

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2025)02-0391-15

论文引用格式: Wang Q T, Li Z L, Ding X, Liu Q and Yang Y. 2025. Gaussian-Wiener-based dense focal stack image synthesis. Journal of Image and Graphics, 30(02):0391-0405(王其腾, 李志龙, 丁新, 刘琼, 杨铀. 2025. 高斯一维纳表示下的稠密焦栈图生成方法. 中国图象图形学报, 30(02):0391-0405)[DOI:10.11834/jig.240249]

高斯一维纳表示下的稠密焦栈图生成方法

王其腾, 李志龙, 丁新, 刘琼, 杨铀*

华中科技大学电子信息与通信学院, 武汉 430074

摘要: 目的 焦栈图像能够扩展光学系统的景深, 并为计算摄影、交互式 and 沉浸式媒体提供灵活的图像表达。然而, 受限于光学系统的物理属性和拍摄对象的动态变化, 人们往往只能拍摄稀疏的焦栈图像。因此, 焦栈图像的稠密化成为当前需要解决的一个难题。为应对上述挑战, 提出了一种高斯一维纳表示下的稠密焦栈图生成方法。方法 焦栈图像被抽象为高斯一维纳表示, 所提出的双向预测模型分别包含双向拟合模块和预测生成模块, 在高斯一维纳表示模型的基础上构建双向拟合模型, 求解双向预测参数并生成新的焦栈图像。首先, 将稀疏焦栈序列的图像按照相同块大小进行分块, 并基于此将相邻焦距、相同位置的块组合成块对, 以块对为最小单元进行双向预测。其次, 在双向预测模块中, 块对将用于拟合出最佳双向拟合参数, 并基于此求解出预测生成参数, 生成新的焦栈图像块。最后, 将所有预测生成得到的块进行拼接, 得到新的焦栈图像。结果 在 11 组稀疏焦栈图像序列上进行实验, 所采用评价指标包括峰值信噪比 (peak signal to noise ratio, PSNR) 和结构相似性 (structure similarity index measure, SSIM)。11 个序列生成结果的平均 PSNR 为 40.861 dB, 平均 SSIM 为 0.976。相比于广义高斯和空间坐标两个对比方法, PSNR 分别提升了 6.503 dB 和 6.467 dB, SSIM 分别提升了 0.057 和 0.092。各序列均值 PSNR 和 SSIM 最少提升了 3.474 dB 和 0.012。结论 实验结果表明, 所提出的双向预测方法可以较好地生成新的焦栈图像, 能够在多种以景深为导向的视觉应用中发挥关键作用。

关键词: 焦栈图像 (FoSI); 预测模型; 高斯一维纳; 表示模型; 双向预测

Gaussian-Wiener-based dense focal stack image synthesis

Wang Qiteng, Li Zhilong, Ding Xin, Liu Qiong, Yang You*

School of Information and Communication, Huazhong University of Science and Technology, Wuhan 430074, China

Abstract: Objective In optical imaging systems, the depth of field (DoF) is typically limited by the properties of optical lenses, resulting in the ability to focus only on a limited region of the scene. Thus, expanding the depth of field for optical systems is a challenging task in the community for both academia and industry. For example, in computational photography, when dense focus stack images are captured, photographers can select different focal points and depths of field in postprocessing to achieve the desired artistic effects. In macro- and micro-imaging, dense focus stack images can provide clearer and more detailed images for more accurate analysis and measurement. For interactive and immersive media, dense

收稿日期: 2024-05-16; 修回日期: 2024-07-04; 预印本日期: 2024-07-11

* 通信作者: 杨铀 yangyou@hust.edu.cn

基金项目: 湖北省技术创新计划重点研发专项 (2023BAB021); 华中科技大学基础研究支持计划资助项目 (2023BR023); 中央高校基本科研业务费专项资金资助—华中科技大学创新研究院技术创新基金项目 (2024JYCXJJ022)

Supported by: Key Research and Development Program of Hubei Province, China (2023BAB021); Fundamental Research Program of Huazhong University of Science and Technology (2023BR023); Fundamental Research Funds for the Central Universities-Technology Innovation Fund Project of Innovation Institute of Huazhong University of Science and Technology (2024JYCXJJ022)

focus stack images can provide a more realistic and immersive visual experience. However, achieving dense focus stack images also faces some challenges. First, the performance of hardware devices limits the speed and quality of image acquisition. During the shooting process, the camera needs to adjust the focus quickly and accurately and capture multiple images to build a focus stack. Therefore, high-performance cameras and adaptive autofocus algorithms are required. In addition, changes in the shooting environment, such as object motion or manual operations by photographers, can also introduce image blurring and alignment issues. These challenges are addressed by introducing the block-based Gaussian-Wiener bidirectional prediction model to provide an effective solution. When the image is embedded into blocks and the characteristics of local blocks for prediction are utilized, the computational complexity can be reduced, and the prediction accuracy can be improved. Gaussian-Wiener filtering can smooth prediction results and reduce the impact of artifacts and noise, which can improve image quality. The bidirectional prediction method combines the original sparse focal stack images (FoSIs) with the prediction results to generate dense FoSIs, thereby expanding the DoF of the optical system. The Gaussian-Wiener bidirectional prediction model provides an innovative method for capturing dense focus stack images. It can be applied to various scenarios and application fields, providing greater creative freedom and image processing capabilities for photographers, scientists, engineers, and artists.

Method This work abstracts the FoS as a Gaussian-Wiener representation. The proposed bidirectional prediction model includes a bidirectional fitting module and a prediction generation module. On the basis of the Gaussian-Wiener representation model, a bidirectional fitting model is constructed to solve for the bidirectional prediction parameters and draw a new focal stack image. First, on the basis of the given sparse focus stack image sequence, the number ranges from near to far according to the focal length. These numbers start from 0 and are incremented according to a certain rule, such as increasing by 2 each time to ensure that all the numbers are even. These results in a set of sparse focus stack images arranged in serial order. Then, all images were divided into blocks according to predefined block sizes. The size of each block can be selected on the basis of specific needs and algorithms. The blocks located in adjacent numbers are combined to form a block pair, which becomes the most basic unit for bidirectional prediction. Before conducting bidirectional prediction, the image was preprocessed by dividing the focus stack image into blocks and recombining them into block pairs. This preprocessing process can be achieved by using image segmentation algorithms and block pair combination strategies. For each block pair, bidirectional prediction was performed to obtain the prediction parameters. These prediction parameters can be determined on the basis of specific prediction models and algorithms, such as the block-based Gaussian-Wiener bidirectional prediction model. In the bidirectional prediction module, block pairs can be used to fit the best bidirectional fitting parameters, and on this basis, the prediction generation parameters can be solved. When the information of the prediction generation parameters and block pairs is applied, new prediction blocks can be generated. Finally, when all the prediction blocks are concatenated according to their positions in the image, new prediction focus stack images can be obtained.

Result This experiment is performed on 11 sparse focal stack images, with evaluation metrics using the peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM). The average PSNR of the 11 sequence generated results is 40.861 dB, and the average SSIM is 0.976. Compared with the two comparison methods of generalized Gaussian and spatial coordinates, the PSNR is improved by 6.503 dB and 6.467 dB, respectively, and the SSIM is improved by 0.057 and 0.092, respectively. The average PSNR and SSIM of each sequence improved by at least 3.474 dB and 0.012, respectively.

Conclusion The experimental results show that the method proposed in this study outperforms both the subjective and objective comparison methods and has good performance on 11 different scene sequences. Combined with ablation experiments, the advantages of bidirectional prediction in the proposed method have been demonstrated. The results indicate that the proposed bidirectional prediction method can effectively generate new focal pile images and play a crucial role in visual applications that target various depths of the field.

Key words: focal stack image(FoSI); prediction model; Gaussian-Wiener; representation model; bidirectional prediction

0 引言

随着计算机视觉和成像技术的发展,焦点堆栈图像(focal stack image, FoSI)的应用范围越来越广泛。焦点堆栈图像是在固定角度下拍摄的一组聚焦在不同平面上的图像序列(王旭等,2023),通过移动镜头(Chang等,2018)或成像探测器来获取(亢新凯等,2020;孙路通等,2020)。将焦点堆栈图像进行融合可以得到全景深图像,这种计算成像技术可以应用于诸多需要扩展景深的领域,例如计算摄影、显微成像(任苏霞等,2022;李四维等,2021)和沉浸式媒体(王涌天等,2016)等。另外,通过焦点堆栈图像还可以计算出深度信息,这使得焦点堆栈图像在场景恢复(Ruan等,2021)、动作识别(Sun等,2023)、三维重建(Hu等,2020)和显著性检测(Wang等,2019)等领域发挥重要作用。然而,从FoSIs中受益的各种应用都是以密集的深度采样为代价的(Wu等,2022)。稠密的焦点堆栈图像包含更多场景信息,这有助于更精确地重建场景的三维结构。通过分析多个焦平面上的图像差异,可以更准确地推断出物体的相对位置和距离,从而在进行深度估计和三维重建时获得更为可靠的结果(Kontogianni等,2017)。此外,在改善图像质量方面,由于稠密焦点堆栈图像包含了多个不同焦点距离的图像,所以在合成全聚焦图像时,可以更好地保留场景中各个部分的细节(Li等,2018)。因此,获取稠密的焦栈图像是至关重要的。

目前获取焦点堆栈图像的方式有两种:一种是通过成像系统直接在同一视点下连续改变对焦距离拍摄获取;另一种是通过先获取光场数据,然后进行数字重聚焦的方式获取(谢柠宇等,2022;解培月等,2017)。然而,通过成像系统直接采集焦栈图像对系统的最短曝光时间有很高要求(任婧荣等,2023),成像系统的性能和拍摄人员的专业性都面临挑战。光场数据可以同时记录场景中光线的强度和方向,从而可以在拍摄之后通过数字重聚焦的方式获取焦点堆栈图像。然而光场数据需要光场相机进行采集,依赖高质量的镜头与高分辨率的探测器,体积大、成本高(王奇和傅雨田,2020)。并且对于单个光场相机,在传感器像元个数一定的情况下,由于要同时获取光线的位置和角度信息,其采集到的图像分辨率较小(张春萍和王庆,2016),进而影响重聚焦后

获取的焦栈图像的分辨率。为了获取更高的分辨率,使用相机阵列来获取光场数据进行重聚焦(Liu等,2016)。但是,搭建相机阵列的成本依然很高,并且由于相机物理尺寸的限制,相机之间始终存在较大的基线距离,因而无法采集密集角度的信息,导致重聚焦后焦栈图像质量下降(方璐和戴琼海,2020),需要经过一系列的前期或后期处理来提升焦点堆栈图像质量,这将增加获取稠密焦点堆栈图像的难度。

在数字重聚焦算法中,除了通过光场数据恢复重聚焦图像的方法外,还有通过深度图估计和卷积的方式来进行重聚焦的方法。基于深度图的方法通过估计场景中不同像素点的深度来恢复出位于不同焦平面的重聚焦图像。向森等人(2022)通过特征匹配找到参考平面后使用因式分解方法恢复深度图,并根据深度图得到重聚焦图像。该方法可以对单摄像头拍摄的图像进行重聚焦,但是在处理纹理较少的区域时,会出现深度图边缘模糊的现象;并且该方法的计算速度慢于一些立体输入的方法。Hog等人(2017)提出了一种基于立体焦平面堆栈计算深度图的方法,不需要计算子孔径图像就可以直接根据深度图渲染出特定深度的重聚焦图像,但是该方法需要对相机特定部件进行调整。由于卷积可以实现模糊图像的效果,而反卷积是对卷积的逆过程,可以从模糊图像中恢复出清晰图像,所以使用特定的卷积算法也可以模拟聚焦和散焦的效果,从而实现图像重聚焦。基于卷积的方法的关键是获取卷积核信息,Anger等人(2018)利用从模糊图像中估计出模糊卷积核就可以使用模糊核计算出恢复清晰图像的反卷积核这一理论,提出了一种利用自然图像统计特性估计模糊核的方法,可以直接从模糊图像的频谱中估计模糊核。但是当图像噪声水平较高时,模糊核的估计将会变得不准确。Liu等人(2021)提出了一种广义的双参数高斯核模型,可以更准确地描述模糊核,并消除振铃伪影,但是该方法没有考虑局部区域的复杂变化。Zheng等人(2022)提出了一种展开深度核估计方法,利用端到端网络学习核先验,可以更好地预测复杂的模糊核,但是该方法的性能依赖于训练数据的质量和多样性。Fu等人(2023)考虑了图像的三维空间特性,可以更准确地描述散焦现象,但是需要针对不同的光学系统进行调整。

为了提升获取稠密焦栈图像的质量,同时降低获取稠密焦栈图像的成本和难度,本文从一个新的

方向出发,提出了一种高斯—维纳表示下的双向预测稠密焦栈图生成方法,可以通过一组稀疏的焦栈图像生成稠密的焦栈图像。该方法采用卷积核估计的方式来模拟相邻稀疏焦栈图像之间的聚焦或散焦变化,并预测出新的卷积核用于生成新的焦栈图像。具体来说,本文设计了一个基于高斯—维纳表示的双向拟合模块,该模块可以通过迭代拟合参数,即高斯卷积标准差 σ 来获得相邻焦距之间的最佳拟合参数,从而模拟出相邻焦距之间的变化。同时,本文还设计了一个预测模块,该模块与双向拟合模块级联,将双向拟合模块的输出作为输入进行预测参数的求解,通过对最佳拟合参数取平均的方法,求解出预测参数用于生成新的焦栈图像。本文方法使用双向预测的模式,充分利用了相邻焦栈图像的信息,提升了预测的准确性。本文方法在 11 组焦栈数据集图像序列(Wu 等, 2020)上进行实验,在生成的焦栈图像和真值图像之间计算峰值信噪比(peak signal-to-noise, PSNR)和结构相似性(structure similarity index

measure, SSIM)两个指标,并且设计了两个对比方法来验证本文方法的优越性。实验结果表明,本文方法生成的焦栈图在 11 个序列上的平均 PSNR 和 SSIM 很高,并且优于对比方法。

1 本文方法

为了应对稠密焦栈图像的采集困难,本文提出了一种基于高斯—维纳表示法(Wu 等, 2022)的双向预测稠密焦栈图生成方法,该方法的总体框架如图 1 所示。本文方法包含 3 个部分:对输入焦栈图进行处理、双向预测模型和拼接部分。首先对输入的稀疏焦栈图像进行分块,并按照焦距和 (x, y) 坐标划分块对。然后将块对输入双向预测模型生成新块。最后将生成的所有新块再按照焦距和 (x, y) 坐标拼接出新的焦栈图像。新的焦栈图与原稀疏焦栈图共同组成稠密的焦栈图。图 1 中使用的操作和参数在表 1 中进行了说明。

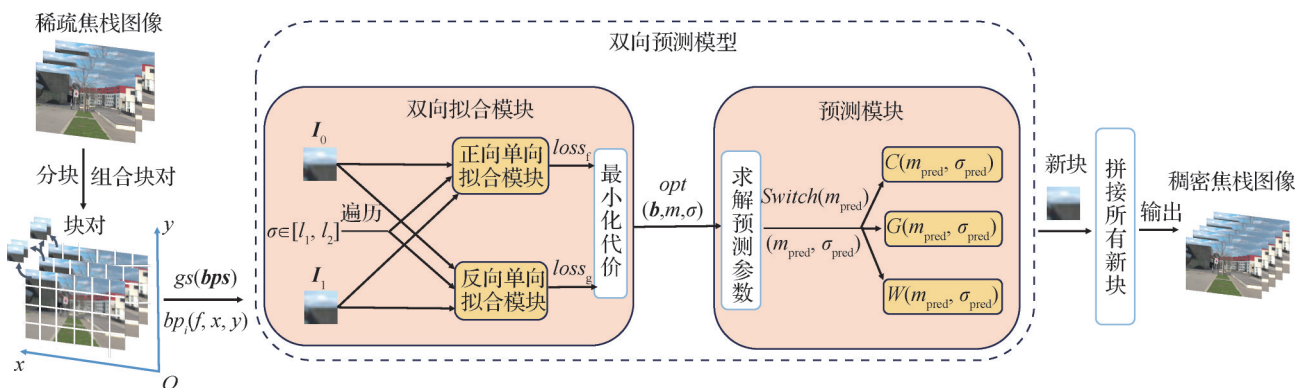


图 1 基于高斯—维纳表示的双向预测稠密焦栈图生成方法整体框架

Fig. 1 Overall framework of bidirectional prediction dense focal stack map generation method based on Gaussian-Wiener representation

1.1 图像分块和块对划分

由于焦栈图像中不同位置清晰度随焦距变化的程度不同,所以本文方法以块为单位进行处理,针对不同位置的块,求解对应的预测参数,并预测出新的块。

为简化计算,本文方法采用固定块大小。对于图像分辨率不能被块大小整除的情况,对图像采用边缘镜像填充的方法进行填充(营永胜 等, 2023),确保分辨率可以被块大小整除。在分块完成后,将相邻焦距的两幅焦栈图像中位于相同位置 (x, y) 的两个块组合成块对。在后续的步骤中,块对中的两个

块将轮流作为参考块和当前块进行预测参数的求解。

1.2 双向预测模型

在本文提出的双向预测模型中,包含双向拟合模块和预测模块。该模型建立在高斯—维纳表示法的基础上。

1.2.1 用于焦栈图像的高斯—维纳表示法

高斯—维纳表示法是通过高斯卷积和维纳反卷积来分别模拟散焦和聚焦变化的。该方法采用分块的形式,对不同块进行处理。同时,在该方法中,对于随焦距变化清晰度变化不大的区域,采用复制的方式进行表示。

表1 变量与参数列表说明

Table 1 Explanation of variables and parameters in this paper

操作和参数名	说明
组合块对	组合块对的规则是相同坐标 (x, y) 、相邻焦距的两个块进行组合。
$bp_i(f, x, y)$	表示不同的块对,以下标 i 来区分。其中, f 表示块对中两个块焦距编号的均值, x 和 y 表示块的坐标。
$gs(bps)$	表示遍历所有块对 $bp_i(f, x, y)$ 组成的块对集合 bps 。
I_0 和 I_1	表示块对中的两个图像块。
$loss_f$ 和 $loss_g$	分别表示正向和反向单向拟合模块的最小损失。
$opt(b, m, \sigma)$	表示最佳双向拟合参数。其中, b, m, σ 分别表示最小代价的单向拟合模块对应的参考块、拟合模式和 σ 系数。
$(m_{pred}, \sigma_{pred})$	表示预测参数对。其中, m_{pred} 和 σ_{pred} 分别表示预测模式和预测系数。
$C(m_{pred}, \sigma_{pred})$	表示对参考块进行复制。
$G(m_{pred}, \sigma_{pred})$	表示对参考块进行高斯卷积。
$W(m_{pred}, \sigma_{pred})$	表示对参考块进行维纳反卷积。
$Switch(m_{pred})$	表示根据预测模式决定进行何种模式的预测。

该方法使用高斯卷积来模拟点扩散函数(point spread function, PSF)。由于光线在到达图像传感器上时,能量会向周围逸散,使得单条光线在图像传感器上形成的并不是一个理想的光点,而是一个光斑。本文将这个光斑称为艾里斑,将其能量分布函数称为点扩散函数(PSF),PSF的函数表达式为

$$P_a(x, y) = \left(\frac{2J_1\left(\frac{\pi D \sqrt{x^2 + y^2}}{\lambda f}\right)}{\frac{\pi D \sqrt{x^2 + y^2}}{\lambda f}} \right)^2 \quad (1)$$

式中, $J_1(\cdot)$ 是一阶第1类贝塞尔函数; λ 、 D 和 f 分别是光的波长、孔径和透镜的焦距。

经过研究发现,高斯函数可以和PSF进行拟合,通过改变高斯函数的标准差 σ 来拟合不同的PSF函数,通过二维高斯函数可以建模出三维的PSF函数。因此,可以使用退化—恢复关系来表示FoSIs的聚焦—散焦过程。FoSIs的退化函数被建模为高斯函数。退化—恢复关系可表示为

$$g(x, y) = f(x, y) \otimes h(x, y) + n \quad (2)$$

式中, $f(x, y)$ 和 $g(x, y)$ 是不同图像的位置 (x, y) 中的区域, $h(x, y)$ 表示退化—恢复关系的算子, n 是加性零均值噪声。注意,在基本退化模型(Bai等, 2020)中, $f(x, y)$ 和 $g(x, y)$ 分别是清晰图像和模糊图像。与基本退化模型中的区域不同, $f(x, y)$ 和 $g(x, y)$ 是本文退化—恢复关系模型中的一般区域,

既可能清晰也可能模糊。如果 $f(x, y)$ 比 $g(x, y)$ 更清晰,则算子 $h(x, y)$ 用来模拟图像退化的过程。相反,如果 $f(x, y)$ 比 $g(x, y)$ 更模糊,则算子 $h(x, y)$ 用来模拟图像恢复的过程。

在这种退化—恢复模型的基础上,提出了高斯—维纳表示法(Wu等, 2022),用于建模这种退化—恢复关系。高斯—维纳表示法包含高斯卷积和维纳反卷积,分别用于建模区域散焦、聚焦过程。本文的区域是固定大小的块。在高斯卷积建模的退化关系中,退化算子为

$$h(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (3)$$

式中, x, y 为二维平面坐标, σ 为二维高斯分布标准差。在本文方法中,通过对参考块进行高斯卷积来模拟散焦的过程,由于图像为二维平面,所以需要使用二维高斯卷积来建模。在维纳反卷积建模的恢复关系中,恢复算子遵循如下计算式,具体为

$$G(u, v) = H_\omega(u, v) \times F(u, v) \quad (4)$$

$$H_\omega(u, v) = \frac{1}{H(u, v)} \times \frac{|H(u, v)|^2}{|H(u, v)|^2 + \frac{1}{SNR}} \quad (5)$$

式中, $F(u, v)$ 和 $G(u, v)$ 分别是 $f(x, y)$ 和 $g(x, y)$ 的傅立叶变换; $H_\omega(u, v)$ 是系统的传递函数,它与PSF和信噪比(signal to noise ratio, SNR)有关。本文在维纳反卷积中采用的PSF是二维高斯函数,而 $H(u, v)$ 是 $h(x, y)$ 的傅里叶变换。在公式中,它惩罚具有较

低SNR的频率,因此噪声信号的频率将被抑制。实际上,精确的SNR几乎不可得到,因此通常采用近似常数。在通过频域计算出 $G(u, v)$ 后,再通过傅里叶反变换得到 $g(x, y)$ 。

1.2.2 双向拟合模块

在双向预测模型中,需要对每一个块对进行双向拟合。在双向拟合中,使用高斯-维纳表示法来建模图像清晰度随焦距变化的趋势。在双向拟合模型中,引入3种模式,分别为复制模式、高斯模式和维纳模式,分别对应复制参考块、对参考块做高斯卷积和对参考块做维纳反卷积这3种不同的操作,并将操作后得到的结果与当前块计算损失,将损失作为各模式的输出。

在双向拟合模块中包含两个单向拟合模块,单向拟合模块的框架图如图2所示。单向拟合即将相邻图像的同一位置的两个块分别作为参考块和当前块,通过参考块向当前块拟合。双向拟合过程中,会将两个块轮流作为参考块和当前块。双向拟合包含正向和反向两个单向拟合模块,图2中通过参数 d 来判断是正向还是反向,然后进行参考块和当前块的判定。在确定好参考块和当前块后,将参考块和当前块输入到后续的拟合部分。

双向拟合的目的是求解最佳双向拟合参数,该参数将用于进一步生成预测参数。最佳双向拟合参数由最佳模式参数 m 、最佳系数 σ ,以及最小代价单向拟合对应的参考块 b 组成,可以通过求解拟合代价的优化问题来获得。在一次单向拟合中,通过拟合失真来计算拟合代价。拟合失真表示当前块与参考块拟合结果之间的失真。对于当前块 I ,复制模式、高斯模式和维纳模式依次应用于参考块 R 。 M

和 N 是当前块和参考块的宽和高。通过最小化所有模式的拟合代价来计算最佳单向拟合参数 (m_d, σ_d) 和最小拟合代价 $loss_d$ 。然后取最小单向拟合代价 $loss_d$ 最小的单向拟合的最佳单向拟合参数作为最佳双向拟合参数。

复制模式的代价函数为

$$J_1(\sigma) = \frac{1}{M \times N} \|I - C(R, \sigma)\|_2^2 \quad (6)$$

式中, $C(\cdot)$ 代表复制操作,由于复制操作不需要 σ ,所以直接设置为0。因此, $C(R, \sigma)$ 表示复制参考块 R 作为当前块 I 的预测, $J_1(\sigma)$ 是复制模式的拟合代价。

在高斯模式下,当参考块 R 比当前块 I 更清晰时,高斯卷积应用于 R ,代价计算为

$$J_2(\sigma) = \min_{\sigma \in [l_1, l_2]} \left(\frac{1}{M \times N} \|I - G(R, \sigma)\|_2^2 \right) \quad (7)$$

式中, $G(\cdot)$ 代表高斯卷积操作, σ 是零均值二维高斯函数的标准差; $G(R, \sigma)$ 表示参考块 R 向当前块 I 的拟合,并且通过将高斯卷积应用于参考块 R 来生成; $J_2(\sigma)$ 是高斯模式的拟合代价。 σ 的搜索范围从 l_1 到 l_2 ,步长为 s 。搜索测试范围内的每个 σ ,并计算代价。搜索完成后选择最小代价 $J_2(\sigma)$ 对应的 σ 作为最佳 σ 。

在维纳模式下,当参考块 R 比当前块 I 更模糊时,通过维纳反卷积生成去模糊预测块。代价计算为

$$J_3(\sigma) = \min_{\sigma \in [l_1, l_2]} \left(\frac{1}{M \times N} \|I - W(R, \sigma)\|_2^2 \right) \quad (8)$$

式中, $W(\cdot)$ 代表维纳反卷积操作, σ 的范围设定为从 l_1 到 l_2 ,搜索方法与高斯模式相同。

在一次单向拟合中,通过比较3种模式各自的最小拟合代价,得出最佳单向拟合参数。在两次单

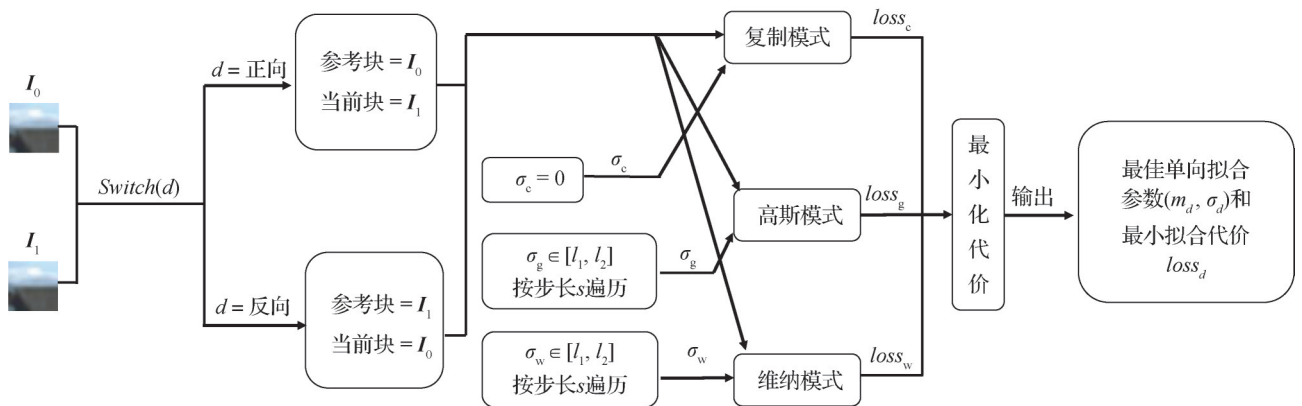


图2 单向拟合框架图

Fig. 2 Unidirectional fitting framework diagram

向拟合后会获得两组最佳单向拟合参数,在这两组最佳单向拟合参数中,通过寻求最小代价的方式获得最佳双向拟合参数。获得最佳双向拟合参数,计算为

$$\text{opt}(\mathbf{b}, m, \sigma) = \arg \min_{m \in \{\pm 1, \pm 2, \pm 3\}} J_m(\sigma) \quad (9)$$

式中, m 是模式参数,用正负表示拟合方向。正向表示从小焦距编号向大焦距编号拟合的方向,负向表示从大焦距编号向小焦距编号拟合的方向。当 m 分别等于 1、2 和 3 时,分别表示正向复制模式、高斯模式和维纳模式。在正向拟合中,将小焦距编号的块作为参考块。当 m 分别等于 -1、-2 和 -3 时,分别表示负向复制模式、高斯模式和维纳模式,在负向拟合中,将大焦距编号的块作为参考块。 σ 表示每种模式的最佳标准差参数, $\mathbf{b}$ 表示最佳双向拟合参数的参考块, $\text{opt}(\mathbf{b}, m, \sigma)$ 是最佳双向拟合参数。

1.2.3 预测模块

在双向预测模型中,预测模块用于求解预测参数,并根据预测参数生成新的块。该模块的输入是上一级双向拟合模块的输出,即最佳双向拟合参数,并且将最佳双向拟合参数对应的参考块一并作为该模块的输入。

通过最佳双向拟合参数来求解生成新焦栈图的预测参数,关键是找到通过参考块向两个块对应焦

距之间的中间块拟合的趋势。预测参数的作用就是通过对参考块进行变换来模拟从参考块向中间块拟合的趋势,这样得到的中间块更加符合中间焦距对应的块。整个预测生成模块的流程如图 3 所示,其中,最佳双向拟合参数为上一级双向拟合模块的输出。参考块为上一级双向拟合模块中代价最小的单向拟合对应的参考块。预测参数通过双向拟合参数求解得到,预测参数计算为

$$\begin{cases} m_{\text{pred}} = m \\ \sigma_{\text{pred}} = \sigma/2 \end{cases} \quad (10)$$

式中, m_{pred} 是预测参数的模式参数, σ_{pred} 是预测参数的标准差参数, m 和 σ 为最佳双向拟合的模式参数和标准差参数。最佳双向拟合参数可以拟合出相邻焦距相同位置块的焦距变化,它最符合相邻焦距相同位置块的焦距变化趋势。通过将最佳双向拟合参数的 σ 除以 2 得到的 σ_{pred} 可以模拟出从参考块向位于相邻焦距相同位置块的焦距正中间焦距块的拟合过程,从而保证输入的稀疏焦栈图和输出的新焦栈图组合成的稠密焦栈图在焦距域均匀分布。通过预测参数 $(m_{\text{pred}}, \sigma_{\text{pred}})$ 来插值出新的焦栈图。在求解出预测参数后,通过判断 m_{pred} 的绝对值来决定通过哪种模式来生成新块。这里使用绝对值判断是因为正负只决定预测的方向。在确定好模式后,将参考块和 σ_{pred} 输入对应模式生成新块。

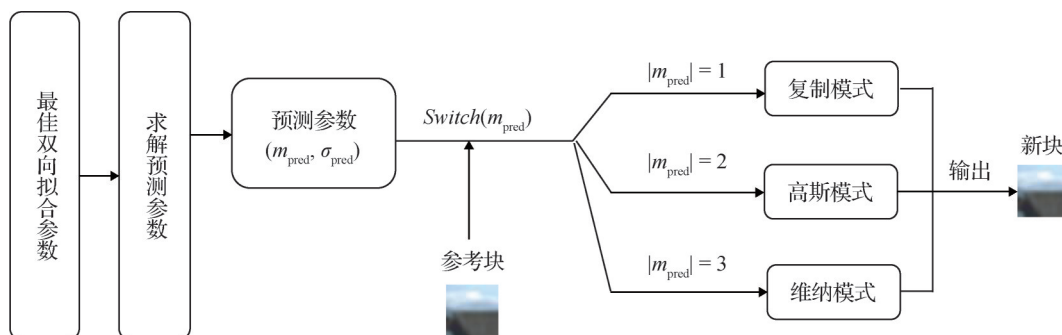


图3 预测生成模块流程图

Fig. 3 Flow chart of prediction generation module

1.3 拼接新焦栈图

在所有块对经过双向预测模型生成新块后,将所有的新块作为拼接模块的输入。在拼接模块中,将所有新块按照焦距进行划分,相同焦距编号的新块划分为一组。然后将同组的新块按照对应块对的坐标位置 (x, y) 进行排列后拼接在一起,组成新的焦栈图像。最后将得到的新焦栈图像与原稀疏焦栈

图像按照焦距编号从小到大顺序排列组成稠密焦栈图像序列。

2 实验结果和分析

2.1 数据集和实验环境

本文使用的焦栈图像(Wu等,2020)数据集,共

有 11 个不同场景。所有场景中的焦栈图像均为 RGB 色彩空间。每个像素采用 24 bit 进行编码。每个场景均包含 21 幅焦栈图像。将 21 幅图像按照焦距从 0 到 20 进行编号,抽取偶数编号组成稀疏焦栈图,抽取奇数编号作为真值。所有实验序列如表 2 所示,其中“序号”为各序列的编号,例如 I01 是第 1 个序列的编号。分辨率表示图像宽高和通道数,例如 $1\,024 \times 512 \times 3$ 表示图像宽 1 024 个像素,高 512 个像素,通道数为 3。图 4 对数据集进行了部分展示,从每个序列中抽取一幅图像进行展示。

本文方法通过 Python3. 7. 16 实现,所有实验均在 64 位 Windows11 平台下完成。计算机处理器参数为 12th Gen Intel(R) Core(TM) i7-12700 2. 10 GHz,内存为 16 GB。

表 2 实验数据信息

Table 2 Information of experimental scenarios			
序号	场景名称	分辨率/像素	数量
I01	AssemblyParts	$1\,024 \times 512 \times 3$	21
I02	Factory	$1\,024 \times 512 \times 3$	21
I03	Fruits	$768 \times 512 \times 3$	21
I04	Greenbox	$512 \times 512 \times 3$	21
I05	IMG_0454	$624 \times 432 \times 3$	21
I06	IMG_0543	$624 \times 432 \times 3$	21
I07	Kobe	$768 \times 512 \times 3$	21
I08	Machine	$1\,024 \times 768 \times 3$	21
I09	Messy_room	$1\,024 \times 512 \times 3$	21
I10	Pipes	$512 \times 512 \times 3$	21
I11	Production_line	$1\,024 \times 512 \times 3$	21

2.2 评价指标和对比方法

为了评价生成新焦栈图的质量,本文采用 PSNR 和 SSIM 两个评价指标。将生成的新焦栈图像与真值分别计算这两个评价指标。两个指标的值越大,说明生成的图像越接近真值。SSIM 主要考量图像的 3 个关键特征:亮度(luminance)、对比度(contrast)和结构(structure),其取值范围介于 0~1 之间。

在对比方法方面,由于本文所提出的方法属于较为新颖的方法,目前没有直接生成焦点堆栈图的方法,所以本文从模拟聚散焦的方向出发,选取了两个使用卷积核估计来模拟清晰和模糊变化的方法。分别是广义的双参数高斯核模型(Liu 等,2021)和根

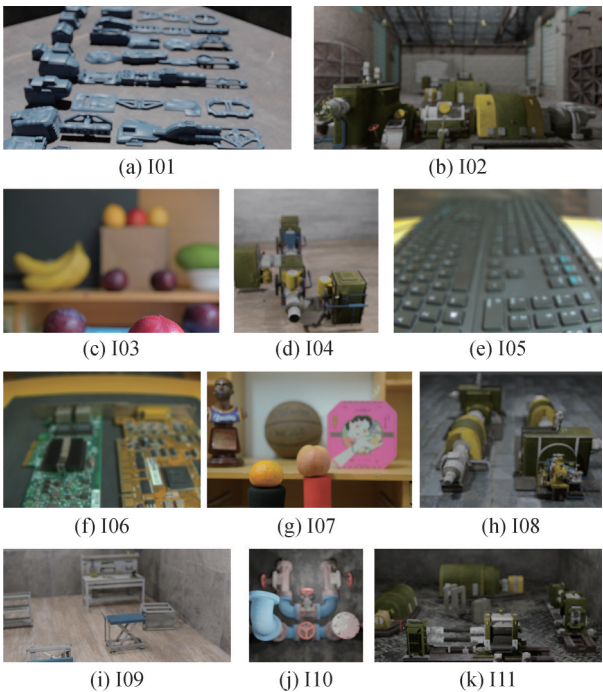


图 4 数据集展示

Fig. 4 Dataset display ((a)—(k) 11 focal stack sequences)

据图像的三维空间特性进行焦点变化模拟(Fu 等, 2023)。前者提出了将双参数简化为单参数的公式;后者通过空间坐标进行计算。由于焦栈图像只在一个空间方向,也就是焦平面进行平移,所以后者也可以简化为一种单参数模型。这两个方法和本文方法在双向拟合阶段采用不同的卷积核参数进行拟合。在预测生成阶段均采用对最佳拟合参数求平均的方式获取预测参数,并使用预测参数生成的卷积核进行焦栈图像的预测生成。

2.3 实验参数确定

在正式实验之前,需要确定的参数有两个:1)分块大小;2)卷积核大小。为了确定这两个参数,选取 I01 和 I07 两个序列进行测试。设置 3 种块大小:64, 32, 16;3 种卷积核大小:7, 5, 3。将块大小和卷积核大小两两组合,共 9 种组合。在两个序列上进行 9 种组合测试。将生成的新焦栈图与真值计算 PSNR 和 SSIM,结果如图 5 所示。其中, b 表示块大小, k 表示卷积核大小。以 $b_{64}k_7$ 为例,其表示块大小为 64×64 ,卷积核大小为 7×7 。

从实验结果看,两组序列在块大小为 64、卷积核大小为 3 时,生成的新焦栈图像的 PSNR 和 SSIM 均优于其他组合。所以确定实验的分块大小为 64,卷积核大小为 3。对于高斯模式和维纳模式,高斯

卷积和维纳反卷积的 σ 的搜索范围为 $[0.1, 10]$, 搜索步长为0.1。

为了证明取最小损失对应模式进行预测生成的有效性, 选取了序列I01进行实验。将3种模式和两种方向共6种组合的损失进行由小到大排序, 依次

生成每个排序位次下的图像, 然后分别计算生成图像与真值之间的PSNR和SSIM, 如图6所示。损失从小到大依次表示为 $loss_0 - loss_5$, $loss_0$ 表示最小损失, $loss_5$ 表示最大损失。从结果可以看出, 采取最小损失生成的结果的PSNR和SSIM是最高的。

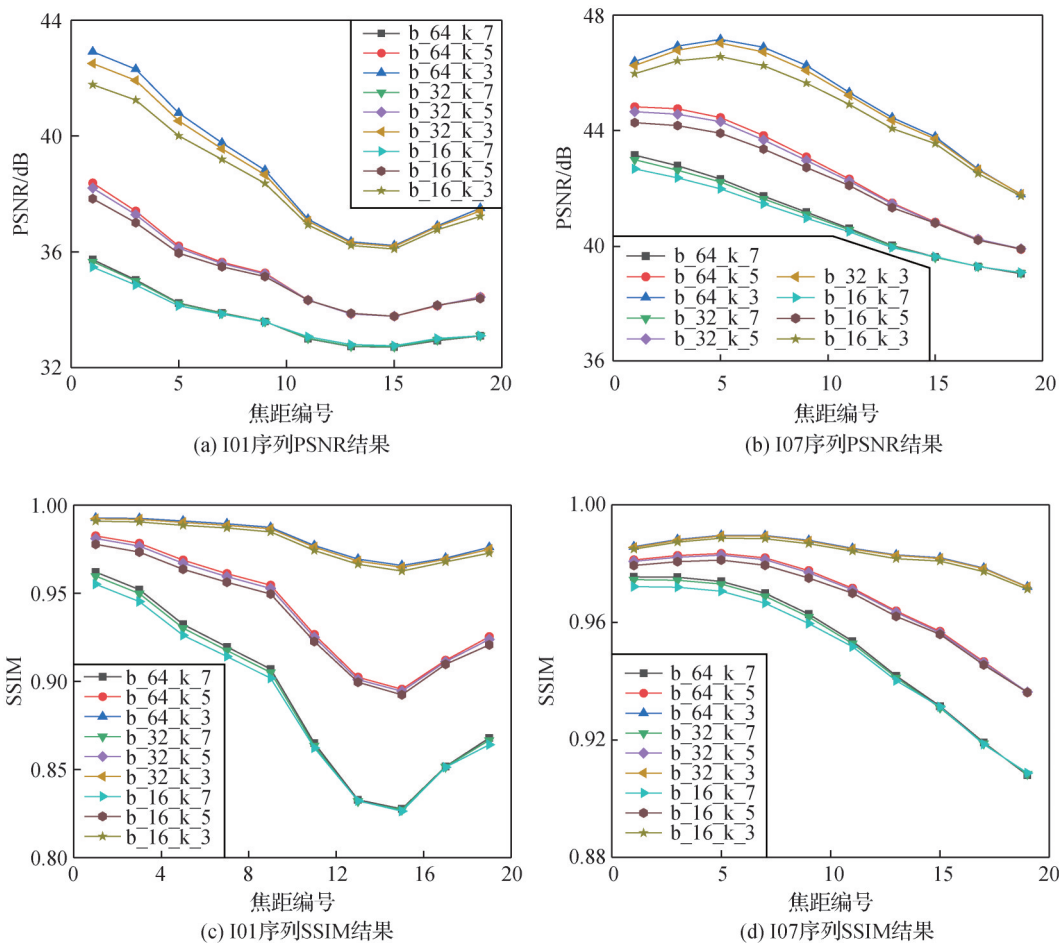


图5 I01和I07序列在9种组合上的测试结果

Fig. 5 Test results of I01 and I07 sequences on 9 combinations ((a) PSNR results of I01 sequence; (b) PSNR results of I07 sequence; (c) SSIM results of I01 sequence; (d) SSIM results of I07 sequence)

2.4 实验结果

客观实验结果如图7所示。图7展示了11个序列生成的新焦栈图的质量, 计算了各序列生成新焦栈图像与真值之间的PSNR和SSIM平均值, 横坐标为序列编号, 纵坐标为各序列生成的所有新焦栈图像的平均PSNR和平均SSIM。从实验数据可以看出, 本文方法在11个序列上均有不错的效果。各序列的平均PSNR均在36 dB以上, 平均SSIM均在0.95以上。所有序列的平均PSNR为40.861 dB, 平均SSIM为0.976。使用广义高斯方法获得的所有序列平均PSNR为34.358 dB, 平均SSIM为0.919。使

用空间坐标方法获得的所有序列平均PSNR为34.394 dB, 平均SSIM为0.884。在所有序列的平均结果上, 本文方法相对于两种对比方法在PSNR上的提升分别为6.503 dB和6.467 dB, 在SSIM上的提升分别为0.057和0.092。并且本文方法在各序列上均优于广义高斯和空间坐标两个对比方法, 相比于两种对比方法, 本文方法在各序列上的平均PSNR和SSIM最少提升了3.474 dB和0.012。

本文从11个序列中按照不同分辨率选取2个序列结果的图像进行主观结果展示, 分别是I07和I10, 主观结果如图8所示。两组序列分别是室内场景和

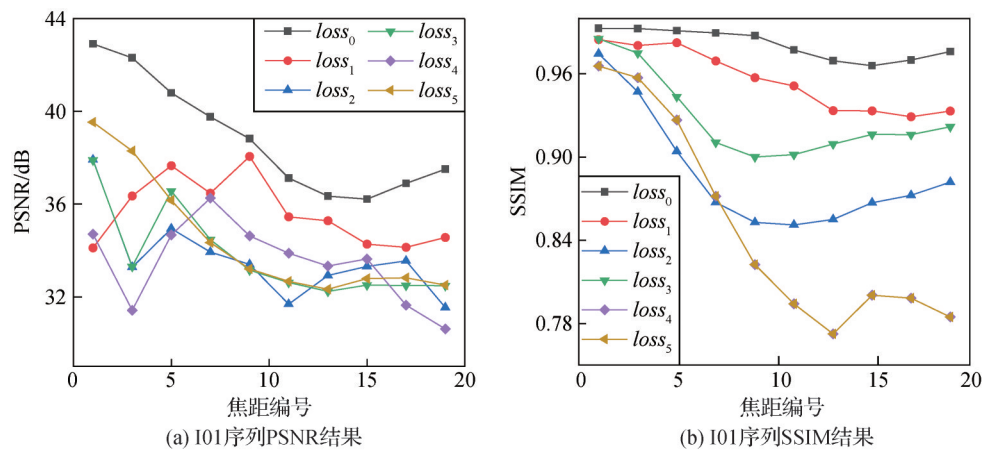


图6 I01序列在各种损失下PSNR和SSIM结果

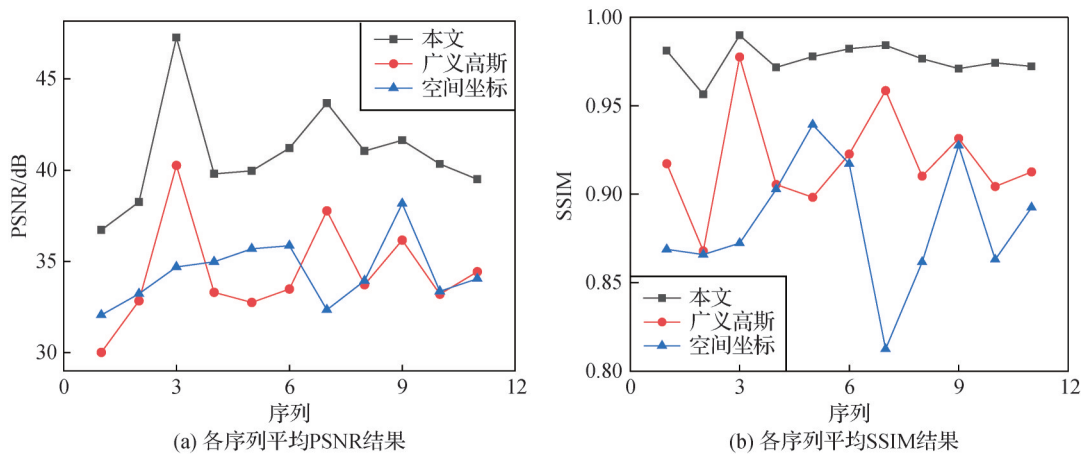
Fig. 6 PSNR and SSIM results of I01 sequence under various losses
((a) PSNR results of I01 sequence; (b) SSIM results of I01 sequence)

图7 各序列平均PSNR和SSIM结果

Fig. 7 Average PSNR and SSIM results for each sequence

((a) average PSNR results for each sequence; (b) average SSIM results for each sequence)

工业场景的焦栈序列。其中,每个序列选取1、11、19共3个焦距编号进行展示。序列号和焦距编号在每行行首位置。如图8第1行行首I07_1表示序列I07,焦距编号为1。每一行从左到右的4幅图像分别表示真值、广义高斯、空间坐标和双向预测的结果。每幅结果图中选取一块区域放大展示,以便于方法之间进行对比。

从主观结果看,广义高斯(Liu等,2021)的结果均存在模糊的现象,原因可能是没有引入维纳反卷积导致对于聚焦的变化无法很好地处理,而本文引入了维纳反卷积就很好地解决了这一问题。空间坐标的结果(Fu等,2023)在某些地方存在伪影,导致了图像失真。从I10的序列更容易发现,放大区域存在伪影并且块效应明显,而本文的结果不存在这么明显的图像失真。与两种对比方法不同,本文方

法在主观结果和客观实验结果上均十分接近真值。原因有两点,一是本文方法通过双向拟合,可以利用两个参考块找寻最优拟合结果,增加了可利用的信息,从而使得拟合结果更接近当前块,预测结果也更接近真值;二是本文方法使用相邻焦距进行拟合,使得参考块和当前块的焦距跨度变小,参考块和当前块之间的相似信息更多,使得拟合出最优结果的可能性增大。

通过客观实验结果和主观实验结果可以看出,本文方法优于两种对比方法。本文方法在客观结果上具有很高的PSNR和SSIM,在主观结果上也十分接近真值。

2.5 消融实验

为了验证本文方法的有效性,针对双向拟合模块设计了3组消融实验,分别是包含正反两个单向

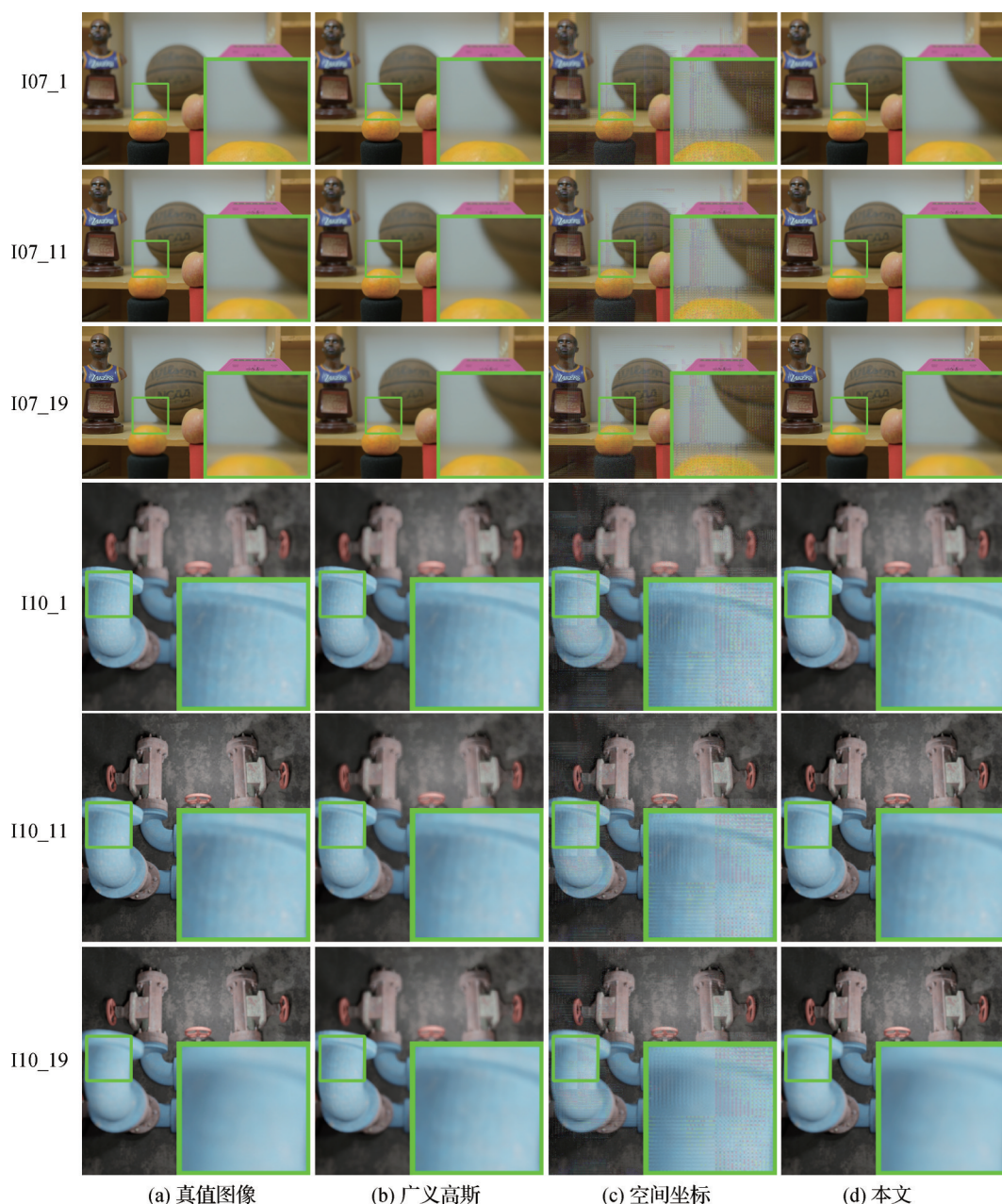


图8 主观实验结果对比

Fig. 8 Comparison of subjective results

((a) ground truth; (b) generalized Gaussian; (c) spatial coordinates; (d) ours)

拟合模块、只包含正向单向拟合模块和只包含反向单向拟合模块3组实验,其中第1组实验即本文方法。通过本文方法与另外两组单向拟合的实验结果进行对比来验证本文双向预测的优势。

消融实验同样采用预测结果与真值图像之间计算PSNR和SSIM两个指标进行评价。实验结果如图9所示。图中PSNR和SSIM为各序列所有新焦栈图预测结果的平均值。结果可以看出,双向拟合的结果的PSNR和SSIM在各序列均优于只包含一种单向拟

合的结果。本实验同时统计了正向预测和反向预测在11个序列预测中的占比情况。经统计,正向预测的占比为57.5%,反向预测的占比为42.5%。从统计结果看,两种预测方向占比接近1:1,由此可见双向预测是有必要的。

同时,本文依次增加输入焦栈图像之间的间隔来测试本文方法对稀疏焦栈序列的效果。选取3个序列: I01、I05和I11进行测试,每次增加间隔为2,确保相邻输入的偶数序号焦栈图像对应正中间一幅

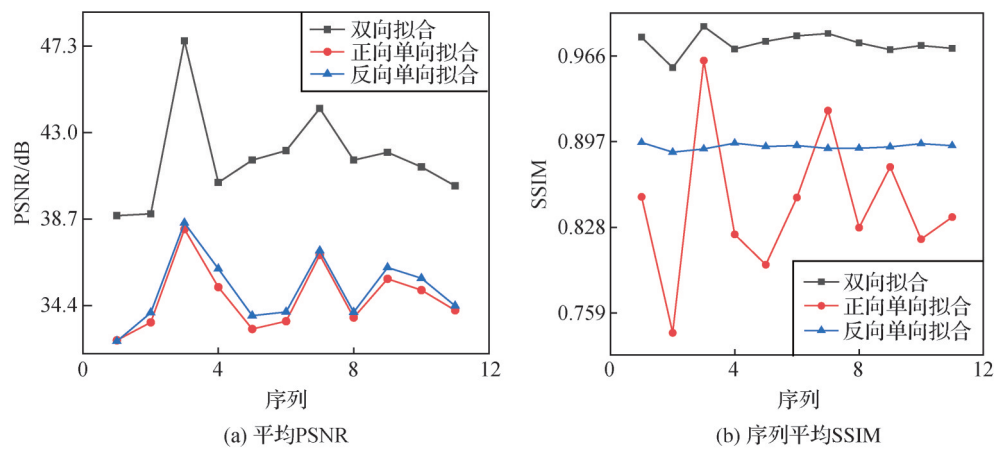


图9 消融实验各序列平均PSNR和SSIM结果

Fig. 9 Average PSNR and SSIM results of each ablation experiment sequence
((a) average PSNR; (b) average SSIM)

奇数生成图像。由于每个序列有21幅图像。所以最大间隔取到20。对每个间隔生成的每个序列结果取平均值,结果如图10所示。从结果可以看出,随着间隔的增加,PSNR和SSIM总体呈现减小趋势。

从图10(a)可以看出,PSNR在间隔超过12的时候低于36 dB。实验结果中,不同焦距间隔下各序列的PSNR均在30 dB以上,SSIM均在0.9以上。由实验结果分析,间隔在12或更小会有较好的效果。

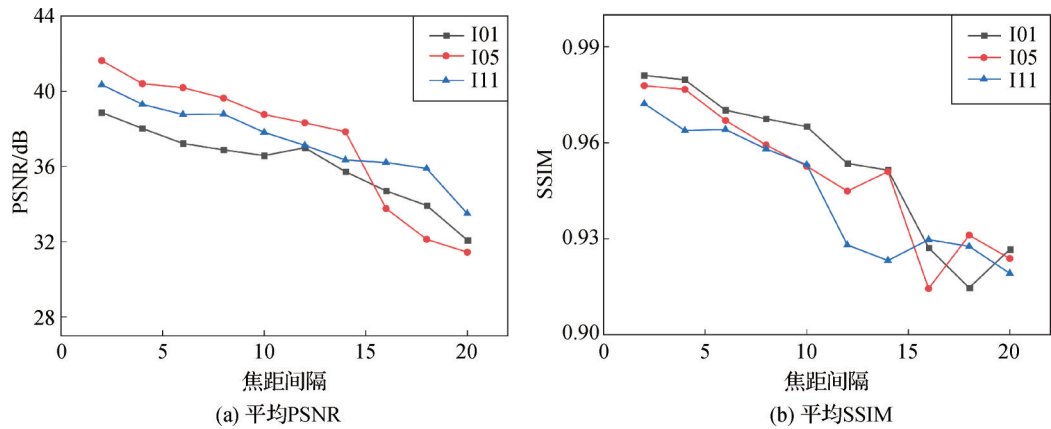


图10 各焦距间隔平均PSNR和SSIM结果

Fig. 10 Average PSNR and SSIM results for each focal length interval
((a) average PSNR; (b) average SSIM)

为了与光场重聚焦的结果(Wu等,2022)进行对比,本文额外使用光场重聚焦后得到的焦栈图像进行实验。选取3组焦栈序列,数据集如图11所示,3组序列所有图像的分辨率均为624×432像素。依然选取21幅,取奇数编号作为真值、偶数编号作为输入。由光场重聚焦的结果作为真值与生成结果计算PSNR和SSIM,从而得出本文方法生成结果与光场重聚焦结果的相似程度。

实验结果如图12所示,图12展示了各序列不同焦距编号下生成的结果与真值之间的PSNR和



图11 光场重聚焦数据集展示

Fig. 11 Display of light field refocus dataset ((a)—(c) 3 focal stack sequences by light field refocusing)

SSIM。可以看出,生成结果与真值之间具有较高的PSNR和SSIM。各焦距编号PSNR均在37 dB以上,SSIM均在0.9以上。

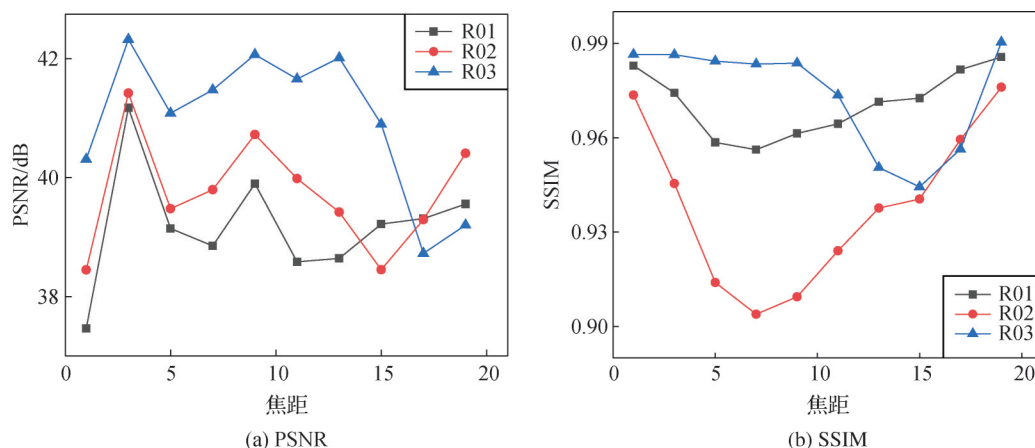


图12 光场重聚焦序列PSNR和SSIM结果

Fig. 12 PSNR and SSIM results of light field refocusing sequence ((a) PSNR; (b) SSIM)

3 结论

本文提出了一种基于高斯—维纳表示的双向预测稠密焦栈图生成方法。通过高斯—维纳表示模型来模拟FoSIs之间聚焦和散焦的变化,通过图像退化—恢复关系来描述FoSIs之间的聚焦—散焦变化特性,在此基础上通过双向预测的方法生成新的焦栈图。该方法的创新点在于通过双向预测的方法可以利用两幅焦栈图像的信息来生成新的焦栈图。其包含两个核心点:1)双向拟合求解最佳拟合参数;2)根据最佳拟合参数求解预测参数。相比于单一的单向拟合,双向拟合通过两幅焦栈图像之间相互拟合可以搜索出最佳双向拟合参数,通过最佳双向拟合参数对焦栈图像进行变换可以模拟出更加接近两幅焦栈图像之间焦距变化趋势。在预测参数求解中,基于之前搜索到的最优变化趋势,通过对最佳双向拟合参数中的标准差参数 σ 求平均,来找到对应变化趋势中间状态的标准差参数,将该参数作为预测参数中的 σ_{pred} ,通过预测参数生成出的新焦栈图像更加接近于稀疏序列中相邻两幅焦栈图像对应焦距的中间值,从而使生成的新焦栈图像和原稀疏焦栈图像在焦距域呈均匀分布,使得稠密焦栈图的变化更加平滑。

本文方法可以通过一组稀疏的焦栈图像生成稠密的焦栈图像。根据实验结果,本文方法在11个焦栈序列上的平均PSNR为40.861 dB,平均SSIM为0.976,具有较好的客观结果。在主观结果上,本文方法也十分接近真值,具有较好的性能。

参考文献(References)

- Anger J, Facciolo G and Delbracio M. 2018. Estimating an image's blur kernel using natural image statistics, and deblurring it: an analysis of the Goldstein-Fattal method. *Image Processing on Line*, 8: 282-304 [DOI: 10.5201/ipol.2018.211]
- Bai Y C, Jia H Z, Jiang M, Liu X M, Xie X D and Gao W. 2020. Single-image blind deblurring using multi-scale latent structure prior. *IEEE Transactions on Circuits and Systems for Video Technology*, 30(7): 2033-2045 [DOI: 10.1109/TCSVT.2019.2919159]
- Chang J H R, Kumar B V K V and Sankaranarayanan A. 2018. Towards multifocal displays with dense focal stacks. *ACM Transactions on Graphics (TOG)*, 37(6): #198 [DOI: 10.1145/3272127.3275015]
- Fang L and Dai Q H. 2020. Computational light field imaging. *Acta Optica Sinica*, 40(1): #0111001 (方璐, 戴琼海. 2020. 计算光场成像. *光学学报*, 40(1): #0111001) [DOI: 10.3788/AOS202040.0111001]
- Fu M L, Zhao Y and Hou Z Y. 2023. An effective interpretation of defocusing and the corresponding defocus convolution kernel. *Optics and Laser Technology*, 160: #109035 [DOI: 10.1016/j.optlastec.2022.109035]
- Hog M, Sabater N, Vandame B and Drazic V. 2017. An image rendering pipeline for focused plenoptic cameras. *IEEE Transactions on Computational Imaging*, 3(4): 811-821 [DOI: 10.1109/TCL.2017.2710906]
- Hu Y Z, Wang Y J and Wang S. 2020. Scene target 3D point cloud reconstruction technology combining monocular focus stack and deep learning. *IEEE Access*, 8: 168099-168110 [DOI: 10.1109/ACCESS.2020.3022630]
- Jian Y S, Zhu D M, Fu Z T and Wen S Y. 2023. Remote sensing image segmentation network based on multi-level feature refinement and fusion. *Laser and Optoelectronics Progress*, 60(4): #0428002

- (菅永胜, 朱大明, 付志涛, 文诗雅. 2023. 多层次特征优化融合的遥感图像分割网络. 激光与光电子学进展, 60(4): #0428002) [DOI: 10.3788/LOP212864]
- Kang X K, Qiu J, Liu C and He D. 2020. Global imaging based on monomer subset of focal stack architecture. *Laser and Optoelectronics Progress*, 57(24): #241101 (亢新凯, 邱钧, 刘畅, 何迪. 2020. 基于聚焦堆栈单体数据子集架构的全局成像. 激光与光电子学进展, 57(24): #241101) [DOI: 10.3788/LOP57.241101]
- Kontogianni G, Chliverou R, Koutsoudis A, Pavlidis G and Georgopoulos A. 2017. Enhancing close-up image based 3D digitisation with focus stacking. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2-W5: 421-425 [DOI: 10.5194/ISPRS-ARCHIVES-XLII-2-W5-421-2017]
- Li S W, Lin D Y, Zhou X H, Zhang W, Chen D N, Yu B and Qu J L. 2021. Multifocal image scanning microscopy based on double-helix point spread function engineering. *Acta Physica Sinica*, 70(3): #038701 (李四维, 林丹樱, 邹小慧, 张伟, 陈丹妮, 于斌, 屈军乐. 2021. 基于双螺旋点扩散函数工程的多焦点图像扫描显微. 物理学报, 70(3): #038701) [DOI: 10.7498/aps.70.20200640]
- Li W T, Wang G J, Hu X W and Yang H Z. 2018. Scene-adaptive image acquisition for focus stacking//*Proceedings of the 25th IEEE International Conference on Image Processing (ICIP)*. Athens, Greece: IEEE: 1887-1891 [DOI: 10.1109/ICIP.2018.8451455]
- Liu D W, Nicolescu R and Klette R. 2016. Stereo-based bokeh effects for photography. *Machine Vision and Applications*, 27(8): 1325-1337 [DOI: 10.1007/s00138-016-0775-5]
- Liu Y Q, Du X, Shen H L and Chen S J. 2021. Estimating generalized Gaussian blur kernels for out-of-focus image deblurring. *IEEE Transactions on Circuits and Systems for Video Technology*, 31(3): 829-843 [DOI: 10.1109/TCSVT.2020.2990623]
- Ren J R, Fu X D, Wang M R, Zhao T Y, Wang Z J, Feng K, Liang Y S, Wang S W and Lei M. 2023. Advances in rapid three-dimensional wide field microscopy. *Chinese Journal of Lasers*, 50(3): #0307104 (任婧荣, 傅相达, 王孟瑞, 赵天宇, 汪召军, 冯坤, 梁言生, 王少伟, 雷铭. 2023. 快速宽场三维显微技术研究进展. 中国激光, 50(3): #0307104) [DOI: 10.3788/CJL221303]
- Ren S X, Zhang C S, Cao H Q, Lin D Y, Yu B and Qu J L. 2022. Three-dimensional multifocal two-photon laser scanning microscopy based on double-helix point spread function engineering. *Acta Optica Sinica*, 42(14): #1411001 (任苏霞, 张晨爽, 曹慧群, 林丹樱, 于斌, 屈军乐. 2022. 基于双螺旋点扩展函数工程的三维多焦点双光子激光扫描显微技术. 光学学报, 42(14): #1411001) [DOI: 10.3788/AOS202242.1411001]
- Ruan L Y, Chen B, Li J Z and Lam M L. 2021. AIFNet: all-in-focus image restoration network using a light field-based dataset. *IEEE Transactions on Computational Imaging*, 7: 675-688 [DOI: 10.1109/TCI.2021.3092891]
- Sun L T, Wang L J, Wang W S and Liu M X. 2020. Design of dual-focal-plane helmet mounted display based on single picture generation unit. *Acta Optica Sinica*, 40(13): #1322004 (孙路通, 王灵杰, 王蔚松, 刘铭鑫. 2020. 基于单图像生成单元的双焦面头盔显示光学系统设计. 光学学报, 40(13): #1322004) [DOI: 10.3788/AOS202040.1322004]
- Sun Z H, Ke Q H, Rahmani H, Bennamoun M, Wang G and Liu J. 2023. Human action recognition from various data modalities: a review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(3): 3200-3225 [DOI: 10.1109/TPAMI.2022.3183112]
- Wang Q and Fu Y T. 2020. Single-image refocusing using light field synthesis and circle of confusion rendering. *Acta Optica Sinica*, 40(1): #0111021 (王奇, 傅雨田. 2020. 利用光场合成与弥散圆渲染的单幅图像重聚焦. 光学学报, 40(1): #0111021) [DOI: 10.3788/AOS202040.0111021]
- Wang T T, Piao Y R, Lu H C, Li X and Zhang L H. 2019. Deep learning for light field saliency detection//*Proceedings of 2019 IEEE/CVF International Conference on Computer Vision (ICCV)*. Seoul, Korea (South): IEEE: 8837-8847 [DOI: 10.1109/ICCV.2019.00893]
- Wang X, Liu Q, Peng Z J, Hou J H, Yuan H, Zhao T S, Qin Y, Wu K J, Liu W Y and Yang Y. 2023. Research progress of six degree of freedom (6DoF) video technology. *Journal of Image and Graphics*, 28(6): 1863-1890 (王旭, 刘琼, 彭宗举, 侯军辉, 元辉, 赵铁松, 秦熠, 吴科君, 刘文予, 杨铀. 2023. 6DoF 视频技术研究进展. 中国图象图形学报, 28(6): 1863-1890) [DOI: 10.11834/jig.230025]
- Wang Y T, Cheng D W and Xu C. 2016. Display technologies in virtual reality systems. *Scientia Sinica Informationis*, 46(12): 1694-1710 (王涌天, 程德文, 许晨. 2016. 虚拟现实光学显示技术. 中国科学: 信息科学), 46(12): 1694-1710 [DOI: 10.1360/N112016-00247]
- Wu K J, Yang Y, Liu Q and Zhang X P. 2022. Gaussian-wiener representation and hierarchical coding scheme for focal stack images. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(2): 523-537 [DOI: 10.1109/TCSVT.2021.3066523]
- Wu K J, Yang Y, Yu M and Liu Q. 2020. Block-wise focal stack image representation for end-to-end applications. *Optics Express*, 28(26): 40024-40043 [DOI: 10.1364/oe.413523]
- Xiang S, Huang N T, Deng H P and Wu J. 2022. Estimation of light field depth based on multi-level network optimization. *Laser and Optoelectronics Progress*, 59(10): #1010009 (向森, 黄楠婷, 邓慧萍, 吴谨. 2022. 基于多级网络优化的光场深度值估计. 激光与光电子学进展, 59(10): #1010009) [DOI: 10.3788/LOP202259.1010009]
- Xie N Y, Ding Y Y, Li M Y, Liu Y, Lyu R M and Yan T. 2022. Light field image re-focusing based on conditional generative adversarial networks leverage. *Journal of Image and Graphics*, 27(4): 1056-1065 (谢柠宇, 丁宇阳, 李明悦, 刘渊, 律睿睿, 晏涛. 2022. 利用条件生成对抗网络的光场图像重聚焦. 中国图象图形学报, 27(4): 1056-1065) [DOI: 10.11834/jig.200471]

- Xie P Y, Yang J F, Xue B and Chen G Q. 2017. Simulation of light field imaging and refocusing models based on matrix transformation. *Acta Photonica Sinica*, 46(5): #0510001 (解培月, 杨建峰, 薛彬, 陈国庆. 2017. 基于矩阵变换的光场成像及重聚焦模型仿真. *光子学报*, 46(5): #0510001) [DOI: 10.3788/gzxb20174605.0510001]
- Zhang C P and Wang Q. 2016. Survey on imaging model and calibration of light field camera. *Chinese Journal of Lasers*, 43(6): #0609004 (张春萍, 王庆. 2016. 光场相机成像模型及参数标定方法综述. *中国激光*, 43(6): #0609004) [DOI: 10.3788/CJL201643.0609004]
- Zheng H Y, Yong H W and Zhang L. 2022. Unfolded deep kernel estimation for blind image super-resolution//Proceedings of the 17th European Conference on Computer Vision. Tel Aviv, Israel:

Springer: 502-518 [DOI: 10.1007/978-3-031-19797-0_29]

作者简介

王其腾,男,硕士研究生,主要研究方向为计算机视觉与新视点合成。E-mail:wangqiteng@hust.edu.cn

杨铀,通信作者,男,教授,主要研究方向为多媒体通信、多媒体技术和计算摄像学。E-mail:yangyou@hust.edu.cn

李志龙,男,博士研究生,主要研究方向为计算成像、图像融合和计算机视觉。E-mail:zhilong_li@hust.edu.cn

丁新,男,博士后,主要研究方向为计算机视觉、图像增强和人脸超分。E-mail:2023511212@hust.edu.cn

刘琼,女,教授,主要研究方向为多媒体通信、3D视频处理和理解和3D计算机视觉。E-mail:q.liu@hust.edu.cn