

基于多光谱融合的端到端计算成像

南亦澜, 申俊飞*, 张启灿

四川大学电子信息学院, 四川 成都 610065

摘要 传统提升低照度成像质量的方法主要包括外界补光和光圈提升,通过增大物理进光量完成源信息增强,但此类方法会造成光源污染和景深缩短等问题。提出一种端到端的多光谱融合方案,在弥补传统方法缺陷的同时有效恢复低照度场景中物体颜色和细节,实现高质量计算成像。通过定制化设计深度学习网络,融合多通道光谱信息,可有效消除场景噪声。所提方法具有较高的自由度,可以根据具体应用场景需求调节通道数和网络参数,同时也能替代传统的相机模块,优化图像信号处理流程。进行详细的消融实验,结果表明,光谱融合后,相比传统基于RGB数据的方法,所提方法得到的图像质量的均方误差(MSE)和感知损失分别降低了54.43%和35.12%。所提方法将为增强现实/虚拟现实(AR/VR)、医疗成像、自动驾驶等新兴技术领域带来新的高质量成像方案。

关键词 计算成像; 多光谱融合; 深度学习; 低照度

中图分类号 TP391 **文献标志码** A

DOI: 10.3788/LOP202259.2011010

End-to-End Computational Imaging Based on Multispectral Fusion

Nan Yilan, Shen Junfei*, Zhang Qican

College of Electronics and Information Engineering, Sichuan University, Chengdu 610065, Sichuan, China

Abstract Traditional methods for improving the quality of low-illumination imaging include external supplementary light and expanding the aperture, and complete source information enhancement by increasing the physical light input. However, these methods have been proposed to cause problems, such as light source pollution and shortening of the depth of field. Therefore, this paper proposes an end-to-end multispectral fusion scheme to achieve high-quality computational imaging, thereby effectively restoring the color and details of objects in low-illumination scenes while making up for the shortcomings of traditional methods. First, the multi-channel spectral information was integrated through a custom-designed deep learning network, after which scene noise was effectively eliminated. As a result, we discovered that the proposed method had a high degree of freedom, could adjust number of channels and network parameters according to the needs of specific conditions, replaced the traditional camera module, and optimized the image signal processing process. Next, we conducted detailed ablation experiments. The results show that after spectral fusion, the mean square error (MSE) and perceptual loss of image quality obtained by the proposed method reduce by 54.43% and 35.12%, respectively, compared with that obtained by traditional methods based on RGB data. Thus, the successful proposal and verification of this study's method should serve as a background for new high-quality imaging solutions within complex application scenarios, such as augmented reality/virtual reality (AR/VR), medical imaging, and autonomous driving.

Key words computational imaging; multispectral fusion; deep learning; low illuminance

1 引言

近年来,在图像处理技术高速发展的背景下,相机的成像优化研究取得了很大进展。然而,在低照度等极端拍摄条件下,大量不规则噪声、颜色失真、运动伪影等情况造成的成像质量下降仍是亟待解决的关键问

题。针对此类问题,提升成像质量的方法主要包括硬件上的改进和图像处理算法上的优化,但此类方案大多面临计算量复杂、设备体积大、在拍摄场景变化时自适应性不足等问题。因此,如何在有限成本下实现智能且精确复原场景中物体的纹理、颜色等细节逐渐成为现代相机成像技术的一项难题与重要挑战。

收稿日期: 2022-08-15; 修回日期: 2022-09-05; 录用日期: 2022-09-09

基金项目: 国家自然科学基金(62105227)、四川省科技计划项目(2022YFS0113)、成都市科技项目(2021-YF05-01990-SN)

通信作者: *shenjunfei@scu.edu.cn

在低照度环境下,由于成像设备接收的光子数少,信噪比低,最终会发生成像质量退化的问题,大量的噪声会淹没场景的细节和颜色信息。现有的解决方法大致分为两个方向:物理成像改进和图像后处理。物理成像改进方法通常利用外界辅助手段来提升成像质量,例如提高成像设备的感光灵敏度来增加输入信号增益,但物理成本较高;或是利用延长拍摄曝光时间、扩大光圈的方法来增加成像设备的进光量,但会面临成像时对象运动或相机抖动而产生的伪影、景深缩短等问题。图像后处理方法大多面向相机图像信号处理(ISP)输出进行后处理,从不同角度弥补低照度拍摄场景下带来的成像缺陷,主要分为图像还原和图像增强。对于图像增强,在基础图像质量尚可的情况下,完成结构化图像信息提取和分析,适应更高级别成像任务。例如:图像超分辨率重建^[1-2]根据图像退化程度来进行配准和细节重建;Chen等^[3]采用深度学习完成图像特征提取和融合,利用U-Net网络对预处理后的RAW数据使用端到端的神经网络,替换传统ISP中RAW数据到sRGB的处理过程;陈清江等^[4]提出一种基于并联残差网络的低照度图像增强算法,对交替残差模块与局部全局残差模块进行并联,有效提高了图像亮度和对比度,减少了边缘细节丢失。对于图像还原,设计图像算法,解决低照度下颜色和噪声影响带来的问题,从而提升成像质量。例如:基于Retinex^[5-8]颜色还原理论的方法可以通过去除场景光照还原物体的颜色;去噪^[9-11]、去模糊^[12]、去雾^[13-14]等算法通过消除噪声还原物体细节。

但目前无论是物理成像改进,还是图像后处理,成像效果严重依赖于输入源信息质量。很多研究通过改进源信息限制完成图像质量提升。例如:通过使用短曝光连续拍摄多帧图像获取稳定的噪声形态,实现对场景噪声的估计,进而对不同曝光下拍摄的图像进行融合^[15],这种方法在一定程度上缓解了常见多帧成像中运动伪影的问题,但多次成像会引入额外的时间与计算成本;华为公司在2019年提出了新的RYYB像素

排列方式^[16],通过拓宽光谱通道来增加感光量,提高了对光线的利用率,但对后续去马赛克操作和白平衡算法提出了更高要求;还有通过夜晚补光灯等外界光源来增加进光量的方法,这种方法可能会导致拍摄场景中局部区域具有高亮度、整体图像具有高亮度动态范围的问题,会引入额外的算法或额外的曝光时间。

对比而言,多光谱数据在低照度成像中能够同时满足以下要求:在同一个场景下的本底噪声的分布具有一定的规律,多个谱段里类似的噪声分布更有利于学习和降噪;不同颜色的物体在不同谱段的响应不同,多光谱数据能够提供更多的可分级信息以供参考,有助于实现物体真彩色的复原预计对场景光源的精确估计;多光谱数据中每个光谱通道的感光量不同,类似于静态的宽动态成像(HDR),能够实现对不同物体颜色和细节的增强与还原。因此本文希望运用多光谱数据,提供更多的可参考噪声形态,解决时域、空域噪声分布差异大的问题,从而更精细地估计噪声,更准确地表达颜色,最终实现对纹理和细节更好的表达。本文利用多光谱数据是希望通过提升通道光信息量结合端到端后处理算法设计,实现高质量计算成像,以增加源信息的方法替代常用的后处理手段,认为通道之间的互相关信息对成像分析也是极为重要的。通过设计端到端的神经网络,直接对原始的多光谱数据进行信息提取与重建,在低照度复杂场景下完成噪声消除和细节还原;通过设计一个变种U-Net框架网络^[17]来替代传统相机成像中的ISP流程进行光谱融合,有效缓解传统三通道积分插值方法导致的光谱信息丢失问题,提升了输入信噪比;同时利用光谱通道间信息的互相关作用,从大量噪声中还原图像细节,最终输出优化后的RGB图像。

2 相关理论

2.1 相机 RGB 感光模型

图1展示了常见的相机物理成像模型。光信号由一个或多个光源发出,经物体表面反射,透过相机镜头,投射到传感器上,光信号由感光二极管转换为电信号,

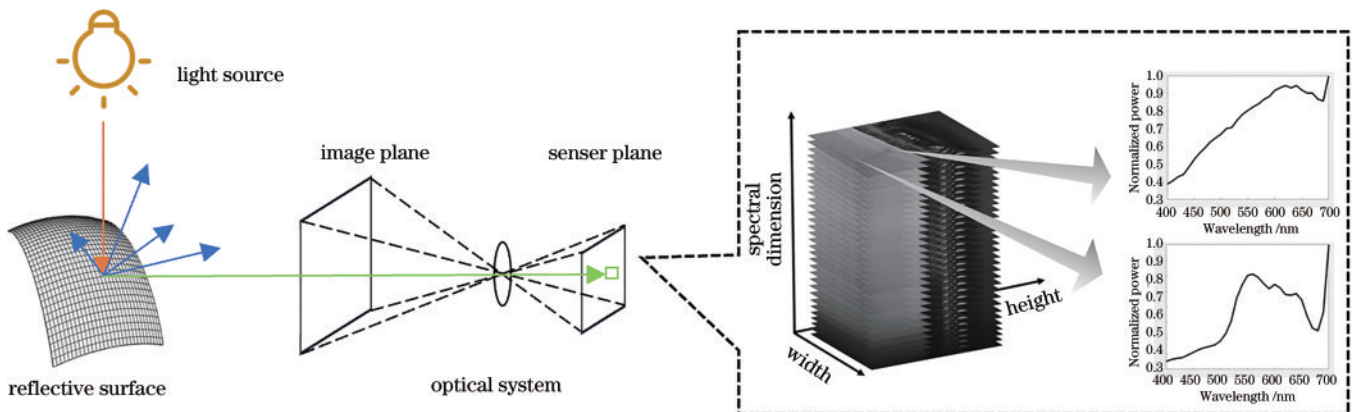


图1 相机成像示例

Fig. 1 Example of camera imaging

同时传感器上覆盖一层色彩滤波阵列用于记录规则分布的色彩信息。

传感器上每个感光二极管能够接收到一定光谱范围内的光辐射。图 1 右侧展示了两个像素点的光谱曲线,横坐标是光谱范围,纵坐标是强度(归一化仅用于展示相对光谱功率分布曲线,实际计算用的曲线 H 是未归一化的)。

$$I_k(x, y) = \int_{\lambda} H(x, y, \lambda) S_k(\lambda) d\lambda, \quad (1)$$

式中: λ 为波长; $H(x, y, \lambda)$ 是像素 (x, y) 在波长 λ 处的像素值; $k \in \{R, G, B\}$ 。式(1)描述的是入射光与相机三通道响应曲线积分后,获取最终输出图像的物理过程。 $S_k(\lambda)$ 为 R、G、B 三通道响应曲线,如图 2 所示。通过式(1),计算波长 λ 范围内接收的多光谱数据 $H(x, y, \lambda)$,得到相同分辨率大小三通道的 RGB 图像 $I_k(x, y)$ 。

2.2 多光谱计算成像方法

基于式(1)所表示的物理成像过程,直接设计深度

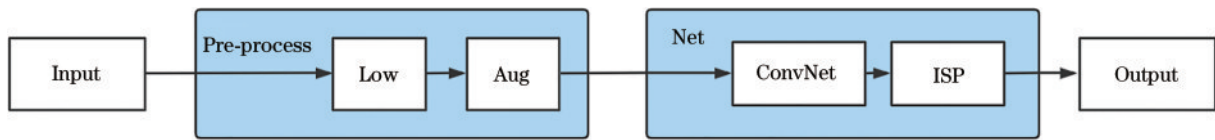


图 3 所提方法实现流程图

Fig. 3 Implementation flow chart of the proposed method

2.2.1 Pre-process 模块

Pre-process 模块用于制作低照度下带噪声的多光谱数据集,并实现数据增强,配合网络训练。首先对干净的多光谱数据实现低照度并引入噪声(图 3 中 Low),添加随机的高斯噪声与椒盐噪声,公式为

$$I'(\lambda) = \alpha \cdot N(I, \lambda) + (G_{\text{noise}} + \beta \cdot P_{\text{noise}}), \quad (2)$$

式中:对于收集的输入多光谱图像 $I, N(I, \lambda)$ 对 I 每个光谱 λ 分别实现归一化;超参数 $\alpha \in (0, 1]$ 按比例整体降低归一化图像后的亮度; G_{noise} 和 P_{noise} 表示所添加的随机高斯噪声(均值为 0, 标准差为 0.1)和椒盐噪声,模拟夜间成像时的大量随机噪声; β 表示相对于原数据生成椒盐噪声的概率,对多光谱数据携带的每一个波长添加相同概率的随机噪声,光谱通道越多,多光谱数据所携带的噪声量越大。式(2)表征了对多光谱数据添加噪声,模拟低照度效果的物理过程。

同时,此模块还在网络训练时负责数据增强功能(图 3 中 Aug),表达式为

$$\text{Aug}(I') = R_r[R_c(I', \gamma_1), \gamma_2], \quad (3)$$

式中: γ_1 和 γ_2 代表两种函数功能实现的概率。对于式(2)所得到的带噪声的多光谱数据 I' ,首先利用 $R_c(x, \gamma_1)$ 进行随机裁剪,获取符合网络输入分辨率大小的图像,再通过 $R_r(x, \gamma_2)$ 进行随机镜面、上下、左右 90° 翻转。式(3)计算得到数据增强后的带噪声的多光

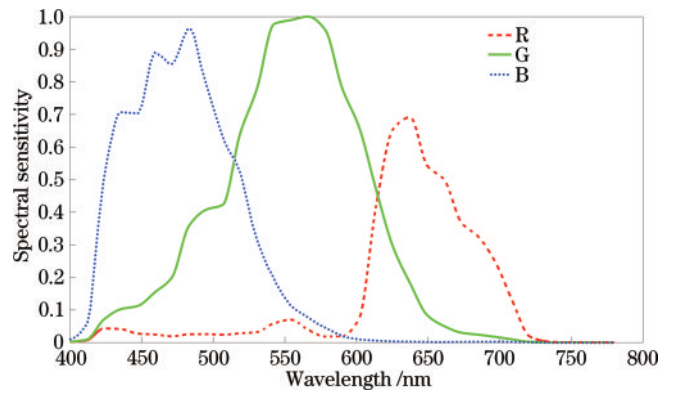


图 2 相机感应曲线示例

Fig. 2 Example of camera response curves

学习网络对多光谱数据进行端到端处理,获得最终 RGB 图像。图 3 为所提方法的模型框架。Input 为输入的多光谱图像,Pre-process 模块在数据训练时模拟低照度及噪声生成,Net 模块为核心多光谱融合算法,Output 为最终输出图。

谱数据。

2.2.2 多光谱融合算法模块

多光谱融合算法模块的具体结构如图 4 所示,可实现对低照度下带噪声的多光谱数据 I (训练时为 I') 进行光谱融合并去噪的功能,主要由卷积网络和标准化 ISP 两个子模块构成。将 ISP 中多个图像处理流程按照网络层的形式添加至网络后端,而 ISP 所包含的网络层中不含有可学习参数。

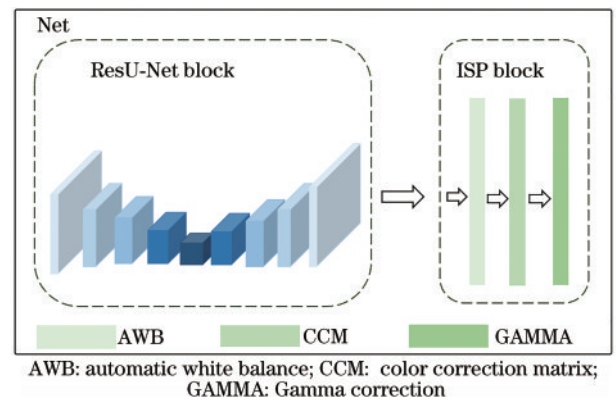


图 4 多光谱融合算法模块

Fig. 4 Multispectral fusion algorithm block

第一个网络模块采用变种 U-Net 模型,如图 5 所示。U-Net 作为卷积网络的一种,因局部性和参数共

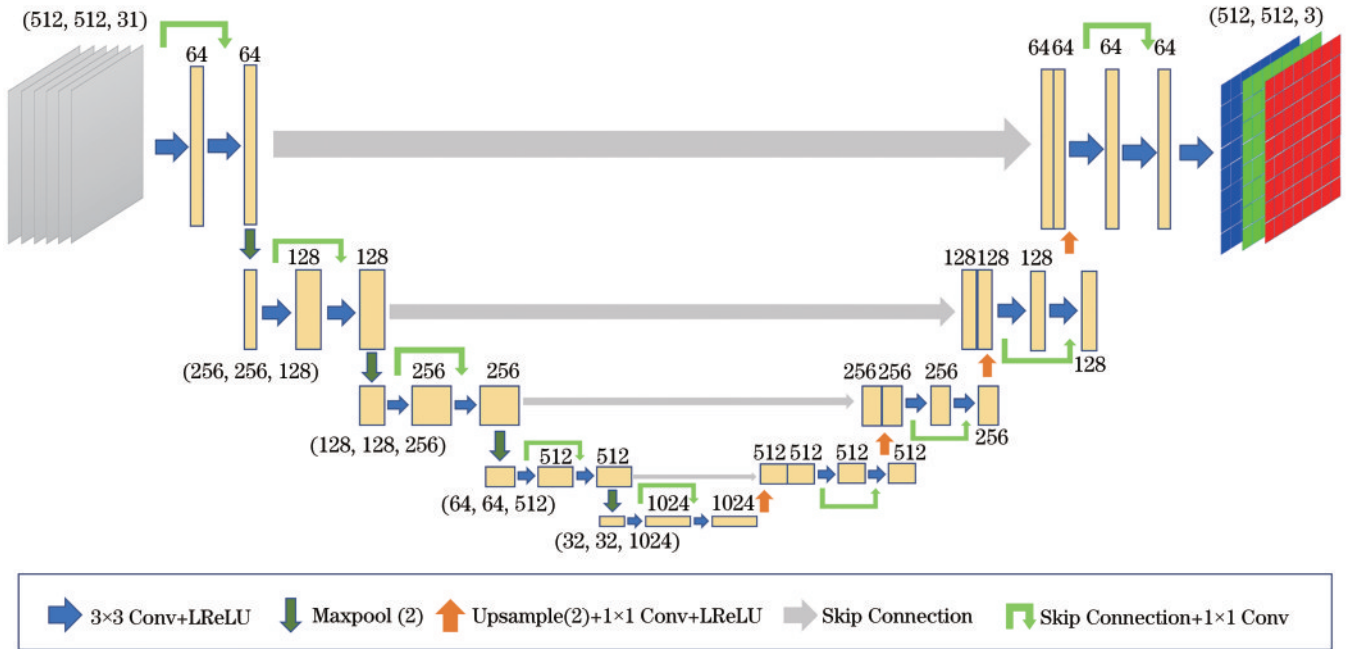


图 5 U-Net模型结构
Fig. 5 Structure of U-Net model

享的特点能有效地关联处理二维图像的空间信息,并降低参数量,使得网络易于优化。而相比于其他卷积神经网络(CNN),U-Net网络中对称U型结构能够同时参考不同分辨率的特征信息;多次下采样的处理能够提取不同尺度的特征,增强网络对输入的鲁棒性;同时 Skip Connection使得网络的输出结合了不同感受野的信息;特征融合时采用通道维堆叠的方式增加特征的维度,提供了更丰富的特征信息;U型右侧,通过多次上采样,在恢复图像分辨率的同时结合左侧浅层特征的补充,使得最终输出在像素级上对图像边缘纹理的表达更精细。

网络从输入图像提取不同层次的语义特征,在每层添加残差连接提升深层网络的效果,并在多次上采样中进行融合,最终输出为3通道RGB图像。每个卷积层后的激活层为Leaky ReLU(LReLU),有效避免训练中梯度消失的问题,且网络中上采样操作得到的图像尺寸是采样输入的2倍。另外,为实现通道数匹配,上采样操作伴随着一个1x1卷积操作,用于改变通道维度。

ISP模块对前向卷积网络输出的RGB数据进行一系列标准化后处理,实现更精准颜色和对比度还原,进一步提高图像质量,增加与人眼感知的相似性。该模块具有图像的自动白平衡(AWB)、color correction matrix(CCM)校正和Gamma校正3个主要功能。自动白平衡通过对图像颜色分布规律进行深入分析,估计场景中白区所记录光源的概率,结合人眼颜色恒常性定律,对图像三通道进行增益调整,从而完成光源颜色校正。CCM校正通过固定的3x3矩阵对图像三通道颜色分量进行进一步调整,校正不同色温对相机表

达颜色的影响,使得机器与人眼对颜色的理解统一。在颜色校正时由于校正矩阵中存在负值,生成的结果中可能存在负值,则需要在这一步Gamma校正前进行数据归一化处理,公式为

$$N(x) = \frac{x - \mu(x)}{\sqrt{\sigma^2(x) + \epsilon}}, \quad (4)$$

式中: ϵ 为一最小项,避免除以0,使得所有像素值位于 $[0, 1]$; $\mu(x)$ 为均值; $\sigma^2(x)$ 为方差。最后再通过Gamma校正,压缩亮部区域,拉伸暗部区域,使得图像灰度变化符合人眼亮度感知。

2.3 损失函数和评价指标

训练网络的损失函数由两种常见的损失函数加权组成,分别为均方误差(MSE)损失和感知损失^[18],表达式为

$$L[f_w(x), y] = \alpha \cdot L_{MSE}[f_w(x), y] + \beta \cdot L_{Per}[f_w(x), y]. \quad (5)$$

两种损失函数分别在图像具体像素值上与高语义特征上进行图像质量评价L。MSE用于统计计算两个相同维度的输入之间差值的平方的期望值,数值越小代表输入数据间像素级的差异越小,计算公式为

$$L_{MSE}[f_w(x_i), y] = \frac{1}{n} \sum_{i=1}^n [f_w(x_i) - y_i]^2. \quad (6)$$

感知损失的流程如图6所示,将两幅图像 $f_w(x)$ 和 y 输入一个相同的预训练网络 $\tau(\cdot)$,分别计算网络输出的高级特征图 $\tau[f_w(x)]$ 和 $\tau(y)$ 。通过MSE计算高级特征图之间的差异来衡量原图像 $f_w(x)$ 和 y 整体语义的相似性,公式为

$$L_{Per}[f_w(x), y] = L_{MSE}\{\tau[f_w(x)], \tau(y)\}. \quad (7)$$

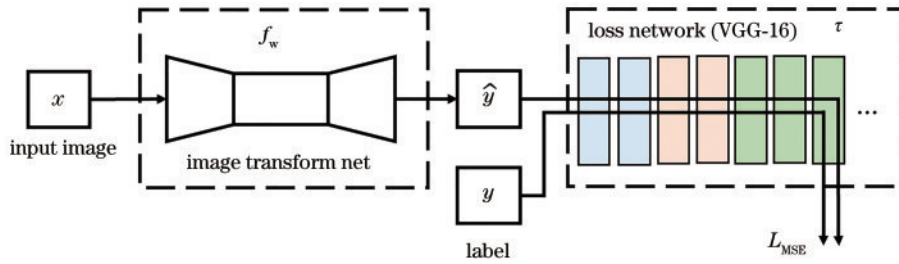


图 6 感知损失流程图

Fig. 6 Flow chart of the perceptual loss

在网络效果评价方面,选择经典的图像评价方法,包括峰值信噪比(PSNR)和结构相似度(SSIM)^[19]。PSNR用于描述两幅图像之间的失真程度,公式为

$$R_{\text{PSN}} = 20 \times \log_{10} \frac{I_m}{\sqrt{E_{\text{MS}}}}, \quad (8)$$

式中: I_m 表示当前图像类型中的最大像素值; E_{MS} 为两幅图像之间像素级的均方差。PSNR值越大,代表图像之间失真程度越小,图像越相似,单位为dB。SSIM从亮度 $l(x, y)$ 、对比度 $c(x, y)$ 、结构 $s(x, y)$ 三个层次分别计算并将它们按一定指数权重结合,来衡量图像之间的相似程度,公式为

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}, \quad (9)$$

$$c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}, \quad (10)$$

$$s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}, \quad (11)$$

$\text{SSIM}(x, y) = [l(x, y)]^a \cdot [c(x, y)]^b \cdot [s(x, y)]^c$, (12) 式中: μ_x 和 μ_y 为图像 x 和 y 的均值; σ_x^2 和 σ_y^2 为方差; σ_{xy} 为协方差; C_1, C_2, C_3 为常数,避免分母为0。计算所得值范围在 $[0, 1]$,值越大,代表两张图像之间的结构特征更相似,将指数加权比设定为 $a=\beta=\gamma=1$ ^[19]。

3 实验内容及结果分析

3.1 数据集

使用的数据集为NTIRE2018^[20]提供的多光谱数据集,并采用Pre-process模块制作携带大量噪声的多光谱数据,模仿真实低照度场景。该数据集使用Specim PS Kappa DX4高光谱相机配合旋转扫描载物台采集而成。该数据集包括200张波长范围从400 nm到700 nm间隔10 nm的31通道光谱数据,分辨率为 1300×1392 ,色深为12 bit。数据集对应的真值图像是真值多光谱数据直接与相机响应曲线积分后,经ISP模块标准化处理后得到的,模拟在正常照度下的高质量成像效果。最终将200张多光谱数据随机地按190:10的比例划分为训练集和测试集。

3.2 实验相关

3.2.1 数据预处理

详细介绍制作低照度下带噪声的多光谱数据集的

方法,配合式(2),其中 $\alpha=0.6$,多光谱数据按比例整体降低亮度, $\phi=31$ 代表多光谱数据的通道数。首先对多光谱数据所有像素值进行归一化,再对每个光谱通道各自添加均值为0、标准差为0.1的随机高斯噪声和随机概率为10%的椒盐噪声(其中椒盐比例各占一半)。

由于提到的多光谱数据集样本量较小,在训练前需通过式(3)进行数据增强。对随机抽取的小批量样本进行随机裁剪,尺寸保持为 512×512 像素(不含黑边),从而满足网络输入大小的要求;再设置 $\gamma_1=\gamma_2=50\%$,也即每一个样本进入网络前都有50%概率旋转和翻转,防止过拟合。为了网络能够正常训练和收敛,真值需与输入训练数据同步完成预处理操作,如图7所示。

3.2.2 参数设置

使用AdamW优化器进行网络训练,内置权重衰减操作,在网络参数迭代更新时,通过添加与参数值有关的惩罚项来降低梯度回传后参数的值,优化权重收敛策略。相比于L2正则化,权重衰减操作在参数更新部分添加惩罚项而不是在损失函数计算部分,避免了对损失函数求梯度时也要对正则项求梯度时的额外计算消耗,AdamW优化器能够在一定程度上缓解数据集样本量少可能造成的过拟合问题。训练Batchsize设为24,共训练800个epoch,初始学习率为 10^{-4} ,采用StepLR策略调整,在训练过程中规律地降低整体学习率,提升优化效果,每经过100次反向传播优化迭代后学习率缩小0.96。网络框架的搭建采用Pytorch 1.8.1软件实现。

3.3 实验结果

提出一种基于变体U-Net结构的Res-UNet神经网络,将传感器上获得的多光谱数据融合成RGB数据,部分替代ISP前端功能,能够给后续固定ISP模块输入更丰富的图像信息,而不是对ISP输出的sRGB图像进行后处理来补偿重建图像信息,单帧图像重建时间约为150 ms。成像结果如图8所示,Input是模拟低照度的多光谱数据,此处为了展示方便,通过相机曲线积分将其映射到RGB图像,值得注意的是,送入网络训练的是多光谱数据;Input-ISP是Input直接经过后续标准化ISP处理的结果,模拟普通相机在低照度

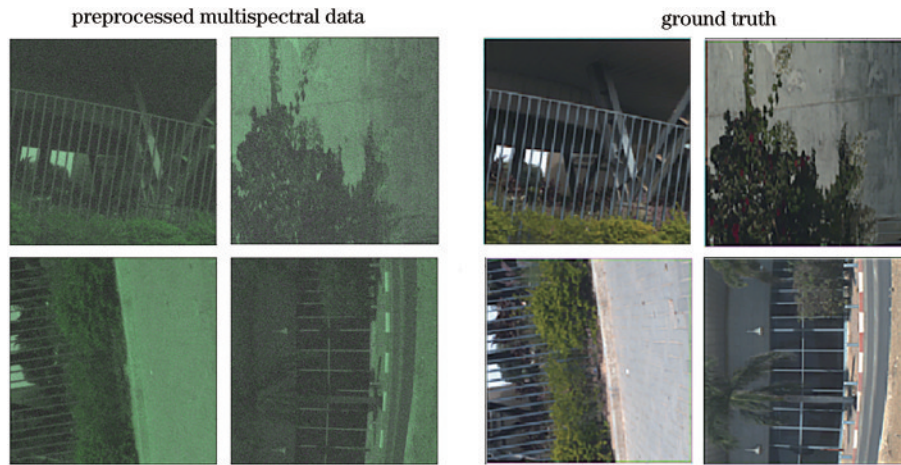


图 7 预处理后的数据和真值图像

Fig. 7 Preprocessed data and ground truth

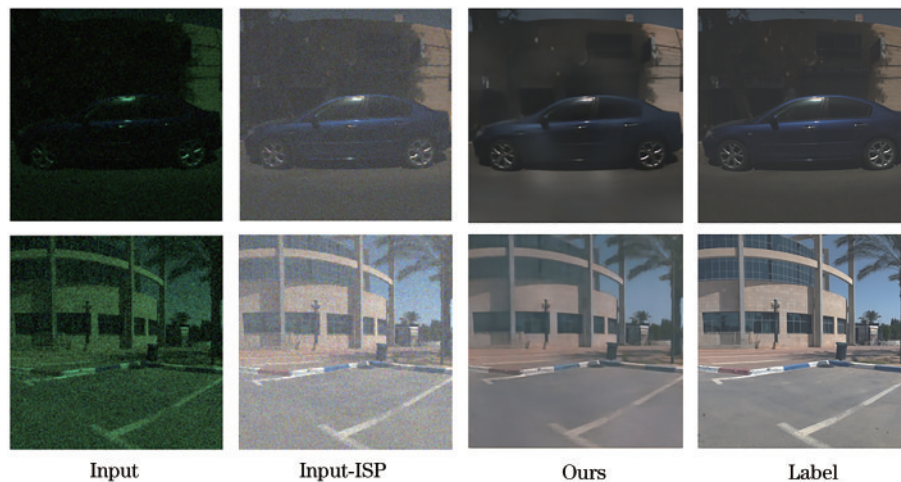


图 8 实验结果对比

Fig. 8 Comparison of experimental results

下的成像结果,图中还是存在ISP无法消除的噪声,纹理细节被严重遮盖,如果加大去噪力度,将会进一步平滑边缘细节,恶化图像质量,现有的大多研究便是针对此结果寻找合适的后处理方法以提高成像质量;Ours则是所提端到端方法对输入的31通道多光谱数据进行融合,输出重建的RGB图像;Label则是对应图像真值。所提方法达到了理想的去噪效果,图像中细节保留更加完整,且在人眼感知上更贴近真值图像。

从图像的边缘、色彩和高动态表现3个方面对所提端到端成像方法进行详细对比展示。图像边缘是图像质量评估的重要因素之一。在严重受污染的信号上,边缘重建是一项具有挑战性的任务。传统图像处理通过对感受野内的信息进行特征提取与推测,完成后续边缘细节复原填充,但这会造成严重的伪细节。所提增加输入光学信息量的方法有效地从多光谱通道中完成特征学习,进而重建边缘特征。如图9方框对应放大图像所示,噪声会影响栏杆细节表现,影响该区域的成像质量,图像整体有颗粒感,所提方法能还

原边缘部分的纹理细节。

提高最终图像的色彩还原度也是相机研发的重要目标。在网络训练中,结构化相似度、感知损失等都可以用于惩罚色偏。如图10所示,普通相机成像中,场景中大量的随机噪声在经过复杂的处理后在最终成像上会以彩色椒盐噪声的形式存在,影响图像的色彩表达,所提方法利用多光谱数据减缓了椒盐噪声的影响,最终在色彩丰富且密集的区域能够恢复出易被噪声淹没的颜色。

在高亮度动态场景中成像时,图像中易存在区域过曝或过暗的情况。如图11所示,在逆光拍摄条件下,由于左上角太阳光的亮度过高,图像中其他区域的成像亮度较暗,图像整体亮度偏低。在方框暗部区域中,有大量细节存在但是和噪声混杂,普通相机在该区域的成像效果会严重地受噪声影响,所提方法能够合理地耦合多光谱数据的光谱通道,在暗部区域里分离噪声的同时也一定程度上保留了高亮度范围图像中暗部区域细节。



图 9 图像边缘评估
Fig. 9 Image edge evaluation

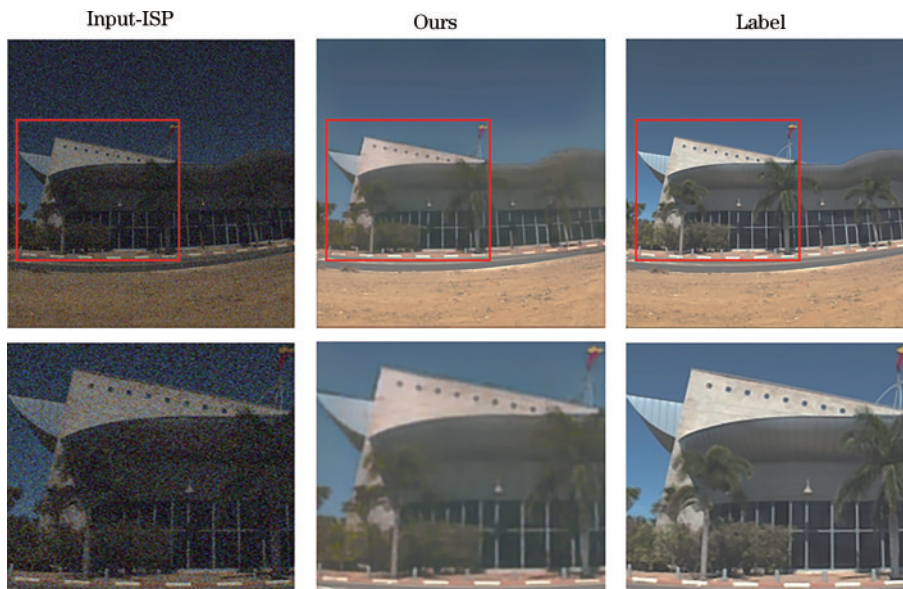


图 10 图像色彩评估
Fig. 10 Color reconstruction evaluation

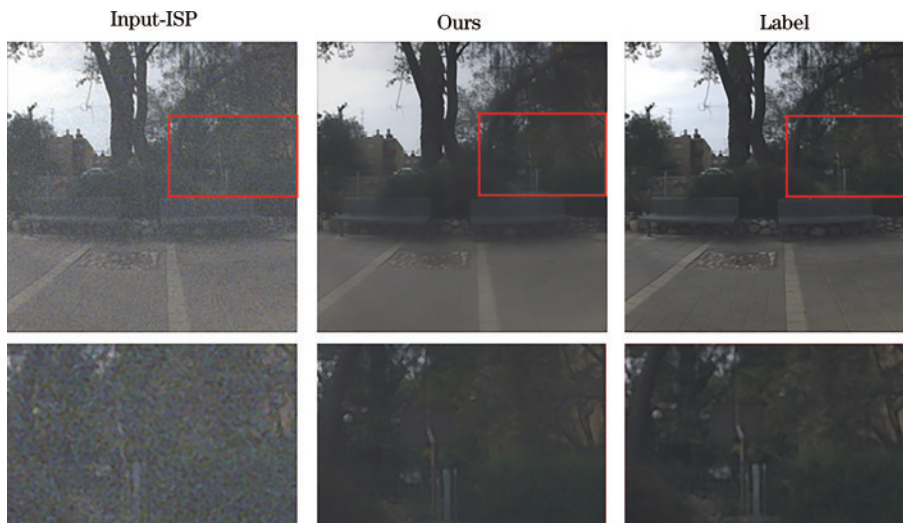


图 11 图像细节评估
Fig. 11 Image detail evaluation

3.4 真实数据测试

所提网络训练时所采用的数据集是根据模拟制作而成的,对干净的多光谱数据添加随机的噪声来模拟低照度下噪声的分布。从研究的实际应用出发,为

了验证该模拟训练集所训练的网络模型在真实应用场景下的有效性,现利用真实低照度下所扫描的多光谱数据对网络模型进行成像测试,成像效果如图 12 所示。



图 12 真实多光谱数据成像测试

Fig. 12 Real multispectral data imaging test

为了方便展示,图 12(a)为多光谱数据通过与相机响应曲线积分到 RGB 格式用于显示的结果,可以看见左侧存在部分高亮区域,整图存在宽动态亮度范围的问题;图 12(b)为对图 12(a)进行 ISP 处理的结果,底部仍存在大量噪声,对左侧高亮过曝区域不当处理导致影子部分成像消失;图 12(c)为所提网络模型处理多光谱数据的成像效果,能够有效解决左侧过曝区域中影子的成像问题,且能从高亮区域中还原人偶嘴部的颜色细节。尽管在真实应用场景下的低照度噪声分布存在本底噪声和环境噪声各异的问题,所提网络能达到一定的消噪效果,同时也验证多光谱数据能够解决在真实低照度下成像的宽动态问题,还原高亮区域的色彩细节。

3.5 消融实验

将通过 3 个消融实验,分别从数据、噪声量(低照度)、网络结构 3 个方面来比较评估所提利用神经网络对多光谱数据进行端到端融合成像,从而优化低照度成像的方法。

3.5.1 输入数据

为了更加客观完整地验证多光谱融合提升源信息量在高质量成像任务中的优势,需要同时对传统 RGB 数据进行实验测试。低照度下,按式(1)对多光谱数据进行积分成像,得到 RGB 图像,所生成的 RGB 数据再送入图 4 的网络结构中,网络只调整输入通道数以匹配 RGB 图像,其余网络结构、参数及训练方法全部与所提融合多光谱方法一致。此时唯一变量便是输入图像通道数,以此验证源信息的增强对成像质量提升的有效性。如图 13 所示:在相同训练条件下,基于多光谱数据的网络能够很好地还原出图像细节和色彩;使用 RGB 数据训练的网络几乎不能从噪声中还原出场景里的色彩信息,且图像纹理复杂的区域还非常模糊,无法分辨其中细节信息,主要原因为受制于有限的输入源信息,噪声影响也比较大,网络无法提取足够特征

进行颜色和细节还原。

进行量化比较,如表 1 所示,基于多光谱数据的网络在测试集上的指标都能优于基于 RGB 数据的网络。在训练过程中,在相同的光谱范围内,基于多光谱数据的网络细分了更多光谱通道的同时携带了更细致的颜色信息,而且不同通道间信息具有互相关作用,相互参考有助于恢复最终纹理和颜色。

3.5.2 噪声等级

对于训练好的网络模型,需要测试其在多种不同低照度情况下的表现情况。采用添加不同级别随机噪声的方法来验证不同数据所训练相同结构网络模型的鲁棒性,设定本文所训练验证集的含噪量为 50%,再分别制作含噪量为 25% 和 75% 的测试集。含噪量为 25%、50%、75% 时所含高斯噪声相同,所含椒盐噪声的随机概率 β 分别为 10%、20%、30%。测试网络的融合去噪优化表现,得到的数值如表 2 所示,其中 Loss 1 为 MSE, Loss 2 为感知损失。

基于多光谱数据的网络在同一水平噪声下的测试均优于基于 RGB 数据的网络,验证了多光谱数据相比 RGB 在融合成像上更具有优势。纵向对比来看,两种数据的网络都是在含噪量为 50% 的测试集上表现最优,主要因为网络是在 50% 含噪量数据集上训练完成的,因此表现稍好一些,但网络结果在 25% 和 75% 噪声量下都只表现出轻微的上升,上升幅度在合理范围内。

3.5.3 网络结构

对于所使用的添加残差的变体网络 Res-UNet,同时还选择了更多结构对 U-Net 网络进行添加修改,分别是在特征提取部分添加通道注意力机制 squeeze and excitation(SE)^[21]和空间注意力机制 convolutional block attention module(Cbam)^[22],测试网络的性能。表 3 是 4 种网络模型的测试表现。

由表 3 可知,基于 U-Net,增加模型的复杂度整体



图 13 消融实验对比不同输入数据

Fig. 13 Comparison of different input data in ablation experiment

表 1 不同数据训练的网络在测试集上的表现

Table 1 Performance of networks trained with different data on test set

Parameter	Multispectral data	RGB
MSE	0.0144	0.0316
PSNR /dB	36.8328	28.0062
SSIM	0.8859	0.6447
Perceptual loss	2.3287	3.5890

表 2 不同网络的鲁棒性测试

Table 2 Robust test of different networks

Noise	Loss	Multispectral data	RGB
25%	Loss 1	0.0245	0.0469
	Loss 2	2.4294	4.1494
50%	Loss 1	0.0144	0.0316
	Loss 2	2.3287	3.5890
75%	Loss 1	0.0389	0.0635
	Loss 2	2.9136	4.6799

有利于在相同训练设定下对多光谱数据的拟合。如图 14 所示,添加残差连接的效果最好,该种网络生成的测试图像色彩更丰富,对比度更高。这是因为残差连接更直接地将浅层特征送入后端网络层,使得后端

表 3 不同结构网络在测试集上的表现

Table 3 Performance of networks with different structures on test set

Network	MSE	PSNR /dB	Perceptual loss
U-Net	0.0215	33.3512	2.5089
SE-UNet	0.0184	34.7036	2.4657
Cbam-UNet	0.0228	32.8413	2.5217
Res-UNet	0.0144	36.8328	2.3287

网络能接收到更多原始信息,提高了网络间的信息流通;训练网络来拟合残差,使得网络参数对输出的变化更加敏感;同时降低模型复杂度,减缓训练时的过拟合,并打破了网络的对称性,提升了网络的表征能力,有效地改善网络训练后期性能退化、梯度消失的问题。在其他两种不同的注意力机制中,只含有通道注意力的 SE-UNet 的结果要略优于同时包含通道与空间注意力的 Cbam-UNet 模型。这是由于通道间的相互作用对图像融合贡献更大,而在低照度下图像整体含大量噪声,Cbam 中会依赖空间局部的信息融合多光谱数据的做法可能会恶化最终的成像质量。总体来看,4 种网络都能实现相对优质的成像,一定程度上对低照度下受污染的数据进行有效的色彩和细节的重建,而所提添加残差连接的 Res-UNet 在最终成像的整体色调上更贴近真值图像。



图 14 不同结构网络输出对比

Fig. 14 Output comparison of networks with different structures

4 结 论

针对低照度下高质量成像任务,利用多光谱数据,替代相机融合彩色图像的过程,快速地从噪声中恢复并重建图像细节。多光谱图像相比现常用的 RGB 数据,含有更多的光谱通道,增加了数据源信息,在低照度等高质量成像困难的场景环境中具有更多的数据选择性;利用深度学习网络,对数据融合过程进行端到端学习,实现了成像时高动态亮度范围的鲁棒性保护。

基于变体 U-Net 网络,对低照度下含大量随机噪声的 NTIRE2018 多光谱数据集实现光谱融合和图像去噪,验证了多光谱数据应用于相机成像领域的可行性;通过充足的消融实验,验证了多光谱数据融合相较于传统 ISP 方法生成 RGB 图像的优势,在 MSE 和感知损失上分别降低了 54.43% 和 35.12%;且在训练中对于归一化后的网络输出,感知损失相比 MSE 的结果差异更明显,更有利于梯度回传,但是感知损失更看重图像的高级特征,对于整体结构,MSE 在像素级上优化图像,能够更好地纠正色偏;通过对网络输出进行 ISP 后处理,模拟了相机成像的物理过程,在自动白平衡和色彩校正的处理下还原并丰富图像色彩;通过详细丰富的消融实验,从客观量化指标和人眼主观效果上验证了所提方法的可行性和成像效果的优势。

利用多光谱数据实现对低照度下相机的快速成像优化,突破了现有研究数据中场景信息过少的局限,转变了大量研究针对 ISP 输出后处理的思路,同时也能对接现有研究成果进一步提升成像质量。但是该网络模型相较于传统基于 RGB 数据的网络参数更多,在训练网络融合处理数据时可能面临 GPU 损耗更大、内存占用过高等计算成本问题,对于具有计算资源约束的实际场景,通过裁剪模型,实现效果和性能的平衡。本文成功提出及验证了利用多光谱数据实现低照度下图像增强的方法,为夜间安防、矿井监控及其他困难条件下的高质量成像提供了新的研究方向,对现有 AI 技术在相机成像任务中的应用具有重要意义,推动了多光谱数据在未来更多的图像处理研究中的智能应用与发展。

参 考 文 献

- [1] 郭林, 陈亮亮, 程德强, 等. 颜色恢复和边缘保持的低照度图像超分辨率重建方法[J]. 计算机应用研究, 2022, 39(4): 1253-1258, 1264.
Guo L, Chen L L, Cheng D Q, et al. Super resolution of low illumination image based on color restoration and edge preserving[J]. Application Research of Computers, 2022, 39(4): 1253-1258, 1264.
- [2] Wei X S, Luo J H, Wu J X, et al. Selective convolutional descriptor aggregation for fine-grained image

- retrieval[J]. *IEEE Transactions on Image Processing*, 2017, 26(6): 2868-2881.
- [3] Chen C, Chen Q F, Xu J, et al. Learning to see in the dark[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 3291-3300.
- [4] 陈清江, 李金阳, 胡倩楠. 基于并联残差网络的低照度图像增强算法[J]. *激光与光电子学进展*, 2021, 58(14): 1410015.
Chen Q J, Li J Y, Hu Q N. Low-illumination image enhancement algorithm based on parallel residual network [J]. *Laser & Optoelectronics Progress*, 2021, 58(14): 1410015.
- [5] Jobson D J, Rahman Z, Woodell G A. Properties and performance of a center/surround Retinex[J]. *IEEE Transactions on Image Processing*, 1997, 6(3): 451-462.
- [6] Zhang Q, Yuan G Z, Xiao C X, et al. High-quality exposure correction of underexposed photos[C]//Proceedings of the 26th ACM International Conference on Multimedia, October 22-26, 2018, Seoul, Republic of Korea. New York: ACM Press, 2018: 582-590.
- [7] Park S, Yu S, Moon B, et al. Low-light image enhancement using variational optimization-based Retinex model[J]. *IEEE Transactions on Consumer Electronics*, 2017, 63(2): 178-184.
- [8] 王奎, 黄福珍. 基于光照补偿的 HSV 空间多尺度 Retinex 图像增强[J]. *激光与光电子学进展*, 2022, 59(10): 1010004.
Wang K, Huang F Z. Multiscale Retinex image enhancement in HSV space based on illumination compensation[J]. *Laser & Optoelectronics Progress*, 2022, 59(10): 1010004.
- [9] Remez T, Litany O, Giryes R, et al. Deep convolutional denoising of low-light images[EB/OL]. (2017-01-06) [2022-03-04]. <https://arxiv.org/abs/1701.01687>.
- [10] Dabov K, Foi A, Katkovnik V, et al. Image denoising by sparse 3-D transform-domain collaborative filtering[J]. *IEEE Transactions on Image Processing*, 2007, 16(8): 2080-2095.
- [11] Gu S H, Zhang L, Zuo W M, et al. Weighted nuclear norm minimization with application to image denoising [C]//2014 IEEE Conference on Computer Vision and Pattern Recognition, September 23-28, 2014, Columbus, OH, USA. New York: IEEE Press, 2014: 2862-2869.
- [12] Hu Z, Cho S, Wang J, et al. Deblurring low-light images with light streaks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, 40(10): 2329-2341.
- [13] He K M, Sun J, Tang X O. Single image haze removal using dark channel prior[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33(12): 2341-2353.
- [14] Dong X, Wang G, Pang Y, et al. Fast efficient algorithm for enhancement of low lighting video[C]//2011 IEEE International Conference on Multimedia and Expo, July 11-15, 2011, Barcelona. New York: IEEE Press, 2011.
- [15] Lakshman P. Combining deblurring and denoising for handheld HDR imaging in low light conditions[J]. *Computers & Electrical Engineering*, 2012, 38(2): 434-443.
- [16] HUAWEI. [P30] P30 RYYB sensor: see through the dark! [EB/OL]. (2019-04-15) [2022-02-03]. https://consumer.huawei.com/en/community/details/P30-P30-RYYB-Sensor-See-Through-the-Dark/topicId_5180/.
- [17] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation[EB/OL]. (2015-05-18)[2022-03-04]. <https://arxiv.org/abs/1505.04597>.
- [18] Johnson J, Alahi A, Li F F. Perceptual losses for real-time style transfer and super-resolution[EB/OL]. (2016-03-27)[2022-03-04]. <https://arxiv.org/abs/1603.08155>.
- [19] Wang Z, Bovik A C, Sheikh H R, et al. Image quality assessment: from error visibility to structural similarity[J]. *IEEE Transactions on Image Processing*, 2004, 13(4): 600-612.
- [20] ICVL. Hyperspectral database-early access[EB/OL]. [2022-03-04]. <https://icvl.cs.bgu.ac.il/hyperspectral/>.
- [21] Hu J, Shen L, Albanie S, et al. Squeeze-and-excitation networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(8): 2011-2023.
- [22] Woo S, Park J, Lee J, et al. CBAM: convolutional block attention module[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11211: 3-19.