

基于动态亮度与感受野扩展的阴影去除方法

吴博文 王圣亚

南京审计大学

DOI:10.12238/acair.v3i1.11931

[摘要] 为解决图像部分区域由于阴影覆盖而导致图像质量差及信息不足的问题,对基于动态亮度调整与感受野扩展的阴影去除方法进行了研究,提出了端到端深度网络设计策略。首先结合动态范围扩展和亮度调整技术,提升了阴影区域的清晰度和准确性;其次,通过感受野扩展技术,捕捉图像中的全局信息,解决局部信息不足和阴影边界伪影的问题;最后,通过输出投影获得去除阴影后的图像,并在ISTD+和SRD数据集上进行了验证,结果表明其相较于其他阴影去除方法有显著提升,保留了图像的关键细节和自然纹理。

[关键词] 阴影去除; 图像增强; 深度学习; 动态范围扩展; 感受野扩展

中图分类号: U285.17 **文献标识码:** A

Shadow removal method based on dynamic brightness and receptive field expansion

Bowen Wu Shengya Wang

Nanjing Audit University

[Abstract] In order to solve the problem of poor image quality and insufficient information caused by shadow coverage in some areas of the image, a shadow removal method based on dynamic brightness adjustment and receptive field expansion was studied, and an end-to-end deep network design strategy was proposed. Firstly, by combining dynamic range expansion and brightness adjustment techniques, the clarity and accuracy of shadow areas have been improved; Secondly, by using receptive field extension technology, global information in the image can be captured to solve the problems of insufficient local information and shadow boundary artifacts; Finally, the shadow removed image was obtained through output projection and validated on the ISTD+and SRD datasets. The results showed significant improvements compared to other shadow removal methods, preserving key details and natural textures of the image.

[Key words] Shadow removal; Image enhancement; Deep learning; Dynamic range expansion; Experience Field Expansion

引言

阴影是光线被物体遮挡或反射引起的区域,导致该区域相对较暗或缺乏细节,常对图像质量和信息获取产生负面影响,限制人类的感知,并影响目标检测^[1]、目标跟踪和识别^[2]等视觉任务。因此,阴影去除的目标是消除或减轻因不均匀光照或物体遮挡导致的阴影,以改善图像质量和可用性。

现有研究提出了多种阴影去除方法,主要分为基于成对阴影与无阴影图像的物理模型方法和基于不成对数据的深度学习。物理模型通过分析阴影与无阴影图像之间的差异,学习阴影特征。Le等^[3]提出通过阴影图像分解和纹理填充修复阴影区域,恢复原始光照。Wang等人^[4]则提出了一种自适应多尺度光照传递的方法,有效应对复杂的阴影场景。尽管物理模型在理解阴影生成方面表现出一定优势,但在复杂场景中的泛化能力仍然

有限。

相较之下,不成对的深度学习方法通过深度神经网络学习阴影去除的映射关系,无需成对图像。基于Transformer的ShadowFormer网络^[5]将阴影退化建模为阴影模型,通过通道注意力增强多尺度编码器-解码器管道,在瓶颈阶段融合全局与局部信息,以更好地恢复阴影区域。然而,该方法在阴影边界仍存在伪影问题。

为解决阴影边界伪影和细节缺失问题,本文提出了一种基于动态亮度调整与感受野扩展的深度阴影去除方法。通过动态亮度调整,使阴影区域更加清晰,提升图像的准确性。感受野扩展则增强了对全局信息的捕捉,有效弥补了阴影区域局部信息不足的问题。该方法实现了端到端的阴影处理,显著减少了边界痕迹的伪影,提高了阴影去除的质量和准确性。实验结果表明,

所提方法在ISTD+和SRD数据集上的表现优于现有方法, 具有更高的有效性。

1 相关工作

在探索阴影去除研究的早期阶段, 研究人员通常依靠手工制作的先验信息来恢复图像中的阴影区域。这些先验信息包括图像光照特性、梯度等。例如, Liu等^[7]提出从阴影生成到阴影去除网络, 通过阴影生成器产生大量伪阴影数据, 并实施联合训练以增强阴影去除的效果。

图像增强通过调整亮度、对比度、色彩平衡、去噪和清晰度等, 改善图像的视觉质量, 旨在提高其在人眼观察或计算机处理中的可用性。尤其在目标检测和识别任务中, 图像增强^[6]通过提升图像质量和目标可见性, 帮助模型更精确地定位和识别目标。

Winograd变换是一种用于加速卷积运算的技术, 它通过改变卷积核和输入数据的表示方式, 减少了乘法运算的数量, 从而间接地提高感受野的效果。在一项研究中, Tong等^[8]提出了基于张量虚拟机和代码重构器实现高性能的Winograd卷积。在Winograd计算阶段中提出了等价变化, 以及利用Tensor Core进行计算的无同步实现, 并进一步提出了利用不同GPU内存层级的部分计算核融合方法。

2 方法

2.1 总体结构

动态亮度调整与感受野扩展的网络Dynamic Brightness Adjustment and Receptive Field Expansion Network (DBA-RFENet), 网络结构如图1所示。首先, 阴影图像IS通过Dynamic Brightness Adjustment (DBA) 模块, 该模块对输入的阴影图像进行动态亮度调整, 得到增强后的输出图像 I_0 。 I_0 与阴影掩模图像IM通过concatenation (cat) 操作进行融合, 融合后的特征经过输入投影和Dropout层, 用于防止过拟合。Dropout后的特征进入带有Receptive Field Expansion Block (RFEB) 块的编码器-解码器和上采样块。RFEB块用于扩大图像感受野, 上采样层用于恢复图像分辨率。特征经过上采样后, 再次进行cat操作, 融合多尺度特征, 以捕获图像的多层次信息。最后再通过输出投影将处理后的特征映射为最终的去阴影图像, 从而得到最后消除阴影的图像。其中IS代表输入的阴影图像, IM代表输入的掩模图像, I_0 代表经过动态亮度调整模块DBA增强后的输出图像。该网络通过DBA模块和带有RFEB块的编码器-解码器两大组件, 实现高效阴影去除。

考虑到阴影的颜色可能会受到外部光照影响而与周围物体接近, 为此, 我们设计了DBA模块。该模块结合对比度增强和动态范围扩展, 显著提高了阴影区域的清晰度和细节, 使得阴影模糊现象得到改善。为进一步增强感知阴影的能力, 我们引入了RFEB模块, 将Winograd可分离卷积技术应用于阴影去除任务中。该模块不仅提高了计算效率, 还加深了对图像信息的理解, 从而最终实现了更出色的阴影去除效果。

2.2 动态亮度调整

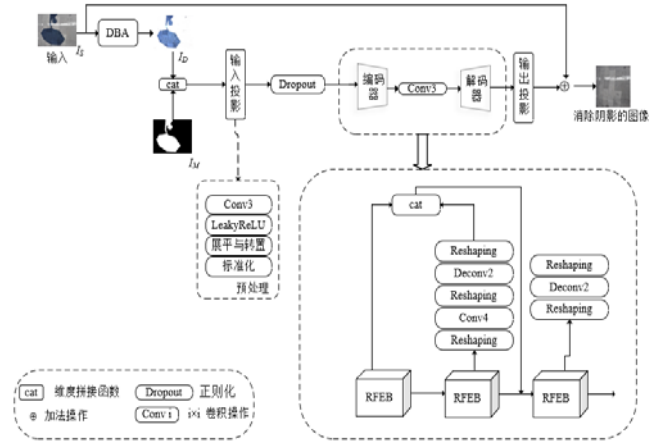


图1 网络总体结构图

通过使用DBA模块对输入图像进行对比度增强和动态范围扩展的操作, 通过动态拉伸阴影区域的亮度范围提高阴影检测。DBA的结构如图2所示。具体操作是首先对输入数据 $X \in \mathbb{R}^{B \times C \times H \times W}$ 应用对比度增强操作, 其中B表示批量大小, C代表通道数, H代表图像高度, W代表图像宽度。通过将图像的像素值乘以对比度增强因子 α , 并通过clamp函数将像素值限制在 $[0, 1]$ 范围内, 以防止像素值超出有效范围。得对比度增强后的输出 X_{ce} , 如式(1)所示:

$$X_{ce} = \text{clamp}(\alpha X, 0, 1) \quad (1)$$

随后对图像进行了动态范围扩展处理。首先计算输入数据的最小值 $X_{\min}$ 和最大值 $X_{\max}$, 通过将图像中的每个像素减去最小值并除以动态范围对每个像素的数值进行标准化, 从而实现了标准化。然后将归一化后的像素值进行动态范围扩展, 即将标准化后的像素值 X_{dre} 的每个元素提升为 δ 次方来扩展动态范围, 这里 δ 表示动态范围扩展因子。图中展示了 δ 取0.8, 1.0, 1.1, 1.2, 1.5, 2.0时图像像素值变化过程。并通过clamp函数将像素值限制在 $[0, 1]$ 范围内, 得动态范围扩展后的输出 X_{dre} , 如式(2)和(3)所示:

$$X_{dre} = \left(\frac{X - X_{\min}}{X_{\max} - X_{\min} + \beta} \right)^{\delta} \quad (2)$$

$$X_{dre} = \text{clamp}(X_{dre}, 0, 1) \quad (3)$$

其中 $\beta > 0$ 用于避免分母为零, 文中 β 取值为 1×10^{-7} 。

将两个处理后的图像相加, 得到融合的增强输出图像, 最后再次通过clamp函数将像素值限制在 $[0, 1]$ 范围内, 其计算方式如式(4)和(5)所示, 最终返回融合的增强输出图像 $X_{fused_enhanced}$ 。这样进一步提升了图像的对比度和动态范围, 从而改善了图像的视觉效果。

$$X_{out} = X_{ce} + X_{dre} \quad (4)$$

$$X_{fused_enhanced} = \text{clamp}(X_{out}, 0, 1) \quad (5)$$

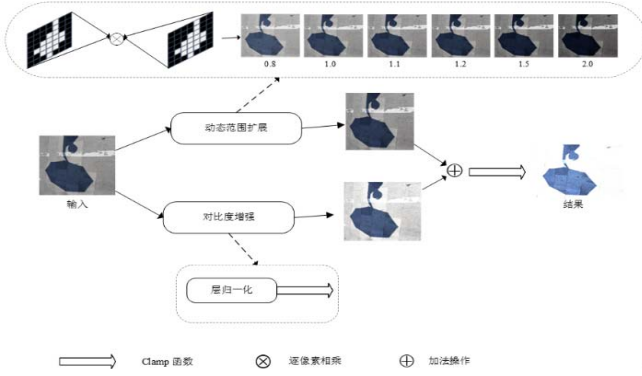


图2 DBA结构图

2.3 感受野扩展

通过使用Winograd^[8]技术的RFEB模块进一步增加感知阴影范围。由于Winograd可以将传统的卷积操作分解为一系列矩阵乘法,所以在RFEB模块中使用Winograd卷积代替传统卷积,使得计算更加高效。模块通过多层卷积和池化操作,进一步扩大感受野,从而有效地去除阴影。RFEB的结构如图3所示。对图像执行Winograd变换,其中U矩阵和V矩阵是用于进行Winograd变换的关键矩阵。U矩阵代表输入数据在Winograd域中的表示。V矩阵代表卷积核在Winograd域中的表示。G代表Winograd基础矩阵,A代表转换矩阵,G和A是固定的。

当卷积核为3时,G矩阵和A矩阵的样式如式(6)和(7)所示:

$$G = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{2}{9} & -\frac{2}{9} & -\frac{2}{9} \\ \frac{2}{9} & \frac{2}{9} & -\frac{2}{9} \\ \frac{1}{90} & \frac{1}{45} & \frac{2}{45} \\ \frac{1}{90} & -\frac{1}{45} & \frac{2}{45} \\ 0 & 0 & 1 \end{bmatrix} \quad (6)$$

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \quad (7)$$

当卷积核为5时,G矩阵和A矩阵的样式如式(8)和(9)所示:

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \frac{2}{25} & \frac{2}{25} & \frac{2}{25} & \frac{2}{25} \\ \frac{2}{25} & -\frac{2}{25} & \frac{2}{25} & -\frac{2}{25} \\ \frac{1}{25} & \frac{1}{50} & \frac{1}{100} & \frac{1}{200} \\ \frac{1}{25} & -\frac{1}{50} & \frac{1}{100} & -\frac{1}{200} \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (8)$$

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 2 & 4 & 8 \\ 1 & -2 & 4 & -8 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (9)$$

Winograd变换旨在通过变换输入和卷积核的表示来优化卷积操作,这一变换分为两个关键步骤:

(1) 计算U矩阵: 将经过形状变换后的输入图像 X_{perm} 与预先计算的Winograd基础矩阵G相乘,然后再与转换矩阵A的转置相乘,得U矩阵。如式(10)所示:

$$U = GA^T X_{perm} \quad (10)$$

(2) 计算V矩阵: 将卷积核的权重矩阵 W_{perm} 与Winograd基础矩阵G相乘,然后再次与Winograd基础矩阵的转置相乘,得到V矩阵。如式(11)所示:

$$V = GW_{perm}G^T \quad (11)$$

最后,通过计算U和V矩阵的乘积,得到经过Winograd变换的输出。使用Einsum函数将Winograd变换后的输入数据与卷积核的权重矩阵 W_{perm} 进行乘法运算,得到最终的卷积结果Y。如式(12)所示:

$$Y = \text{Einsum}(W_{perm}, U, V) \quad (12)$$

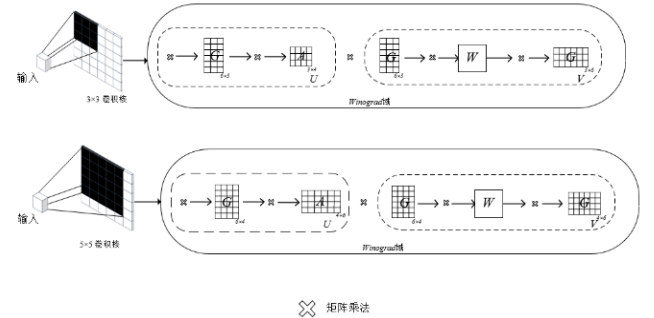


图3 RFEB结构图

3 实验

所提出的DBA-RFENet阴影去除模型是通过PyTorch实现。该模型使用AdamW优化器进行训练,动量设置为(0.9, 0.999),并应用权重衰减为0.02。从初始学习率 2×10^{-4} 开始,通过使用余弦退火降至 1×10^{-6} 。在实验中,网络经过了AdamW优化器的训练,完成了500多次迭代,输入图像的大小调整为 480×640 ,并且批量大小为16。在实验设置中,采用3倍尺度的编码器-解码器结构,确保最终输出维度与原始图像比例一致。

3.1 数据集及度量指标

在数据集ISTD+和SRD上进行训练和评估。ISTD+数据集包含1330个训练三元组和540个检测三元组,每个三元组由一个有阴影的图像、相应的掩码和一个无阴影的图像组成。SRD数据集则包括2680对训练图像和408对检测图像。

使用PSNR和SSIM这两个指标来度量不同方法在RGB颜色空间中的阴影去除性能。在这两个指标中,数值越高代表着阴影去

除效果越好。

3.2 实验对比分析

分别在ISTD+和SRD这三个通用的阴影去除数据集上进行实验,并与最新的几种阴影去除方法进行了比较,如表2、3和4所示。例如SRD数据集上,从表1数据中可以看出,所提方法Ours在所有评估指标上均显著优于其他对比方法。在阴影区域,该方法的PSNR为34.77,显著高于STC-GAN的32.60,这是第二高的结果。SSIM值也达到了0.983,超过了AEF的0.960。在非阴影区域,该方法的PSNR为33.25,高于DSC的31.94。同样,SSIM值为0.974,也高于DSC的0.965。在整体性能上,该方法的PSNR为29.84,高于DC-ShadowNet的28.76,SSIM值为0.950,优于所有其他方法。

图中的对比结果清晰地展示了所提出的方法在ISTD+和SRD数据集上的显著优势,如图4和5所示,其中GT代表地面真实图。例如在SRD上的可视化图4中可以看出,其他方法在处理阴影时,仍存在明显的阴影伪影。然而,该方法几乎完全消除了这些伪影,显著提升了图像的质量。该方法在阴影区域的细节保留和非阴影区域的亮度均衡方面表现优异,这得益于感受野扩展与动态亮度调整技术的有效结合。该方法不仅在定量指标上领先,在视觉效果上也优于其他阴影去除方法,验证了其有效性和广泛适用性。

表1 SRD数据集上与几种阴影去除方法的比较

方法	Shadow		Non-Shadow		ALL	
	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑
DSC ^[9]	30.65	0.960	31.94	0.965	27.76	0.930
AEF ^[16]	32.26	0.966	31.87	0.945	28.40	0.893
DC-ShadowNet ^[11]	32.00	0.966	33.53	0.945	28.76	0.893
STC-GAN ^[13]	32.60	0.960	25.01	0.758	23.95	0.708
SIM ^[6]	29.44	0.937	26.67	0.879	24.16	0.779
PR ^[21]	29.19	0.954	15.13	0.753	15.48	0.699
Ours	34.77	0.983	33.25	0.974	29.84	0.950

表2 ISTD+数据集上与几种阴影去除方法的比较

方法	Shadow		Non-Shadow		ALL	
	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑
SID ^[3]	35.08	0.984	36.38	0.979	31.89	0.953
DSC ^[9]	35.97	—	35.76	—	32.05	—
AEF ^[16]	36.04	—	31.16	—	29.45	—
DC-ShadowNet ^[10]	32.00	—	33.53	—	28.76	—
DHAN ^[10]	32.92	—	27.15	—	25.66	—
WSD ^[16]	33.09	0.983	35.26	0.977	30.12	0.950
Ours	39.40	0.991	38.75	0.985	35.38	0.974

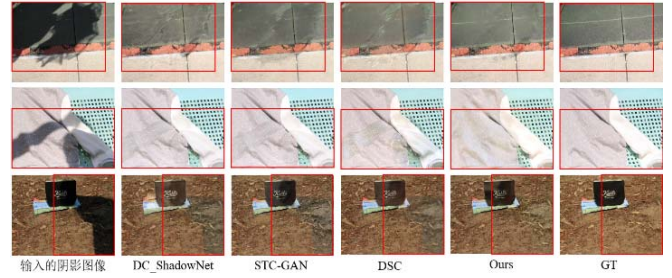


图4 SRD数据集上阴影去除结果的可视化图

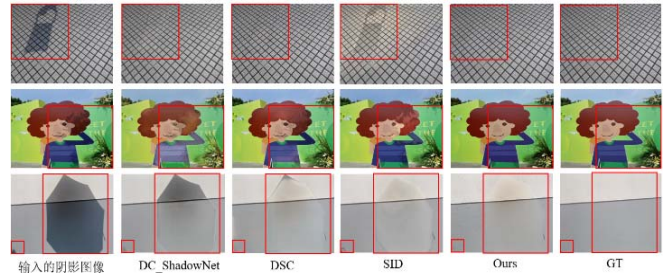


图5 ISTD+数据集上阴影去除结果的可视化图

3.3 消融实验

在对比分析中,实验数据如表3所示,可视化结果如图6所示。所提出的方法通过添加DBA和RFE模块,在阴影去除任务上展现了显著的性能提升。从数据上看,当单独添加DBA模块后,阴影区域的PSNR提高到37.69,SSIM提高到0.988,而非阴影区域的PSNR和SSIM也分别达到了34.00和0.980。类似地,单独添加RFE模块后,阴影区域的PSNR达到了37.98,SSIM为0.989,非阴影区域的PSNR和SSIM也分别为33.81和0.978。结合使用DBA和RFE模块后,阴影区域的PSNR提升至38.28,SSIM提升至0.990;非阴影区域的PSNR提升至34.64,SSIM提升至0.980;这些指标均明显高于单独使用任一模块的情况且远超其他阴影去除方法。例如,SPA-Former^[17]的PSNR为33.51,SSIM为0.982;清晰导向阴影网络CNSNet^[16]的PSNR和SSIM分别为36.67和0.980。观察消融实验的可视化图,可以看到单个模块相对于其他阴影去除方法能够有效去除阴影伪影,而将这些模块结合起来则进一步消除残余的阴影。

通过全面的消融实验,深入探讨了各个模块对模型整体性能贡献的贡献。结果显示,DBA模块在动态调整亮度以应对不均匀光照方面发挥了重要作用,而RFE模块通过扩大感受野有效捕捉了更多的上下文信息。当这两个模块结合在整个网络使用时,能够显著提升模型的综合性能。这种多模块协同工作的设计理念,证明了在复杂任务中采用多种方法相结合的重要性,为未来的阴影去除研究提供了有价值的参考。

4 结束语

本文提出了一种基于动态亮度与感受野扩展的深度阴影去除方法,采用端到端深度网络设计,显著提升了阴影区域的清晰度和准确性。通过添加动态范围扩展和亮度调整技术,增强了阴影区域的细节表现;以及利用感受野扩展技术,有效捕捉图像的全局信息,从而解决了局部信息不足和阴影边界伪影的问题。

表3 消融实验各组件对比数据

组件	Shadow		Non-Shadow		ALL	
	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑
CNSNet [16]	36.67	0.980	32.15	0.976	30.29	0.962
SPA-Former [17]	33.51	0.982	30.16	0.956	27.73	0.931
Added DBA	37.69	0.988	34.00	0.980	31.87	0.965
Added RFEB	37.98	0.989	33.81	0.978	31.81	0.964
Added DBA+RFEB	38.28	0.990	34.64	0.980	32.48	0.967

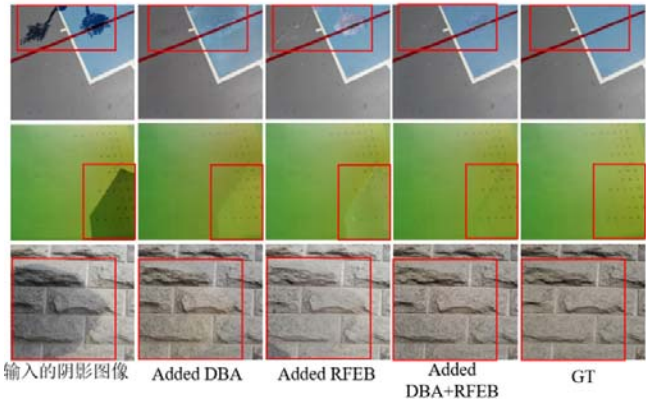


图6 消融实验可视化图

未来的研究可进一步探索雨雪天气条件下的阴影去除技术及其在医学影像中的应用。针对当前阴影边界伪影尚未完全解决的问题，建议优化动态调整算法，并结合更先进的生成对抗网络，推动阴影去除技术在各个领域的应用与发展。

[参考文献]

[1]徐梦溪,施建强,郑胜男,等.模拟复眼视叶神经网络的目标运动方向检测模型[J].智能系统学报,2024,19(3):546–555.

[2]张歆羽,杨钟亮,周哲画.面向多目标医疗垃圾分类的智能识别分拣系统设计[J].智能系统学报,2024,19(3):584–597.

[3]Le H,Samaras D.Shadow removal via shadow image decomposition[C].IEEE,2019,8578–8587.

[4]Wang P, Zhao X. Multi-scale adaptive light propagation for shadow reduction[C].IEEE/CVF,2020,5032–5039.

[5]Guo L,Huang S,Liu D,et al. ShadowFormer:Global Context Helps Image Shadow Removal[J].AAAI,2023.

[6]刘波,田广粮,肖斌,等.利用自适应光照初始化的弱光图像增强方法[J].电子与信息学报,2024,46(02):643–651.

[7]Liu Z,Yin H,Wu X,et al. From shadow generation to shadow removal[C].IEEE/CVF,2021,4927–4936.

[8]童敢,黄立波,吕雅帅.面向现代GPU的Winograd卷积加速研究[J].电子学报,2024,52(01):244–257.

[9]Hu X,Fu C,Zhu L,et al. Direction-Aware Spatial Context Features for Shadow Detection and Removal[J].IEEE,2020,42(11):2795–2808.

[10]Fu L,Zhou C,Guo Q,et al. Auto-exposure fusion for single image shadow removal[C].IEEE,2021,10571–10580.

[11]Jin Y,Sharma A,Tan R T.Dc-shadownet:Single-image hard and soft shadow removal using unsupervised domain-classifier guided network[C].ICCV,2021.

[12]Cun X,Pun C M,Shi C.Towards Ghost-free Shadow Removal via Dual Hierarchical Aggregation Network and Shadow Matti ng GAN[J].arXiv,2019,1911.08718.

[13]Zhang H, Xu Y, Zhu S, et al. STC-GAN: Shadow transfer and removal via a generative adversarial network[J].IEEE,2022,31:3844–3857.

[14]Zhu Y, Xiao Z, Fang Y, et al. Efficient model-driven network for shadow removal[C].AAAI,2022,36(3):3635–3643.

[15]Le H, Samaras D. From shadow segmentation to shadow removal[C].IEEE,2020.

[16]YU Qianhao,ZHENG Naishan, HUANG Jie, et al. CNSNet: A Cleaness-Navigated-Shadow Network for Shadow Removal[J]. European Conference on Computer Vision. Cham: Springer Nature Switzerland,2022,221–238.

[17]ZHANG Xiaofeng,ZHAO Yudi,GU Chaochen,et al.SPA-Former:An Effective and lightweight Transformer for image shadow removal[J].In IJCNN,2023.

作者简介:

吴博文(1999--),女,汉族,吉林长春人,硕士研究生,研究方向:模式识别。