

用于简单透镜成像系统的深度迭代滤波网络

黄毅*, 熊涛

武汉光电国家研究中心华中光电技术研究所, 湖北 武汉 430223

摘要 提出一种基于迭代自适应滤波原理的端到端深度神经网络。该网络旨在解决由简单透镜的光学结构引起的显著图像边缘模糊问题。利用具有大视场的单个胶合透镜,提出一种像素级去模糊滤波器,该滤波器可有效地适应模糊的空间变化,从而恢复输入图像的模糊特征。通过模拟和在原型摄像机系统上进行的实验验证了所提方法的有效性。

关键词 计算成像技术; 图像退化模型; 图像重建; 大视场; 深度学习

中图分类号 O436

文献标志码 A

DOI: 10.3788/LOP232176

Deep Iterative Filter Adaptive Network for Simple Lens Imaging System

Huang Yi*, Xiong Tao

Huazhong Institute of Electro-Optics, Wuhan National Laboratory for Optoelectronics,
Wuhan 430223, Hubei, China

Abstract Herein, an end-to-end deep neural network based on iterative adaptive filtering principle is proposed. This network aims to solve the significant image edge blurring caused by the optical structure of simple lenses. A pixel level deblurring filter is proposed, using a single glued lens with a large field of view, to effectively adapt to the spatial changes of blur and restore the blurry features of the input image. The effectiveness of the proposed method is verified through simulation and experiments conducted on a prototype camera system.

Key words computational imaging technology; image degradation model; image reconstruction; large field of view; deep learning

1 引言

在现实世界的成像系统中,捕获图像的质量经常受到光学像差的影响。为了解决这个问题,传统的相机镜头包括多个精心设计的组件,以减轻这些像差。然而,这种方法导致相机系统既昂贵又笨重^[1]。

在最近几十年里,图像处理算法的快速进展引发了研究人员对计算成像技术(CIT)的兴趣^[2]。CIT结合光学设计和图像处理模块,可实现卓越的成像结果。与引入复杂的复合透镜设计的方法不同,CIT表明图像处理能够有效校正几乎所有光学像差^[3]。模糊种类可以分为大气模糊、运动模糊、离焦模糊,以及本文主要处理的光学像差导致的模糊。作为一种替代方法,计算光学像差校正利用算法处理简单透镜成像并减轻像差。然而,现有方法要么依赖于预测的像差模型,要么在有限的 10° 视场(FOV)内操作^[1,4],从而限制了其适用性。

近年来,深度学习由于能够建模复杂的数据分布,已广泛应用于去模糊问题^[5-7]。深度学习方法可以直接学习模糊图像到其相应清晰图像的映射^[8]。滤波自适应网络(FAN)被提出,以促进各种任务中特征的空间自适应处理^[9-15]。FAN通常由两部分组成:预测空间自适应滤波器和使用预测滤波器的特征转换,后者被称为滤波自适应卷积(FAC)。各种FAN相继被提出,并应用于不同的任务,如帧插值^[9-10]、去噪^[11]、超分辨率^[12]、语义分割^[13]和点云分割^[14]。对于去模糊任务,Zhang等^[15]提出用于运动去模糊的像素级自适应卷积。然而,他们的方法需要显著的计算成本,因为递归神经网络必须在每个像素上运行。

深度学习方法在处理具有大FOV和非均匀模糊的图像时遇到挑战^[16]。为了解决这个问题,本文提出一种创新的深度迭代滤波自适应网络(DIFAN),专门处理大视场和空间变化引起的光学像差。DIFAN的一个关键组成部分是新颖的深度迭代自适应卷积

收稿日期: 2023-09-21; 修回日期: 2023-10-06; 录用日期: 2023-11-01; 网络首发日期: 2023-11-21

通信作者: huang2020bit@163.com

(DIAC)层,该层生成逐像素的去模糊滤波器并应用于输入的模糊图像,可有效减轻光学像差。通过深度可分卷积并迭代应用可分离滤波器,DIAC层显著提高了去模糊的质量,同时最小化了计算开销。

本文的主要贡献包括:1)提出一种用于处理大FOV(具体来说,20°)下数百像素大小模糊的离轴像差滤波自适应网络;2)引入一种新颖的DIAC层,增强了网络的有效感受野(ERF),从而提高去除大模糊的能力;3)使用屏幕捕获实验制作了现实世界模糊与清晰图像数据集,并成功使用DIFAN进行图像复原。

2 实验原理

2.1 光学像差校正

自制的单片胶合光学透镜具有50 mm的焦距,使用一个保持环来固定透镜并充当简单的光圈,如图1所示。自制的光学透镜的参数可以在ZEMAX源项目文档中找到。图1(a)中的光路分析和后续的模拟是基于无限物体距离的前提条件进行的。由于其简单的结构,其光学像差会在视角增大后,急剧增大。后续的仿真实验及现实实验的目的都是校正其在20°FOV下的光学像差。

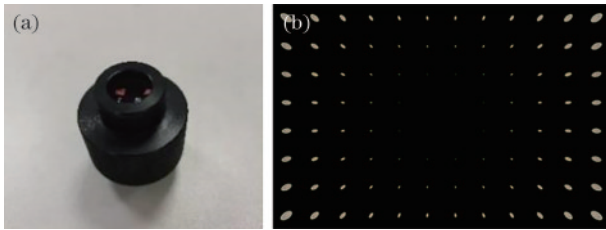


图1 制作光学透镜的过程:(a)自行设计的单片胶合透镜;(b)20°FOV下镜头的点扩散函数(PSF)

Fig. 1 The process of making the optical lens. (a) Self-designed single glued lens; (b) PSF of lens at 20° FOV

2.2 像差分析和图像退化模型

简单透镜会引入降低成像质量的像差。这些像差可以分为轴上和轴外像差。轴上像差主要包括沿透镜轴的纵向色差和球差。轴外像差包括横向色差、彗差、场曲率、散光和畸变。这些像差综合影响成像质量,正如PSF的复杂性所示。PSF在不同的RGB通道中表现出变化,如图像中的非均匀性以及大内核尺寸。

$$I_c(x, y) = \sum_{(m, n)} I_s(m, n) K(m, n, x, y) + N_c(x, y), \quad (1)$$

式中: $I_c(x, y)$ 代表成像系统捕获的源图像, (x, y) 代表观察图像空间的像素位置; $I_s(m, n)$ 代表真实地面图像, (m, n) 代表真实地面图像空间的像素位置; $K(m, n, x, y)$ 为运动模糊的PSF; $N_c(x, y)$ 是噪声。观察到的模糊图像 $I_c(x, y)$ 是已知的,而真实地面图像 $I_s(m, n)$ 、点扩散函数 $K(m, n, x, y)$,以及噪声 $N_c(x, y)$ 都是未知的,这是一个典型的逆问题。

从数学角度来看,病态问题可以通过在优化框架中添加先验约束来解决。然而,传统的基于优化的恢复方法通常涉及多次迭代^[4],这意味着图像处理需要更多的时间。幸运的是,深度学习理论和GPU加速硬件的进步促进了基于深度学习原理的神经网络去模糊方法的发展,其在图像恢复方面表现出了令人印象深刻的性能^[8]。本文提出一种依赖于深度学习原理的恢复方法,实现了端到端的图像恢复过程。所提方法为解决病态问题和基于优化方法所面临的挑战提供了有希望的解决方案。

2.3 深度迭代自适应卷积层

所提方法中的DIAC层在很大程度上依赖于可分离自适应卷积(SAC)。先前的研究,如Xu等^[17]的工作已经证明,使用一维滤波器的卷积网络可以有效近似大型反卷积任务中的逆滤波器。在类似的思路下, Lee等^[18]通过迭代应用可分离滤波器(IAC)来解决这个问题。

鉴于这些发现,所提DIAC层采用一维滤波器而不是二维卷积,这使我们能够在不增加额外的内存和计算成本的情况下实现更大的感受野。为了进一步提高性能,在DIAC层中引入深度大内核卷积作为额外的组成部分,这使网络能够捕获更全面的信息,同时保持效率。

如图2所示,DIAC以特征图 e 和空间可变卷积滤波器 F 作为输入。 e 和 F 具有相同的空间大小 $h \times w$ 。输入特征图 e 具有 c 个通道。 F 中的每个空间位置都是一个 $c(2k+1)$ 向量。对于每个空间位置 (x, y) ,根据每个通道,它们会分为2个 ck 向量外加一个偏置 c 。这样在SAC层重新塑形后变为两个一维滤波器。然后,该层以通道方式将这些滤波器应用于 e 中以 (x, y) 为中心的特征,生成一个输出特征图 \hat{e} ,其大小为 $\mathbf{R}^{h \times w \times c}$ 。

网络中的深度大核卷积块(DLKD)设计的动机是以完全卷积方式探索局部和全局依赖关系,如图3所示。第一部分类似于深度可分卷积,但在初始阶段采用了异常大的内核大小(例如 9×9),然后是内核大小为 1×1 的点卷积,并在它们之间添加了内部捷径。与标准卷积层不同,后者通过3D滤波器同时混合空间和通道维度,而我们的特征混合器分别在内部空间通道和深度通道特征上操作。这允许混合远距离的空间位置,并且与大内核一起,实现了大的有效感受野。我们的设计受到MLP-Mixer^[19]和Conv-Mixer^[20]的启发,它们也分别混合了空间和通道维度。前者强调多层感知器(MLP)的重要性,而后者侧重于对视觉任务的有效性进行分块嵌入。

后半部分包括具有 3×3 内核的深度卷积层,用于高效编码局部信息。与文献^[21]类似,为了实现特征的高效传播和融合,通过门控机制^[22]添加了一个额外的路径,随后以GeLU激活函数^[23]作为门控。

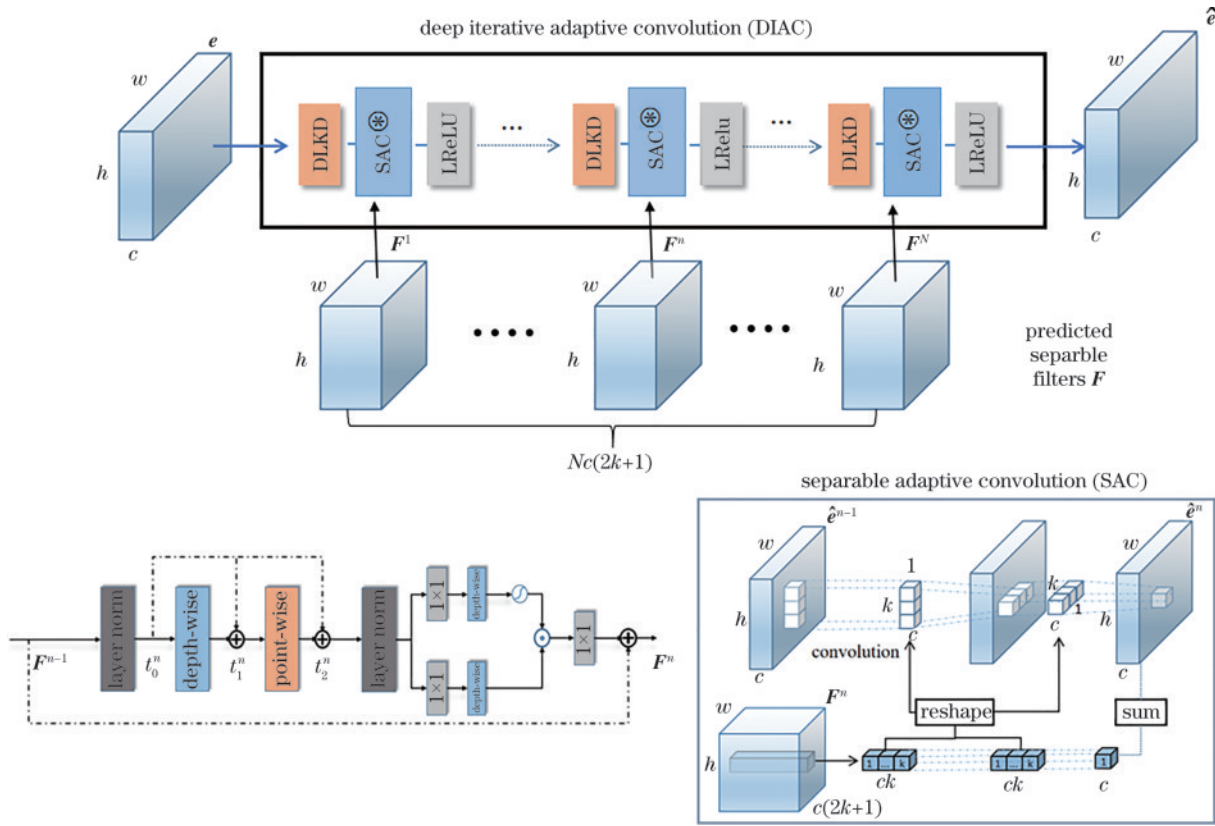


图 2 深度迭代自适应卷积层
Fig. 2 Deep iterative adaptive convolution

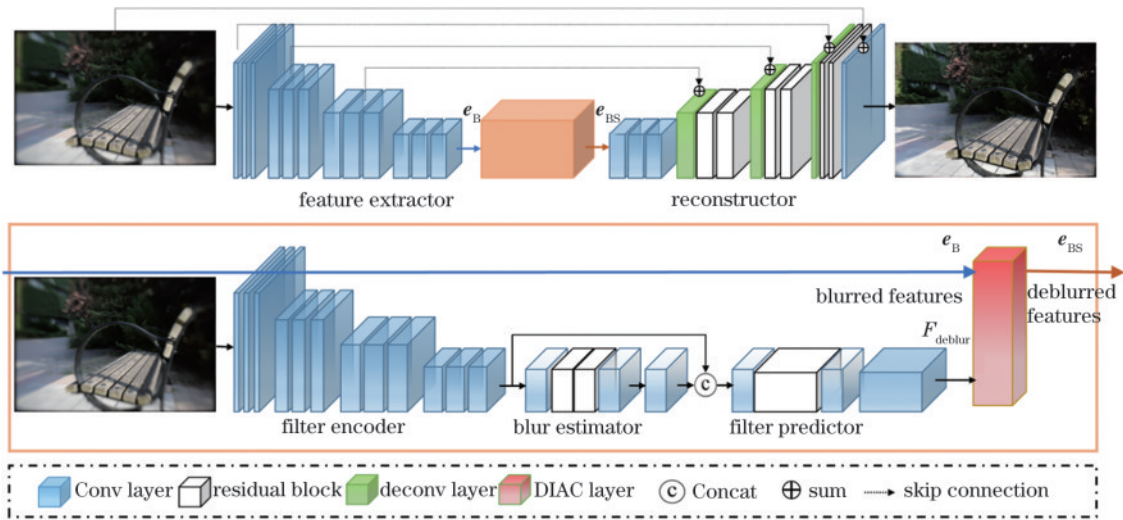


图 3 深度迭代自适应去模糊网络
Fig. 3 Deblurring network with DIFAN

首先介绍特征混合模块,该模块输出第 n 个 DLKD 中的特征 t^n :

$$\begin{cases} t_0^n = \text{LN}(F^{n-1}) \\ t_1^n = t_0^n + \delta(t_0^n), \\ t_2^n = t_1^n + \varphi(t_0^n) \end{cases} \quad (2)$$

式中: δ 和 φ 分别表示深度可分离卷积和点卷积;LN 是图 2 所示的层归一化。特征融合过程的输出特征 F^n 可表示为

$$F^n = F^{n-1} + \text{LN} \left\{ \text{GeLU} \left\{ \delta \left[C_1(t^n) \right] \odot \delta \left[C_2(t^n) \right] \right\} \right\}, \quad (3)$$

式中: C_1 和 C_2 是两个独立的 1×1 卷积层,后跟元素逐个相乘(表示为 \odot),然后是 GeLU 激活函数; δ 为深度可分离卷积,使用 3×3 内核。所提网络包括以分层方式分布的 DLKD,允许更大的 ERF,并对精细细节恢复作出重要贡献。

2.4 深度迭代自适应去模糊网络

图 3 为基于 DIFAN 的去模糊网络。所提网络以模糊图像作为输入,记为 I_b ,生成去模糊的输出,记为 I_o 。网络基于简单的编码器-解码器架构,由特征提取器、重构器和中间的 DIFAN 模块组成。特征提取器提取出模糊特征,记为 $e_b \in \mathbb{R}^{h \times w \times 3}$,并将它们提供给 DIFAN。DIFAN 通过预测空间变化的去模糊滤波器并利用 DIAC 将它们应用于 e_b 来在特征域中消除模糊,去模糊特征记为 e_{BS} 。 e_{BS} 随后传递给重构器以恢复模糊图像。

网络以模糊特征 e_b 作为输入,生成去模糊特征 e_{BS} 。为了生成去模糊特征,DIFAN 将 I_b 视为辅助输入,并从 I_b 中预测去模糊滤波器。滤波器编码器将 I_b 编码成一个表示,然后传递给模糊估计器。模糊估计器是专门设计的子网络,用于有效地训练 DIFAN,使用单透镜模糊图像的逐块最小像素(PMP)值。在模糊估计器之后,通过滤波器预测器来预测去模糊滤波器图,记为 $F_{deblur} \in \mathbb{R}^{H \times W \times C}$ 。最后,DIAC 层使用预测的滤波器 F_{deblur} 来转换输入特征 e_b ,生成去模糊特征 e 。

3 网络训练及仿真实验结果

3.1 仿真数据准备

使用 DPDD 数据集^[24]作为真实地面真实图像的数据源,图像大小为 1680×1120 。使用 ZEMAX 软件进行光线追迹仿真,得到光学像差模糊图像,形成一个

名为 DPDD-SL 的数据集,包括 450 对清晰和模糊图像。将数据集按照 7:1.5:1.5 的比例划分为训练集、验证集和测试集。

使用 PyTorch 框架^[25]实现了所提模型。在训练过程中,使用修正的 Adam 优化器^[26],参数设置如下: $\beta_1=0.9, \beta_2=0.99$,权重衰减率为 0.01。为了控制梯度爆炸,应用了梯度范数剪切,经验值为 0.5。网络经历了 60 万次训练迭代,初始学习率为 1.0×10^{-4} 。对于每次迭代,随机从训练集中选择一个图像批次。批次大小设置为 8,为了增强数据,应用了随机的高斯噪声、灰度图像转换和缩放。随后,图像被裁剪到 256×256 的大小以确保一致性。以去模糊结果与相应地面实图像之间的峰值信噪比(PSNR)、结构相似性(SSIM)、平均绝对误差(MAE)和学习感知图像块相似度(LPIPS)^[27]作为评估指标来评估去模糊性能。

3.2 仿真训练结果分析与对比

3.2.1 消融实验

比较了各种层和模块对所提网络的去模糊效果的影响,具体如表 1 所示。其中,第 1 行对应于基准模型,FP、DIAC、BPR 和 RBN 分别代表滤波预测器、深度迭代自适应卷积、模糊估计器和再模糊网络。为了公平评估并获得基准模型,将所提模型组件替换为具有相似模型参数数量(Params)和计算成本(MACs)的常规卷积层和残差块。

表 1 消融研究对比

Table 1 Comparison of ablation studies

DIFAN	Evaluations on the DPDD-SL dataset ^[24]				Computational costs	
	PSNR \uparrow	SSIM \uparrow	MAE \downarrow	LPIPS \downarrow	Params / 10^6	MACs / 10^9
FP	23.88	0.723	0.041	0.368	10.57	364.3
FP+DIAC	25.78	0.789	0.035	0.280		
FP+DIAC+BPR	26.37	0.824	0.032	0.232	10.47	419.5
FP+DIAC+BPR+RBN	26.94	0.847	0.029	0.221		

3.2.2 与其他方法的实验结果比较

在 DPDD-SL 测试数据集上进行了不同算法的结果,具体如表 2 所示。对于非学习方法,使用相同的测试数据集。对于基于学习的方法,使用相同的数据集进行训练,其他基准参数进行了调整以保持一致性,所有训练都进行 60 万次迭代。批处理大小为 256×256 。

图 4 为不同图像重建算法的结果。在大视场退化

模型的影响下,模糊图像呈现出严重的空间变化特征,如图 4(a)所示。图 4(b)为基于 ISD 的传统方法的处理结果,通过自适应正则化对图像进行处理^[28]。可以看到,传统方法可以清晰地恢复强边缘,但纹理和亮度仅部分恢复,存在明显的锐化,因为它不能处理大变化的模糊核。此外,基于 GAN 的重建算法^[7]的性能也有了显著改进,如图 4(c)所示。然而,与图 4(e)

表 2 重建性能的定量比较

Table 2 Quantitative comparisons of reconstruction performance

Model	Evaluations on the DPDD-SL dataset				Computational costs		
	PSNR \uparrow	SSIM \uparrow	MAE \downarrow	LPIPS \downarrow	Params / 10^6	MACs / 10^9	Time / s
Original	22.15	0.717	0.049	0.331			
ISD-Deblur ^[28]	23.78	0.75	0.039	0.369			45
Deblur-Ganv2 ^[7]	26.01	0.82	0.028	0.319	33.15	858.5	2.67
Proposed model	26.94	0.847	0.029	0.221	10.47	419.5	0.677

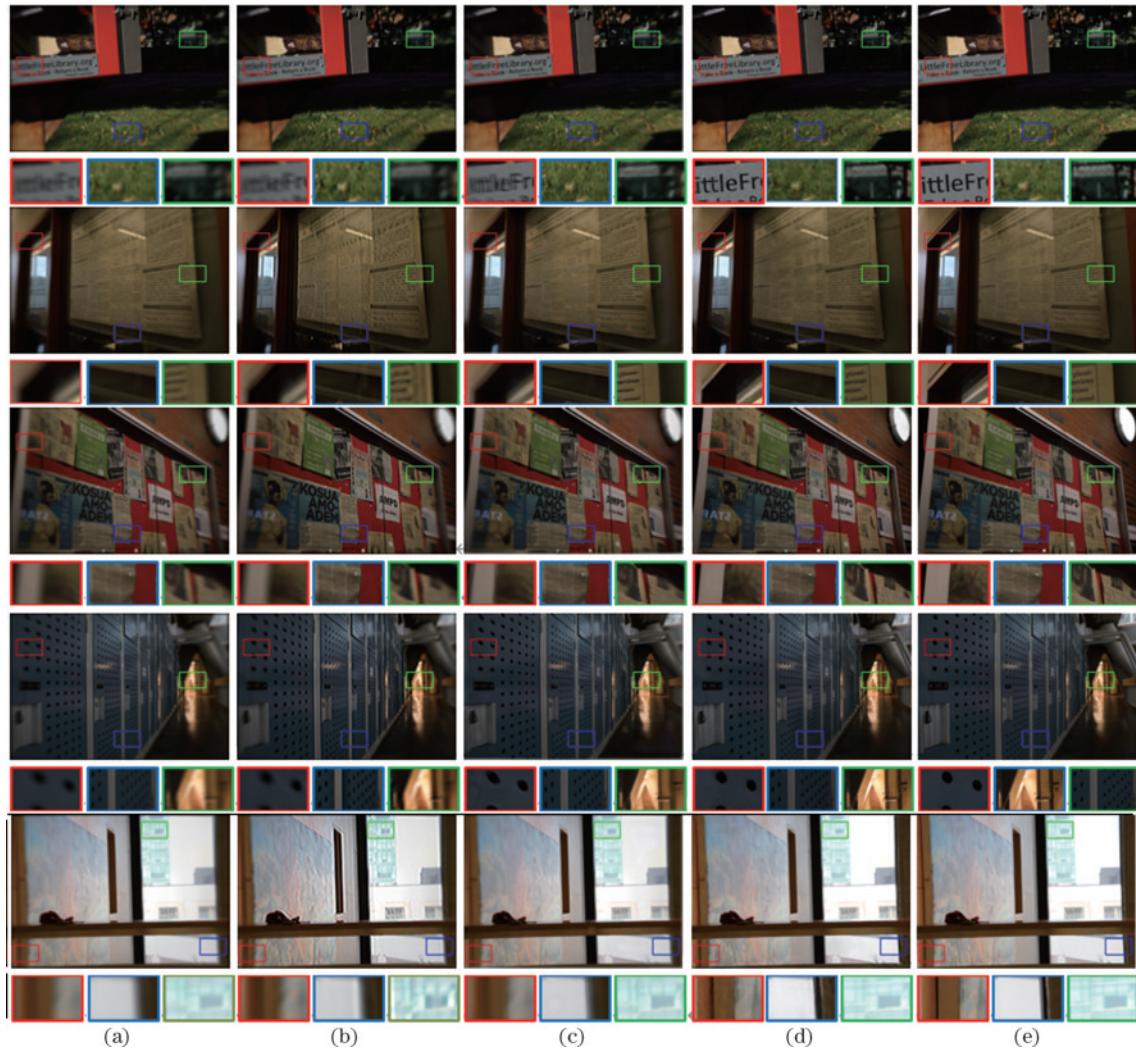


图 4 基于模拟数据集的不同复原方法的比较。(a)原始模糊图像;(b) ISD-Deblur复原结果;(c) DeblurGAN-v2复原结果;(d)所提 DIFAN复原结果;(e)地面真实图像

Fig. 4 Comparison of different restoration methods based on simulation dataset. (a) Original blurred image; (b) recovery results of ISD-Deblur; (c) recovery results of DeblurGAN-v2; (d) recovery results of proposed DIFAN; (e) ground truth images

中的地面实图像相比,其边缘细节仍有偏差,并且在处理边缘上的大内核模糊时表现不如所提网络。所提 DIFAN 的处理结果如图 4(d)所示,获得了更平滑的平面区域、更清晰的纹理细节和更准确的颜色像差校正。

通过计算 PSNR、SSIM 和 LPIPS^[27] 指标来客观评估模型的性能。根据网络参数数量、在 1680×1120 图像上计算的乘积累加操作数量,以及测试图像的平均计算时间来测量计算成本,具体如表 2 所示。对于不基于学习的传统方法,如 ISD-Deblur,只测量其计算时间。可以看出,所提 DIFAN 获得的客观评估指标优于其他重建算法,与图 4 中的主观结果一致。实验结果表明,所提模型在保持较低的计算成本时性能显著优于 DeblurGan-v2 和传统算法。

3.2.3 DIFAN 复原泛化性实验

为了验证网络泛化性,进行了相关实验,结果如图 5 所示。可以看出,36° 视场角下的单胶合透镜的

PSF 在边缘处要比 20° 视场角增大接近 2 倍,是超大尺寸的模糊核,也就导致图片边缘处的光学像差非常严重,模糊图片在边缘处出现严重变形以及明显的黑边效应。但所提网络仍然具有一定的复原效果。

还进行了 15° 视场角下的平凸透镜实验对比,使用 Wu 等^[16] 给出的光学退化仿真数据集,用网络训练后,得到的复原效果见图 6。

不同视场角下的 DIFAN 复原对比实验的定量比较见表 3,可以看出,所提 DIFAN 在不同视场角下都具有良好的复原效果,并且在视场角为 20° 时性能达到最佳。

平凸透镜 15° 视场角下的不同网络复原方法对比实验的定量比较见表 4,使用相同数据集比较了几个具有代表性的网络以及 Wu 等^[16] 的 RRG-GAN。在相同数据集下,所提网络的复原指标都高于其他网络,证明了所提网络对简单透镜光学退化图像的复原能力。

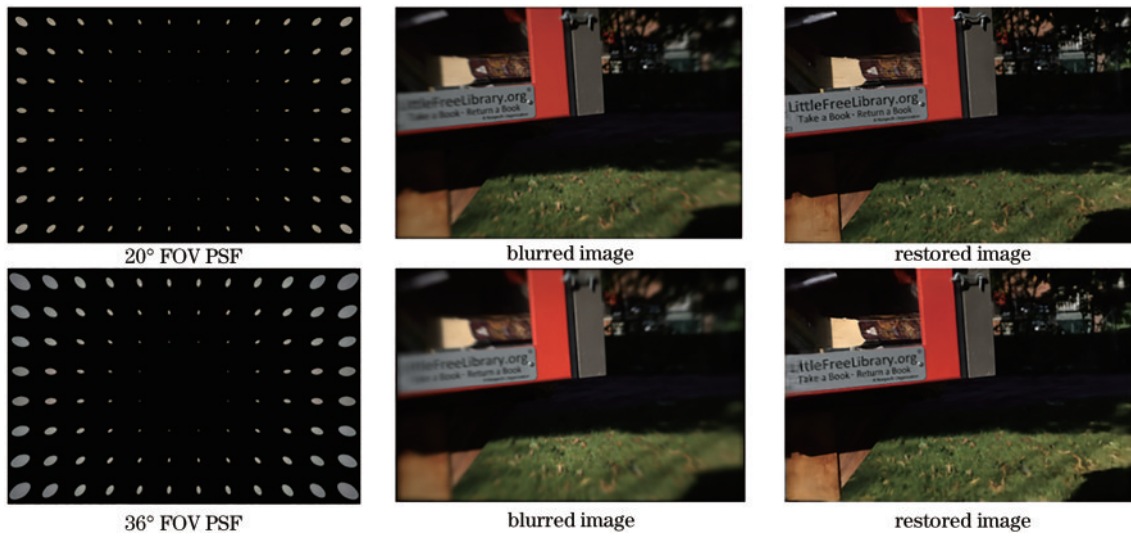


图 5 基于不同视场角下的模拟数据集的网络复原效果比较

Fig. 5 Comparison of network restoration effects based on simulation datasets at different field angles



图 6 平凸透镜在 15° 视场角下仿真数据集的网络复原效果图

Fig. 6 Network restoration effect of simulation data set with a plano-convex lens at 15° field of view angle

表 3 不同视场角下的 DIFAN 复原对比实验的定量比较

Table 3 Quantitative comparison of DIFAN restoration comparison experiments under different field of view angles

Experiment	Evaluations on the DPDD-SL data			
	PSNR ↑	SSIM ↑	MAE ↓	LPIPS ↓
15° FOV flat-convex lens	30.5241	0.88427	0.01859	0.10433
20° FOV single glued lens	26.9425	0.84679	0.02904	0.22172
36° FOV single glued lens	23.6107	0.76617	0.04471	0.32384

表 4 平凸透镜 15° 视场角下不同网络复原方法对比实验的定量比较

Table 4 Quantitative comparison of comparative experiments of different network restoration methods under 15° field of view angle of plano-convex lens

Model	PSNR ↑	SSIM ↑
Original	24.4386	0.6278
Multiscale	25.0048	0.6859
Fov-GAN	25.2653	0.7486
Deblur-GAN ^[7]	28.0252	0.7843
RRG-GAN ^[16]	30.4102	0.8650
Proposed DIAN	30.5241	0.8842

4 现实实验

实际情况中的图像受到多种因素的影响,不仅包括



图 7 畸变矫正和显示捕获实验

Fig. 7 Distortion correction and display-capture experiments

在开始图像捕获程序之前,进行以下设置:

- 1) 使用 i1 Pro 校准套件来校准 LCD 监视器的色调曲线和色彩再现;
- 2) 通过捕获白色校准图像,校准系统的均匀性,包括 LCD 显示器的亮度均匀性和捕获相机的成像渐晕;
- 3) 通过捕获 LCD 显示器上显示的几种已知棋盘图案[图 7 左图],获得了捕获图像的粗畸变校正参数

光学像差,还包括色移和杂散光。色移通常由 CMOS 相机的白平衡问题引起,而明亮的背景干扰则是由于杂散光引起的。由于单片胶合透镜成像系统没有专门的孔径,压盖环到孔径的过程会增加杂散光的产生,从而直接影响成像质量。杂散光会导致背景明亮,进一步降低光学系统的调制传递函数(MTF)。这两种额外的干扰因素导致模拟数据集与实际场景的数据分布存在差异,因此需要进一步构建一个更符合实际场景的数据集。

4.1 显示捕获实验

由于现实拍摄样本较少,考虑到数据集的不足可能导致网络的过拟合等问题,与 Peng 等^[6]类似,构建了一个显示捕获实验室设置,如图 7 所示。将数据集中的 900 张图像显示在一个 LCD 设备上,并通过单胶合透镜成像系统进行采集。

以及捕获图像与显示器上显示的真实图像之间的对齐转移矩阵。

对于屏幕捕获实验,使用部分 MIT-Adobe Five K 数据集和 DIV2K 数据集,总共拍摄了 900 对清晰和模糊图像对。由于 CMOS 拍摄软件的输出分辨率为 1920×1080,因此显示器的分辨率也设置为 1920×1080。最终,获得 900 对混合现实的清晰和模糊图像对,图 8 为其中一组。



图 8 使用显示捕获设备拍摄的图像对

Fig. 8 Image pairs taken using display capture devices

从图 8 可以观察到,在实际场景中,使用仿真测试集训练的网络参数并不表现出良好的效果。这一问题的主要原因在于仿真数据集的数据分布与真实场景存在不一致性。除了传统的光学像差导致的图像模糊

外,还存在色移和明亮背景干扰这两个额外的退化因素,这些因素在 ZEMAX 软件中并未进行模拟。

4.2 实验结果

与前文的仿真实验网络训练过程相同,只不过使

用现实拍摄采集的数据集 Real-DIV 来获取现实情况下各种不确定因素对网络的影响。在经过 1800 轮训练后,

复原图像如图 9 所示。可以看出,所提方法不仅可以消除光学模糊,还可以纠正颜色偏差和明亮背景干扰。



图 9 实验结果。(a)原始模糊图像;(b)网络复原图像结果;(c)地面真实图像

Fig. 9 Experimental results. (a) The original blurred images; (b) image recovery results from the network; (c) ground truth images

5 结 论

提出一种用于从简单透镜成像系统捕获的模糊图像中恢复潜在图像的新型 DIFAN。所提方法解决了由具有 20° 视场的单片胶合透镜引起的严重像差所带来的挑战。DIFAN 包括多个关键组件,包括模糊估计模块、深度大内核卷积模块和深度迭代自适应卷积。这些组件使 DIFAN 适用于大视场非均匀去模糊任务。

为了评估所提方法的性能,分别使用 ZEMAX 模拟和现实实验收集了两组具有不同锐度和像差水平的模糊图像数据集。大量的实验比较和实际测试结果证实了所提方法的有效性和泛化性。此外,该方法可以集成到现有的光学设计方法中,从而克服大视场设计中的 PSF 一致性限制,并实现更高分辨率的成像。

参 考 文 献

- [1] Heide F, Rouf M, Hullin M B, et al. High-quality computational imaging through simple lenses[J]. ACM Transactions on Graphics, 2013, 32(5): 149.
- [2] 刘飞, 吴晓琴, 赵琳, 等. 广域高分辨率计算光学成像系统研究进展[J]. 激光与光电子学进展, 2021, 58(18): 1811001.
- [3] Liu F, Wu X Q, Zhao L, et al. Research progress of wide-field and high-resolution computational optical imaging system[J]. Laser & Optoelectronics Progress, 2021, 58(18): 1811001.
- [4] Li W L, Liu Y, Yin X Q, et al. A computational photography algorithm for quality enhancement of single lens imaging deblurring[J]. Optik, 2015, 126(21): 2788-2792.
- [5] Schuler C J, Hirsch M, Harmeling S, et al. Blind correction of optical aberrations[M]//Fitzgibbon A, Lazebnik S, Perona P, et al. Computer vision-ECCV 2012. Lecture notes in computer science. Heidelberg: Springer, 2012, 7574: 187-200.
- [6] 邸江磊, 唐唯, 吴计, 等. 卷积神经网络在光学信息处理中的应用研究进展[J]. 激光与光电子学进展, 2021, 58(16): 1600001.

[1] Heide F, Rouf M, Hullin M B, et al. High-quality

- Di J L, Tang J, Wu J, et al. Research progress in the applications of convolutional neural networks in optical information processing[J]. *Laser & Optoelectronics Progress*, 2021, 58(16): 1600001.
- [6] Peng Y, Sun Q, Dun X, et al. Learned large field-of-view imaging with thin-plate optics[J]. *ACM Transactions on Graphics*, 2019, 38(6): 219.
- [7] Kupyn O, Martyniuk T, Wu J R, et al. DeblurGAN-v2: deblurring (orders-of-magnitude) faster and better[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea (South). New York: IEEE Press, 2020: 8877-8886.
- [8] Zhang H G, Dai Y C, Li H D, et al. Deep stacked hierarchical multi-patch network for image deblurring [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: 5971-5979.
- [9] Niklaus S, Mai L, Liu F. Video frame interpolation via adaptive convolution[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 2270-2279.
- [10] Niklaus S, Mai L, Liu F. Video frame interpolation via adaptive separable convolution[C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 261-270.
- [11] Mildenhall B, Barron J T, Chen J W, et al. Burst denoising with kernel prediction networks[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: 2502-2510.
- [12] Wang X T, Yu K, Dong C, et al. Recovering realistic texture in image super-resolution by deep spatial feature transform[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: 606-615.
- [13] Su H, Jampani V, Sun D Q, et al. Pixel-adaptive convolutional neural networks[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: 11158-11167.
- [14] Xu C F, Wu B C, Wang Z N, et al. SqueezeSegV3: spatially-adaptive convolution for efficient point-cloud segmentation[M]//Vedaldi A, Bischof H, Brox T, et al. *Computer vision-ECCV 2020. Lecture notes in computer science*. Cham: Springer, 2020, 12373: 1-19.
- [15] Zhang J W, Pan J S, Ren J, et al. Dynamic scene deblurring using spatially variant recurrent neural networks [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: 2521-2529.
- [16] Wu X T, Li J C, Zhou G X, et al. RRG-GAN restoring network for simple lens imaging system[J]. *Sensors*, 2021, 21(10): 3317.
- [17] Xu L, Ren J S, Liu C, et al. Deep convolutional neural network for image deconvolution[C]//Annual Conference on Neural Information Processing Systems 2014, December 8-13, 2014, Montreal, Quebec, Canada. [S. l.: s.n.], 2014: 1790-1798.
- [18] Lee J Y, Son H, Rim J, et al. Iterative filter adaptive network for single image defocus deblurring[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 20-25, 2021, Nashville, TN, USA. New York: IEEE Press, 2021: 2034-2042.
- [19] Tolstikhin I, Houlsby N, Kolesnikov A, et al. MLP-mixer: an all-MLP architecture for vision[EB/OL]. (2021-05-04)[2023-03-05]. <https://arxiv.org/abs/2105.01601>.
- [20] Trockman A, Kolter J Z. Patches are all You need? [EB/OL]. (2022-01-24)[2023-03-05]. <https://arxiv.org/abs/2201.09792>.
- [21] Yuan L, Hou Q B, Jiang Z H, et al. VOLO: vision outlooker for visual recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, 45(5): 6575-6586.
- [22] Dauphin Y N, Fan A, Auli M, et al. Language modeling with gated convolutional networks[EB/OL]. (2016-12-23)[2023-05-06]. <https://arxiv.org/abs/1612.08083>.
- [23] Hendrycks D, Gimpel K. Gaussian error linear units (GELUs)[EB/OL]. (2016-06-27)[2023-05-06]. <https://arxiv.org/abs/1606.08415>.
- [24] Abuolaim A, Brown M S. Defocus deblurring using dual-pixel data[M]//Vedaldi A, Bischof H, Brox T, et al. *Computer vision-ECCV 2020. Lecture notes in computer science*. Cham: Springer, 2020, 12355: 111-126.
- [25] Wen F, Ying R D, Liu Y P, et al. A simple local minimal intensity prior and an improved algorithm for blind image deblurring[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2021, 31(8): 2923-2937.
- [26] Paszke A, Gross S, Chintala S, et al. Automatic differentiation in pytorch[EB/OL]. [2023-03-05]. <https://gwern.net/doc/www/openreview.net/54b149cfe1fbbe841975209b4840fa04086a701.pdf>.
- [27] 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.
- [28] Xu L, Jia J Y. Two-phase kernel estimation for robust motion deblurring[M]//Daniilidis K, Maragos P, Paragios N. *Computer vision-ECCV 2010. Lecture notes in computer science*. Berlin, Heidelberg: Springer, 2010, 6311: 157-170.