberg: Springer: 221-246 [DOI: 10.1007/11414353]
- Daw E. 2020. What is visual reasoning? [EB/OL]. [2020-05-09]. <https://www.wisegeek.com/what-is-visual-reasoning.htm>
- Fan S J, Ng T T, Koenig B L, Herberg J S, Jiang M, Shen Z Q and Zhao Q. 2018. Image visual realism: from human perception to machine computation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(9): 2180-2193 [DOI: 10.1109/TPAMI.2017.2747150]
- Feest U. 2021. Gestalt psychology, frontloading phenomenology, and psychophysics. *Synthese*, 198: 2153-2173 [DOI: 10.1007/s11229-019-02211-y]
- Forbus K D, Liang C and Rabkina I. 2017. Representation and computation in cognitive models. *Topics in Cognitive Science*, 9(3): 694-718 [DOI: 10.1111/tops.12277]
- Hou R L. 2019. Marching ahead in exchanges and mutual learning. *China Today*, (6): #2
- Huang J, Wang C, Liu Y and Bi T T. 2019. The progress of monocular depth estimation technology. *Journal of Image and Graphics*, 24(12): 2081-2097 (黄军, 王聪, 刘越, 毕天腾. 2019. 单目深度估计技术进展综述. 中国图象图形学报, 24(12): 2081-2097) [DOI: 10.11834/jig.190455]
- Hubbard T L. 2018. *Spatial Biases in Perception and Cognition*. Cambridge: Cambridge University Press [DOI: 10.1017/9781316651247]
- Juang L H and Wu M N. 2015. Fall down detection under smart home system. *Journal of Medical Systems*, 39: #107 [DOI: 10.1007/s10916-015-0286-3]
- Kirschner P A, Sweller J, Kirschner F and Zambrano R J. 2018. From cognitive load theory to collaborative cognitive load theory. *International Journal of Computer-Supported Collaborative Learning*, 13(2): 213-233 [DOI: 10.1007/s11412-018-9277-y]
- Koch C and Ullman S. 1985. Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology*, 4(4): 219-227
- Kovalerchuk B. 2017. Visual cognitive algorithms for high-dimensional data and super-intelligence challenges. *Cognitive Systems Research*, 45: 95-108 [DOI: 10.1016/j.cogsys.2017.05.007]
- Li N, Zhao X B, Ma B Y and Zou X C. 2018. A visual attention model based on human visual cognition//Proceedings of International Conference on Brain Inspired Cognitive Systems. Xi'an, China: Springer: 271-281 [DOI: 10.1007/978-3-030-00563-4_26]
- Liu Y J, Yu M J, Fu Q F, Chen W F, Liu Y and Xie L X. 2016. Cognitive mechanism related to line drawings and its applications in intelligent process of visual media: a survey. *Frontiers of Computer Science*, 10(2): 216-232 [DOI: 10.1007/s11704-015-4450-1]
- Ma G, Yang X, Zhang B, Qi B Y and Shi Z Z. 2015. An environment visual awareness approach in cognitive model ABGP//Proceedings of the 27th IEEE International Conference on Tools with Artificial Intelligence. Vietri sul Mare, Italy: IEEE: 744-751 [DOI: 10.1109/ICTAI.2015.111]
- Martinez-Hernandez U and Dehghani-Sanji A A. 2018. Adaptive Bayesian inference system for recognition of walking activities and prediction of gait events using wearable sensors. *Neural Networks*, 102: 107-119 [DOI: 10.1016/j.neunet.2018.02.017]
- Martinez-Villaseñor L and Ponce H. 2020. Design and analysis for fall detection system simplification. *Journal of Visualized Experiments*, #158 [DOI: 10.3791/60361]
- Nazemi K. 2018. Intelligent visual analytics—a human-adaptive approach for complex and analytical tasks//Karwowski W and Ahram T, eds. *Intelligent Human Systems Integration*. Dubai, United Arab Emirates: Springer: 180-190 [DOI: 10.1007/978-3-319-73888-8_29]
- Nussbaumer A, Verbert K, Hillemann E C, Bedek M A and Albert D. 2016. A framework for cognitive bias detection and feedback in a visual analytics environment//Proceedings of 2016 European Intelligence and Security Informatics Conference. Uppsala, Sweden: IEEE: 148-151 [DOI: 10.1109/EISIC.2016.038]
- Pan Y H. 2016. Heading toward artificial intelligence 2.0. *Engineering*, 2(4): 409-413 [DOI: 10.1016/J.ENG.2016.04.018]
- Puchkin N and Spokoyny V. 2020. An adaptive multiclass nearest neighbor classifier. *ESAIM: Probability & Statistics*, 24: 69-99 [DOI: 10.1051/ps/2019021]
- Rao H Y. 2002. *Dynamic Memory Representation of Visual Information and Cognitive Neuroscience Research on Visual Pathway*. Hefei: University of Science and Technology of China (饶恒毅. 2002. 视觉信息的动态记忆表征及视通路的认知神经科学研究. 合肥: 中国科学技术大学)

- Shi J W, Zhu Q G, Chen Y J, Wu J and Xiong R. 2019a. Human visual perception based image quality assessment for video prediction//Proceedings of 2019 Chinese Automation Congress (CAC). Hangzhou, China; IEEE: 3205-3210 [DOI: 10.1109/CAC48633.2019.8996234]
- Shi J X, Zhang H W and Li J Z. 2019b. Explainable and explicit visual reasoning over scene graphs//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA; IEEE: 8368-8376 [DOI: 10.1109/CVPR.2019.00857]
- Shi Z. 2017. Chapter 5: Visual Perception//Mind Computation. New Jersey: World Scientific: 183-217 [DOI: 10.1142/9789813145818_0005]
- Taffar M and Miguet S. 2019. Local appearance modeling for objects class recognition. Pattern Analysis and Applications, 22(2): 439-455 [DOI: 10.1007/s10044-017-0639-2]
- Wang W G, Shen J B and Jia Y D. 2019. Review of visual attention detection. Journal of Software, 30(2): 416-439 (王文冠, 沈建冰, 贾云得. 2019. 视觉注意力检测综述. 软件学报, 30(2): 416-439) [DOI: 10.13328/j.cnki.jos.005636]
- Wu W C, Zheng Y X, Chen K Y, Wang X Y and Cao N. 2018. A visual analytics approach for equipment condition monitoring in smart factories of process industry//Proceedings of 2018 IEEE Pacific Visualization Symposium. Kobe, Japan; IEEE: 140-149 [DOI: 10.1109/PacificVis.2018.00026]
- Xi Y, Zheng J B, He X J, Jia W J, Li H H, Xie Y F, Feng M C and Li X X. 2020. Beyond context: exploring semantic similarity for small object detection in crowded scenes. Pattern Recognition Letters, 137: 53-60 [DOI: 10.1016/j.patrec.2019.03.009]
- Yang X G, Ma M S, Li W P and Xie X L. 2019. Intelligent visual enhancement system. Journal of Physics: Conference Series, 1168(2): #022002 [DOI: 10.1088/1742-6596/1168/2/022002]
- Zhang D, Lu N, Li Y Z, Teng F and Wang L. 2018. Research situation analysis of intelligent visual perception and understanding. Computer Engineering and Applications, 54(19): 18-25, 33 (张迪, 鲁宁, 李宜展, 滕飞, 王丽. 2018. 智能视觉感知与理解研究态势分析. 计算机工程与应用, 54(19): 18-25, 33) [DOI: 10.3778/j.issn.1002-8331.1808-0125]
- Zhang H, Huang J, Tian F, Dai G Z and Wang H A. 2019. Trajectory prediction model for crossing-based target selection. Virtual Reality and Intelligent Hardware, 1(3): 330-340 [DOI: 10.3724/SP.J.2096-5796.2019.0017]
- Zhang H B, Lan X G, Zhou X W, Tian Z Q, Zhang Y and Zheng N N. 2018. Robotic grasping in multi-object stacking scenes based on visual reasoning. Scientia Sinica Technologica, 48(12): 1341-1356 (张翰博, 兰旭光, 周欣文, 田智强, 张杨, 郑南宁. 2018. 基于视觉推理的机器人多物体堆叠场景抓取方法. 中国科学(技术科学), 48(12): 1341-1356) [DOI: 10.1360/N092018-00169]
- Zheng N N, Liu Z Y, Ren P J, Ma Y Q, Chen S T, Yu S Y, Xue J R, Chen B D and Wang F Y. 2017. Hybrid-augmented intelligence: collaboration and cognition. Frontiers of Information Technology and Electronic Engineering, 18(2): 153-179 [DOI: 10.1631/FITEE.1700053]
- Zhou J J and Wang P Y. 2019. Development of the visual perception technology and some intelligent suggestions. Video Engineering, 43(5): 91-97 (周建军, 王培元. 2019. 视觉感知技术发展及智能化建议. 电视技术, 43(5): 91-97) [DOI: 10.16280/j.videoe.2019.05.025]
- Zibafar A, Saffari E, Alemi M, Meghdari A, Faryan L, Pour A G, RezaSoltani A and Taheri A. 2019. State-of-the-art visual merchandising using a fashionable social robot: RoMa. International Journal of Social Robotics, 13: 509-523 [DOI: 10.1007/s12369-019-00566-3]

作者简介



王培元, 1985年生, 男, 讲师, 主要研究方向为计算机视觉。

E-mail: wangpy629@163.com

关欣, 女, 教授, 主要研究方向为信息对抗和态势感知。

E-mail: gxtongwin@163.com