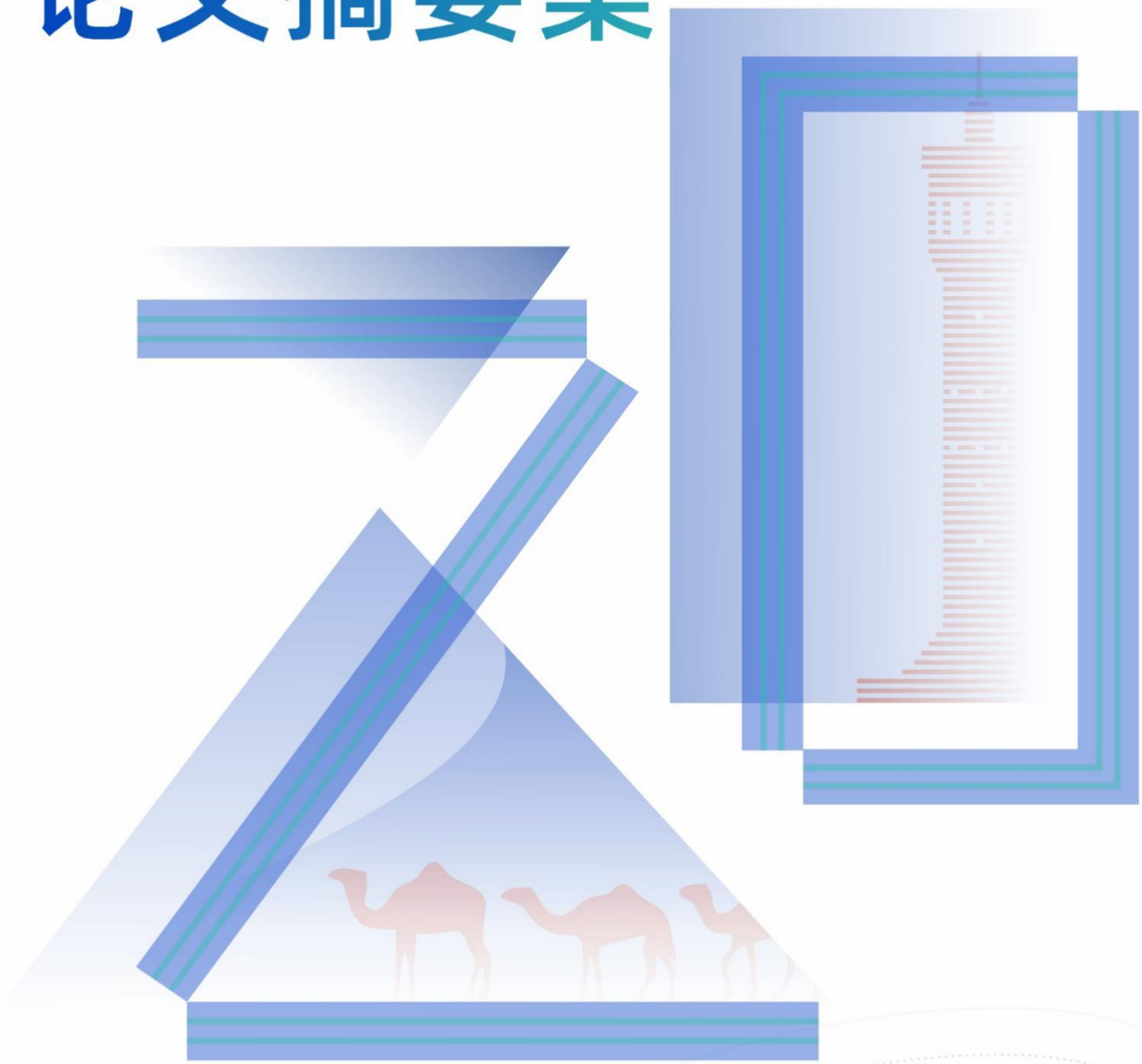
NCIG 2020

第二十届全国图象图形学学术会议

论文摘要集



6月28日-30日 线上召开

The 20th National Conference on Image and Graphics

(NCIG2020)

第二十届全国图象图形学学术会议

论文摘要集

中国图象图形学学会　编

组委会

大会主席

谭铁牛　　中国科学院自动化研究所

吾守尔 斯拉木　新疆大学

程序委员会主席

赵　耀　　北京交通大学

王　亮　　中国科学院自动化研究所

汪烈军　　新疆大学

组织委员会主席

马惠敏　　北京科技大学

孔祥维　　浙江大学

马洪兵　　新疆大学

论坛主席

彭宇新　　北京大学

程　健　　中国科学院自动化研究所

赞助主席

俞能海　　中国科学技术大学

王　涛　　爱奇艺

展览主席

毋立芳　　北京工业大学

唐金辉　　南京理工大学

奖励委员会主席

吴一戎　　中国科学院空天信息创新研究院

章毓晋　　清华大学

赖剑煌　　中山大学

出版主席

罗　斌　　安徽大学

林春雨　　北京交通大学

宣传主席

程明明　　南开大学

马占宇　　北京邮电大学

马思伟　　北京大学

王楠楠　　西安电子科技大学

本地主席

李秀红　　新疆大学

库尔班 吾布力　新疆大学

讲习班主席

刘青山　　南京信息工程大学

鲁继文　　清华大学

竞赛主席

林　倞　　中山大学

刘　偲　　北京航空航天大学

李冠彬　　中山大学

博士生论坛主席

郑伟诗　　中山大学

徐　昆　　清华大学

网站主席

许永超　　华中科技大学

程序委员会委员（按姓氏拼音排序）

安高云、安　如、白　琮、柏连发、鲍秉坤、蔡兴泉、操晓春、曹聪琦、曹　刚、岑翼刚、陈北京、陈　斌、陈纯毅、陈恩庆、陈仁杰、陈善雄、陈伟锋、陈熙霖、陈　阳、陈泽睿、

陈志波、程　俊、程述立、储　珺、丛润民、邓　成、丁　凯、董　晶、董军宇、董天阳、董天祯、董笑菊、杜培军、樊　彬、樊永显、范　超、方贤勇、方志军、冯国瑞、冯结青、冯前进、付冬梅、付　莹、高常鑫、高海昌、葛仕明、耿则勋、顾广华、郭栋梁、郭丰俊、郭英凯、郭振华、韩　苛、何　冰、何　川、何高奇、和红杰、侯北平、呼　延、胡东辉、胡　栋、胡峻林、胡瑞敏、胡　涛、胡小强、胡　英、黄方军、黄　浩、黄　华、黄建军、黄林江、黄小明、黄晓霖、纪　雯、贾云得、蹇木伟、姜　光、姜求平、姜　涛、姜志国、蒋秉川、蒋朝辉、蒋慧琴、蒋　杰、蒋　敏、蒋兴浩、解　翠、金　鉴、金连文、金　枝、荆　雅、康显桂、柯　晶、柯永振、孔　军、孔庆杰、孔祥维、库尔班.吾布力、赖惠成、赖剑煌、雷　涛、冷　璐、李安虎、李　波、李　策、李昌华、李　晨、李　华、李金屏、李　军、李俊山、李瑞瑞、李　腾、李　玺、李晓龙、李秀红、李学龙、李学明、李永杰、李云松、李　治、梁荣华、廖　康、廖理心、廖　频、廖　鑫、林　川、林春雨、林　巍、林　一、林宙辰、刘　斌、刘　东、刘复昌、刘富强、刘赶超、刘　刚、刘光杰、刘国英、刘汉烨、刘　浩、刘华锋、刘建毅、刘　琚、刘　凯、刘　丽、刘利刚、刘良云、刘　龙、刘美琴、刘青山、刘世光、刘文予、刘晓慧、刘　勇、刘　越、刘振广、卢　伟、鲁　坚、陆慧娟、罗　斌、罗正雄、骆伟祺 、吕　培、吕　岳、马　华、马华东、马惠敏、马建华、马锦华、马利庄、马占宇、闵卫东、牟轩沁、缪永伟、聂礼强、宁纪锋、牛　凯、欧卫华、欧阳建权、潘　纲、庞　勇、彭　飞、彭瑞东、彭宇新、普园媛、钱育蓉、强　彦、强振平、秦斌杰、秦　川、秦继伟、秦姣华、邱学军、屈玉福、任桐炜、任文琦、桑　军、桑　农、邵明文、单光存、邵绪强、邵允学、沈琳琳、沈项军、石　刚、石陆魁、石争浩、史振威、司晨阳、宋　勇、隋晨红、孙建德、孙建国、谭舜泉、谭　云、唐　宏、唐　鑫、田华伟、仝明磊、汪国兴、汪烈军、汪　萌、王　栋、王　皓、王金伟、王俊祥、王坤峰、王培珍、王旗龙、王胜科、王世刚、王卫宁、王文珂、王醒策、王兴刚、王　勋、王永芳、王永红、王涌天、王　朕、王智慧、魏振忠、温杰彬、文　冬、文益民、毋立芳、吴　丹、吴恩华、吴汉舟、吴建斌、吴宪云、吴　晓、吴亚东、吴志泽、武玉伟、夏道勋、夏定元、夏佳志、夏时洪、肖伯祥、肖化超、肖　良、肖鹏峰、熊礼治、徐　凯、徐　明、徐文梅、许倩倩、许威威、许增朴、薛向阳、严宝平、严冬明、阳　兵、杨春芳、杨　关、杨　健、杨健华、杨　猛、杨　夙、杨文明、杨文忠、杨　鑫、杨余久、姚　睿、姚　涛、姚远志、叶继华、印　弘、于俊清、于仕琪、余　卿、俞宏远、俞能海、袁强强、原　达、曾焕强、曾　琼、曾　伟、查正军、战荫伟、张东波、张凤军、张桂梅、张怀清、张建勋、张军平、张乐飞、张　磊、张丽清、张　强、张庆明、张　涛、张天柱、张维刚、张伟伟、张小国、张晓鹏、张新晨、张永飞、张玉金、张玉书、张兆翔、张宗华、赵　晖、赵佳琦、赵世杰、赵铁松、赵文清、赵永强、赵振兵、郑　锋、郑海峰、钟　平、周　刚、周琳娜、周　诠、周　涛、周志光、朱　明、朱向炉、朱学芳、朱永贵、朱远平、左旺孟

前 言

第二十届全国图象图形学学术会议（NCIG2020）得到了领域内科技工作者的广泛响应和有关单位的大力支持，共收到投稿论文 479 篇。经过相关领域专家的认真审阅，最后录用论文 206 篇（录用率为 43%），其中，40 篇录用为口头报告论文（8.4%），166 篇录用为墙报论文（34.6%）。

本次会议论文摘要集的论文内容基本涵盖了图像图形领域的各个方向，包括：图像处理和编码、图像分析和识别、图像理解和计算机视觉、计算机图形学与可视化、虚拟现实和增强现实、人工智能与图像图形、图像图形应用领域等。

会议邀请了 4 位本领域的知名专家作大会特邀报告，并设置了 2 个讲习班、1 个优秀博士与青年学者论坛、4 个专题论坛、1 个图像图形技术挑战赛和前沿成果展览。

本次会议由中国图象图形学学会主办，新疆大学承办，南京理工大学协办，会议筹备工作得到了相关单位和同行专家们的大力支持，在此谨向本次会议论文评审专家和全体论文作者表示衷心感谢!向关心和支持本次会议的单位、专家和同行们致以诚挚的谢意！

程序委员会主席 赵耀 王亮 汪烈军

目录

分组报告 1.....1
中国图像工程 25 年.....1
NSST 域多尺度信息蒸馏网络的医学图像超分辨率.....2
基于改进 U-Net 的二阶段 DCE-MRI 乳腺癌分割.....3
基于深度 K-SVD 的沙漠地震勘探图像降噪.....4
基于文本引导的注意力图像转发预测排序网络.....5
基于 DDPG 的三维重建模糊概率点推理.....6
基于逆向暗通道先验的水下图像增强网络.....7
面向倾斜摄影测量数据的跨越式 LOD 模型 out-of-core 调度机制.....8
应用双通道 Faster R-CNN 与注意力机制模型的图像篡改检测.....9
基于对偶代偿网络的共融人脸对齐与超分.....10
融合多模态注意力机制的视频概述技术.....11
分级特征反馈融合的深度图像超分辨率.....12
分组报告 2.....13
基于 Res2-UNeXt 的医学图像分割.....13
一种具有光照鲁棒性的行人再识别方法.....14
基于子集划分迭代投影集成的标签恢复算法.....15
基于 one-step-ahead 策略的重定向算法.....16
视觉惯性里程计中异常视觉测量的检测与处理.....17
基于深度残差注意力网络的视频表情识别.....18
双光流网络指导的视频目标检测.....19
基于高低分辨率图像融合的快速立体匹配算法.....20
基于极化相关方向图的极化 SAR 图像舰船检测.....21
基于单眼视觉和超声波的室内障碍物检测系统在智能机器人车上的应用.....22
分组报告 3.....23
基于多 GPU 的高精度烟雾模拟并行求解方法的研究.....23
一种基于汉明码、图像矫正和特征值分解的彩色图像盲水印方法.....24
基于宏块复杂度的视频隐写算法.....25
一种将渲染误差高效地转化为屏幕空间蓝噪声的过滤方法.....26
基于径向曲线的颅面统计复原.....27
IIH-GAN:基于生成对抗网络的图像信息隐藏.....28
基于迭代对抗攻击的 JPEG 隐写改进.....29
基于跨尺度字典学习的图像盲解卷积算法.....30
分组报告 4.....31
基于混合高阶图卷积的半监督分类.....31
基于 D-UNet 的遥感图像分割.....32
一种基于人脸和人体特征融合的漫画人物身份识别方法.....33
基于对齐特征表示的跨模态人脸识别.....34
通过风格迁移的雨雾天气条件下无人机图像目标检测方法.....35
18F-FDG PET/CT 全身显像肺部肿瘤诊断：划分样本空间下的集成 E-ResNet-NRC 模型.....36
基于梯度预测网络的雷达回波图像外推研究.....37
基于注意力机制的多尺度车辆行人检测算法.....38

场景视点偏移改善点云分割.....39
基于空间自适应 EPLL 的沙漠地震图像去噪.....40
墙报 1-1.....41
一种基于空间编码结构光的稠密三维重建算法.....41
AV-GAN：基于 Arousal-Valence 情感模型的表情图像生成及其在表情识别中的应用.....42
基于多尺度目标运动辅助的视觉跟踪方法.....43
基于文本注意力指导的双线性融合及 Tucker 分解双模态情感特征提取.....44
基于特征子空间直和的跨年龄人脸识别方法.....45
基于深度回归网络的实时人脸跟踪算法.....46
墙报 1-2.....47
三维烟雾流场超分辨率数据生成网络模型的研究.....47
基于款式变换和局部渲染相结合的虚拟试衣.....48
基于 WGAN-GP 的人脸素描画像-照片转化方法研究.....49
基于多尺度重叠滑动池化的 SSD 果冻杂质检测方法.....50
GCNAS:面向图分类的网络架构搜索.....51
基于 SmsGAN 的对抗样本修复.....52
基于注意力机制的人脸姿态生成.....53
基于信息补偿的缺陷检测方法.....54
基于改进的 SSD 模型手机违规使用目标检测.....55
面向图像分类的基于注意力引导的 Dropout.....56
基于 Transformer 的证件图像无检测文字识别.....57
墙报 1-3.....58
基于跨模态信息迁移融合的人脸情感分析.....58
基于注意力机制的视频对象分割优化算法.....59
基于注意力的短视频多模态情感分析.....60
基于直推式支持向量机和像素级特征的膀胱肿瘤准确分割.....61
基于曲线轮廓匹配的甲骨拓片缀合研究.....62
一种手写体古文字检测模型.....63
三维空间中的鲁棒对抗攻击方法.....64
基于多样性约束和离散度分层聚类的无监督视频行人重识别.....65
感受野形态对卷积神经网络性能影响的研究.....66
墙报 1-4.....67
基于损失函数灵敏度的视频分割网络通道剪枝方法.....67
结合特权信息与注意力机制的场景识别.....68
拉普拉斯超分辨率印刷电路板瑕疵检测方法.....69
超分辨率对抗网络的电路板瑕疵检测.....70
基于改进的 3D-CNN 的高光谱遥感图像地物分类.....71
融合双神经网络的 PCB 缺陷检测方法.....72
一种改进的分层双线性池化图像行为识别方法.....73
深度语义哈希算法.....74
基于跟踪候选目标分类的红外弱小目标检测算法.....75
一种距离交并比回归的孪生网络目标跟踪算法.....76
墙报 1-5.....77
基于本质造型属性的汽车前脸识别.....77

可变形深度人脸特征插值编解码网络.....78
用于角度变化下人脸检测的区域式渐进校准网络.....79
YOLOv3 与点偏移估计相结合的车牌定位.....80
BNNeck 改进批特征消除网络的行人重识别研究.....81
基于三角建模的困难样本三元组损失的行人再识别方法.....82
一种通用的头像卡通化方法.....83
基于自注意力循环网络的无监督单目深度估计.....84
基于深度卷积与全局特征的图像密集字幕描述.....85
一种用于 COVID-19 预测的图卷积神经网络时空数据学习方法.....86
墙报 1-6.....87
基于密集深度插值的 3D 人体姿态估计方法.....87
基于边缘检测生成对抗网络的红外与可见光图像融合.....88
全卷积神经网络的手写历史文档图像版面分析.....89
基于贝叶斯粗糙集和布谷鸟算法的肺部肿瘤高维特征选择算法.....90
基于 DenseNet-Attention 模型的高光谱图像分类.....91
基于深度学习的甲骨文偏旁与合体字的识别.....92
T2AD:一种面向自动驾驶的测试与测试控制语言.....93
支持力反馈的沉浸式物理学习环境.....94
基于沉浸式增强现实的流场可视化方法研究.....95
一种基于虚拟现实系统的船舶数字孪生框架.....96
一种基于三维模型设计的增强现实手术引导方法.....97
墙报 2-1.....98
基于 CRNN+CTC 印刷体维吾尔文识别系统的设计与实现.....98
基于语义-表观特征融合的无监督前景分割.....99
印刷体藏文文档图像行字粘连切分与识别.....100
支持差分隐私的图像数据挖掘方法研究.....101
面向智慧生物实验室的弱外观多目标轻量级跟踪网络.....102
基于残边稀疏机制的红外弱小目标检测.....103
一个噪声鲁棒的轻量级人脸识别网络.....104
Integration Cosaliency Prior and Random Walks for Image Cosegmentation.....105
融合边缘计算和改进 Faster R-CNN 的变电站设备典型视觉缺陷检测系统研究.....106
基于 GIOU 损失的轻量级人车检测模型.....107
墙报 2-2.....108
一种双跟踪器动态级联的视觉跟踪方法.....108
基于聚合局部相关性和全局网络流的多目标跟踪.....109
面向图像检索的特征选择方法.....110
基于多尺度特征融合的火灾检测模型.....111
一种基于卷积神经网络的轨道交通场景人群计数模型.....112
基于镜像蒸馏的人体姿态估计算法.....113
基于小波变换结合显著性检测分割的热带气旋客观定位方法.....114
一种改进的脱机手写体文本行切分算法.....115
墙报 2-3.....116
形态学重构滤波和多阈值区域生长的眼底血管三分图提取方法.....116
融合外观及深度特征的无监督视频目标分割算法.....117

基于置信度加权融合与视觉注意机制的　前景检测方法.....118
基于类脑模型与深度神经网络的目标检测与跟踪算法研究.....119
结合目标局部和全局特征的 CV 遥感图像分割模型.....120
一种基于 U-Net 架构的指静脉轻量级融合 分割算法.....121
基于超像素 Log-Euclidean 高斯度量的传统文化图案分割.....122
基于加权局部特征选择的目标跟踪方法.....123
利用通道剪枝技术的实时实例分割方法.....124
基于匹配解码的视频目标分割.....125
墙报 2-4.....126
一种基于 Mask RCNN 的融合几何特征的冠状动脉分割方法.....126
基于双流 Faster R-CNN 的像素级图像拼接篡改定位算法.....127
融合 Hough 直线检测和 Grab-cut 算法的自适应风机叶片图像分割方法.....128
基于深度学习的藏文现代印刷物版面检测技术研究.....129
遮挡对于目标检测的影响分析.....130
融合图像深度的抗遮挡目标跟踪算法.....131
遮挡图像数据生成系统.....132
多指标融合的人物图像美感评价算法研究.....133
墙报 2-5.....134
一种基于色彩补偿的水下图像综合增强算法.....134
基于双通道卷积神经网络的视频目标移除取证算法.....135
无人系统多模态视觉感知单元测试与评估方法的研究.....136
固定翼无人机杆塔类障碍物探测技术.....137
基于食物检测和类别注意力监督的食物识别.....138
词嵌入技术和端对端的遥感图像描述生成模型.....139
基于光照方向一致性的换脸视频检测.....140
融合稀疏点云补全的 3D 目标检测算法.....141
墙报 2-6.....142
基于视觉的激光雷达数据增强.....142
基于近似后验信息的高光谱海上舰船目标检测.....143
复杂背景下苹果目标的识别定位方法.....144
一种基于图像识别的增量式自聚焦透镜数值孔径测量方法.....145
一种利用弧支撑线段进行瞳孔识别的方法.....146
基于注意力机制和深度监督的全分辨率残差网络乳腺超声图像分割.....147
一种水下鱼类动态视觉序列运动目标检测方法.....148
基于深度卷积网络的海洋涡旋检测模型.....149
基于小样本学习的 X 光图像违禁物品识别.....150
基于 LTP 和 HOG 纹理特征融合的中亚文档图像文种识别.....151
墙报 3-1.....152
基于颜料褪色光谱色差模型的壁画颜色复原方法研究.....152
基于悬链线的曲线保长度伸缩变形算法.....153
基于分形图案的数字凹浮雕生成研究.....154
基于格子玻尔兹曼方法的涡量修正高效烟雾仿真算法.....155
基于 IFS 码的可控分形图算法.....156
面向火电机组运行优化的复杂参数可视分析.....157

基于稀疏超采样的时间性反走样算法.....158
基于样例的祥云图案生成.....159
基于非对称码率分配的立体视频传输算法.....160
墙报 3-2.....161
面向通用空域错误隐藏的自适应混合填充算法.....161
基于图像全位面的高嵌入率可逆数据隐藏.....162
基于混沌置乱和 DWT 变换的无载体图像信息隐藏研究.....163
基于稀疏采样的多媒体传感网图像压缩算法.....164
局部边缘结构引导的深度图像超分辨算法.....165
基于自适应参数引导与边缘点辅助的深度图像超分辨率重建.....166
基于生成对抗网络的 HDR 图像生成方法.....167
针对伪影改善的图像拼接方法.....168
基于 Bregman 迭代字典学习方法的定量磁化率成像.....169
墙报 3-3.....170
基于二元优化模型的自适应直方图均衡.....170
完美可恢复的近似缩略图保持加密技术.....171
多粒度深度神经网络的图像超分辨率重建方法.....172
一种基于邻域像素差值最大化的空域隐写嵌入方法.....173
双加权 LP 范数 RPCA 模型及其在椒盐去噪中的应用.....174
一种基于物理光流与细节增强的湍流图像恢复方法.....175
基于样本错配训练的图像 PRNU 噪声提纯方法.....176
数据驱动的 AVS3 像素域 JND 预测模型.....177
纸质包装品精密检测及评价系统.....178
基于激光—相机系统的语义栅格建图和路径规划.....179
墙报 3-4.....180
基于相关熵的多视角彩色点云配准.....180
融合八度卷积与残差学习的 JPEG 内容自适应隐写分析.....181
基于 DMIR 算法的多重复纹理图像拼接.....182
基于多视角图像的作物果实三维表型重建.....183
基于形变网络的草图人体建模.....184
一种基于无监督学习的多光谱图像配准方法.....185
连通域先验的水下坝体裂缝图像拼接.....186
基于偏好概率加权采样的鲁棒模型拟合方法.....187
墙报 3-5.....188
基于双向加权特征金字塔网络的多人姿态估计.....188
基于多监督学习的跨域草图检索算法.....189
一种基于反距离加权融合的图像拼接算法.....190
基于眼动跟踪的人机交互应用研究.....191
基于改进 L-K 光流的 WebAR 信息可视分析方法.....192
基于多任务学习的初始图像对选取方法.....193
基于点云增强的室内含弱支撑平面物体的三维重建方法.....194
基于 GA-SURF 的多通道图像拼接应用研究.....195
前视钻孔图像的全景图展开与拼接.....196
墙报 3-6.....197

遥感影像变化检测综述.....197
面向多源遥感图像的自适应目标检测.....198
属性约束的人脸素描-照片转换与识别.....199
基于稀疏量化电平和线性规划的 PET 探测器时间标记.....200
一种适形 X 线透视成像系统校正方法.....201
基于组合曝光的计算光学成像方式及其复原算法.....202
结合自适应超像素的极化 SAR 舰船检测.....203
基于 CNN 和 VLAD 的指静脉描述子提取方法.....204
基于循环一致性对抗网络的动漫线稿图像提取.....205
VW-SAE：一种改进的光谱数据特征表示方法.....206

论文编号：6.....161
论文编号：8.....98
论文编号：9.....23
论文编号：13.....54
论文编号：14.....180
论文编号：15.....41
论文编号：22.....24
论文编号：23.....55
论文编号：27.....25
论文编号：36.....99
论文编号：37.....89
论文编号：38.....179
论文编号：40.....100
论文编号：41.....83
论文编号：47.....152
论文编号：54.....96
论文编号：56.....56
论文编号：58.....49
论文编号：59.....31
论文编号：61.....50
论文编号：69.....84
论文编号：71.....102
论文编号：75.....143
论文编号：76.....182
论文编号：78.....162
论文编号：80.....57
论文编号：82.....90
论文编号：85.....103
论文编号：88.....183
论文编号：90.....1
论文编号：92.....163
论文编号：95.....32
论文编号：96.....104
论文编号：98.....102
论文编号：100.....58
论文编号：101.....52
论文编号：102.....13
论文编号：104.....184
论文编号：106.....43
论文编号：107.....164
论文编号：108.....195
论文编号：112.....133
论文编号：113.....165

论文编号：115.....192
论文编号：117.....85
论文编号：119.....105
论文编号：120.....106
论文编号：121.....91
论文编号：122.....144
论文编号：125.....59
论文编号：127.....60
论文编号：128.....197
论文编号：131.....107
论文编号：133.....134
论文编号：136.....108
论文编号：137.....109
论文编号：138.....61
论文编号：139.....53
论文编号：140.....110
论文编号：141.....33
论文编号：142.....92
论文编号：143.....198
论文编号：144.....97
论文编号：145.....14
论文编号：146.....166
论文编号：151.....185
论文编号：160.....15
论文编号：166.....62
论文编号：168.....16
论文编号：172.....87
论文编号：173.....95
论文编号：174.....42
论文编号：178.....193
论文编号：181.....199
论文编号：183.....86
论文编号：185.....34
论文编号：186.....153
论文编号：187.....145
论文编号：188.....3
论文编号：189.....35
论文编号：190.....4
论文编号：192.....146
论文编号：193.....26
论文编号：194.....63
论文编号：198.....200
论文编号：202.....154
论文编号：203.....27

论文编号：204.....201
论文编号：205.....36
论文编号：207.....112
论文编号：208.....17
论文编号：209.....64
论文编号：211.....113
论文编号：216.....147
论文编号：217.....28
论文编号：318.....37
论文编号：219.....114
论文编号：220.....202
论文编号：221.....101
论文编号：224.....43
论文编号：229.....115
论文编号：232.....148
论文编号：237.....206
论文编号：238.....136
论文编号：240.....137
论文编号：243.....186
论文编号：248.....18
论文编号：250.....116
论文编号：255.....5
论文编号：258.....111
论文编号：261.....6
论文编号：262.....155
论文编号：266.....167
论文编号：269.....168
论文编号：270.....29
论文编号：272.....149
论文编号：274.....7
论文编号：280.....117
论文编号：281.....19
论文编号：284.....118
论文编号：286.....150
论文编号：287.....194
论文编号：290.....65
论文编号：291.....170
论文编号：295.....66
论文编号：296.....171
论文编号：297.....67
论文编号：299.....205
论文编号：303.....8
论文编号：306.....20
论文编号：307.....172

论文编号：308.....187
论文编号：309.....119
论文编号：311.....120
论文编号：312.....68
论文编号：318.....37
论文编号：320.....30
论文编号：321.....47
论文编号：323.....138
论文编号：325.....69
论文编号：326.....121
论文编号：327.....70
论文编号：331.....71
论文编号：333.....72
论文编号：336.....156
论文编号：337.....73
论文编号：340.....21
论文编号：342.....122
论文编号：343.....123
论文编号：346.....124
论文编号：350.....178
论文编号：351.....203
论文编号：354.....74
论文编号：356.....173
论文编号：357.....174
论文编号：358.....48
论文编号：361.....125
论文编号：362.....75
论文编号：363.....93
论文编号：364.....44
论文编号：365.....126
论文编号：367.....76
论文编号：368.....139
论文编号：370.....77
论文编号：372.....157
论文编号：373.....158
论文编号：374.....191
论文编号：378.....188
论文编号：379.....189
论文编号：382.....9
论文编号：386.....38
论文编号：388.....78
论文编号：389.....45
论文编号：390.....175
论文编号：392.....160

论文编号：394.....204
论文编号：397.....140
论文编号：399.....190
论文编号：400.....127
论文编号：404.....128
论文编号：407.....129
论文编号：409.....196
论文编号：413.....169
论文编号：417.....94
论文编号：420.....22
论文编号：429.....46
论文编号：432.....39
论文编号：434.....159
论文编号：436.....10
论文编号：437.....88
论文编号：441.....79
论文编号：444.....181
论文编号：445.....141
论文编号：447.....135
论文编号：452.....40
论文编号：458.....176
论文编号：461.....142
论文编号：462.....132
论文编号：463.....80
论文编号：464.....11
论文编号：465.....81
论文编号：466.....130
论文编号：467.....151
论文编号：469.....177
论文编号：473.....131
论文编号：476.....82
论文编号：479.....12

分组报告 1

论文编号：90

中国图像工程 25 年

章毓晋

清华大学电子工程系，北京 100084

（E-mail: zhang-yj@tsinghua.edu.cn， Homepage: oa.ee.tsinghua.edu.cn/~zhangyujin/）

摘　要：图像工程是一个系统地研究各种图像理论，设计和开发各种图像技术和方法，以及研制和使用各种图像设备的综合学科，主要可分成紧密联系又有区别的 3 个层次：图像处理、图像分析和图像理解，另外还包括相关的技术应用。本文是对从 1996 年开始的、已有 25 年的中国图像工程年度文献综述系列的回顾。近 25 年来，为了使国内广大从事图像工程研究和应用的科技人员能够较全面地了解国内图像工程研究和发展的现状，能够有针对性地查询有关文献，并向期刊编者和作者提供有用的参考，笔者每年都对上一年度图像工程相关文献进行统计和分析。25 年间，一共从国内 15 种有关图像工程重要期刊发行的共 2964 期上发表的 65014 篇学术研究和技术应用文献中，选取出 15850 篇属于图像工程领域的文献，并根据各文献的主要内容将其分别归入图像处理、图像分析、图像理解、技术应用和综述评论 5 个大类，然后进一步分入 23 个专业小类，并在此基础上分别进行各期刊各类文献的统计和分析。此次回顾，除汇总和分析了这 25 年的统计分类情况，包括文献总数（期刊登载的文献数量）、选取总数（所选取的图像工程文献数量）和（各大类和各小类的）选取率，还着重对一些主要的研究方向进行了具体的分析和讨论，以提供更全面和可信的各个研究方向发展趋势的信息，并描绘中国图像工程四分之一个世纪不断发展前进的“图像”。

关键词：图像工程；图像处理；图像分析；图像理解；技术应用；文献综述；文献分类；文献计量学

论文编号：96

NSST 域多尺度信息蒸馏网络的医学图像超分辨率

桑雨 1，金海波 1，王杨 2，彭晏飞 1

1.辽宁工程技术大学 电子与信息工程学院,辽宁 葫芦岛 125105； 2.辽河油田宝石花医院, 辽宁 盘锦 124010

摘　要：高分辨率医学磁共振（MR）图像或计算机断层扫描（CT）图像可以提供更清晰的人体解剖细节，有利于疾病的早期诊断。然而，由于成像环境、成像系统和人为等因素，获取清晰的高分辨率医学图像并不容易。近年来，深度学习技术在图像分类、目标检测等任务中表现出了优越的性能。本文提出一种非下采样剪切波变换（NSST）域多尺度信息蒸馏网络（MSIDN）的医学图像超分辨率，即 NSST-MSIDN。MSIDN 主要由一系列级联的多尺度信息蒸馏块组成，自适应地检测出不同尺度下的长路径和短路径信息，充分提取医学图像的多尺度特征，有效地将低分辨率图像恢复为高分辨率图像。此外，以往的方法大多是在空间域预测高分辨率图像，丢失了纹理细节的同时产生了过平滑的输出。因此，将医学图像的超分辨率问题描述为 NSST 系数的预测问题，使得 MSIDN 进一步保持比空间域更丰富的结构细节。最后，在由头部、肺部、腹部和骨部四个身体部位构建的图像数据库上的实验结果表明，与其他先进的方法相比，NSST-MSIDN 能够获得更高的 PSNR 和 SSIM 值，更好地保持了全局拓扑结构和局部纹理细节；进一步，通过消融实验评价了 MSID 块数量对网络性能的影响以及 NSST 的有效性。

关键词：医学图像；超分辨率；卷积神经网络；多尺度信息蒸馏网络；非下采样剪切波变换

论文编号：188

基于改进U-Net的二阶段DCE-MRI乳腺癌分割

秦传波，宋子玉，曾军英(*)，甘俊英，柯凡辉，翟懿奎，朱伯远

五邑大学智能制造学部，广东江门,529020

摘　要：乳腺癌的准确分割是自动化计算机辅助诊断系统的关键基础，对于早期预防、术中引导和预后具有重大临床研究价值。通常来讲，肿块形乳腺癌相对容易辨识和分割，而非肿块乳腺癌病变区的形态变化不一，且病变区存在弱边界、强度不一和较低的对比度，容易造成放射科医生的漏检和错判。针对乳腺癌的自动化分割问题，我们提出一种基于改进U-Net的二阶段乳腺癌分割模型。首先使用FCN模型完成乳房区域的粗分割，然后构建改进U-Net模型UTB-Net实现乳腺东准确分割。所做的工作和贡献如下：

首先，重新设计了该网络的编码单元和解码单元、同时采用PReLU激活函数减少网络非线性表达的信息丢失。针对医学影像分割需要庞大的计算量，我们将归一化层由BatchNorm更换成GroupNorm，以此来减少计算量不足对分割精度的影响。

其次，提出多尺度并行卷积融合模块，其采用多尺度的并行卷积结构，通过不同核大小的卷积操作提取不同尺度的特征信息，丰富模型提取的边缘信息与精细信息。

再次，提出注意力、残差混合解码模块，更换了注意力模块的非线性整流单元，将其与使用残差连接的解码单元结合。以此来改善网络信息流向、提高特征提取能力。

综上，采用深度监督的UTB-Net模型可以提取更加丰富的细节信息，结合全局注意力和局部注意力对细小的乳腺癌病灶进行更准确的定位，最后通过深度监督机制加强网络对前景的响应。实验表明，相较于当前主流分割模型，乳腺癌区域的精确分割结果提升了4个百分点，同时减少钙化区域错误分割，增强了对小面积乳腺癌区域的正确分割。

关键词 ： 乳腺癌；DCE-MRI；医学影像；分割；U-Net

论文编号：190

基于深度 K-SVD 的沙漠地震勘探图像降噪

林红波 1，吕永正 1，邓小英 2

1. .吉林大学通信工程学院，吉林省长春市 130012； 2. 北京理工大学信息与电子学院,北京市 100081

摘　要：地震勘探是一种利用地震图像蕴含的信息对地下结构成像，进而探查地下油气资源的有效方法，地震图像质量直接影响成像精度。然而，地震勘探图像不可避免地收到噪声的干扰，压制地震勘探图像中的随机噪声是提高地震图像质量，辨识和提取有效信号的关键环节，复杂的地震勘探信号结构特征和噪声特性使地震图像降噪面临巨大的挑战。本文提出一种基于快速迭代收缩阈值（FISTA）的深度 K-SVD 的地震勘探图像降噪方法(FDK-SVD)。这种数据驱动的 FDK-SVD 算法构建端到端的神经网络结构学习参数和字典，能准确捕捉地震勘探信号的形态特征，从非平稳的沙漠随机色噪声中恢复有效信号。此外，结合 FISTA 算法利用前两次迭代获得的稀疏系数，不仅加快收敛速率，且提高图像的信噪比，降低信号能量损失。合成沙漠地震图像和实际地震勘探图像降噪结果表明，这种结合深度学习和稀疏表示的快速深度 K-SVD 算法能同时提高信噪比和保留地震勘探信号细节，与其他常用地震降噪算法相比，本文方法在信噪比提高和信号保真方面更具优越性，能够更有效地抑制沙漠勘探图像中非平稳低频色随机噪声，具有实际应用价值。

关键词：地震勘探；沙漠勘探随机噪声；深度学习；K-SVD；快速迭代收缩阈值算法

论文编号：255

基于文本引导的注意力图像转发预测排序网络

潘文雯 1，赵洲 1，俞俊 2，吴飞 1

1 浙江大学计算机科学与技术学院，杭州，310027； 2 杭州电子科技大学计算机科学与技术学院，杭州，310018

摘　要：转发预测在社交媒体网站（SMS）中是一个很有挑战性的问题。现有的转发预测方法大多涉及到媒体信息的选择和表征，但这些模型缺乏探索图像推特与其标题或评论之间关系的能力。基于这方面的缺陷，我们使用多模态神经网络从多模态内容中学习联合图像的推特表征，从而可提供不同形式的互补信息。本文研究了 SMS 中的图像转发预测问题，预测用户再次转发图像推特的图像共享行为。与现有的研究不同，本文首先提出异构图像转发建模网络（IRM），所利用的是用户之前转发图像推特中的相关内容、之后在 SMS 中的联系和被转发者的偏好三方面的内容。同时，引进由文本引导的两个子网络多模态神经网络，其中递归神经网络学习图像推特上下文信息的语义表征，卷积神经网络学习视觉表征。在此基础上，提出文本引导的多模态神经网络，构建新型多方面注意力排序网络学习框架，从而学习预测任务中的联合图像推特表征和用户偏好表征，使多面排序的度量值隐含在用户偏好表征中来进行图像转发预测。在 Twitter 的大规模数据集上进行的大量实验表明，在所有情况下，AMNL+方法都能达到最佳性能。这表明利用多模态图像推特的联合图像推特表征及其关联上下文和多维度排序度量的注意多维度排序网络学习框架可以进一步提高图像转发预测的性能。

关键词：图像转发预测；多模态学习；文本引导；注意力机制

论文编号：261

基于 DDPG 的三维重建模糊概率点推理

李雷，吴素萍

宁夏大学信息工程学院，银川 750021

摘　要：随着人工智能、计算机视觉和图像处理的发展，三维重建在机器人、自动驾驶、虚拟现实和增强现实的需求日益增加。近期，在三维重建任务中应用深度学习方法取得了显著的效果，并且基于隐式地表示物体三维形状作为深度网络分类器的决策边界的方法也得到了广泛的应用，但是对于复杂拓扑的物体以及一些表面细节信息仍然难以准确进行预测，即存在模糊概率点。为了解决上述问题，本文提出了一种基于深度强化学习算法 DDPG(Deep Deterministic Policy Gradient)对三维重建中模糊概率点进行再推理，实现了高质量的单视图物体三维重建。本文的方法是端到端的，包括以下四个部分：拟合物体三维形状的网络学习过程、聚合模糊概率点周围点信息的邻域路由机制、注意力机制引导的信息聚合和基于深度强化学习算法的概率调整。具体来说，本文通过聚合模糊概率点的局部信息和图像全局信息，再通过深度强化学习算法 DDPG(Deep Deterministic Policy Gradient)训练的智能体给出相应的动作使其跳出模糊概率区间。本文在公开的三维形状数据集上进行了大量的定性和定量评估证明了该方法的有效性。本文提出的方法结合了强化学习和深度学习，并且聚合了模糊概率点周围局部信息和图像全局信息以更加准确的对模糊概率点进行再推理，进而有效的提升了模型对复杂拓扑和细节信息的重建能力。

关键词：三维重建；深度学习；注意力机制；信息聚合；深度强化学习算法

论文编号：**274**

基于逆向暗通道先验的水下图像增强网络

申月 1，赵浩然 1，孙鑫 1，董军宇 1

1.中国海洋大学,山东省青岛市 266100

摘　要：水下图像在水下机器人领域和海洋研究领域发挥着重要作用。然而，复杂的水下成像环境会导致水下图像出现颜色失真、细节模糊和对比度降低等退化问题。本文提出了一种新的水下图像增强模型，可以有效恢复多种不同水质中的图像，尤其是浑浊现象严重的水下图像。本文模型由两部分组成：第一部分设计逆向暗通道先验算法，通过设置最大值滤波暗通道阈值可分离质量下降现象严重的图像；并采用水下恢复算法获得的适应性透射率，对分离出的质量下降严重图像进行恢复。第二部分将恢复得到的图像及非浑浊水质图像输入以生成对抗网络为主体的框架，端到端地进行着色转换和增强，并添加增加边缘加深模块，学习多尺度边缘信息以进行边缘恢复。所提算法可以从浑浊水质图像恢复得到高质量图像，并具有很好的适应性。

通过实验进行定量和定性分析，本文所提出模型在在恢复模糊现象严重的水下图像时有较大提升，在SSIM和PSNR得分上优于现有方法。本文与其他现有方法在真实水下数据集U45上进行了对比实验，在多种水质中的图像颜色及结构恢复方面都有很好的效果，在实际场景中更具有普适性。

关键词：水下成像模型；图像恢复；图像增强；暗通道先验；生成式对抗网络；

论文编号：303

面向倾斜摄影测量数据的跨越式 LOD 模型 out-of-core 调度机制

曾红梅 1，杨沁七 1，刘子桓 1，张严辞 1

1.四川大学视觉合成图形图像技术国家级重点实验室,四川省成都市 610065

摘　要：大规模城市场景实时可视化对智慧城市的建设具有重大意义，但是场景LOD模型数据巨大，往往是几十GB乃至几百GB，不能常驻内存。对于实时加载无人机倾斜摄影采集的大规模场景数据，场景LOD模型out-of-core调度机制能否快速加载显示用户期望的LOD节点至关重要。本文的提出跨越式LOD模型out-of-core调度机制能够跨过无需加载的中间层次LOD模型，减少数据加载量。基本思想是，通过预计算生成LOD节点树保存在外存，系统运行时载入的方法，使得系统能够跨层次加载LOD节点。同时，在限制时间内无法完全加载用户期望的LOD节点时，利用多选择背包问题方法，求解生成替代效果最优的LOD节点集合，并且跨越所有中间层次的节点进行加载显示。 与此同时，为了满足实时性，本文离线生成物品模板，通过从物品模板中生成少量物品的方法，降低求解多选择背包问题的时间复杂度。实验结果表明，相比与OpenSceneGraph三维渲染引擎的LOD模型调度机制，在相同的限制加载时间内，本文算法加载的LOD节点数量为OpenSceneGraph的30%左右，节点数据总加载量为OpenSceneGraph的50%左右，并且本文算法的最终渲染效果比OpenSceneGraph更加精细，更加接近期望渲染结果。在大规模场景可视化系统中，本文提出的跨越式LOD模型out-of-core调度机制，能在限制内存容量下，对大规模场景数据进行实时调度显示，并且最终渲染效果更加精细。

关键词：LOD(level of detail)；out-of-core 调度机制；倾斜摄影；多选择背包问题；OSG(OpenSceneGraph)

论文编号：382

应用双通道 Faster R-CNN 与注意力机制模型的图像篡改检测

陈晓铭　李斌*

（深圳大学 电子与信息工程学院 广东省智能信息重点实验室,广东 深圳 518060）

摘　要：基于深度学习的模型在图像篡改检测领域表现出良好的性能。然而，大多数基于深度学习的图像篡改检测方法采用通用的特征提取网络，这些网络不能有效地提取弱特征，而待检测的篡改信号通常非常微弱，所以这些网络对篡改图像的特征表达具有局限性。为了解决这个问题，本文提出了一种基于双通道Faster R-CNN的图像篡改检测模型，该模型利用注意力机制促使特征提取网络专注于提取图像的篡改区域。由于通道注意力模块和空间注意力模块分别关注“什么特征”更重要和“哪些位置的特征”更重要，所以结合了通道及注意力模块的特征提取网络能更好地区分篡改特征，并抑制无关特征的影响。同时，模型将RGB通道的信息以注意力机制的形式反馈给噪声通道，增强不同通道间的信息交互，从而提高了卷积神经网络对篡改特征的表达能力。此外，该模型采用双线性插值法对特征图进行缩放，避免了池化过程中的取整操作，减少了定位误差，从而改善了RoI池化层定位篡改区域的性能。本文在合成数据集和四个常用篡改图像数据集上进行了实验，实验结果表明，模型所采用的注意力模块有效提高了卷积神经网络对篡改图像特征的提取能力，并且基于双线性插值实现的RoI池化层更有利于对篡改区域的检测。综上所述，本文方法有效地提高了卷积神经网络对篡改图像的检测精度。

关键词：深度学习；图像篡改检测；Faster R-CNN；注意力机制；双通道网络

论文编号：436

基于对偶代偿网络的共融人脸对齐与超分

季则欣 1，祝洪宇 1，刘昊 1，2

1.宁夏大学 信息工程学院，宁夏银川 750021；

2.宁夏大数据与人工智能省部共建协同创新中心，宁夏银川 750021

摘 要 ：近年来，人脸对齐和图像超分辨率重建这两类任务一直受到广泛关注，本文提出了一种共融人脸对齐和图像超分辨率重建的对偶代偿网络，该方法旨在学习重建低质量人脸图像和精确定位人脸关键点任务间的互补信息。传统方法通常由于分开优化上述两类任务，容易导致模型陷入局部最优解。本文方法通过设计一种迭代共融的深度学习框架，进一步对超分辨率重建与人脸对齐两者间互补的特性进行建模。具体而言，本文设计了两个共融式的网络模块，分别为人脸对齐沙漏网络和超分辨率编码器-解码器网络。人脸对齐沙漏网络基于生出的超分图像定位人脸关键点的坐标，超分辨率编码器-解码器网络利用人脸对齐网络预测出的关键点坐标计算出的注意力掩膜恢复低分辨率图像上的人脸细节。由于这两个网络模块的执行具有对偶性和代偿性，使得这两个网络模块的性能得到相互提升。为了评估本文提出方法的有效性，本文在标准人脸数据集上进行了大量的实验，最后从定量、定性和消融三个方面分析得到的实验结果。实验结果表明，人脸对齐沙漏网络和超分辨率编码器-解码器网络是相互促进的，并且与当前方法相比性能具有明显提升，在视觉效果方面，在生成的超分图像具有更加清晰的纹理细节的同时具有更好的人脸对齐效果。

关键词：人脸对齐；人脸超分辨率；人脸去模糊；深度学习；多任务学习

论文编号：464

融合多模态注意力机制的视频概述技术

张维刚 1，张梦琴 1，周延森 2，黄庆明 3

1.哈尔滨工业大学(威海)计算机科学与技术学院,山东省威海市 264209

2.国际关系学院信息科技学院,北京市 100091

3.中国科学院大学计算机科学与技术学院,北京市 101408

摘　要：视频概述技术的目标是为一段视频自动生成一个较完整且通顺的内容描述语句。该任务不仅需要计算机视觉领域的知识来分析和理解视频的视觉信息，还需要自然语言处理领域的知识来将理解后的视觉信息转换为自然语句，因此极具挑战性。主流的视频概述方法主要是采用编码器-解码器框架，即编码器从视频中提取丰富的视频特征，然后解码器逐步将该视频特征解码为自然描述语句。为了更好地拟合编码器-解码器框架在视频特征与语义特征之间的鸿沟，本文提出一种融合多模态注意力机制的视频概述方法：在编码器阶段，通过多标签分类提取视频的先验类别属性，然后将视频的多类别属性特征与视频帧的全局视觉特征共同作为编码的视频特征；在解码阶段，融合视觉注意力与属性注意力机制为每个时间步的解码器挑选感兴趣的视觉特征与属性特征，为生成每个单词提供特定的视频特征，并且还将当前时间步挑选的属性特征与解码器的输出进行融合来生成更加准确的单词；此外，还建立基于语义注意力机制的视觉重建框架，反向训练将生成的概述语句重建为视觉特征的过程。本方法在两个标准数据集MSVD和MSR-VTT上的测试结果表明，所提技术可以提高编码器-解码器框架的性能，与当前主流的方法相比具有较强竞争力，说明本方法将多标签属性特征作为视频的先验特征，并融合多模态注意力机制来实现视频概述是切实有效的，能够充分利用原视频的特征来增强对其内容的准确概述。

关键词：视频概述；编码器-解码器；多模态注意力机制；多标签分类；视觉重建

论文编号：479

分级特征反馈融合的深度图像超分辨率

张帅勇 1,2，刘美琴 1,2，林春雨 1,2

1. 北京交通大学信息科学研究所，北京 100044；

2. 现代信息科学与网络技术北京市重点实验室，北京 100044

摘　要：在计算机视觉领域中，深度信息可以显著地提升远程医疗、无人驾驶、安防监控等产品的性能。然而受深度相机采集装置的限制，采集的深度图像存在分辨率较低、易受噪声干扰等问题，难以满足实际需求。因此，如何由低分辨率重建高分辨率的深度图像已经成为多媒体领域的研究热点之一。本文提出了一种基于分级特征反馈融合网络（Hierarchical Feature Backward Network，HFBN）的深度图像超分辨率重建算法。HFBN 网络由单深度图像重建网络和深度-纹理融合特征增强网络组成，利用单层卷积分别提取深度图像和纹理图像的浅层特征；级联残差块构建深层网络，分别提取深度图像和纹理图像的深层特征；利用并行金字塔结构构建深度图像和纹理图像的分层特征表示，获取不同尺度下的结构信息。为了提高不同尺度特征的相容性，本文设计了反馈式融合策略，实现金字塔不同层特征的反馈式融合。在此基础上，结合深度-纹理边缘特征，获得深度图像的边缘引导信息，生成边缘引导图像。采用残差学习的方式，边缘引导图像与单深度图像重建网络的重建结果相加，进一步提升深度图像的重建质量。实验结果表明，与state-of-the-arts方法相比，本文方法获得了主、客观质量均有提升的深度图像超分辨率重建结果。

关键词：深度图像；超分辨率重建；特征融合；残差学习；

分组报告 2

论文编号：102

基于 Res2-UNeXt 的医学图像分割

产思贤 1,4，黄诚 1，丁维龙 1，白琮 1,4，周小龙 1,2，陈胜勇 3

1.浙江工业大学，浙江杭州 320023； 2.衢州学院，浙江衢州 32400； 3.天津理工大学，天津 300000；

4. 浙江省可视媒体智能处理技术研究重点实验室 杭州 310023

摘　要：近年来，深度学习在医学图像分割领域中取得了很大的进展，无论是在分割的精度和模型的高效方面。为了提高模型的性能，目前常用的方法是增加网络的深度，但很少有将关注点放在提升网络模型的多尺度信息提取的能力。另一个提升模型性能的常用方法是通过人工标注的方式来增加训练图像数量，但该数据增强方法的成本昂贵。因此，本文针对以上问题，首先提出了一个多尺度的深度学习网络结构——Res2-UNeXt，一种编解码器类型的网络结构。为了能够更好的获取图像的多尺度信息，本文设计了网络的核心单元结构块Res2XBlocks ；然后，提出了一个简单且有效的医学图像数据增强方法，该方法是根据医学图像中细胞的移动以及形变来设计的，具有一定的生物意义；最后，选取ISBI cell tracking challenge 2019 中三种不同的细胞图像为数据集，分别与UNet++、CE-Net和LadderNet基于U-Net的变体进行医学图像分割的实验对比。 通过分析实验结果可知，本文的方法无论在分割的精度还是分割的视觉质量上看，都是要优于目前这几个比较流行的U-Net变体，主要是因为本文方法能够有效的针对了现有的多数模型无法很好的提取多尺度信息的问题和医学图像缺少标签（few-shot）问题。此外，本文设计的消融性实验进一步证明了本文所提出方法的合理性与有效性。

关键词：医学图像分割；数据增强；深度学习；图像处理

论文编号：145

一种具有光照鲁棒性的行人再识别方法

张培熙 1，赖剑煌 2,3，冯展祥 2，谢晓华 2,3

1.中山大学电子与信息工程学院，广州 510006； 2.中山大学数据科学与计算机学院，广州 510006；3.广东省信息安全技术重点实验室，广州 510006；

摘　要：行人再识别是计算机视觉领域中一项富有挑战性的任务，在智慧城市安防监控中有着重要的应用价值。近年来，尽管许多有效的行人再识别模型相继被提出，但是在实际应用场景中，这些模型未能很好地应对低光照问题，光照不足导致模型提取的行人外观特征显著退化，严重影响行人检索的准确度。本文针对该问题，提出了一种具有光照鲁棒性的行人再识别方法，该方法的主要创新由低光照增强网络构造和行人特征提取器联合调优方法两部分构成，低光照增强网络能够自适应地对不同的低光照图样本进行亮度提升，显著提升行人外观的清晰度，拉近与正常光照样本的分布距离；本文还采用了低光照增强网络和孪生行人特征提取器联合调优方法，通过协同学习的训练策略，模型能够充分利用低光照样本，对齐低光照样本和正常光照样本的特征，有效地提升模型的光照鲁棒性。本文通过充分的实验证明了该方法中各个模块的有效性，以及模型在低光照行人再识别领域具有出色的性能。

关键词：行人再识别；光照增强；协同学习；光照鲁棒性

论文编号：**160**

基于子集划分迭代投影集成的标签恢复算法

应晓清，刘浩

东华大学 信息科学与技术学院,上海 201620

摘　要：在图像特征提取中，样本标签通常并非是完全真实有效的。然而现有图像归类框架的执行效果在很大程度上取决于样本标签的质量，在噪声环境下的分类精度将大幅下降，这导致其在实际应用中可能受到诸多限制。针对含噪标签，目前虽已提出较多恢复算法可供参考，但现有标签恢复算法仍面临含噪样本难以高效再利用的瓶颈问题。为此，本文提出一种基于子集划分迭代投影集成的标签恢复算法，首先随机多次地提取小规模子集信息，其次综合主成分分析、邻域图正则化等技术进行样本图像的可靠降维与全局、局部数据结构的同时保留，从而构建更为准确可靠的低秩投影矩阵，然后通过K-近邻算法进行标签预估与迭代集成，最后遵循多数投票原则实现类别标签的有效复原。本文选取两大代表性的人脸数据库，说明了所提算法的最优参数取值区间，并对多种标签恢复算法进行了不同指标下的大量对比分析。实验结果证明，所提算法能够有效地校正样本的含噪标签，在同一图像归类框架下针对Yale B与AR数据库分别使分类精度提升了16.9%与8.4%，卡帕系数提升了0.173与0.082。相较于目前最好的标签恢复算法，所提子集划分迭代投影集成算法可以提升4.3% ~ 4.7%的分类精度与0.045~0.047的卡帕系数。此外，该算法原则上可灵活运用于各种图像归类框架，在确保样本数据完整性的同时具备了一定的可扩展性。

关键词：图像归类；特征提取；含噪标签；标签恢复；可靠降维

论文编号：168

基于 one-step-ahead 策略的重定向算法

刘淦 1，刘云秋 1，崔嘉 1，戚萌 1

1.山东师范大学，信息科学与工程学院，济南，山东，250014

摘　要：在虚拟现实中，如何在一个较小的跟踪空间中实现对较大虚拟空间的漫游是一个有趣且具有挑战性的问题。当人蒙上眼睛走直线时，其实际运动轨迹并非直线而是沿着一个大圆不停的绕圈。受这种运动错觉的启发，本文提出了一个新的重定向算法，该算法将虚拟空间中由直线段组成的路径映射为实际空间中的弧线路径，采用one-step-ahead策略，对用户下一步路线进行分析，并设置缓冲区，引导用户在快接近跟踪区域边界时进行转向，以减少用户碰撞边界的次数并提升用户的沉浸感。我们针对迷宫和无限长公路的情境进行了仿真模拟实验。实验结果表明，相比于传统的重定向算法，该算法能够有效减少碰撞次数和漫游用时，使整个漫游过程更加流畅。足迹分布图显示采用本文算法时用户的足迹在整个跟踪区域分布比较均匀，这表明我们的算法在解决多用户漫游问题时具有潜在的优势。此外，本文实现的虚拟漫游系统是基于cardboard的，用户可以用比较低的成本实现在较小跟踪区域中漫游较大虚拟场景的需求，具有一定的商业价值。

关键词：人机交互；虚拟现实；虚拟漫游；重定向算法

论文编号：208

视觉惯性里程计中异常视觉测量的检测与处理

朱　涛 1,3，马惠敏 2*，柴后青 3，张胜虎 1

1.清华大学电子工程系,北京 100084；　2.北京科技大学计算机与通信工程学院,北京 100083

3.中国人民解放军 31401 部队,长春 130022

摘　要：对于基于特征跟踪和非线性优化的视觉惯性里程计（VIO），紧耦合的数据融合方式可以短时解决特征跟踪丢失时的位姿估计问题。但是，视觉遮挡、运动物体等复杂场景干扰可能会带来异常的视觉测量，导致系统定位精度急剧下降甚至无法正常运行。对此，本文提出了一种新的视觉惯性里程计中异常视觉测量的检测和处理方法。首先，选取位置差异、特征点数量和IMU角速度作为检测指标。通过有无干扰场景下的对比实验分析，选取了各项指标的先验阈值。设计了异常视觉测量检测分类器，实现了对异常视觉测量的有效分类与检测。在此基础上，提出多传感器融合策略，根据是否存在异常视觉测量来选择不同的传感器方案。设计了自适应误差加权算法，通过在优化目标函数中增加视觉项的权重因子，减小可能的异常视觉测量对定位精度的影响，实现了对异常视觉测量的有效处理。最后，将异常视觉测量检测和处理算法整合到OKVIS系统中，提出了EDS-VIO系统框架。在复杂场景仿真数据集上的评测结果表明，EDS-VIO比OKVIS取得了更好的性能，定位误差均值从1.045m下降到0.437m，估计的轨迹更接近于真值，在处理较长时间的异常视觉测量时精度提升明显，全面验证了算法的有效性。本文提出的检测和处理异常视觉测量的方法可以应用到基于非线性优化的视觉惯性里程计算法设计中，进一步提升视觉惯性里程计在复杂场景中的定位精度和鲁棒性。

关键词：视觉惯性里程计；异常检测与处理；多传感器融合；异常视觉测量

论文编号：**248**

基于深度残差注意力网络的视频表情识别

陈港 1，2，张石清 2，赵小明 1，2

1. 浙江理工大学 机械自动控制学院，浙江 杭州 310018

2. 台州学院 智能信息处理研究所，浙江 台州 318000

摘　要：考虑到视频中的手工特征与主观性情感之间存在语义鸿沟的问题，识别视频序列中的人脸表情是一项挑战性的任务。针对此问题，提出一种基于深度残差注意力网络的视频表情识别方法。首先，考虑到人脸图像中各局部区域情感表示强度方面的差异性，提出采用一种基于空间注意力机制的深度残差注意力网络实现对视频中每帧表情图像进行高层次的人脸表情特征的学习，并采用平均池化计算得到固定长度的全局性视频特征参数。然后，将该全局性视频特征参数输入到一个多层感知器，用于实现表情的分类任务。在BAUM-1s和RML视频情感数据集上的实验测试结果表明，该方法分别取得了68.50%和56.72%的正确识别率，优于其他现有文献报道的结果。可见，本方法能够有效提高视频中人脸表情的识别性能。

关键词：深度残差注意力网络；注意力机制；表情识别；多层感知器

论文编号：**281**

双光流网络指导的视频目标检测

尉婉青 1，禹晶 1，史薪琪 1，肖创柏 1

1. 北京工业大学信息学部，北京 100124

摘　要：目前卷积神经网络广泛应用于目标检测中，静态图像目标检测旨在判断一幅图像中目标的类别及其位置，而视频目标检测的任务是在序列图像中定位运动目标，并为各个目标分配指定的类别标签。现有的大部分视频目标检测方法在静态图像目标检测的基础上，利用视频特有的时间相关性来解决运动目标遮挡、模糊等现象导致的漏检和误检问题。为了准确利用近邻帧的特征补偿当前帧的特征，本文提出了一种双光流网络指导的视频目标检测模型，在两阶段目标检测的框架下，对于不同间距的近邻帧，利用两种不同的光流网络估计光流场进行多帧图像特征融合。根据近邻帧与当前帧的间距，对于较近的近邻帧，利用小位移运动估计的光流网络估计光流场，对于较远的近邻帧，利用大位移运动估计的光流网络估计光流场。双光流网络能够更准确地估计帧间的光流场，进而能够更准确地从近邻帧补偿当前帧的特征。在ImageNet VID数据集上的实验结果表明，本文的模型利用视频特有的时间相关性，结合两阶段目标检测和双光流网络指导的时间特征融合提高了视频目标检测的准确率，较好地解决了视频目标检测中目标模糊和遮挡导致的漏检和误检的问题。

关键词：目标检测；卷积神经网络（CNN）；运动估计；光流网络

论文编号：306

基于高低分辨率图像融合的快速立体匹配算法

陈家鑫 1，赖剑煌 1，谢晓华 1，尹冬生 2

1.中山大学,广东省广州市 510000； 2.天棣网络科技（上海）有限公司,上海市 200000

摘　要： 双目视觉是计算机视觉领域的一个研究方向，利用左右两个摄像头同时拍摄得到的照片，可以恢复出画面的三维场景。立体匹配是双目视觉最为关键的两个环节之一，在双目视觉的整个流程中发挥重要的作用。随着测距距离的增加，双目视觉的测距误差在大幅度上升。为了解决这个问题，必须使用高分辨率图片来做立体匹配，而高分辨率图片的输入，会导致立体匹配更大的耗时，算法难以实际落地。针对立体匹配算法计算高分辨率图像耗时大的问题，提出了一种新的基于高低分辨率图像融合的快速立体匹配算法。该算法首先在匹配代价计算步骤使用高分辨率的图片，以获取准确的匹配代价矩阵。接着添加一个匹配代价矩阵下采样步骤，将图像进行下采样以降低后续计算复杂度。最后按照低分辨率图片处理流程完成代价聚合、视差计算或视差优化，以及视差求精步骤，以缩短这三个步骤的耗时。通过Middlebury公共评估平台的准确度实验，证实了本文算法相比于输入高分辨率图片，准确率平均下降了5.18%，相比于输入低分辨率图片，准确率平均提升了14.61%。通过全志H3开发板上的耗时实验，证实了本文算法的耗时仅为75ms，而输入高分辨率图片的耗时为220ms，输入低分辨率图片的耗时为50ms。本文算法可以在准确率和速率之间取得一个更好的平衡，有利于立体匹配算法的实际落地应用，特别是在嵌入式等低计算力设备上的应用。

关键词：双目视觉；立体匹配；高分辨率；匹配代价矩阵下采样

论文编号：340

基于极化相关方向图的极化SAR图像舰船检测

崔兴超，粟毅，陈思伟

国防科技大学电子信息系统复杂电磁环境效应国家重点实验室,湖南长沙 410073

摘　要：对于极化SAR数据，极化通道之间的相关值敏感于目标姿态及其与雷达传感器飞行方向的夹角。这种散射多样性给极化SAR散射解译带来困难，但同时又蕴含着丰富的极化信息值得进一步挖掘。为了探究不同夹角情况下极化相关值的取值变化，并进一步挖掘有效信息，本文提出极化相关方向图的概念，并导出新的极化特征并用于极化SAR舰船检测。其核心思想是将一定成像角度下的极化相关值扩展到绕雷达实现的极化旋转域，并由此导出包括原始相关值、极化相关最大值、最小值等在内的一系列新的极化特征。通过对新导出的极化特征进行对比度（target-to-clutter ratio, TCR）分析，优选TCR最高的三个极化特征，即$\left|\hat{\gamma}_{\text{(HH-VV)-(HV)}}(\theta)\right|_{\text{org}}$、$\left|\hat{\gamma}_{\text{HH-HV}}(\theta)\right|_{\text{org}}$和$\left|\hat{\gamma}_{\text{(HH-VV)-(HV)}}(\theta)\right|_{\min}$用于极化SAR舰船检测。本文所提方法主要包括三个步骤：特征优选、门限处理和形态学后处理。其中，门限处理是基于舰船样本选取的，可由此得到初始的舰船检测结果。形态学后处理通过剔除像素点少的小区域来抑制虚警，由此得到最终的检测结果。基于Radarsat-2和高分三号全极化数据开展对比实验，本文所提方法对比于其他方法有更好的检测性能，特别是对于舰船密集区域。

关键词 ：极化SAR 极化相关方向图 旋转域 舰船检测

论文编号：420

基于单眼视觉和超声波的室内障碍物检测系统在智能机器人车上的应用

马连杰 1，唐明 1，李明法 1，李媛媛 1

上海工程技术大学,上海市 201620

摘　要：随着智能机器人技术的发展，室内移动机器人技术受到广泛关注，其中移动机器人在复杂的室内环境中准确识别障碍物是一项关键技术。在室内障碍物检测技术中，基于视觉的移动机器人自主导航是室内模式识别研究热点。地面障碍物的检测和距离测量是机器人自主行走的关键，也是后续避障和轨迹规划的基础。针对室内复杂环境中的障碍物检测和距离测量，提出了一种基于单眼视觉和超声波测距的Raspberry Pi智能机器人车检测静态和动态障碍物的方法。该方法采用改进的单眼视觉障碍物检测算法，对室内的静态和动态障碍物进行轮廓检测，并利用超声波测距传感器在静态和动态障碍物检测过程中进行实时测距。与大多数工作不同，在静态障碍物检测系统中，本文在预处理阶段添加直方图均衡化图像增强功能，使用HSV图像变换以减少复杂因素（例如照明）的影响。根据Canny边缘检测算法，对障碍物的边缘特征进行检测，最后对障碍物轮廓进行标记，以实现对静态障碍物的检测。在动态障碍物检测系统中，通过将背景差分法与3D图像显示相结合以检测动态障碍物。实验结果表明，所提方法可有效减少障碍物检测的平均消耗时间以及障碍物位置信息的错误率，并在将来实现移动机器人的路线规划。

关键词：单眼视觉；超声波传感器；室内障碍物检测；图像增强；背景差分

分组报告 3

论文编号：9

基于多 GPU 的高精度烟雾模拟并行求解方法的研究

杜金莲 1，蒋博 2，金雪云 3

1.北京工业大学计算机学院,北京 100124； 2. 北京工业大学计算机学院,北京 100124

摘　要：高精度、真实感、实时性的烟雾模拟在影视、动画以及仿真领域中应用很多。本文在目前基于物理的烟雾模拟算法研究成果的基础上，针对规模大、精度要求高、实时性要求高的需求，结合计算着色器的计算架构，设计了一种基于多GPU的分布式烟雾流场并行模拟算法。首先，通过计算中断处理及中间变量转存机制实现流场空间的状态同步，有效解决了计算流程中烟雾流场阶段性同步问题保证了烟雾流场模拟状态的正确性。其次，基于重叠分割和同步控制机制，针对超出单个GPU的计算能力时的高精度、大规模烟雾模拟计算问题，提出了基于多GPU协同计算的烟雾流场模拟算法框架，该算法可对烟雾流场模拟进行自适应的分割和任务分配，并通过边界数据同步机制保证分割后每个流场子域中处于分割边界的重叠数据在多GPU协同计算环境下的一致性，从而为精度高、规模大的烟雾流场模拟提供高实时性解决方案。该算法在烟雾流场模拟上具有一定的通用性，即使对于大规模流场只有一个GPU的情况下，同样可以通过循环求解实现高精度大规模烟雾模拟。最后，通过对不同初始条件下的烟雾数据对比、时间效率对比和扩散效果对比，验证本文提出的烟雾求解算法的正确性和在模拟规模大的烟雾流场时效率性。

关键词：烟雾模拟；N-S 方程；GPU 计算；大规模

论文编号：22

一种基于汉明码、图像矫正和特征值分解的彩色图像盲水印方法

刘得成 1，苏庆堂 1，袁子涵 1，张雪婷 1

1.鲁东大学,山东省烟台市 264025

摘　要：随着互联网技术的飞速发展，作为大信息量载体的彩色数字图像在给人们带来了极大便利的同时，也带来了网络上一些篡改、剽窃等的侵权行为；同时，图像几何处理对载体中所含水印的破坏使得盲检测水印的难度增加，如何设计一种水印容量大、隐蔽性好、安全性高、鲁棒性强的彩色图像数字水印算法成为目前亟待解决的版权保护问题之一，故本文设计了一种基于汉明码、图像矫正和特征值分解的彩色图像盲水印方法。嵌入水印时，使用仿射变换加密彩色水印信息，并借助汉明码理论将已加密的信息编为汉明码，然后利用特征值分解计算出彩色宿主图像中像素块的全部特征值，并通过量化特征值绝对值之和来完成水印的嵌入；提取水印时，利用含水印图像的几何属性对多种几何攻击后的含水印图像进行矫正，并借助量化技术提取水印。衡量水印不可见性的PSNR达到40dB以上，多种攻击后的NC接近于1，度量安全性的密钥空间为2432，最大水印嵌入容量为0.25bit/pixel，平均运行时间为19.36秒。实验结果表明，本方法不仅具有较好的水印不可见性和较大的水印嵌入容量，而且具有较强的鲁棒性和较高的安全性。

关键词：彩色水印技术；汉明码；图像矫正；仿射变换；特征值分解

论文编号：**27**

基于宏块复杂度的视频隐写算法

唐洪琼[1]，杨晓元[1,2*]，钮可[1]，张英男[1]

1. 武警工程大学,西安 710000； 2. 中国人民武装警察部队网络与信息安全重点实验室,西安 710000

摘　要：基于运动矢量的视频隐写算法会引起视频质量的下降和比特率的增加，还会破坏运动矢量的局部最优性和相关性，可能导致隐蔽通信事实的暴露。如何选取合适的运动矢量作为信息嵌入的载体对隐写算法的性能具有显著影响，因此研究隐写载体的选择具有重要意义。本文根据视频编码的特点，形式化地分析了运动矢量修改对视频质量与局部最优性的影响，指出基于运动矢量修改的隐写算法选择平滑宏块作为隐写载体能更好的保持局部最优性。而后，引入Hilbert曲线来衡量重建宏块的复杂度，根据嵌入需求选取复杂度阈值，小于阈值的宏块的运动矢量作为隐写载体，对运动矢量进行随机LSB匹配修改，提出一种基于宏块复杂度的隐写算法。通过与现有基于运动矢量的隐写算法的对比，本文提出的隐写算法在不可见性、比特率变化、抗隐写分析能力指标具有一定优势。整体来看，提出的视频隐写算法具有很高的视频质量，比特率增长小，较好的抗隐写分析能力。

关键词：运动矢量；视频隐写；Hilbert 曲线；宏块复杂度；隐写分析

论文编号：**193**

一种将渲染误差高效地转化为屏幕空间蓝噪声的过滤方法

刘泓立，韩红雷

中国传媒大学 动画与数字艺术学院, 北京, 100024

摘　要：基于蒙特卡洛的积分器是最为常用的辐射度近似方法。在屏幕空间中将蒙特卡洛积分的误差分布为蓝噪声，可以显著提升视觉质量。然而，在任意样本数和维度下，对于复杂场景，快速地将积分噪声分散为蓝噪声还难以实现。我们首先使用蓝噪声种子扰乱积分的采样序列，使积分在低采样时具备基本的蓝噪声特性。然后，对渲染结果引入一个有偏的过滤，使得渲染结果在更高维度也基本保持蓝噪声的视觉效果。实验对比了多个最新相关工作，本文的方法在多种不同类型的场景中，都可以快速得到更加稳定的蓝噪声分布的误差。特别地，在视觉信息相对低频变化的场景中，本文的方法比其他方法能够产生更好的蓝噪声特征。该方法不仅在任意样本数和维度的情况下能够基本保持蓝噪声特征，而且有很高的计算效率。虽然引入了偏差，但是该方法易于实现，计算成本很低，适用于现代基于蒙特卡洛方法的渲染器中。

关键词：蒙特卡洛；采样；蓝噪声；渲染

论文编号：**203**

基于径向曲线的颅面统计复原

王琳 1，赵俊莉 1，黄瑞坤 1，李淑娴 2，李守哲 3

1.青岛大学数据科学与软件工程学院，青岛 266000； 2.大连大学机械工程学院，大连 116622

3.威斯康星大学麦迪逊分校文理学院，美国威斯康星州麦迪逊 53706-1380

摘 要：颅面复原是一项依据人的颅骨恢复其对应面貌的研究，在法医学、人类学、公安刑侦等领域都有重要的应用价值。针对复杂的颅面形态，本文依据其几何结构，构建了一种基于径向曲线的颅面统计复原模型，通过由这种模型得到的先验知识联合待复原颅骨数据,来复原颅骨面貌。首先，通过在颅面模型上定位鼻尖点，求取过鼻尖点的颅面对称面，从而确定第一条径向曲线，然后将颅面对称面依次旋转相等的角度间隔提取所有的径向曲线，并进行重采样后得到颅面的径向曲线表示。接着，使用颅骨和对应面皮的径向曲线作为训练样本，建立基于径向曲线的 PCA 统计模型，复原出未知颅骨对应人脸的径向曲线；最后，根据训练人脸数据建立人脸统计模型,通过复原出的径向曲线估计人脸统计模型参数 ,进而估计出待复原颅骨对应的面貌。在 192 套颅面数据上分别采用本文提出的基于径向曲线的颅面统计复原方法和经典的 PCA（主成分分析）统计复原方法进行实验，利用欧氏距离计算复原面貌与原始面貌的误差，结果显示基于径向曲线的颅面统计复原方法复原面貌误差更小，该方法降低了模型中样本数据的维度，提高了颅面复原的精度和速度，复原效果得到进一步改善。

关键词：颅面复原；三维颅面模型；径向曲线；主成分分析(PCA)；统计模型

论文编号：217

IIH-GAN:基于生成对抗网络的图像信息隐藏

戈辉 1，胡东辉 1，郑钢 1，郑淑丽 1

合肥工业大学计算机与信息学院,安徽合肥 230009

摘　要：隐写术是一种将秘密信息隐藏在载体（如数字图像）中的技术。传统的隐写算法依赖于人工设计的复杂特征,需要专业的领域知识。近年来,研究者们试图利用生成对抗网络(Generative Adversarial Nets，GANs）进行隐写算法设计，但所设计的算法在隐写信息提取准确率、隐写图像质量、隐写安全性、实用性等方面存在不足。本文提出了一种基于GANs的新型端到端隐写系统，设计基于SE-ResNet块的编码器和解码器，通过优化网络通道之间的相互依赖性及提高全局特征自动选择的能力，实现更加准确、高质量的信息嵌入和提取；并通过判别器与编码器-解码器的对抗训练，保持载体图像的分布不改变，提高隐写对抗安全性。为解决对真实场景中的真实图像进行解码的问题，我们首次在编码器和解码器之间添加了一个四舍五入层。为了弥补基于GANs的隐写在抵抗目前强大的基于深度学习的隐写分析算法检测的不足，我们首次将对抗样本添加到基于GANs的隐写模型中。实验结果证明所设计方法相对当前方法在性能方面具有较大改善；并且当添加对抗样本时，本文算法被基于深度学习的隐写分析模型检测出来的准确率从97.43％降低到48.69％，从而极大提高了隐写安全性。

关键词：隐写术；图像信息隐藏；生成对抗网络；对抗样本

论文编号：270

基于迭代对抗攻击的 JPEG 隐写改进

宋婷婷，莫华晓，陈勃霖，骆伟祺

中山大学广东省信息安全重点实验室,广东省广州市 510006

摘　要：近几年，大量研究报告表明基于神经网络的图像隐写分析要优于基于人为设计特征的隐写分析方法，但是这些基于深度网络的方法往往容易受到对抗样本的攻击而导致其性能大幅度下降。据此特性，本文设计了一种基于迭代对抗攻击的JPEG图像隐写，在我们原有工作JS-IAE的基础上提出了几点改进。首先，为提升模型的收敛速度，我们有效地结合了MinMax策略以选择每一轮迭代的载密图像。此外，通过大量的比较实验分析，我们提出了一种新的代价更新函数以及修改单元的选择策略，使得隐写倾向于图像更安全区域并且修改元素更少。实验表明：此方法的JPEG隐写安全性达到目前国际领先水平。如在质量因子为75情况下，我们可使得J-UNIWARD隐写在分析器SRNet的检测率从85%降至70%，安全性提升了近15%，在质量因子为95的情况下也提升了10%以上。与我们原有工作JS-IAE相比，本文隐写算法能在修改更少元素的前提下提升了其安全性能，同时加快了收敛速度。在质量因子为75时，我们的方法在分析器GFR和SRNet上的准确率分别降低了5%和6%；在质量因子为95时，检测准确率下降的幅度则相对较少。

关键词：JPEG 隐写；JPEG 隐写分析；对抗样本；迭代对抗

论文编号：**320**

基于跨尺度字典学习的图像盲解卷积算法

彭天奇 1，禹晶 1，郭乐宁 1，肖创柏 1

1.北京工业大学,北京 100124

摘　要：模糊是一种常见的图像降质现象，在数字成像设备采集图像的过程中，拍摄环境差、设备的抖动、散焦以及被拍摄场景本身的瞬时变化都会导致模糊的产生，它给图像的实际应用带来了很大的挑战。图像解卷积算法研究的是如何从模糊图像中复原出原始的清晰图像，而在模糊核未知情况下利用模糊图像对清晰图像进行复原称为图像盲解卷积问题，这是一个欠定逆问题，现有的大部分算法通过引入模糊核和清晰图像的先验知识来约束问题的解空间。本文提出了一种基于跨尺度字典学习的图像盲解卷积算法，采用降采样图像训练稀疏表示的字典，并将图像纹理区域在该字典下的稀疏表示作为正则化约束引入盲解卷积目标函数中。图像降采样过程减弱了图像的模糊程度，且图像中存在冗余的跨尺度相似块，利用更清晰的图像块训练字典能够更好地对清晰图像进行稀疏表示，减小稀疏表示误差；同时，由于在纹理区域清晰图像的稀疏表示误差小于模糊图像的稀疏表示误差，在该字典下对图像中的纹理块进行稀疏表示，使重建图像偏向清晰图像。在大量模糊图像上的实验验证了本文的算法能够有效解决大尺寸模糊核的复原，并具有良好的鲁棒性。

关键词 ：盲解卷积；稀疏表示；字典学习；跨尺度；自相似性

分组报告 4

论文编号：59

基于混合高阶图卷积的半监督分类

雷方元 1，刘勋 2，戴青云 1，李争名 3

1．广东省知识产权大数据重点实验室，广州 510665

2．广东技术师范大学电子与信息学院，广州 510665

3．广东技术师范大学工业中心，广州 510665

摘　要：图结构网络作为一种典型的非欧式空间数据集，普遍存在于社会的各个层面，如电子商务、社交网络等。从复杂结构的网络中挖掘出网络中隐含的内在规律，对于分析研究网络行为和网络节点之间的隐形联系有重要的意义和应用前景。图卷积神经网络是卷积神经网络在非规则的图网络中的推广。信号在图网络中的传递受到网络节点的邻域节点的影响，通过捕获更多的邻域信息来扩大感受野将有助于提高分类精度。图网络中的高阶邻域信息是扩大感受野的一种方式，而目前的高阶图卷积方法存在参数量大，计算复杂度较高等问题。为此，提出一种混合高阶邻域的图卷积半监督分类模型(HGCN)，该模型利用高低阶权重共享机制来降低模型的网络参数量，同时，提出了两种全新的融合池化层HGCNM和HGCNA来实现高低阶邻域信息融合，降低计算复杂度，保持网络拓扑结构。在模型构建过程中分析研究了图卷积网络层数带来的影响。从理论上，对比分析了HGCN与当前最先进模型的模型复杂度。最后在大规模引用网络数据集上对HGCNM模型和HGCNA模型及其多种高阶混合邻域分类进行了验证。结果表明，HGCN与经典GCN模型的计算复杂度和训练参数量相当，分类精度平均提高1%；与Mixhop高阶方法相比，HGCN的待训练参数量降低60-80%，计算复杂度降低30-40%。未来的研究将关注于网络层数过多带来过平滑问题。

关键词：图卷积；权重共享；信息融合池化；混合高阶模型；半监督学习

论文编号：95

基于 D-UNet 的遥感图像分割

龙丽红 1，朱宇霆 1，闫敬文 1，刘敬瑾 1，王宗跃 2

1.汕头大学工学院电子系,广东省汕头市 515063； 2. 集美大学计算机工程学院,福建省厦门市 361021

摘　要：高分辨率遥感图像语义分割，尤其是建筑物分割，在航空图像分析领域具有重要的理论和应用价值。近年来深度学习理论及应用取得重要进展，尤其是 FCN 实现了像素级的语义分割，取得较好的效果。但是，由于高分辨率遥感图像中楼房语义的丰富性和图像背景的复杂性，以往的分割方法往往产生边缘模糊、细节信息丢失和分辨率低等缺点，高精度的楼房分割仍然是一个具有挑战性的问题。为了解决高分辨率卫星图像语义分割边界模糊和信息丢失的问题，本文提出一种端到端的卷积神经网络 Dilated-UNet（D-UNet）。首先通过改进 U-Net 网络结构，采用 Dilation 技术拓展四通道的多尺度空洞卷积模块，每个通道分别采用不同的卷积扩张率来识别多尺度语义信息，从而提取更丰富的细节信息。其次设计了一种交叉熵和 Dice 系数的联合损失函数，更好的训练模型以达到预期分割效果。最后，在 Inria 航空图像数据集上对模型进行综合实验评估与检验，同时，为了充分调查模型中空洞卷积模块的效果，本文进行有无空洞卷积模块的消融实验，最终筛选出最有效的四通道空洞卷积结构。实验结果表明提出的遥感图像分割方法能够准确提取高级语义信息，有效地从高分辨率遥感图像中进行像素级的城市建筑物分割，分割精度更高，优于其它方法，具有较高的实际应用价值。

关键词 ： 深度学习；语义分割；遥感图像；图像处理

论文编号：141

一种基于人脸和人体特征融合的漫画人物身份识别方法

程海杰 1，赖剑煌 2,3，谢晓华 2,3

1.中山大学电子与信息工程学院，广州 510006； 2.中山大学数据科学与计算机学院，广州 510006；

3.广东省信息安全技术重点实验室，广州 510006；

摘　要：近年来，随着深度学习的兴起，识别系统的性能得到了飞速的提升。但是，已有的识别算法大都是围绕真实人物图像进行研究的，其性能在一定程度上也很难再提升，很少有专门的算法针对漫画人物图像进行研究。当前，已有的漫画人物身份识别方法往往是简单的套用人脸识别的方法，仅仅利用单一人脸特征对漫画人物的身份进行识别。为此，本文提出了一种基于人脸与人体特征融合的漫画人物身份识别方法，该方法首先通过人脸检测对齐得到对应人物图像的人脸图像，而后通过双流网络综合利用人脸和人体的特征对漫画人物的身份进行识别，这样不仅可以充分利用图像的有用信息，也能进一步提升漫画人物身份识别的精确度。同时，针对当前公开的漫画数据集不仅少而且小，难以充分验证当前漫画识别方法的有效性，本文提出了一个新的、更大的真实人物图像和漫画人物图像对应的数据集，记为TC数据（True Caricature Dataset），并在WebCaricature数据集和TC数据集上进行了实验。实验结果表明，该方法不仅可以充分利用图像的有用信息，更能辅助解决漫画人物身份识别风格差异大的困难，特别是在漫画人物与真实人物面部相差较大的情况下，提升效果更为明显。

关键词：人脸特征；人体特征；特征融合；漫画人物身份识别

论文编号：185

基于对齐特征表示的跨模态人脸识别

王绍颖，范春晓，明悦

北京邮电大学电子工程学院，北京 100876

摘　要：跨模态人脸识别一直是人脸识别领域的研究热点，在安防、刑侦等现实场景中具有极高的应用价值和发展潜力。现有的跨模态人脸识别算法通常在图像空间或潜在空间建立不同模态人脸的联系，却忽略了二者的内在关联性，容易导致跨模态信息的丢失。针对这一问题，本文提出基于对齐特征表示的跨模态人脸识别算法（Cross-Domain Representation Alignment，CDRA）。CDRA算法在人脸图像空间和潜在空间、模态内和模态间探索不同模态人脸数据间的关联性：首先，为减少信息损失，CDRA算法通过对单一模态内人脸的重建，学习到包含判别信息的模态内潜在特征表示；然后，在图像空间，CDRA算法通过从不同模态的潜在特征表示中跨模态重建图像间接对齐不同模态的潜在特征表示，在潜在空间，CDRA算法通过对齐不同模态数据的潜在高斯分布直接对齐不同模态的潜在特征表示。因此，在图像空间和分布空间同时进行对齐的CDRA算法，能够获得不同模态人脸间多个空间维度中的不同层次的关联关系，有利于提取到更具判别能力的跨模态关联信息。CDRA算法在跨模态人脸数据集Multi-Pie和CASIA NIR-VIS 2.0上进行跨模态人脸识别实验，实验结果表明，CDRA算法能够有效提高人脸识别的准确率，并且适用于不同的跨模态人脸识别任务。

关键词：跨模态人脸识别；变分自动编码器；人脸合成；潜在子空间

论文编号：189

通过风格迁移的雨雾天气条件下无人机图像目标检测方法

殷旭平 1，钟平 1，薛伟 1，2

1.国防科技大学电子科学学院 ATR 重点实验室，长沙 410073 ；2.安徽工业大学计算机科学与技术学院，马鞍山 243032

摘　要：雨雾恶劣天气条件下，无人机获取的图像细节丢失严重，直接对降质后的图像进行目标检可能导致目标漏检和错检问题。传统的先去雾后检测的方法，虽然可以在一定程度上提高目标检测模型的性能，但该方法过度依赖图像中的内容信息，稳定性较差。而风格迁移化的方法可以依靠风格变换，降低内容变换对目标检测模型的影响，同时有效地增强图像中目标物的轮廓信息，使目标检测模型使从学习目标物纹理、颜色转变为学习目标物轮廓这一更高难度的任务。因此提出采用图像风格迁移化的方法，完成雨雾条件下的无人机图像目标检测任务。本文在无人机数据集Visdrone2019上进行实验，首先生成18种不同程度的雨雾图像及32种的不同风格图像；其次分析不同程度雨雾浓度以及不同迁移风格对图像目标检测性能的影响，并挑选出最优的迁移风格；最后和传统的雨雾图像目标检测方法进行对比。实验结果表明，相比传统方法，通过风格迁移的目标检测方法对无人机雨雾图像目标检测的鲁棒性更高，整体性能更优。

关键词 ：无人机图像；雨雾降质图像；风格迁移；目标检测

论文编号：205

18F-FDG PET/CT 全身显像肺部肿瘤诊断：划分样本空间下的集成 E-ResNet-NRC 模型

霍兵强 1，周涛 1,3，陆惠玲 2，董雅丽 1

1.北方民族大学 计算机科学与工程学院 宁夏 银川 750021

2.宁夏医科大学 理学院 宁夏 银川 750004

3.宁夏智能信息与大数据处理重点实验室 宁夏 银川 750021

摘 要：在18F-FDG PET/CT多模态全身显像肺部肿瘤诊断的背景下，针对卷积神经网络训练时的网络退化、特征维度较高等问题，结合不同模态医学图像PET/CT在样本空间中的差异性，从划分样本空间的角度出发，提出基于集成非负表示分类的残差神经网络肺部肿瘤的识别模型E-ResNet-NRC。该模型包括：1)使用参数迁移法将预训练残差神经网络模型初始化参数；2)根据多模态医学图像PET/CT在样本空间中的差异性把样本划分成三个不同的样本空间(CT、PET、PET/CT)，提取病灶ROI区域；构建三模态数据集(CT-ROI、PET-ROI、PET/CT-ROI)；3)采用残差神经网络ResNet模型提取ROI区域特征，得到特征向量；4)在全连接层采用非负表示NRC构造个体分类器ResNet-NRC；5)采用“相对多数投票法”得到集成分类器E-ResNet-NRC。最后通过AlexNet+NNC、AlexNet+Softmax、AlexNet+NRC、ResNet-50+NNC、ResNet-50+Softmax、ResNet-50+NRC、E-AlexNet+NNC、E-AlexNet+Softmax、AlexNet+NRC、E-AlexNet+NRC、E-ResNet-50+NNC、E-ResNet-50+Softmax、E-ResNet-50+NRC共12种模型进行五折交叉对比实验，实验结果表明，集成E-ResNet-NRC模型在精确度为99.57%，模型整体分类性能优于个体分类器，且特异性、灵敏度、F值和MCC等评价指标也较高，该方法具有较好的鲁棒性和泛化能力。

关键词：残差神经网络；PET/CT；多模态医学图像；集成学习；非负表示分类

论文编号：318

基于梯度预测网络的雷达回波图像外推研究

张晋文 1，李腾 1，闫庆 1，苗开超 2，张亚力 2，汪翔 2

1.安徽大学,合肥 230601； 2. 安徽省公共气象服务中心,合肥 230031

摘　要：精准的雷达回波图像外推技术可以提供后面2小时内的精确雷达回波图像。根据这些外推的雷达回波强度，我们可以预计区域降水强度，来预测路况、发布强降水预警等。然而，现有基于循环神经网络的大多数检测算法因其相对简单的状态转换函数，难以学习雷达回波变化中过于复杂的情形。本文提出了一个新的时空预测学习的递归网络——GhuLSTM，用于外推预测雷达回波。我们利用时空记忆单元，同时提取和传递雷达图像空间和时序上的表征。在已提取的高层特征信息中，我们引入时空记忆单元变体，以利用相邻的递归状态之间的差分信号处理雷达回波演变的高阶非平稳性。此外，为了缓解雷达图序列外推中长时序梯度的消失，网络中加入了梯度公路。对比韩等提出的模型PredRNN，我们的模型在均方误差mse上效果提升了18.9%，在绝对误差mae上效果提升了2.8%，在峰值信噪比PSNR上效果提升了5.5%。我们还做了消融实验来各自验证了引入的两个结构在GhuLSTM中的重要性。这些实验结果都表明了我们所提出的网络在雷达回波外推上应用上的有效性。这证实了我们所提出的网络模型有效拓展了基于深度学习方法在雷达回波图像外推上的应用前景。

关键词：图像序列；雷达回波图像；外推预测；深度学习

论文编号：386

基于注意力机制的多尺度车辆行人检测算法

李经宇 1,2,3,孔斌 1,3，杨静 1,3，王灿 1,3，张露 1,2,3,

1.中国科学院合肥智能机械研究所，合肥　230031，中国；2.中国科学技术大学，合肥　230026，中国；

3.安徽省仿生感知与先进机器人技术重点实验室，合肥　230031，中国

摘　要：近年来由于深度学习的发展，无人驾驶汽车受到越来越多的关注。在复杂多变的交通场景中能提前且准确检测到路面车辆行人的动态信息尤为重要。然而，目前通过在无人驾驶平台上运行的目标检测进行实时准确的场景解析仍具有很大的挑战。无人驾驶场景下存在相机快速运动、尺度变化大、目标遮挡和光照变化等问题。为了应对这些挑战，本文提出一种基于注意力机制的多尺度目标检测算法。在YOLOv3网络的基础上，通过添加改进的SPP（空间金字塔池化）模块，对多尺度局部区域特征进行融合和拼接，使网络能够更全面地学习目标特征；其次，利用空间金字塔缩短通道间的信息融合，构造了YOLOv3-SPP+-PAN网络；最终，得到了基于注意力机制的高效的目标检测器SE-YOLOv3-SPP+-PAN。在KITTI实验数据集下，SE-YOLOv3-SPP+-PAN网络车辆行人检测结果为86.2%mAP比YOLOv3网络提升了2.2%mAP，然而推理速度仍然保持智能驾驶平台下实时的要求。证明了所提出的SE-YOLOv3-SPP+-PAN网络比YOLOv3更高效、更准确，因此更适合于智能驾驶场景下的目标检测。

关键词 ：车辆行人检测；计算机视觉；神经网络；注意力机制

论文编号：**432**

场景视点偏移改善点云分割

郑阳，廖康，林春雨，赵耀

北京交通大学信息科学研究所,北京市 100093

摘　要：当前，由于成熟的网络架构的应用以及大量的训练数据标注，3D 激光雷达点云分割方法已取得了较好的分割结果。点云学习任务主要分为直接方法和变换方法，直接处理方法会减少信息损失，但是通常计算量大，处理复杂；变换方法更符合人们认知，也更容易与现有的2D检测网络结合，但是投影过程中不可避免会产生大量信息丢失。目前的很多方法忽略了由激光雷达成像特点以及点云变换引起的信息丢失和遮挡问题，这导致分割性能降低，尤其是在小样本数据以及行人和骑行者等小物体场景中。受人类观察机制的启发，本文提出了一种场景视点偏移方法，以改善3D激光雷达点云分割结果。首先，本文利用球面投影将3D点云转换为2D球面正视图（SFV）。然后，水平移动SFV的原始视点以生成多视点序列，进而解决点云变换引起的信息丢失和遮挡的问题。最后，考虑到多视图序列中的冗余，本文提出了一种场景视点偏移预测模块来预测最佳场景视点移位。与经典点云分割算法相比，本文的方法在“行人”和“骑行者”的分割结果上有较大提高，且易于适配不同点云分割算法。本文的方法在较小的数据集上对于各个类别的分割性能均有一定程度的提升，在与SqueezeSeg相同的数据集上也在一些类别上的分割结果上也有一定程度的改善。

关键词：点云分割；球面正视图；场景视点偏移；场景视点偏移预测

论文编号：452

基于空间自适应 EPLL 的沙漠地震图像去噪

林红波 1　马阳 1　邓小英 2

（1.吉林大学通信工程学院,吉林省长春市 130012；2.北京理工大学信息与电子学院,北京市 100081）

摘 要：在沙漠地区采集的地震勘探图像中，非平稳地震勘探信号的强度在时间和空间上剧烈变化，弱的有效反射信号在强信号和弱相似随机噪声的干扰下很难被辨识。期望块对数似然 EPLL（Expected Patch Log Likehood，EPLL）去噪算法利用学习外部干净图像结构特征作为先验约束降噪问题，利用与噪声方差有关的正则化参数平衡局部细节保持和图像全局特征的恢复，但在处理非平稳地震信号时易造成弱信号的损失。针对地震勘探信号非平稳问题，本文在 EPLL 框架下提出空间自适应 EPLL（SA-EPLL）方法。该方法利用图像块信噪比调节正则化参数，使其随地震勘探信号强度的时空变化而自适应变化，从而对包含强信号的图像块设置较大的正则化参数，以保留更多信号的细节信息；对弱反射信号图像块设置较小的正则化参数，以滤除更多的弱相似随机噪声。此外，本文采用高斯混合模型（Gaussian mixture model，GMM）学习地震勘探图像的纹理特征，实现弱相似随机噪声干扰下的图像块恢复，并利用正则化参数对图像块加权平均重构信号。合成和实际地震勘探图像去噪结果表明， SA-EPLL 方法在抑制非平稳图像中的随机噪声方面优于其它竞争性降噪方法，这种学习的先验增强了对复杂形态地震勘探信号的恢复能力，减少非平稳信号失真。为沙漠地震图像降噪提供了一种可靠的方法。

关键词：EPLL; 随机噪声压制; 非平稳信号；沙漠地震勘探图像；块信噪比

墙报 1-1

论文编号：15

一种基于空间编码结构光的稠密三维重建算法

陈荣 1，许宏丽 1,2，杨东学 1，黄华 1,3

1.北京交通大学计算机与信息技术学院，北京 100044;

2.轨道工程北京市重点实验室，北京 100044;

3.交通数据分析与挖掘北京市重点实验室，北京 100044

摘　要：基于主动立体技术的三维重建方法，由于具有非接触、精度高、鲁棒等优点，已经被广泛应用于许多实际三维扫描系统中。其中，空间编码结构光技术只需要单幅结构光编码图案就可以获取被测物体表面的三维数据，因此适合用于动态场景和快速运动目标的三维重建。在本文中，我们基于空间编码结构光提出了一种新颖、简单、快速的稠密三维重建方法。首先，针对彩色条纹不适用于带颜色物体的问题，提出了一种由正弦条纹和伪随机点组合而成的结构光编码模式，所有的条纹采用相同颜色进行编码，通过引入伪随机点来解决条纹识别问题；其次，针对传统条纹方法得到的点云空间分辨率低的问题，提出了一种基于局部相位匹配的稠密三维重建方法，先基于条纹中心点提取算法，准确提取条纹的中心坐标，再利用傅立叶变换获取条纹间的相位信息，最后对相邻条纹间的所有像素点进行匹配，实现稠密三维重建。实验结果表明，在被测物体表面结构复杂的条件下，本文提出的算法依然能够有效地获取物体表面的三维信息，且具有较好的精度和较高的分辨率。

关键词：三维重建；结构光；投影仪-相机系统；傅立叶变换

论文编号：174

AV-GAN：基于 Arousal-Valence 情感模型的表情图像生成及其在表情识别中的应用

杨静波 1，赵启军 1，2，吕泽均 1

1.四川大学计算机学院,四川省成都市 610065； 2.西藏大学信息科学技术学院，西藏自治区拉萨市 850000

摘　要：基于深度学习的人脸表情识别方法的准确率有了显著提升，但是仍然受到非均衡训练数据的限制。一方面，现有方法一般将表情分成离散的若干个类别，忽略了人类情绪表达的连续性和复杂性；另一方面，训练数据覆盖的表情类别有限，而且不同表情下的数据规模往往也不平衡。为了解决这一问题，本文提出了Arousal-Valence（AV）空间中基于生成对抗网络（GAN）的表情图像生成方法AV-GAN，用于生成更多样且均衡的表情识别训练数据。该方法在训练时将表情标签（V, A）转化为标记分布形式，使用标记分布表示表情图像，通过引入身份控制和表情控制模块分别对生成图像的身份和表情进行控制，在对抗学习方法的基础上实现在AV空间中随机采样和生成表情图像。本文实现了在小规模数据库中的应用，生成了较高分辨率的图像，并且进行了生成图像视觉效果的对比实验进行验证。同时在Oulu-CASIA数据库上的评估实验也证明了本文方法能有效提升非均衡训练数据下的表情识别准确率。

关键词：Arousal-Valence 情感模型；生成对抗网络；数据增强；人脸表情识别

论文编号：**224**

基于多尺度目标运动辅助的视觉跟踪方法

薛万利 1,2，李昂 2，张程伟 3，陈胜勇 1,2

1.学习型智能系统教育部工程研究中心，天津理工大学，天津 300384； 2.天津理工大学计算机科学与工程学院，天津 300384； 3.大连海事大学信息科学技术学院，大连 115026

摘　要：当前的多数目标跟踪研算法均能较好地处理背景干扰、光照变化、目标形变等挑战，但是对于目标遮挡往往效果不甚理想。鉴于此，本文从全局运动轨迹与局部运动惯性两个角度，提出利用多尺度目标运动辅助跟踪的思路来应对目标遮挡。针对全局目标轨迹获取问题，提出基于图像拼接的全局目标轨迹获取算法。具体地，采用基于SIFT的图像拼方法接将所跟踪的视频序列拼接成一幅“全景图”，获得目标在不同帧的位置图，进而拟合后得到目标的运动轨迹。针对局部运动惯性问题，利用运动物理模型，即短时运动加速度，获得目标在局部时空尺度上的运动惯性。然后，通过全局目标运动轨迹以预测目标粗略的全局运动方向，再辅以细致的局部运动惯性，共同调整区域内目标候选样本的置信度，形成合理的多尺度目标运动辅助模型。最后，将多尺度目标运动辅助模型与基于检测的跟踪框架相结合。特别的，针对目前跟踪研究中数据集较为缺乏目标遮挡挑战性视频序列的客观问题，同时也为验证上述所提思路的正确性，提出了一个全新的以目标遮挡为主要挑战的数据集OFV-25，包含25个视频序列，共计3900帧。在该数据集及OTB-2015基准测试集上的实验结果验证了基于多尺度目标运动辅助视觉跟踪思路的有效性。

关键词：目标跟踪；运动轨迹；目标遮挡；图像拼接

论文编号：364

基于文本注意力指导的双线性融合及 Tucker 分解双模态情感特征提取

贾熹滨，安迅

北京工业大学，北京市 100124

摘　要： 随着大数据迅速发展，基于文本、图像等多模态融合的情感分析渐成为热点之一，然而由于文本和图像的异构特征，双模态情感分析中的融合特征提取仍是一个难点问题。针对该问题，本文利用文本情感特征更显著的特点，提出基于文本注意力指导的双线性融合特征，并正则化处理实现多模态融合的特征提取。在基于文本注意力机制的实现情感图像特征对齐、提取基础上，针对所提取视觉和文本情感特征的融合问题，克服现有的融合方法采用将两种模态特征进行拼接、相加，相乘的融合方法，考虑到在融合方面利用不同模态之间更加深层的关联关系，本文提出了利用双线性模型来融合不同模态特征，进而考虑到双线性模型融合特征维度高，分类计算代价大问题，本文利用Tucker分解方法进行降维计算，利用Tucker分解去除冗余信息，挖掘出图像和文本特征之间的高层信息。本文提出的双线性融合及Tucker 分解特征提取方法（BTF），有效解决了不同模态之间特征对齐的问题，提升了多模态特征融合性能，在Flickr图文数据集实验，与主流多模态融合情感分析方法相比取得了较好的分类效果，实验结果表明，所提出特征提取方法，预测情感倾向的准确率为84.5%。

关键词：情感分析；多模态；注意力机制；张量分解；特征融合

论文编号：389

基于特征子空间直和的跨年龄人脸识别方法

叶继华 1，郭祺玥 1，江爱文 1

江西师范大学计算机信息工程学院,江西省南昌市 330022

摘　要：跨年龄人脸识别问题作为人脸识别的一个分支，具有巨大的理论价值和现实意义，因此引起了越来越多的学者关注。近年来，学术界提出了许多判别性方法用于解决跨年龄人脸识别任务，并取得了可观的成果。然而，这些方法都在一定程度上忽视了年龄因子与身份因子之间的相关性。因此，本文在同时进行人脸身份识别和年龄分类这两个任务的多任务卷积神经网络的基础上引入直和模块，提出了一种新颖的基于特征子空间直和的多任务卷积神经网络 (Feature Subspace with Direct Sum CNN, FSDS-CNN)用以解决上述问题。该网络利用两个并行子网分别从卷积单元共享的深度特征中提取出身份相关特征和年龄相关特征，并对这两个相关特征所对应的特征子空间施加直和约束，使得身份相关特征与年龄相关特征尽可能无关。通过多损失的联合监督学习，该网络可以获得随年龄变化鲁棒的年龄无关人脸身份特征。本文在Morph Album 2、CACD-VS和Cross-Age LFW这三个公开的基准老化数据集上进行了大量实验，并对直和模块的两个重要超参数（特征向量个数与直和损失权重）的设置进行了探索实验。实验结果证明了本文提出的模型能够有效地提升跨年龄人脸识别的性能并具有一定的优越性。

关键词：人脸识别；跨年龄；多任务；子空间直和；特征子空间。

论文编号：429

基于深度回归网络的实时人脸跟踪算法

陈旭 1，祁亚楠 2

1.清华大学电子工程系,北京 100084； 2.江西师范大学,江西南昌 330099

摘　要：人脸跟踪是计算机视觉领域的一项重要研究课题，在诸如公共安全、视频监控等领域具有广泛的应用前景。当前主流的跟踪算法通常采用深度学习方法，虽然相较于传统特征方法能够获取更高的准确率，但由于深度网络参数的复杂度，导致跟踪速度较慢。此外，受限于跟踪问题本身的特性，现有方法往往存在误差的累积传递，从而造成无法连续跟踪多帧图像。针对上述问题，本文提出一种基于深度回归网络的人脸跟踪算法（TBR: tracking by regression），TBR算法充分利用跟踪问题本身的特性以及人脸具有区别于其他物体的明显的特征，一方面根据跟踪目标前后两帧的相对位置关系，对当前帧目标位置的确定提供了易于获取且较为准确的大致位置估计；另一方面利用一个轻量级的网络就能够提取出人脸特征，并且根据人脸特征就能够在候选区域准确地定位人脸的位置。实验结果表明，其处理速度在单核GPU上能够达到1700FPS以上，在CPU上能够达到500FPS以上。这说明本文所提出的TBR算法在使用深度特征的前提下达到了实时跟踪的要求。同时，由于TBR独特的特征选取方式，使得下一帧回归框的位置不会过分依赖于前一帧的跟踪结果，从而避免了大多数算法存在的误差累积传递问题，进而达到了较高的跟踪准确率和跟踪时间长度。

关键词：人脸跟踪；深度学习；回归网络；端到端；多任务

墙报 1-2

论文编号：321

三维烟雾流场超分辨率数据生成网络模型的研究

杜金莲，李淑飞 ，金雪云

北京工业大学计算机学院,北京市朝阳区 100029

摘 要：基于物理的烟雾模拟的缺点是 N-S 方程求解过程复杂，模拟效率较低，使得许多应用场景的特效制作周期都十分漫长，制作成本较高。本文通过深入分析已有的基于深度学习的超分辨率生成方法与使用 N-S 方程求解烟雾模拟的过程，探索设计了一种能基于 N-S 方程求解的低分辨率烟雾流场数据生成高分辨率烟雾流场数据的深度学习模型，以代替复杂的微分方程求解。模型以生成式对抗网络架构为基础，借鉴图像超分辨率生成中亚象素的思想构建基于亚体素卷积层的烟雾数据重建网络层，使用优化的Earth-Mover(EM)距离的对抗损失来稳定神经网络的训练过程，同时匹配生成模型和判别模型的学习进度；通过对烟雾模拟过程进行分析设计基于平流运动的平流损失，将烟雾流场时间域上的语义信息通过平流损失融入到算法中，从而实现高精度烟雾的模拟。最后本文将图像超分辨率质量评价指标进行扩展，构建了烟雾密度数据峰值信噪比用于对重建的高分辨率三维烟雾流场数据质量进行评价。实验表明，将本文设计的深度学习模型与烟雾流场 N-S 方程相结合而重建的烟雾数据在数值分布、准确性以及视觉效果上都表现良好。

关键词：深度学习；烟雾模拟；超分辨率；亚体素卷积

论文编号：358

基于款式变换和局部渲染相结合的虚拟试衣

徐俊1，普园媛1,2*，徐丹1，赵征鹏1，钱文华1，吴昊1，阳秋霞1

1.云南大学信息学院，云南昆明 650504； 2.云南省高校物联网技术及应用重点实验室，云南昆明 650504

摘　要：目前，人们主要依赖于传统试穿来观察服装效果，针对如何通过计算机让服装虚拟试穿到人物图像上进行的研究逐渐成为热点话题。现有方法的虚拟试衣结果，会出现人物细节特征丢失、非试衣区域特征发生改变等情形，影响了服装观感以及试穿者对服装的评判。针对上述问题，提出了款式变换和局部渲染相结合的试衣方法。首先，为了确保试衣区域的准确性，提出了先对人物图像进行像素级语义分割的策略，寻找出服装区域的位置，同时也为局部渲染提供实施条件；其次，为了实现不同服装款式间的试衣，构造了可学习的款式变换模块；再次，为了让服装更好地契合试穿者身材姿态，提出了可学习的服装形变模块；最后，为了保留人物原始细节特征和维持试衣图像观感，提出了局部渲染的策略，仅对服装区域进行渲染，保证了非试衣区域的细节特征不会丢失。实验结果表明，本文方法在人物手臂、手部等细节保持方面效果显著，除试衣区域外，更好地保留了非试衣区域的特征信息，试衣后图像更加写实保真。此外，还解决了试穿者姿态复杂时无法正确试衣的问题。

关键词 ：虚拟试衣；图像保真；语义分割；款式变换；局部渲染

论文编号：58

基于 WGAN-GP 的人脸素描画像-照片转化方法研究

闵佳媛 1，王世刚 2，江春晖，高鹏，王健，韦健

1.吉林大学通信工程学院研究生,吉林长春 130012； 2.吉林大学通信工程学院教授,吉林长春 130012

摘　要：公安部门通常利用犯罪现场和目击证人的证词作为依据，聘请专业的心理画像师画出犯罪嫌疑人的素描画像，但素描画像所包含的人脸信息不如照片包含得丰富、直接，因此常常以素描画像为基础生成嫌疑人照片，供刑警追捕使用。针对这个现实问题，提出了基于WGAN-GP网络的生成模型，引入U-NET网络作为生成器模型，patchGAN作为判别器模型。该算法利用了U-NET网络可以直接生成图片的特点，加上L1损失函数鲁棒性较强的性质，可以较快的训练网络。实验采用人类肉眼观察判断和图像质量评价标准来评估算法性能：本文在人类肉眼观察中设置了直接观察对比图和志愿者判断真假两个环节，实验结果表明本文生成的人脸照片较为接近真实照片，实验取得初步进展；本文引用了SSIM、FSIM和VIF三个图像质量评价指标作为客观评价标准，通过与其他方法对比可知，本算法生成的人脸照片质量较高、与真实人脸照片的风格较为相似。综合以上主观评价和客观评价结果，本文提出的基于WGAN-GP框架的人脸素描-照片转化方法可以很好的将素描照片转化为人脸照片，在保留了大量的细节特征的同时还保留了真实照片的风格。

关键词：生成对抗网络；异质图像；图像合成；素描画像；卷积神经网络

论文编号：61

基于多尺度重叠滑动池化的SSD果冻杂质检测方法

张芳慧1，章春娥1，张琳娜2*，岑翼刚1，阚世超1，金文3

1.北京交通大学，北京，100044； 2.贵州大学，贵州贵阳，550025；

3.江苏金海星导航科技有限公司,江苏镇江 212002

摘　要：果冻内部的杂质检测直接关系到食品安全问题，也是企业产品质量的保证。由于果冻行业生产效率的提高，传统的人工杂质检测方法已不能满足果冻生产商的需求。随着深度学习技术与计算机视觉的不断发展，其在工业产品表面缺陷检测方面得到了很好的应用。因此，研究基于视觉图像的果冻杂质检测系统是果冻生产企业一个亟需解决的问题。本文提出一种基于SSD的杂质检测方法，用于检测生产线中果冻内部的杂质，并标注出杂质的类型和位置。在预处理阶段，提出滑动图像块分割方法，将整张果冻图像分割成若干分割块图像，避免图像过大、目标过小，造成准确率低的现象。使用迁移学习的方法，将神经网络在ImageNet上学习到的特征迁移到果冻数据库中，加快网络收敛速度，在一定程度上避免了过拟合现象。提出多尺度重叠滑动池化(SOSP)，用于取代SSD网络中第五层池化以取得更加鲁棒的特征池化。最后，将一幅图下的所有分割块图像上的检测结果进行整合，得到整张图像的检测结果。实验结果表明，本文提出的方法有效可行，准确率达到0.7271。相比其他先进的方法，本文的算法更具鲁棒性，可应用到果冻杂质检测生产线中。

关键词：果冻杂质检测；深度学习；多尺度重叠滑动池化；迁移学习；目标检测算法

论文编号：106

GCNAS:面向图分类的网络架构搜索

张文强，王兴刚，冯镔，刘文予

华中科技大学 电子信息与通信学院 武汉 430074

摘　要：最近，神经网络架构搜索（Neural Architecture Search，NAS）技术获得研究人员的广泛关注，搜索得到的网络性能表现超过许多现有的人工设计架构（MobileNet, ResNet, VGG等），在许多大型图像数据集上都表现出很好的性能。图神经网络对处理非欧式数据有着得天独厚的优势，通过聚合邻居节点的信息，可以实现特征信息的逐级传递，实现对全图特征的提取。研究者们根据图的数学性质和网络设计知识，提出了不同的图卷积形式和图网络结构，在节点分类和图分类数据集上超越了传统方法。然而，人工设计网络需要耗费巨大的人力和时间成本，且无法充分挖掘图神经网络（GNNs）的表现能力。为此，本文提出基于神经网络架构搜索的图分类网络框架GCNAS，旨在获得表现更好的用于图分类神经网络架构。针对图分类任务，我们设计了路径级的搜索空间，利用二元优化的梯度下降算法和维特比算法来获得最终的网络架构。搜索得到的网络在三个图分类公开数据集都上取得了更优的性能。其中，GUNAS-B相对人工设计的Graph U-Net,在PROTEINS数据集上提升2.92%，在COLLAB上提升1.87%。

关键词：图神经网络；图分类；神经网络架构搜索

论文编号：**101**

基于 SmsGAN 的对抗样本修复

赵俊杰 1，王金伟 1，2

1.南京信息工程大学计算机与软件学院,江苏南京 210044；

2. 中国科学院信息工程研究所,信息安全国家重点实验室,北京 100093

摘　要：由于深度卷积神经网络的强大性能，它被广泛运用在文本语义识别、图像分类、目标检测等各个领域。对抗样本给神经网络的安全性带来了极大的威胁，然而对抗样本通常不具有鲁棒性，这使得对抗样本的高质量修复成为一项重要的研究课题。本文提出一种基于生成对抗网络(GAN)的方案实现对抗样本的修复，将基于深度学习的样本重构技术引入对抗样本的修复过程。判别器采用能对对抗样本进行高精度取证的随机多滤波特征统计网络（SmsNet）。在SmsNet中，我们提出了用于获取特征图多维数值特性的特征统计层，并使用随机多滤波连接（SmsConnection）将每个卷积核输出的特征图直接送到特征统计层，然后将这些特征值送入全连接层进行分类。生成器采用多尺度卷积核并行结构，联结后再进行一次卷积，输出最终的归一化图像，有效避免了反卷积带来的棋盘效应。生成器的损失函数由判别损失和引导损失两部分构成，前者用于监督生成器，使其重构的样本不再具有对抗性，后者用于缩小修复样本与原始样本之间的差异，两者与生成器共同形成目标引导生成器，实现生成器的固定目标优化。对抗样本经过下采样网络获取局部统计特征，再输入SmsGAN得到修复的样本。实验表明，采用SmsGAN修复的对抗样本，不仅在视觉质量上超过了均值滤波、颜色位深缩减等传统信号处理方案，而且保证了高修复率。

关键词：深度学习；对抗样本；样本修复；生成对抗网络

论文编号：139

基于注意力机制的人脸姿态生成

卢钦伟 1,3,4，赖剑煌 1,2,3,4，冯展祥 1,3,4，谢晓华 1,3,4

1.中山大学数据科学与计算机学院，广州，510006； 2.中山大学新华学院信息科学学院，广州，510520 3.广东省信息安全技术重点实验室，广州，510006；4.机器智能与先进计算教育部重点实验室，广州，510006

摘　要：近年来人脸识别精度随着深度学习的发展得到迅速提升，而深度学习需要依赖大量的高质量训练样本才能发挥出理想的性能。然而现有的人脸数据集规模虽然较大，但数据集中包含的人脸姿态信息有限，而带姿态的人脸数据生成提供了一种简单高效的数据增强以及人脸表征学习方法，可以很好地弥补这一不足。由于两张不同人脸图像的之间面部姿态差异巨大，对于生成任务来说仍然是一个重要挑战，因此本文提出了一种基于注意力机制的生成对抗网络方法，将人脸图像输入主干网络，再引入面部关键点信息到注意力分支，通过注意力分支的关键点信息引导学习生成目标姿态的人脸图像，并在一系列实验中证明了本文方法能够完成多对多的人脸姿态生成，以及高效地生成目标姿态人脸图像。

关键词：人脸；注意力机制；生成对抗网络；深度学习

论文编号：13

基于信息补偿的缺陷检测方法

李丰位[深圳大学，广东省深圳市 518061]，孔庆杰[精锐视觉智能科技（上海）有限公司，上海市 200000；东莞理工学院，广东省东莞市 523000]

摘　要：在缺陷检测领域，基于机器视觉的检测方法以其高效、无损的特点逐渐成为了主流检测方法之一。随着深度学习的发展，传统的图像处理方法逐渐被精度高、鲁棒性好的深度学习模型所取代，深度模型在缺陷检测领域已经成为了普遍的基于机器视觉的的检测方法。然而，在利用深度学习模型进行缺陷检测时，往往会遇到两个困难：（1）由于采样条件的限制，采取到能用于模型训练的样本数量较少；（2）采集到的不同缺陷的数量差距较大。为了克服这些困难，我们提出了一种利用单个缺陷各部分之间的联系来补偿缺陷特征细节的新深度模型，称为深度随机链条（Deep Random Branches，DRBS）。DRBS模型相较于现在通用的深度检测框架的改进之处在于其增加了一个随机链条生成器和一个特征序列学习器。首先，经过预处理的图片被传输到模型的特征提取层得到特征图，然后区域候选网络在特征图上得到粗略的缺陷候选区域，接下来随机链条生成器为每个缺陷根据缺陷内部块之间的联系生成随机链条，随机链条携带了每个缺陷的结构信息；最后，特征序列学习器利用随机链条重构补偿每个缺陷的语义信息细节。从纵向的角度观察，补偿了语义信息的缺陷在进入最后的检测前，有效信息得到了增强；从横向的角度来讲，不同类别的缺陷间的语义信息量的差距极大被拉近。两组缺陷实验数据证明了该DRBS检测子在缺陷检测领域的优异性能。

关键词：深度学习；缺陷检测；深度随机链条；信息补偿

论文编号：23

基于改进的SSD模型手机违规使用目标检测

李智伟1，杨亚莉1，钟卫军2，杨聪1，包壮壮1

1.空军工程大学基础部,陕西 西安 710051； 2.宇航动力学国家重点实验室,陕西 西安 710043

摘　要：随着智能手机占有率的爆发式增长，手机泄密正在成为新的泄密途径，手机的拍照、录音等功能成为了保密场所的重大安全隐患。针对某些对手机使用有特殊规定的场所时常面临难以准确、高效地识别手机违规使用的问题，文中提出了一种基于改进的SSD模型来检测手机的违规使用。由于实验中SSD模型对手机目标检测的定位精度高但识别精度低，为避免改进SSD模型时对定位精度造成不利影响，本文采取在不改动原有SSD模型的情况下，为模型添加二级网络用来提高识别精度的策略。首先，利用SSD模型获取初次目标位置及区域分类；其次，为提升边缘特征提取能力，根据迁移学习的原理，设计了改进的Densenet模型，将其作为二级网络对初次目标候选框进行二次识别；最后，结合初次目标位置和二次识别的置信度，获取精确的手机检测边界框。为改进数据预处理流程，文中采用了数据扩增与图像质量改善相结合的策略。在自建的手机检测数据集上的实验结果验证了这些改进策略的有效性，改进的SSD模型定位精度可达90.9%，识别精度高达99%，相比原有SSD模型提升了36%。文中改进的SSD模型同时具有识别精度高和定位精度高的特点，可为智能识别违规使用手机的行为提供理论依据和技术支持。

关键词：目标检测；深度学习；SSD；迁移学习；数据预处理

论文编号：56

面向图像分类的基于注意力引导的 Dropout

常东良 1，尹军辉 1，谢吉洋 1，孙维亚 2，马占宇 1

1. 北京邮电大学人工智能学院 北京 100876　　2. 南水北调中线信息科技有限公司 北京 100176

摘　要：当一个较大的神经网络在较少的训练数据上训练时，不可避免的会遭遇过拟合问题，进而在测试数据集上泛化性能较差。因此，现有的研究提出了多种基于Dropout的正则化方法来缓解这个问题。这些技术通过随机关注或丢弃一部分特征或特征描述算子来缓解它们之间的协同自适应问题，从而增加模型分类准确率。然而，这些方法往往不能直接促使模型关注有较低判别力的特征，这些特征对减少过拟合同样非常重要。为了解决这个问题，本文提出了一种基于注意力引导的Dropout（Attention-guided Dropout，AD），该技术利用自监督机制更高效的缓解特征描述算子之间的协同自适应问题。AD包含两个独特的组件：特征的重要性度量机制和基于可学习丢弃概率的Dropout。特征的重要性度量机制通过SE-Block度量得到每个特征的重要程度。基于可学习丢弃概率的Dropout能够通过丢弃“较好”的特征检测算子，强迫“较差”的特征检测算子学习到一个较好的特征表示，从而缓解特征检测算子之间的协同自适应并促使模型学习拥有较低判别力的特征。实验结果表明提出的方法可以很容易地被应用到各种卷积神经网络（Convolutional Neural Network，CNN）结构里。同时，我们也证明了，当训练样本数大量减少时，通过选择合适的删除率（Dropout Rate），提出的方法能够比一些现有的方法更好地抑制过拟合。

关键词：深度神经网络；过拟合；Dropout；自注意力机制

论文编号：80

基于Transformer的证件图像无检测文字识别

肖慧辉1，2，张东波1，王旺2，王家奎2

湘潭大学信息工程学院,湖南省湘潭市，411100 2. 武汉唯理科技有限公司，湖北省武汉市，430000

摘　要：近年来，深度学习在图像文档文字识别方面的应用已较为成熟，但是现存的模型都有检测和识别两个过程，且需要借助复杂的网络结构、大量的文本框标注来提高识别准确率。但无论是两步策略还是端到端的识别模型，都离不开文字区域检测过程。本文提出了一个简单且鲁棒性强的证件图片无检测文字识别方法，本文方法不需要定位文字区域，而是直接通过嵌入二维特征图中不同序列位置的水平、竖直方向位置编码，将不同子空间的特征表达连接到基于Transformer的序列解码器，由于不采用递归神经网络，本文的网络模型能并行训练并可以快速收敛，且仅需要文字级别的标注，显著降低了标注成本，缩短了识别时间，而且，本文方法可以将特殊符号插入到识别的文本序列中，大大简化了文本信息后处理的流程，可以直接得到结构化的字段数据，这个优势是传统的文字识别处理方法难以比拟的。测试结果表明，本文方法在身份证扫描件文本图像识别表现出优越的性能。

关键词：Transformer；端到端模型；无检测文字识别；全局上下文。

墙报 1-3

论文编号：100

基于跨模态信息迁移融合的人脸情感分析

付彦喆 1，王耀平 1，江子文 1，冯镔 1，刘文予 1

华中科技大学 电子信息与通信学院，湖北省 武汉市 430070

摘　要：人脸情感分析技术是近年来计算机视觉中研究的热门方向。随着多媒体技术的发展，基于单模态的人脸情感分析技术无法满足实际应用的需要。针对在实际应用中多模态特征采样率不同而导致的融合效率低的问题，本文提出一种基于人脸特征的跨模态信息迁移融合模型，在无需特征对齐预处理的情况下，用于迁移文本和语音中与人脸相关的特征，提高多模态融合效率。另外，针对人脸特征在模态融合前后产生的来自其他模态的噪声，本文提出一种跨模态残差融合结构，用于提取完整的全局融合特征，提升多模态融合特征的表达能力。最后在CMU-MOSI和CMU-MOSEI情感分析数据集上进行实验，在七分类精度、二分类精度和F1值上，较已有的方法取得了最好的结果，在MOSI数据集上分别提升0.9%、1.2%、1.2%。在MOSEI数据集上分别提升0.7%、1.7%、1.9%。

关键词：人脸情感分析、跨模态信息迁移融合模型、多模态融合、跨模态残差融合结构

论文编号：125

基于注意力机制的视频对象分割优化算法

钱晶晶 1，孙力娟 1,2，曹莹 1，郭剑 1,2 ，任恒毅 1

1. 南京邮电大学计算机学院 江苏 南京 210003 2.江苏省无线传感网高技术研究重点实验室 江苏 南京 210003

摘　要：近年来，深度学习推动了计算机视觉的发展。在视频对象分割方向，我们基于OSVOS（One-Shot Video Object Segmentation）提出了一个基于注意力机制的视频对象分割优化算法(Attention and Morphology Video Object Segmentation,AMVOS)算法。首先由于OSVOS对所有特征信息采用相同的方式，因此，背景的区域特征可能会影响目标分割算法的性能，针对这一问题，AMVOS采用Non-Local机制（一种自注意力机制）。其次，Non-Local在捕获视频帧关键区域的同时，通过计算像素点之间的关系，能够快速捕获长范围依赖，并将目标信息引入到OSVOS模型中。最后，在OSVOS算法中，分割视频帧因噪声、光照、遮挡物等问题，使得分割目标出现非目标连通区域，从而造成分割目标质量下降问题，因此，本文基于形态学提出优化算法，来有效删除非目标连通区域。所提AMVOS在Davis2016，GyGo，SegTrack-v2三个数据集上测试，在davis2016上，AMVOS区域相似度J得分82.1，轮廓准确度F得分84.1。平均每个数据集区域相似度J提高了1.4%，轮廓准确度F上升了2.2%。实验结果证明了AMVOS较OSVOS性能更优越。

关键词：视频对象分割；OSVOS；注意力机制；形态学

论文编号：127

基于注意力的短视频多模态情感分析

黄欢 1,孙力娟 1,2,曹莹 1,韩崇 1,2 ,任恒毅 1

1.南京邮电大学计算机学院 江苏 南京 210003 2.江苏省无线传感网高技术研究重点实验室 江苏 南京 210003

摘　要：随着社交媒体的发展，短视频已成为人们情感和生活状态的主要表达方式之一。基于短视频的情感分析已经变得越发重要并引起了研究人员的关注。现有的情感分析方法缺乏对短视频中信息的充分考虑，从而导致不恰当的情感分析结果。基于此，本文提出了音视频多模态情感分析（AV-MSA）模型，该模型通过利用视频帧图像中的视觉特征和视频中的音频信息来完成短视频的情感分析。该模型分为两个分支，即视觉分支和音频分支，在音频分支中，我们采用CNN架构来提取音频图谱中的情感特征，实现情感分析的目的；在视觉分支中我们采用3D卷积操作来增加视觉特征的时间相关性。此外，在Resnet的基础上，为了突出情感相关特征，我们添加了注意力机制，以提高模型对信息特征的敏感性。本文设计了一种交叉投票机制用于融合视觉分支和音频分支的结果，产生情感分析的最终结果。所提AV-MSA在IEMOCAP和微博视听(微博视听，WB-AV)数据集上进行了评估。实验结果表明，与现有算法相比，本文提出的AV-MSA在分类精确度上有了较大的提升。

关键词 ：多模态情感分析；残差网络；3D 卷积神经网络；注意力；决策融合

论文编号：138

基于直推式支持向量机和像素级特征的膀胱肿瘤准确分割

杜鹏 1，徐肖攀 1，郑豪杰 1，梁继民 2，卢虹冰 1，刘洋 1

1 第四军医大学军事生物医学工程学系.,陕西西安 710032； 2. 西安电子科技大学生命科学技术学院,陕西西安 710126

摘　要：膀胱癌是泌尿系统最常见且致死率最高的恶性肿瘤。术前准确评估膀胱癌患者的分级、分期，对患者手术方案的制定具有重要影响。近年来，影像学诊断在膀胱癌的分期、分级评价中得到极大关注，而膀胱癌的分割问题是其中的关键步骤。然而，膀胱肿瘤的准确分割仍面临许多实质性的挑战，比如在磁共振成像上肿瘤区域的强度不均匀、肿瘤区域与膀胱壁存在交叠。为了解决这些问题，本文提出一种基于高通量像素特征和直推式支持向量机分类器的算法提升膀胱肿瘤的分割效果。本文将收集的20名膀胱癌患者的T2加权磁共振影像（T2W）数据，随机分成10名患者组成训练组，10名组成测试组。训练组中每名患者利用双向水平集算法提取候选区，并提取三个感兴趣区域（ROI），分别为：肿瘤ROI、壁组织ROI、以及混合ROI（壁组织和肿瘤组织混合交叠的区域）。随后对感兴趣区域进行二维离散小波变换后，进行三个ROI像素级特征的提取，并基于SVM的递归式特征消除法进行特征选择并构建最优特征子集。基于最优特征子集利用直推式支持向量机训练分类器模型，并在测试组验证分类器效能。实验结果表明，半监督算法能充分利用混合区域的信息，提升模型的分割效能。此外，利用本文提出的方法得到的Dice相似性系数为88.41%，高于最新的膀胱肿瘤深度学习分割模型得到的结果。综上所述，本文提出的基于高通量像素级特征和直推式支持向量机分类器的膀胱肿瘤分割算法，能有效分割肿瘤区域和膀胱壁区域并获得较好的分割效果。

关键词：膀胱癌；磁共振影像；肿瘤分割；像素级特征；直推式支持向量机

论文编号：166

基于曲线轮廓匹配的甲骨拓片缀合研究

田瑶琳 1，2，高未泽 2，刘绪兴 2，陈善雄 2，莫伯峰 3

1.西南大学商贸学院,重庆市 402460； 2.西南大学计算机与信息科学学院,重庆市 400715；3.首都师范大学甲骨文研究中心,北京市 100048

摘　要：甲骨拓片的缀合是甲骨考释研究的基础，而准确缀合碎片化严重的甲骨拓片却并非易事。针对甲骨拓片缀合中存在的小碎片缀合难、人工标注细节丢失、断裂口曲线匹配准确度低等问题，本文提出一种曲线拟合度分析（CFDA）和相关性分析（PCA）结合算法（CFDA&PCA），可提高甲骨拓片自动缀合的准确度。首先，基于曲线特征分析构建了正交化体系进行轮廓线局部特征提取；其次，通过提取轮廓线坐标序列构建了全局特征描述向量，实现了曲线轮廓的跟踪及拟合；随后，结合轮廓线特征集和CFDA算法筛选了匹配项，使得甲骨拓片匹配候选项的范围大大缩小；最后，利用PCA算法生成了基于PCA相似度排名的缀合图像库，以评估缀合效果。实验结果表明，本文提出的CFDA&PCA算法能使甲骨拓片的缀合准确度高达84%以上，比传统算法高出至少25%，有效地提高了小样本上甲骨拓片缀合的准确度和自动化水平，从而为解决甲骨碎片缀合的研究难点提供了一种有效的思路。

关键词：甲骨拓片；曲线特征描述；轮廓线匹配；缀合算法；Pearson 系数

论文编号：194

一种手写体古文字检测模型

李然康，陈善雄，赵富佳，林小渝，韩旭

西南大学计算机与信息科学学院，重庆，400715

摘　要：古籍中的手写体文字由于含有模糊、缺失、风化、污渍等特征，一直是手写体文字检测的难点，这些特征会严重影响检测精度。我们设计了一个有效的古文字检测模型（ATD模型，Ancient Text Detection Model），该模型包含基于CNN的文本框生成网络，即ATD网络，和基于NMS的MSER文本框生成模型，即MSER模型。CNN可以很好学习到图像的深度特征，对于噪点不敏感，但是需要大量时间和预处理，而MSER可以快速生成精度极高的文本框，但是噪点对于精确度影响极大。因此，我们结合了CNN和MSER的优点，利用ATD网络和MSER模型对同一字符生成两个不同的文本框，合并算法对两个输出文本框进行合并，得到了更高精度的文本框定位。在彝文、英文、拉丁文、意大利文等古籍数据集上的测试结果表明，ATD模型具有较好的精度，而且ATD模型在不同语言的数据集上的实验表明本模型具有普适性，为手写体古文字的检测研究迈出了坚实的一步。

关键词：文本检测；卷积神经网络；最大稳定极值区域；非极大值抑制;文本框合并

论文编号：209

三维空间中的鲁棒对抗攻击方法

魏红阳 1，韦世奎 1，阮涛 1，楚浩宇 1，赵玉凤 2，赵耀 1

1、北京交通大学 信息科学研究所，北京 100044

2、中国中医科学院 中医药数据中心，北京 100700

摘　要：研究对抗样本攻击方法对于认知深度神经网络缺陷具有重要理论和应用价值，并吸引了学术界的广泛关注。由于现有的大多数对抗攻击方法如FGSM算法、CW算法等都是通过修改二维图像空间的部分像素实现攻击，其不能有效抵抗现实世界物体的3D物理特性（如旋转、平移和光照变化等），从而造成攻击失败。为此，本文拟研究三维空间中的鲁棒对抗攻击，提出了一种新的攻击方法：通过特定的优化算法对3D模型纹理图像进行人类视觉不可见的微弱修改，并且在3D渲染过程中对三维空间中的变换分布进行建模，使得最终三维渲染器生成该模型的不同形态下的二维图片可以成功攻击一个深度神经网络分类器并误导其错分为任一指定类别。在自建的3D模型数据集上的实验结果表明，该方法具有较高的攻击成功率，其产生的对抗样本在面对三维空间中的变换时具有鲁棒性。并且该方法产生的对抗样本在不同的深度网络模型之间具有迁移能力，其可以攻击未知的黑盒模型，对现实中AI系统能够造成安全威胁。此外通过在3dsmax中模拟真实场景验证了本文所提方法生成的对抗样本在现实世界场景中同样具有有效性。

关键词：对抗样本；对抗攻击；攻击失败；三维空间；鲁棒；迁移能力

论文编号：290

基于多样性约束和离散度分层聚类的无监督视频行人重识别

曹亮，王洪元，戴臣超，陈莉，刘乾

常州大学信息科学与工程学院,江苏常州 213164

摘　要：视频行人重识别是一项应用非常广的计算机视觉任务。目前的视频行人重识别方法通常是基于监督学习的，该方法需要手工标记大量的数据，代价非常高且并不适用于现实场景。本文提出了一种从底向上的基于多样性约束和离散度分层聚类的无监督视频行人重识别方法。首先，将每个样本当作是一个不同的类，利用卷积神经网络提取视频序列特征，然后结合类内间离散度和多样性约束进行从底向上的分层聚类，类内和类间离散度都小的类别将被优先合并，同时多样性约束能够平衡每类中的样本数量，防止数量过多的类被错误合并，提升聚类效果。在每次聚类的过程中，利用一个线性变化的特征存储器自动更新每类的特征，从而动态的更新模型。在两个大型的视频行人重识别数据集上的实验表明，相比于目前先进的无监督视频行人重识别方法，本文方法在性能上有一定的提升。在Mars数据集上，首次识别率（Rank-1）和平均精度均值（mAP）达到了63.2%和41.3%，在DukeMTMC-VideoReID数据集上，首次识别率（Rank-1）和平均精度均值（mAP）达到了75.4%和67.4%，同时在两个数据集上分别进行了消融实验,实验结果验证了所提出方法的有效性。因此本文提出的基于多样性约束和离散度分层聚类方法可以有效提升聚类效果，从而提高无监督视频行人重识别的性能。

关键词：无监督视频行人重识别；多样性约束；离散度；聚类；特征存储器

论文编号：295

感受野形态对卷积神经网络性能影响的研究

葛一帆 1，吴建鑫 1

1.南京大学软件新技术国家重点实验室,南京 210023

摘　要：本文研究卷积神经网络的感受野。不同于大部分的工作，我们分析感受野的形态，而不仅仅是尺寸。我们基于从输入到输出的路径数目给出了感受野的一种形式化定义，并利用神经网络的可微性对其进行计算，发展出一套分析感受野形态的可视化方法。通过对感受野形态的分析，发现广为使用的一些卷积神经网络存在感受野中心偏移和呈“网格”状的问题。我们认为这些问题对神经网络的性能有害，于是针对这些问题给出了修改方案，即，将步长为2的卷积替换成一个步长为2的2×2平均池化加上一个步长为1而卷积核大小不变的卷积。从可视化结果可以看到修改方案消除了这些问题，并通过在分类，物体检测和实例分割等计算机视觉任务上进行实验验证修改方案的有效性。本文所提修改方案在基于ResNet-50的模型上能够有效地提高性能，并且在对感受野形态敏感的物体检测和实例分割任务上比在对感受野形态不敏感的分类任务上效果更显著。我们认为，感受野形态的缺陷在其它模型中一样会带来不利影响，其解决方案值得进一步研究。

关键词：卷积神经网络；感受野；可视化；计算机视觉

墙报 1-4

论文编号：297

基于损失函数灵敏度的视频分割网络通道剪枝方法

张茂军 1，龙鑫 1，贲宗承 2，曾向荣 1，刘煜 1

1.国防科技大学系统工程学院,湖南长沙 410073； 2.国防科技大学计算机学院, 湖南长沙 410073

摘　要：近年来，随着计算机视觉相关任务的精细化，所使用的卷积神经网络（CNN）模型结构日益复杂化，其计算成本和存储量成指数级别增长，严重阻碍了它们在有限资源设备（小型移动端以及FPGA等）上的应用。因此迫切需要通过某些方法来对深度网络模型进行压缩和加速。在本文中，我们提出了一种基于损失函数灵敏度的方法来剪枝CNN模型的冗余通道。在给定训练充分的模型之后，该方法基于深度网络常用批量标准化层（BN）中的缩放和移位因子的泰勒展开式来识别网络中重要性程度较低的通道，之后按设定的百分比阈值剪枝这些通道，以获得参数量和FLOP消耗更少的紧凑型网络结构，最后对裁剪后的网络进行微调以弥补剪枝而带来的性能损失。在实验部分，我们主要在常见的视频分割网络上（OSVOS，One-Shot Video Object Segmentation）评估了该方法，统计发现剪枝后的网络模型能够在减少约75%的网络参数以及52%的计算量的同时不失分割模型的性能（整体的分割性能均值相比较于全精度模型仅仅约有1个百分点的损失）。

关键词：卷积神经网络；视频分割网络；泰勒展开；通道剪枝

论文编号：312

结合特权信息与注意力机制的场景识别

孙宁 1,2，王龙玉 1,2，刘佶鑫 1，韩光 1

1. 南京邮电大学 宽带无线通信技术教育部工程研究中心，南京 210003； 2.南京邮电大学 通信与信息工程学院，南京 210003.

摘　要：在场景识别中，RGB图像呈现场景外观信息，深度图像包含场景几何信息，两者互为补充。因此，基于RGB-D图像的场景识别研究成为近年来的热点。相比RGB图像，深度图像不易获取，特别是在实际应用中，往往只有RGB图像可以利用。为了在只有RGB图像的测试阶段也能利用深度图像与RGB图像所包含的互补信息，本文以深度图像为特权信息（训练时可用，测试时不可用的信息），提出了一种结合特权信息和注意力机制的端到端可训练的深度神经网络模型，称为PIA-SRN (scene recognition network using privilege information and attention mechanism)模型。在该模型中，以图像编码到特征解码再到图像编码为架构，建立了由RGB图像到深度图像再到深度图像高层语义特征的映射关系。通过注意力机制，将RGB图像高层语义特征与对应的深度图像高层语义特征进行融合，输入分类网络，最终得到预测结果。在测试时，只需要输入RGB图像，便可在本文模型获取的深度图像特权信息的帮助下，提升场景识别的性能。在SUN RGB-D和NYUD2两个RGB-D场景识别图像数据库上进行了大量实验，分别取得了51.5%和65.4%的正确率，优于多数RGB-D多模态融合场景识别算法的结果，验证了本文方法的有效性。

关键词：场景识别；特权信息；注意力；卷积神经网络

论文编号：325

拉普拉斯超分辨率印刷电路板瑕疵检测方法

施玉娟 1，林璐颖 1，杨文元 2

1.漳州职业技术学院,福建 漳州 363000； 2.闽南师范大学 福建省粒计算及其应用重点实验室，福建 漳州 363000

摘　要：瑕疵检测是印刷电路板生产中一个重要环节。电路板复杂度高，瑕疵尺寸小，工业采集的图像分辨率低，传统检测方法无法满足高精度及高召回率的要求。针对此问题，提出结合拉普拉斯金字塔超分辨率网络和YOLOv3目标检测算法的超分辨率瑕疵检测方法(Super-Resolution Defect Detection Method, SRDDM)。首先，使用卷积神经网络提取PCB图像特征，将特征逐级放大，同时预测每一级的残差，将放大后的特征图与该级预测的残差相加，得到对应级别的超分辨率图像。经过两级超分辨率处理，瑕疵纹理更加清晰，边缘细节更为丰富。其次，超分辨率后的图像连同瑕疵标注文件输入到Darknet53网络中，使用卷积层提取特征图，通过残差网络融合各个层级的特征信息，生成三个尺度的特征图。然后，针对PCB瑕疵微小，YOLO设计的先验框尺寸不合适问题，使用K-Means聚类方法对数据集真实瑕疵框进行聚类分析，根据聚类结果调整先验框的大小，利用先验框分别在三个特征图上定位瑕疵，生成候选框。最后，使用非极大值抑制合并候选框，预测瑕疵的位置及分类。实验表明，使用SRDDM对瑕疵数据集进行检测，平均精确度均值为99.41%，召回率为99.31%。

关键词：目标检测；卷积神经网络；拉普拉斯超分辨率网络；YOLO；瑕疵检测

论文编号：327

超分辨率对抗网络的电路板瑕疵检测

吴慧君 1，施玉娟 1，杨文元 2

1.漳州职业技术学院，福建漳州 363000；2.闽南师范大学 粒计算及其应用重点实验室,福建漳州 363000

摘　要：随着电子信息高速发展，电路板集成规模越来越大，生产难度直线上升，出现瑕疵的几率也大大增加。面对大量的 PCB 板，人工目测已无法快速完成检测任务。采用自动光学检测虽然可以提高检测速度，但 PCB 图像中的瑕疵只占整个图像区域的很小比例，传统的深度学习神经网络对于小目标的检测精度并不是很高。针对以上问题，提出超分辨率生成式对抗网络的电路板瑕疵小目标检测方法，利用超分辨率生成式对抗网络提高 PCB 数据集的图像分辨率，再用深层网络的多尺度进行小目标瑕疵检测，以提升瑕疵检测的精确率。首先，将电路板数据集图片通过超分辨率生成式对抗网络提高分辨率，使小目标瑕疵放大至 4 倍尺寸，得到标注好 6 类瑕疵位置的超分辨率数据集 PCB-SR；其次，在 Darknet-53 网络中的不同层级通过卷积、采样和融合提取 19×19、38×38、76×76 三种粒度的特征图，将对应的高层特征上采样后融合低层特征映射，使高层能使用低层的细粒度特征以及更多语义信息，以取得各尺度下细致的检测结果；最后，聚类出的 9 种长宽比不同的先验框，对特征图进行多尺度预测，调整边界框的位置和大小并输出边界框坐标及对应的分类。实验表明，超分辨率生成式对抗网络的电路板瑕疵小目标检测方法在 PCB-SR 的 6 个瑕疵类别平均精确率达到了 99.38%，比相关检测网络均有提高。

关键词：目标识别；生成式对抗网络；超分辨率；YOLO；瑕疵检测

论文编号：331

基于改进的 3D-CNN 的高光谱遥感图像地物分类

谢幸雨 1，贺辉 1，邢海花 2

1 北京师范大学珠海校区,广东省珠海市， 519087

2 海南师范大学，海南省海口市，571158

摘　要：高光谱遥感影像波段众多，包含丰富的辐射、空间和光谱信息，是多种信息的综合载体，应用广泛。但是传统的高光谱影像地物分类方法多着重于光谱维度的特征提取，却忽略了空间维度上的特征，进而影响了分类的准确性。三维卷积核具有可以同时在三个维度上对数据进行卷积处理的特性，故本文采用3D-CNN深度网络进行高光谱影像地物分类，并针对3D-CNN网络存在的问题，提出了一种基于改进的3D-CNN的高光谱遥感影像地物分类算法。本方法对提取到的空间和光谱特征实现融合复用，尽可能发挥特征的价值。此外，本文引入浅层细节特征保存网络思想，将浅层特征细节与融合复用后的特征结合，减少由于特征复用而产生的误差，进一步提高了高光谱影像地物分类的准确度。在Tensorflow框架下对两个常用的高光谱遥感影像数据集（Indian Pines高光谱遥感影像数据集、Pavia University高光谱遥感影像数据集）的实验结果表明，相比基础的3D-CNN网络，本文方法分类精度提高了近2个百分点。

关键词：高光谱图像分类；3D-CNN 网络；特征融合浅层；浅层特征细节保存

论文编号：333

融合双神经网络的 PCB 缺陷检测方法

林璐颖 1，吴慧君 1，杨文元 2

（1.漳州职业技术学院,福建 漳州 363000；2.闽南师范大学 福建省粒计算及其应用重点实验室,福建 漳州 363000）

摘　要：印刷电路板（PCB）缺陷检测是电子工业领域的一个重要研究课题。PCB电路密集复杂，缺陷非常小，传统检测方法检测精度低，漏检率高。针对以上问题，提出一种融合YOLOv3和VGG19神经网络的双网络小目标检测方法(Dual Nerve Defect Detection Network,DNDD)，实现对PCB缺陷的高质量检测。DNDD算法首先对PCB缺陷数据集DE进行预处理，根据标注文件将缺陷目标裁剪出来形成数据集DE1。将DE和DE1分别输入YOLOv3和VGG19网络进行训练，生成缺陷检测模型M1和M2。其次，使用YOLOv3检测模型M1对PCB图像进行检测，目标置信度大于等于50%为缺陷确认目标R1，置信度小于50%为缺陷待确认目标R2。最后，将R2裁剪出来，使用VGG19缺陷检测模型M2对其进行二次检测，检测结果中目标置信度大于等于50%的为R3，将R1和R3进行融合，提高了目标的置信度。实验表明，双神经网络缺陷检测方法在PCB缺陷小目标上的检测精度提高至99.44%，有效地减少了由于置信度低而导致的漏检问题。

关键词：目标识别；深度学习；神经网络；印刷电路板；缺陷检测

论文编号：337

一种改进的分层双线性池化图像行为识别方法

吴伟 1，于嘉乐 1

内蒙古大学计算机学院,内蒙古自治区呼和浩特市 010021

摘　要：基于图像的行为识别由于受到类内图像背景信息的差异性和类间行为的相似性干扰，至今仍然是一项极具挑战性的任务。某些行为类别在人物姿态和表情动作方面十分相似，因此提取区分性特征或者对各种行为语义部件进行建模对于基于图像的行为识别至关重要。许多方法使用昂贵的手动注释来学习可区分的部件信息以进行动作识别，这严重影响了这些方法在实际生产和生活中的广泛应用。近年来，双线性池化模型由于其自动学习显著性特征而显示出其在图像分类中的有效性。借鉴双线性池化模型在细粒度图像分类中的优势，同时为避免该模型包含大量嘈杂的背景信息而影响行为识别的精度，论文提出了一种改进的双线性池化模型用于图像的行为识别。该模型利用通道和空间注意力机制关注图像中重要的目标区域，并且通过集成多层注意力掩码图来生成ROI，这可以有效的抑制图像中的背景噪声信息，从而有效提高了行为识别的准确性。更具体地说，该方法允许网络更好地关注图像中关键目标的区分区域，提取更具有代表性的特征。在图像行为识别数据集Stanford-40和自定义的60类行为数据集上的实验结果表明，论文提出的方法可以有效地提升图像行为识别的精度，相比于传统的行为识别方法和最新的方法取得了最好的识别性能。

关键词：图像行为识别；分层双线性池化；注意力机制；掩码聚合

论文编号：354

深度语义哈希算法

姜少臣 1，汪烈军 2，李永明 2，杜安钰 2，程述立 2

1、新疆大学软件学院,新疆维吾尔自治区乌鲁木齐 830046；

2、新疆大学信息工程与科学学院,新疆维吾尔自治区乌鲁木齐 830046

摘　要：哈希技术能减少数据存储空间并提高检索性能,因此在计算机视觉中的图像检索领域被广泛应用。随着深度学习研究的进一步开展，深度哈希方法在图像检索范畴内的性能不断提高，但在语义信息利用上的欠缺，仍然可能会产生次优的哈希码。本文中我们提出的深度语义哈希算法（DSHA）基于监督学习，使用成对的标签信息在端到端框架中学习哈希码。 DSHA模型包含三个关键组件：（1）捕捉图像表示的卷积神经网络（CNN）；（2）二进制哈希码通过CNN中修改后的全连接层生成；（3）同时使用成对损失和对比损失对实值输出进行正则化以逼近所需的离散值。在标准图像检索数据集上的实验表明，所提出的DSHA模型与最新的哈希方法相比有实质性的改进。

关键词：深度学习；深度哈希；图像检索；成对标签

论文编号：362

基于跟踪候选目标分类的红外弱小目标检测算法

张荔哲，刘凯

西安电子科技大学计算机科学与技术学院，陕西西安 710071

摘　要：红外弱小目标检测跟踪算法在当下已广泛应用于红外制导等军事和民用领域。对于远距离运动的飞行目标，由于其具有良好的连续性，故基于运动轨迹建模的卷积神经网络适用于判断出目标与噪点的区别。良好的红外弱小目标检测算法是发现目标的基石，鲁棒的目标跟踪算法是得到准确运动路径的屏障，精准的轨迹分类算法是判断目标的决定因素。方法　本文结合红外弱小目标的特征，针对现有算法的一些缺陷，在分割预处理的图像上应用多方向梯度矢量场和通量密度的特征提取算法，在此基础上提出了一种基于对候选目标跟踪所得轨迹分类的红外弱小目标检测方法，通过卷积神经网络对轨迹坐标进行分类以达到筛除噪点保留真目标的目的。结果 通过目标跟踪OTB评估平台评测了算法性能，所提出跟踪算法的性能优于当下最新流行算法，且容易实现。通过对比提出算法与热门红外小目标检测算法的探测率与召回率，本文算法明显可以在保证高探测率的前提下明显减少虚警率，鲁棒性强。结论　本文跟踪轨迹分类算法思路可将特征与目标过于相似的无法单用检测算法区分出的噪点与真实目标准确地区分开，红外弱小目标检测性能优越且工程部署便捷。

关键词：红外弱小目标检测；目标跟踪；卷积神经网络；多方向梯度矢量场；通量密度

论文编号：367

一种距离交并比回归的孪生网络目标跟踪算法

黄智慧，詹瑾*，赵慧民，郑伟俊，吕巨建，李伟健，郑鹏根 1

广东技术师范大学,广东省广州市 510630

摘　要：视频目标跟踪是视频分析和理解视频语义任务的前提和基础，是计算机视觉的一个重要研究分支。在基于深度学习的视频跟踪算法中，损失函数则是模型优化最关键一步，可反映预测模型与实际数据之间的差异，从而衡量训练模型的性能好坏。但是，大多数跟踪算法采用交并比（IOU, Intersection over Union）度量方法定位候选边界框时，无法优化真实目标和候选边界不重叠情况，容易造成目标丢失，导致跟踪失败。为解决上述问题，提出了一种距离交并比（DIOU, Distance Intersection over Union）回归的孪生网络目标跟踪算法。与IOU度量方式比较，DIOU不仅关注真实目标和候选边界的距离，还考虑回归损失的重叠率与中心点距离因素，训练过程收敛速度快且不易发散，提高模型训练效率。本文将DIOU度量方法与损失函数结合，引入孪生网络的回归分支中，通过优化回归损失函数解决交并比度量存在的问题。同时，使用区域候选子网络（RPN, Region Proposal Network）产生候选目标区域，大大降低了时间复杂度，保证跟踪实时性。在OTB 2015标准数据集上与较流行的IOU、GIOU和L1-smooth损失算法对比性能差异，并根据实验结果进行定量和定性分析，评估结果验证了本文算法有效，并在视频包含目标尺度变化、低分辨率、光照变化等干扰情况下，具有较强鲁棒性。

关键词：目标跟踪；距离交并比（DIOU）；孪生网络；深度学习

墙报 1-5

论文编号：370

基于本质造型属性的汽车前脸识别

陈峰蔚，李宝军，孙旭生，宋明亮

大连理工大学,辽宁省大连市 116024

摘　要：针对深度卷积神经网络具有一定的数据依赖性，而实际交通场景中经常出现遮挡以及车标篡改导致汽车前脸识别精度大幅度下降的问题，提出了通过数据定向增强提取汽车本质造型属性的汽车前脸识别方法。本文首先采用深度卷积神经网络进行汽车前脸敏感区域检测，根据事先设置好的锚点位置，通过遍历随机遮挡对比前后识别精度损失分布进行敏感区域定位；针对定位到的敏感区域，通过添加遮挡块以及擦除对应语义部件纹理进行两次数据增强。最后基于增强后的数据进行深度网络模型二次训练，得到本质造型属性识别模型。经过数据增强，本文建立了涵盖国内22类主流品牌、共计22742张图像的FVD22汽车前脸数据库，其中包括原图像数据集、遮挡不变数据集、造型属性数据集以及挑战性测试集。实验结果表明，本文方法在各类深度卷积神经网络上，针对最敏感区域遮挡、车标篡改、真实大面积遮挡情况，分别取得37.08%、34.78%以及5.99%的平均精度提升，较大幅度提升了汽车前脸品牌识别的鲁棒性和性能。其中基于ResNet50的模型效果最佳，在真实遮挡情况下，识别精度达到90.23%。

关键词：汽车前脸识别；本质造型属性；数据增强；深度卷积神经网络（DCNN）

论文编号：388

可变形深度人脸特征插值编解码网络

高谦 1，刘嘉润 1，刘文杰 1，李瑞瑞 1

1.北京化工大学信息科学与技术学院,北京 100029

摘　要：人脸属性编辑是一个具有挑战性的图像处理任务，使用传统的图像处理软件手工编辑，操作成本高。深度特征插值方法是通过深度学习将图像空间非线性的语义映射到隐空间线性特征再进行语义属性编辑的技术。该方法通用性较强，不需要设计特定网络结构，图像处理的速度较快，效果也比较好。遗憾的是，该方法由于难以将特征完全解耦，导致线性插值之后生成图像模糊或者存在伪影。针对这一问题，本文采用可变形的编解码器架构，通过无监督的训练方式，将图像分离为与外观相关的纹理信息和与形状相关的变形信息，解耦图像变形和纹理信息。同时使用与编码器对称的解码器结构，并引入从编码器到解码器的带过滤条件的信息通道帮助重建。本文提出的方法，可以获得耦合度更低、更有效的特征表示方法，重建的图像更加清晰；将该技术应用于人脸语义属性编辑，也能有效地消除伪影现象。

关键词：人脸风格迁移；特征插值；可变形自编码器；无监督解耦；对称解码器。

论文编号：441

用于角度变化下人脸检测的区域式渐进校准网络

孙劲光，黄胜，杨忠伟

辽宁工程技术大学 电子与信息工程学院,辽宁 葫芦岛市 125105

摘　要：任意平面角度下的人脸检测是人脸检测中的开放性挑战问题之一，若实现图像中任意角度下人脸精准检测具有极大的困难。现有的高精度检测模型都需要巨大的参数支撑，才能解决大幅度角度变化的人脸面部检测问题。本文提出的区域式渐进校准网络（Regional Progressive Calibration Network，RPCN），用于任意平面角度变化下的人脸检测，RPCN通过级联网络结构实现逐步降低人脸平面角度变化、提升检测精度，通过降低网络复杂度提升检测速度。具体方法是：首先构造特征提取网络，对人脸图像进行特征表示；然后利用区域生成网络，产生高质量的候选区域；构造级联检测网络，逐步缩小面部平面角度范围，并实现由粗到细地对候选区域进行人脸面部检测；最后调整模型整体复杂度降低检测时耗。其中，我们将特征提取网络的中间层融合以在参数量较少时更好的表示面部特征，调整区域生成网络中锚设置以更好的解决小尺度面部问题。我们在角度增强的FDDB数据集与WIDER FACE数据集上，采用RPCN进行任意平面角度下的人脸检测实验。实验表明，所提出的方法在两个数据及上分别取得了88.3%与87.4%的平均召回率，在与Faster RCNN具有相同召回率且保持检测出小尺度面部的情况下，提升了对平面角度变化下的检测鲁棒性，且检测速度在角度变化增加的情况下，提升了4-5FPS。

关键词：人脸检测；神经网络；机器视觉；级联网络；旋转不变

论文编号：463

YOLOv3 与点偏移估计相结合的车牌定位

徐光柱 1，2，李兴维 1，万秋波 1，匡婉 1，雷帮军 1，2，石勇涛 1，2

1.三峡大学计算机与信息学院,宜昌 443002； 2.湖北省水电工程智能视觉监测重点实验室,宜昌 443002

摘　要：车牌是车辆身份的重要标识，其作用类似于公民身份证。在当今交通状况日益复杂的情况下，快速鲁棒的车牌辨识系统的作用尤显突出，而车牌定位技术是其关键核心之一。尽管小区出入口、ETC收费站等有约束环境下的车牌定位研究目前已获得了长足进步并投入使用；但对于非约束环境下的车牌定位，由于易受到倾角、光照条件、拍摄距离以及抖动模糊等多种因素的影响，到目前为止仍是一项具有挑战性的开放课题。近年来，快速发展的深度学习技术能够通过CNN自动学习多种特征，并表现出了强大的性能优势，因此该项技术被广泛地应用到了车牌等多种目标检测中。但是，目前基于深度学习技术的车牌定位算法大多使用矩形框进行检测，无法精确定位车牌区域。为了实现非约束场景下的车牌精定位，本文在YOLOv3模型的基础上进行了改进，提出了一种YOLOv3与点偏移估计相结合的车牌定位算法，该算法以YOLOv3输出的矩形框的四个顶点为基准点，并行地预测车牌的四个顶点与对应基准点之间的偏移量，然后通过基准点以及相应偏移量计算得到车牌的四个顶点坐标，从而间接实现车牌区域精定位。实验结果表明，本文提出的算法在多种非约束场景下均具有良好的表现，和其它车牌定位算法相比，不仅能够准确地预测出车牌的四个顶点坐标，且在CCPD车牌数据集中的Base、Tilt和Weather子集上的准确率均在99%以上。本文提出的方法除了应用于车牌定位外，还可以扩展到其它领域，这种方法通过同时预测同于定位目标的矩形框以及特征点有效地扩展了YOLO的应用范围。

关键词：深度学习；非约束车牌定位；视觉目标检测；点偏移估计；YOLOv3

论文编号：465

BNNeck 改进批特征消除网络的行人重识别研究

曹 洁 1,2，冯 帆 1，李晓旭 1

1. 兰州理工大学 计算机与通信学院,甘肃 兰州 730050；2. 甘肃省城市轨道交通智能运营工程研究中心，甘肃 兰州 730050

摘　要：行人重识别任务始终存在姿势，遮挡和视角变化的问题，这导致卷积网络训练时一些细小的局部特征往往会被忽略。批特征消除（BDB）网络可以通过分支学习这些被忽略的局部特征，结合全局显著特征获得更具全面性的特征表示。但是，BDB网络中的三元组损失函数和Softmax交叉熵损失函数的优化目标并不在一个相同嵌入空间中，并且BDB网络的三元组损失忽略了类内的紧凑性。本文提出使用BNNeck和中心损失函数改进批特征消除网络。BNNeck使BDB网络在不同的嵌入空间优化三元组损失和Softmax交叉熵损失，中心损失直接约束嵌入空间的类内距离来进一步增强BDB网络所提特征的辨别力。在Market1501、DukeMTMC-reID和CUHK-Detect数据集验证所提方法的效果。BNNeck结构改进BDB网络的方法，Rank-1精度在较少样本的CUHK-Detect数据集有4.7%的明显提升，这表明BNNeck可以显著提高模型的鲁棒性，尤其在防止小样本数据集过拟合方面具有极好的效果。中心损失函数改进BDB网络的全局分支，在Market1501和DukeMTMC-reID数据集明显提升识别精度，这表明中心损失约束特征类内距离，能够提升网络所提判别特征的辨别力。所提方法同时添加BNNeck和中心损失在三个数据集上均取得较好的识别性能，在DukeMTMC-reID和CUHK03数据集上取得了最高识别精度，在Market1501数据集实现了94.9％的Rank-1精度，相比BDB基线网络提高了0.5％。

关键词：行人重识别；卷积网络；损失函数；中心损失；BNNcek

论文编号：476

基于三角建模的困难样本三元组损失的行人再识别方法

曹洁 1,2，丁悦 1，李晓旭 3

1.兰州理工大学电气工程与信息工程学院,甘肃省兰州市 730050； 2. 甘肃省城市轨道交通智能运营工程研究中心,甘肃省兰州市 730050； 3. 兰州理工大学计算机与通信学院，甘肃省兰州市 730050

摘　要：行人再识别是图像和计算机视觉领域的一个重要课题。良好的相似性度量可以提高行人再识别任务图像间的搜索性能，提高模型的再识别准确率。针对行人再识别任务中采用三元组损失所存在的困难样本对的距离区分度不高的问题，本文结合角度损失函数的建模思想，提出一种基于三角建模的困难样本三元组损失函数。该损失函数利用难样本挖掘原理选定三元组损失中的锚样本、正样本和负样本，对这三个样本点建立三角形模型，根据三角形的角与边的对应关系，重新定义样本对之间的距离，使得困难样本对之间的距离更容易区分。由实验结果可知，在与其他损失函数比较时，本文提出的损失函数在数据集Market1501上Rank-1精度可达91.7%，mAP结果可达80.8%；DukeMTMC-reID的Rank-1精度可达85.5%，mAP结果可达71.7%；cuhk03-label的Rank-1结果达到了80.0%，mAP达到了76.8%；cuhk03-detect的Rank-1结果达到了75.7%，mAP结果达到了72.6%。在与其他行人再识别现有方法进行比较时，本文所提出的方法也取得了较好的识别结果。实验表明本文提出的损失函数可以有效提升行人再识别的识别精度，并且在多个数据集上有较好的识别性能。因此，本文所提出的对困难样本三元组损失进行三角形建模的方法可以有效提升模型的识别性能，优化难样本对之间的距离度量，为计算机视觉任务中相似性度量方法的研究提供了新的思路。

关键词：行人再识别；三元组损失；困难样本挖掘；三角模型

论文编号：41

一种通用的头像卡通化方法

刘洁，王涛，刘倩，杨龙飞，卜琪

北京爱奇艺科技有限公司，100080

摘　要：卡通是在日常生活中我们能够广泛接触到的一种艺术形式，除了艺术本身之外，它的适用范围从出版媒体到儿童教育等领域，另外将人物真人头像转换成卡通头像在个人头像生成，用户照片隐私保护，动漫素材制作和娱乐等领域存在较大实用价值，人工制作卡通素材耗费大量时间精力，目前已有研究用真实景物的照片作为源图片，采用图像翻译方式生成任意风格的漫画，同时也有研究采用图像翻译将人物真人头像转换成卡通头像，不过真人头像转换成卡通头像存在转换不稳定，转换的卡通结果和输入源图片不像或者二次元风格缺失等问题，因此目前还缺乏高质量的头像卡通转换方法，我们致力于解决这个问题。我们基于一种新提出的以端到端的方式结合了新的注意模块和新的可学习的归一化功能无监督图像到图像翻译方法U-GAT-IT，提出使用StyleGAN进行人脸属性定制生成各类真人头像数据，同时提出一种男女卡通头像混合训练方法生成和女性卡通风格一致的男性卡通头像数据，实现了一种效果稳定千人千面的人物头部肖像卡通化方法，通过在我们构建的多类测试数据集上进行大量实验，从实验效果的定性分析和定量分析可以看出我们提出的方法在各个数据集各种复杂条件下（比如侧脸、闭眼、肤色变化、光照变化）都能进行稳定的卡通头像生成。

关键词：头像卡通化；无监督图像到图像翻译；注意力；生成对抗网络；U-GAT-IT

论文编号：69

基于自注意力循环网络的无监督单目深度估计

刘香凝 1，2，赵 洋 2，3，王荣刚 1，2

1.北京大学深圳研究生院,广东省深圳市 518055； 2.鹏城实验室,广东省深圳市 518055；

3.合肥工业大学,安徽省合肥市 230009

摘　要：单幅图像的深度估计是场景几何理解过程中的一个重要步骤，但由于尺度模糊，也被计算机视觉领域普遍认为是一个典型的不适定问题。近年来，尽管监督学习方法在单目深度估计中取得了基本令人满意的效果，但需要对数据集进行大量真实深度值的标记，这是一项成本较高的工作。此外，由于物体的运动、遮挡、光照等常见问题，单目深度估计的表现并不尽如人意，尤其是在物体边缘和弱纹理区域。为了解决这些问题,本文提出了一种基于自注意力循环网络的无监督单目深度估计方法。该方法具有以下特点：1）循环网络对训练过程中的深度估计具有较强的约束和监督作用；2）通过掩模加权重构损失和左右视差一致性损失对网络进行优化；3）采用自注意力机制捕捉更多上下文信息，进而提升预测结果。实验结果表明，该方法在 KITTI 数据集上取得了优于现有方法的效果。

关键词：无监督学习；单目深度估计；循环网络；自注意力

论文编号：117

基于深度卷积与全局特征的图像密集字幕描述

武文博 1,2，刘青茹 1,2，顾广华 1，2*，赵志明 1,2

1.燕山大学 信息科学与工程学院，河北 秦皇岛 066004；

2. 河北省信息传输与信号处理重点实验室，河北 秦皇岛 066004

摘 要：密集字幕描述是一种采用自然语言描述图像中不同目标的新兴计算机视觉任务。其主要内容是在图像中找到感兴趣区域，并使用短语或句子来描述。本文研究了密集字幕描述的两个主要挑战，一个是感兴趣区域（ROI），另一个是区域描述。为了解决这两个问题，首先，本文使用深度卷积残差网络来提取图像特征，通过将特征提取网络的层数加深，以便提取具有完整视觉描述的图像特征，从而可以更好地表示图像中存在的高级语义；其次，为了解决计算机难以描述图像中视觉概念以及视觉概念之间的关系的问题，本文提出了一种并行LSTM网络结构来生成图像密集字幕描述，通过将全局、局部和上下文三种不同特征组合，验证了综合特征的有效性。本文在Visual Genome数据集上设计了对比实验，分析了基于Faster R-CNN中NMS不同阈值，特征提取网络层数、proposal数量的影响，结果表明深度特征提取网络有助于更全面特征的获取。同时基于数据集V1.0，与单纯使用局部特征相比，融合特征的实验结果中mAP分数从8.96％提高到10.03％。此外，本文还基于数据集V1.2进行了类似的对比试验，融合特征的实验结果中mAP分数达到了10.39％，进一步证明了本文模型的优势，并且证明特征融合极大地提高了模型的整体性能，从而生成更完整的视觉表示。

关键词：密集字幕生成；并行长短时记忆网络；Faster R-CNN；感兴趣区域；特征融合

论文编号：183

一种用于 COVID-19 预测的图卷积神经网络时空数据学习方法

杨成意 1，2，刘峰 1*，齐佳音 1*，段妍 1，2，吕润倩 1，2，肖子龙 3

1. 上海对外经贸大学 人工智能与变革管理研究院，上海，200336；

2. 上海对外经贸大学 统计与信息学院，上海，201620；

3. 中山大学 数据科学与计算机学院，广州，510006

摘　要：目前对于疾病传播的研究主要集中于时序数据和传染病模型，缺乏运用空间信息提升预测精度的探索和解释，本文为此提出了一种基于图卷积神经网络(GCN)的时空数据学习方法。在研究过程中发现了疫情与地理空间、高铁线路、飞机航线之间的正相关关系，遂将国内各城市之间的空间分布和交通连接关系映射成网络图，并编码成了地理邻接矩阵、高铁线路直达矩阵、飞机航线直达矩阵、飞机航线或高铁线路直达矩阵。按时间窗对疫情数据进行切片后形成张量，依次分批输入到图卷积神经网络中参与卷积运算，为训练和预测的过程提供了更多先验信息，改变了处理时空数据依赖于模型集成的方法，实现了端对端学习的训练框架。通过反向传播和梯度下降训练模型，分别在验证集和测试集上调优参数和进行预测。在新型冠状病毒肺炎疫情数据集上的实验结果显示，采用GCN学习这一时空分布特征，相较于RNN、GRU、LSTM等循环神经网络模型，有着更高的训练精度、更低的训练时间与运算成本。本文的实验及结论将为图深度学习方法在处理时空数据问题上的更广泛应用提供参照，下一步将对图深度学习方法用于回归问题时的性能优化展开探索和研究，最终为更多领域的工商业应用提供理论依据和技术支撑。

关键词：图卷积神经网络；时空特征；新冠肺炎；深度学习；疫情预测

墙报 1-6

论文编号：172

基于密集深度插值的 3D 人体姿态估计方法

陈梦婷 1，王兴刚 1，刘文予 1

1.华中科技大学电子信息与通信学院,湖北省 430074

摘　要：3D人体姿态估计是计算机视觉任务中一直非常具有挑战的任务。由于样本标注难度大，往往只能获得有限场景下的离散关键点数据。而且深度的估计往往是三个维度的预测中最难的。我们发现，虽然人体是一个非常灵活的结构，但是单个躯干可以看做刚体。这意味着当我们只知道躯干两端的深度时，整个躯干的深度都可以通过密集插值得到估计值。因此，我们将每个躯干的密集深度插值特征图作为我们的中间监督。这个特征图为深度的估计提供了更加密集，更加结构化的学习目标，而不仅仅是直接对离散关键点的深度进行回归。为了降低学习难度，我们采用了两个辅助特征图来学习2D信息。通过在数据集Human3.6M上的结果验证了仅仅通过简单的网络结构，模型就能获得先进的性能。在数据集MPI-INF-3DHP上进行的跨域实验进一步证明了模型强大的泛化能力。

关键词：3D 视觉；人体姿态估计；深度密集差值；跨域泛化；

论文编号：437

基于边缘检测生成对抗网络的红外与可见光图像融合

赵康成 1，程江华 1，刘通 1，孙振华 2，邓华付 1

1.国防科技大学电子科学学院,湖南长沙 410005； 2.军事科学院系统工程研究院,北京 100039

摘　要：红外图像拥有物体轮廓清晰、热度信息丰富等优点，但其视觉性模糊、细节信息缺乏。可见光图像却与之相反，因此将二者融合可获得更具有价值的图像。因此提出了一种基于边缘检测生成对抗网络的红外与可见光图像融合方法。该方法思路是在生成融合图像的生成器与判别融合图像是否达标的判别器间建立了一项对抗性博弈。通过该网络生成的融合图像，既有红外图像的宏观特征，又有更加符合人体视觉特征的可见光图像细节特征。通过sobel算子所提取出的梯度信息构造网络的损失函数以约束上述网络得到所需要的效果，本文将上述损失函数与其他融合网络的损失函数相比较，证明使用该种损失函数在源图像线条提取，最终融合后图像线条还原方面的优越性。网络的生成器采用边界平衡对抗网络（BEGAN）的结构，判别器则是采取了类似VGG的网络结构加以判别。实验结果证明在主观视觉评价上使用该种模型辅以基于边缘检测损失函数的融合网络生成的图像边界信息丰富，细节及背景纹理清晰，生成结果较为理想。在客观数据指标分析上，使用该方法可以较好地还原源图像，融合后图像有着较多的信息量。

关键词：边缘检测；sobel 算子；边界平衡对抗网络；VGG 网络

论文编号：37

全卷积神经网络的手写历史文档图像版面分析

应自炉，赵毅鸿，甘俊英，宣晨，王发官

五邑大学智能制造学部,广东省江门市 529020

摘　要：相较于印刷体文档资料，手写体文档资料保存的时间更为悠久，数量更为丰富，并且历史文献资料记载着早期人们创作的宝贵信息，这些宝贵的文献信息对研究社会、自然等发展规律有着重要作用。但由于手写历史文档图像属于非曼哈顿结构，版面结构复杂，在分割上面临书写风格、局部字体偏斜以及文档纸张老化等问题，此外，对象级的目标检测网络难以处理非曼哈顿版面结构的文档图像。针对上述问题，提出空洞残差全卷积神经网络。首先，该模型是像素级的端到端的目标检测网络，不包含复杂的图像预处理过程，模型结构简单；其次，结合空洞卷积和图像级特征对残差块进行改进，从而提出Atrous-Res2Net模块，提高残差块获取感受野的能力；最后，受全卷积神经网络的启发，本文将改进的Atrous-Res2Net模块作为基本网络结构，并将添加了图像级特征的Atrous-Res2Net模块作为整体模型的最后部分，进而提出空洞残差全卷积神经网络，实现对目标对象即文本主体、注释、装饰的定位与识别任务。在DIVA-HisDB历史手写文档图像数据集上的实验表明，提出的算法在均交占比(mIOU)与像素准确率(PA)上分别达到84.50%和97.26%。与改进前相比，本文方法在mIOU与PA上分别有0.45%与0.25%的提升，并且优于本文对比的文档分割算法。

关键词：文档图像处理；版面分析；目标检测；深度学习；语义分割

论文编号：82

基于贝叶斯粗糙集和布谷鸟算法的肺部肿瘤高维特征选择算法

周涛 1,2，陆惠玲 3，张飞飞 4

1.北方民族大学 计算机科学与工程学院 宁夏 银川 750021；2.宁夏智能信息与大数据处理重点实验室 宁夏 银川 750021；3.宁夏医科大学 理学院 宁夏 银川 750004；4.中国电信股份有限公司宁夏分公司 宁夏 银川 750002

摘　要：在高维特征选择过程中最优特征子集生成和分类器参数优化方面，提出一种基于贝叶斯粗糙集（BRS）、遗传算法（GA）和布谷鸟算法（CS）的两阶段优化高维特征选择算法。该算法首先分析 3000 例肺部肿瘤 CT 图像的形状、灰度和纹理特征，提取 104 维特征分量共同量化 ROI；然后进行两阶段优化：（1）从全局相对增益函数的角度分析了属性重要度，结合属性约简长度和基因编码权值函数的加权和构造适应度函数，通过选择、交叉和变异等遗传操作生成最优特征子集，在不降低分类精确度的前提下降低特征维度；（2）利用 CS 对支持向量机（SVM）参数进行全局寻优；最后通过实验验证本文算法的可行性和有效性。实验结果表明，第一阶段采用 BRSGA 进行属性约简克服了 VPRSGA 参数人工设置的局限性，并且在属性约简和分类阶段都表现出较为理想的效果。第二阶段采用 CS 算法优化 SVM 的参数，相比 GS-SVM、GA-SVM 和 PSO-SVM 算法在一定程度上提高了分类精确度，分类时间只高于 GS-SVM，但是相比 GA-SVM 和 PSO-SVM 分类时间显著降低。同时本文的各项评价指标均是多次结果平均值，有效避免片面性的评价，具有一定的科学性。因此，本文提出的两阶段优化高维特征选择算法有效提升了肺部肿瘤良恶性识别能力，降低了算法的时间复杂度，为医学图像特征级处理提供新的思路。

关键词：遗传算法；贝叶斯粗糙集；布谷鸟算法；支持向量机；特征选择

论文编号：121

基于 DenseNet-Attention 模型的高光谱图像分类

张永鹏，张春梅，白静

北方民族大学，宁夏回族自治区银川市 750021

摘　要：针对高光谱图像标记样本量少，提取特征不充分以及提取到的特征不区分贡献度的问题，提出一个新型的 DenseNet-Attention 网络模型（DANet）。首先，该模型利用三维卷积核同步获取联合光谱空间特征，同时 DenseNet 的稠密连接块除了能够充分提取更加鲁棒的特征外，还减少了大量参数；其次，self-attention 机制作为一个模块加入到稠密连接块中，可以使上层提取到的特征在进入下一层网络之前，经过该模块对其进行权重分配，使具有丰富地物类别信息的特征得到加强，进而区分特征的贡献度。该算法以原始高光谱图像邻域块作为输入，无需任何预处理，是一个端对端的深度学习框架。在印第安松树林和帕维亚大学数据集上进行对比试验，结果表明本文所提出的网络模型的分类精度分别能够达到 99.43% 和 99.99%，有效提高了高光谱图像分类精度。

关键词：三维卷积；高光谱图像分类；稠密网络；自注意力机制；

论文编号：142

基于深度学习的甲骨文偏旁与合体字的识别

林小渝 1，陈善雄 1，高未泽 1，邱小刚 1，莫伯峰 2，焦清局 3

1.西南大学,重庆 400715； 2.首都师范大学,北京 100048； 3. 安阳师范学院,河南 455000

摘　要：由于字形结构复杂异体字较多，甲骨文字的识别一直是一个难于解决的问题。本文首次提出以甲骨文偏旁为识别的基本构件，建立单偏旁和合体结构的甲骨文字符识别方法，提升甲骨文识别的精度。方法一：根据甲骨文偏旁字形特点，对甲骨文拓片上的合体字进行甲骨文单偏旁最大极值稳定区域的选取，然后，通过改进的BN-Lenet模型识别甲骨文各个偏旁；方法二：针对甲骨文合体字拓片稀缺的问题，本文提出一种直接对甲骨文合体字进行整体识别的OraNet模型，该模型采用迁移学习的训练策略，对在脱机手写汉字HCL2000数据集预训练的卷积神经网络模型进行参数和结构上的微调，实现迁移得到低层表示和甲骨文合体字集上高层表示的特征融合，以此来提取甲骨文合体字的高级特征。实验结果表明，BN-Lenet网络对甲骨文单偏旁识别率为96.24%，微调的OraNet模型对甲骨文合体字识别率为98.58%，从而表明从甲骨文单偏旁的角度进行甲骨文字形识别，可以获得较高的识别精度。同时本文将甲骨文视为偏旁组合而非整字识别，这使得算法能够识别从未见过的甲骨文新字，即零样本学习，对甲骨文研究有着重要的应用意义。

关键词：甲骨文字符识别；迁移学习；偏旁分析；非极大值抑制算法；零样本学习

论文编号：363

T2AD:一种面向自动驾驶的测试与测试控制语言

董乾 1，2，杨光 1，薛云志 1，2，孟令中 1，2，师源 1，丁明明 1

1. 中国科学院软件研究所，北京 100190； 2. 计算机科学国家重点实验室，北京 100190

摘　要：自动驾驶作为环境感知、行为决策与车辆控制的综合集成性技术，对其进行充分的、系统的测试和验证，是其能够大规模应用的技术前提之一。但目前的测试与验证中存在测试用例固定受限、不同模拟器间不兼容、形式复杂不便迁移的问题。针对上述问题，本文提出了一种用于自动驾驶测试与测试控制表达的语言T2AD（Testing and Test Control Notation of Autonomous Driving），并结合仿真测试环境中实现了T2AD的解释器与执行器。T2AD能够在抽象语义层面描述所需的测试场景，通过时序关系的声明控制所定义物体的行为，并灵活描述测试判据。本文实验表明，T2AD能够在仿真环境中生成用户定制的测试用例，可对利用仿真环境构建自动驾驶领域基于模型的测试与验证方面发挥至关重要的作用。

关键词：仿真环境；自动驾驶；语言设计；测试和测试控制

论文编号：417

支持力反馈的沉浸式物理学习环境

郑明钰，李家和，张晗，骆岩林*，申佳丽，朱小明*

北京师范大学人工智能学院，北京市海淀区，100875

摘　要：随着虚拟现实技术的发展，沉浸式学习环境在教育教学领域的应用需求日益增大，前景日趋广阔，如物理实验仿真教学。然而，现有虚拟学习环境大多只能提供视觉与听觉的交互，不支持力觉交互，逼真感和沉浸感不足。本文将力反馈技术应用于虚拟环境，描述了一个支持力反馈的沉浸式物理学习环境。项目组首先对中学物理实验教学进行充分调研，梳理相关知识点并设计出虚拟学习环境的场景；然后利用Touch力反馈设备替换HTC Vive虚拟现实设备中的操控手柄，在虚拟现实环境中为用户提供力觉交互方式，同时借助Unity3D软件和OpenHaptics力觉函数库，实现典型案例:重力和摩擦力实验；最后项目组参考相关文献，建立了虚拟物理学习环境的评价指标体系，从多方面对沉浸式物理学习环境和当今使用较为广泛的“NoBook” 虚拟物理实验室进行对比考察，实验结果表明：相比于传统的虚拟物理学习环境，支持力反馈的沉浸式物理学习环境具有更好的沉浸感和自然交互性，可加深学生对抽象物理概念和规律的理解，提高学习兴趣和自主性，同时利用计算机技术构建的虚拟仿真试验环境，可以为学生提供更多的实验机会，为解决物理教学中的“实验一遍过”甚至“无实验”等问题提供新思路，是虚拟现实和力反馈技术的有效尝试与创新。

关键词：沉浸式学习环境；力反馈；虚拟现实；中学物理教学；

论文编号：173

基于沉浸式增强现实的流场可视化方法研究

许世健 1，吴亚东 1，3，赵丹 2，王赋攀 1

1.西南科技大学,四川绵阳 621010； 2.中国空气动力研究与发展中心,四川绵阳 621000； 3.四川轻化工大学，四川自贡，643002

摘　要：针对目前二维环境中流场可视化应用深度信息不足、交互不自然等问题，提出一种基于沉浸式增强现实的流场可视化方法，在现实空间中通过流线展示流场变化特征，利用沉浸式增强现实技术为用户提供有效的深度信息，并通过手势、凝视等自然交互方式进行流线布置。由于增强现实设备不具备流场可视化计算所需要的计算能力，该方法采用服务器/客户端架构，以服务器端进行流线计算，以头戴增强现实设备进行流线定位、结果绘制及用户交互。在此基础上，研究开发了原型系统，设计了基于用户凝视的流线布置和基于凝视与手势的种子点放置两种交互方式，利用手势与凝视交互的自然性，提高用户的流线布置及分析效率。实验结果表明该方法能够在增强现实环境中准确完成流线绘制，利用沉浸式增强现实环境中的深度信息，反映流场的变化情况，并提供快捷、自然的交互。

关键词：流场可视化；沉浸式；增强现实；流线法；人机交互

论文编号：54

一种基于虚拟现实系统的船舶数字孪生框架

景乾峰 1，神和龙 1，尹勇 1

1. 航海动态仿真和控制实验室，大连海事大学，辽宁，大连 116026

摘　要：数字孪生是真实船舶的动态虚拟再现，它利用来自传感器数据和物理模型的综合知识来模拟船舶的运行状态。作为典型的虚拟现实系统，模拟器为船舶数字孪生应用的落地提供了良好的培育土壤。现阶段基于物理模型的船舶运动预测主要在固定的模式下进行，例如旋回和Z形操纵，由于缺乏实船数据，鲜在自由操纵下的预测。同时，现有预测中的环境通常基于恒定假设或单一影响，未能体现真实的海洋情况和环境耦合影响。本文提出一种基于虚拟现实系统的船舶数字孪生框架，首先基于海洋数值模型构建实海域环境数据库，通过时空变化的风、浪和流场以重现真实海洋环境，确保虚拟海洋环境真实海洋环境的物理一致性，提升虚拟海洋环境的物理真实感；建立船舶水动力数据库来保持船舶运动特性的一致性；创新性地将从传感设备测量的离散控制信号引入物理模型，保持虚拟和实际船舶的在控制输入的一致性；最后通过传感器数据和物理模型分别建立基于数据驱动和基于模型预测的虚拟船舶，并在船舶模拟器中对实海域环境和虚拟船舶进行渲染，实现船舶数字孪生。此外，展示了基于该数字孪生框架对一艘商船的航程重现分析，并利用实船采集的数据进行了对比验证，证明了模型预测结果的准确性。航程分析结果说明了该框架能够有效地对航行状态进行实时监控，其中物理模型的预测结果可与恶劣海况下的工况极限等知识进行集成，以提供早期预警信息，增强智能船舶的态势感知能力。

关键词：虚拟现实系统；数字孪生；传感器；实海域数据库；物理模型；智能船舶

论文编号：144

一种基于三维模型设计的增强现实手术引导方法

陈琳 1，李衡 1*，邵龙 1，崔建民 1，王涌天 1，杨健 1

北京理工大学 光电学院，北京 100081

摘　要：增强现实(Augmented Reality，AR)技术是医学研究的热门课题，它可以通过直观地为医生提供信息来辅助手术。然而，关于AR辅助医疗的研究缺乏虚拟图像与现实场景准确便捷的融合方法，因而限制了其临床应用。为了探究虚实融合在医学应用中局限性的解决办法，本文提出一种利用头骨三维模型针对性设计咬合夹板，以此来限定标志物位姿实现稳定虚实融合的方法，取代了术中佩戴咬合夹板二次拍摄CT的常用方案。创新性的提出了根据CT影像建立头骨三维模型，以此为依据针对性的设计能与牙齿完全契合的咬合夹板并打印，进而固定头骨三维模型与定位标识物间位置关系，使其与实际空间中患者头骨与咬合夹板间位置关系保持一致，通过识别定位标识物位姿实现将虚拟模型精准的融合到头骨上。此外本文与现有基准方法进行了比较，在实验头骨模型上选取11个点，在-60度到60度的区间内每间隔10度分别测标识点 的虚实融合误差，并记录整个验证实验过程中跟踪丢失次数，最终数据显示两种方法误差均基本控制在一毫米内，429次取点误差实验中丢失跟踪次数分别为本文提出方法13次，基准方法11次。实验结果表明，该方法在明显简化操作过程、减少患者受辐射剂量以及缩短患者在医院所需时间的情况下，达到与基准方法精度及跟踪操作稳定性相当的效果，具有临床应用价值。

关键词：增强现实;手术引导;3D 打印;咬合夹板

墙报 2-1

论文编号：8

基于 CRNN+CTC 印刷体维吾尔文识别系统的设计与实现

熊黎剑 1，2，3 吾守尔·斯拉木 1，2，3 许苗苗 1，2,3 汤敬 1，2，3

（1.新疆大学 信息科学与工程学院，乌鲁木齐 830046；2.新疆多语种信息技术实验室，乌鲁木齐 830046；3.新疆多语种信息技术研究中心，乌鲁木齐 830046）

摘　要：光学字符识别技术（OCR）在图书数字化，文献管理和古籍保护等诸多领域得到了广泛应用，相比于已十分成熟的中文、英文印刷体识别系统，小语种（如维语）印刷体识别还有研究空间，尤其是在国家“一带一路”和“西部大开发”战略的推动下，民文信息化已成为了现实和迫切的需求。针对维吾尔文有效识别的问题，本文结合日益兴起的深度学习技术，采用 Python 语言编程，Django 框架设计前后端，选用卷积循环神经网络和连接时序分类器（CRNN+CTC）作为识别算法核心。实验合成了包含 32 个维文字母在内，共约 10 万张单词图片及对应的标签数据作为训练数据，又从天山网（维文版）上获取了真实数据制作成测试数据，结果表明，端到端的印刷体维吾尔文识别系统在保证数据均衡的同时，测试识别精度达到 94.1%，平均速度达到 10fps。该系统能实现在网页上的端到端维文整词识别，具有良好的应用价值。

关键词 ：OCR；图像识别；CRNN；CTC；维吾尔文；端到端

论文编号：36

基于语义-表观特征融合的无监督前景分割

李熹 1，马惠敏 2，*，马洪兵 1，3，王弈冬 1

1.清华大学,北京市 100084； 2.北京科技大学,北京市 100083；3.新疆大学，新疆省 830046

摘　要：前景分割是图像理解领域中的重要任务，在无监督条件下，由于不同图像、不同实例往往具有多变的表达形式，这使得基于固定规则、单一类型特征的方法很难保证稳定的分割性能。针对这一问题，本研究提出了一种基于语义-表观特征融合的无监督前景分割方法（SAFF）。在研究中发现，从图像类别信息中学习到的语义特征能够对前景物体关键区域产生精准的响应，而以显著性、边缘为代表的表观成像特征则提供了更丰富的细节表达信息。为了融合了二者的优势，研究建立了融合语义、表观信息的一元区域特征和二元上下文特征编码的方法，实现了对两种特征表达的全面描述。接着，一种图内自适应参数学习的方法被设计，用于计算最适合的特征权重，并生成前景置信分数图。进一步的，分割网络被使用来学习不同实例间前景的共性特征。通过融合语义和表观特征，并级联图内自适应特征权重学习、图间共性特征学习的模块，研究在PASCAL VOC 2012数据集上取得了显著超过基线的前景分割性能。

关键词：前景分割；无监督学习；语义-表观特征融合；自适应加权

论文编号：40

印刷体藏文文档图像行字粘连切分与识别

陈洋 1,2，胡鹏飞 1,2，蔡正琦 1,2，王维兰 1,2,*

1.中国民族语言文字信息技术教育部重点实验室（西北民族大学），中国 兰州 730000

2.西北民族大学 数学与计算机科学学院，中国 兰州 730000

*通信作者:wangweilan@xbmu.edu.cn

摘　要：文档图像中粘连断裂字符的切分与识别，是很多OCR应用系统中存在的主要困难之一。本文针对行、字粘连以及笔画断裂的印刷体藏文，提出了一种结合基线与连通域重心的藏文文档行字切分方法。首先，获取文本行基线，依据行基线信息完成相邻行粘连的连通域分割，并结合连通域重心完成藏文文档的行切分；其次，对切分出来的行进行连通域分析滤除音节点及标点，并估算剩余连通域的平均宽度，依据平均宽度完成粘连字切分之后恢复滤除音节点及标点完成文档图像的字切分。同时，构建了一个印刷体藏文字丁数据集（简称TCDB）用于学习训练，基于TCDB搭建卷积神经网络模型（CovNet），采用CovNet模型对文档图像行字切分结果进行字丁识别。实验表明，所提出的方法能够较好的克服印刷体藏文字符的粘连断裂问题，在实际样本上的行切分率达到99.07%，字切分率为98.89%，识别率为99.15%。

关键词：藏文识别；基线检测；行字切分；TCDB；卷积神经网络

论文编号：221

支持差分隐私的图像数据挖掘方法研究

杨云鹿 1，周亚建 2

1,2.北京邮电大学,北京 100876

摘　要：随着大数据与人工智能技术的广泛应用，数据背后蕴藏的巨大价值得以挖掘，同时带来也带来了不容小觑的隐私泄漏问题。如何在保障数据安全的前提下，实现大数据开放共享、高效挖掘日益成为重要的研究领域。针对数据挖掘模型中存在的隐私泄漏问题及现有隐私保护技术的不透明性，本文将差分隐私与图像生成模型(GAN)相结合，提出了一种更具普适性的支持图像数据隐私保护的数据挖掘模型(IDP-GAN)，能够有效地实现数据安全性与可用性之间的平衡。IDP-GAN采用在深度学习模型训练过程中引入差分隐私的方法，通过差分隐私的拉普拉斯实现机制，将拉普拉斯噪声合理地分配到判别器的仿射变换层的输入特征以及输出层的损失函数的多项式近似系数中，以提高GAN的隐私性。更重要的是，IDP-GAN在实现差分隐私保护的同时，有效地减少了训练过程中隐私预算的消耗，提高了生成数据的效用。通过在标准数据集MNIST和CelebA上的实验，验证了IDP-GAN可以生成更高质量的图像数据，此外用成员推理攻击实验证明了IDP-GAN具有提供较好的抗攻击能力。

关键词：差分隐私；数据挖掘；图像生成；生成对抗网络

论文编号：71

面向智慧生物实验室的弱外观多目标轻量级跟踪网络

宗佳平 1，吴妍 2，陈建强 3，张琳娜 3，张悦 1，岑翼刚 1

1. 北京交通大学计算机与信息技术学院，北京，100044
2. 贵州省贵阳市公安司法鉴定中心，贵州贵阳，550025
3. 贵州大学机械工程学院，贵州贵阳，550025

摘　要：基于监控视频的弱外观多目标跟踪是建设智慧生物实验室的一个重要内容。但是，由于遮挡、目标外观差别细微等因素的影响，容易出现漏检、误检等问题，导致跟踪失败。此外，基于深度学习的相关算法需要大量的计算量，在嵌入式平台上难以达到实时性。因此，本文提出了一种新的轻量级多目标跟踪算法，以YOLOv3作为基础目标检测网络，提出基于归一化层权重评价的层剪枝算法压缩检测网络计算量，以提高该算法在嵌入式平台上的运算速率。同时，基于已有的跟踪结果，对当前帧检测结果进行校正，实现对漏检目标的补偿校正，用于提高检测的准确性。最后利用卷积神经网络来提取目标特征，融合目标特征及候选框与预测框间的交并补(IOU)，进行数据关联。实验结果表明，本文提出的轻量级多目标跟踪算法较已有的多目标跟踪算法取得了较好的跟踪结果，且在检测网络压缩率达到95%的情况下，网络大小、参数量及推理速度大大减少，而mAP、查准率及召回率仅是略有降低。实现了在仅损失较少精度的情况下，网络压缩率高，适于嵌入式平台前端实现。

关键词：多目标检测；多目标跟踪；压缩剪枝；智慧实验室；嵌入式平台

论文编号：85

基于残边稀疏机制的红外弱小目标检测

薛伟 1，2，钟平 2，张文生 3，喻高航 4

1.安徽工业大学计算机科学与技术学院,安徽 马鞍山 243032； 2.国防科技大学 ATR 重点实验室,湖南 长沙 410073； 3.中国科学院自动化研究所,北京 100190； 4.杭州电子科技大学理学院,浙江 杭州 310018

摘　要：红外弱小目标检测是红外探测领域的研究热点之一。由于红外检测具有抗干扰性强等优点，已经成为搜索跟踪、导弹制导、空中预警等多个作战任务的关键技术。受大气辐射等外部因素的影响，经系统成像后，红外图像中的目标呈现“弱”和“小”的特征，给复杂背景下的目标检测带来了许多困难和影响。本文提出了一种残边稀疏化的红外弱小目标检测模型。方法 首先，利用红外块图像模型将待检测红外图像建模成背景块、目标块和噪声块三部分之和；然后，利用矩阵的-范数和-范数的加权组合对残边进行约束，并结合加权目标块图像技术以更好地去除目标图像中的残边；最后，考虑到模型具有多变量和非光滑等特性，采用交替优化策略对所提模型进行求解，进而获得目标图像。结果 通过在海天、海面、浓云和空天等4种不同场景下的红外图像进行实验验证，绘制了目标图像的3D强度分布，给出了背景抑制因子值和噪声块图像稀疏度。从结果可以看出，利用本文算法得到的噪声块图像更稀疏，且相较于对比算法，本文算法提取到的目标效果更好。结论 本文提出的基于残边稀疏机制的红外弱小目标检测方法，利用矩阵的加权组合范数有效抑制了复杂背景中的噪声和杂波，有效避免了目标图像中残边的出现，在不同场景下均能得到较好的检测效果。

关键词：红外图像；弱小目标检测；红外块图像模型；稀疏表示；交替方向乘子法

论文编号：98

一个噪声鲁棒的轻量级人脸识别网络

孟祥邦 1，2，严严 1，2，王菡子 1，2

1(厦门大学 信息学院，厦门 福建 361005)

2(厦门大学 福建省智慧城市感知与计算重点实验室，厦门 福建 361005)

摘　要：在现实环境中，人脸图片往往受到噪声的污染，导致人脸识别性能严重降低。本文提出了一个基于注意力机制的噪声鲁棒的轻量级人脸识别网络，其包含去噪子网络和人脸识别子网络。具体来说，我们使用级联的方式将两个子网络进行端到端的联合训练。同时，我们采用了通道注意力模块来显著增强去噪子网络中有效特征图的利用率，并强化去噪效果。此外，为了有效减小网络模型的参数量，我们基于深度可分离卷积设计了一个轻量级的人脸识别子网络。在多个公开的人脸数据集的实验结果表明，我们的方法在识别准确率和模型参数量等方面取得了良好性能。

关键词：卷积神经网络；人脸识别；图片去噪；轻量级网络；注意力机制

论文编号：119

Integration Cosaliency Prior and Random Walks for Image Cosegmentation

Haiping Xu1, 2, Geng Lin1, Meiqing Wang3

1.College of Mathematics and Data Science, Minjiang University, Fuzhou 350108;

2.The School of Electronic Engineering, University of Electronic Science and Technology of China, Chengu 611731;

3.College of Mathematics and Computer Science, Fuzhou University, Fuzhou 350108.

摘　要：In recent years, image cosegmentation has become an emerging issue in image segmentation. Different from the traditional single image segmentation, image cosegmentation aims at discovering the shared objects from an image group containing two or more relevant images, while the shared objects are entirely unknown. Most image cosegmentation approaches need the constraint of significant foreground and dramatic background in images. To relieve this constraint, this paper proposes a random walk based cosegmentation technique, which utilizes the relationship between cosaliency images as prior information to find corresponding seed points to achieve higher cosegmentation accuracy. According to the prior knowledge of the class membership of seed nodes, seed nodes in the same class are marked by the same walker, and then the supervision of seed nodes is introduced into the random walk process through the interaction of multiple walkers. Walkers are propagated from the seed nodes to every node in the graph. To get satisfactory interactions among these walkers, a restart rule is designed, which is composed of a repulsive interaction and a concurrence relation. The goal of repulsive interaction is to find the dominant walker for every node. The concurrence relation represents the similarity of foreground objects among images. Experiment results demonstrate that the proposed technique achieves promising cosegmentation performance.

关键词：Image Cosegmentation；Multiple Random Walkers；Cosaliency；Random Walk with restart

论文编号：120

融合边缘计算和改进 Faster R-CNN 的变电站设备典型视觉缺陷检测系统研究

尹子会 1，孟蓉 1，范晓丹 1，李冰 2，赵振兵 3

国网河北省电力有限公司检修分公司,河北石家庄 050070； 2.华北电力大学科技学院,河北保定 071000；
3.华北电力大学电子与通信工程系，河北保定 071000

摘　要：随着巡检机器人和巡检无人机在变电站巡检作业中的应用，其搭载的传统巡检装置在变电设备典型视觉缺陷检测中快速定位及缺陷检测的自动化程度和效率较低。针对该问题，提出了一种融合边缘计算和改进Faster R-CNN的变电站设备典型视觉缺陷的检测系统。首先，对采集到的存在典型视觉缺陷的图像通过反转、平移、旋转、缩放等操作对数据集进行扩充，通过对Faster R-CNN方法中的非最大值抑制（NMS）进行改进，通过加入一个权重系数λ减缓置信度降权时的衰减；对锚点框生成机制进行了改进，设置了不同的锚框长宽比和锚框尺寸，并对改进后的检测算法模型在服务器上利用NVIDIA GeForce RTX 1080Ti GPU进行训练后存储到边缘设备中。巡检时，将边缘设备部署到巡检机器人或巡检无人机上，直接对采集的变电站设备视频图像进行典型缺陷检测，并将检测出的缺陷进行标记后发送至云服务器。与传统方法的直接将图像信息传送至云服务器进行检测方式相比，数据传输量减少约90%以上，大大降低了变电站巡检工作人员的工作量。实验结果表明，该系统在变电站设备典型视觉缺陷中具有检测准确率高、实时性强、占用带宽小等特点，有效提高了变电站巡检的工作效率和智能化水平。

关键词：边缘计算；改进 Faster R-CNN；变电站巡检；视觉缺陷检测

论文编号：131

基于 GIOU 损失的轻量级人车检测模型

马俊才 1,2，赵耀 1,2，庞华鑫 1,2，韦世奎 1,2

1.北京交通大学 信息科学研究所，北京 100044；2.现代信息科学与网络技术 北京市重点实验室，北京 100044；

摘　要：随着通用目标检测算法的发展，行人检测算法有了很大的进步，然而当前的行人检测算法采用的特征提取网络参数量很大，这导致训练的模型虽然在精度上达到了工业需求，但是网络计算量巨大让训练的模型难于部署，如何让模型在满足精度要求的同时仍能满足部署需求成为当前的一大难点。为了解决这一问题，本文提出了一种轻量级人车检测网络，通过采用通道剪枝策略对特征提取网络进行网络压缩，同时为了提升压缩后网络的效果，我们在训练阶段引入了GIOU损失函数代替原有的均方误差损失，使预测框回归的优化目标更为合理，在后处理阶段为了缓解遮挡问题带来的影响引入了Soft NMS代替传统的NMS。我们在自制的行人车辆数据集上验证了我们方法的有效性，在模型FLOPs小于10G的情况下mAP达到了85.58%，同时满足了工业部署的实时性、计算量、精度的要求。

关键词：人车检测；网络压缩；GIOU；Soft NMS

墙报 2-2

论文编号：136

一种双跟踪器动态级联的视觉跟踪方法

黄文举，董艳梅，武玉伟，贾云得

北京理工大学 计算机学院 智能信息技术北京市重点实验室，北京 100081

摘　要：为了提升视觉目标跟踪准确性，本文提出了一种双跟踪器动态级联的视觉目标跟踪方法，通过粗跟踪器和细跟踪器之间的相互协同实现准确的目标跟踪，其中粗跟踪器用以获取鲁棒的跟踪结果，细跟踪器用以修正粗跟踪器输出的边界框。为了保证算法的鲁棒性，采用了动态的级联策略，通过一个评估网络评估每个跟踪器的跟踪结果，并根据评估结果删减病态跟踪器。在离线训练集上训练细跟踪器和评估网络，学习不同目标的先验知识，测试时用视频的初始帧学习特定目标的表观模型。本文方法在OTB-2013数据集上取得0.966的DP分数和0.723的AUC分数，在OTB-2015数据集上取得0.925的DP分数和0.696的AUC分数，在VOT2017数据集中取得0.2956的EAO分数、0.5307的Overlap分数和17.05的Failures分数，超越了当前许多公认的方法，验证了本文方法可以有效地提升视觉目标跟踪的准确性。

关键词：视觉目标跟踪；目标定位；双跟踪器；动态级联；强化学习

论文编号：137

基于聚合局部相关性和全局网络流的多目标跟踪

赵大伟 1，戴斌 1，肖良 1，刘帅 1，许娇龙 1，彭鹏 1

1.国防科技创新研究院,北京 100071

摘　要：针对多目标跟踪中的两个关键问题：1）提出一种计算相邻两帧之间目标相似性的聚合局部相关性描述子；2）利用全局网络流方法在线估计目标轨迹。首先，在二维图像和三维点云中进行目标检测，为了提高图像中弱小目标和遮挡目标检测的准确率，本文引入循环卷积神经网络，在图像检测的基础上，三维目标可以在二维目标投影的锥体空间内进行检测。其次，为了提升鲁棒性，利用聚合局部相关性描述子对局部相关性进行建模，该描述子由视觉和几何信息组成，包括目标的表征信息，二维/三维形状和位置信息。最后，将多目标关联建模为最小费用流问题，利用最短路径算法求解。算法在国际公开数据库KITTI上进行测试（缩写为MCT），实验结果表明，本文算法性能优异。

关键词：多目标跟踪；目标检测；局部相关性；网络流

论文编号：140

面向图像检索的特征选择方法

李国祥 1,2，王继军 2,3，马文斌 1,2

1.广西财经学院 教务处，南宁 530003；2.广西多源信息挖掘与安全重点实验室，广西师范大学，桂林 541004； 3.广西财经学院 信息与统计学院，南宁 530003

摘　要：针对大规模图像检索和分类所带来的时间和空间挑战，本文从减少特征数量这一简单思路出发，提出了一种有效的连通图特征点选择方法。结合最近邻单词交叉核、特征距离、特征尺度等属性，构建包含若干个连通分支和平凡图的像素级特征分离图，利用子图特征点的逆文本频率修正边权值，从各连通分量的节点数量和孤立点最近邻单词相关性两个方面开展特征选择，将问题转化为在保证图像匹配精度情况下，最小化特征分离图的阶。实验证明了大连通域的冗余性和孤立点的可选择性，约简特征在Oxford5k和Paris中的检索精度基本与原始特征持平，特征约简50%以上，且部分类别效果优于原始特征和其他编码方法。

关键词：词袋模型；特征选择；图像检索；连通分量

论文编号：258

基于多尺度特征融合的火灾检测模型

张坚鑫，郭四稳，张国兰，谭琳

广州大学计算机科学与网络工程学院,广东省广州市 510006

摘　要：火灾的发生对于人民的生命及财产安全都会造成较大的危害，随着科技的进步，火灾检测设备与方法层出不穷。为了提高火灾检测的准确率，降低火灾的漏报率，使用图像型火灾检测方法逐渐成为主流的检测方法，深度学习的发展更是为火灾检测提供了新的思路与方法。改进的两场景检测模型 Faster R-CNN能够在火灾检测中发挥巨大的作用。首先使用 Resnet101 作为特征提取网络，接着使用特征金字塔结构 FPN 提取了 Resnet101 的浅层特征和高层特征，将 Resnet101 的浅层特征图输入 Inception Module 结构提取多种尺寸的卷积特征，最终使用像素注意力机制和信道注意力机制对目标位置进行强化并弱化其余部分，使得检测目标更加精确。该网络避免了主干网络特征提取不充分的问题，融合了多种尺度的特征来区分火灾区域和非火灾区域，有效提高了火灾图像数据集的检测准确率。该模型的平均检测准确率 MAP 相对于Resnet101 特征提取网络有了9个百分点的提升，相对于 Resnet101 加入特征金字塔结构网络 FPN 提升了5个百分点。该模型的检测效率与 Resnet101 特征提取网络、Resnet101 加入特征金字塔结构网络 FPN 都接近，检测一张火灾图像所需时间在0.8秒到0.9秒之间。该模型能够检测出室内和室外火灾的目标区域，精度相对原始模型得到较大的提升，效率与原始模型接近。

关键词：深度学习；火灾检测；卷积神经网络；多尺度特征；特征金字塔结构

论文编号：207

一种基于卷积神经网络的轨道交通场景人群计数模型

杨路辉 1，湛忠义 1，潘尚考 1，刘光杰 2

1.南京理工大学自动化学院,南京 210094； 2.南京信息工程大学电子与信息工程学院,南京 210044

摘　要：随着城市轨道交通的发展，越来越多的人开始使用乘坐轨道交通，尤其是地铁出行。为了防止地铁站内发生踩踏等安全事故，需要对地铁站内实时客流进行实时感知和分析。现有的人群计数方法应用在轨道交通场景中，往往存在场景不适应且计算复杂度太高的问题，针对上述问题，本文提出了一种基于卷积神经网络的人群计数算法，可适用于摄像头低角度与行人互相遮挡的地铁站内场景。本文模型采用VGG16作为前端网络提取浅层特征，使用基于Inception改进的M-Inception结构结合空洞卷积构成后端网络，可增大感受野，适应不同监控角度下不同尺寸的行人目标。为了提高模型训练精度，本文提出一种融合行人总数估计损失和密度图损失的加权损失函数。本文提出的算法在公开数据集和本文收集的地铁场景数据集上与两种现有模型进行了对比实验，实验结果表明，本文提出的人群计数算法的平均绝对误差和均方误差表现均优于两种对比模型。考虑到模型的实际应用，本文将模型部署到海思嵌入式芯片上进行实验测试，结果表明，本文模型能够在嵌入式芯片上取得较高的计算速度和准确率，可满足实际应用场景的需求。

关键词：人群计数；地铁场景；空洞卷积；嵌入式实现

论文编号：211

基于镜像蒸馏的人体姿态估计算法

宾言锐，桑农

华中科技大学人工智能与自动化学院图像信息处理与智能控制教育部重点实验室，湖北武汉，430074

摘　要：人体姿态估计，指根据人体解剖结构定义的人体骨架特征，通过在图像上定位人体关键点像素坐标来构建人的整体姿态，是一个关键点位置的定位问题。当前的方法大多采用多阶段的结构。每个阶段都是一个简单的轻量级网络。各个阶段之间的特征图（含热图）保持高分辨率，并被同时监督，以进行端到端的训练。在测试阶段，往往将最后一个阶段的预测作为模型最终的预测。我们观察到，当我们为每个关键点从多个阶段的预测中挑选出最优的预测结果时，网络整体的性能将得到显著的提升。这说明当前的多阶段结构的性能还存在提升空间。基于上述观察，提出一种新的多阶段网络训练策略，称为镜像蒸馏策略。它基于知识蒸馏的思想，该思想已经成功地用于图像分类深度模型。我们顺序的训练一个镜像模型和一个本体模型。镜像模型和本体模型具有完全相同的网络结构。镜像模型的训练遵循常规的多阶段结构训练流程，即对每个阶段的输出热图计算与真值热图的L2损失。随后，我们根据真实关键点坐标，从镜像模型多个阶段的预测热图中，为每个关节点挑选出最优预测热图。这些最优的预测热图中隐含了镜像模型学到的姿态结构信息。最后我们利用一个姿态知识蒸馏目标函数，将镜像模型学到的姿态结构信息迁移到本体模型，以提高本体模型的性能。我们在MPII数据集和LSP数据集上进行了实验并证明了镜像蒸馏策略的有效性。

关键词：人体姿态估计；知识蒸馏；多阶段结构；

论文编号：219

基于小波变换结合显著性检测分割的热带气旋客观定位方法

王晓洁[1]，张长江[1]，薛利成[1]，马雷鸣[2]，鲁小琴[3]

1.浙江师范大学物理与电子信息工程学院,金华 321004； 2.上海中心气象台，上海 200030；3.中国气象局上海台风研究所,上海 200030

摘　要：本文提出了一种改进的偏差角方差热带气旋中心定位的方法。首先，从原始红外卫星云图中截取包含TC主体云系的感兴区域,并利用显著性检测结合小波变换的方法进行图像预处理。显著性检测分割算法分割得到云系前景以及其中心，小波变换分割算法提取红外亮温变化剧烈的位置。用显著性检测分割算法得到前景中每个超像素块的分类特征值与对应每个像素块平均位置值相乘求和，再除以所有前景超像素块特征值之和，得到位置值代表该算法确定的前景中心，然后以该中心为质心选择合适尺寸的区域作为TC中心检测区域。接着，将显著性检测后得到的二值图像与小波变换分割后的二值图像做相与计算，确定密闭云区所在的区域，便得到仅含有亮温变化剧烈位置的密闭云区的二值图像。最后，在检测区域设定一个参考中心计算密闭云区的偏差角矩阵，遍历检测区域每一个点得到检测区域偏差角方差矩阵，矩阵中最小值的位置即为TC中心位置。**结果：**直接用偏差角定位方法测得的有眼热带气旋中心位置对比中国气象局（CMA）、日本气象厅（JMA）和美国台风预警中心（JTWC）最佳路径数据的平均偏差为44km，无眼较强热带气旋的平均定位偏差为71km。本文方法对有眼热带气旋中心位置的平均定位误差为28km，无眼较强热带气旋的平均定位误差为47km。**结论：**实验表明，本文方法与现有同类方法的定位效果相比，是一种很有效并且具有较强潜力的TC定位方法。与没有经过显著性检测和小波变换分割等预处理的定位方法相比，本文的定位方法能明显降低定位误差。本文方法定位与CMA和JMA的偏差小，与JTWC的偏差相对较大，因为本文方法所用的实验图像均来自西北太平洋，因此证明了本文方法具有较高的可信度。

关键词：热带气旋定位；图像分割；显著性检测；小波变换；偏差角方差；

论文编号：229

一种改进的脱机手写体文本行切分算法

伊力亚斯·苏来曼 1，帕力旦·吐尔逊 1，阿不都萨拉木·达吾提 1，

艾斯卡尔·艾木都拉 2

1.新疆大学软件学院,新疆乌鲁木齐市 830046； 2. 新疆大学信息科学与工程学院,新疆乌鲁木齐市 830046

摘　要：手写体文本行切分在手写体文字识别和手写文档关键词检索中都是最重要和基本的一步。切分质量的好坏直接影响了后续工作。简单的投影方法对于印刷体文本行切分非常有效，但是在手写文档中存在着大量不规范的书写形式，所以其对于手写体来说效果并不理想。通过分析现有投影法的不足之处，本文提出一种分段式自适应阈值的手写体文本行切分算法。方法 首先在预处理阶段，经过四个步骤对原始的手写体文档进行优化，抑制手写体文本图像中的离散噪声点以及可能影响切分的复杂笔迹等杂波信号。之后通过计算文本行中最长一行的长度，然后根据此长度把手写体文档分割成等长的子文档并且将子文档中的每一行根据改进的自适应阈值进行切分。结果 实验在三种不同的数据集上与现有的方法做了对比。在维吾尔文数据集中达到了99.3%的正确切分率，与最新的方法相比提高了1.6个百分点。在波兰文数据及中，切分效果有显著的提升并且达到了92.06%的准确切分率。在公开数据集IAM英文手写体数据集中正确切分率达到了98.43%，与最新方法相比提高了0.38%。数据表明此算法对于手写体文本可以准确的切分。结论 本文所提出的手写体文本行切分算法，根据现有投影方法的缺点改进并且融合了两种方法的优点是文本行切分更加准确。

关键词：手写文档；行切分；自适应阈值；投影；脱机。

墙报 2-3

论文编号：250

形态学重构滤波和多阈值区域生长的眼底血管三分图提取方法

李童 1，2，付冬梅 1，2，杨焘 1，2

1.北京科技大学自动化学院，北京，100083

2.北京市工业波谱成像工程技术研究中心，北京，100083

摘　要：针对计算机辅助诊断系统发展中的实际需要，提出一种自动提取眼底视网膜血管三分图的方法。眼底三分图是将图像像素划分为三部分：确定的前景即血管部分、确定的背景即非血管部分和未知区域即有待确定的部分。首先，采用阈值法对预处理后毛细血管增强的眼底图像进行二值化获得种子区域；然后，采取形态学重构滤波增强眼底图像的血管结构，利用多阈值二值化得到多张图像，接着依据与这些二值化图像的八邻域连通准则对种子区域迭代生长；最后通过调整生长迭代的次数生长出不同的血管图并合成为三分图。实验表明，该算法需要调节的参数少，在眼底血管三分图的完整性、抗噪性与算法效率上取得平衡，能够快速获得眼底图像三分图。

关键词：视网膜血管；三分图；形态学重构；多阈值；区域生长

论文编号：280

融合外观及深度特征的无监督视频目标分割算法

陈宁 1，2，尹辉 1，2，朱宏梁 1，3，许宏丽 1，3

1. 北京交通大学计算机与信息技术学院，北京 100044；

2. 北京交通大学交通数据分析与挖掘北京市重点实验室，北京 100044；

3. 北京交通大学轨道工程北京市重点实验室，北京 100044

摘　要：视频序列的复杂性，给无任何目标标注或先验的无监督视频目标分割带来了巨大的挑战。利用目标的外观显著性和运动显著性进行无监督视频目标分割时常常由于运动估计的不准确性造成分割错误，加之目标和背景之间存在的外观和运动混淆，严重影响分割精度。本文考虑场景中的三维信息，将深度特征引入无监督视频目标分割框架。一方面，前景目标和背景深度差异的引入可有效解决外观和运动混淆问题，提高分割精度；另一方面，深度图中的三维信息有利于克服二维运动估计偏差引起的目标错误分割。基于此，本文提出了一个融合外观和深度特征的双流无监督视频目标分割网络DNet-FAD（Dual Network Fusing Appearance and Depth）。该网络由外观分支、深度分支及融合模块构成，其双分支结构可同时有效提取外观特征和深度特征，并通过融合模块融合两类特征，最后通过级联不同网络阶段的特征预测分割。为了提高双流分割网络的性能，本文重点研究了深度特征与外观特征融合的层次、方式，并提出了三种双分支特征融合策略。本文在通用数据集SegTrack v2和DAVIS 2016上进行了实验，通过不同特征融合策略的对比实验和分析，确定DNet-FAD-2为最优的融合策略；在DAVIS 2016数据集上，本文方法取得了与主流方法相当的结果；在基网络预训练阶段未使用到的数据集Segtrack v2上，本文方法取得了最优的结果，分割精度超过第二名12.8%。实验结果证明，有效的外观和深度特征融合策略对提高无监督视频目标分割网络精度具有极大意义。

关键词：深度学习；视频目标分割；无监督；深度特征；双流网络

论文编号：284

基于置信度加权融合与视觉注意机制的前景检测方法

成科扬 1，孙爽 1，王文杉 2，师文喜 3，李鹏 3，詹永照 1

(1.江苏大学计算机科学与通信工程学院，江苏 镇江 212013；

2.社会安全风险感知与防控大数据应用国家工程实验室，北京 100846；

3.新疆联海创智信息科技有限公司，乌鲁木齐 830011)

摘　要：在视频前景检测中，像素级的背景减除法检测结果轮廓清晰，灵活性高。然而，基于样本一致性的像素级分类方法不能有效利用像素信息，遇到颜色伪装和出现静止前景等复杂情形时无法有效检测前景。为了有效解决这一问题，提出了一种基于置信度加权融合和视觉注意的前景检测方法。首先通过加权融合样本的颜色置信度和纹理置信度之和判断前景，其次进行自适应更新样本的置信度和权值；通过划分子序列结合颜色显著性和纹理差异度构建视觉注意机制判定静止前景目标，最后使用更新置信度最小的样本的策略保持背景模型的动态更新。实验结果表明，与其他五种主流算法相比，提高了在颜色伪装和存在静止前景等复杂情形下前景检测的精度和召回率，在公开数据集上得到更高的检测效果。

关键词：前景检测；置信度；颜色伪装；视觉注意；静态前景

论文编号：309

基于类脑模型与深度神经网络的目标检测与跟踪算法研究

赵宇飞 1，宋勇 1，*，杨昕 1，李国齐 2，蒋玉蓉 1，郭拯坤 1，王枫宁 1，张子烁 [1]

1.北京理工大学光电学院,北京市海淀区 100081； 2.清华大学精密仪器系,北京市海淀区 100084

摘　要：目标检测与跟踪技术已广泛应用于交通、医疗、安保和军事等多个领域。然而，目前目标检测与跟踪技术存在一些挑战，如目标微弱、背景复杂、目标被遮挡等。作为自然界中高效的生物智能系统，人脑视觉系统在计算机视觉领域具有显著优势。本文结合神经工程导向的类脑模型和计算机工程导向的DNNs，提出三种基于类脑模型和DNNs的目标检测与跟踪算法，具体包括：（1）基于演算侧抑制的目标检测算法，利用图像信息量将图像自适应地分成多个并行处理通道，利用演算侧抑制模型实现单通道图像的运动区域检测，对各通道图像进行融合处理并结合改进的模板匹配方法提取运动目标，实现动态背景下的运动目标检测；（2）基于双层视觉注意模型的目标检测算法，在特征提取层，通过两个通道分别提取图像的灰度和对比度特征，并进行融合获得初级显著图，在概率估计层，依据初级显著图计算目标和背景的先验概率和似然函数，利用贝叶斯公式获得最终显著图，实现复杂背景图像的显著性目标检测；（3）基于记忆机制与分层卷积特征的目标跟踪算法，该算法在相关滤波框架下利用分层卷积特征对目标进行定位，并融合人脑记忆机制利用响应图分析网络估计目标状态，通过调整分类器更新的学习率与间隔，实现自适应分类器更新，从而解决目标跟踪过程中目标被遮挡等问题。对比实验结果表明，将类脑模型和深度神经网络应用于目标检测和跟踪领域，有利于实现复杂条件下的高精度目标检测和鲁棒性目标跟踪。

关键词：人脑视觉系统；深度神经网络；类脑模型；目标检测与跟踪

论文编号：311

结合目标局部和全局特征的 CV 遥感图像分割模型

李晓慧 1， 汪西莉 1*

1. 陕西师范大学计算机科学学院，陕西 710119；

摘　要：随着遥感卫星技术的发展，不同尺寸、不同分辨率以及不同场景的的高分辨率卫星及遥感影像不断涌现。高效地从含有较多信息且背景复杂的遥感影像中自动提取目标，特别是小目标，是一个亟待解决的难题。传统图像分割方法主要依赖图像光谱、纹理等底层特征，容易受到图像中遮挡和阴影等的干扰。在采用图像分割提取目标的方法中，基于水平集的 CV 图像分割方法最大的优势在其处理对象拓扑变化的能力。但面对复杂的遥感图像，仅依靠图像颜色、纹理等底层信息，当图像中存在目标部分被遮挡、目标与背景杂糅等情况时，往往得不到正确的分割结果。为此，本文提出结合目标局部和全局特征的 CV(Chan Vest)图像分割模型，目标全局特征为其形状特征，采用深度学习生成模型--卷积受限玻尔兹曼机来表达目标形状特征；局部特征为目标边缘特征，采用 Canny 算子提取边缘信息；通过“点乘”的方式结合目标全局和局部特征，得到的特征一方面减少了边缘信息中虚假的信息，另一方面弥补了深度学习生成模型生成形状时丢失的目标边缘细节特征，因而提取的目标特征在全局和细节上都更为可靠，从而更好地约束演化曲线向目标轮廓不断逼近，得到正确的分割结果。在遥感数据集 Levir-oil drum、Levir-ship 和 Levir-airplane 上的实验结果表明:本文提出的分割模型不仅可以克服传统分割方法对噪声敏感的缺点，且在训练数据有限、目标尺寸较小、遮挡和背景复杂的情况下仍然能够快速且精准地分割出目标。

关键词：图像分割；形状先验；卷积受限玻尔兹曼机；深度学习；CV 模型

论文编号：326

一种基于 U-Net 架构的指静脉轻量级融合分割算法

曾军英，朱伯远， 秦传波， 朱京明，甘俊英， 翟懿奎，黄宇杰，王璠

五邑大学，广东省江门市， 529020

摘 要：基于深度学习的语义分割方法虽然已成功应用到图像分类、分割和检测等任务上，但是针对于现有的手指静脉分割网络大都过于庞大、参数过多而不适合在移动终端实现，且目前主流的轻量化网络在指静脉应用时会出现由于参数的减少而降低分割指标的问题。本文提出了一种基于 U-Net 的指静脉轻量级融合分割算法。首先通过在每张图像中随机选择其中心获得子块来进行数据扩充，然后从 U-Net 网络架构出发，通过 MobileNetV2 中的深度可分离卷积方法对普通的卷积进行替换完成了轻量级网络的初步实现，同时在深度可分离结构中融合了 InceptionV2 的部分结构在减少计算量的情况下获得更大的感受野，继而在深度可分离卷积之后引入 SENet 中的注意力模块来学习通道之间相关性，增加了少量的计算量的同时提升了网络性能。实验表明本文提出的指静脉分割网络在模型大小仅为同结构 U-Net 五分之二、Mult-Adds 为同结构 U-Net 百分之九的情况下，分割性能超越了 U-Net 网络，且相对更为复杂且庞大的网络 DU-Net 与 R2U-Net，在 SDU-FV 的 F1 值分别提高了 8.65%、15.42%，在 MMCBNU_6000 上 F1 值分别提高了 1%和 2.58%。

关键词：手指静脉分割；轻量级网络；深度可分离卷积； SE 模块；算法融合

论文编号：342

基于超像素Log-Euclidean高斯度量的传统文化图案分割

侯小刚 1， 赵海英 2

1. 北京邮电大学网络技术研究院，北京 100876； 2. 北京邮电大学计算机学院，北京 100876

摘　要：为提高传统文化图案分割中目标特征表征的有效性从而实现更准确的相似性度量，同时解决分割数目的自适应确定问题，本文提出了一种采用Log-Euclidean高斯进行超像素相似性度量的自适应图像分割方法。首先采用超像素算法对待分割的图案进行超像素预处理。然后基于超像素Log-Euclidean高斯度量(superpixel Log-Euclidean Gaussian metric，SLEGM)对形成的超像素进行有效特征表征从而实现其更准确的相似性度量。由于高斯样本协方差矩阵分布空间不是一个线性空间而是一个黎曼流形，通过矩阵对数运算将这个流形映射到一个线性空间，从而实现用简单的Euclidean运算处理高斯问题。最后通过在谱聚类算法中引入快速搜索和寻找密度峰值方法实现了传统文化图案自适应分割。在伯克利分割数据集(Berkeley Segmentation Data Set，BSDS500)和传统文化图案数据集(Traditional Cultural Pattern Database，TCPD)上的实验结果表明，本文算法分割性能优于目前经典的同类图像分割算法。

关键词：图像分割；超像素；Log-Euclidean高斯度量；流形；谱聚类

论文编号：343

基于加权局部特征选择的目标跟踪方法

郑鹏根，赵慧民，詹瑾，吕巨建，黄智慧，赵若妍，李伟健

广东技术师范大学计算机科学学院，广东省广州市 510630

摘　要：基于稀疏表示的全局特征跟踪模型在目标遮挡和背景干扰等情况下容易出现稀疏特征选择非最优的问题，为解决该问题，提出了基于加权局部特征选择的目标跟踪方法。该方法将目标划分为多个局部块，根据目标局部和目标整体的结构关系，综合多个局部模型的决策信息来选择最佳的目标位置。构建目标外观模型中，先在初始目标区域获取正负学习样本及其局部特征，然后基于LC-KSVD方法分别构建多个相互独立的局部判别字典作为目标跟踪模型。在目标跟踪中，先计算样本局部稀疏特征与初始目标局部稀疏特征的距离，并综合所有距离信息作为样本的相似性，选出相似性最佳的样本作为候选样本；然后基于样本局部稀疏系数的背景系数和琐碎系数的总量来衡量局部的扰动强度和可靠性，并将扰动强度作为局部权重构造加权重构误差选择最佳样本，以提高最后特征选择的稳定性。在模型更新中，采用周期与自适应结合的方式实现增量式字典更新，以提高目标模型对外观变化的适应性，设置扰动强度的动态阈值来衡量局部遮挡和跟踪漂移等情况，局部扰动大于阈值则取消周期更新操作，以减少误差引入。在OTB100数据集上的实验表明，本文方法在定性和定量分析中优于同类方法，在背景干扰、遮挡和目标外观变化等干扰因素影响下具有较好的跟踪鲁棒性。

关键词：目标跟踪；特征选择；局部加权；扰动强度；判别字典；自适应字典更新

论文编号：346

利用通道剪枝技术的实时实例分割方法

宁欣 1，刘江宽 2，李卫军 1，石园 2，支金林 2，蔡尉尉 3，南方哲 4

1.中国科学院半导体研究所,北京 100089； 2.威富集团形象认知计算联合实验室,北京 102200；
3.中南林业科技大学，长沙 410004；4.新疆大学软件学院 乌鲁木齐 830091

摘　要：随着实例分割技术在各种场景中的应用越来越广泛，运行速度和硬件资源占用是该技术在应用中需要考虑的两个重要因素。当前实例分割研究的主要方向更加注重分割精度的提升，对运行速度、资源占用的问题研究较少。最近提出的基于图像原型掩码系数的实例分割网络（YOLACT）在运行速度方面和分割精度上都获得了不错的性能，在COCO评测集上实例分割平均精度mAP可以达到30以上，同时在单GPU上可以做到实时的处理效果。但是，YOLACT需要设置较大的特征提取网络才能保证分割精度，这就导致了模型占用的硬件资源较多，同时运行速度也受到了限制。本文在YOLACT基础的上，提出了一种新的模型，对实例分割的特征提取网络进行了优化。首先，使用基于批量归一化层放缩因子的通道剪枝方法对YOLACT网络进行压缩，然后对压缩后模型的卷积层和批量归一化层进行融合，进行进一步的加速处理，最后，在COCO val2017上对本文提出的方法进行了评估。实验结果表明，在同样设置ResNet101作为Backbone时，与原始的YOLACT网络相比较，本文方法的模型文件大小可以减少56.9%，运行速度提升28.6%，运行时显存占用也降低了13.6%，有效地减少了硬件资源占用，并且提升了运行速度，更加有利于实例分割模型的实际应用和部署。

关键词：实例分割；模型压缩；通道剪枝；运行效率

论文编号：361

基于匹配解码的视频目标分割

钟航世，储琪，刘斌，俞能海

中国科学技术大学，中国科学院电磁空间信息重点实验室，安徽省合肥市

摘　要：随着互联网技术的不断发展，计算机视觉技术正在从方方面面影响着人们的生活。视频目标分割作为其重要的组成部分，对视频编码、自动驾驶、视频编辑和智能监控等领域的发展起到了关键作用。早期的视频目标分割算法通过在线学习提升视频的分割精度，然而这些算法通常具有较慢的分割速度。近期的算法更加注重分割精度和分割速度的平衡，进而出现了一批基于特征匹配的视频目标分割算法。这类算法利用解码网络将特征匹配的结果恢复到图像原尺度从而获得分割掩码，但是多数基于特征匹配的算法在解码过程中对所有的图像特征具备相同的计算权重，未能更加关注目标区域的分割。同时，多数算法利用双线性插值方式构建解码网络，导致了较多细节信息的丢失。本文针对这些问题，提出一种基于匹配解码的视频目标分割算法，利用所提出的两个模块完成对图像特征信息的充分利用：特征增强模块用于提升图像特征在目标区域的显著性，提高目标区域的分割权重；特征优化模块用于在解码过程中融合图像高层语义信息，恢复更多图像细节信息。本文在DAVIS-16验证集的消融实验验证了所提出模块的有效性，给出的分割结果图也表明了本文算法得到的分割掩码具备更多的细节信息；在DAVIS-16验证集上的对比实验表明本文提出的算法具备较高的分割精度和分割速度，具备实际应用价值。

关键词：特征匹配；特征增强；特征优化；视频目标分割

墙报 2-4

论文编号：365

一种基于 Mask RCNN 的融合几何特征的冠状动脉分割方法

邵凯 1,2，张云峰 1,2，包芳勋 3，郑勇 1，秦超 1,2，张彩明 4

1.山东财经大学,济南 250014；

2.山东省数字媒体技术重点实验室（山东财经大学），济南 250014；

3.山东大学数学学院,济南 250100；

4.山东大学计算机科学与技术学院，济南 250100

摘　要：冠状动脉分割是冠心病计算机辅助诊断系统中的一个重要步骤，其目的是保证在后续步骤中只对冠状动脉区域进行处理。冠状动脉CT血管造影（Coronary Computed Tomography Angiograph, CCTA）图像具有边界不清、结构复杂、特征不明显等内在特征，这些特点导致CCTA图像分割成为一项困难的任务。针对这一问题，融合图像的几何特征可以增强网络的特征学习能力，以此提高分割精度。因此，本文提出了一种将几何特征融合到Mask RCNN网络中的冠状动脉分割方法。首先，为了减少边界不清晰对分割精度的影响，提出了一种边界提取方法来提取图像的边界特征。其次，冠状动脉具有不同尺度的自相似性，并且分形维数可以表示区域的相似性。为了提高分割精度，本文将分形维数表示的分形特征引入到Mask RCNN中。最后，在网络中添加一个特征融合层，将得到的几何特征融合到原特征图（Feature Map）中，网络利用新的特征图进行分割。本文使用冠脉数据集来评估所提出的方法。在评估指标中，本文方法的平均精度（AP）和Dice系数达到83%和84.0±10.1%。试验结果表明，本文方法具有较高的精度和鲁棒性。

关键词：冠状动脉分割；Mask RCNN；几何特征；分形维数

论文编号：400

基于双流 Faster R-CNN 的像素级图像拼接篡改定位算法

吴鹏，陈北京，郑雨鑫，高野

南京信息工程大学计算机与软件学院，南京 210044

摘　要：近年来，由于深度学习模型表现出较强的图像特征学习能力，因此其被广泛应用于图像拼接篡改领域。然而，以往这些深度学习模型仅仅采用单流输入作为图像特征的提取方式，所以，相关学者考虑采用多流输入的方式提取篡改图像的特征。其中，基于双流Faster R-CNN的算法因综合考虑彩色图像和其噪声图像作为输入而获得了良好的性能。但是，它仍存在以下两个不足：(1) 定位精度只是块级；(2) 经过SRM滤波器产生的噪声图容易夹杂大量冗余得非篡改区域的信息，这不利于深度学习模型对特征的提取。因此，本文提出了一种基于双流Faster R-CNN的像素级拼接篡改定位模型。针对第一个缺陷，本文增加了一个全卷积神经网络分支用于实现像素级定位。针对第二个缺陷，采用错误等级分析(Error Level Analysis, ELA)噪声模型替代SRM滤波器用于提取噪声图，ELA算法能够更加突出拼接区域的轮廓，由其生成的图像能较为显著地区分图像中的拼接区域和非拼接区域，更有利于模型对边缘特征的提取。此外，为了增强主干网络对特征的提取能力，本文采用基于残差网络的特征金字塔网络多尺度提取图像特征。在三个公开篡改数据集上的实验结果表明，与现有的深度学习定位算法相比，提出算法获得最好的性能。

关键词：图像拼接；Faster R-CNN；错误等级分析；特征金字塔；全卷积神经网络

论文编号：404

融合 Hough 直线检测和 Grab-cut 算法的自适应风机叶片图像分割方法

唐震 1，彭业萍*1，，王伟江 1，，曹广忠 1，，吴超 1，

广东省电磁控制与智能机器人重点实验室，深圳大学机电与控制工程学院，广东省深圳市，518060

摘　要：完整的边缘信息对风力发电机叶片的边缘缺陷检测至关重要，但由于户外采集的风机叶片图像背景复杂多样，现有图像分割算法的分割精度不足，无法保证边缘缺陷的完整性。因此，本文提出一种基于图像边缘特征与颜色信息的自适应图像分割方法实现风机叶片边缘检测。首先，使用Hough直线检测初步定位叶片直线边缘；其次，根据直线数目自适应调整图像预处理参数与概率阈值，准确提取叶片目标区域；然后，在目标区域应用基于Otus阈值分割和形态学运算的Grab-cut算法，实现叶片图像的自适应分割；最后，采用无人机采集多个场景的图像作为测试样本，对本文分割方法与其他方法进行定性和定量对比分析。结果表明，本文方法能自适应且准确地实现复杂背景下风机叶片图像分割，并保留边缘缺陷的完整性，其边缘覆盖率和边界位移误差指标均优于其他方法，对风机叶片的边缘缺陷检测具有重大潜在应用价值。

关键词 ：风机叶片；图像分割；Grab-cut；Hough直线检测

论文编号：407

基于深度学习的藏文现代印刷物版面检测技术研究

吴燕如 1，2,珠杰 1,2,管美静 1,2

（1. 西藏大学信息科学技术学院,西藏拉萨 850000;2.藏文信息技术国家地方联合中心,西藏拉萨 850000）

摘　要：针对藏文现代图书版面中的文本行分布不均匀，现代藏文字体差异较大的问题，本文提出了一种基于Faster R-CNN的版面文本行检测算法，一是，在整理标注的数据集上训练，用ResNet-50网络提取出藏文现代图书版面特征信息；二是，为了有效提高模型的泛化能力，在coco数据集下的网络模型中进行迁移学习。实验结果表明，该方法可以对藏文现代印刷物的版面实现文本行的定位，检测准确率为83%，召回率为95%，明显提高了版面检测的精确度。

关键词：深度学习；藏文现代印刷物；Faster R-CNN；版面检测

论文编号：466

遮挡对于目标检测的影响分析

张胜虎 1，马惠敏 2*

1.清华大学电子工程系,北京 100084； 2.北京科技大学计算机与通信工程学院,北京 100083

摘　要：当前目标检测任务中遮挡问题是一项具有挑战性的工作，由于存在遮挡导致物体的整体特征结构遭到破坏，在检测过程中容易发生漏检误检等问题。常见遮挡处理方法在很大程度上提高了遮挡检测效果，然而对遮挡构成因素和不同遮挡比例对于检测性能的影响情况，目前并没有具体量化分析。对此本文从数据驱动方法出发，通过仿真方式构建生成大量均匀分布的遮挡数据集More than Common Object Detection(MOCOD)，在此数据集上分析不同遮挡比例下的检测性能，量化分析了不同遮挡对于检测性能的影响情况，在分析的基础上，通过按遮挡比例引入衰减权重方式来筛选高质量的正样本参与模型训练，有效提升了遮挡情况下的检测性能。

关键词：深度神经网络；目标检测；遮挡处理；遮挡数据集

论文编号：473

融合图像深度的抗遮挡目标跟踪算法

王希鹏，李永，李智，杨凯凯，张妍，梁起明

武警工程大学 信息工程学院,西安 710086

摘　要：由于视频信息的局限性，在遮挡情况下的目标跟踪依然是一个很难解决的问题。人在对视频中目标进行视觉跟踪时，会估计出视频中场景的层次关系，判断目标和干扰物的位置关系，减小跟踪过程中遮挡对目标跟踪的影响。根据上述思路，针对目标跟踪过程中的遮挡问题，提出将图像深度引入单目标跟踪，构建遮挡判别模块，利用目标深度信息的变化判断遮挡情况并修正跟踪结果。首先应用单目图像深度估计算法对图像进行深度估计，获取图像的深度信息；其次，将基于孪生网络的目标跟踪算法与图像深度相结合，构建遮挡判别模块，利用目标深度信息的变化判断遮挡情况；最后，对目标跟踪器的候选框重新排序避免被遮挡物干扰。采用OTB-2015数据集中存在目标遮挡的11个视频进行测试，其中在10个视频序列中，本文算法的精确度高于基准跟踪器SiamRPN，在平均成功率和平均精确度上分别提高了1.7%和0.9%。实验结果表明，在目标遮挡场景下，与6种主流跟踪算法相比，本文具备更优的跟踪性能。同时也说明，图像深度信息可以辅助提升目标跟踪的性能。下一步研究工作可针对遮挡判别策略进行改进，进一步提高算法的跟踪性能。

关键词：孪生网络；　目标跟踪；　单目图像深度估计；　抗遮挡

论文编号：462

遮挡图像数据生成系统

梅若恒，马惠敏

北京科技大学, 北京市 100083

摘　要：在计算机视觉领域，遮挡问题一直以来是一个比较困难的问题。现实中存在大量的遮挡现象，使得图像中目标特征缺失，从而加大了目标检测的难度，同时部分图像数据在现实中难以获取，使得可训练样本量难以满足训练需求。当前在遮挡问题上，仍然缺乏相应的数据集和评估标准对目标检测算法进行系统的评价。本文提出了一个遮挡图像数据生成系统来生成遮挡图像以及对应标注信息，并利用该系统构建遮挡图像数据库MOCOD。系统生成图像的遮挡类型主要包括两种：1）不透明目标实体间的互相遮挡，2）半透明烟雾对目标的遮挡。本系统提供了丰富的场景和自由的样本拍摄方式，为图像内所有目标和遮挡物提供了实例分割级的语义标注以及遮挡强度等级划分。相比于KITTI等数据集对于遮挡的简单划分，本系统生成的数据集能够更好的评价目标检测算法在遮挡问题下的瓶颈和性能。

关键词：遮挡；仿真；数据集；计算机视觉

论文编号：112

多指标融合的人物图像美感评价算法研究

朱瑞 1，王美丽 1，2，3

1.西北农林科技大学信息工程学院,陕西杨凌 712100；

2.农业农村部农业物联网重点实验室,陕西杨凌 712100；

3.陕西省农业信息与智能服务重点实验室，陕西杨凌 712100

摘　要：图像可计算美感评价属于艺术、心理学及计算机科学等多学科交叉的创新性研究课题，它模拟人类感知及审美思维，以可计算的方法对图像美感做出恰当的判断，具有很大的复杂性和主观性，是图像处理领域极具挑战性的一个难题。目前人物图像还未有统一且完善的美学质量评价标准，而且由于人物图像的关注度越来越高，针对这一具体类别的研究成果也具有越来越大的应用价值。本文针对人物图像，提出一种根据人类视觉及审美习惯来评价其美学价值的可计算方法，从技术特征、感知特征和社会特征三个方面对图像进行特征提取并作为评价指标，融合各指标建立美感评价模型，采用在线调研的方式确定各指标的影响因子，从而实现人物图像美感自动评价。在香港中文大学人物图像美学质量评价数据集（CUHKPQ）上的实验测试结果表明，本文提出的方法准确率可达到98.08%，具有一定的有效性，并且各指标的评价结果清晰直观，可以从不同角度促进图像美学质量的提高。本文将模型集成到手机应用程序中，能够对用户上传的多张人像照片进行美感评价和打分，并依据评分结果选择保留或删除的照片，从而简化用户操作，为其生活提供便利。

关键词：人物图像；美感评价；美学特征；指标融合

墙报 2-5

论文编号：133

一种基于色彩补偿的水下图像综合增强算法

杨淼 1，2，胡珂 1，王海文 1，胡金通 1，殷歌 1

1. 江苏海洋大学 电子工程学院，江苏 连云港 222005，中国；2. 江苏省科技大学海洋装备研究院，江苏 镇江 212000，中国

摘　要：由于水下图像的失真是由多种退化混合叠加而成，因此单一的水下图像复原、增强的算法无法综合的解决各个方面的水下图像质量退化。针对退化的水下图像在高级视觉分析任务中无法进行有效的目标检测及识别的问题，本文提出了一种基于色彩补偿的组合水下图像增强算法，通过色彩补偿和对比度拉伸，HSV空间伽马校正和亮度通道去模糊系列方法实现了对水下图像的色彩校正，色彩对比度、饱和度和细节清晰度的综合提高。其中，本文提出了基于高斯滤波的亮度通道去散射方法，并对典型水体水下图像综合增强参数进行了分析。同时，对比了本文提出的方法和其他增强方法对偏蓝、偏绿、偏黄、白色近岸浅滩水下图像的处理结果。本文还通过目标检测网络模型对经过7种算法增强后的水下图像数据集进行训练与测试，对比了平均水下目标识别准确率和检测到的目标数量与实际目标数量的比值来评估各个增强算法对于水下目标识别和检测任务中的作用。实验说明，本文提出的方法与现有方法相比，不仅可以有效的实现各类水下图像清晰度提高和色彩增强，适用范围广，而且获得了最高的93.3%的目标检测率和85.2%的目标识别准确率，可以有效的提高水下图像目标识别任务的准确率和检测数量。

关键词：水下图像增强；亮度通道去散射；目标检测；水下图像质量评价

论文编号：447

基于双通道卷积神经网络的视频目标移除取证算法

白珊山 1，2，倪蓉蓉 1，2，赵耀 1，2

北京交通大学信息科学研究所，北京 100044；

现代信息科学与网络技术北京市重点实验室，北京 100044

摘　要：数字视频通常被认为是对事件进行公正可靠记录的载体，被广泛应用于交通、医学、司法等领域以提供相关事实的基本证据。但是，经过恶意篡改的伪造视频如果被传播，将会对人们的日常生活和正常的政治、经济和社会秩序造成危害。针对现有数字视频目标移除取证算法的伪造帧识别准确率低的问题，本文提出了一种基于双通道卷积神经网络的视频目标移除取证算法。该算法利用双通道结构，分别提取视频绝对帧差图像的RGB特征和噪声特征，并利用双线性池化对二者进行特征融合，而后通过分类层输出视频帧的分类结果，从而有效地识别经过篡改的视频帧。其中，RGB通道能够发现绝对帧差图像中不自然的篡改边界和对比度，噪声通道能够发现原始区域和篡改区域之间噪声的不一致性。此外，算法在网络前端增加了预处理层来放大篡改视频帧的伪造痕迹。通过在目前最大的基于对象的伪造视频数据集SYSU-OBJFORG上进行实验，验证了本文所提算法的有效性。与现有两种主流算法的对比实验结果显示，所提算法在各项评价指标中，尤其是伪造视频帧的识别准确率（达到96.04%）上有明显提升，且相对于传统的单通道网络结构，双通道特征融合的方式取得了更好的检测性能。

关键词：数字视频取证；视频目标移除取证；双通道卷积神经网络；Inception-v3 网络

论文编号：238

无人系统多模态视觉感知单元测试与评估方法的研究

王伟 1，2，郑国强 3，刘海春 3,4，李扬 3，裴凌 3,4

1.陆军装甲兵学院,北京 100072；2.北京特种车辆研究所,北京 100072；

3.上海北斗导航创新研究院,上海 201702；4.上海交通大学,上海市北斗导航与位置服务重点实验室,上海 200240

摘　要：本文提出了一种通用的无人系统多模态视觉感知平台的测试与评估的方法。研究的内容主要分为可见光波段和红外光波段下的动态测距实验方法与复杂光变测距实验方法。针对动态测距实验分为室内环境及室外环境，室内环境下将无人系统多模态视觉感知平台安装到高速导轨进行测试，室外测试将无人系统多模态视觉感知平台安装到测试车辆上进行。针对复杂光变测距实验，测试的场所为室内环境，主要分为光照瞬变实验、局部高亮实验及局部阴影实验。整个过程中采用了基于北斗卫星实时授时的时间同步技术。针对时间同步技术，主要采用了基于NTP/PTP协议的网络对时方法，其中由于部分设备不支持网络通讯，因此采用了基于硬件触发的高精度时间同步策略，通过构建时钟驯服系统进行校时，有效的保证了整个系统的时间同步精度，使不同设备之间的时间同步误差小于10毫秒。针对真值采集设备，对比已有的高精度采集设备，根据测试条件及限制，最终采用了高速导轨、激光全站仪、激光跟踪仪等设备，有效的保证了测试真值数据的准确性。同时在不同设备间进行不同坐标系间的关联与统一，保证输出的数据均在统一坐标系下。最后，对比测量数据和真值数据，给出测试设备的性能与功能的量化评估。根据实际的测试结果看，本文提出的测试评估方法有效地获取了真值，并进行了时间同步，满足了实际的需求，具备推广性，为后续的测试提供了借鉴参考的价值。

关键词：测试；评估；多波段；立体视觉；时间同步

论文编号：240

固定翼无人机杆塔类障碍物探测技术

梁欣凯，张航，赵佳佳，宋闯

复杂系统控制与智能协同技术重点实验室，北京 100074

摘　要：非预期的杆塔类障碍物（如信号塔、高压线塔等）严重影响了固定翼无人机的全自主飞行安全，因此需要赋予飞行器自主探测障碍物的能力。杆塔类障碍物细长的几何形状特征导致其成像呈现弱纹理、低显著性，它极易湮没于复杂的背景中难以被有效识别，更加难进在其上提取特征点进行精确空间位置解算。除此之外，基于飞行器的障碍物自主探测建立于有限机载计算资源条件下对障碍物目标的准确二维图像识别与精确三维空间定位，对算法鲁棒性、算法运行时间等也提出了严重的考验。针对以上难题，本文提出了一类固定翼无人机杆塔类障碍物探测技术，通过基于改进的YOLOv3-tiny障碍物目标识别技术、几何敏感性障碍物目标特征点提取与匹配技术和多帧障碍物目标关联与距离解算技术构建了固定翼无人机障碍物探测系统，利用卷积神经网络提取高维抽象特征的能力同时提升障碍物检测性能与障碍物特征点提取性能，完成了强实时条件下的障碍物检测与相对距离解算，最终通过固定翼无人机实际飞行试验对于所提出算法进行验证，实现了全自主飞行状态下对于杆塔类障碍物目标实时探测与定位，识别准确率达到97.3%，平均定位误差不大于15%。

关键词：固定翼无人机；障碍物探测；特征点；深度学习

论文编号：323

基于食物检测和类别注意力监督的食物识别

刘欢 1，2，张淳杰 1，2，赵耀 1，2

1.北京交通大学信息科学研究所,北京 100044；

2.“现代信息科学与网络技术”北京市重点实验室,北京 100044

摘　要：食物识别技术的发展能够促进农业、工业、医疗健康和新零售等领域的自动化和智能化。近年来，基于图像内容分析的食物识别受到了研究者们越来越多的关注。但是大多数研究往往针对现有的深度神经网络的使用与联合展开研究，而忽视了食物图像本身固有的特点。针对食物图像存在的空间分布不确定性与不同地域之间的复杂背景干扰，本文提出了一种基于食物检测与类别注意力监督的食物识别方法。模型主要由食物检测网络和类别注意力监督网络两部分组成。该方法实现了由粗到细的注意力监督机制，在去除干扰背景的同时关注细粒度的食物特征。本文创新点如下：1)针对食物图像背景的复杂地域性差异，提出了由特征金字塔与区域生成网络组成的食物检测网络，去除复杂背景噪声干扰；2)针对食物图像的空间不确定性，提出了由空间变换网络实现的类别注意力监督网络，关注细粒度特征并缓解空间不确定性；3)构建了基于食物检测与类别注意力监督的食物识别模型。在食物数据集ETH Food-101和VireoFood-172上进行了实验，结果表明本文提出的基于食物检测和类别注意力监督的食物识别方法能够有效利用食物类别标注信息，与现有方法相比识别准确率有明显提高。并通过进一步对食物检测网络和类别注意力监督网络的对比实验，证实了模型各部分能有效提高食物识别准确率。

关键词：食物识别；鲁棒性；食物检测；类别注意力

论文编号：368

词嵌入技术和端对端的遥感图像描述生成模型

王媛 1，2，马洪兵 2，阿里甫·库尔班 3，艾里西尔·库尔班 4，玉米提·哈力克 5，吕亚龙 1

1.新疆大学信息科学与工程学院，乌鲁木齐 830046；2.清华大学电子工程系，北京 100084；3.新疆大学软件学院，乌鲁木齐，830091 4.中国科学院新疆生态与地理研究所，乌鲁木齐，830011 5 新疆大学资源与环境学院，乌鲁木齐，830046

摘　要：近年来，受人造卫星和遥感技术发展的影响，遥感图像在很多应用中引起了广泛的关注。然而，目前大多数遥感图像的研究仍然集中于目标识别、图像分割和场景分类等[1][2][3]领域。虽然这些研究识别了遥感图像中待识别的物体，并获取了遥感图像中待识别物体类型的类标签，但忽略了待识别对象的属性和对象之间的关系。为了描述图像的内容，研究学者设计并实现了许多方法用于图像的描述和表示。但是研究学者生成的句子都是简单的自然语言句子，并不能有效的描述复杂的遥感图像内容，且通常应用在小规模数据集上。因此，本文结合词嵌入技术，构建了一种新的端对端模型将图像解码成自然语言句子，并完成遥感图像的描述任务。为了用精确简洁的自然句子来描述复杂遥感影像中的胡杨和柽柳，本文结合词嵌入技术，提出了一种端对端的图像描述生成模型，以实现胡杨和柽柳的分类识别。首先，由于遥感图像中大尺度区域内类别模糊性的特点，引入共现矩阵(Co-occurrence matrix)和Glove生成了待识别物体的词向量特征；其次，为进一步描述遥感图像的内容，构建了一个多层级的端对端生成模型，来更好地推进遥感图像中胡杨和柽柳的描述任务。实验结果表明，相比于传统深度学习方法，该方法生成的自然语言句子，可以更好的描述遥感图像中的胡杨和柽柳。

关键词：词嵌入技术；类别模糊性；共现矩阵；词向量；深度学习；端对端生成模型

论文编号：397

基于光照方向一致性的换脸视频检测

李纪成 1，刘琲贝 1,2，胡永健 1,2，王宇飞 2，廖广军 3

1. 华南理工大学 电子与信息学院，广东 广州 510640； 2. 中新国际联合研究院，广东 广州 510000

3. 广东警官学院 刑事技术系，广东 广州 510230

摘　要：光照条件是视频外部成像环境的重要因素，可反映视频成像时的客观物理条件，即使采用复杂的计算机图形学模型，也难以完全反映光照的真实场景。针对大部分Deepfake换脸视频与真实视频在外部成像环境上的差异，提出一种基于光照方向一致性的换脸视频检测方法。本文利用Lambert光照模型逐帧计算待测视频的二维光照方向，通过判断整段待测视频二维光照方向的角度变化是否平滑确定视频真伪。实验结果表明该方法在公开的测试数据库TIMIT和FaceForensics++上具有较好的检测性能，可以有效区分真实视频和换脸视频。由于不需要训练模型，该方法具有计算复杂度低、实时性好的特点。

关键词：视频篡改检测；换脸视频；成像环境；光照方向

论文编号：445

融合稀疏点云补全的 3D 目标检测算法

徐晨 1，2，倪蓉蓉 1，2，赵耀 1,2

1.北京交通大学信息科学研究所,北京 100044；

2.现代信息科学与网络技术北京市重点实验室,北京 100044

摘　要：基于雷达点云的3D目标检测方法有效地解决了基于RGB图像的2D目标检测易受光照、天气等因素的影响的问题。但是，由于雷达的分辨率以及扫描距离等问题，激光雷达采集到的点云往往是稀疏的，这将会影响3D目标检测精度。针对这个问题，提出一种融合稀疏点云补全的目标检测算法，采用编码、解码机制构建点云补全网络， 由输入的部分稀疏点云生成完整的密集点云，根据级联解码方式的特性，定义了一个新的复合损失函数。复合损失函数中除了原有的折叠解码阶段的损失之外，还增加了全连接解码阶段存在的损失，以保证解码网络的总体误差最小，从而使得点云补全网络生成信息更完整的密集点云，并将补全的点云应用到3D目标检测任务中。实验结果表明，所提算法能够很好地将KITTI数据集中稀疏的汽车点云补全，并且有效地提升目标检测的精度，特别是针对中等和困难等级的数据效果更佳，提升幅度分别达到6.81%和9.29%。

关键词：目标检测；雷达点云；点云补全；复合损失函数

墙报 2-6

论文编号：461

基于视觉的激光雷达数据增强

宋绪杰 1，戴孙浩 2，林春雨 2，詹书涛 2，赵耀 2

1.北京交通大学交通运输学院,北京市 100044；2.北京交通大学计算机与信息技术学院,北京市 100044

摘　要：激光雷达在自动驾驶中扮演着重要的角色，在其价格昂贵的同时，产生的激光线束数量仍然较少，造成采集的点云密度较为稀疏。为了更好地感知周围环境，本文提出了一种基于视觉的激光雷达数据增强算法，利用视觉实现了点云的稠密化处理。其不针对特定的3D目标检测网络结构，而是一种通用的点云稠密化方法。该算法首先利用双目RGB图像生成深度图像，并根据先验的相机参数计算出每个像素点在激光雷达坐标系下的大致三维坐标。为了更好地分割地面，解决复杂场景下普通RANSAC算法失效的问题，本文设计了循环RANSAC算法，改进了地面分割效果。然后将原始点云进行地面分割后插入k-dimensional tree (KDTree)，在KDTree中搜索大致三维坐标点的若干近邻点，基于这些近邻点进行曲面重建，并设计一种计算几何方法导出各像素对应的激光雷达坐标系下的精确坐标点。最后将这些生成的精确坐标点与原始点云融合得到稠密化点云。实验结果表明，稠密化的点云在视觉上具有更加完整的形状和轮廓，物体的特征更加明显；并且稠密化的点云在KITTI数据集上提升了3D目标检测精度，在使用该数据增强方法后，AVOD的AP3D-Easy提升了8.25%，AVOD-FPN的APBEV-Hard提升了7.14%。

关键词：激光雷达；数据增强；地面分割；点云稠密化；3D 目标检测

论文编号：75

基于近似后验信息的高光谱海上舰船目标检测

李宁阳 1，王兆晖 1

1.海南大学计算机与网络空间安全学院,海南省海口市 570228

摘　要：舰船检测在海上交通管制、渔业监管、海战形势预测等领域应用广泛。高光谱图像光谱分辨率高、图谱合一，基于高光谱图像的舰船检测主要利用舰船目标与周围海水背景的光谱差异较大的特性，不仅能检测出舰船目标的位置、形状、尺寸等空间特征，又能对舰船的表面材质进行初步分析，有利于后续舰船类型识别工作的开展。但在近海区域，尤其在海岸线附近，小岛屿、码头等目标与舰船目标的尺寸、红外波段光谱特性相似，造成舰船检测率低、虚警率高。针对这一问题，本文提出一种基于近似后验信息的舰船检测方法。首先对近、短波红外波段灰度影像进行阈值分割和形态学处理，完成海陆预分割；然后通过顶点成分分析从陆地区域获取陆地近似后验光谱，并利用其对海域进行约束能量最小化检测，去除海域中的疑似陆地像元，对海陆再次进行分割，进一步缩小检测海域；最后对再分割后的海域进行海水背景抑制，再次利用顶点成分分析从新海域中获取舰船近似后验光谱，再次对海域进行约束能量最小化检测，提取出舰船目标。实验结果表明：相比于只进行一次海陆分割，未去除海岸线附近疑似陆地像元的算法，本文提出的算法能够在场景复杂且没有任何先验信息的情况下，有效去除海岸线附近的疑似陆地像元，完成海陆二次分割，检测出舰船目标，有效降低虚警率。

关键词：高光谱；舰船检测；近似后验光谱；顶点成分分析；约束能量最小化

论文编号：**122**

复杂背景下苹果目标的识别定位方法

姜玉良，王鹏*，姚刚

哈尔滨理工大学 电气与电子工程学院，哈尔滨，150080

摘　要：果实采摘环节由于其本身要求较高的时效性和劳力投入性，成为了苹果的生产作业中十分重要的一部分。针对在复杂背景下苹果的自动采摘过程中出现的光照影响、枝叶遮挡和果实间相互重叠的问题，提出一种基于K-means和最大外接矩形内切圆算法的苹果目标的识别定位方法。以自然环境下生长的红苹果为对象，通过Retinex算法消除光照的影响；选用Canopy算法与K-means算法相结合对图像中的苹果目标进行分割，使用形态学方法对苹果轮廓进行提取；利用基于距离变换的分水岭图像分割算法对可能重叠的苹果轮廓进行切分；根据最大外接矩形内切圆法对苹果目标进行重建，然后计算各个重建结果的圆心距离和半径大小，剔除计算结果小于阈值的苹果目标，最终确定各个苹果目标在图像中的位置和大小。Retinex算法可以较好的削弱在图像处理过程中由于光照造成的高亮和阴影影响，而K-means算法和最大外接矩形内切圆法结合提取的苹果目标轮廓与真实苹果目标轮廓重合度较高。对100张自然环境下苹果目标图像的识别定位在OpenCV平台下的验证结果表明：该方法对苹果目标进行识别定位的平均分割误差为9.67%，苹果目标识别准确率为90.85%，平均识别定位时间为13.57s。应用本文方法可以对自然环境下的苹果目标进行准确识别定位，提高识别效率。

关键词：苹果；识别定位；复杂背景；K-means；OpenCV

论文编号：**187**

一种基于图像识别的增量式自聚焦透镜数值孔径测量方法

邢春蕾 1，2，3，许峰 1，2，3*，胡正文 1，2，3，邹快盛 1，2，3

1. 苏州大学光电学院,江苏 苏州 215006；2. 江苏省先进光学制造技术重点实验室,江苏 苏州 215006；3. 江苏省现代光学技术重点实验室,江苏 苏州 215006

摘　要：针对自聚焦透镜关键光学参数数值孔径的精密测量问题，本文提出了一种借助计算机图像识别增量式的测量方法。测量原理基于几何光学对自聚焦透镜数值孔径的定义，利用成像法对入射光线在径向圆心处具有的最大光线接收角的正弦值进行测量。增量式测量方法指自聚焦透镜对发光网格物在不同物距下两次成像，经过对两次成像的物高与两次物面相对位移量进行计算得到最大接收角正弦值。由于自聚焦透镜成像时会导致像面畸变，本文利用计算机图像识别算法来提高测量精度，具体方法是通过自聚焦透镜像面上网格与视场边缘的特征来判断增量式测量中的两次物高值。一个完整的物高值包括中心物高和边缘物高，判断物高首先需要提取纵方向截线，利用截线上所有条纹数判读中心物高并利用外插算法确定视场边缘到最外侧条纹对应的边缘物高。本文通过误差模型的建立对测量系统进行了误差分析，确定了整个测量系统误差的主要来源；并通过搭建实验证明此测量方法的稳定性可以满足实际需求。本文提出的图像识别与增量式测量结合的自聚焦透镜数值孔径测量方法不仅可以减少绝对式测量带来的测量误差，而且可以减少直接用人眼判读物高时带来的测量误差，是一种简单、快速、准确的测量方法。

关键词：梯度折射率光纤；光学参数；测量系统；测量算法；图像识别；误差分析

论文编号：192

一种利用弧支撑线段进行瞳孔识别的方法

王朕 1，2，肖盼 1

1. 天津财经大学理工学院，天津，300222

2. 国家海洋局数字科学技术重点实验室，天津，300171

摘　要：视线追踪技术通过视线估计对人在特定场景中的注意力进行客观有效的分析，并在心理学、医学、界面可用性分析、航空军事、市场研究、汽车工业等多个领域有着广泛的应用，高效、准确地检测瞳孔中心是视线估计的关键任务。近年来涌现出许多瞳孔检测算法，但由于光照、遮挡、离轴和误记等原因，在真实场景中对瞳孔中心的检测依然存在挑战。并且在基于瞳孔-角膜反射法的头戴式跟踪系统中，瞳孔特征提取算法通常只适用于瞳孔检测，普尔钦光斑的识别则需重新设计算法。本文提出了一种利用Haar特征检测器粗定位瞳孔位置的检测方法，并利用圆弧支撑线段提取可能属于瞳孔轮廓的椭圆圆弧。然后根据弧段之间的距离、极性和方向，将这些弧段组合成不同的组，并进行椭圆拟合，产生初始椭圆集，最后依据椭圆置信度选择最佳椭圆作为瞳孔边界，依据瞳孔和光斑的位置关系及在眼图中的亮度差异，该算法在检测瞳孔的同时也可检测光斑中心。通过在公共眼睛图像数据集（LPW）上进行比较实验，新算法与现有5种算法中的最高平均检测准确率相比提升了11%。

关键词：瞳孔检测；椭圆弧段；椭圆拟合；眼动追踪

论文编号：216

基于注意力机制和深度监督的全分辨率残差网络乳腺超声图像分割

侍尧 1，屈晓磊 1，2，单光存 1，储冰峰 3，张静雅 1

1.北京航空航天大学,仪器科学与光电工程学院,北京 100191； 2.北京航空航天大学大数据精准医疗高精尖创新中心,北京 100191； 3.中国人民解放军总医院,口腔综合科 100039

摘　要：乳腺癌在女性癌症中具有较高的发病率。医学超声成像是一种常用的乳腺成像方法。在超声图像上进行乳腺肿瘤的自动分割可以帮助医生客观评价肿瘤的尺寸。然而，超声图像的伪影和噪声给分割任务带来一定困难。针对这一问题，我们在本文中提出了一种基于注意力机制和深度监督的全分辨残差神经网络(ASFRRN)来分割超声图像上的乳腺肿瘤。在本文模型中，我们将注意力机制和深度监督引入全分辨率残差网络(FRRN)，并使用两个数据集（数据集B和D）对模型进行评估。其中数据集B包含163张超声图像，而数据集D内有980张超声图像。其中所有乳腺超声图像都由专业医生进行手动标注，给出了肿瘤所在区域。同时本文中使用Dice系数，杰卡德相似系数和F1分数作为分割精度的评估指标。本文模型最终在数据集B上得到了比原有方法高的Dice系数(84.0±10.1%),杰卡德相似系数(74.0±11.3%)和F1分数(84.0±10.1%)，而在数据集D上，本模型也得到了较高的Dice系数(89.1±13.3%)，杰卡德相似系数(74.0±11.3%)和F1分数(89.1±13.3%)。

关键词：深度学习；肿瘤分割；乳腺癌；超声图像

论文编号：232

一种水下鱼类动态视觉序列运动目标检测方法

张明华 1，龙腾 1，宋巍 1，黄冬梅 2，梅海斌 1，覃学标 1

1.上海海洋大学,上海市 201306； 2.上海电力大学,上海市 200090

摘　要：我国是世界水产养殖大国，水产养殖也是我国国民经济收入的一项重要来源，随着计算机视觉技术的快速发展，其在运动视频目标检测领域应用愈发广泛，对于水下环境的目标检测术已成为海洋技术领域的研究热点。为了克服水下视频质量不高、视频模糊不清甚至很难辨认的缺点，从水下环境的视频中检测出运动的鱼类目标，利用计算机视觉技术对水下鱼类目标进行快速目标检测，提出了一种基于背景去除的水下视频目标检测方法，设计适合水下环境的鱼类目标检测框架，使用偏最小二乘（PLS）分类器进行目标检测。利用水下拍摄的鱼类数据集收集输入的视频序列，并提取单独的帧。将帧进行RGB转HSI格式转换以及中值滤波器去噪的预处理，应用于GMG背景去除过程，提取了基于局部二值模式（LBP）纹理和灰度系数的重要特征，最后将所提取到的特征，利用PLS分类器，实现了分别对白天环境以及夜晚环境中的水下鱼类目标检测。结果表明，该方法利用水下拍摄的鱼类视频数据集目标检测精度可达到96.89%，查准率可达到93.32%，提高了对于水下鱼类目标的检测效率，降低了人工成本。为水下鱼类等生物资源的监测、保护和可持续开发等工程应用方面提供了一定的参考价值。

关键词：偏最小二乘；背景去除；鱼类；目标检测；动态视觉序列

论文编号：272

基于深度卷积网络的海洋涡旋检测模型

张盟 1，杨玉婷 1，孙鑫 1，董军宇 1，梁瑶 1

摘　要：中尺度海洋涡旋的自动识别对监测其动态变化有着重要的意义，高效的涡旋检测对于提高我们对海洋动力学的认识至关重要。传统的利用遥感数据检测涡旋的方法通常是基于物理参数、几何特征、手工特征或专家知识。深度学习技术近年来得到了广泛地应用，本文重点研究基于深度学习技术从海表面高度图中识别海洋涡旋的方法。针对海洋卫星拍摄的海洋表面高度图中的涡旋检测问题，本文提出了一种基于卷积神经网络的多涡旋检测模型，该模型能够准确提取涡旋的特征信息，拟合语义信息与海面高度之间的关系。同时，本文公开了一个新的用于涡旋检测的训练数据集——SCSESA-Eddy，以评估基于人工智能的涡旋检测方法的性能。该数据集涵盖了15年来位于中国南海及其东部部分海域的每日卫星遥感海表面高度数据。涡旋检测实验结果表明，与现有的方法相比本文的模型取得了更好的检测结果，能够更好地区分相邻的多个涡旋。同时本文也统计了不同检测方法的涡旋检测数量以及检测半径，以评估它们的性能。

关键词：卫星遥感；海表面高度图像；涡旋检测；深度学习

论文编号：286

基于小样本学习的 X 光图像违禁物品识别

邢琛聪 1，吕淑静 2，吕岳 2

1.华东师范大学计算机科学与技术学院,上海 200062；

2.华东师范大学，上海市多维度信息处理重点实验室,上海 200241

摘　要：X光图像违禁物品筛查是维护公共安全的重要手段，现已广泛应用于多个行业。目前，X光图像违禁物品筛查主要依赖人工检查。随着筛查要求的提高，自动化X光分析技术的应用是大势所趋。X光图像违禁物品识别的研究可分为基于图像分类、目标检测、图像分割等方法。然而，违禁品的定义常根据不同应用场景和安检级别变化，现有研究仅针对训练集包含的类别，模型不能直接处理新类别。同时，数据集不平衡的问题也会影响模型的训练。本文从小样本学习角度分析问题，并针对此任务提出一种新的图像分割模型。该模型将测试图像（查询图像）与多张标注图像（支持图像）映射至相同的特征空间，依据图像特征之间的相似度与标注信息，分割查询图像内含有违禁物品的区域。不同于自然场景图像，X光图像具有透视重叠等问题。稠密特征编码模块通过融合底层特征与中层特征的方法提取X光图像的特征。图像间相似度的度量由双路比较模块实现。该模块同时捕捉细粒度的像素级相似度与粗粒度的区域级相似度。针对数量不定的支持集图片，模型采取一种基于卷积化门控循环单元（ConvGRU）的操作，将不同测试图像与支持图像的双路比较特征融合。最终，多分辨率解码模块逐级放大、融合双路比较模块的输出特征与查询图像的底层特征并生成分割结果。SIXray数据集的实验结果显示，在1-shot和5-shot情况下，本文的模型分别取得48.2%、50.0%的平均交并比（MeanIoU），相比现有最优的小样本图像分割模型分别提高了2.0%和2.2%。

关键词：X光安检图像；小样本学习；图像分割；双路比较模块；卷积化门控循环单元

论文编号：467

基于 LTP 和 HOG 纹理特征融合的中亚文档图像文种识别

吴正健 1，木特力铺·马木提 2，吾尔尼沙·买买提 1，阿力木江·艾沙 2，库尔班·吾布力 1+

1．新疆大学 信息科学与工程学院，新疆 乌鲁木齐 830046；2．新疆大学 网络与信息中心，新疆 乌鲁木齐 830046

摘　要：多文种识别作为 OCR 的前端技术，对实现文档自动化有着比较重要的意义。由于中亚地区存在一些相似度较高的文种，提出基于一种具有旋转不变性的统一局部三值模式（riu2-LTP）和方向梯度直方图（HOG）特征交叉融合的方法。通过交叉融合结合 riu2-LTP 和 HOG 的纹理特征，均衡利用它们的特性，使用 SVM 分类器对包含十个文种共 10000 张图片的数据库进行实验；为了提高多文种识别效果，采用贝叶斯优化 SVM 的超参数。采用 5 折交叉验证避免过拟合，以 5 比 5 的比例划分训练集和测试集。该方法首先提取文档图像半径为 1，采样点为 8 的 riu2-LTP，再提取 64*64 的像素为一个 cell 的 HOG 特征，最后，将 20 个 riu2-LTP 特征与 36 个 HOG 特征分别依次融合到新的特征集。针对阿拉伯文，俄文，中文，哈萨克文，藏文，蒙文，土耳其文，维吾尔文，英文和吉尔吉斯坦文共十个文种，实验得出其查全率分别为 100%、100%、99.67%、97%、100%、100%、98.67%、99.67%、100%和 95%。通过对比实验表明，本文方法平均识别率达到 99%，相较于单一 LTP、riu2-LTP 和 HOG 方法有更好的性能；哈萨克文、土耳其文和吉尔吉斯坦文这样的相似文种不易区分的问题，得到了一定程度的解决。通过与近些年的其他多文种识别方法比较，显示了本文方法的有效性。

关键词：LTP; HOG; 特征融合; 贝叶斯优化

墙报 3-1

论文编号：47

基于颜料褪色光谱色差模型的壁画颜色复原方法研究

王可 1，王慧琴 1，王展 2，赵丽娟 1

1.西安建筑科技大学 信息与控制工程学院,陕西省西安市 710055；

2.陕西省文物保护研究院，陕西省西安市 710025

摘　要：古代壁画由于受到环境和人为因素影响导致颜料颜色褪变，辨识壁画初始颜色信息较为困难，现阶段通常使用的真彩色方法对古代壁画进行颜色复原容易受到同色异谱现象的干扰，影响复原颜色的真实性。针对该情况，本文基于颜料的光谱反射率“指纹”数据，通过光谱与色度的空间转换进行颜色信息精准表征再现，克服同色异谱的干扰；并针对降尘影响导致颜料褪色的问题进行研究，分析降尘导致颜料褪色的特性及其光谱特征变化的规律，建立颜料降尘样本库及褪色光谱的数学模型，得到降尘重量与颜料色差的拟合关系；在此基础上重建颜料光谱色差，计算壁画的初始真实颜色，并以真实的墓室与殿堂壁画作为实验案例进行分析。实验结果表明，与真彩色以及反射光谱直接复原的颜料颜色相比，褪色光谱色差模型能够消除降尘因素对颜料褪色的影响，可以复原壁画初始色彩，具有较好的色度精度和视觉主观评价效果，在人眼可接受范围之内，能够为壁画本体的颜料层修复提供辅助支持，并为认知及研究壁画原貌提供新的依据。

关键词：同色异谱；光谱色度空间；降尘褪色模型；壁画颜色复原

论文编号：186

基于悬链线的曲线保长度伸缩变形算法

李运平，王翠香

盐城工学院信息工程学院,盐城市,224051

摘　要：为实现曲线伸缩变形过程中长度保持不变，基于悬链线柔性力学特性以及几何非线性属性，将悬链线理论运用于曲线的伸缩变形，提出了一种保长度伸缩变形算法。依据等高悬链线的标准方程，分别创建具有单峰值和区域峰值特性的悬链线形状方程，用于驱动曲线实现单峰和区域峰值的保长度伸缩变形。在变形过程中，首先依据变形需求确定待变形曲线的伸缩区间，以伸缩区间所定义的待变形曲线段的两端点所在的直线为x轴，中点为原点建立局部坐标系；然后根据待变形曲线段的长度，建立局部坐标系下以待变形曲线段两端点为悬点的曲线段--悬链线形状控制方程，通过调整控制方程中的伸缩系数和伸缩因子等形状参数，完成局部坐标系下的待变形曲线段变形结果的计算；最后通过坐标变换和几何变换计算整体坐标系下的曲线变形结果。文中算法依据悬链线理论建立待变形曲线段的形状控制方程，使得曲线伸缩变形过程中曲线长度保持不变。为实现曲线的复杂变形和周期变形，通过将一条曲线划分成多段，分别依据悬链线理论编辑其形状，最后通过平移变换计算整条曲线的变形结果。数值实例验证了该方法的有效性和可控性。

关键词：自由变形；代数曲线；参数曲线；伸缩因子

论文编号：202

基于分形图案的数字凹浮雕生成研究

姜金岑 1，杨丽莹 1，王美丽 1，2，3

1. 西北农林科技大学信息工程学院，陕西杨凌 7121200

2. 农业农村部农业物联网重点实验室，陕西杨凌 712100

3. 陕西省农业信息与智能服务重点实验室，陕西杨凌 712100

摘　要：浮雕是一种历史悠久的雕塑艺术，能够在较少的空间内展示更多的信息。其中凹浮雕是在背景平面刻除相应图案得到的一种浮雕，镂空浮雕是凹浮雕的一种表现形式，它刻除了全部的背景平面。分形常用于生成具有相似特征的复杂场景，利用分形图案作为输入生成凹浮雕既拓宽了分形的应用范围，为分形图像提供一种新的呈现方式，又丰富了凹浮雕模型的多样性。本文针对从一般图像生成的凹浮雕模型设计感不强的问题，研究利用分形规则生成多样的、具有艺术设计感的图像，提出一种基于分形图案的凹浮雕生成方法。在经典分形算法基础上利用交互式Bezier曲线提高了分形图案中线条的光滑性，然后重用现有的艺术图案，通过分形拼贴将多个图案组合生成一个新的图像。研究基于距离变换的分形图案凹浮雕生成方法，以分形图案作为输入，利用距离图像能够反映深度信息的特点，对图像进行距离变换，实现凹浮雕表面的高度变化，再利用灰度映射将高度压缩到合适的范围使其符合凹浮雕的高度特点。最后通过三角化生成具有真实感的三维凹浮雕模型，并在凹浮雕基础上，通过刻除背景平面的三角面片生成镂空浮雕模型。本文将分形与凹浮雕进行了创新性的结合，对凹浮雕和镂空浮雕生成研究，增加了浮雕模型的多样性及真实感。

关键词：分形几何；图像处理；凹浮雕；三维重建；特征增强

论文编号：262

基于格子玻尔兹曼方法的涡量修正高效烟雾仿真算法

朱静远 1，马惠敏 2，*，袁坚 1

1.清华大学,北京市 100084； 2.北京科技大学,北京市 100083

摘　要：基于流体动力学模型的烟雾仿真研究是计算机图形学和计算机仿真领域的重要分支，涡量修正是烟雾仿真的重要步骤，通过合理的涡量修正可以得到高真实度的烟雾仿真效果。本文提出通过追踪涡量的大小和方向并添加涡量修正力进行涡量补偿的格子玻尔兹曼算法实现烟雾的高效真实仿真。算法从纳维-斯托克斯方程以及亥姆霍兹方程出发分别对烟雾仿真过程中涡量的大小和方向进行追踪，并利用追踪到的涡量信息和对应时刻的速度场信息添加涡量修正力进行涡量补偿和烟雾细节的补充，保证烟雾仿真符合物理规律并提升仿真的真实度。这一烟雾仿真算法基于格子玻尔兹曼方法实现，对动量分布函数，涡量分布函数以及温度分布函数进行离散化，通过三个分布函数和涡量方向场的迭代更新实现烟雾仿真。与格子玻尔兹曼方法相同，该烟雾仿真算法通过分布函数项在格点之间的迁移仿真流体的流动过程，通过添加碰撞项和外力项仿真流体粒子之间的相互作用以及流体粒子受外力作用。本文将这一算法的实验结果与之前提出的烟雾仿真算法进行对比验证模型的有效性，并通过计算帧率证明该算法在合理分辨率下可以达到实时仿真的效率，展示了该烟雾仿真算法的真实性和高效性。

关键词：涡量修正；烟雾仿真；格子玻尔兹曼方法；高效仿真

论文编号：336

基于 IFS 码的可控分形图算法

陶雪娇 1，柏森 1，何睿 1

1.重庆工程学院 重庆 400056

摘　要：分形图形仿真生成在信息隐写、艺术装饰、影视场景等领域均有用，但分形图形生成机理不是很明确，生成参数作用不是很清楚，所生成的分形对象可控性不足，使应用受到局限。因此研究IFS系统中各种仿射变换参数的作用，得到自由可控的分形图形生成方法，是非常有意义的。本文基于迭代函数系统生成分形图形的原理，分析了IFS码的每一个仿射变换生成时对分形图形外观的影响。通过分析仿射变换的类型及其变化规律以生成具有不同变换特性的分形图形。调整IFS码的可控性参数来仿真不同形态分形树和分形山，并以此构造更加自然的分形森林与分形群山，为信息隐写等应用领域提供可控性更强的分形图形。本文第2节介绍了迭代函数系统仿射变换公式，IFSP码拼贴生成分形图形的原理；第3节分析了真实分形图形整体与局部相似性，列出对应变换方程组，并用克莱姆法则求解IFS码，计算出各个仿射变换的概率集，实现具有分形特征分形物体仿真；第4节从仿射变换在分形图形生成中的作用入手，研究了分形图形形态变化机理，增减仿射变换的算法，探讨了基于IFS码基本的平移、旋转、缩放、拼贴等线性变换的可控分形树生成。本文的仿真结果表明本文的方法能够生成细节丰富、造型变化多端、像素细节可控的分形图形，以期为如信息隐藏的其他分形应用领域提供形态可控的分形图形。

关键词：分形；仿射变换；可控 IFS 码；分形森林；分形群山

论文编号：372

面向火电机组运行优化的复杂参数可视分析

纪连恩 1，2，王炎林 1，2，杨云 1，2，王艺 1，2

1.中国石油大学（北京）石油数据挖掘北京市重点实验室，北京 102249；

2.中国石油大学(北京)计算机科学与技术系,北京 102249

摘　要：火力发电机组累积的历史监控数据对其运行优化有着重要的分析价值。历史数据中参数众多， 它们之间既蕴含着复杂的机理关系，又具有一定的随机性，导致传统分析方法和工具难以对其深入发掘，为此，引入可视分析技术，设计iaDCS（interactive analysis of DCS）可视分析系统辅助优化机组运行过程。

本文首先对火电机组运行控制优化领域中的数据处理、分析需求和设计挑战进行抽象和提炼，提出了面向复杂参数控制优化的可视分析数据流模型。然后从展示、探索和评估三个角度建立iaDCS 可视分析系统，通过参数矩阵热力图支持多角度观察工况段数据，并对单一参数和成组参数进行相关性分析；通过内外部交互增强的平行坐标辅助用户从参数维度和参数区段两个方向展开探索，并结合多级联动的柱状图进行多级参数组合探索；通过平行坐标与旭日图、盒图的联动交互对参数优化方案提供多侧面评估，并将评估贯穿于整个探索过程。通过单参数探索及多参数探索，从历史控制数据中找到提高经济性的关键参数状态或参数组合，使评价指标达到最优。最后与领域专家紧密合作，采用真实电厂历史数据进行火电机组热经济性案例分析，揭示了参数间复杂相关性和变化规律，完成了参数组合寻优和多侧面评估，还发现了一些实际运行与理论设想不符的变化特征，验证了该可视分析工具的有效性和适用性。

关键词：可视分析；火电机组；复杂参数；控制优化；相关性

论文编号：373

基于稀疏超采样的时间性反走样算法

李根，陈文倩，张严辞

四川大学 视觉合成图形图像技术国防重点学科实验室　成都 610065

摘　要：时间性反走样算法因其帧间复用的思想，能够以较低的时间成本获得较高的反走样质量，在实时渲染引擎中广泛应用。针对时间性反走样算法在处理帧间复用时，若场景中有较多高频颜色区域或精细模型时，存在复用历史帧信息会失效的情况，会造成重影、模糊、闪烁及子像素细节丢失的问题，本文提出了基于稀疏超采样的时间性反走样算法。本文的核心思想是，在时间性反走样算法的基础上，对于无法复用历史帧的像素，重新引入空间域的超采样，在同一帧中生成多个采样点，有效地解决时间性反走样算法的复用问题。为了避免不必要的绘制开销，本文对复用成功的网格和片元进行预剔除，只对复用失效的网格和片元进行超采样，实现了对整个场景的稀疏超采样，同时提高本文算法的渲染效率。实验结果表明，本文算法能有效地处理时间性反走样算法中无法复用历史帧的情况，避免了重影、模糊、闪烁以及子像素细节丢失问题；本文算法能高效地剔除复用成功的网格和片元，实现对场景进行稀疏超采样的目的，避免了不必要的绘制开销；本文算法还能得到与超采样反走样算法媲美的反走样效果，并且具有更高的渲染效率。本文算法为反走样技术在实时渲染应用中提高渲染画面的视觉质量提供了一种可行的解决方案。

关键词：时间性反走样；稀疏超采样；剔除；复用

论文编号：434

基于样例的祥云图案生成

方建文 1，周小龙 1，曹文明 1，李勋祥 2

1.衢州学院 电气与信息工程学院,浙江省衢州市 324000;2.温州大学 美术与设计学院,浙江省温州市 邮编 325035

摘 要：祥云图案在中国悠久历史发展中传承和演变，成为最富有中国民族特色的装饰图案之一。它造型飘逸独特，婉转优美，具有极高的审美价值、文化价值和应用价值，已经被广泛装饰在古代建筑、雕刻、服饰、器具及各种工艺品上。祥云图案的创作目前主要依赖手工绘制，现有的大部分研究工作主要针对真实感云和烟的模拟和绘制，难以通过物理模型的参数调控和过程化方法的参数设置来生成艺术家需要的祥云图案形状，对生成祥云图案的研究工作还很少。本文提出一个基于样例和交互式草图的祥云图案生成模型。首先对手绘的祥云图案样例进行特征线的提取并归一化后形成特征库；用户输入简单的二维草图来分别表示云头、云体和云尾的骨架曲线；算法对云体和云头的草图轨迹进行非均匀自动划分得到生成线段集，利用这些线段到特征库中寻找相应的特征曲线并安装来得到云体和云头部分；通过对云尾骨架曲线进行平移和缩放来过程化生成云尾部分。实验结果表明通过该生成模型，用户通过很少的交互和参数控制就可以生成大量具有手绘风格特征的祥云图案。

关键词：祥云；建模；图案；中国传统风格

论文编号：392

基于非对称码率分配的立体视频传输算法

翟宇轩 1，林丽群 1，陈炜玲 1，房颖* 1,赵铁松 1

1. 福建省媒体信息智能处理与无线传输重点实验室，福州大学物理与信息工程学院,福建 福州,350108

摘　要：近年来，网络技术的发展推动了在线立体视频服务的兴起，但立体视频的大数据量使得传输问题更加严峻。为了保证网络服务的用户体验质量，立体视频的传输引入了基于HTTP的动态自适应流媒体技术，根据不同网络带宽自适应调整视频码率；同时结合人眼双目视觉特性，优化立体视频的传输算法。本文结合用户体验质量重点研究了基于非对称码率分配的立体视频自适应传输算法。首先，本文通过主观实验，构建左、右视点平面视频质量线性加权模型，利用左右视点的视频分辨率及量化参数确定左、右视点感知质量权重，实现对立体视频感知质量的预测；其次，以该线性加权模型为基础，结合网络吞吐量及客户端的缓存状态，设计适用于立体视频的自适应传输算法，指导左右视点不同的码率分配。立体视频传输系统客户端获取视频片段左、右视点的视频分辨率、量化参数、空间信息及时间信息等，根据预测带宽、缓存状态，利用该算法计算左、右视点最佳码率组合，使得用户体验质量最佳。本文搭建了立体视频的自适应流媒体系统进行算法性能验证，实验结果表明该算法能够为立体视频的左右视点分配非对称码率，从而适应带宽的动态变化，保证立体视频流畅播放且视频质量平滑，系统传输性能稳定，提高网络视频用户体验质量。

关键词：立体视频；视频传输；体验质量；码率分配

墙报 3-2

论文编号：6

面向通用空域错误隐藏的自适应混合填充算法

邓开连，刘浩，黄荣，袁浩东

东华大学信息科学与技术学院， 上海 201620

摘 要：在现有的空域错误隐藏算法中，运行快速的多方向插值算法在处理多种丢失模式时效果不甚理想，而恢复质量好的迭代类算法又难以满足实时性需求。为此，本文提出了一种面向通用空域错误隐藏的自适应混合填充（AHF）算法，以更好地进行多性能指标的联合优化。对于当前受损块的错误隐藏，AHF算法首先通过各向同性梯度检测器学习延拓区域的邻域梯度特征，根据收缩填充次序执行局部预测过程，从外层素组到内层素组逐一地恢复受损块的各个素组；若预测相关性足够低，AHF算法将转而执行非局部片匹配过程，利用相似片对同样位置的未隐藏像素进行填充。实验结果表明相比于其它代表性的算法，所提AHF算法较好地平衡了通用性、计算复杂度和恢复质量等多指标性能，在各种丢失模式下取得了具有竞争力的综合性能。

关键词：通用空域错误隐藏；混合填充；局部预测；非局部片匹配

论文编号：78

基于图像全位面的高嵌入率可逆数据隐藏

周旭 1，吴福虎 1，陈志立 1，任帅 1

1.安徽大学计算机科学与技术学院,安徽省合肥市 230601

摘　要：近年来，数据隐藏技术引起了学术界的广泛关注，提供服务者将隐蔽信息嵌入到数字图像中进行安全传输。然而，数据嵌入往往会给原始图像带来破坏，在数据传输完成后无法完全恢复图像。因此，密文域的可逆数据隐藏技术成为学术界的研究热点。具体做法是首先对载体图像进行加密，然后利用可逆数据隐藏将隐蔽信息嵌入到加密图像中，通过隐蔽传输，发送给接收者。算法大体上可分为两类：vacating room after encryption（VRAE）和reserving room before encryption（RRBE）。之前的可逆数据隐藏方法中，加密图像冗余空间少，位平面利用率不高，嵌入容量小，翻转一些像素会导致图像失真。本文把灰度图像的的八个位平面都用来进行数据嵌入，将每个位平面划分成为不重叠的块，分为非连续块（块内像素值0,1都存在）和连续块（块内为全0或全1像素值），按块进行重排列且将排列前的块标签嵌入到图像中，使用流密码对排列后的图像进行加密。在数据嵌入阶段，提出了带修正信息的像素预测方法。连续块中，保持块内右下角像素值不变，用于连续块的恢复，其他位置嵌入数据；非连续块中，对预测正确的像素嵌入数据，预测错误的像素保持不变。实验过程实现了多种密文域可逆数据隐藏算法，做了大量的对比实验，并在BOSSbase数据集上做了验证实验，结果表明本方案实现了完全可逆，具有较高的嵌入容量。

关键词：像素预测；图像加密；可逆数据隐藏；图像恢复

论文编号：92

基于混沌置乱和DWT变换的无载体图像信息隐藏研究

谭云1，秦姣华1*，向旭宇1，刘强1，周青阳1

1.中南林业科技大学计算机与信息工程学院，湖南长沙，410004

摘　要：无载体信息隐藏通过提取载体特征，并与信息序列进行映射，从而无需修改载体即可实现对秘密信息的隐藏，因此，具有极强的抗隐写分析能力，是近年来的研究热点。但是已有无载体信息隐藏算法在隐藏容量方面仍然有限，且大多需要构建大样本图像库。本文提出一种基于混沌置乱和离散小波变换（DWT）的无载体图像信息隐藏算法，首先，对载体图像进行混沌块置乱，随后，对置乱后的图像分块进行DWT变换，根据相邻图像块的低频DWT系数之间的关系生成对应的哈希序列，实现秘密信息的承载。不同的混沌序列参数可以生成不同的置乱图像，可映射为不同的秘密信息，因此通过设置不同的秘钥，每张载体图像可以承载的秘密信息容量获得极大提升。将载体图像及秘钥发送给接收方，即可实现秘密信息的传递。实验结果表明，本文提出的算法针对JPEG压缩、噪声、滤波等攻击具有较强的鲁棒性。与现有算法相比，本算法在容量和隐藏成功率上获得了较大提升，极大改善了目前已有方案中的信息容量不够和需构建大样本数据库的问题。同时，本算法方案简单，传输负载小，具有较强的实用价值。

关键词：无载体信息隐藏，混沌映射，图像置乱，DWT变换

论文编号：**107**

基于稀疏采样的多媒体传感网图像压缩算法

郭　剑 1,2，韩　崇 1,2，施金宏 1，薛浩天 1，孙力娟 1,2

1.南京邮电大学计算机学院　南京　210003

2. 江苏省无线传感网高技术研究重点实验室　南京　210003

摘　要：无线多媒体传感器网络（WMSNs）是一种由大量低成本的传感器节点通过无线通信组成的自组织网络系统。这些节点通常配备有麦克风、摄像头以及其他传感器来收集多媒体数据。随着互补金属氧化物半导体（CMOS）和低成本无线网络技术的日益成熟，WMSNs现已被广泛应用于图像采集，视频监控等领域。由于WMSNs可以捕获到更多维的数据，其获得的数据量更大，处理和传输时所需的节点能耗也会随之激增。然而，对于部署在恶劣自然环境中的WMSNs，其节点能量有限，且难以进行补充。一旦电池耗尽，节点将无法正常工作。因此，有限的能源存量已经成为限制WMSNs发展的瓶颈之一，设计适合的压缩方案来降低处理和传输的能耗是非常有必要的。现有的相关研究成果大多数设计能耗较少的数据压缩算法，然而这些方案在数据采集的编码端依然包含有较复杂的计算工作，这会导致有限供电的传感器节点在图像处理过程仍旧需要大量的能源消耗。本文结合WMSNs的特点，通过引入压缩感知技术，提出了一种基于傅里叶域稀疏采样的分布式多媒体传感网图像压缩算法。该算法考虑多传感器节点的协作性，实现了图像在傅里叶域上的分块自适应压缩采样与传输。与现有其他的分布式压缩感知算法的仿真对比实验表明了本文所提算法的有效性。

关键词：无线多媒体传感网；分块压缩感知；傅里叶变换；稀疏采样；自适应采样

论文编号：113

局部边缘结构引导的深度图像超分辨算法

王春兴 1，祖兰晶 1，万文博 2*，任艳楠 3，葛川 4，孙建德 2

1.山东师范大学物理与电子科学学院，济南 250358

2.山东师范大学信息科学与工程学院，济南 250358

3.山东交通学院信息科学与电气工程学院，济南 250358

4.中国移动通信集团山东有限公司，济南 250001

摘　要：由于高分辨率深度图像的广泛应用，人们越来越关注深度图像超分辨的研究。目前的深度图像超分辨算法均存在未有效利用像素间的相关性以及未对深度图与彩色图在边缘区域的结构实现细致比较等问题。为此本文提出局部边缘结构引导的深度图像超分辨算法。首先，在低分辨率深度边缘图的指导下，对低分辨率深度图像完成不可靠像素的修正。结合高分辨率彩色图像的梯度矩阵，基于以某个像素点为中心的3×3像素块完成深度图与彩色图的结构一致性判断，并完成像素点分类。最后，根据像素间的影响以及空间位置约束，利用存在的有效深度完成深度修正。将本文的算法与当前存在的算法分别进行客观数据与主观视觉上的对比，实验结果证明本算法可输出结构清晰、坏点率降低、细节边缘区域有效增强的高分辨率深度图像。

关键词 ：深度图像；边缘结构；超分辨；深度修正

论文编号：146

基于自适应参数引导与边缘点辅助的深度图像超分辨率重建

王春兴 1，祖兰晶 1，万文博 2*，任艳楠 3，葛川 4，张凯 2

1.山东师范大学物理与电子科学学院，济南 250358

2.山东师范大学信息科学与工程学院，济南 250358

3.山东交通学院信息科学与电气工程学院，济南 250358

4.中国移动通信集团山东有限公司，济南 250001

摘　要：当前TOF等深度相机仅能获取低分辨率的深度图像，无法满足三维视觉系统的需求。高分辨率深度图像可通过深度图像超分辨算法获得，但目前存在的算法输出的实验图像存在边缘结构不清晰，以及纹理复制伪影等问题。为此本文提出了一种基于自适应参数引导与边缘点辅助的深度图像超分辨率重建算法。本文利用低分辨率深度图的边缘点完成对低分辨率深度图像中不可靠像素点的标记与修正以及像素点聚类，并对初始化深度图进行深度重建。对处于结构复杂区域的像素点利用已加入多个自适应参数的自适应自回归模型完成深度预测。将本文的算法与当前存在的算法分别进行客观数据与主观视觉上的对比。实验结果证明该算法能够有效降低深度不连续区域的模糊性，获得性能稳定、可靠的高质量的高分辨率深度图像。

关键词：深度图像；超分辨率重建；自适应参数；边缘点检测；深度修正

论文编号：266

基于生成对抗网络的 HDR 图像生成方法

贝悦 1，王琦 1，程志鹏 1，潘兴浩 1，汪立民 2，雷健 2，甄广坤 2，

丁丹丹 2*

1. 咪咕视讯科技有限公司　上海　201201

2. 北京观止创想科技有限公司　北京　100036

摘　要：高动态范围(High Dynamic Range，HDR)图像相比低动态范围(Low Dynamic Range，LDR)图像有更宽的色域和更高的亮度范围，更符合人眼视觉效果。 但由于目前的图像采集设备大都是LDR设备，导致HDR图像资源匮乏，解决该问题的一种有效途径是是通过逆色调映射将LDR图像映射为HDR图像。本文提出一种基于条件对抗网络(Generative Adversarial Network，GAN)的逆色调映射算法，以重建HDR图像。为此，设计了基于多分支的生成网络与基于鉴别块的鉴别网络。在生成网络中使用多分支结构，提取图像局部、中频、全局特征，并将不同尺度的特征进行融合得到最终的HDR图像。在鉴别网络中，使用组合的鉴别块完成鉴别任务。利用GAN的数据生成能力和特征提取能力，将单张LDR图像从BT.709色域映射到对应的BT.2020色域。实验结果表明，与现有方法相比，所提出的网络能够获得更高的客观与主观质量。特别是针对低色域中的模糊区域，所提出的方法能够重建出更清晰的纹理与细节。所提出的方法在客观质量上超过了以往方法，尤其是在直接反应HDR重建图像质量的MS-SSIM与HDR-VDP指标上取得了较高性能。在主观质量上，所重建的HDR图像看上去自然并且无显著色彩失真。

关键词：生成对抗网络；卷积神经网络；逆色调映射；色域转换；特征提取

论文编号：269

针对伪影改善的图像拼接方法

姚莉 1，2，杜俊康 1

1.东南大学计算机科学与工程学院,江苏省南京市 211189； 2.计算机网络和信息集成教育部重点实验室（东南大学）,江苏省南京市 211189

摘　要：图像拼接技术的发展已经日趋成熟了，并且也广泛应用到了智能手机拍摄、虚拟现实、航天航空和视频监控等诸多领域当中，其目的都是为了获取到宽视野、高分辨率和观感相对自然的图像。然而目前的图像拼接技术在应用到不同场景时，图像的特征数目相对较少、特征种类单一，且特征分布存在局限性，同时图像对重叠区域中深度不一致。本文针对图像拼接中的伪影问题，提出了一整套的拼接解决方案。图像拼接方法分为两大阶段：图像配准阶段和图像融合阶段。在图像配准阶段，提出了一种基于网格优化的图像配准方法。通过使用点特征和线特征相结合的特征提取算法，提高特征的数量和质量。同时提出了基于先验的RANSAC特征点对筛选方法，剔除误匹配点对，并提高了模型计算的速度。为了进一步提升对齐效果和优化伪影问题，提出了多项图像对齐约束。最后采用了柱面投影模型，将图像对投影到同一平面。在图像融合阶段，使用了基于图像对重叠区域细节增强的缝合线检测算法，优化多焦段场景下的拼接效果。利用渐入渐出融合算法对同一平面的投影图像进行融合，得到一幅宽视野的高分辨率图像拼接结果。通过将现有的拼接方法和本文方法进行实验数据集进行主观和客观验证对比，结果都表明本文的方法能够更好地应对伪影问题，在拼接质量上有着更优的表现。

关键词：图像拼接；伪影；网格优化；图像配准

论文编号：413

基于Bregman迭代字典学习方法的定量磁化率成像

杜嘉程1，申云鹏1，陈阳2

东南大学生物医学与工程学院,南京 210096；

东南大学计算机科学与工程学院,南京 210096；

摘　要：定量磁化率成像(Quantitative Susceptibility Mapping)技术作为一种新颖的磁共振成像技术，可以从包含了丰富磁化率信息的相位信号中重建出组织磁化率分布图像，并能定量的分析体内铁含量、钙化，血氧饱和度等相关信息，可以为许多疾病提供诊断信息，具有重要的临床应用价值。

磁化率分布图像的重建过程包括场图拟合，相位解卷绕，背景场去除以及偶极核反演这四步。由于偶极核在K空间中与主磁场方向呈54.7°角的位置值为0，使得偶极核反演是一个病态的逆问题，目前，有多种方法用于解决这个逆问题，例如多方向采集磁化率计算方法，单方向的方法有傅里叶域的阈值截断法，有空间域中的全变分方法，闭式L2约束方法以及基于形态学相似的偶极核反演方法，然而这些方法均有一定的局限性。在磁化率分布图像的重建中，如何有效的抑制伪影，提高重建质量仍然是一个重要的挑战。

近些年来，基于字典学习的方法在低剂量CT，降采样MRI的重建中，可以很好的抑制噪声及伪影。本文基于字典学习的方法，提出基于Bregman迭代字典学习(BI-DL)的磁化率分布图像重建方法用于解决偶极核反演这一逆问题，在重建过程中，通过Bregman迭代方法在重建过程中更新字典并用于约束磁化率分布图像的重建。通过3T人脑数据验证算法的有效性，结果表明该方法相较于主流的重建方法在伪影抑制的效果以及定量分析的结果上均有较大的提升。

关键词：定量磁化率成像；偶极核反演；Bregman迭代；字典学习

墙报 3-3

论文编号：291

基于二元优化模型的自适应直方图均衡

汪子玉 1，戴声奎 1，朱益铭 1，陈世龙 2

1.华侨大学信息科学与工程学院，福建省厦门市 361021；

2.泉州科缔恒电子科技有限公司，福建省泉州市 362000

摘　要：直方图均衡(HE)方法具有简单快速和效果显著的优点，在图像增强领域得到了广泛的应用。但是HE方法的缺点也是明显的，其结果图像容易出现过度增强、亮度均值漂移和细节丢失等问题。因此诸多种类的HE改进算法被相继提出，其中一类方法是将图像直方图进行伽马变换调整后再进行均衡操作。目前许多算法的伽马参数为固定值或者经验值，因此，此类HE改进算法缺少自适应性。本文设计并实现了一种求取最优伽马参数的方法：首先将一张输入图像生成图像序列（每张图像对应不同的伽马值），统计每张图像的信息熵和对比度，通过提出的二元优化模型计算得到最优伽马值；此外，本文根据灰度分辨率曲线对直方图进行钳位处理，使得直方图均衡后的图像视觉感知更舒适；最后，针对HE算法细节丢失的通病，本文提出了直方图二次伽马校正方法，加强了图像细节的保护能力。在多个数据库的测试中，本文算法除了具有良好的主观视觉效果之外，客观指标结果也领先于其他对比算法。同时，本文算法实际上不需要真的生成序列图像，仅通过直方图变换来计算序列的统计参数即可，因此本文算法继续保留了HE算法计算复杂度低的优点。

关键词：伽马变换；对比度增强；优化模型；视觉特性；直方图均衡

论文编号：296

完美可恢复的近似缩略图保持加密技术

张玉书*，叶茜，赵若宇

（南京航空航天大学 计算机科学与技术学院 南京 211106）

摘　要：随着智能终端的普及，人们愈发习惯于将图片存储在云端，从而没有了后顾之忧，不用担心大量照片的存储问题。人们在享受图片云存储的好处时却面临着泄露个人隐私的风险，为此很多学者提出了阻止隐私泄露的方法。然而，绝大多数方法只重视隐私性的保护却忽略了可用性。近几年来，Marohn等人提出了近似缩略图保持的加密方法，一定程度上平衡了隐私性和可用性，但是该技术中对图像是实施了不可逆的有损操作，加密后的图像无法完美恢复。为此，本文设计了一种可以完美恢复的近似缩略图保持的加密发方法，将信息隐藏技术与加密技术相结合，首先利用差分扩展将调整区数据嵌入到嵌入区，然后对整幅图像实施传统的图像加密技术，最后调整调整区像素。最终，加密后图像的缩略图与原始图像的缩略图近似，从而很好地平衡了隐私性和可用性。经过实验可以发现，该方法不仅很好地保护了用户的隐私，而且保留了基本的可用性，如图像智能分类等功能。人们也能通过加密调整后图像的缩略图，识别出原来的图像，继而在线进行一系列操作，而无需下载原图。

关键词：差分拓展；隐私保护；完美可恢复；近似缩略图；缩略图保持加密

论文编号：307

多粒度深度神经网络的图像超分辨率重建方法

杨文元 1，2，沈泽君 1，刘建然 1

1.福建省粒计算及其应用重点实验室(闽南师范大学)，漳州 363000；

2.闽南师范大学计算机学院，漳州 363000

摘　要：图像超分辨率重建技术（Single Image Super-Resolution, SISR）是解决在给定的低分辨率图像下生成高分辨率的问题。SISR是数字图像处理中的一个重要研究方向，在图像处理软件、医疗诊断、卫星图像遥感和机械工程领域中有着广泛的应用。然而传统的超分辨率重建方法未能充分考虑不同像素级间的粒度关系，多在算法上进行细究、对模型参数进行调整或者与其它算法模型进行结合。本文提出一种结合粒计算思维的多粒度深度神经网络的图像超分辨率重建方法（Multi-granularity deep neural network for super-resolution reconstruction, MGSR）。MGSR算法能够分析人脑思维模式，以人脑的思维模式更高效的模拟人脑行为模式来综合分析不同像素级的结果并筛选最优。首先，对输入的低分辨率图像进行粒度划分，划分成若干个不同粒度大小的粒度块。然后对不同粒度层下的粒度块进行深度神经网络运算，得出各粒度层的高分辨率图像块。最后，进行粒度拼接输出高分辨率图像。MGSR与HEVC、SRCNN和VVC算法进行比较，在Y-PSNR评价标准上分别高出1.4365、2.6248和0.6649；在U-PSNR评价标准上分别高出2.3069、1.3584和1.7481；在V-PSNR评价标准上分别高出1.6826、0.9673和1.0254；在YUV-PSNR评价标准上分别高出1.47、2.6525和0.739，同比高出3.99%、7.44%和1.97%。在信噪比上，SRCNN算法整体表现好，MGSR算法在FourPeople和SlideEditing等图像上取得最高值。实验结果显示MGSR算法在峰值信噪比上取得更高的处理能力。

关键词：高效视频编码；超分辨率重建；卷积神经网络；粒计算；残差学习

论文编号：356

一种基于邻域像素差值最大化的空域隐写嵌入方法

郑淑丽，余聪，胡东辉

合肥工业大学计算机与信息学院,安徽合肥 230009

摘　要：信息隐藏以图像、视频、音频、文本等载体为掩护，在其中通过特殊的编码算法嵌入秘密信息，使得含密载体在公开信道传输时可以规避第三方检测，以期进行安全的秘密通信。在图像中进行隐写时，需要抵御视觉、统计分析、特征学习等隐写分析手段。当前基于深度学习隐写分析方法对像素残差特征进行学习。为了抑制图像内容造成的干扰，保留更多有效隐写分析特征，信号截断和激活成为隐写分析模型的常用结构。针对基于深度学习的隐写分析攻击，本文提出一种空域隐写嵌入方法，通过引入嵌入方向的差异，放大嵌入点的像素差值，一方面在预处理阶段通过残差截断丢失部分隐写特征，另一方面使得嵌入信号的激活值向梯度为零的区域趋近，对基于深度学习隐写分析模型进行干扰。嵌入过程中对图像进行分组，提高目标嵌入点数量，降低基于深度学习隐写分析的检测准确率。在实验中，我们以 HILL 作为目标隐写算法，使用 Xu-Net 和 Ye-Net 的检测准确率作为评估标准。在嵌入率 0.4pb 时，本文方法可使 Xu-Net 和 Ye-Net 的检测准确率分别降低 4.05%和 3.04%。随着嵌入率减小，本文方法逐渐退化为 HILL 算法，因此优势逐渐变小。实验表明，本文方法可以有效提高空域隐写方法的抗检测能力。

关键词：隐写；残差特征；截断；梯度；抗检测。

论文编号：357

双加权 LP 范数 RPCA 模型及其在椒盐去噪中的应用

董惠雯 1，禹晶 1，郭乐宁 1，肖创柏 1

1. 北京工业大学信息学部，北京市 100124

摘　要：鲁棒主成分分析（RPCA）模型旨在分离观测数据中潜在的低秩成分与稀疏成分，RPCA 模型中秩函数和 ℓ_0 范数的求解是 NP 难问题，凸近似模型的求解通常导致过收缩问题，现有的改进 RPCA 模型采用各种方法对秩函数和 ℓ_0 范数进行松弛。本文提出了一种基于双加权 ℓ_p 范数的 RPCA 模型，结合加权方法和 ℓ_p 范数，利用加权 S_p 范数低秩项和加权 ℓ_p 范数稀疏项分别建模 RPCA 框架中的低秩恢复问题和稀疏恢复问题，使其更接近秩函数和 ℓ_0 范数最小化问题的解，提升了矩阵秩估计和稀疏估计的准确性。为了验证本文模型的性能，本文利用图像的非局部自相似性，结合相似图像块组的低秩性与椒盐噪声的稀疏性，将双加权 ℓ_p 范数鲁棒主成分分析模型应用于椒盐噪声的去除。定量与定性的实验表明，本文算法与现有算法相比均有明显的提升，并通过奇异值过收缩分析验证了本文模型能够有效抑制秩成分的过度收缩。

关键词 ：图像去噪；鲁棒主成分分析（RPCA）；低秩；稀疏；非局部自相似性（NSS）

论文编号：390

一种基于物理光流与细节增强的湍流图像恢复方法

刘亮 1，蔡泽民 1，赖剑煌 2

1.汕头大学工学院电子工程系,广东汕头 515063； 2. 中山大学数据科学与计算机学院,广东广州 510275

摘　要：针对远距离成像系统在大气湍流中压强、温度等因素的影响下，得到的图像序列会产生图像畸变与模糊等图像质量下降的问题，为了解决湍流下图像失真的问题，目前最常用的方法主要有三种，一是利用图像配准和图像融合思想，二是稀疏表示的思想重建去模糊图像，三是使用各种先验思想计算出模糊核。本文提出了一种基于物理光流与引导图滤波器细节增强的方法，从受湍流影响图像序列中恢复出清晰图像。首先，利用物理光流计算序列帧与参考图像之间的位移信息差，通过位移校正获得单张去扭曲的图像；其次，采用归一化稀疏表示进行去模糊；最后，利用引导图滤波器将细节增强图像从单张去扭曲图像中提取出来。实验结果表明，本文方法能有效去除湍流图像的畸变和模糊，获得相对清晰的图像，与其他方法相比，本文实验结果在平均结构相似性（MSSIM）和PSNR值上的表现有较大提升，证明了本文方法的优越性，并且，物理光流法对像素位移差的计算更加准确，而且速度更快，极大的降低了对有效序列帧数量的要求。在本文的实验中，我们使用5至10帧湍流图像就可以恢复出去扭曲图像。在细节增强阶段，本文采用引导图滤波器对图像进行增强，使图像细节更加清晰，恢复效果更理想。

关键词：去扭曲；物理光流；稀疏表示；盲反卷积；引导图滤波器

论文编号：458

基于样本错配训练的图像 PRNU 噪声提纯方法

郝昕泽 1，肖延辉 1，田华伟 1，张明旺 2

1.中国人民公安大学国家安全学院,北京 10038； 2.四川警察学院科研所,四川 泸州 646000

摘　要：基于传感器模式噪声的图像来源鉴别算法的关键在于如何获取高质量的PRNU噪声，目前大多数增强PRNU噪声质量的算法以及出于实际应用的目的对其压缩的算法几乎是在人工假设的模型的基础上实现的。鉴于近些年来深度学习技术在图像特征提取、图像内容识别等领域取得的革命性突破，并且已经成功解决了一些复杂的应用问题，本文提出了一种基于样本错配训练的图像PRNU噪声提纯方法。本文首先设计了一种深度堆叠自编码器网络，并设计了一种与其对应的“样本错配策略”的训练方法，即改变传统自编码器输入与自身对应训练的方式，采用输入与其同类相似样本交叉训练的方法。通过在Dresden图像数据集上进行的比较实验结果表明，本文提出的图像PRNU噪声提纯方法对基于小波变换去噪算法提取的PRNU 噪声和基于BM3D去噪算法提取的PRNU噪声均具有一定的提纯作用，并且在64×64的图像分辨率下，对基于小波变换去噪算法提取的PRNU 噪声的Kappa系数提升了0.0077，可以有效提升其在图像来源鉴别任务中的准确率。这种基于样本错配训练的端到端深度神经网络的使用有效避免了人工设计算法的局限性，对图像PRNU噪声进行了有效提纯，进而提升了其用于图像来源鉴别时的性能。

关键词：图像来源鉴别；光照响应不一致；样本错配；深度堆叠自编码器；

论文编号：469

数据驱动的 AVS3 像素域 JND 预测模型

李兰兰 1，刘晓琳 1，吴珂欣 1，林丽群 1，魏宏安 1，赵铁松 1

1. 福建省媒体信息智能处理与无线传输重点实验室，福州大学物理与信息工程学院,福建 福州,350108；

摘　要：AVS3作为我国第三代国家数字音视频编码技术标准，在消除视频时域/空域冗余信息方面做了大量的工作，但在消除感知冗余方面仍存在进一步优化的空间。本论文提出了一种数据驱动的AVS3像素域JND预测模型，在尽量保证视觉主观质量的前提下，对AVS3视频编码器进行优化。首先基于现有的主流大型JND主观数据库（VideoSet），获取符合人眼视觉特性的像素域JND阈值；然后基于深度神经网络构建深度像素域JND预测模型；最后通过构建的深度像素域JND预测模型所预测的像素域JND阈值建立残差滤波器，来消除AVS3的感知冗余，降低编码比特率。为了验证提出模型的性能，本论文通过客观实验来比较编码比特率的节省情况，PSNR与MS-SSIM客观指标显示无明显差异。然后采用进行主观实验的方法来进一步验证人眼对于所提出模型算法所编码的视频主观感知情况，结果显示参与测试人员无法区别出标准测试模型HPM5.0编码视频与所提算法编码视频，因此可认为所提算法模型编码视频是感知无损的。实验结果表明：与AVS3的标准测试模型HPM5.0相比，在保证人眼主观感知质量几乎无损的情况下，所提出的像素域JND模型最高可节省21.52%的码率，平均可节省5.11%的码率。

关键词：视觉感知特性；恰可察觉失真（JND）；AVS3；残差滤波

论文编号：350

纸质包装品精密检测及评价系统

龚健 1，李海鹏 1，徐丹 1，李杰 1，张浩 1，张豪远 1，张超 2，罗忠 2

1. 云南大学信息学院计算机科学与工程系，云南省 昆明市 650500；
2. 昆明超泰经贸有限公司，云南省 昆明市 650051
3. 云南锐驰数控技术有限公司，云南省 昆明市 650051

摘　要：为了解决压印包装材料在不同温度和湿度条件下和高速包装机的上机适应性问题，本文设计了一套纸质印刷品无接触式精密检测及评价系统。使用高精度激光微距扫描设备采集样本表面原始数据，搭载负压吸附平台及双轴直线电机对样本表面进行全幅面采样，实验平面均匀采样密度高于1100点/平方毫米，深度数据理论误差小于0.02毫米，解决了接触式机械测量方案无法完整覆盖样品表面，以及量具接触样本造成样品表面受压形变的问题。本文对烟标全幅面的采样数据进行拼接及可视化重构，实验表明本文方案测量精度高于业界精度要求，并提出了基于局部频率变化的压痕饱满度、压痕宽度等参数检测标准，测试结果的方差能有效反应机械测量无法还原的同批样品中不同样本间的微小个体差异，定位到样品表面结果易波动区域，而且对于不同批次的合格样品，本文方法检测出的样本工艺控制差异更显著，具有较好的分类性能。测试版系统提供了基于烟标印刷生产线需求设计的数据报告打印接口，数据报告可用于指导压印生产线工作，以及基于原始数据重构的等高线着色三位模型展示窗口，并提供交互式角度和长度测量功能。从而帮助解决纸质压印制品工业生产中残次品检测、上机适应性验证、不同生产环节的物理参数沟通等困难。

关键词：数据可视化；精密检测；压痕测量；轮廓拟合；三维重建

论文编号：38

基于激光—相机系统的语义栅格建图和路径规划

丁梦远 1，郭迟 2，3，黄凯 1

1. 武汉大学计算机学院,武汉 430072； 2.武汉大学卫星导航定位技术研究中心,武汉 430072；

3.武汉大学人工智能研究院,武汉 430072

摘 要 ：智能化移动机器人广泛应用于工业、军事、物流、家用服务等领域。SLAM是移动机器人在未知环境进行探索、感知和导航的关键技术。激光SLAM测量精确，便于机器人导航和路径规划，但缺乏语义信息。而视觉SLAM的图像能提供丰富的语义信息，特征区分度更高，但其构建的地图不能直接用于路径规划和导航。为了实现移动机器人构建语义地图并在地图上进行路径规划，本文提出一种语义栅格建图方法。首先，建立可以同步获取激光和语义数据的激光-相机系统，根据联合标定将这些数据转换为同一坐标系下，并对采集的激光分割数据与目标检测算法获得的物体包围盒进行匹配，得到各物体对应的语义激光分割数据。然后，将连续多帧语义激光分割数据同步融入占据栅格地图，并根据语义栅格更新得到各物体置信度。最后，对具有不同语义类别的栅格进行聚类，得到标注物体类别和轮廓的语义栅格地图。此外，针对语义栅格地图发布相关导航任务，利用路径搜索算法进行全局路径规划，采用语义赋权算法减少易移动物体对全局路径的影响，并根据语义信息对路径进行优化。多种环境下的实验证明本文提出的方法能获得与真实环境一致性较高、标注环境中物体类别和轮廓的语义栅格地图，且实验硬件结构简单、成本低、性能良好，适用于智能化机器人的导航和路径规划。

关键词：智能机器人；语义栅格地图；激光 SLAM；目标检测；路径规划

墙报 3-4

论文编号：14

基于相关熵的多视角彩色点云配准

刘荧 1，李垚辰 1，刘跃虎 2，付靖文 3，马飞 1

1.西安交通大学软件学院,西安，710049；2.西安交通大学人工智能与机器人研究所,西安，710049；3.西安交通大学计算机科学与技术系,西安，710049；

摘　要：与传统图像数据相比，三维彩色点云数据能够更加细致的表征物体的空间位置、颜色信息等。因此，运用点云数据来解决问题已成为计算机视觉领域的热点之一。由于存在遮挡问题，每次使用设备采集时只能获取物体的部分信息，需要在多个角度进行采集，并将采集的数据进行配准，才能得到完整的三维点云数据。迭代最近点算法是最经典的点云配准算法，通过不断迭代建立点云间对应关系并求解空间变换，最小化对应关系中对应点对之间的距离，直到算法收敛得到最优的匹配参数。本文提出的算法在ICP算法的基础上进行了改进，用传统的三维坐标与HSV颜色空间的色调值相结合，形成四维的特征向量，建立点云之间的对应关系，以提高配准精度。同时，结合相关熵以减小离群值和噪声对配准造成的影响，从而提高算法的鲁棒性。进行双视角配准后，将该算法所计算的结果作为多视角配准的初值，然后通过运动平均算法减小累计误差，完成更加精确的多视角配准。实验结果表明，本文的算法在精度、鲁棒性方面与同类算法相比，均具有明显优势。此外，该算法的计算结果作为初值与运动平均算法良好结合，能得到可靠的多视角点云配准结果。

关键词：相关熵；彩色点云；迭代最近点算法；运动平均算法

论文编号：444

融合八度卷积与残差学习的 JPEG 内容自适应隐写分析

兰鑫玥 1，2，倪蓉蓉 1，2，赵耀 1，2

1.北京交通大学信息科学研究所，北京 邮编 100044；

2.北京市现代信息科学与网络技术重点实验室，北京 邮编 100044

摘　要：目前JPEG图像内容自适应隐写分析方法难以有效建模图像中纹理复杂或者噪声区域的复杂统计特性，检测精度低。本文提出融合八度卷积与残差学习的隐写分析方法能够有效针对内容自适应隐写嵌入改变集中在纹理复杂的高频区域的特性，提取到更利于分类的特征。首先进行DCT滤波预处理，增强图像的信噪比，抑制图像内容对隐写弱信号的干扰；其次设计一种融合了八度卷积与残差学习的OC-Block结构，分离高低频特征并压缩低频，重点提取与嵌入位置相关性较高的高频区域特征，并且可以在深层中保留由消息嵌入生成的微弱信号。实验对JC-UED和J-UNIWARD两种内容自适应隐写方法进行检测，与SCA-GFR、PNet、VNet、XuNet、CNN-SCA-GFR这5种方法进行了比较，针对JCUED隐写，在质量因子为75时，0.4bpnzac以上的高嵌入率下，检测的准确率可以达到98.83%以上，提升了0.79~2.29%。在0.1bpnzac的低嵌入率下，提升了1.98~7.41%。模型在不同时期训练和测试的检测错误率接近，证明模型有很好的泛化能力。本文所提出的融合八度卷积与残差网络的方法能够有效针对内容自适应隐写嵌入集中在纹理复杂的高频区域的特性，提取到更利于分类的特征，提升隐写的检测准确率。

关键词：隐写分析；内容自适应；残差网络；八度卷积

论文编号：76

基于 DMIR 算法的多重复纹理图像拼接

张琳娜 1，陈建强 1，吴研 2，张悦 3，岑翼刚 3

1.贵州大学机械工程学院，贵州省贵阳市 550025；2.贵阳市公安局，贵州省贵阳市 550025；3 北京交通大学计算机与信息技术学院，北京市 100044

摘　要：图像拼接技术广泛应用于全景图像拼接、虚拟现实、无人机遥感图像拼接、多图像视频拼接等领域。目前已经存在的方法大部分都是关于特征提取，单应矩阵和图像融合方面，在特征匹配方面涉及很少。然而现实生活中有很多重复纹理的场景，比如超市的货架、高楼上的窗户、地面的地砖等，利用传统的方法对多重复纹理的图像进行拼接，拼接效果并不理想，究其原因并不是因为提到的特征点不对，而是其特征点进行匹配后，误匹配率非常高。针对目前图像配准算法对于多重复纹理图像配准位置偏差的问题，提出图像内自匹配与图像间互匹配相结合的双匹配配准算法 (Double Match Image Registration, DMIR)。首先在对待匹配图像提取SIFT特征之后，通过KNN算法进行特征匹配，分别得到同一张图片的自匹配点对和不同图像间的初始互匹配点对；然后对初始互匹配点对进行相关性计算得到最正确的匹配点对，并根据最正确的匹配点对与自匹配点对的位置关系确定更多的正确匹配点对，最后计算仿射矩阵对图像进行拼接.实验结果显示经过DMIR算法获得的正确匹配点对更多，且分布更均匀，更准确，拼接的图像更好。

关键词：图像配准; SIFT; KNN; DMIR; 匹配; 图像拼接

论文编号：88

基于多视角图像的作物果实三维表型重建

杨会君 1，王瑞萍 1，王增莹 3，王昕 2

1.西北农林科技大学信息学院，陕西杨凌 712100；

2.西北农林科技大学外语系，陕西杨凌 7121002；

3.三只松鼠南京研发创新中心，江苏南京 210019

摘　要：针对基于设备获取点云存在操作复杂、成本高、难以被普及等问题，本文通过建立集 SFM 算法、PMVS 算法以及半自动化去噪方法优势为一体的三维重建架构，研究了基于普通图像的复杂背景中作物果实三维表型重建问题。以一组多视角目标作物果实二维图片为输入源，首先研究了基于 SIFT 算子的比例和旋转不变性参数提取多幅二维图像特征信息的方法，结合 FLANN 算法实现不同角度的数据匹配，提出了基于二维图片关键点和相机参数等信息的稀疏点云的快速生成方法。为精确表示果实、植株的表型特征，需要获取包含更多三维信息的稠密果实点云，因此，在稀疏获取作物果实点云的基础上进行稠密解析是三维重建过程的重要步骤。本文通过基于 PMVS 算法初始特征匹配的种子面片提取、扩散得到密集面片，进一步利用可见性约束来过滤不正确匹配导致的错误面片，实现了复杂点云模型的生成。最后，提出了交互式选择和滤波器相结合的果实表型点云离群点半自动化去除方法，解决了作物果实模型的准确性问题。对从不同角度拍摄的苹果、菠萝、榴莲、桃子等果实的图像序列进行处理，完成了复杂实验背景中作物果实的三维表型重建。结果表明，本文的方法能有效解决复杂背景中果实表型数据的低成本、准确、方便快捷获取问题。

关键词：三维果实表型；稀疏点云算法；稠密点云算法；去噪

论文编号：104

基于形变网络的草图人体建模

卢锦涛，陈雪锦

中国科学技术大学类脑智能及应用国家工程实验室，安徽省合肥市，邮编 230026

摘　要：基于草图的三维建模旨在从手绘线条图得到三维立体形状，对此本文重点探索基于单视角草图的人体建模。从单视角草图构建高质量三维人体模型的难点在于二维草图缺乏几何和纹理的细节信息，并且存在遮挡。此前的草图建模方法只能构造较为粗糙或者欠真实感的玩具模型。本文提出一种基于形变深度神经网络的手绘人体建模方法。该方法从单视角草图出发，以标准人体模板作为初始模型，提取草图和初始化人体模型的特征并串联成一个全局特征。全局特征指导初始模型上各顶点坐标形变到目标位置，得到形变后的人体模型。接着，采用SMPL参数回归网络获得更加自然和高质量的人体模型。本文构建了一个草图和模型配对的数据集，在此数据集上与最新的单视角图像重建方法相比本方案建模精度和可视化效果达到最佳。本文深入分析比较了使用单级形变解码器和多级形变解码器以及使用关节点损失函数监督对于本方案结果的影响，验证了当前方案的设置可以达到最优结果。总的来说，本方案可以根据简单的手绘草图获得到高质量人体模型，使得草图建模有很大应用价值，能够为普通用户设计不同姿态的人体模型提供便捷途径。

关键词：草图；人体；三维建模；深度学习

论文编号：151

一种基于无监督学习的多光谱图像配准方法

张鎏锟 1，陈书界 1，颜城城 1

1.浙江工商大学,浙江杭州 310000

摘　要：由于多光谱成像系统对不同波段的光学特性响应存在差异，不同波段下同一个物理空间点在相机平面的成像位置有所不同，宏观表示为不同波段图像间存在偏移、放缩、斜切等形变。因此在多光谱图像采集过程中，为得到高置信度的光谱信息，需要对多光谱图像做配准。传统多光谱图像配准方法一般采用迭代优化计算坐标变换模型的参数，时间成本较高。近年来，越来越多的研究工作表明用深度学习配准图像是一种高效、快速的方法。本文提出了一种基于无监督学习的多光谱图像配准方法。在空间变换网络的基础上，以归一化总梯度（Normalized Total Gradient）为两幅不同模态图像特征的距离损失，构建无监督的配准网络模型，输出最终的参数化坐标变换。本文方法的优势在于不需要真值标注，训练样本可通过给定变换参数随机生成，大大提高了网络模型的实用性。为进一步提高配准精度，本文在配准网络的基础上结合传统迭代优化方法，构建了一个基于无监督学习与迭代优化的混合配准框架。实验结果表明，本文提出的混合配准框架能快速配准两幅多光谱图像，其配准精度可达亚像素级别。

关键词：多光谱图像配准；无监督学习；混合配准框架；归一化总梯度；

论文编号：**243**

连通域先验的水下坝体裂缝图像拼接

黄渊博 1，2，张学武 1，2，徐志峰 3，黄建红 3，徐晓龙 1，2

1. 河海大学物联网工程学院,江苏省常州市 213022；

2. 江苏省输配电装备技术重点实验室,江苏省常州市 213022；

3. 南京瑞迪建设科技有限公司 Nanjing R&D Tech Group Co.,Ltd 210029;

摘　要：水下坝体裂缝是威胁大坝安全运行的重大隐患之一，根据国际坝工委员会的调查报告，坝体裂缝占溃坝原因的12.4%。对水下坝体裂缝的检测评估成为了刻不容缓的问题。为提高大坝安全运行评估的准确性，完整的裂缝图像是一项重要指标。针对传统方案在水下裂缝图像拼接的遇到的问题，提出连通域先验的水下坝体裂缝图像拼接方法。首先，通过实验得到适合水下大坝的图像增强处理方法；其次，探索裂缝图像特征，发现特征点提取与连通域直接的关系；最后，结合图像连通域先验的知识，将复杂的特征点筛选和匹配问题转化为简单的区域划分问题。真实场景的裂缝图像为研究对象，通过与经典的特征点匹配算法和传统的图像拼接算法进行对比实验，本文提出的特征点匹配算法使正确率得到了提升（57.3%和33.3%），且明显提升了拼接结果的质量。实验结果与已有方法比较，本文提出的图像拼接方法在牺牲了较少运行时间的情况下，极大的改善了拼接的质量，得到了高质量裂缝拼接图像，为大坝安全性能评估提供定性定量科学依据。

关键词：连通域；水下坝体；裂缝；图像拼接；

论文编号：308

基于偏好概率加权采样的鲁棒模型拟合方法

江斓 1，2 ，林舒源 1，2 ，郭翰林 1，2，严严 1，2，王菡子 1，2

1.厦门大学 信息学院，福建省厦门市 361005；

2.厦门大学 福建省智慧城市感知与计算重点实验室，福建省厦门市 361005

摘　要：鲁棒模型拟合是计算机视觉中的一个基础性的研究任务。在实际应用中，视觉数据会受噪声和离群点的污染，鲁棒模型拟合的任务是在包含噪声和离群点的数据中估计真实模型参数。基于数据聚类的鲁棒模型拟合方法通常将模型拟合问题转化为聚类问题，然而此类方法对于不平衡数据比较敏感。对此，本文提出了一种基于偏好概率加权采样的鲁棒模型拟合方法。首先，我们提出了一个模型假设更新策略，通过在更新过程中均衡评估模型假设质量来获取有效模型假设。然后，我们提出了一个基于偏好概率加权采样的策略，通过引入偏好概率的统计信息来有效地区分数据点中的内点与离群点，以此动态调整数据点的权重来指导后续的采样过程。该策略可以提高对小结构模型实例的采样精度，缓解数据不平衡的干扰。最后，我们通过结合这两种策略来生成模型假设集，并根据模型假设对应的残差集所构成的点与点的相似度矩阵，应用谱聚类技术对数据进行聚类，进而估计多结构模型实例。在两个模型拟合任务上的实验结果表明，本文所提方法与当前流行的一些模型拟合方法相比，在拟合精度与速度上都表现出有竞争力的结果。

关键词 ：鲁棒模型拟合；多结构数据；指导采样；谱聚类；

墙报 3-5

论文编号：378

基于双向加权特征金字塔网络的多人姿态估计

肖文福，张陶宁，陈恩庆，孙燚

郑州大学信息工程学院,河南郑州 450001

摘　要：人体姿态估计是计算机视觉领域中的一个重要研究方向，用来实现从图像或视频中检测人体若干关键点位置。多人姿态估计是当前人体姿态估计研究的主流趋势，主要应用于人机交互系统、视频监控、体育训练、辅助医疗等诸多方面。级联金字塔网络（Cascaded Pyramid Network，CPN）模型是一种基于卷积神经网络的多人姿态估计方法。其存在的问题是网络中的特征图尚未被充分有效利用。针对此问题，本文对 CPN 模型进行两方面的改进，首先是使用双向加权特征金字塔网络改进 GlobalNet，使特征金字塔中不同尺度的特征图按照不同比例进行融合，并且特征金字塔除了原有的从上到下的连接方式，又加入了从下到上的连接方式，从而使不同尺度的特征图被有效利用；其次是把 DenseNet 中的网络连接方式加入到 RefineNet，使特征图每经过一层卷积得到的新特征图与原输入特征图进行拼接，这样经过多个卷积层后特征图能够被多次重利用，增加特征的多样性，使网络模型能够学习到更多的信息。实验结果表明，本文提出的改进 CPN 模型运行时间比原 CPN 模型增加了 4.6%，模型精度比原 CPN 模型高 1.5 个 AP 值，比原 CPN 模型提高了 2.08%。改进 CPN 模型可以保证不显著增加模型运行时间的情况下有效提高模型的预测精度。

关键词：多人姿态估计；CPN；Bi-FPN；双向加权特征金字塔网络；DenseNet

论文编号：379

基于多监督学习的跨域草图检索算法

闫康明，周圆，李绰

天津大学, 天津 300072

摘　要：基于草图的跨域图像检索旨在根据输入的手绘草图与数据库中彩色图像的相似性度量返回彩色图像排序结果。该检索可以在基于文本的图像检索不准确或基于内容的图像检索中彩色图像难以获得的情况下，通过输入手绘草图完成图像检索任务。虽然相对于传统的图像检索算法，现有的草图检索算法效果得到了较好的提升。但是草图与彩色图像的特征映射与跨域检索等问题仍然没有得到较好的解决。为了更好的实现草图跨域检索任务，本文提出了一种基于混合网络与多监督学习的草图检索算法。本算法采用多子网络联合监督优化的方式最大化域内和域间特征的相关性，学习跨域特征间的公共表示。针对域内学习，本文提出在特征提取网络后加入分类子网络与重建学习子网络，以实现域内类别可区分性以及域内高层语义特征一致性；针对域间学习，本文提出在特征提取网络后加入跨域同分布学习子网络与三元组排序子网络，以保证草图与彩色图像在同一空间中的同分布映射以及域间同类特征的聚类。最后通过对模型分析并与其他方法对比实验验证，证明了提出的基于多监督学习的跨域草图检索模型的显著优越性。

关键词：草图；跨域检索；混合网络；多监督学习

论文编号：399

一种基于反距离加权融合的图像拼接算法

潘倩，杨爽，任洋甫，杨培，刘晓静

青海大学 计算机技术与应用系,青海 西宁 810016

摘　要： 图像拼接作为一种重要的图像处理技术被广泛运用于图像、视频编辑、军事、遥感等众多领域。图像融合是图像拼接的重要步骤，其性能对拼接结果有重要影响。但包括均值图像融合、渐入渐出图像融合、三角函数加权图像融合等典型的加权融合算法存在拼接痕迹明显的问题，本文针对这一问题提出了一种基于反距离加权融合的图像拼接算法。该算法在权重的计算过程中，基于所有有效边界进行权重建模，并使用重叠区每个像素点到边界距离的倒数构建该像素点的权重。同时针对边界中存在权重不连续点的问题，提出了基于不连续点的重叠区域划分及边界更新方法。为验证本文算法性能，在包括曝光差异、旋转、边界权重不连续等四种典型拼接场景对本文算法及对照算法进行试验分析，并使用定性和定量两种方式对实验结果进行评价。结果表明，本文算法相较于其他三种对比方法可有效地消除融合区域边界上的拼接痕迹，且在边界及融合区像素的像素值过渡更平缓。此外，本文算法在四组测试场景上的PSNR和SSIM值均优于渐入渐出和三角函数两种加权融合算法。融合结果的视觉评价及基于PSNR和SSIM的定量评价结果均表明基于反距离加权融合的拼接算法具有良好的拼接效果，验证了本文算法的有效性。

关键词：图像拼接；图像融合；均值图像融合；渐入渐出融合；反距离加权

论文编号：374

基于眼动跟踪的人机交互应用研究

黄君浩 1，贺辉 1

1.北京师范大学珠海校区,珠海 519087

摘　要：传统的人机交互方式存在诸多局限，已经不足以完成人们的需求。作为人机交互研究领域的一个分支，眼动交互有着广泛的应用前景。而目前借助专用眼动识别设备的交互应用系统普遍搭建复杂价格昂贵，无法得到普及。本文旨在使用价格低廉的网络摄像头作为视频源，研究基于眼动识别和视线跟踪在计算机上的交互办法，使简便且低成本的眼动交互有广泛应用的可能。本文在前期眼动行为识别研究的基础上，提出一种基于外观人眼视线方向估计的人机交互方法：使用一个简单快速的卷积神经网络模型粗略估计人眼在屏幕上的注视点，进而将眼动识别和不精确视线跟踪结果用于实现在计算机上的精确人机交互。在此基础上开发了眼动鼠标与眼动打字的功能以替代鼠标和键盘硬件。经过测试，熟练的用户使用本文的眼动鼠标模拟点击0.5cm×0.5cm目标的准确率在90%左右，更大目标物体的准确率则更高，使用本文的眼动打字长时间输入字符的准确率在84%左右。精确点击一次目标物体和打出单个字符的平均时间都为8秒。实验结果表明，本文的方法让使用者仅用一个普通的单目摄像头，就可以实现眼动跟踪，并利用眼动控制完成计算机上大部分精确的交互。

关键词：视线追踪；人机交互；眼动鼠标；眼动打字

论文编号：115

基于改进L-K光流的WebAR信息可视分析方法

裴云强[1]，吴亚东[2]，王赋攀[1]，张晓蓉[1]，蒋宏宇[3]，许世健[1]，

唐文生[1]

1.西南科技大学,四川 绵阳 621010； 2.四川轻化工大学,四川 自贡 643002；

3. 法政大学,计算机信息科学学院,日本 东京 184-8584

摘　要：信息可视化结合MAR（Mobile Augmented Reality，移动端增强现实技术）技术在目标跟踪领域存在着两个主要的困境。其一，大多数MAR应用虽有良好的模型渲染效果但计算过于密集，且对于移动设备的计算能力而言负载过大，目标跟踪时所渲染的三维模型在画面中容易出现剧烈抖动、卡顿或运动滞后于目标物的现象；其二，在信息可视化结合MAR的应用研发中，若仍坚持采用同跟踪平面图像的方案跟踪现实场景中的立体对象，则需要获取立体对象多角度的参照图像，跟踪过程计算压力加重，故难以达到相同或较好的效果。针对上述问题，该工作基于改进的L-K (Lucas-Kanade) 光流算法将特征点的跟踪问题转化为光流估计问题。在此基础上，本文提出了一种基于改进的L-K算法的WebAR（基于Web的移动增强现实技术）解决方案和一种优化的三维信息可视化交互策略。实验结果表明所提出的方法能够有效提高MAR在跟踪目标时的计算效率和跟踪稳定性，丰富信息可视化的呈现效果和交互方式。

关键词：移动增强现实；WebAR；信息可视化；Lucas-Kanade 算法；光流跟踪

论文编号：178

基于多任务学习的初始图像对选取方法

刘宇翔 1，张茂军 2，颜深 3，彭杨 4，李京蓓 4

1.2.3.4 国防科技大学,湖南省长沙市 410073

摘　要：近年来随着无人机与高清相机的广泛应用，基于图像的大规模三维场景重建技术得到了广泛关注，其主要运用多视图几何原理，通过不同视点所拍摄的图像计算出相机姿态和场景三维结构。其中，增量式运动恢复结构（SfM）是目前最普遍、最稳定的无序图像三维重建方法。而初始图像对选取是增量式的运动恢复结构（SfM）的一个关键环节，若初始图像对选取不佳，可能会直接导致重建失败。传统的基于几何关系的初始图像对选取方法存在计算效率低、对特殊场景不鲁棒的问题。因此，本文提出基于多任务学习的初始图像对选取网络以提高该过程的效率，网络中一个子任务学习估计相机位姿，另一个子任务学习图像全局匹配。通过设计交叉连接网络和联合损失函数，将两个子任务有效结合。此外，本文针对一些特殊场景提出结合场景连接图的选取策略，进一步提高重建的鲁棒性。通过与传统初始图像对选取方法进行实验对比，本文提出的方法不仅速度大幅提升，而且能够注册最多的图像，生成最多的点云，重投影误差最小，实验结果证明了所提方法的有效性。

关键词：增量式 SfM；初始图像对选取；多任务学习；场景连接图

论文编号：287

基于点云增强的室内含弱支撑平面物体的三维重建方法

彭杨，汪光跃，张茂军

国防科技大学系统工程学院,湖南长沙 410073

摘　要：虽然当前基于图像的三维重建已经取得了较大的成功，但是有一类物体的重建效果仍然不理想，这类物体在室内环境下非常常见，主要包括镜面物体和弱纹理物体。在传统的三维重建框架中，该类物体表面在稠密点云重建过程后会出现点云缺失的情况，造成没有足够的点云支撑进行表面重建，所以称之为弱支撑表面。本文主要针对含此类物体表面中平面形状的场景（即“弱支撑平面”）的三维重建问题进行研究，提出在传统三维重建框架中加入点云增强算法辅助弱支撑表面的正确重建，通过利用弱支撑平面在二维图像中的平面分割信息作为先验，结合相机映射模型来约束三维点云增强，从而实现弱支撑平面的强支撑化，改善该类平面的三维重建结果。为了验证本文方法确实优于传统三维重建方法，本文在DTU标准数据集与自行采集的数据上进行实验对比，结果表明本文方法能够准确完整地重建出含弱支撑平面物体的三维模型。

关键词：弱支撑平面；三维重建；多视图重建；分割掩模；点云增强

论文编号：108

基于 GA-SURF 的多通道图像拼接应用研究

李燕苹 1，石毅杰 2

1. 上海海事大学信息工程学院，上海，201306；2.上海大学通信与信息工程学院，上海，200444

摘　要：现有多光谱图像拼接应用大多基于传统矢量代数，先将图像转换成灰度图像，再利用灰度图像进行特征提取、配准，如SIFT、SURF等，不能很好地利用多光谱图像中不同图层间信息的关联性。几何代数（GA）具有多维统一与坐标无关特性，可形成多维统一、不依赖于坐标的空间物体几何表示与计算模型；可从多个角度呈现光谱层之间的联系特征，为解决上述问题提供了一个新的方向。本文将几何代数特征提取（GA-SURF）算法引入多光谱图像拼接应用，充分利用图像多光谱信息，进一步提升多通道图像的拼接效果。首先将多通道图像表示成几何代数多矢量结构，利用GA-SURF算法对待拼接的图像进行图像特征提取，基于RANSAC算法计算出待拼接图像之间的仿射变换矩阵，最后通过改进的渐入渐出算法完成多通道图像的拼接。针对不同拍摄角度、不同亮度的多通道图像拼接进行实验分析，通过特征匹配数量、RANSAC算法去误后的特征匹配数量、准确率、消耗的时间、拼接后的图像等方面的对比，验证了GA-SURF的有效性与合理性，在保持高质量图像拼接效果的同时有较快的运算速度。

关键词：图像拼接；特征提取；GA-SURF；渐入渐出算法；几何代数

论文编号：409

前视钻孔图像的全景图展开与拼接

杨彦从 1，周 洲 2，赵磊磊 2，彭瑞东 2

1.中国矿业大学（北京）中国矿业大学（北京）机电与信息工程学院，北京 100083

2.中国矿业大学（北京）中国矿业大学（北京）力学与建筑工程学院，北京 100083

摘　要：研究实现钻孔窥视图像的孔壁全景图展开，对于分析研究钻孔壁特征有着重要意义。通过分析前视钻孔视频图像的成像原理和特点，给出了生成钻孔壁全景图的处理方法和流程，并对其中一些关键技术问题进行了分析和讨论。首先通过对感兴趣区域的所有前视钻孔图像累加求出平均值图像，对其采用 Hough 变换确定圆心后，利用对称性和坐标映射关系，根据1/8区域的计算结果生成位置映射矩阵，实现原始环形图与展开条带图的像素对应，另外根据该平均图像的展开图确定径向灰度校正公式。然后逐一导入各帧图像，采用 Gamma 变换和双侧滤波的方法对图像进行预处理，以便在平滑降噪的同时尽可能保持细节，之后借助位置映射矩阵将图像环状区域展开为条带图像，并根据径向灰度变化规律对条带图像进行校正。最后对相邻图像的行、列灰度投影曲线分别进行 Pearson 统计相关分析，确定图像配准时的行重叠区域和列偏移角度，实现相邻图像的配准，并采用加权累加的方法实现相邻图像的融合，从而拼接得到平滑过渡的全景展开图。采用上述方法对实际拍摄得到的钻孔视频图像进行了展开和拼接测试。在得到的钻孔壁全景图中，裂缝位置对接准确，图像亮度均匀，完整体现了钻孔壁的岩性特点和裂缝分布情况。测试图像的处理结果表明，该方法稳定可靠，实现了钻孔壁图像的快速展开与精确拼接，为钻孔图像的进一步分类识别与定量计算奠定了基础。

关键词：钻孔成像；全景图；图像展开；图像配准；图像融合；双侧滤波；Hough 变换；灰度投影法

墙报 3-6

论文编号：128

遥感影像变化检测综述

任秋如 1，杨文忠 1，汪传建 2

1.新疆大学信息科学与工程学院,新疆省乌鲁木齐市 830000； 2.石河子大学信息科学与技术学院,新疆省石河子市 832000

摘　要：在科学技术迅速发展的今天，遥感影像变化检测越来越被人们所重视，变化检测技术也越发的成熟，被广泛的应用在灾害监测、城市规划、土地利用/土地覆盖等领域。然而在变化检测流程中仍存在一些问题有待解决，因此现在尚未提出一种综合性强、适用性广的变化检测方法，现有的变化检测效果距理想的自动化的、高效的变化检测也还有一定的差距。为了能对变化检测领域有一个较为清楚的认识，更加方便变化检测方法的研究，本文通过对大量有关变化检测的研究进行分析、比较、整理等工作，对变化检测进行了较为全面的综述。首先阐述了变化检测的发展历程；然后从数据选择及预处理、变化检测技术、后处理及精度评价三个方面详细归纳了变化检测的研究进展；最后对变化检测各个阶段存在的问题进行了总结并提出了未来的发展方向。

关键词：遥感；变化检测；发展现状；检测方法；综述

论文编号：143

面向多源遥感图像的自适应目标检测

张天坤 1，李汶原 1，史振威 1

北京航空航天大学宇航学院,北京 102206

摘　要：近年来，目标检测已经在含有大量标注的数据上展现出了良好的效果，但当真实测试数据与标注数据存在域间差异时，往往会导致训练好的目标检测模型性能降低。在遥感图像领域，随着我国卫星不断发射，各种卫星型号的遥感图像数据越来越多。传统的方法需要将多源图像数据重新标注，这将消耗大量人力和时间，因此需要研究域自适应的目标检测算法。由于相比于自然图像，多源遥感图像在成像方式和分辨率等方面存在特有的差异，因此在遥感图像上实现自适应目标检测面临特有的挑战。针对以上问题，本文提出了一种面向多源遥感图像的自适应目标检测算法，在图像级别和语义级别上对网络进行对抗训练，挖掘多源遥感图像在两个级别上的公共信息。此外，通过结合超分辨网络，进一步缩小了图像级别的差异，实现了自适应目标检测。本文在两个多源遥感数据集上进行实验，结果表明本文方法有效提升了目标域上的检测效果。

关键词：目标检测；域自适应；多源遥感图像；超分辨

论文编号：181

属性约束的人脸素描-照片转换与识别

黄婕 1，赵启军 1,2，吕泽均 1

1.四川大学计算机学院,成都市 610000; 2.西藏大学信息科学技术学院,拉萨市 850000

摘　要：目击证人描述和嫌疑人画像是侦破刑事案件的重要依据,然而尽管人脸识别技术近年来发展迅速，但是根据嫌疑人画像自动识别嫌疑人身份仍然极具挑战。大部分现有素描人脸识别方法或者在特征空间中减小嫌疑人画像与照片之间的模态差异，或者先在嫌疑人图像和照片之间进行转换再进行人脸比对，而都忽略了目击证人对嫌疑人的属性描述。本文提出利用目击证词中对人脸特点的描述作为人脸属性引入素描与照片间的转换过程，在照片域和素描域中同时进行人脸比对，并融合两者结果以识别嫌疑人身份。本文给出了基于条件和环形生成对抗网络（CycleGAN）的实现方案，选择了21种人脸属性作为附加信息与人脸图像一起输入生成器中，得到相应的生成图像。本文在香港中文大学的素描人脸库（CUFS）上进行实验，从生成图像质量和人脸识别准确率两个角度进行评估。在生成图像质量评估实验中，由于本文方法的主要目标是提升素描图像人脸的识别准确率，本文方法与现有方法相比，并不会显著提升图像质量。在人脸识别准确率实验中，本文采用了Dlib中的人脸识别模块，分别在照片域和素描域中的识别率都比现有方法有明显的提升，再融合两个域的识别结果后，识别率又有进一步提升，这证明了本文方法在素描人脸识别方面的优势以及融合不同域识别结果的必要性。

关键词：素描人脸识别；素描-照片转换；生成对抗网络；风格迁移

论文编号：198

基于稀疏量化电平和线性规划的PET探测器时间标记

邓贞宙1，邓玉姗1，曹香珠1

1.南昌大学PET实验室,江西省南昌市 330031

摘　要：在PET探测系统中，常用时间标记方法标记闪烁脉冲的阈值到达时间，并通过计算符合脉冲阈值到达时间的差值，求出系统的符合时间分辨率（CTR），而传统的时间标记方法精度较低，计算出的符合时间分辨率往往较差。稀疏量化电平（SQL）数字化方法是一种低功耗的采样方法，该方法利用一些参考电压可调的比较器确定闪烁脉冲的阈值到达时间,在此基础上,可以利用各种复杂的算法提高系统的时间分辨率。在先前实现的SQL数字化方法中，闪烁脉冲的阈值到达时间通过线性前沿假设（LLA）获得，该方法只涉及了闪烁脉冲上升沿的渡越时间，并将闪烁脉冲的上升沿视为线性处理。为了得到更好的符合时间分辨率，本文提出在SQL采样样本上升沿和下降沿的基础上使用线性规划（LP）的方法对组合参数进行优化。在研究过程中，为了评估LP/SQL与传统时间标记方法的性能，使用LaBr3：Ce晶体和滨松光电倍增管R9800建立两个伽马射线探测器，并利用Tektronics DPO 71604数字存储示波器读出闪烁脉冲。当两个探测器输出的闪烁脉冲都使用129个阈值时，SQL/LP获得了159.8438 ps的CTR。实验结果表明，SQL/LP对提高系统的符合时间分辨率具有潜在的优越性。同时，证明了以LLA为基础标记闪烁脉冲阈值到达时间的方法在数据分析中是不合理的。

关键词：正电子发射断层成像；符合时间分辨率；稀疏量化电平；线性前沿假设

论文编号：204

一种适形 X 线透视成像系统校正方法

李阳 1，牟轩沁 1

西安交通大学,陕西省西安市 710049

摘　要：相比于传统X射线源的成像模式,平板射线源的小体积和可寻址特性使它可以通过联合多个点源，使用小锥束完成对感兴趣区域的极低剂量成像。该成像模式利用了基于投影重排的图像恢复计算方法，需要对投影图像做进一步的处理才能得到最终我们需要的图像。新的成像方案需要根据成像物体的大小和形状将阵列射线源和平板检测器临时放置在适当的位置。在放置设备期间，几何误差的出现无法避免。当图像恢复中使用的系统几何参数与实际参数不同时，最终恢复的图像将具有几何畸变。多点源联合成像系统有它独特的几何关系，利用这种几何关系，本文提出了一种应用于新成像方案的几何校正方法。在此方法中，使用单点模体来估计新成像系统中的角度参数。在特殊的坐标系下，它仅需要对模体进行一次投影，就可以推导出关于三个角度误差的解析公式，从而计算出角度误差的大小，可以避免陷入局部最小值。本文设计了三组仿真实验，分别验证了提出的校正方法的可行性，以及探究了模体摆放位置偏差和探测器读数精度对最终校正结果的影响。在不同的参数条件下，校正方法的仿真结果的误差都较小，在可接受范围之内，证明了该方法的有效性和准确性。当模体摆放位置偏差较小时，最终结果误差很小，这一点提高了该方法在实际应用中的可操作性。文章给出的仿真结果证明了我们提出的校正方法可以用于适形X线透视成像系统角度误差的几何校正。

关键词：几何校正；透射成像；未对准；适形成像

论文编号：220

基于组合曝光的计算光学成像方式及其复原算法

吴笑天 1，2，刘博 2，吕博 2，杨航 2

长春理工大学光电工程学院，吉林省长春市，130022；

中国科学院长春光学精密机械与物理研究所，吉林省长春市，130022；

摘　要：计算光学是近年来新兴的交叉学科，强调成像链路环节上的协同配合及联合设计。本文提出一种基于组合曝光图像的计算光学成像方式，旨在解决低照度条件下高速运动物体的清晰成像。传统光学成像方式在低照度条件下拍摄高速运动物体时，通过积分时间的配置难以平衡能量获取和高速运动模糊之间的矛盾。 对此本文提出一种基于组合曝光的计算光学成像方式，通过高帧频相机采集临近两帧形成组合曝光图像对，第一帧图像为短积分时间，信噪比较低但强边缘可清晰成像，第二帧图像为长积分时间，信噪比较高但高速运动目标形成明显图像模糊。通过两帧之间信息的互补，估算运动模糊点扩散函数，并最终获取高信噪比的复原图像。试验结果表明，本论文所述的计算成像方法能够良好的解决传统成像在低照度条件下高速运动成像的模糊问题，通过复原算法获取的图像清晰可靠，具有良好的主观视觉效果。

关键词：计算光学；图像复原；组合曝光；非均匀点扩散函数

论文编号：351

结合自适应超像素的极化 SAR 舰船检测

李铭典，崔兴超，陈思伟

国防科技大学电子科学学院电子信息系统复杂电磁环境效应国家重点实验室,湖南长沙 410073.

摘　要：舰船检测是极化合成孔径雷达的重要应用，获得持续关注。当前，近岸密集舰船目标的稳健检测仍面临着挑战。恒虚警检测器（CFAR）易受近邻目标影响，导致检测性能下降。针对这一问题，本文提出一种结合自适应超像素的恒虚警检测(Superpixel-CFAR,SP-CFAR)方法，优化了传统CFAR检测器滑窗设置的拓扑结构。首先，利用超像素的均值和方差特性进行非监督聚类，提前确定可能含有目标及受目标影响严重的超像素。其次，在进行恒虚警检测时，通过迭代方式自适应的选取背景超像素用于门限求解，并在此过程中排除前期标记的超像素。最后，进行门限判决得到二值化检测结果。对于舰船密集的场景，通过这种方式选取到的背景超像素将不再局限于待检测超像素周围，从而有效地避免门限的过高估计。结合Radarsat-2实测数据的对比实验表明，相比于传统CFAR检测器和结合超像素的CFAR检测器，本方法具有明显优势，特别是在舰船密集区域能够减少漏检，品质因数(FoM)指标能够达到90%以上。

关键词：极化合成孔径雷达;舰船检测;恒虚警检测;超像素

论文编号：394

基于 CNN 和 VLAD 的指静脉描述子提取方法

文东霞 1,2，胡永健 1,2，王宇飞 2，刘琲贝 1,2，陈光 1,3

1. 华南理工大学，广东 广州 510640； 2. 中新国际联合研究院，广东 广州 510700； 3. 广电运通金融电子股份有限公司，广东 广州 510663

摘　要：为了提高指静脉描述子的鲁棒性同时降低网络参数量，通过修改VGGFace-Net并引入局部聚合描述子向量（Vector of Locally Aggregated Descriptors，VLAD）得到参数量仅0.3M的指静脉图像描述子提取网络，VLAD编码过程实现了局部描述子的聚类和重组，使描述子对手指姿势变化更加鲁棒；由于公开的指静脉训练数据库规模通常不够大，提出基于三元组和难分负样本挖掘策略进行网络的训练，并针对三元组损失没有约束样本对距离的类内方差的问题，提出一种样本对中心约束损失函数用于训练指静脉描述子提取网络，通过将正负样本对视为两个类别，分别约束正、负样本对靠近其类中心，这样更容易找到一个合适的二分类阈值来划分正负样本对。在三个公开数据库FV-USM，SDUMLA，MMCBNU上的指静脉验证结果表明，所提取的描述子在基于欧氏距离进行匹配的情况下，指静脉验证的结果均优于现有方法，且在图像发生随机平移的情况下具有更好的鲁棒性。

关键词：指静脉识别；卷积神经网络；局部聚合描述子向量；三元组；损失函数

论文编号：299

基于循环一致性对抗网络的动漫线稿图像提取

王素琴，张加其，石敏，赵银君

华北电力大学 控制与计算机工程学院，北京 102206

摘　要：动漫制作中线稿绘制与上色耗时费力，为此很多研究致力于动漫制作过程自动化。目前基于数据驱动的自动化研究工作快速发展，但并没有一个公开的线稿数据集可供使用。针对真实线稿图像数据获取困难，以及现有线稿提取方法效果失真问题，提出基于循环生成对抗网络的线稿图像自动提取模型。该模型基于循环生成对抗网络结构，采用非对称数据训练网络参数，以解决动漫彩色图像与对应真实线稿数据获取困难。为了进一步提升网络线稿提取效果，提出掩码指导卷积单元与掩码指导残差单元。其通过将不同尺度的输入图像与对应边界图低频信息输入对应网络单元层，以自主选择网络中间输出特征，确保网络关注于目标图像线稿区域特征信息。同时提出的边界一致性约束损失函数，通过约束生成图像与输入图像对应梯度图上差异，并添加增强边界区域的正则项，最终实现放大边界区域、平滑非边界区域，进一步提升动漫彩色图像线稿提取能力。最终在公开动漫彩色图像数据集Danbooru2018上，应用本文模型提取出相比于现有线稿提取方法，噪声更少，线条更清晰且接近真实漫画家绘制的线稿图像。此外该模型不需要大量真实线稿图像训练数据，实验中仅采集1000张左右真实线稿图像。模型结构不仅为后续动漫绘制与上色研究提供数据支持，同时也为图像边缘提取方法提供了新的解决方案。

关键词：动漫线稿图像生成；非对称数据训练；掩码指导卷积单元；循环生成对抗网络；卷积神经网

论文编号：237

VW-SAE：一种改进的光谱数据特征表示方法

胡晓勇，王海荣，刘午杨

北方民族大学，宁夏回族自治区银川市，750000

摘　要：针对高光谱图像的维度高、目标特征提取不准确的问题，本文提出了一种可变加权的堆叠式自编码器（Variable-Wise Weighted Stacked Autoencoder，VW-SAE）的光谱数据特征表示方法。VW-SAE方法在堆叠式自编码器（SAE）的基础上，通过对输出变量的相关性分析，从每个AE的输入层中识别出重要的变量，然后将输出信息在AE的目标函数中对不同的变量引入不同的权值而纳入到训练中，逐层提取获得与输出相关的特征，并将其堆叠形成深网络。通过对每层网络权重的调控，在降低光谱数据维度的过程中，同时更好地提取了光谱数据中的特征信息，进而提高了预测模型的精度。为验证方法的有效性，使用已采集的10248多张水稻图像，在堆叠式自编码器结合全连接神将网络（SAE-FNN）的基础上，搭建了VW-SAE-FNN模型对水稻氮元素进行检测，实验结果表明该方法相对于SAE方法的准确率明显提升。

关键词：高光谱技术；可变加权；堆叠式自编码器；权重；氮元素

中国图象图形学学会
电话：010-82544661/82544676
邮箱：info@csig.org.cn
官网：www.csig.org.cn