

基于自监督学习的光场空间域超分辨成像

梁丹, 张海苗, 邱钧*

北京信息科技大学理学院, 北京 100101

摘要 针对光场成像的空间域图像分辨率限制, 提出一种基于自监督学习的光场空间域超分辨成像方法。利用深度学习中的自编码器, 对全部光场子孔径图像同步进行空间域的超分辨重构。设计一种基于多尺度特征结构和全变差正则化的混合损失函数, 约束模型输出图像与原始低分辨率图像的相似度。数值实验结果表明, 所提方法对噪声具有抑制作用, 在光场成像的不同数据集上的超分辨结果平均值超过基于有监督学习的光场空间域超分辨方法。

关键词 光场; 超分辨; 自监督学习; 深度学习

中图分类号 TP391 文献标志码 A

DOI: 10.3788/LOP231188

Self-Supervised Learning for Spatial-Domain Light-Field Super-Resolution Imaging

Liang Dan, Zhang Haimiao, Qiu Jun*

School of Applied Science, Beijing Information Science and Technology University, Beijing 100101, China

Abstract This paper proposes a self-supervised learning-based method for the super-resolution imaging of spatial-domain resolution-limited light-field images. Using deep learning self-encoding, a super-resolution reconstruction of the spatial-domain is performed simultaneously for all light field sub-aperture images. A hybrid loss function based on multi-scale feature structure and total variation regularization is designed to constrain the similarity of the model output image to the original low-resolution image. Numerical experiments show that the newly proposed method has a suppressive effect on noise, and the resultant average super-resolutions for different light field imaging datasets exceed those of the supervised learning-based method for light field spatial domain images.

Key words light field; super-resolution; self-supervised learning; deep learning

1 引言

光场相机获取的光场数据存在空间分辨率与角度分辨率之间制约的限制, 通过计算成像的方法利用光场数据在空间和角度域的约束信息构建数学模型和算法, 可以将光场在空间域、角度域或空-角域联合进行超分辨成像^[1-4]。本文针对光场的空间域超分辨率成像问题进行研究, 利用深度学习中的生成模型实现光场超分辨重构。

现有关于光场超分辨成像的技术手段可分为 3 大类: 基于几何投影的方法、基于优化的方法和基于深度学习的方法^[5]。

基于几何投影的方法主要依据光场相机的成像原理, 通过获取不同视角子孔径图像之间的视差(几何

信息对目标视图进行超分辨。Lim 等^[6]将二维子孔径图像在空间维度上的亚像素位移投影到凸集上来获取高分辨率图像。Nava 等^[7]利用重聚焦原理, 将其他视图的像素投影到中心视图, 得到当前场景的超分辨率全聚焦图像。

光场超分辨率重构的优化方法利用 4D 光场结构信息和对实际拍摄场景的先验假设来构建成像的物理模型, 将问题转化为对目标泛函进行优化求解的数学问题。Wanner 等^[8-9]使用结构张量方法从极平面图像(EPIs)估计视差图, 将基于深度的光场重构问题表述为基于全变差正则化的能量最小化问题, 对估计的低分辨率视差图进行上采样, 来实现对光场的空间域和角度域超分辨重构。Rossi 等^[10]利用不同光场视图信息并结合图正则化器来对光场的几何结构进行约束实

收稿日期: 2023-04-27; 修回日期: 2023-06-07; 录用日期: 2023-06-28; 网络首发日期: 2023-07-08

基金项目: 国家自然科学基金(12101061, 61931003)

通信作者: *qiu jun@bistu.edu.cn

现光场的空间超分辨重构。

基于深度学习的光场超分辨成像技术在近几年取得了令人瞩目的进展。Yoon 等^[11]提出基于卷积神经网络(CNN)的超分辨方法 LFCNN,首次将深度卷积神经网络用于光场的空间域和角度域超分辨,采用有监督学习的方法获得高分辨率的空间和角度超分辨光场数据。Zhang 等^[12]提出一种基于残差结构的卷积神经网络(ResLF)用于光场空间域子孔径图像超分辨。Liang 等^[13]通过设计角度和空间 Transformers,充分提取光场的有效信息,实现光场角度域和空间域超分辨重构。为了有效利用耦合在空间域和角度域中的视差信息,Wang 等^[14]设计了空间、角度、极平面等 3 种解耦卷积运算,并基于解耦机制设计了光场空间域超分辨重构神经网络模型 DistgSSR。

上述基于深度学习的光场超分辨重构方法均为有监督学习的方法,模型的超分辨性能依赖于在大量的低分辨率和高分辨率光场数据对所构成的数据集上的训练结果。实际应用中,采集大量场景的高分辨率光场数据较困难,并且在光场相机成像参数和场景有限的条件下训练获取的深度学习模型在新场景和新成像参数下的泛化性能较弱。针对这些问题,大量关于单张图像超分辨的自监督学习模型被提出^[15-18],其中,Lempitsky 等^[16]以神经网络结构本身作为图像表示学习的先验,不需要大规模训练数据集,仅以原始低分辨率图像作监督即可实现单张图像超分辨率成像。

受 deep image prior(DIP)^[16]启发,本文将自监督学习方法引入光场数据的空间域超分辨重构任务,在神经网络模型架构设计时引入多尺度结构,并由此定

义多尺度损失函数,与全变差正则化和均方误差损失函数结合,设计出一种新的基于自监督学习的光场空间域超分辨方法,记为 DIP-LFSR。所提方法基于自监督学习对光场空间域进行超分辨率成像,不依赖大规模的成对训练数据集,采用多尺度神经网络结构提取各子孔径图像中不同尺度的图像特征^[19],构建多尺度损失函数,使模型在重构高分辨率的子孔径图像时能够利用耦合在子孔径图像中的视差信息,在多个添加噪声的数据集上数值结果和视觉效果均优于普通插值方法和有监督方法。

2 基于自监督学习的光场空间域超分辨成像

DIP-LFSR 是一种生成模型,利用神经网络来表示从低维输入数据到高维的高分辨率图像之间的映射关系。这种基于神经网络的数据表示方式本身包含良好的结构先验,能有效捕获图像中的低频信息,神经网络的强大表示能力使得模型经过训练后能够有效表示图像中的高频信息^[20]。DIP-LFSR 对神经网络生成的高分辨率图像进行降采样获得与原始采集到的低分辨率图像同样维度的数据,并将其作为模型训练的监督信息,这种基于自监督学习的方法将超分辨图像重构问题转换化成了一个有条件的图像生成问题。

实现光场空间域超分辨重构的 DIP-LFSR 方法主要包含两部分:1)获取高分辨率光场生成模型的网络架构;2)自监督学习的损失函数。

2.1 超分辨模型框架

图 1 展示了基于自监督学习的光场空间域超分辨

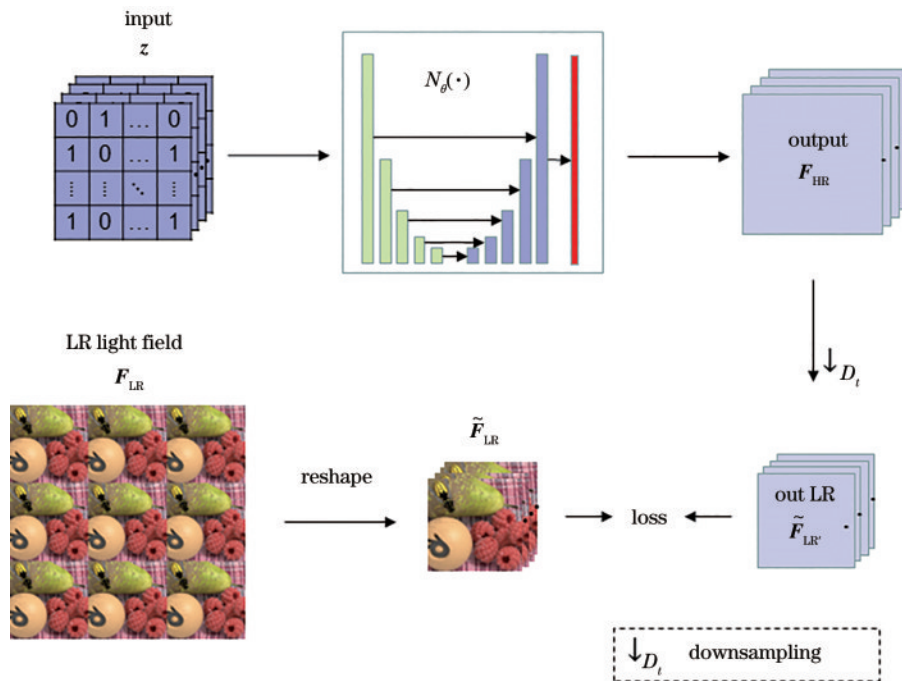


图 1 基于自监督学习的光场空间域超分辨流程图

Fig. 1 Flow chart of light field spatial-domain super-resolution based on self-supervised learning

率重构的流程。网络的输入是一个随机的编码向量,网络的输出是高分辨率的光场子孔径图像数据堆栈(子孔径图像沿 RGB 通道维度堆叠),训练网络仅需要原始的低分辨率数据作为监督信息。该流程图展示了神经网络将随机编码数据映射到高分辨率光场数据的过程,然后用采集到的低空间分辨率的光场数据定义损失函数,由此监督模型的学习过程。图 1 体现了采用自监督学习方法实现光场空间域超分辨的整体思路。与其他监督学习相比,用自监督学习的方法对光场空间域进行超分辨重构不需要大规模的成对训练数据,如低空间分辨率光场-高空间分辨率光场数据,训练光场超分辨网络时只需要以原始的低分辨率光场数据堆栈作为监督数据即可实现光场的空间域超分辨。该流程的具体细节如下:

采用双平面法表示光场数据,对于采集到的低分辨率光场 $\mathbf{F}_{\text{LR}} \in \mathbf{R}^{3 \times U \times V \times H \times W}$,其角度域的数据维度为 $U \times V$,空间域各子孔径图像(彩色图像)的维度为 $3 \times H \times W$ 。为了对光场数据 \mathbf{F}_{LR} 的空间维度进行 t 倍超分辨重构,以随机编码向量 $\mathbf{z} \in \mathbf{R}^{C \times H \times tW}$ 作为编解码器神经网络 $N_\theta(\cdot)$ 的输入,通过神经网络生成高分辨率的光场子孔径图像堆栈 $\mathbf{F}_{\text{HR}} \in \mathbf{R}^{3 \times UV \times tH \times tW}$ 的过程如下:

$$\mathbf{F}_{\text{HR}} = N_\theta(\mathbf{z}), \quad (1)$$

式中:编码向量 \mathbf{z} 中的元素从标准正态分布中采样获得,即 $z_{i,j,k} \sim N(0, 1)$, $i = 1, \dots, C, j = 1, \dots, tH, k = 1, \dots, tW$;编解码器神经网络 $N_\theta(\cdot)$ 的参数集合为 θ 。

为了利用采集到的低分辨率光场数据 $\mathbf{F}_{\text{LR}} \in \mathbf{R}^{3 \times U \times V \times H \times W}$ 对生成模型 $N_\theta(\cdot)$ 的训练过程进行监督,将其在角度域中的 UV 个子孔径图像进行通道叠加,得到低分辨率的子孔径图像堆栈 $\tilde{\mathbf{F}}_{\text{LR}} \in \mathbf{R}^{3UV \times H \times W}$ 。将神经网络 $N_\theta(\cdot)$ 生成的高分辨率光场子孔径图像堆栈 \mathbf{F}_{HR} 在空间维度进行 t 倍下采样,得到模拟的低分辨率子孔径图像堆栈 $\tilde{\mathbf{F}}_{\text{LR}'}$,即

$$\tilde{\mathbf{F}}_{\text{LR}'} = D_t(\mathbf{F}_{\text{HR}}), \quad (2)$$

式中: $D_t(\cdot)$ 表示对输入数据 \mathbf{F}_{HR} 在空间维度进行 t 倍下采样。则模型的损失函数为

$$E(\mathbf{F}_{\text{HR}}, \tilde{\mathbf{F}}_{\text{LR}}) = L_{\text{MSE}}(\mathbf{F}_{\text{HR}}, \tilde{\mathbf{F}}_{\text{LR}}) = \frac{1}{3UV} \|D_t(\mathbf{F}_{\text{HR}}) - \tilde{\mathbf{F}}_{\text{LR}}\|^2. \quad (3)$$

为了提高模型的稳定性和抗噪声干扰能力,在损失函数中引入正则化项,则光场超分辨率重构任务可表述为如下优化问题:

$$\mathbf{F}_{\text{HR}}^* = \underset{\mathbf{F}_{\text{HR}}}{\operatorname{argmin}} E(\mathbf{F}_{\text{HR}}, \tilde{\mathbf{F}}_{\text{LR}}) + R(\mathbf{F}_{\text{HR}}), \quad (4)$$

式中: $E(\mathbf{F}_{\text{HR}}, \tilde{\mathbf{F}}_{\text{LR}})$ 为数据保真项; $R(\mathbf{F}_{\text{HR}})$ 为正则化项,在图像重构任务中常见的图像先验信息有图像的光滑性、稀疏表示、低秩表示^[21]。

将式(3)中的目标函数作为自监督学习神经网络的损失函数,得到的高分辨率的光场数据堆栈 \mathbf{F}_{HR}

经通道维度的重排后即具有 t 倍空间域分辨率的光场数据,记为 \mathbf{F}_{HR}^* 。当 $L_{\text{MSE}}(\mathbf{F}_{\text{HR}}, \tilde{\mathbf{F}}_{\text{LR}})$ 趋于 0 时,由高空间分辨率的子孔径图像经下采样后得到的模拟低分辨率图像与采集到的低分辨率子孔径图像的拟合程度最高,此时带来的问题是:高分辨率光场拟合到了低分辨光场数据中的噪声。为了减轻这种过拟合现象,可以采取两种策略:1)设计适当的停机准则,即仅对神经网络模型训练有限步;2)在损失函数中引入正则项 $R(\mathbf{F}_{\text{HR}})$ 。

Hessian 正则化是目前较好的高阶正则化方法,但是由于 Hessian 正则项的高度非线性和不可微性,图像去模糊和去噪过程耦合度高,求解算法的复杂度高^[22-23]。本研究选择深度学习中常用的一种正则项——全变差(TV)正则化来对神经网络生成的高分辨率图像的分布进行约束。Rudin 等^[24]的研究表明:受噪声污染的图像的全变差比无噪声图像的全变差明显更大,最小化全变差正则化损失函数可以抑制图像超分辨重构带来的噪声,从而获得较为平滑的输出图像。由此构建的光场空间域超分辨重构模型不仅利用了神经网络强大的低频和高频信息表示能力,而且通过在损失函数中加入全变差正则化 L_{TV} 使图像在平滑区域的特征保持一定的光滑性。

对于一组光场数据 \mathbf{f} ,全变差正则化损失函数为

$$L_{\text{TV}}(\mathbf{f}) = \frac{1}{3UV} \sum_{u=1}^U \sum_{v=1}^V (\|\nabla_x \mathbf{f}_{uv}\|^2 + \|\nabla_y \mathbf{f}_{uv}\|^2)^{\beta/2}, \quad (5)$$

式中: $\nabla_x \mathbf{f}_{uv}$ 表示 (u, v) 视点的子孔径图像 \mathbf{f}_{uv} 在 x 轴方向上的梯度; $\nabla_y \mathbf{f}_{uv}$ 表示 \mathbf{f}_{uv} 在 y 轴方向上的梯度。数值实验中关于上述 x 和 y 方向离散形式梯度的计算采用前向欧拉差分格式。根据 Mahendran 等^[25]的分析,当 $\beta > 1$ 时,可以去掉图像中的伪影。本研究在数值实验部分比较不同取值的 β 对光场空间域超分辨重构质量的影响。

2.2 多尺度网络结构

在图 1 中,用于生成高分辨率图像堆栈的生成网络 $N_\theta(\cdot)$ 采用 U-Net 架构^[26] 作为骨干网络,对 U-Net 的架构进行改进使其适用于光场的空间域超分辨成像任务,其结构如图 2 所示。该网络主要包含 3 大模块:编码模块(d_i)、跳跃连接(skip connection)模块(s_i)和解码模块(u_i)。编码模块和解码模块都为 4 个,编码模块的每一个子模块的结构为 Conv $3 \times 3 \rightarrow$ BN \rightarrow ReLU \rightarrow Conv $3 \times 3 \rightarrow$ BN \rightarrow ReLU,其中,Conv 3×3 使用尺寸为 3×3 的卷积核,BN 为归一化层^[27],激活函数选用 ReLU^[28]。解码模块的结构为 BN \rightarrow Conv $3 \times 3 \rightarrow$ BN \rightarrow ReLU \rightarrow Conv $1 \times 1 \rightarrow$ BN \rightarrow ReLU,这里的 Conv 1×1 包含 4 个尺寸为 1×1 的卷积核,最后一层为上采样操作。网络的 5 个跳跃连接模块结构为 Conv $1 \times 1 \rightarrow$ ReLU \rightarrow BN,网络的最后一层为解码模块的前 7 层再加上 Conv $1 \times 1 \rightarrow$ Sigmoid。

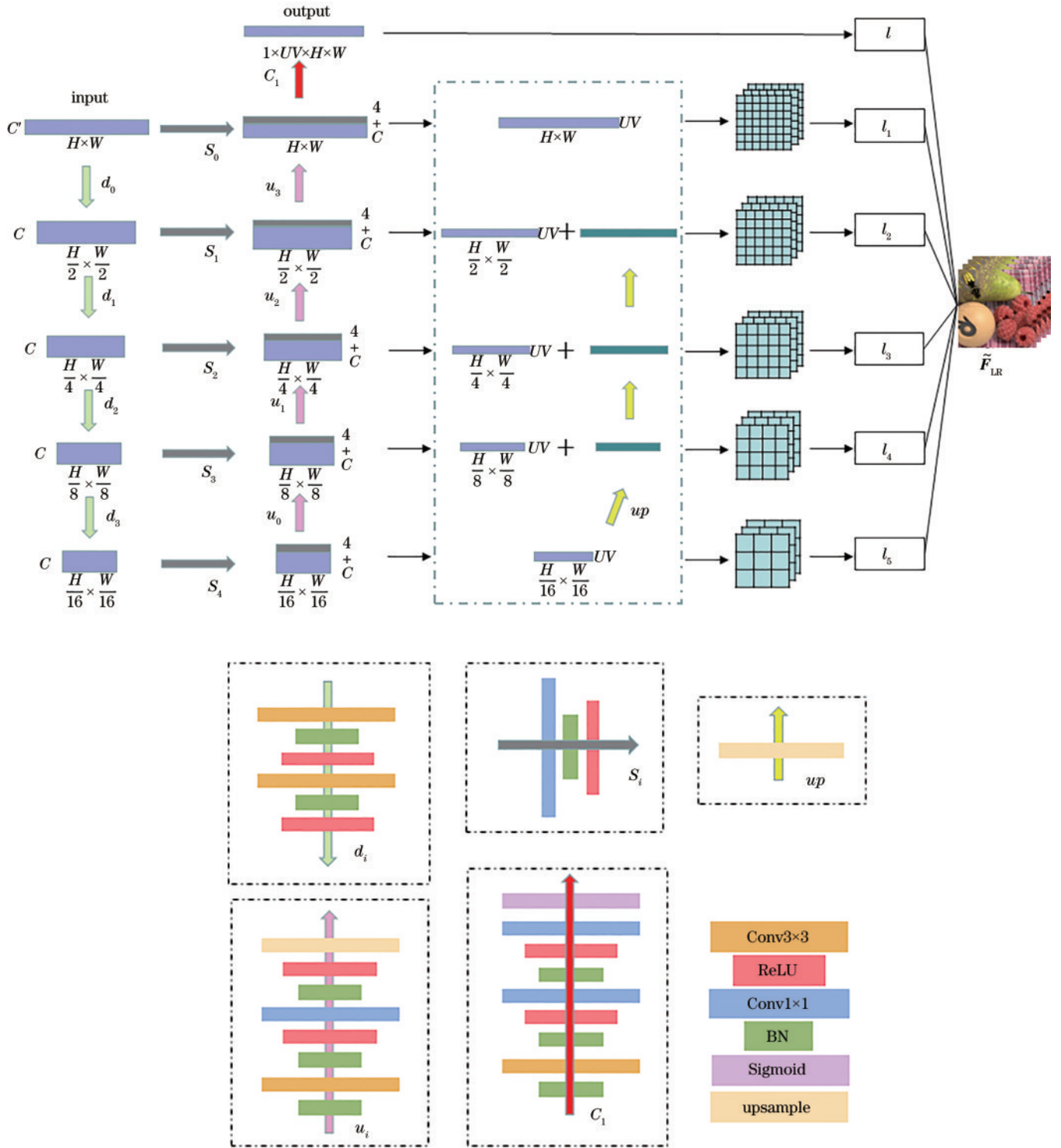


图 2 基于多尺度损失的 U-Net 结构图

Fig. 2 U-Net structure diagram based on multi-scale loss

编码模块 $\{d_i\} (i=0, 1, 2, 3)$ 对输入的随机编码向量 z 顺次进行下采样, 得到 5 个不同尺度的编码特征图, 各层编码数据通过跳跃连接模块 $\{s_i\} (i=0, 1, \dots, 4)$ 与解码模块 $\{u_i\} (i=0, 1, 2, 3)$ 连接。编码模块最终生成的特征图作为跳跃连接模块的输入, 将跳跃连接模块的输出与相同维度的编码模块生成的特征图进行通道融合, 可以将编码层的信息融入解码层; 解码模块对通道融合之后的输出逐级进行上采样; 将

解码模块的每一层单独输出获得 6 个不同尺度的图像特征序列, 最后一层为最终的输出 (output)。多尺度的结构如图 2 第 1 行整体模型中的虚线框所示, 先将解码模块输出的第 1 个 $3UV$ 通道数的特征图进行上采样, 然后与浅层的特征图逐像素相加, 继续重复以上操作可以得到 5 个不同尺度的输出光场数据堆栈。这些不同尺度的图像堆栈将会用于下一小节中的多尺度损失函数设计。

2.3 损失函数设计

采用最小均方误差(MSE)作为损失函数训练获得的图像容易出现边缘模糊或者过于平滑的问题,为了提升重构高分辨率光场数据的质量,对于损失函数进行改进。

为了利用神经网络中间层不同尺度特征图中的信息,将图像分割任务中的多尺度特征提取思想^[29]用于定义光场空间域超分辨成像的多尺度损失函数,这种多尺度思想在图像去模糊的任务中表现了良好的性能^[30]。以MSE损失函数为基础构建多尺度损失 L_{MS} ,约束模型从子孔径图像中隐式地提取视差信息,从而提升重构图像质量。

根据在U-Net加入的多尺度结构中不同层输出的图像特征定义多尺度损失函数^[29],具体为:将U-Net的解码器中每一层输出作为多尺度的超分辨重构图像的预测结果,都与已有的低分辨率图像构建损失,由此定义模型损失函数。在没有引入多尺度损失之前,普通的DIP方法进行光场超分辨任务时构建的损失就是网络最终的输出 \mathbf{F}_{HR} 经过下采样与 $\tilde{\mathbf{F}}_{LR}$ 构建的损失 $L_{MSE}(\mathbf{F}_{HR}, \tilde{\mathbf{F}}_{LR})$ 。引入多尺度损失以后, $l_i (i=1, 2, 3, 4, 5)$ 为网络中的每一层输出分别与处理之后的 $\tilde{\mathbf{F}}_{LR}$ 构建的损失:

$$l_i = \frac{1}{3UV} \text{MSE} [D_{m_i}(\mathbf{F}_{out_i}), D_{n_i}(\tilde{\mathbf{F}}_{LR})], \quad (6)$$

式中: \mathbf{F}_{out_i} 为网络多尺度结构的第*i*层的输出; D_{m_i} 为对网络的每一层输出进行 m_i 倍的下采样操作; D_{n_i} 为对原始的低分辨率的空间域数据堆栈 $\tilde{\mathbf{F}}_{LR}$ 进行 n_i 倍的下采样操作。具体的下采样倍数由超分辨任务决定,将网络输出与原始低分辨率图像堆栈处理之后结合起来定义多尺度损失函数。

为了调整不同尺度空间中图像堆栈在损失函数中的贡献,对各层的损失函数 l_i 引入权重。由于各层的网络输出对于超分辨重构图像的影响不同,越靠近最终层的输出包含的信息越丰富,对结果的影响越大。在光场空间域超分辨任务中,采用如下定义的多尺度

损失函数:

$$L_{MS} = \lambda_i^6 - l_i^2, \quad (7)$$

式中: λ_i 为权重,通过对不同输出层的损失函数赋值不同权重来调整多尺度特征损失函数对超分辨率图像重构的影响。当 $\lambda_i = 0 (i=1, 2)$, $\lambda_i = \frac{1}{3UV} (i=3, 4, 5)$ 时,多尺度损失函数充分利用不同尺度的信息,可以同时优化所有层的损失。

综合以上的MSE损失函数、全变差损失函数和多尺度损失函数,DIP-LFSR模型的训练采用如下损失函数:

$$L = L_{MSE} + \alpha_1 L_{MS} + \alpha_2 L_{TV}, \quad (8)$$

式中:系数 $\alpha_1, \alpha_2 > 0$ 。为了调整不同损失函数的权重:当 $\alpha_1 = \alpha_2 = 0$ 时, L 退化为原始的MSE损失;当 $\alpha_1 = 1, \alpha_2 = 0$ 时, L 退化为MSE损失加多尺度损失。由于全变差正则化损失函数会使得复原的图像过于光滑,重构图像中的纹理细节丢失,为此数值实验中,取 $\alpha_2 = 0.001$ 。当 $\alpha_1 = 1, \alpha_2 = 0.001$ 时, L 退化为MSE损失、多尺度损失和全变差正则化损失。

3 数值实验

3.1 实验设定

3.1.1 数据集

数值实验使用合成和真实场景采集的光场数据集来评估DIP-LFSR模型在光场空间域超分辨重构任务上的效果,所采用的数据集包括:HCI(new)^[31]、HCI(old)^[32]、Stanford^[33]、EPFL^[34]、Synthetic^[32]、INRIA_Lytro^[35]。每个数据集均选取两个场景(scene)的光场数据用于测试。如果没有特殊说明,每个场景都选取9个视角($U=3, V=3$)的子孔径图像。为了适应所采用的U-Net中下采样操作的次数,将每个场景的图像裁剪为尺寸为16的整数倍的图像(ground truth),并使用lanczos2方法^[36]对高分辨率光场数据分别下采样2倍、4倍和8倍来获得模拟的低分辨率光场子孔径图像。表1展示了所选用的6个测试光场数据集的场景以及裁剪之后的不同场景图像的尺寸(size)。

表1 不同测试数据集、场景及光场子孔径图像的尺寸

Table 1 Dimensions of different test data sets, scenes and light field sub-aperture images

Dataset	HCI(new)		HCI(old)		Stanford	
Scene/size	herbs	512×512	stilllife	512×512	lego truck	512×512
	origami		medieval		lego knights	
Dataset	EPFL		INRIA_Lytro		Synthetic	
Scene/size	books	432×432	bee_1	432×432	fishi_camera	512×512
	graffiti		building		shrubbery_camera	

3.1.2 模型训练设定与对比方法

DIP-LFSR的输入数据 $\mathbf{z} \in \mathbf{R}^{C' \times U \times V}$ 的通道数 C' 为32,网络中间层的通道数 C 为256,下采样层采用lanczos2方法,上采样层使用双线性插值(Bilinear)方

法。神经网络的整体设计基于PyTorch框架、Python编程语言。硬件计算设备配置为RTX 3090(24 GB)的GPU。采用Adam优化器^[37],学习率的初始值设为0.0001,学习率调节器采用指数型下降的形式^[38]。

在上述模型参数设定下进行对比实验,通过对不同方法在不同参数设置下对光场图像空间域的超分辨率结果,验证 DIP-LFSR 的有效性,对比方法如下:

1) 插值方法 Bicubic。传统的插值方法有邻近插值、双线性插值和双三次插值,其中,双三次插值的超分辨率结果比其他两种插值方法保持更平滑的图像边缘,因此选用双三次插值方法 Bicubic 与所提方法进行对比。

2) DIP。将原始的自监督学习的方法 DIP 记为基准算法。

3) DIP-LFSR 方法。DIP-LFSR 进行光场图像超分辨率时对所有视点图像同时超分辨率,只需要进行一次超分辨率就能得到整个光场图像的超分辨率结果,所需要的时间更短,同时图像空间域和角度域信息相互耦合,因此可用信息更多。

4) DistgSSR^[14]方法。Wang 等^[14]所提出的有监督学习的方法,需要使用大量成对的数据集进行训练。

5) DIP-SAISR 方法。用最基本的方法对光场空间域进行超分辨率重构,对每一张子孔径图像逐张用 DIP 超分辨率,记为 DIP-SAISR。

3.1.3 量化评价指标

选用图像量化评价指标峰值信噪比(PSNR)^[39]、structural similarity(SSIM)^[40]和 LPIPS^[41]来对超分辨率重构的光场质量进行评估。PSNR 用来评价两组图像像素值差异,对于超分辨率重构光场数据 x 和真实光场数据 y ,PSNR 的表达式为

$$Q_{\text{PSNR}}(x, y) = 10 \cdot \log_{10} \left[\frac{I_{\text{max}}^2}{\text{MSE}(x, y)} \right], \quad (9)$$

式中: I_{max} 为光场数据像素的最大值。

SSIM 用来衡量两幅图像的结构相似程度,定义为

$$Q_{\text{SSIM}}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}, \quad (10)$$

式中: μ_x, μ_y 分别为 x, y 的均值; σ_x^2, σ_y^2 分别为 x, y 的方差; $c_1 = (K_1L)^2, c_2 = (K_2L)^2, L$ 为像素值的最大值, $K_1 = 0.01, K_2 = 0.03$ 。

LPIPS 用来评价两幅图像的感知相似度。

3.2 实验结果

3.2.1 不同方法对比

本小节展示不同光场空间域超分辨率重构方法在不同高斯噪声水平(标准差记为 σ)的光场数据集上实现超分辨率 2 倍、4 倍和 8 倍的平均量化评价结果和视觉效果。

3.2.1.1 评价指标定量结果

通过数值实验来对 DIP-LFSR 与有监督学习方法 DistgSSR^[14]和传统的插值方法进行对比。具体比较的方法有: Bicubic、DIP、DIP-LFSR 和 DistgSSR。表 2 展示了噪声水平为 0.05 时,不同方法在 12 个场景的光场数据集上进行 2 倍、4 倍和 8 倍空间域超分辨率重构的量化评价结果,包括 PSNR、SSIM 和 LPIPS。表 3 以 2 倍光场空间域超分辨为例,分别展示每个数据集的评价指标。同一实验设定下的最佳结果加粗突出显示,次佳结果以斜体显示。

表 2 $\sigma = 0.05$ 时,不同方法在不同数据集上对光场空间域分别进行 2 倍、4 倍和 8 倍超分辨率重构的平均结果

Table 2 When $\sigma = 0.05$, the average results of different methods for 2 times, 4 times and 8 times super-resolution reconstruction of the light field space domain on different datasets

σ	Steps	Scale	Method	PSNR	SSIM	LPIPS
0.05	1500	$\times 2$	Bicubic	25.6376	0.8698	0.2563
			DIP	27.3224	0.9797	0.0181
			DIP-LFSR	28.0067	0.9826	0.0134
		DistgSSR	24.8324	0.9664	0.0238	
		Bicubic	23.4645	0.8204	0.4873	
		DIP	24.8368	0.9643	0.0351	
	$\times 4$	DIP-LFSR	25.2508	0.9678	0.0293	
		DistgSSR	23.1276	0.9505	0.0380	
		Bicubic	21.4069	0.7572	0.5743	
	1000	$\times 8$	DIP	22.8220	0.9317	0.0591
			DIP-LFSR	22.8236	0.9319	0.0586
			DistgSSR	21.2621	0.9050	0.0665

从表 2 可以看出:在低分辨率光场含有噪声的情况下,在 2 倍、4 倍和 8 倍光场空间域超分辨率任务上,DIP-LFSR 方法的性能最好。表 3 结果表明:进行 2 倍超分辨率时,在 6 个数据集上,DIP-LFSR 的 3 种量化评

价指标结果均高于其他方法,其中,有监督方法的性能最差,甚至没有超过插值方法。这是由于有监督学习的模型在无噪声的光场数据上进行训练时,不能较好地泛化至含噪声的光场数据,如需获得良好的光场超分

表 3 以 2 倍超分辨为例,不同方法在每个数据集上超分辨的量化评价指标结果

Table 3 Taking 2 times super-resolution as an example, the quantitative evaluation index results of different methods super-resolution on each dataset

Method	Parameter	Dataset					
		HCI(new)	HCI(old)	Stanford	EPFL	Synthetic	INRIA_Lytro
Bicubic	PSNR	25.9698	25.7012	26.7195	26.1549	24.0982	25.1822
	SSIM	0.8937	0.8975	0.7932	0.8589	0.9351	0.8402
	LPIPS	0.3016	0.3199	0.2707	0.2521	0.1812	0.2122
Baseline	PSNR	28.3245	27.4327	29.6461	28.2012	23.1102	27.2195
	SSIM	0.9793	0.9742	0.9830	0.9835	0.9688	0.9897
	LPIPS	0.0226	0.0274	0.0152	0.0063	0.0294	0.0075
DIP-LFSR	PSNR	28.7187	27.8434	30.5655	29.1421	24.1197	27.6507
	SSIM	0.9807	0.9761	0.9868	0.9876	0.9738	0.9907
	LPIPS	0.0162	0.0196	0.0110	0.0040	0.0234	0.0060
DistgSSR	PSNR	24.9760	24.7544	25.6203	25.1893	23.8128	24.6414
	SSIM	0.9590	0.9484	0.9609	0.9709	0.9773	0.9817
	LPIPS	0.0406	0.0324	0.0323	0.0009	0.0141	0.0136

辨效果,需要将有监督学习方法在每个特定的超分辨倍数和噪声水平下重新训练。此外,从数值结果可以看出,引入多尺度损失函数和全变差损失函数使DIP的性能得到提升,受噪声的影响更小。表2结果表明,当用不同方法进行光场空间域2倍超分辨时,DIP-LFSR的PSNR比DIP高出约0.7 dB,比插值方法高出约2.4 dB,比有监督方法DistgSSR高出约3.2 dB。当进行4倍超分辨时,DIP-LFSR的PSNR比DIP高出约0.4 dB,比插值方法高出约1.8 dB,比有监督方法DistgSSR高出约2.1 dB。当进行8倍超分辨时,DIP-LFSR的PSNR比DIP高出约0.001 dB,比插值方法高出约1.4 dB,比有监督方法DistgSSR高出1.6 dB。这些结果表明了所提自监督学习方法DIP-LFSR在抗

噪声干扰和泛化性方面的优势。

3.2.1.2 视觉展示

以 origami 场景为例,加入 $\sigma=0.05$ 的噪声,图3展示了不同超分辨方法对光场空间域进行2倍超分辨的视觉效果。图4为在 books 场景上的视觉效果。从图3右图的局部放大图可以观察到,与真实值(ground truth)相比,双三次插值方法的效果最模糊,不能可靠地恢复缺失的细节且受噪声的影响比较大。DIP和DIP-LFSR比普通插值方法和有监督学习的方法都清晰而且噪点明显少于其他方法,相比于DIP,DIP-LFSR的纹理更加清晰,细节更加丰富,视觉上超过了DistgSSR的效果。因此,所提DIP-LFSR在视觉效果上优于对比方法。

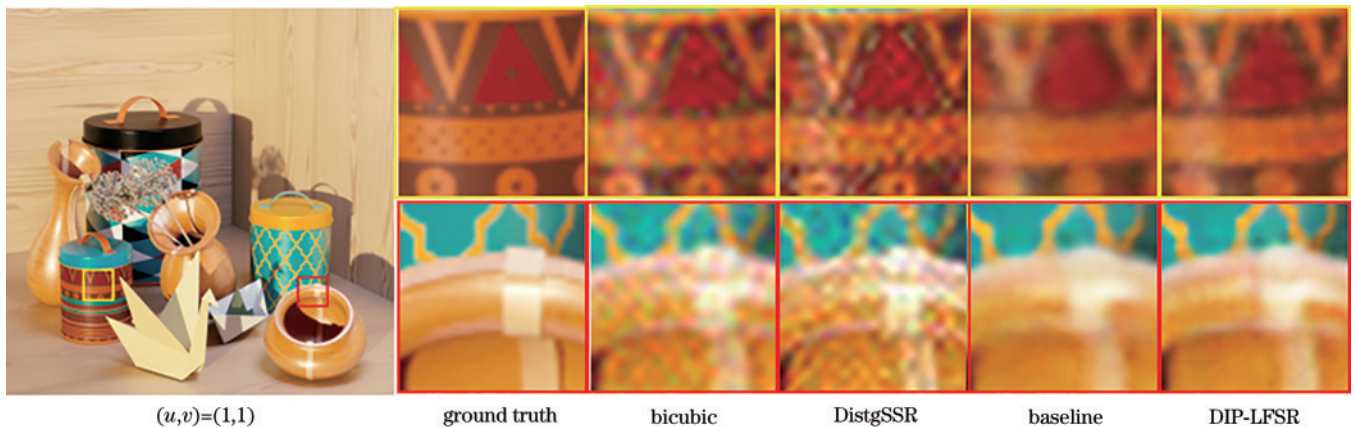


图3 不同超分辨方法在含噪声($\sigma=0.05$)的origami场景上对光场空间域进行2倍超分辨重构的视觉效果

Fig. 3 Visual effects of two times super-resolution reconstruction of light field spatial-domain in noisy origami scene ($\sigma=0.05$) using different super resolution methods

3.2.2 噪声对超分辨重构的影响

对于选取的12个场景的光场数据集,通过对其分别添加 $\sigma=0.1$ 和 $\sigma=0.2$ 的噪声来模拟生成含噪声

的光场。用不同的超分辨图像重构方法对光场在空间域实现2倍和4倍超分辨率重构,研究不同噪声大小对模型性能的影响。为了获得良好的光场空间域超分辨

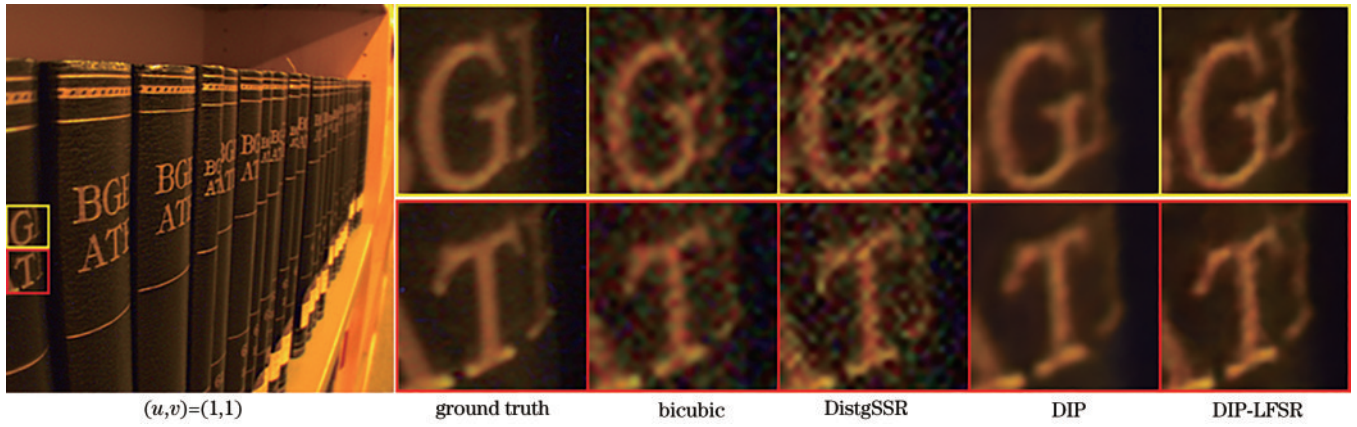


图 4 不同超分辨方法在含噪声($\sigma=0.05$)的books场景上对光场空间域进行2倍超分辨重构的视觉效果

Fig. 4 Visual effects of two times super-resolution reconstruction of light field spatial-domain in noisy books scene($\sigma=0.05$) using different super-resolution methods

重构效果, DIP-LFSR 模型在不同任务上选取不同停机步数(steps)。具体如表 4 所示。

表 4 在不同程度的噪声影响的情况下, 不同方法在所选 6 种数据集上对光场空间域分别进行 2 倍和 4 倍超分辨重构的平均结果

Table 4 In the case of different degrees of noise influence, the average results of 2 and 4 times super-resolution reconstruction of the light field space domain by different methods on the selected six datasets

σ /steps	Scale	Method	PSNR	SSIM	LPIPS
0.1/500	×2	Bicubic	21.2011	0.7155	0.4981
		DIP	25.3143	0.9674	0.0278
		DIP-LFSR	25.4202	0.9681	0.0270
	×4	DistgSSR	19.8475	0.8991	0.0782
		Bicubic	20.1957	0.6736	0.6588
		DIP	23.2744	0.9476	0.0486
0.2/200	×2	DIP-LFSR	23.6202	0.9530	0.0390
		DistgSSR	19.1030	0.8819	0.0920
		Bicubic	16.1815	0.4895	0.7349
	×4	DIP	21.8439	0.9239	0.0533
		DIP-LFSR	22.5179	0.9397	0.0396
		DistgSSR	14.6421	0.7265	0.1950
		Bicubic	15.7796	0.4581	0.7770
		DIP	20.3775	0.8922	0.0850
		DIP-LFSR	20.4059	0.8930	0.0845
		DistgSSR	14.4191	0.7124	0.2056

结合 $\sigma=0.05$ 时的超分辨结果, 表 4 结果表明: DIP-LFSR 在 $\sigma=0.1$ 和 $\sigma=0.2$ 不同程度的噪声水平下, 对光场在空间域进行 2 倍和 4 倍超分辨率重构的数值结果量化评价指标都优于对比方法。当噪声水平增加时, DIP-LFSR 性能下降最慢。以光场在空间域进行 2 倍超分辨为例, 当 σ 由 0.1 增加为 0.2 时, Bicubic

和有监督方法 DistgSSR 的 PSNR 下降约 5 dB, DIP 的 PSNR 下降约 3.5 dB, 而 DIP-LFSR 的 PSNR 只下降约 3 dB, 说明 DIP-LFSR 对噪声具有抑制作用。

图 5 和图 6 展示不同方法对光场在空间域进行 2 倍超分辨率重构结果的角度结构一致性, 图 5 以 stilllife 场景为例, 图 6 以 Lego Knights 场景为例, 模拟噪声水平为 $\sigma=0.1$, 选取 9×9 的子孔径图像进行 2 倍超分辨, 并展示不同方法超分辨率重构结果的子孔径图像在 $v=5, w=50$ 处的 EPI 图像。有监督方法和普通的插值方法对带有噪声的光场数据进行超分辨的结果存在很多噪点, 而 DIP-LFSR 的 EPI 图像更加光滑, 线条更加清晰, 表明 DIP-LFSR 方法在实现超分辨的同时能抑制低分辨率图像中的噪声。

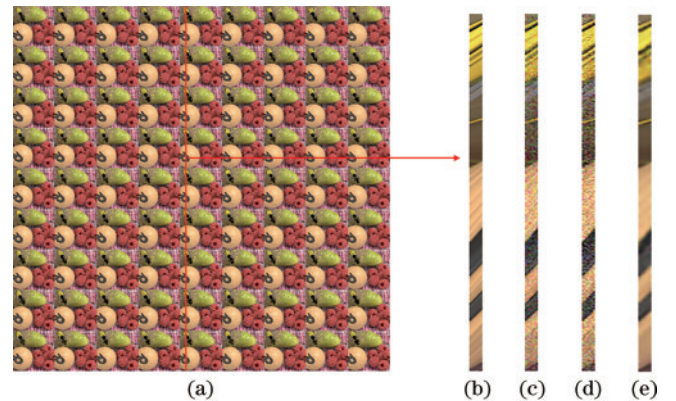


图 5 stilllife 场景的 9×9 子孔径图像及对比方法进行光场空间域超分辨重构结果的 EPI 展示, 模拟噪声水平 $\sigma=0.1$ 。(a) 光场子孔径图像; (b) 真实 EPI; (c) 双三次插值的超分辨结果的 EPI; (d) DIP-LFSR 超分辨结果的 EPI; (e) 有监督方法 DistgSSR 超分辨结果的 EPI

Fig. 5 The sub-aperture images of the stilllife scene and the EPI of the super-resolution results, the simulated noise level $\sigma=0.1$. (a) Light field sub-aperture images; (b) ground truth of EPI; (c) EPI of the Bicubic interpolated light field; (d) EPI of the DIP-LFSR; (e) EPI of the DistgSSR

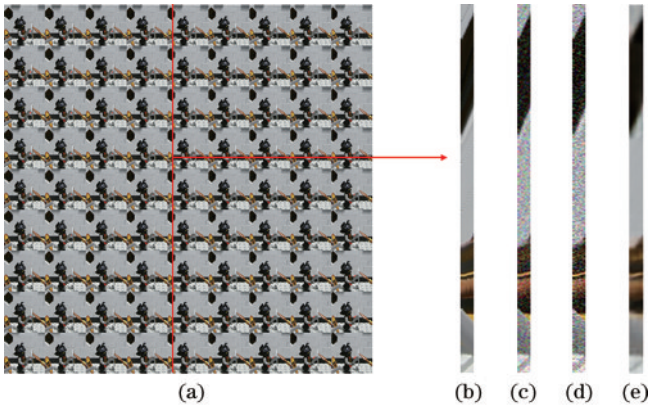


图6 Lego Knights场景的 9×9 子孔径图像及对对比方法进行光场空间域超分辨重构结果的EPI展示,模拟噪声水平 $\sigma=0.1$ 。(a)光场子孔径图像;(b)真实EPI;(c)双三次插值的超分辨结果的EPI;(d)DIP-LFSR超分辨结果的EPI;(e)有监督方法DistgSSR超分辨结果的EPI

Fig. 6 The sub-aperture images of the Lego Knights scene and the EPI of the super-resolution results, the simulated noise level $\sigma=0.1$. (a) Light field sub-aperture images; (b) ground truth of EPI; (c) EPI of the Bicubic interpolated light field; (d) EPI of the DIP-LFSR; (e) EPI of the DistgSSR

3.2.3 DIP-SAISR和DIP-LFSR方法的对比

比较DIP-SAISR对光场子孔径图像逐张用DIP做超分辨重构和DIP-LFSR方法对光场的所有视点图像同时做超分辨的效果。以herbs场景和medieval为例,模拟加入 $\sigma=0.05$ 的噪声,两种方法的网络参数通过网格搜索的方法调至最优。图7和图8展示了真实高分辨图像、Bicubic插值方法、DIP-SAISR和DIP-LFSR 2倍超分辨的结果,选取光场 $(u,v)=(1,1)$ 位置的子孔径图像上的两个区域的局部放大图进行展示。

图7和图8表明:对子孔径图像逐张进行超分辨的效果优于双三次插值方法,DIP-LFSR的效果优于DIP-SAISR。DIP-SAISR超分辨的结果存在很多噪点,用DIP-LFSR对光场子孔径图像同时超分辨的结果噪点明显减少,图像更加光滑,表明DIP-LFSR能够隐式利用到光场各子孔径图像中所蕴含的视差信息。

3.2.4 全变差正则化损失函数中参数 β 的影响

本组实验研究全变差正则化损失函数在选取不同参数 β 时对DIP-LFSR模型性能的影响。以herbs场景的光场数据为例,加入 $\sigma=0.2$ 的噪声,DIP-LFSR模型进行2倍超分辨重构。参数 β 的取值范围为 $[4, 16]$,模型训练步数steps设置为1000。图9展示了DIP-LFSR在不同 β 取值下的超分辨结果,其中,横坐标为 β 的取值,纵坐标为DIP-LFSR在herbs场景上的PSNR。

图9表明: β 增大时,PSNR取值随之变化,但在 β 为10时,超分辨结果的评价指标值是最高的,因此为了简化模型测试,数值实验在所有数据集上都取 $\beta=10$ 。

3.2.5 网络参数数量的对比

表5表明,在DIP-LFSR多尺度网络结构中,中间通道数 C 设置为128时,DIP-LFSR模型的参数量远远低于有监督方法DistgSSR模型的参数量,当中间通道数 C 设置为256时,DIP-LFSR模型的参数量与DistgSSR模型的参数量相当,此时DIP-LFSR对光场空间域进行超分辨率重构的优势能够表现出来,并且在有噪声的情况下超过有监督方法DistgSSR模型的性能,因此DIP-LFSR网络结构中的中间通道数设置为256。

3.2.6 网络子模块数对超分辨结果的影响

本组实验比较网络编码模块和解码模块的子模块

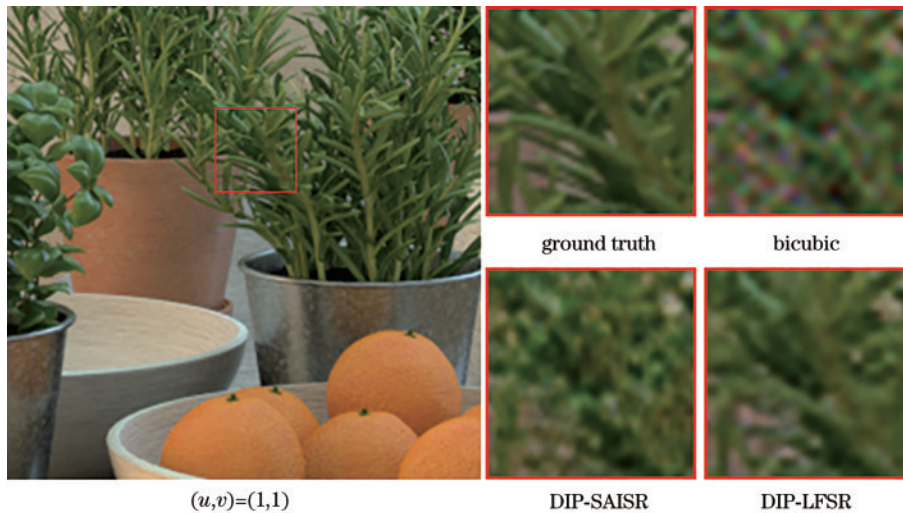


图7 在herbs场景上模拟 $\sigma=0.05$ 的噪声,用不同方法对其空间域进行2倍超分辨的结果对比

Fig. 7 The herbs scene with added noise $\sigma=0.05$, comparison of results of 2 times super-resolution in its spatial domain using different methods

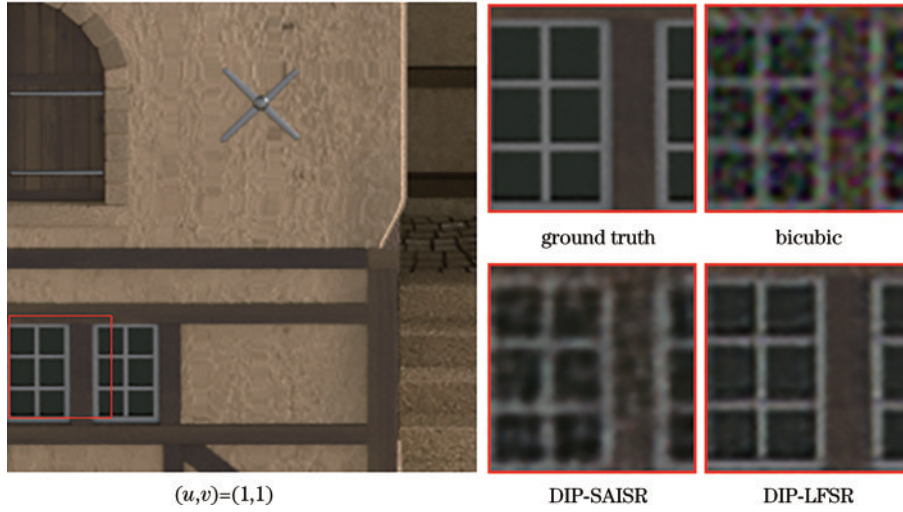


图 8 在 medieval 场景上模拟 $\sigma = 0.05$ 的噪声,用不同方法对其空间域进行超分辨 2 倍的结果对比

Fig. 8 The medieval scene with added noise $\sigma = 0.05$, comparison of results of 2 times super-resolution in its spatial domain using different methods

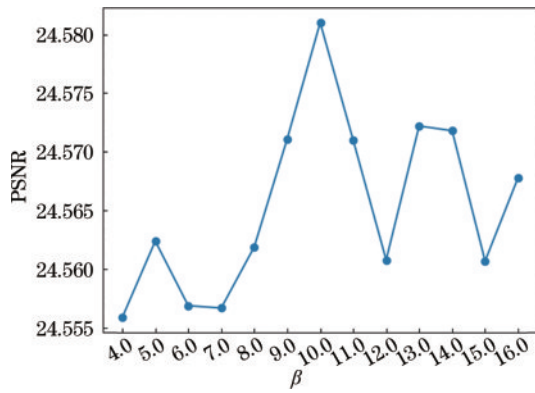


图 9 光场超分辨质量(PSNR)随 β 取值的变化

Fig. 9 Variation of light field super-resolution quality (PSNR) with β value

表 5 不同方法的不同网络参数量对比

Table 5 Comparison of different network parameters in different methods

Method	Channels	#Parameters
DIP-LFSR	128	1922943
	256	2696832
DistgSSR		2646912

数(number)分别设置为 3、4、5、6 时,对 DIP 和 DIP-LFSR 超分辨性能的影响。以 HCI 数据集为例,模拟噪声水平 $\sigma = 0.05$,进行光场空间域 2 倍超分辨。表 6 为不同方法超分辨的量化评价指标。

表 6 数据表明,用 DIP 和 DIP-LFSR 进行超分辨,网络子模块数设置为 4 时,不同量化评价指标最好,因此数值实验中统一将网络模块数设置为 4。

3.2.7 不同权重的多尺度损失函数的影响

本组实验对比了设置不同权重得到的多尺度损失函数(分别记为 L_1, L_2, L_3, L_4, L_5)对超分辨结果的影

表 6 网络编码模块和解码模块的子模块数设置为 3、4、5、6 时的光场空间域 2 倍超分辨结果

Table 6 Results of 2 times light field spatial-domain super-resolution when the number of encoding and decoding modules is set to 3, 4, 5, 6

σ	Steps	Scale	Method	Number	PSNR	SSIM	LPIPS	
0.05	1500	$\times 2$	DIP	3	28.5070	0.9788	0.0157	
				4	28.7270	0.9808	0.0158	
				5	28.6043	0.9795	0.0176	
				6	28.3891	0.9786	0.0205	
				DIP-LFSR	3	28.5834	0.9793	0.0160
					4	28.7703	0.9810	0.0161
	5	28.5562	0.9793		0.0181			
	6	28.2962	0.9780		0.0205			

响,设置 L_1 的权重为 λ_1^{i+1}, L_2 的权重为 λ_1^{i-1} ,其中 $\lambda_i = 0 (i = 1, 2), \lambda_i = \frac{1}{3UV} (i = 3, 4, 5), L_3$ 的权重为 $\lambda_i^i (\lambda_i = 0.03), L_4$ 的权重为 $\lambda_i^i (\lambda_i = 0.05), L_5$ 的权重为 λ_i^i ,其中, $\lambda_i = 0 (i = 1, 2), \lambda_i = \frac{1}{3UV} (i = 3, 4, 5)$ 。以 HCI 数据集为例,模拟噪声水平 $\sigma = 0.05$,进行光场空间域 2 倍超分辨,表 7 展示出了 DIP-LFSR 加入不同尺度损失函数超分辨的量化评价指标的平均结果。

表 7 多尺度损失函数设置不同权重对超分辨结果的影响

Table 7 Influence of multi-scale loss function with different weights on super-resolution results

σ /steps	Method	Scale	PSNR	SSIM	LPIPS	
0.05/1500	DIP-LFSR	$\times 2$	L_1	28.7295	0.9802	0.0164
			L_2	28.7402	0.9803	0.0160
			L_3	28.7371	0.9801	0.0162
			L_4	28.7548	0.9803	0.0162
			L_5	28.7703	0.9810	0.0161

表 7 表明, 权重设定为 $\lambda_i = 0 (i = 1, 2)$ 、 $\lambda_i = \frac{1}{3UV} (i = 3, 4, 5)$ 时的超分辨结果是最好的, 因此本文实验部分统一采用该设定的权重选择规则。

3.2.8 不同损失函数对超分辨性能的影响

本组实验将 L_{MSE} 损失函数分别与多尺度损失函数 L_{MS} 和全变差正则化损失函数 L_{TV} 结合构成新的混合损失函数的方法分别记为 DIP+MS 和 DIP+TV, 并与 DIP 以及采用 L_{MSE} 、 L_{MS} 和 L_{TV} 构成的混合损失函数的方法 DIP-LFSR 进行对比。以 HCI 数据集为例, 模拟噪声水平设定为 $\sigma = 0.05$, 进行光场空间域 2 倍超分辨。表 8 展示了高分辨率光场的量化评价指标。

表 8 中的实验结果表明, DIP 模型的 MSE 损失函

表 8 不同混合损失函数对光场空间域进行 2 倍超分辨的结果
Table 8 2 times light field spatial-domain super-resolution results on different hybrid loss functions

σ /steps	Method	Scale	PSNR	SSIM	LPIPS
0.05/1500	DIP	×2	28.7270	0.9808	0.0158
	DIP+TV		28.7252	0.9815	0.0159
	DIP+MS		28.7277	0.9801	0.0163
	DIP-LFSR		28.7703	0.9810	0.0161

表 9 不同上采样方法对光场空间域进行 2 倍超分辨的结果

Table 9 2 times light field spatial-domain super-resolution results on different upsampling methods

σ	Steps	Method	Scale	Unsample	PSNR	SSIM	LPIPS
0.05	1500	DIP	2	Bilinear	28.7270	0.9808	0.0158
				ConvTranspose2d	23.3155	0.9323	0.0785
	500	DIP-LFSR		Bilinear	28.7703	0.9810	0.0161
				ConvTranspose2d	23.3862	0.9310	0.0780

4 结 论

DIP-LFSR 将自监督学习方法用于光场的空间域超分辨任务中, 超分辨过程只需要原始的低分辨率图像, 通过在模型损失函数中加入全变差正则化损失和多尺度特征损失来隐式利用光场数据的子孔径图像蕴含的视差信息。实验结果表明, 在有噪声影响的情况下, 所提自监督学习模型在加噪声以后的多个公开光场数据集上进行空间域超分辨的重构结果在主观视觉效果和量化评价指标上优于传统的插值方法和有监督学习的方法。综上所述, 所提自监督光场空间域超分辨重构方法对噪声有抑制作用, 后续工作可以考虑用 DIP 模型对光场进行去噪, 同时对光场角度域进行超分辨。

参 考 文 献

[1] 徐欣宜, 邓慧萍, 向森, 等. 基于特征交互融合与注意力的光场图像超分辨率[J]. 激光与光电子学进展, 2023, 60(14): 1410017.

数分别再加入全变差正则化损失以及多尺度损失之后的超分辨结果并没有明显提升 DIP 的效果。添加全变差正则化以及多尺度损失之后的 DIP-LFSR, 性能相比 DIP 得到提升。因此, 实验中的 DIP-LFSR 采用加入全变差正则化和多尺度损失的混合损失函数。

3.2.9 不同上采样方法对模型性能的影响

本组实验对比在网络中使用不同的上采样方法对超分辨结果的影响。典型的上采样方法有双线性插值、PixelShuffle^[42] 和 ConvTranspose2d^[43]。由于用 PixelShuffle 方法进行特征图上采样会引入大量新的可学习参数, 仅将双线性插值与 ConvTranspose2d 进行对比, 说明不同上采样方法对模型性能的影响。以 HCI 数据集为例, 模拟噪声水平设定为 $\sigma = 0.05$, 进行光场空间域 2 倍超分辨。表 9 展示出超分辨结果的量化评价指标。

表 9 中数据表明, 上采样方法选取 Bilinear 时, 模型的性能明显超过上采样方法选取 ConvTranspose2d 时的模型性能。出现该现象的原因是在采用 ConvTranspose2d 实现上采样时模型中引入部分新的可学习参数, 且该运算与原始低分辨率图像的降采样过程不匹配。关于上采样运算更合适的实现方式的研究需要进一步探讨。

Xu X Y, Deng H P, Xiang S, et al. Light field image super-resolution based on feature interaction fusion and attention mechanism[J]. Laser & Optoelectronics Progress, 2023, 60(14): 1410017.

- [2] 葛鹏, 游耀堂. 基于稀疏表示的光场图像超分辨率重建[J]. 激光与光电子学进展, 2022, 59(2): 0210001.
Ge P, You Y T. Super-resolution reconstruction of light field images via sparse representation[J]. Laser & Optoelectronics Progress, 2022, 59(2): 0210001.
- [3] 吕天琪, 武迎春, 赵贤凌. 角度差异强化的光场图像超分辨率网络[J]. 光电工程, 2023, 50(2): 0220185.
Lü T Q, Wu Y C, Zhao X L. Light field image super-resolution network based on angular difference enhancement[J]. Opto-Electronic Engineering, 2023, 50(2): 0220185.
- [4] 许娇, 袁三男. 增强型多尺度残差网络的图像超分辨率重建算法[J]. 激光与光电子学进展, 2023, 60(4): 0411002.
Xu J, Yuan S N. Image super-resolution reconstruction algorithm based on enhanced multi-scale residual network[J]. Laser & Optoelectronics Progress, 2023, 60(4): 0411002.

- [5] Cheng Z, Xiong Z W, Chen C, et al. Light field super-resolution: a benchmark[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), June 16-17, 2019, Long Beach, CA, USA. New York: IEEE Press, 2020: 1804-1813.
- [6] Lim J, Ok H, Park B, et al. Improving the spatial resolution based on 4D light field data[C]//2009 16th IEEE International Conference on Image Processing (ICIP), November 7-10, 2009, Cairo, Egypt. New York: IEEE Press, 2010: 1173-1176.
- [7] Nava F P, Luke J P. Simultaneous estimation of super-resolved depth and all-in-focus images from a plenoptic camera[C]//2009 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video, May 4-6, 2009, Potsdam, Germany. New York: IEEE Press, 2009.
- [8] Wanner S, Goldluecke B. Spatial and angular variational super-resolution of 4D light fields[C]//Proceedings of the 12th European conference on Computer Vision-Volume Part V, October 7-13, 2012, Florence, Italy. New York: ACM Press, 2012: 608-621.
- [9] Wanner S, Goldluecke B. Variational light field analysis for disparity estimation and super-resolution[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, 36(3): 606-619.
- [10] Rossi M, Frossard P. Graph-based light field super-resolution[C]//2017 IEEE 19th International Workshop on Multimedia Signal Processing (MMSp), October 16-18, 2017, Luton, UK. New York: IEEE Press, 2017.
- [11] Yoon Y, Jeon H G, Yoo D, et al. Learning a deep convolutional network for light-field image super-resolution[C]//2015 IEEE International Conference on Computer Vision Workshop (ICCVW), December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2016: 57-65.
- [12] Zhang S, Lin Y F, Sheng H. Residual networks for light field image super-resolution[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2020: 11038-11047.
- [13] Liang Z Y, Wang Y Q, Wang L G, et al. Light field image super-resolution with transformers[J]. *IEEE Signal Processing Letters*, 2022, 29: 563-567.
- [14] 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*, 2023, 45(1): 425-443.
- [15] Luo Z X, Huang Y, Li S, et al. Learning the degradation distribution for blind image super-resolution [EB/OL]. (2022-03-09)[2023-02-06]. <https://arxiv.org/abs/2203.04962>.
- [16] Lempitsky V, Vedaldi A, Ulyanov D. Deep image prior [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: 9446-9454.
- [17] Luo Z W, Huang H B, Yu L, et al. Deep constrained least squares for blind image super-resolution[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 18-24, 2022, New Orleans, LA, USA. New York: IEEE Press, 2022: 17621-17631.
- [18] 彭晏飞, 张曼婷, 张平甲, 等. 聚合残差注意力网络的单图像超分辨率重建[J]. *激光与光电子学进展*, 2023, 60(10): 1010017.
- [19] Peng Y F, Zhang M T, Zhang P J, et al. Single-image super-resolution reconstruction aggregating residual attention network[J]. *Laser & Optoelectronics Progress*, 2023, 60(10): 1010017.
- [20] 吴洛冰, 谷玉海, 吴文昊, 等. 基于多尺度特征提取的遥感旋转目标检测[J]. *激光与光电子学进展*, 2023, 60(12): 1228010.
- [21] Wu L B, Gu Y H, Wu W H, et al. Remote sensing rotating object detection based on multi-scale feature extraction[J]. *Laser & Optoelectronics Progress*, 2023, 60(12): 1228010.
- [22] Sitzmann V, Martel J N P, Bergman A W, et al. Implicit neural representations with periodic activation functions[EB/OL]. (2020-06-17)[2023-02-03]. <https://arxiv.org/abs/2006.09661>.
- [23] Zhang H M, Dong B. A review on deep learning in medical image reconstruction[J]. *Journal of the Operations Research Society of China*, 2020, 8(2): 311-340.
- [24] 刘鹏飞, 肖亮. 基于 Hessian 核范数正则化的快速图像复原算法[J]. *电子学报*, 2015, 43(10): 2001-2008.
- [25] Liu P F, Xiao L. A fast algorithm for image restoration based on Hessian nuclear norm regularization[J]. *Acta Electronica Sinica*, 2015, 43(10): 2001-2008.
- [26] Lefkimmiatis S, Ward J P, Unser M. Hessian Schatten-norm regularization for linear inverse problems[J]. *IEEE Transactions on Image Processing*, 2013, 22(5): 1873-1888.
- [27] Rudin L I, Osher S, Fatemi E. Nonlinear total variation based noise removal algorithms[J]. *Physica D: Nonlinear Phenomena*, 1992, 60(1/2/3/4): 259-268.
- [28] Mahendran A, Vedaldi A. Understanding deep image representations by inverting them[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 5188-5196.
- [29] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation [M]//Navab N, Hornegger J, Wells W M, et al. Medical image computing and computer-assisted intervention-MICCAI 2015. Lecture notes in computer science. Cham: Springer, 2015, 9351: 234-241.
- [30] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift [EB/OL]. (2015-02-11)[2023-03-05]. <https://arxiv.org/abs/1502.03167>.
- [31] Glorot X, Bordes A, Bengio Y. Deep sparse rectifier neural networks[C]//Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, April 11-13, 2011, Fort Lauderdale, USA. Cambridge: JMLR, 2011: 315-323.
- [32] Zhao B N, Zhang X S, Li Z, et al. A multi-scale strategy for deep semantic segmentation with convolutional neural

- networks[J]. *Neurocomputing*, 2019, 365: 273-284.
- [30] 虞志军, 王国栋, 张镡月. 基于增强多尺度特征网络的图像去模糊[J]. *激光与光电子学进展*, 2022, 59(22): 2215007.
- Yu Z J, Wang G D, Zhang X Y. Image deblurring based on enhanced multiscale feature network[J]. *Laser & Optoelectronics Progress*, 2022, 59(22): 2215007.
- [31] Honauer K, Johannsen O, Kondermann D, et al. A dataset and evaluation methodology for depth estimation on 4D light fields[M]//Lai S H, Lepetit V, Nishino K, et al. *Computer vision-ACCV 2016. Lecture notes in computer science*. Cham: Springer, 2017, 10113: 19-34.
- [32] Wanner S, Meister S, Goldlücke B. Datasets and benchmarks for densely sampled 4D light fields[EB/OL]. [2023-02-03]. <http://diglib.eg.org/handle/10.2312/PE.VMV.VMV13.225-226>.
- [33] Raj A S, Lowney M, Shah R, et al. Stanford lytro light field archive[EB/OL]. [2023-03-02]. https://scholar.google.com/scholar_lookup?title=Stanford+Lytro+Light+Field+Archive&author=Raj,+A.S.&author=Lowney,+M.&author=Shah,+R.&author=Wetzstein,+G.&publication_year=2016.
- [34] Rerabek M, Ebrahimi T. New light field image dataset [EB/OL]. [2023-03-02]. <https://infoscience.epfl.ch/record/218363>.
- [35] Le Pendu M, Jiang X R, Guillemot C. Light field inpainting propagation via low rank matrix completion[J]. *IEEE Transactions on Image Processing*, 2018, 27(4): 1981-1993.
- [36] Duchon C E. Lanczos filtering in one and two dimensions [J]. *Journal of Applied Meteorology*, 1979, 18(8): 1016-1022.
- [37] Kingma D P, Ba J. Adam: a method for stochastic optimization[EB/OL]. (2014-12-22) [2023-02-03]. <https://arxiv.org/abs/1412.6980>.
- [38] Li Z Y, Arora S. An exponential learning rate schedule for deep learning[EB/OL]. (2019-10-16) [2023-03-02]. <https://arxiv.org/abs/1910.07454>.
- [39] Huynh-Thu Q, Ghanbari M. Scope of validity of PSNR in image/video quality assessment[J]. *Electronics Letters*, 2008, 44(13): 800-801.
- [40] 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.
- [41] Zhang R, Isola P, Efros A A, et al. The unreasonable effectiveness of deep features as a perceptual metric[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: 586-595.
- [42] Shi W Z, Caballero J, Huszár F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 1874-1883.
- [43] Gao H Y, Yuan H, Wang Z Y, et al. Pixel transposed convolutional networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(5): 1218-1227.