

面向 JPEG Pleno 光场编码失真的盲伪影去除

赖嘉隆¹, 蒋志迪², 蒋刚毅¹

(1. 宁波大学 信息科学与工程学院, 浙江 宁波 315211; 2. 宁波大学 科学技术学院信息工程学院, 浙江 宁波 315212)

摘要: JPEG Pleno 压缩会在压缩后的 4D 光场图像中产生伪影, 这不仅会降低其视觉质量, 还会影响其他图像处理任务的性能。目前尚缺少对 JPEG Pleno 光场编码失真修复的研究和可供网络训练的 JPEG Pleno 编码失真光场数据集。本文提出了一种端到端灵活的盲卷积神经网络, 即 JPLARNet。它充分考虑 JPEG Pleno 编码特点, 且通过预测其编码压缩质量因子, 以在伪影去除和细节保留之间进行权衡。具体而言, JPLARNet 首先使用空角解耦模块对光场图像进行初步特征提取, 再经过多尺度的解耦器得到预测的压缩因子和高级语义特征, 然后通过压缩因子融合模块将预测的压缩因子嵌入到随后的重构器模块中, 以此引导压缩光场图像伪影去除, 生成高质量光场图像。此外, 构建了大核下采样模块和混合注意力增强模块, 分别用于下采样和将重构的图像进行图像增强。实验结果表明, 在所构建的数据集 JPL-DATA 压缩因子从 5000 到 200000 的六个压缩质量上, 伪影去除后的光场图像相较于去除前其 YUV-PSNR/Y-SSIM 平均增益为 0.81dB/0.025。由于考虑了 JPEG Pleno 光场图像编码的特点, 相较于 2D 图像 JPEG 伪影去除方法, 所提出的方法获得了具有更高主客观质量的光场图像。

关键词: JPEG Pleno 伪影去除; 光场图像修复; 卷积神经网络; 多尺度信息

中图分类号: TP391.41; O438 **文献标识码:** A **DOI:** 10.3969/j.issn.1001-5078.2025.06.020

Blind artifact removal for JPEG Pleno light field coding distortion

LAI Jia-long¹, JIANG Zhi-di², JIANG Gang-yi¹

(1. Faculty of Information Science and Engineering, Ningbo University, Ningbo 315211, China;

2. College of Information Engineering, College of Science and Technology, Ningbo University, Ningbo 315212, China)

Abstract: JPEG Pleno compression can introduce artifacts in the compressed 4D light field images, which not only degrade their visual quality but also affect the performance of other image processing tasks. At present, there is a lack of research on JPEG Pleno light field coding distortion repair and JPEG Pleno coding distortion light field dataset for network training. In this paper, an end-to-end flexible blind convolutional neural network, namely JPLARNet, is proposed, which fully considers the characteristics of JPEG Pleno encoding and can predict the coding quality factor to control the trade-off between artifact removal and detail preservation. Specifically, JPLARNet initially employs a spatial-angle decoupling module to perform preliminary feature extraction on light field images. The extracted features are then processed by a multi-scale decoupler to obtain predicted compression factors and high-level semantic features. Subsequently, a compression factor fusion module is utilized to embed the predicted compression factors into the sub-

基金项目: 国家自然科学基金项目 (No. 62271276) 资助。

作者简介: 赖嘉隆 (2000 -), 男, 硕士研究生, 研究方向为光场图像编码失真修复。E-mail: 1303756217@qq.com

通讯作者: 蒋刚毅 (1964 -), 男, 博士, 教授, 博士生导师, 主要从事多媒体信号处理与通信、计算成像、视觉感知与编码、图像与视频质量评价等方面的研究工作。E-mail: jianggangyi@126.com

收稿日期: 2024-08-15

sequent reconstructor module, and then embeds the predicted compression factors into the subsequent reconstruction module through the compression factor fusion module, thereby guiding artifact removal in compressed light field images. In addition, two modules, namely the LK Down-sample module and the mixed attention enhancement module, are constructed for downsampling and image enhancement of the reconstructed images, respectively. The experimental results show that on the six compression qualities of the constructed JPL-DATA dataset with compression factors ranging from 5000 to 200000, the average gain in YUV-PSNR/Y-SSIM of the light field image after artifact removal compared to before removal is 0.81 dB/0.025. By taking into account the characteristics of JPEG Pleno encoded light field image, the proposed method achieves light field images with higher subjective and objective quality than the JPEG artifact removal method of 2D images.

Keywords: JPEG Pleno artifact removal; light field image restoration; convolutional neural networks; multi-scale information

1 引言

光场成像能同时捕获场景中的光线强度与光线方向,为用户提供更真实的场景信息。与2D图像不同,光场图像(Light Field Image, LFI)不仅可以包含图像的空间域信息,同时还包括独特的角度域信息。光场图像可以由4D数学模型 $L(u, v, x, y)$ 表示,其中 (u, v) 平面和 (x, y) 平面分别对应于角度域信息和空间域信息^[1]。然而,由于记录更多信息而产生的光场图像具有巨大的数据量,这对存储设备和传输通道构成了巨大的负担^[2]。因此,必须对光场图像进行压缩。在光场图像压缩中,广泛采用有损图像压缩方法用于图像存储和传输。然而,有损压缩不可避免地会在压缩图像中产生失真,例如块效应和振铃效应。这些失真会降低压缩图像的视觉质量和后续计算机视觉任务的性能。因此,迫切需要探索减少压缩光场图像伪影的方法,以提高压缩光场图像的视觉感知效果,并且推动光场图像在各个领域更广泛的应用^[3]。

2D图像JPEG去伪影任务是图像恢复领域中的一个重要任务,目前这类任务已经有许多优秀的研究。例如,Dong等人^[4]提出的ARCNN。其是基于深度学习的JPEG伪影去除的开山之作,第一个将卷积神经网络用于JPEG伪影去除。这是一个4层卷积网络,灵感来自他们之前的图像超分辨率工作SRCNN。Zheng等人^[5]提出的S-Net引入了一种可扩展的架构,在推理过程中可以调整网络深度。Li等人^[6]提出的RGSF分为两个阶段,第一阶段恢复JPEG图像,第二阶段对JPEG图像进行上采样,超分和伪影去除同时进行。SB Ko等人^[7]提出了JRCNN,利用深度密集残块和并行卷积层联合处理JPEG伪影去除和超分辨率任务。Sun等人^[8]通过预测DCT系数,从而利用DCT域的信息,有效地去除压缩伪影。最近,Yang等人^[9]将空域和频域同时

考虑,提出双域学习网络,以交互学习空间信息与频域信息。尽管这类方法成功地减少了块效应和振铃效应,但它们在处理未知质量因子(QF)的图像仍面临困难。

因此,针对不同压缩质量的图像进行恢复的灵活性成为最近工作的主要关注点之一,特别是在QF未知的盲伪影去除场景中。这类方法通过预测QF,并把其融入特征图中,以作为先验知识引导图像进行修复。例如,Galteri等人^[10]提出了一个GANs集合,每个GAN用于一个压缩质量级别。他们使用了一个额外的QF预测网络来选择使用哪个GAN网络。Kim等人^[11]通过首先估计QF,然后从模型集合中选择一个模型,提供了类似的解决方案。Ehrlich等人^[12]提出了一种由量化矩阵(QM)参数化的量化引导JPEG AR网络。其用于完全在频域上处理图像,但未取得很好的效果。而Yoon J等人^[13]提出的DOGNet网络则从输入图像本身提取质量信息,而不是用JPEG头文件中提供的QF,以获取更准确的质量信息引导伪影去除任务。此外,还有些学者利用特殊网络进行JPEG AR任务。例如,Kawar等人^[14]使用扩散模型对编码失真图像进行修复。这类方法很好地完成了在不同QF下的盲伪影去除任务。

JPEG Pleno光场编码方式是JPEG委员会推出的标准。JPEG Pleno编码标准作为新颖的标准,目前仍有大量的研究基于JPEG Pleno光场编码^[15-16]。JPEG Pleno支持两种编码方式,分别为变换模式(4DTM)与预测模式(4DPM)。4DTM基于四维离散余弦变换(4D-DCT)和块划分,而4DPM利用光场数据之间的相关性进行预测编码^[17]。无论是4DTM还是4DPM,其都是利用了光场的4D冗余进行编码。根据文献[18]的结果显示,对于角度密度极高的光场,如来自Lenslets数据集的光场,

4DTM 相对于 4DPM 与伪视频光场编码方式 (HEVC 光场编码) 表现更优。因此, 本文主要关注于低码率条件下 JPEG Pleno 4DTM 编码的失真修复问题。

JPEG Pleno 4DTM 的基本原理与 2D 图像的 JPEG^[19] 编码方式相似。首先将 4D 光场数据分成固定大小的 4D 块。其次, 根据率失真准则得到最优分区参数, 并根据此参数将初始块进一步划分成一组不重叠的 4D 子块。然后让每一个不重叠的子块通过 4D DCT 变换进行独立的变换。最后, 利用十六树位平面分解和自适应算术编码对变换后的块进行量化和熵编码, 得到光场的压缩表示。然而这种分块的编码方式在低码率下会产生严重的块效应。由于 4DTM 编码块的大小不是固定的, 所以表现在图像上为不同大小的块效应失真, 这也是 JPEG Pleno 4DTM 光场编码与基于 HEVC 的伪视频光场编码和 2D 图像 JPEG 编码的不同之处。

综上所述, 因为 JPEG 编码与 JPEG Pleno 4DTM 编码虽存在共性, 但更多是差异, 所以不能简单地将 2D JPEG 去伪影方法用于去除光场图像 JPEG Pleno 伪影。此外, 目前缺乏相应的 JPEG Pleno 伪影去除方法与相应的 JPEG Pleno 4DTM 光场失真数据集。针对上述问题, 本文主要研究 JPEG Pleno 4DTM 在低码率下产生的编码失真修复。首先创建了相应的

光场 JPEG Pleno 编码失真数据集以供网络进行训练与测试。提出了一个端到端网络对 JPEG Pleno 编码失真的光场图像整体进行伪影去除任务, 实验结果证明该网络相比 2D JPEG 伪影去除网络更加符合光场的结构特性和 JPEG Pleno 编码特点, 取得了更好的结果。

2 所提出方法

2.1 JPLARNet 总体网络结构

本文针对 JPEG Pleno 4DTM 光场编码方法在不同低码率下产生的不同程度的失真, 从 JPEG Pleno 编码原理和光场的结构特点考虑, 提出了一个端到端的灵活网络 JPLARNet, 对不同压缩质量的 JPEG Pleno 失真光场进行盲伪影去除, 其网络结构如图 1 所示。该方法主要包括浅层特征提取模块、深度多尺度特征提取模块、质量增强模块。

输入经过 JPEG Pleno 压缩的光场子孔径图像 (Sub-Aperture Image, SAI) 阵列 I_C , 得到修复后的光场 SAI 阵列 I_R 。即:

$$I_C = \text{JPEG Pleno}(I_{gt}) \quad (1)$$

$$I_R = \text{JPLARNet}(I_C) \quad (2)$$

其中, JPEG Pleno 表示 JPEG Pleno 变换编码模式, I_{gt} 表示未经编码的光场 SAI 阵列, 即真值光场图像。

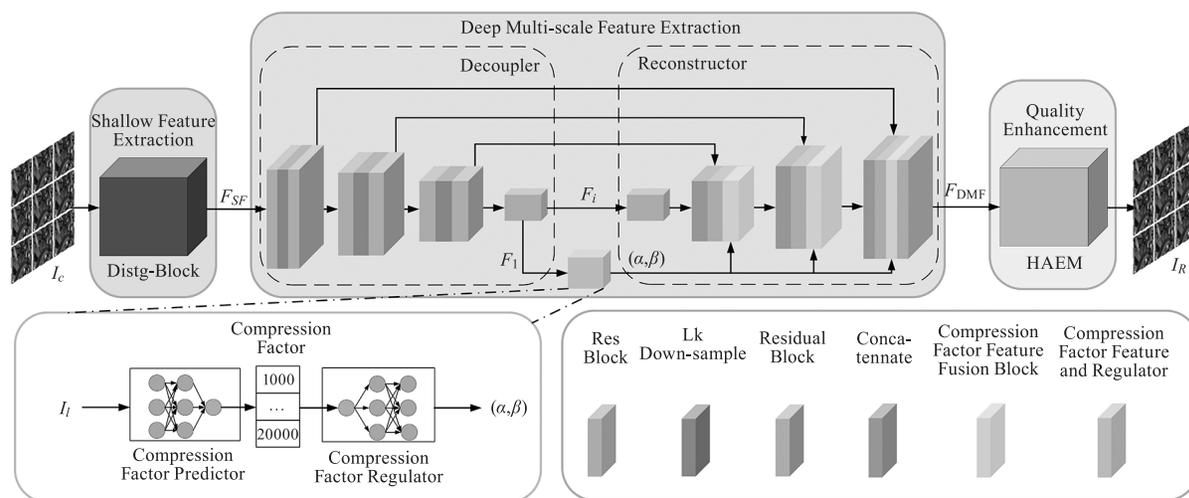


图 1 所提出的 JPLARNet 整体框架

Fig. 1 Overall framework of the proposed JPLARNet

浅层特征提取模块采用文献[20]中的光场空角解耦模块 (Disentangling block, Distg-Block) 提取压缩后的图像特征, 具体来说, 对于给定的压缩图像 I_C , 使用光场空角解耦模块提取浅特征 F_{SF} , 浅层特征提取模块使模型能够更有效地从输入压缩图像中捕获空间信息。这对于在压缩伪影减少任务中保留输出的更精细的细节和纹理是必不可少的。Distg-Block

模块使用 SFE (Spatial Feature Extractor)、AFE (Angular Feature Extractor)、EFE (EPI Feature Extractor) 子块分别提取光场图像的空间信息、角度信息、极平面信息, 其结构与基本工作原理如图 2 所示, 图中光场的角度分辨率为 3×3 (即 $A = 3$)。本文使用该模块时去除了其最外层残差连接, 因为其会把压缩图像中的大量伪影信息代入特征 F_{SF} 中。

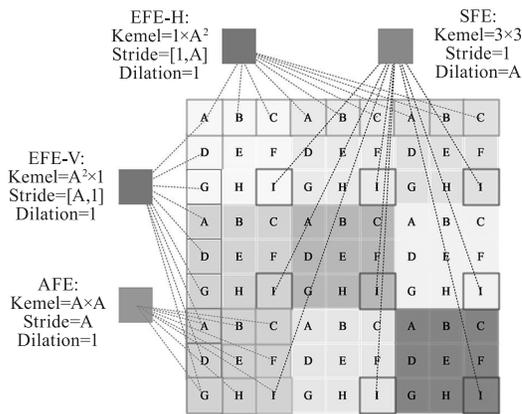


图2 SFE、AFE 和 EFE 基本工作原理的示意图

Fig. 2 Diagrams of the basic working principles of SFE, AFE, and EFE

随后,进入深度多尺度特征提取模块,从压缩图像中提取多尺度特征 F_{DMF} 。该模块包括多尺度特征解耦与重建模块、压缩因子预测器与调节器,以充分提取压缩图像中的有效信息。最后,通过质量增强模块对多尺度特征 F_{DMF} 进行处理,以增强图像质量。

2.2 多尺度特征提取模块

2.2.1 多尺度特征解耦

该模块的输入 I_C 是经过浅层特征提取模块处理后的特征图 F_{SF} ,其目标是提取图像的深层特征,并分解图像特征 F_i 和压缩因子特征 F_1 。充分利用多尺度特征提取方式,以应对 JPEG Pleno 4DTM 产生的失真光场上不同大小的块效应,最大程度提取失真光场有效信息,去除块效应与振铃效应,从而提高重建图像质量。它涉及四个尺度,从第一到第四尺度的每一层中的输出通道数目分别被设置为 64、128、256、512。每个尺度都有一个到重建器的跳跃连接。经过四层 Res Block 递进式地提取有效信息,并解耦出深度图像特征和压缩因子特征,然后通过压缩因子预测器预测出压缩因子,同时通过压缩因子特征融合块用预测出来的压缩因子来引导每个尺度的图像重建。

此外,考虑到较大的感受野有利于区分伪影和图像内容以及去除宽范围伪影,所以如何有效地扩大感受野是问题的关键。而多尺度的特征提取会涉及下采样,其一般使用膨胀卷积以此扩大感受野,但它主要依赖于特征图的稀疏采样,这可能导致像素信息的丢失,进而可能导致恢复结果粗糙或参数浪费,从而影响网络的效率。为了克服这个缺点,本文提出一个基于大核的卷积代替膨胀卷积进行下采样,命名为 LK Down-sample,以此用相似的成本和计

算量,更好地保留图像有效像素信息。LK Down-sample 模块框架如图 3 所示。其中 LKDSM Block 对图像特征进行下采样,首先通过 1×1 卷积将输入特征图的通道提高到 t 倍,然后应用内核大小为 k 、步长为 2 的大核依赖卷积,并且应用 1×1 卷积以减少通道。

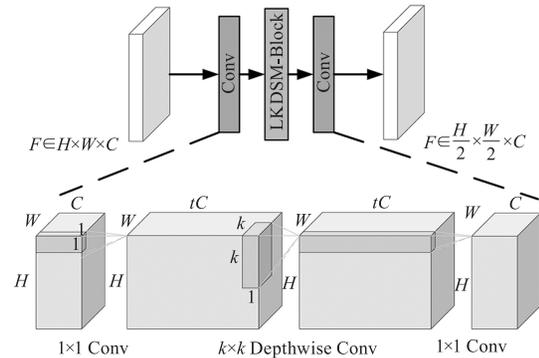


图3 大核下采样模块

Fig. 3 LK Down-sample module

2.2.2 多尺度图像重建

与多尺度特征学习部分相对应,图像重建部分包括三个尺度,并接受来自卷积得到的图像的高级语义特征和压缩因子调节器输出的参数对 (α, β) 以去除伪影,恢复高质量光场图像。其中压缩因子 (λ) 特征融合块是图像重建重要组成部分,每个尺度的压缩因子特征融合块的数量被设置为 4。通过网络学习的参数对 (α, β) 输入到每个尺度的压缩因子特征融合块中,以对每个尺度重建光场中间特征图应用仿射变换来引导光场图像进行重建,以便获得高质量的图像。在从调节器获得 (α, β) 之后,通过缩放和移位特定层的特征图来执行变换,如图 4 所示。该具体操作为:

$$f'_{out} = (1 + \alpha) \otimes f'_{in} \oplus \beta \quad (3)$$

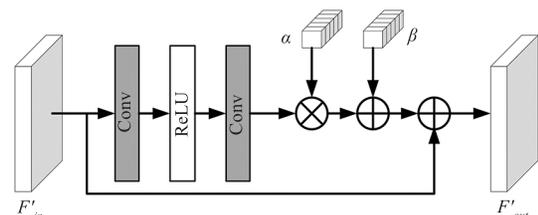


图4 压缩因子特征融合块

Fig. 4 Compression factor features fusion block

其中, f'_{in} 和 f'_{out} 分别表示经过压缩因子特征融合块前的图像特征和之后的图像特征, \otimes 为元素乘法。

给定一批内的 N 个训练样本,训练的目标是最小化重建图像 I_r 和真值图像 I_{gt} 之间的以下 L_1 损失函数:

$$L_g = \frac{1}{N} \sum_{i=1}^N \|I_R^i - I_{gt}^i\|_1 \quad (4)$$

总的来说,训练的总损失项设置为:

$$L = L_g + \sigma \cdot L_\lambda \quad (5)$$

其中, σ 控制图像重建和 λ 预测之间的平衡。 L_λ 为压缩因子估计的损失,具体在章节 2.2.3 中介绍。

2.2.3 压缩因子(λ)预测器

压缩因子预测器是一个如图 1 所示的三层线性层组成的多层感知器。它将 512 维 λ 特征作为输入,并产生压缩图像的估计压缩因子 l_{est} 。具体操作为:

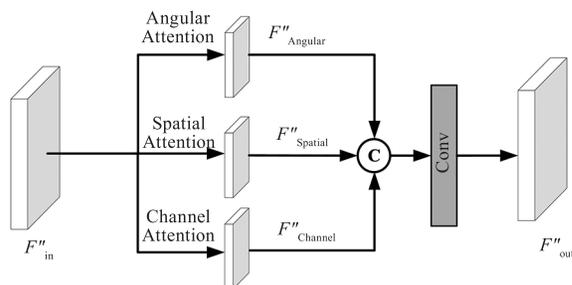
$$l_{est} = MLP(F_1) \quad (6)$$

其中,MLP 表示由三层线性层组成的多层感知器, F_1 表示从图像中提取的压缩因子特征。

为了更好地预测,本文将每个隐藏层中的节点数设置为 512。在训练过程中,小尺寸的补丁可能只包含有限的信息,并且对应于多个质量因子,使得质量因子不能被准确地估计,这可能导致不稳定的训练过程。因此,本文使用 L_1 损失函数约束训练,每个批次中压缩因子估计的损失函数为:

$$L_\lambda = \frac{1}{N} \sum_{i=1}^N \|l_{est}^i - l_{gt}^i\|_1 \quad (7)$$

其中, N 表示训练的批次数, l_{gt} 表示该 λ 的真值。



(a) 混合注意力增强模块总框架

2.2.4 压缩因子(λ)调节器

同样,压缩因子调节器是一个如图 1 所示的三层线性层组成的多层感知器。该调节器接收压缩因子预测器预测的压缩因子,并根据压缩因子学习输出调制参数对 (α, β) 的映射函数,以此来引导网络对压缩图像进行伪影去除。同时,通过压缩因子特征融合块将调制参数融入到图像重建的三个尺度中。具体操作可表示为:

$$(\alpha_i, \beta_i) = MLP(F_{est}) \quad (8)$$

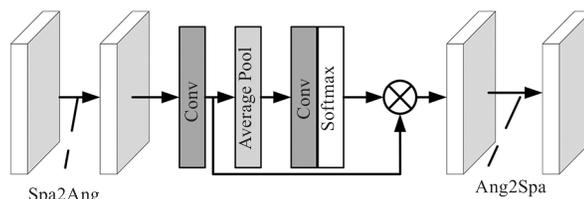
其中, F_{est} 是 l_{est} 经过预测压缩因子编码器得到的特征。 $i=1, 2, 3$, 在最后一层 MLP 中 α 的产生使用 Sigmoid 激活函数, β 的产生使用 Tanh 激活函数。

2.3 混合注意力增强模块

本文提出利用混合注意力机制重建光场的空间与角度之间的联系,以此加强光场图像的质量。混合注意力增强模块框架如图 5 所示,其中混合注意力增强模块总框架如图 5(a)所示,其包括通道注意力、空间注意力、角度注意力。具体操作为:

$$F''_{out} = Conv(Cat(F''_{Channel} + F''_{Spatial} + F''_{Angular})) \quad (9)$$

其中, $F''_{Channel}$ 、 $F''_{Spatial}$ 、 $F''_{Angular}$ 分别为通道注意力、空间注意力、角度注意力的结果。



(b) 角度注意力

图 5 混合注意力增强模块框架

Fig. 5 The framework of hybrid attention enhancement module

这里,使用文献[21]中的通道和空间注意力结构。角度注意力结构则如图 5(b)所示,首先将子孔径阵列转换为微透镜图像,其中每个宏像素的特征由其在空间中的尺寸 (x, y) 和通道数目 (C) 确定,整体空间大小为 (u, v) 。通过应用空间卷积来提取角相关性,然后沿通道轴压缩通道数目。在经过 3×3 卷积处理后,每个通道的宏像素特征被拉伸成一维数组,即被平坦化。最后,利用 Softmax 激活以生成注意力权重,以便在数据处理中对不同部分赋予不同的权重。混合注意力机制能够更好地提取压缩图像中的有效信息,更加精确地定位图像失真严重的部分,进而提高失真修复的效率和有效性。此外,其能进一步考虑光场的角度信息,更加符合光场

的结构特点。

3 实验结果与分析

3.1 数据库 JPL-DATA

由于缺少公开的 JPEG Pleno 编码失真的光场图像数据集,本文根据 JPEG Pleno 白皮书第三部分-通用测试标准^[22]及 CTC 通用测试条件^[23]的要求,使用 JPEG Pleno 官方提供的 EPLF 数据集作为数据源,建立了相应的光场编码失真数据集 JPL-DATA。由于 JPEG Pleno 在相同码率下对不同场景可能出现不同的编码效果,并且 EPLF 数据源中有些场景过于简单,所以,这里对 EPLF 数据源中的场景进行了筛选,筛选后数据集 JPL-DATA 场景如图 6 所示。



(a) JPL-DATA测试集光场图像中心SAI



(b) JPL-DATA部分训练集光场图像中心SAI

图6 JPL-DATA数据集部分光场图像中心SAI

Fig. 6 Central SAIs of some LFI in JPL-DATA dataset

具体而言,首先使用 JPEG Pleno 4DTM 对 EPFL 数据源中的所有 91 个场景进行编码。再计算出所有场景的 YUV-PSNR 与 Y-SSIM,并进行主观视觉对比实验。最后选取其中 40 个场景作为训练集,4 个场景作为测试集。所创建的失真光场其角度分辨率为 7×7 ,空间分辨率为 434×625 ,位深为 10 位。JPEG Pleno 4DTM 使用压缩因子 λ 控制压缩图像的质量,图像质量与 λ 成反比,即图像质量随 λ 增大而降低。 λ 的取值范围为 $[1, \infty]$,本文仅考虑压缩因子 $\lambda \in [5000, 200000]$ 的压缩图像,从中取 6 个失真等级的光场图像,分别为 $\lambda = 5000、10000、20000、50000、100000、200000$ 。JPL-DATA 的测试集与部分训练集光场中心 SAI 缩略图分别如图 6(a)、图 6(b)所示。

3.2 实施细节以及评价标准

在实验中,在训练过程中,使用自建数据集 JPL-DATA 作为训练数据,其中包括 40 个场景,每个场

景拥有 6 个质量等级的图像,用 Adam 优化器,学习率从 1×10^{-4} 开始,并且将 σ 设为 0.1, k 与光场的角度分辨率相同,设为 7。在一个 NVIDIA GeForce GTX 3090 GPU 上使用 PyTorch 训练模型。在测试过程中,测试评估 JPL-DATA 测试场景中 6 个质量等级的图像($\lambda = 5000、10000、25000、50000、100000、200000$)。

由于 JPEG Pleno 编码只支持 10 bit 的光场数据,JPL-DATA 数据集也均为 10 bit 光场图像。使用 CTC 通用测试条件中测试要求和相关计算指标代码计算测试图像的 YUV-PSNR 与 Y-SSIM,以此来定量评估所提出的 JPEG Pleno 伪影去除模型的性能。YUV-PSNR 与各个通道的 PSNR 之间的关系为^[23]:

$$PSNR_{YUV} = \frac{6 \times PSNR_Y + PSNR_U + PSNR_V}{8} \quad (10)$$

3.3 总体性能比较

考虑到缺少 JPEG Pleno 去伪影的相关方法进行对比,为了进行基准测试,从 2D JPEG 伪影去除方法中选择了 ARCNN^[4],DOGNet^[13]。同时,考虑到光场结构的特点,为加强对比方法的完整性,因而也选取了 DistgSSR^[20]、EPIT^[24] 两个光场超分网络进行对比。为了使网络可以兼容 JPEG Pleno 去伪影任务,这里去除掉其网络上采样层,将其角度分辨率设为 7×7 。

3.3.1 定量比较

按照 10 bit 光场图像的评估标准计算相应修复光场的 YUV-PSNR 和 Y-SSIM。表 1 展示了所提出的方法与对比方法的实验结果,性能最优为黑体加粗。从实验结果来看,所提出的方法在数据集 JPL-DATA 的不同压缩质量的光场图像修复上取得了最优的结果。

表 1 JPL-DATA 数据集上不同方法 YUV-PSNR/Y-SSIM 结果

Tab. 1 YUV-PSNR/Y-SSIM values of different algorithms on the JPL-DATA dataset

Methods	λ					
	5000	10000	25000	50000	100000	200000
JPEG Pleno	36.35/0.945	35.16/0.928	33.34/0.892	32.36/0.868	31.10/0.824	30.16/0.790
ARCNN ^[4]	36.39/0.945	35.21/0.932	33.45/0.899	32.46/0.875	31.22/0.833	30.24/0.797
DOGNet ^[13]	36.59/0.943	35.49/0.938	33.61/0.902	32.59/0.877	31.30/0.835	30.31/0.799
DistgSSR ^[20]	36.48/0.946	35.53/0.940	33.31/0.907	32.66/0.886	31.50/0.847	30.37/0.809
EPIT ^[24]	36.68/0.952	35.70/0.940	33.86/0.910	32.89/0.888	31.69/0.801	30.58/0.814
Proposed	36.99/0.955	36.01/0.943	34.25/0.917	33.25/0.896	31.97/0.860	30.85/0.823

此外,不同方法在数据集 JPL-DATA 上相对于原始解码光场图像 IC 在 YUV-PSNR 上的增益如图 7 所示。可以看到所提出的算法在 JPEG Pleno 伪影

去除任务上有明显的优势。在前两个质量分数($\lambda = 5000、10000$)下 IC 无明显失真,图像质量较高,所以在此情况下 YUV-PSNR 增益较小。而在后四个

质量分数 ($\lambda = 25000, 50000, 100000, 200000$) 下 IC 失真明显, 块效应严重, 并且随着压缩因子 λ 的增加, 光场图像失真程度越高, 图像中包含的有效信息也就越少, 所以 YUV-PSNR 的增益幅度减少。例如, 所提出的方法在 $\lambda = 200000$ 时的质量增幅比其他几个质量小。另一方面, DistgSSR 方法在不同压缩等级下的增幅呈现出 不稳定状态, 这证明光场超分网络不能直接用于 JPEG Pleno 去伪影任务, 进一步证明所提方法网络的稳定性。

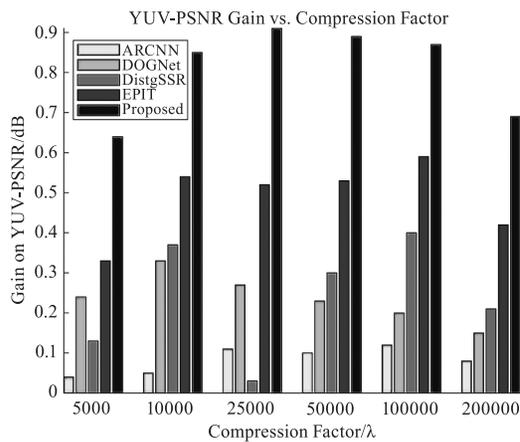


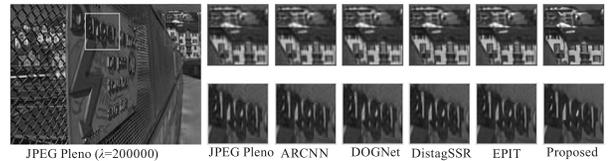
图 7 不同方法在数据集 JPL-DATA 上 YUV-PSNR 增益

Fig. 7 YUV-PSNR gain of different algorithms on JPL-DATA dataset

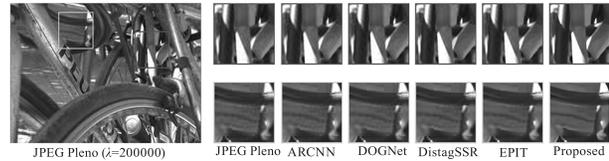
3.3.2 定性比较

图 8 展示了两个场景下使用不同方法在压缩光场图像质量最差的情况下的可视化结果, 即 $\lambda = 200000$ 时的结果。其中, “JPEG Pleno” 为原始解码图像 (I_c), 而其他则分别对应着五种去伪影方法。图 8(a) 展示了 Danger de Mort 场景测试可视化结果, 可以看到由所提出方法重建的图像的红框放大区域中窗户及屋顶处更加清晰, 纹理细节恢复得更好; 同时, 绿框内的字母相较于 I_c 以及由其他对比方法重建的图像, 显示出更加连贯的视觉效果, 割裂感减弱, 而且块效应也明显减弱, 从而大幅提升了视觉体验。此外, 图 8(b) 展示了 Bikes 场景测试可视化结果, 可以看出由所提出方法重建出的右侧柱子区域的视觉效果更加清晰, 图像质量得到提升, 同时, 可以看到木板上的线恢复更加清晰, 这也进一步说明了所提方法的优越性。

从上述两个不同场景的可视化结果中, 可以明显观察到所提出的方法相对于其他方法更有效地消除了由 JPEG Pleno 低码率压缩产生的块效应, 显著提升了重建光场图像的视觉效果。这也进一步验证了所提出的方法更符合于 JPEG Pleno 4DTM 压缩光场的特性。



(a) Danger de Mort 场景测试可视化结果



(b) Bikes 场景测试可视化结果

图 8 $\lambda = 200000$ 下中心 SAI 测试可视化结果

Fig. 8 Visualization results of central SAIs tested under $\lambda = 200000$

3.3.3 角度一致性的测试

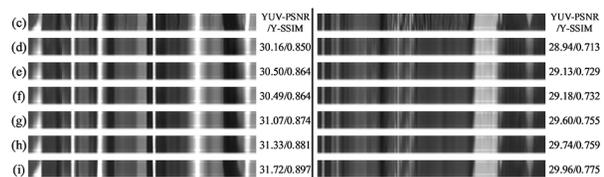
作为 4D 数据, 光场有着其独有的角度一致性特性, 能够实现数字重聚焦、全景深扩展和密集视点渲染等功能。对于光场图像修复而言, 不仅要修复出空域质量高的图像, 还需要保持光场的角度一致性。

图 9 展示了 $\lambda = 200000$ 下不同方法重建光场图像中的级平面图像 (Epipolar Plane Image, EPI)。图 9(a)、(b) 中的线段标记了所选取的 EPI 的位置, 其跨越场景中的前景与背景。图 9(c) 为这一段 EPI 的真值, 图 9(d) 为从原始解码图像 I_c 中提取的这一段 EPI 的结果, 图 9(e) 到图 9(i) 为不同对比方法的 EPI 结果, 图 9(i) 为所提方法的 EPI 结果。在 Bikes 场景中的 EPI 真值左侧有个白色的半圆弧, 只有所提方法将其恢复出来, 而该 EPI 右侧黑色的一条斜线在其他方法中已经变成竖直的直线, 表明对应区域的视差由于编码失真导致了明显的误差, 但所提方法较好地恢复出与真值 (图 9(c)) 类似的结果。类似的, Danger de Mort 场景的 EPI 真值右侧有一个倒立的三角形, 其他方法对应的 EPI 均表现为矩形, 只有所提方法与真值相接近。



(a) Bikes 场景 EPI 选取位置

(b) Danger de Mort 场景 EPI 选取位置



(c) Ground truth (d) JPEG Pleno (e) ARCNN (f) DOGNet (g) DistagSSR (h) EPIT (i) Proposed

图 9 $\lambda = 200000$ 下不同方法重建光场中的 EPI

Fig. 9 EPIs of light fields reconstructed by different methods at $\lambda = 200000$

除主观对比外,图9中也给出了这段EPI相较于其真值的YUV-PSNR与Y-SSIM。可以看到,所提出的方法取得了最好的质量效果。2D JPEG去伪影方法未考虑光场的角度信息,而光场超分方法网络不能很好地针对JPEG Pleno编码产生的伪影。所提的方法通过混合注意力增强模块增强重建光场空角特征之间的联系,进而增强了重建光场的角度一致性。

3.4 消融实验

为了评价浅层特征提取模块中的Distg-Block、深度多尺度特征解耦中的LK Down-sample模块(LKDSM)和混合注意力增强模块(HAEM)的影响,在相同的测试集上进行了消融实验。实验分三组,一组使用单纯的空间卷积代替Distg-Block;一组实验使用普通的膨胀卷积代替LK Down-sample模块;另一组实验则去除了混合注意力增强模块。

浅层特征提取模块中的Distg-Block可以充分提取光场图像的信息,使用它对压缩光场图像进行

光场特征提取是必不可少的,因为它相比单纯的空间卷积能够更好地提取压缩光场图像中的有效信息,这对后续的伪影去除和保证更加细致的纹理输出是至关重要的。如表2第一行所示,使用单纯的空间卷积代替Distg-Block,不同质量下的评价指标均有所下降,证明了使用Distg-Block作为浅层特征提取模块的必要性。

使用LK Down-sample模块进行下采样旨在有效地权衡有效信息和感受野之间的平衡,以更好地进行JPEG Pleno伪影去除。如表2第二行所示,使用普通的膨胀卷积代替LK Down-sample进行下采样,各质量评价指标亦均有不同程度的下降,证明了LK Down-sample模块的有效性。

混合注意力增强模块旨在增强重建光场的空间与角度之间的联系,从而提高重建光场的质量,减少JPEG Pleno伪影。如表2的第三行所示,在移除混合注意力增强模块后,所有质量评价指标都表现出降低。这一表现证明了该模块的有效性。

表2 消融实验YUV-PSNR/Y-SSIM的结果

Tab.2 YUV-PSNR and Y-SSIM results of the ablation experiments

CONFIG	Distg-Block	LKDSM	HAEM	λ					
				5000	10000	25000	50000	100000	200000
(Scheme I)	×	√	√	36.48/0.950	35.53/0.940	33.84/0.910	32.87/0.888	31.66/0.852	30.63/0.815
(Scheme II)	√	×	√	36.65/0.952	35.73/0.943	34.01/0.914	33.04/0.891	31.70/0.855	30.60/0.814
(Scheme III)	√	√	×	36.78/0.954	35.74/0.943	34.07/0.915	33.09/0.893	31.72/0.856	30.63/0.820
Proposed	√	√	√	36.99/0.955	36.01/0.943	34.25/0.917	33.25/0.896	31.97/0.860	30.85/0.823

4 结论

本文提出了一种新的端到端JPEG Pleno伪影去除网络(JPLARNet)。通过卷积从编码失真图像中预测出压缩因子,然后通过压缩因子特征融合模块将压缩因子映射为特征引导图像重建,以生成高质量图像。同时,使用空角解耦模块在不破坏光场本身结构的情况下初步提取失真图像中的有效信息。在网络设计中还考虑了有效信息与感受野大小之间的权衡。此外,在重建图像过程中使用混合注意力增强模块以进一步增强重建图像的质量,加强光场空间与角度之间的联系。实验结果表明,所提出的方法在数据集JPL-DATA六个不同失真等级($\lambda = [5000, 10000, 25000, 50000, 100000, 200000]$)上其YUV-PSNR/Y-SSIM相较于未进行伪影去除的解码重建光场图像平均增益为0.81dB/0.025,相比其他方法取得了更好的效果。

参考文献:

- [1] Xiang J J, Jiang G Y, Yu M, et al. No-Reference light field image quality assessment using Four-Dimensional sparse transform[J]. IEEE Transactions on Multimedia, 2023, 25: 457 - 472.
- [2] Jamil S, Piran M J, Rahman M U, et al. Learning-driven lossy image compression: a comprehensive survey[J]. Engineering Applications of Artificial Intelligence, 2023, 123: 106361.
- [3] Gao W, Fan S L, Li G, et al. A thorough benchmark and a new model for light field saliency detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(7): 8003 - 8019.
- [4] Dong C, Deng Y B, Loy C C, et al. Compression artifacts reduction by a deep convolutional network[C]//Proceedings of the IEEE International Conference on Computer Vision. Santiago: CVF, 2015: 576 - 584.

- [5] Zheng B L, Sun R, Tian X, et al. S-Net: a scalable convolutional neural network for JPEG compression artifact reduction [J]. *Journal of Electronic Imaging*, 2018, 27(4):043037.
- [6] Li B, Shi Y, Wang B, et al. RGSr: a two-step lossy JPG image super-resolution based on noise reduction [J]. *Neurocomputing*, 2021, 419:322 – 334.
- [7] Amaranageswarao G, Deivalakshmi S, Ko S B. Joint restoration convolutional neural network for low-quality image super resolution [J]. *The Visual Computer*, 2022, 38(1): 31 – 50.
- [8] Sun M D, He X H, Xiong S H, et al. Reduction of JPEG compression artifacts based on DCT coefficients prediction [J]. *Neurocomputing*, 2020, 384:335 – 345.
- [9] Yang G, Lin L, Wu C, et al. Dual-Domain learning for JPEG artifacts removal [C]//*Neural Information Processing*. Singapore: Springer, 2023:556 – 568.
- [10] Galteri L, Seidenari L, Bertini M, et al. Deep universal generative adversarial compression artifact removal [J]. *IEEE Transactions on Multimedia*, 2019, 21(8):2131 – 2145.
- [11] Kim Y, Soh J W, Park J, et al. A Pseudo-Blind convolutional neural network for the reduction of compression artifacts [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2020, 30(4):1121 – 1135.
- [12] Ehrlich M, Davis L, Lim S N, et al. Quantization guided JPEG artifact correction [C]//*Computer Vision – ECCV 2020*. Cham: Springer, 2020:293 – 309.
- [13] Yoon J, Cho N I. JPEG artifact reduction based on deformable offset gating network controlled by a variational autoencoder [J]. *IEEE Access*, 2023, 11:30282 – 30291.
- [14] Kawar B, Song J, Ermon S, et al. Jpeg artifact correction using denoising diffusion restoration models [DB/OL]. (2022 – 09 – 23). <https://arxiv.org/abs/2209.11888>.
- [15] De Carvalho M B, Pagliari C L, De O. E. Alves G, et al. Supporting wider baseline light fields in JPEG pleno with a novel slanted 4D – DCT coding mode [J]. *IEEE Access*, 2023, 11:28294 – 28317.
- [16] Li Y, Mathew R, Taubman D. JPEG pleno light field encoder with mesh based view warping [C]//*2023 IEEE International Conference on Image Processing (ICIP)*. Kuala Lumpur: IEEE, 2023:2100 – 2104.
- [17] Stepanov M, Valenzise G, Dufaux F. Compression of light fields [M]//Valenzise G, Alain M, Zeman E, et al. *Immersive Video Technologies*. Amsterdam: Academic Press, 2023:201 – 226.
- [18] Astola P, Da S C L A, Da S E A B, et al. JPEG pleno: standardizing a coding framework and tools for plenoptic imaging modalities [J]. *ITU Journal: ICT Discoveries*, 2020, 3(1).
- [19] Wallace G K. The JPEG still picture compression standard [J]. *Communications of the ACM*, 1991, 34(4):30 – 44.
- [20] Wang Y Q, Wang L G, Wu G C, et al. Disentangling light fields for super-resolution and disparity estimation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 45(1):425 – 443.
- [21] Song C H, Han H J, Avrithis Y. All the attention you need: global-local, spatial-channel attention for image retrieval [C]//*Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. Waikoloa: CVF, 2022:2754 – 2763.
- [22] ISO/IEC. Information technology—Plenoptic image coding system (JPEG Pleno)-Part 3: Conformance testing: ISO/IEC 21794 – 3; 2021 [R/OL]. <https://www.iso.org/standard/74533.html>.
- [23] Pereira F, Pagliari C, da Silva E A B, et al. JPEG Pleno light field coding common test conditions v3.2 [J]. *Doc. ISO/IEC JTC*, 2019, 1.
- [24] Liang Z Y, Wang Y Q, Wang L G, et al. Learning Non-Local Spatial-Angular correlation for light field image Super-Resolution [C]//*Proceedings of the IEEE/CVF International Conference on Computer Vision*. Paris: CVF, 2023: 12376 – 12386.