

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2024)12-3501-28

论文引用格式: Wang P, Zhu Y, Yan Q S, Sun J Q and Zhang Y N. 2024. Real-world image deblurring: challenges and prospects. Journal of Image and Graphics, 29(12):3501-3528(王珮, 朱宇, 闫庆森, 孙瑾秋, 张艳宁. 2024. 真实场景图像去模糊: 挑战与展望. 中国图象图形学报, 29(12):3501-3528)[DOI:10.11834/jig.230802]

真实场景图像去模糊: 挑战与展望

王珮¹, 朱宇^{1*}, 闫庆森¹, 孙瑾秋², 张艳宁¹

1. 西北工业大学计算机学院, 西安 710072; 2. 西北工业大学航天学院, 西安 710072

摘要: 图像去模糊是计算机视觉的基础任务, 对医学影像、监控摄像及卫星图像等领域具有重要意义。对于真实场景去模糊任务, 由于场景内可能存在多个目标以及复杂的运动, 成像过程容易受到许多外界因素的干扰, 例如噪声、光照等, 使图像去模糊问题更复杂。早期的研究主要针对仿真降质, 但由于仿真模型受到多种假设限制, 例如高斯噪声和全局一致运动等, 难以在真实场景下展现出良好的复原效果。因此, 越来越多的学者着手研究真实场景去模糊问题, 以提升去模糊方法在现实生活中的使用价值。当前对真实场景下去模糊问题的综述性研究尚处于空白阶段, 为此本文对真实场景去模糊任务进行了系统调研, 分析其中存在的挑战, 从降质模型的角度出发, 由浅入深, 由易到难, 将真实场景下的去模糊问题拆解开, 归纳为单一模糊去除方法、复合模糊去除方法以及真实场景下未知模糊去除方法, 全方位描述了当前学术界在该问题上的研究内容和方法, 总结和对标了各类方法的优缺点, 阐述了阻碍复原性能进一步提升的难点问题, 并对常用的一些数据集和评价指标进行了整理总结。最后, 对真实场景去模糊任务的未来发展前景和研究热点进行了展望, 并给出了可能的解决方法。

关键词: 图像去模糊; 真实场景; 非均匀模糊; 复合模糊; 未知降质表征

Real-world image deblurring: challenges and prospects

Wang Pei¹, Zhu Yu^{1*}, Yan Qingsen¹, Sun Jinqiu², Zhang Yanning¹

1. School of Computer Science and Engineering, Northwestern Polytechnical University, Xi'an 710072, China;

2. School of Astronautics, Northwestern Polytechnical University, Xi'an 710072, China

Abstract: Image deblurring is a fundamental task in computer vision that holds significant importance in various applications, such as medical imaging, surveillance cameras, and satellite imagery. Over the years, image deblurring has garnered much research attention, leading to the development of numerous dedicated methods. However, in real-world scenarios, the imaging process may be subject to various disturbances that can lead to complex blurring. Certain factors, such as inconsistent object motion, camera lens defocusing, pixel compression during transmission, and insufficient lighting, can lead to a range of intricate blurring phenomena that further amplify the deblurring challenges. In this case, image deblurring in real-world scenarios becomes a complex ill-posed problem, and conventional image deblurring based on simulated blurry degradations methods often falls short when confronted with these real-world deblurring challenges. These limitations are ascribed to the extent of assumptions on which these conventional methods depend. These assumptions include but are not limited to 1) traditional methods often assume that the noise in the image follows a Gaussian distribution; 2) spatially invariant uniform blur assumption; and 3) independence of the blurring phenomena assumption. Although conve-

收稿日期: 2023-11-20; 修回日期: 2024-03-13; 预印本日期: 2024-03-20

* 通信作者: 朱宇 yuzhu@nwpu.edu.cn

基金项目: 国家自然科学基金项目(61901384, 61871328, U19B2037)

Supported by: National Natural Science Foundation of China(61901384, 61871328, U19B2037)

nient for theoretical analysis and algorithm development, these assumptions prove to be restrictive when applied to the complex deblurring problems encountered in real-world scenarios. Consequently, there is a pressing need to conduct specialized research tailored to the challenges of real-world image deblurring and to enhance the effectiveness of image restoration methods. Real-world image deblurring is an intricate task that requires the development of innovative algorithms and techniques that are capable of accommodating the diversity of blurring factors and the complexities present in practical environments. This paper attempts to create unique deblurring solutions that can efficiently handle real-world scenarios and enhance the practical applicability of deblurring methods. One approach to addressing the above challenges is to design algorithms that are robust to various types of noise and are capable of handling non-uniform and coupled blurring effects. Additionally, machine learning and deep learning have emerged as powerful tools for addressing complex real-world deblurring problems. Deep learning models, such as convolution neural networks and generative adversarial networks, have shown remarkable adaptability in learning from diverse data and producing high-quality deblurred images. Furthermore, researchers are exploring the integration of multiple sensor inputs, including depth information, to improve deblurring accuracy and effectiveness. These multi-modal approaches leverage additional data sources to disentangle complex blurring effects and enhance deblurring performance. As real-world image deblurring continues to gain attention, the research community is expected to contribute valuable insights and develop innovative solutions to further improve image restoration in complex scenarios. The ongoing collaboration among researchers from different fields, including computer vision, machine learning, optics, and imaging, will likely yield breakthroughs in addressing real-world deblurring challenges. In conclusion, real-world image deblurring is a multifaceted problem that requires tailored solutions to overcome the limitations of conventional deblurring methods. By acknowledging the complexities of real-world blurring phenomena and harnessing the power of advanced algorithms, machine learning, and multi-modal approaches, researchers are working toward enhancing image restoration in practical, challenging environments. Despite the growing interest in real-world deblurring, there remains a dearth of comprehensive surveys on the subject. To bridge this gap, this paper conducts a systematic review of real-world deblurring problems. From the perspective of image degradation models, this paper delves into various aspects and breaks down the associated challenges into isolated blur removal methods, coupled blur removal methods, and methods for unknown blur in real-world scenarios. This paper also provides a holistic overview of the state-of-the-art research in this domain, summarizes and contrasts the strengths and weaknesses of various methods, and elucidates the challenges that hinder further improvements in image restoration performance. This paper also offers insights into the prospects and research trends in real-world deblurring tasks and offers potential solutions to the challenges ahead, including the following:

- 1) Shortage of paired real-world training data: Acquiring high-quality training data with blur and sharp images that accurately represent the diversity of real-world scenarios is a significant challenge. The scarcity of comprehensive, real-world datasets hinders the development of supervised deblurring tasks. To address the lack of real-world data, researchers are exploring data synthesis and unsupervised learning techniques. By generating synthetic data that simulate real-world scenarios, algorithms can be trained on highly diverse data, and unsupervised learning is particularly suitable for improving adaptability to real-world conditions.
- 2) Efficiency of complex models: Recent deblurring algorithms rely on complex deep learning models to achieve high-quality results. However, the models often result in computational inefficiency, making these algorithms impractical for real-time or resource-constrained applications. The computational overhead and memory requirements of these models also limit their deployment in practical scenarios. Researchers are developing highly efficient model architectures, such as lightweight neural networks and model compression techniques, to strike a balance between computational efficiency and deblurring performance, making them suitable for real-time applications and resource-constrained environments.
- 3) Overemphasis on degradation metrics: Many deblurring methods prioritize optimizing quantitative metrics related to the reconstruction of image details. While these metrics provide a quantitative measure of image quality, they may not align with the perceptual quality as perceived by the human visual system. Therefore, a narrow focus on these metrics may neglect the importance of achieving results that are visually realistic and aesthetically pleasing to human observers. There has also been a growing emphasis on perceptual quality metrics, which evaluate the visual quality of deblurred images based on human perception. Integrating these metrics into the evaluation process can help ensure that deblurred images are not only quantitatively accurate but also visually pleasing to humans. As research on real-world image

deblurring continues, the above challenges are expected to be gradually addressed, thus leading to effective and practical deblurring solutions. This survey aims to provide a comprehensive understanding of the current landscape of research in real-world deblurring and offers a roadmap for further advancements in this critical area of computer vision.

Key words: image deblurring; real-world scenes; non-uniform blur; coupled blur; unknown degradation representation

0 引言

当今社会,图像成为人们日常生活的重要组成部分,无论是医学影像、监控摄像、卫星图像还是科学研究,都涉及到获取、传输和处理大量的视觉信息。然而,这些数据往往受到各种因素的影响,例如摄像机与目标之间存在的晃动,导致图像出现模糊现象,对于数据可用性构成了严重挑战,极大影响了对信息的获取和探知,阻碍了下游任务的发展。因此,去模糊方法的研究和应用至关重要,对各个领域的应用都具有重要意义。随着科技的不断发展,图像去模糊方法已经取得了令人瞩目的进展。

自20世纪七、八十年代以来,越来越多的图像处理技术用于图像去模糊任务。最经典的方法是逆滤波,直接通过逆向解方程的方式估计原始图像,适用于不存在噪声或噪声较小的情况。然而,该方法过于理想化,实际情况下由传感器和成像环境产生的噪声是不可避免会存在的,为克服逆滤波对噪声比较敏感的缺点,Wiener(1964)在假定图像和噪声是随机的基础上,提出维纳滤波方法,试图找到图像的估计,使其与清晰图像之间的均方误差最小,提升了抗噪能力,使图像去模糊研究向着真实场景迈进了一步。Richardson-Lucy滤波(Richardson, 1972)方法在消除噪声对复原过程的影响上展开了相关研究,能够处理不同程度的噪声,但是该算法需要人为设置迭代次数,无法在实际中灵活使用。

除了噪声的影响,真实场景中成像环境还会受到如目标运动形式不一致、相机镜头散焦、传输过程像素压缩以及光照不足等因素的干扰,呈现出更加复杂的模糊现象,如非均匀模糊、噪声与模糊耦合等现象。此时,以往的图像去模糊方法,例如基于模糊核估计的去模糊方法(Levin, 2006; Shan等, 2008; Krishnan和Fergus, 2009; Levin等, 2011; Pan等, 2014, 2016b)都不足以应对真实场景下的复杂去模糊问题。究其原因,主要是其受到很多限制,如假设图像的噪声是高斯分布的、模糊是空间均匀分布以

及模糊是独立存在的等,导致其无法适配于更真实场景的模糊去除问题,因此,需要专门针对真实场景去模糊任务进行研究,提升图像复原效果。

针对上述问题,学者开展了深入的研究,先后推出数篇关于图像去模糊方法的综述。Wang和Tao(2014)全面回顾了2014年前的图像去模糊方法,从如何处理不适定性的角度将大约190篇文献提出的方法分为:贝叶斯推理框架、变分方法、基于稀疏表示的方法、基于单应性的建模和区域模型,并对代表性方法进行了分析。Lai等人(2016)首次使用真实世界的模糊图像对单图像盲去模糊进行了全面的研究和分析,总结了近49篇文献,表明仿真模糊图像和真实模糊图像之间存在的性能差距,构建了新的真实场景去模糊数据集,评估了从2006年到2014年间提出的13种方法在该数据集上的复原结果。Koh等人(2021)整理了113篇相关文献,回顾了2019年之前的基于深度学习的去模糊方法,并定量和定性比较了它们的性能,对基于深度学习的单图像去模糊技术进行了总结,特别是对其中的7种非盲和7种盲去模糊方法以及4种优化技术进行了整理分类,并在仿真数据集上进行性能评估,表明了深度学习的应用使去模糊技术远远超出了以前优化技术的限制。Zhang等人(2022)对2015年至2022年发表的169篇基于深度学习的图像去模糊方法进行了全面调查,讨论了图像模糊的常见原因,分析了常用网络模型和损失函数,介绍了基准数据集和各方法在其上的性能,并总结了一些特定领域的去模糊应用。

上述综述大多集中在对仿真场景的研究上,尽管Lai等人(2016)和Zhang等人(2022)总结了一些真实图像的去模糊方法,但是缺少系统的讨论,特别是Lai等人(2016)参考的文献都较为久远,Zhang等人(2022)缺少对真实场景的分析。为更新真实场景下去模糊方法的研究进展,充实真实场景去模糊综述研究,使读者对真实场景的研究有更全面深入的了解,本文对真实场景去模糊任务进行了调研和系统性的讨论和总结。针对广泛存在于成像过程中的各类降质现象,以恢复出清晰图像为根本目的,本文

以模糊图像的物理退化模型作为切入点,按照由浅入深的概念、由易到难的任务程度,将真实场景下的去模糊问题进行拆解,归纳为单一模糊去除方法、复合模糊去除方法以及真实场景未知降质下的模糊去除方法,为研究人员提供全面的真实场景下去模糊方法的信息,鼓励不同学科领域间的合作,更好地迎接真实世界图像模糊的挑战,引领去模糊方法未来的发展方向。表1展示了已有综述与本

综述对比。

本文主要贡献如下:1)分类总结了在不同条件下真实场景去模糊方法的发展历程,介绍各类方法的核心思想,对比方法的优缺点,分析存在的问题及可能的解决方案;2)整理总结了真实场景去模糊方法常用的数据集和评价指标;3)系统梳理真实场景去模糊任务仍需解决的关键性问题,对未来研究方向及发展趋势进行展望。

表1 已有综述与本综述对比
Table 1 Comparisons with existing surveys and this paper

综述文献	数据类型	研究背景和内容	引用文献年份	引用文献数
Wang和Tao(2014)	仿真图像	主要针对仿真图像去模糊的传统优化方法进行了总结	2014年之前	190余篇
Koh等人(2021)	仿真图像	主要针对仿真图像去模糊的深度学习方法进行了总结	2019年之前	113余篇
Lai等人(2016)	真实图像	首次对真实世界的模糊去除方法进行综述,并提出真实去模糊数据集,但只考虑了13种方法	2006年至2014年	49篇
Zhang等人(2022)	仿真图像和真实图像	总结了近期基于深度学习的去模糊网络结构、损失函数、训练数据集、评价指标	2015年至2022年	169篇
本综述	真实图像	主要针对真实场景去模糊方法进行总结,按照场景真实的程度进行分类,总结了真实场景去模糊常用的数据集和评价指标	2023年之前	126篇

1 图像模糊退化模型

去模糊的目的是恢复出包含丰富结构和细节的清晰图像。模糊图像的退化过程可以被建模为“一个退化函数和一个加性噪声项的和”,从而形成图像退化模型。从数学角度抽象来讲,一般认为图像模糊的过程是由表征模糊信息的模糊因子,即模糊核 k 和原始图像 x 在空间域上所作的卷积运算(对应频域上的点乘),模糊图像退化过程的具体模型可以表示为

$$y(a,b)=k(a,b)\otimes x(a,b)+n(a,b) \tag{1}$$

式中, a,b 表示空间坐标, k 表示退化系统 H 的模糊核,也称点扩散函数(point spread function, PSF), $\otimes$ 表示二维卷积运算。

图像复原是图像退化的逆过程,传统去模糊方法大多是基于贝叶斯推论进行的,通过统计学理论,对图像数据进行建模,人为设计不同的先验,用以约束问题的解空间。基于贝叶斯的图像去模糊问题可以表示成如下形式,即在给定模糊图像 y 的情况下,估计清晰图像 x 和模糊核 k 的概率,具体为

$$p(x,k|y)=\frac{p(y|x,k)p(x)p(k)}{p(y)} \tag{2}$$

式中, $p(x,k|y)$ 和 $p(y|x,k)$ 分别表示潜在图像的后验概率和似然概率, $p(x)$ 、 $p(k)$ 和 $p(y)$ 分别表示清晰图像、模糊核以及模糊图像的先验。最大后验概率是常用的贝叶斯推断方法之一,以最大化清晰图像的后验概率 $p(x,k|y)$ 为目的,当给定模糊图像 y 时,先验概率 $p(y)$ 可被视为固定值,因此,模糊核和清晰图像的估计可表示为

$$(k^*,x^*)=\arg\max_{x,k}p(x,k|y)=\arg\max_{x,k}p(y|x,k)p(x)p(k) \tag{3}$$

以往传统方法(Krishnan和Fergus,2009;Levin等,2011;Pan等,2014)大多假设噪声是服从高斯分布的,即 $n\sim\mathcal{N}(0,\sigma^2)$,则 $p(y|x,k)\propto\exp(-1/2\sigma^2\|y-x\otimes k\|^2)$ 。因此,在高斯噪声假设下,清晰图的估计模型为

$$(k^*,x^*)=\arg\min_{x,k}\|y-x\otimes k\|^2+\lambda\Omega(x)+\eta\Psi(k) \tag{4}$$

式中, $\|y-x\otimes k\|^2$ 是与噪声模型有关的数据拟合项, $\Omega(x)$ 和 $\Psi(k)$ 是与先验 $p(x)$ 和 $p(k)$ 对应的正则

项, λ 和 η 为正则项系数。此前研究方法的主要工作在于对先验 $\Omega(\mathbf{x})$ 和 $\Psi(\mathbf{k})$ 的设计, 取得了很多良好的进展, 极大改进了当时的复原效果。

真实场景下, 由于缺少模糊核的相关信息, 同一场景内的模糊形式常常是全局不一致的, 且还会受到其他诸如噪声、光照等外界因素的干扰, 致使图像去模糊问题更加复杂, 未知变量(潜在图像、模糊核、

噪声等)个数远多于已知量(模糊图像), 是一个严重病态的问题。本文从图像退化模型的角度出发, 在真实场景下, 当给定模糊图像 $\mathbf{y}$, 按照退化模型中未知变量的个数以及复原任务的难易程度, 将求解清晰图像 $\mathbf{x}$ 的方法归纳为3大类, 如图1所示, 包括单一模糊去除方法、复合模糊去除方法和未知模糊去除方法。

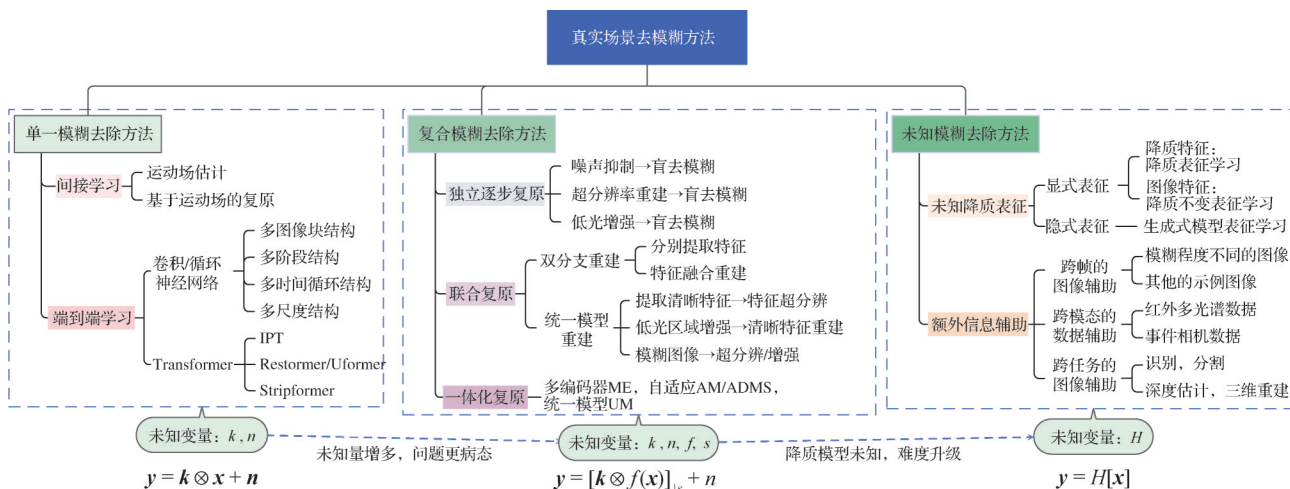


图1 真实场景去模糊方法分类

Fig. 1 The classification of deblurring methods in real-world scenarios

2 单一模糊去除方法

对于单一模糊存在的场景, 首先假设只存在一种形式的运动, 此时模糊核是空间不变的, 从而会产生全局均匀模糊(uniform blur), 可以用一个模糊核进行建模, 如图2(b)所示, 很多学者已经对该问题展开了大量的研究(Fergus等, 2006; Levin等, 2007, 2009; Zhang和Zhang, 2009; Pan等, 2014; Qiu等, 2019), 并取得了令人瞩目的效果。

然而, 在真实世界中, 场景中的物体存在不同的运动, 导致同一幅图像的模糊在空间上存在全局不一致的状态, 形成非均匀模糊(non-uniform blur)。如图2(c)所示, 由相机以左下角为中心旋转产生模糊, 模糊随着空间位置变化而变化, 左下角模糊程度最小, 右上角模糊较大, 模糊的运动方向不同。此时, 全局一致去模糊方法难以处理这种复杂的情况, 退化公式中的 $\mathbf{k}$ 不再如图2(b)所示的均匀模糊核一般, 而是呈现为图2(d)形式的运动场。因此, 需要开发更适用于真实场景的图像去模糊方法。

最初关于非均匀模糊的研究来源于Levin(2006), 假设模糊是由匀速运动产生的, 将边缘梯度建模为模糊核宽度的函数, 利用该函数区分不同模糊的区域。Gupta等人(2010)提出运动密度函数, 记录每一个相机姿态在空间中所占的时间比例, 以描述空间变化的模糊核, 用很多小模糊核 $\mathbf{k}_j$ 组建全图模糊核 $\mathbf{k} = \sum_j \alpha_j \mathbf{k}_j$, 则非均匀一致模糊模型可表示为

$$\mathbf{y} = \sum_j \alpha_j (\mathbf{k}_j \otimes \mathbf{x}) + \mathbf{n} \quad (5)$$

式中, α_j 表示相机姿态在 j 处的密度, $\mathbf{n}$ 表示噪声, 在此基础上, 可以采用以往全局一致的算法进行复原。

自2010年以来, 随着深度学习技术的兴起, 对于真实场景下的图像复原问题, 基于深度学习的图像去模糊方法凭借其出色的性能受到了学者的持续关注, 并得到了长足的发展。针对非均匀模糊去除问题, 图1左边框图对当前基于深度学习的单一模糊去除方法进行了总结, 分为基于运动场间接学习的方法和基于端到端学习的方法。

2.1 基于运动场间接学习的方法

起初的方法先估计模糊核, 再进行反卷积的间

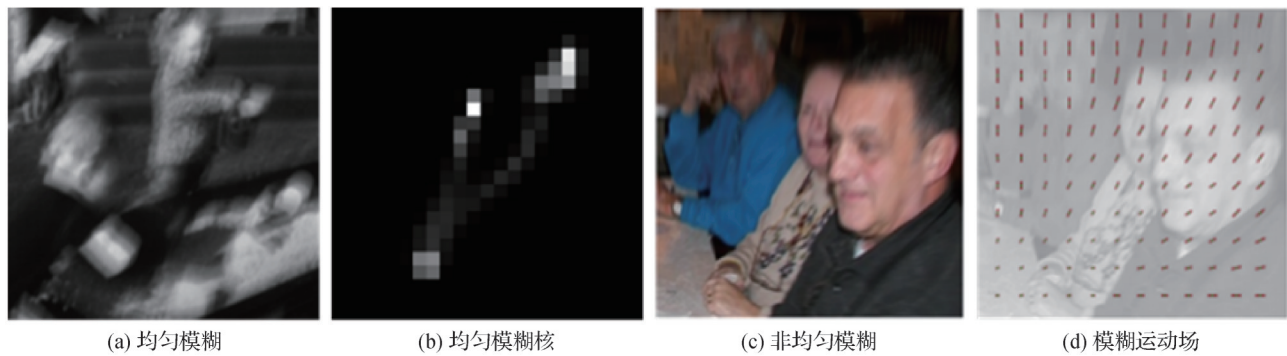


图2 均匀模糊与其模糊核(Levin等,2009)和非均匀模糊图像及其模糊运动场(Sun等,2015)
Fig. 2 Uniform blur with kernels (Levin et al. , 2009) and non-uniform blur with motion flow (Sun et al. , 2015)
((a) uniform blur; (b) kernel of uniform blur; (c) non-uniform blur; (d) blur motion flow)

接方式求解清晰图像,Sun 等人(2015)使用卷积神经网络(convolutional neural network,CNN)来估计局部图像块的模糊核,构建了第一个用于网络模型训练的模糊核数据库。Gong 等人(2017)提出用全卷积网络(fully connected network,FCN)估计出用于复原的模糊运动场,并根据该模糊场进行清晰图像估计。Kaufman 和 Fattal(2020)构建了分析—合成网络,前者用于模糊核估计,后者将模糊核合并到 UNet 网络进行复原。Fang 等人(2023)提出基于归一化流的不确定性模糊核估计网络,在估计出模糊核后,利用反卷积操作或者深度网络恢复出最终的清晰图像。这些方法在当时都取得了很好的效果,但是由于需要独立估计模糊核,对于模型训练和模糊核数据集的构建来说,都较为繁杂,不方便实际任务使用,而且,如果模糊核估计不够精确,则会将误差传输给后续的反卷积步骤,影响最终的复原效果。

2.2 基于端到端学习的方法

除了利用网络学习模糊核外,另一些方法尝试训练端到端的网络,依靠大规模的数据集,直接从模糊图像中恢复潜在的清晰图像,省略了中间估计模糊核的步骤,从而不受模糊核参数和形式的限制,加速了图像盲复原算法的发展。

2.2.1 卷积/循环神经网络

在该领域,Nah 等人(2017)提出的 DeepDeblur 模型被视为开山之作,如图 3(a)所示。他们发现模糊图像在不同分辨率上具有不同的特征,将图像尺寸整体缩小之后,原本严重的模糊现象会减弱。基于该性质,提出了基于多尺度的卷积网络(multi-scale),每一个尺度由 19 个残差块(residual blocks)组成,以由粗到细的方式逐渐恢复出不同分辨率下

的清晰图像。由于较小尺度的模型具有较大的感受野,可以更好地覆盖不同模糊区域,从而提升了复原效果,实验证明了该方法明显优于以往传统方法的结果。陈加保等人(2023)提出了双尺度网络,通过高低尺度通路同时增强模型的去模糊效果,使得去模糊图像的细节信息恢复得更加显著。

受该结构的启发,Tao 等人(2018)提出 SRN(scale-recurrent network)模型,以跨尺度的权值共享方式降低网络的训练难度,将 Nah 等人(2017)方法中同一尺度下的残差块改为权值共享的 UNet 模型,如图 3(b)所示,并在编码器和解码器中间加入循环神经网络(recurrent neural network,RNN)模块实现跨尺度递归,学习长期依赖,该方法以更少的参数和更易训练的特性显著提升了 DeepDeblur 的去模糊性能。Gao 等人(2019)提出参数选择性共享和嵌套跳层连接模块改进了 SRN,使得网络可以更有针对性地学习不同模糊模型的参数,从而训练更快,效果更好。

由于不同空间尺度的模型具有很强的依赖性,Park 等人(2020)提出多时间循环神经网络(multi-temporal recurrent neural network,MT-RNN),如图 3(c)所示,单个 UNet 网络按时间顺序被重复多次,并将上一次迭代中来自解码器的特征图传输到下一次迭代的编码器中,将严重的非均匀模糊建模为一系列小模糊的叠加,从而逐步消除模糊现象。与多尺度结构不同,多时间结构中的每个去模糊子问题仍然处于原始空间分辨率,因此,可以使用高频信息进行可靠的去模糊处理。

DMPHN(deep stacked hierarchical multi-patch network)(Zhang 等,2019)分析了多尺度结构存在的弊端,如上下采样会增加模型训练难度和时间,图像

在降采样过程中会存在信息损失等,提出将图像分成不同的图像块(multi-patch)进行处理,如图3(d)所示。分块处理可以将网络的感受野限制到特定的局部区域内,学习不同局部特有的模糊运动信息,相比于multi-scale模型,multi-patch能更准确地描述全局不一致的运动信息,更适合非均匀去模糊任务。Zamir等人(2021)在DMPHN的基础上,提出基于多阶段(multi-stage)的网络模型MPRNet(multi-stage progressive restoration network),第1、2阶段分别采用4分块和2分块的编码器—解码器模型来学习多尺度上下文信息,第3阶段对原始分辨率的图像进行操作,以保留精细的空间细节。基于MPRNet结构

(Zamir等,2021),HINet(Chen等,2021c)修改了模型的归一化层,提出半实例归一化操作,进一步提升了包括MPRNet在内的多个去模糊模型的效果。

MIMO-UNet(multi-input multi-output network)(Cho等,2021)在总结以往方法的基础上,提出了一个高效的去模糊算法,如图3(e)所示。不同于多个模型级联或者多阶段处理的方法,该方法充分利用从输入图像中提取的多尺度特征,仅用一个具有多输入和多输出的框架,模拟多级联网络,在编解码器中构建多尺度信息传输,大大简化了模型复杂度,降低了计算成本,实现了更简单但更有效的去模糊效果。

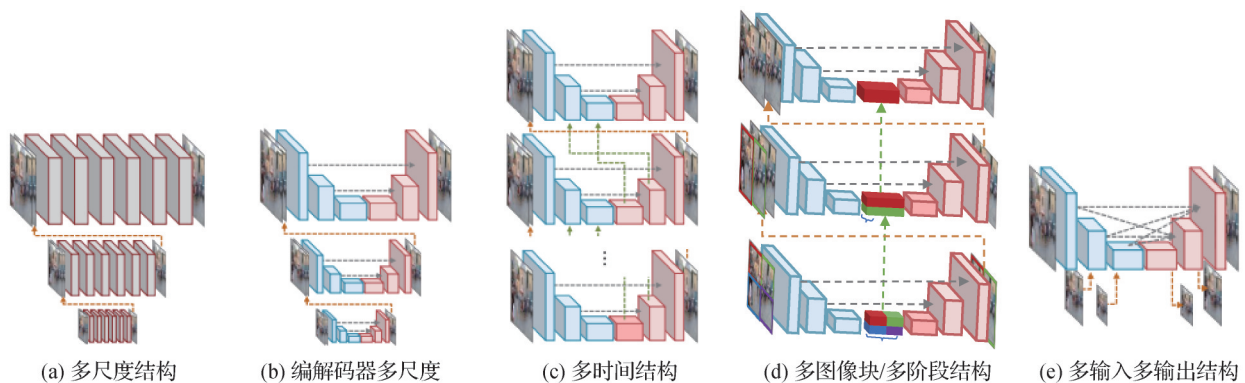


图3 基于端到端深度学习的去模糊方法模型(Cho等,2021)

Fig. 3 The overview of end-to-end deep learning-based image deblurring methods (Cho et al. , 2021)

((a) multiple scales; (b) multiple scales with encoder and decoder; (c) multiple temporal; (d) multiple patches/multiple stages; (e) multiple inputs and multiple outputs)

2.2.2 Transformer网络

以上方法大多基于CNN或RNN设计的,然而CNN只能处理固定长度的特征向量,使特征在解码时无法获得足够的全局信息,限制了其在非均匀模糊去除任务上的性能提升。虽然RNN能处理序列数据,但训练和推理速度较慢,在长序列上易面临梯度消失或爆炸的问题。

为解决上述问题,注意力模型(attention mechanism)被提出,起初用于自然语言处理,后来凭借其在处理长序列时仍能够有效捕捉到不同位置的相关信息特性,被广泛用于图像复原任务。相比于CNN和RNN,注意力模型能够动态地关注输入图像的不同部分,根据当前需要的信息自适应性地调整注意力,适用于需要全局信息、多尺度处理或自适应特征分配的任务。2017年,Google发表的《Attention is All You Need》抛弃了RNN和CNN等网络结构,仅

采用Attention机制实现机器翻译任务,取得了很好的效果,自此,注意力机制成为研究热点,特别是基于自注意力(self-attention)的Transformer模型更是得到广泛应用。自注意力机制的核心思想是将输入序列中的每个位置都看做查询(query)、键(keys)和值(values),然后计算每个查询与所有键之间的关联度,并根据关联度分配权重来加权平均对应位置的值,从而生成输出,实现对不同位置的关系建模,而不受序列长度的限制。

IPT(image processing Transformer)(Chen等,2021b)是第一个用于复原任务的Transformer模型,由包含卷积层的多个头(heads)和尾(tails)构成,不同的头和尾可以处理不同的任务,但共享一个Transformer主体,从而可以提取更泛化的特征。然而,IPT依赖于大量参数、大规模数据集和多任务学习来获得良好的性能,且其计算复杂度随空间分辨

率呈二次方增长,无法应用于大多数涉及高分辨率图像的图像复原任务,因此很难在实际中使用。为了减小计算量并能够处理高分辨率数据,Restormer (Zamir 等, 2022)和 Uformer (Wang 等, 2022)几乎在同一时期相继提出。Restormer引入基于深度卷积 (depth-wise convolution, Dconv)的多头转置注意力模块,能够聚合局部和非局部像素,可以有效处理高分辨率图像,提出一种门控 Dconv 前馈网络允许有用的信息通过网络层次结构,同时抑制信息较少的特征。Uformer采用非重叠的基于窗口的自注意力模型降低计算量,同时在前向网络中采用深度卷积改善捕获局部上下文的能力。Stripformer (Tsai 等, 2022a)表明 IPT 直接将整幅图像展成 token,破坏了图像空间像素之间的相关性,提出通过计算带内和带间注意力,在水平和垂直方向上重新调整图像特征的权重,以捕捉不同方向的模糊特征,并且通过不断堆叠两种注意力模块来表征模糊的程度。与其他方法相比,Stripformer 的内存占用、参数量和计算成

本都较小,且无需大量的训练数据做预训练,达到目前最高的复原水平。

2.3 小结

综上所述,对于单一情况下非均匀模糊问题,基于运动场间接学习的方法和基于端到端学习的方法都能在一定程度上去除全局不一致模糊,但是性能上各有利弊,总结如表 2 所示。综合来看,当前非均匀模糊去除方法发展的难点在于,较复杂的网络模型、缓慢的计算速度以及有限的计算资源限制了模型的有效训练,从而影响了算法性能的提高及算法在实际生活中的应用。将来的研究可以在保持复原效果的基础上,向轻量级和实时性算法方向发展。例如, Hu 等人 (2021)在标准去模糊数据集上的实验证明了对 720 p 图像的实时去模糊算法效率达到约 58 帧/s。Zhao 等人 (2022)提出的方法模型参数比最先进的方法 (state-of-the-art, SOTA) (Lu 等, 2019)小 25 倍,计算速度快 5 倍,且不受图像域和分辨率的限制,同时保持了先进的去模糊性能。

表 2 单一情况下非均匀模糊去除方法优缺点总结
Table 2 Summary of advantages and disadvantages of independent non-uniform blur removal methods

方法(代表性文献)	优势	劣势
基于分块建模的传统方法 (Gupta 等, 2010)	依靠目标运动轨迹建模,复原模型精准	依赖过多的假设,先验信息需人工设计
基于运动场间接学习的方法 (Sun 等, 2015)	强大的表示和学习能力提升了模糊核估计效果	模糊核数据集构建困难,不精确的模糊核估计结果影响后续复原效果
卷积神经网络 (Nah 等, 2017)	直接学习清晰图像,局部特征捕获能力强	需要构造大规模数据集,只能处理固定长度的特征向量,无法构建全局依赖关系
基于端到端学习的方法 循环神经网络 (Tao 等, 2018)	模型可学习长期依赖,参数少,易训练	模型训练和推理速度较慢,在长序列上容易面临梯度消失或者爆炸问题
Transformer (Tsai 等, 2022a)	局部和全局特征相融合,根据输入的不同部分动态地分配关注度,不受序列长度的限制	注意力机制导致计算量较大,计算速度慢,需要大量的存储和内存空间

3 复合模糊去除方法

虽然前述方法在单一模糊情况下取得了显著的进展,但当放宽假设,考虑更真实世界的模糊问题时,由于未知量个数增加,问题变得更加病态,这些方法无法继续保持优异的效果。受成像环境干扰,真实场景经常包含其他类型的降质,例如噪声、低光、低分辨率压缩等,且多种降质往往是耦合在一起

的,形成多因素复合模糊,这种模糊情况更贴近于现实世界的真实模糊,严重影响图像质量。

复合模糊去除是指利用各种方式对复杂的模糊现象进行去除。相比于单一模糊,复合模糊更难去除,因为退化过程不再服从初始的退化模型,退化系统 H 不再是单一的模糊核卷积操作,可能还包含降采样、动态范围压缩等操作,需要构建新的退化模型进行求解,一种常用的建模方式可表示为

$$y = [k \otimes f(x)]_l + n \tag{6}$$

式中, f 表示相机响应函数, $\downarrow$ 表示下采样过程, n 为噪声。对于盲复原情况, 只有降质图像 y 已知, 需要求解出 x , 相比于单一模糊模型(1)而言, 该问题的求解更加困难。对于复合模糊的情况, 根据算法步骤的设计, 处理方法可大致分为: 独立逐步复原、联合复原以及一体化复原, 如图1中间的框图所示。

3.1 独立逐步复原方法

独立逐步复原针对不同降质, 按一定顺序将不同的降质分别去除, 例如对于模糊和噪声耦合存在情况, 先对噪声分布进行建模, 如脉冲噪声(Cai等, 2010)、椒盐噪声(Schmidt等, 2011)等, 然后基于噪声模型, 对图像中的噪声信号进行抑制, 去除不利于模糊核估计的外点, 能够实现在去噪的同时, 保留良好的结构信息。在去噪基础上, 参考第2.1节提到的方法进行去模糊处理。需注意的是, 对于不同的噪声分布, 复原模型也会有些许区别, 需单独建模, 例如泊松噪声下的图像去模糊任务, 可采用如下形式的复原模型(Carlavan和Blanc-Feraud, 2012), 具体为

$$\mathbf{x}^* = \arg \min_{\mathbf{x}} \mathbf{1}^T (\mathbf{k} \otimes \mathbf{x}) - \mathbf{y}^T \log(\mathbf{k} \otimes \mathbf{x}) + R(\mathbf{x}) \quad (7)$$

式中, $\mathbf{1}$ 表示所有元素均是1的向量, $R(\mathbf{x})$ 表示清晰图像的正则项。式(7)常用的求解策略是Richardson-Lucy算法(Dey等, 2006), 由于其可以对泊松统计量进行计算, 因而在泊松噪声图像去模糊中得到了广泛应用, 但是其性能受到迭代次数的限制, 不足以防止反卷积过程中的噪声放大现象, 可能导致图像出现伪影。之后的研究大多数依靠正则项对问题进行约束, 改善了复原效果(Dupe等, 2009)。

对于低分辨率与模糊及低光与模糊耦合情况下的复原, 因为在超分辨的过程中会不可避免地导致图像变模糊, 而低光增强会使原先不明显的模糊区域变得显著, 大多数方法遵循先超分辨或低光增强后, 再进行图像的去模糊。例如, Hu等人(2014)利用在低光情况下拍摄到的光线轨迹作为模糊核估计的参考值, 实现在低光场景下的去模糊任务。Zhou等人(2022)利用最先进的低光增强方法RUAS(retinex-inspired unrolling with cooperative prior architecture search)(Liu等, 2021)和最先进的去模糊方法MIMO-UNet(Cho等, 2021)以串行的方法进行图像复原, 复原的图像在模糊去除方面还存在较大提升空间。

独立逐步复原方法的优势在于可以更有针对性

地对不同降质单独设计算法, 但是前后步骤之间容易存在误差问题, 并会随着处理步骤逐渐累积传递, 前一步不精确的估计结果会对后面的模型产生较大影响, 并且两个任务之间的信息难以互相共享, 不能相互利用, 导致信息缺失, 使得最终的复原结果往往陷入次优解。然而, 由于操作简单, 该类方法经常用于各种工程实践中, 更具实用价值。

3.2 联合复原方法

联合复原用一个单独的复原框架直接去除多种降质, 优势在于避免了独立逐步复原方法的误差累计, 但是由于需要设计的算法同时具有处理多种降质的能力, 面临更多的挑战。

以模糊和超分辨耦合的情况为例, 近年来有学者提出用联合优化的方式对低分辨率模糊图像进行处理, 对一幅经过下采样的运动模糊图像同时进行去模糊和超分辨率处理。早在1996年, Bascle等人(1996)利用物体的运动轨迹来估计模糊核, 并建立了多帧超分辨率的子像素(sub-pixel)之间的对应关系, 然而, 这种方法仅适用于易于跟踪的目标, 不适用于整幅图像。基于深度学习的处理方法按照算法框架可分为双分支重建模型和单支路重建模型。

3.2.1 双分支重建模型

双分支重建模型常常是构建两个单独的复原支路, 对输入图像分别进行去模糊和超分辨, 最后再将两个分支的输出结果进行融合。Zhang等人(2018a)提出两个任务共享一个编码器, 但解码器是单独构建的网络模型, 在融合部分只是进行了简单的相加操作。同年, Zhang等人(2018b)提出以并行的双分支网络模型解决图像去模糊和超分辨问题, 并开发了一种门控融合机制来聚合去模糊特征和超分辨特征, 最终重建高分辨的清晰图像。Yun等人(2020)提出一种基于生成对抗网络(generative adversarial network, GAN)框架的模型, 采用5个卷积层从低分辨(low resolution, LR)人脸图像中提取特征图, 并将该特征提供给两分支的编码器—解码器网络, 生成包含模糊和不包含模糊的超分辨图像, 并提出局部和全局相结合的判别器, 重建图像的细节结构。

Liang等人(2019)使用去模糊模块进行模糊去除, 并将去模糊特征提供给超分辨重建模块, 旨在从去模糊特征中产生最终的超分辨率结果。同时, 如图4所示, 该方法基于双分支监督学习策略, 对去模糊的低分辨率图像进行上采样用以监督超分辨结

果,而对超分辨率后的高分辨率图像进行下采样用以监督去模糊结果,从而实现循环损失监督,约束不同任务生成的特征具有内容一致性。Zhang 等人(2020a)改进了Liang 等人(2019)的结构,用额外的双模块充分挖掘了LR 和超分辨率(super-resolution, SR)图像之间的关联性,提高了网络的学习性能。

双分支重建方法的主要优势在于通过并行的两个模块,网络可以不受干扰地单独提取有利于不同任务的清晰特征,但同时由于缺乏任务之间的信息传递,一些有利于对向任务的信息被损失了。因而,该任务的难点在于如何在保持各自任务性能的前提下,进行有效的信息利用和特征融合。

3.2.2 单支路重建模型

单支路重建方法直接构建一个独立的网络模型,以端到端的方式进行训练,重点在于设计具有多种降质同时复原潜能的特殊网络模块。Albluwi 等人(2018)提出将高分辨的清晰图像作为监督信号,直接训练单独的模型进行低分辨率去模糊任务。Xu 等人(2017)针对模糊的人脸和文本图像进行超分辨,基于GAN 模型学习特定类别先验,并引入像素级训练损失函数,恢复精细的图像细节。吴迪等

人(2020)在密集连接卷积网络中引入生成式对抗框架,生成与感官更加接近的清晰图像。Bai 等人(2021)提出一种级联的非局部残差网络,先使用去模糊模块生成中间清晰特征,然后利用非局部残差网络作为超分辨率模块,从中间特征生成高分辨率图像,并且开发了一种图像梯度约束来保持边缘清晰。针对具有模糊的低分辨率空间卫星图像,Yang 等人(2022)提出基于GAN 的框架提高图像质量,使用对比学习提取退化特征实现去模糊,并向超分辨网络添加对称下采样和上采样模块以恢复纹理信息。对于低光和模糊耦合的情况,Zhou 等人(2022)提出独立的UNet模型,其编码器部分对暗区域进行增强,解码器重建清晰特征,编解码器通过自适应动态滤波器进行跳层连接,并在编码器结尾处使用真值图像作为监督计算重建损失,实现了两个任务的联合优化。

单支路的方法省略了双分支模型需要单独构建的融合模块,简化了网络设计步骤,但也存在一些问题,如两个任务的优化方向可能存在不同,从而导致训练难度增加。因此,该类方法的难点在于如何设计更有效的重建模型以及合适的优化策略。

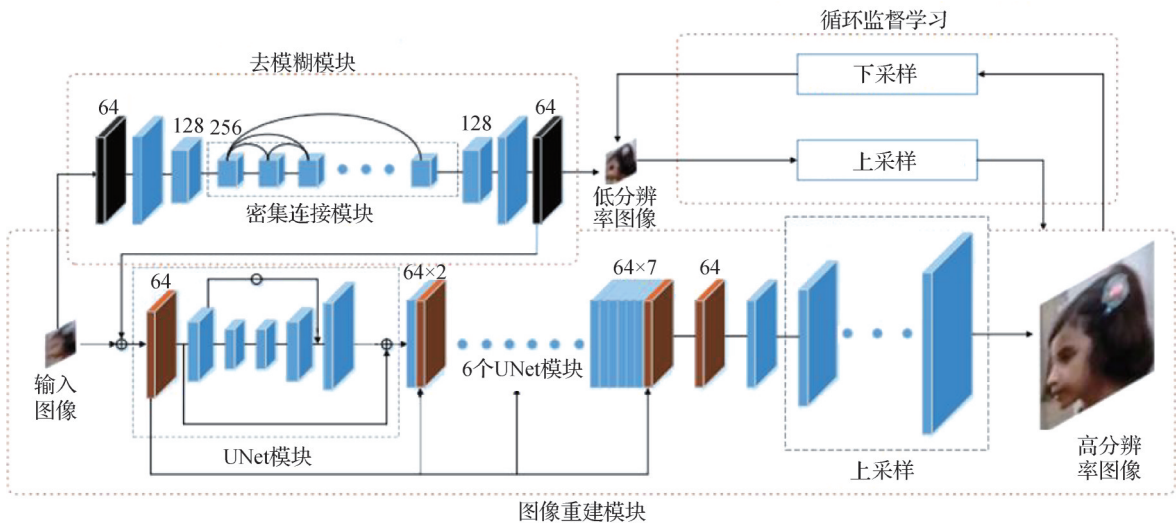


图4 去模糊和超分辨率的双分支模型(Liang等,2019)

Fig. 4 The details of dual-branch framework for image deblurring and super-resolution(Liang et al. ,2019)

3.3 一体化复原方法

对于只考虑两种因素耦合的问题,联合复原方法在复原效果和处理效率上都有所提升。但是,当面临更多降质因素耦合的情况时,针对不同降质独立设计的方法则不太适用,此时,需要设计能够处理

各类降质的统一的复原方法。Li 等人(2022a)提出一种可以从各种降质类型和降质程度中恢复图像的一体化方法,称为一体化处理模式,主要优点有:一体化解决方案可以用同一个网络恢复各种退化的图像;其不受降质类型和降质程度的影响,仅使用捕获

到的降质图像进行处理。由于在现实世界的退化过程很难预先知道,且退化会随空间和时间的变化而变化,因此该模式能在真实世界的复原任务中具有更好的灵活性和更高的经济性。

图5展示了一些一体化复原方法的模型。Li等

人(2020b)使用具有多个独立编码器,但共享解码器的结构(model with independent encoders, ME)进行多降质复原,与每个降质独立复原的结构(independent model, IM)相比,该方法具有较小的网络参数。

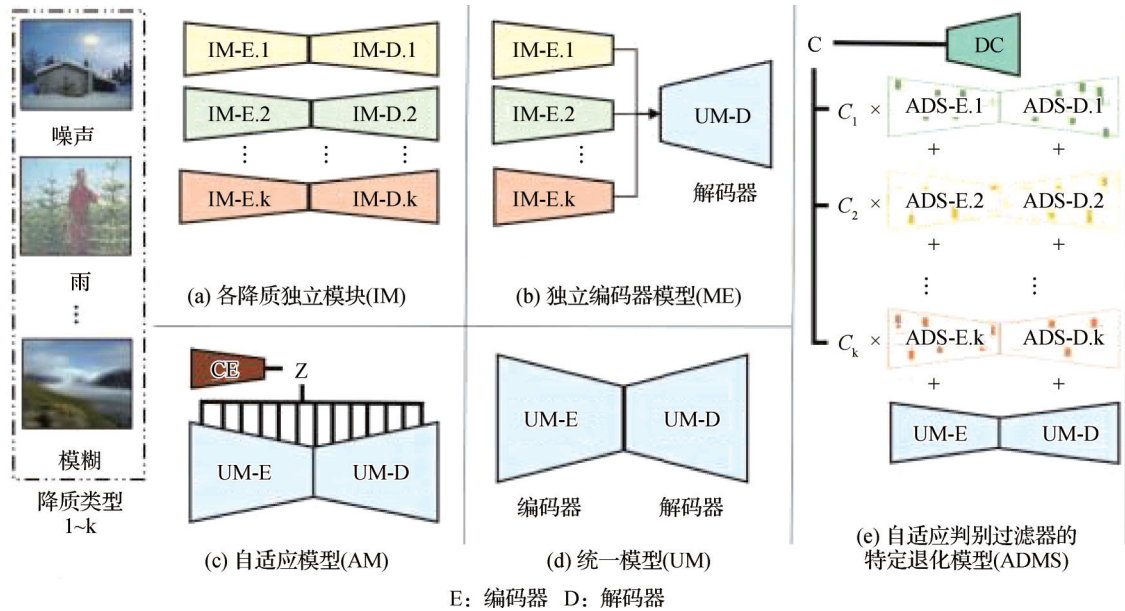


图5 一体化复原方法模型(Park等,2023)

Fig. 5 All-in-one-based restoration methods(Park et al.,2023) ((a) independent model; (b) model with independent encoders; (c) adaptive model; (d) unified model; (e) adaptive discriminative filter-based model for specific degradation)

然而,不同降质在共享的解码器处理之间可能存在相互干扰,为此,Chen等人(2022)使用一对共享的编码解码器的统一模型(unified model, UM)进行复原,通过多种降质之间的知识蒸馏来完成模型训练,但是考虑到处理多个退化任务时的不同恢复特性,这种方式可能并不理想,例如去噪主要依靠低通滤波,而去模糊则更类似于高通滤波处理,两种处理方式相互干扰。Li等人(2022a)提出基于对比学习的降质编码器(contrastive encoder, CE)和降质引导下的自适应复原网络(adaptive model, AM),经过隐式训练表示多种退化过程,以便在给定相应退化信息的情况下执行复原操作,能够自适应地提高各种降质,包括雨、噪声、模糊等因素的性能。Park等人(2023)充分结合了高性能IM和内存高效UM的优势提出了一种基于自适应判别过滤器的特定退化模型(adaptive discriminative filter-based models for specific degradation, ADMS)来恢复具有未知退化的图像,该方法允许网络在每次退化时,通过降质专用滤波器(adaptive discriminant filters for specific degra-

dations, ADS)和降质分类器(degradation classifier, DC),自适应地解耦网络的多种降质因素。

相比于前面两节的方法,一体化复原方法能够同时处理多种降质,实现多降质耦合下的复原任务,更加符合真实世界模糊场景的处理要求,例如对雨天模糊等现象,均可以达到良好的复原效果。

3.4 小结

综上所述,对于复合模糊去除的问题,本文将上述提到的方法进行总结,如表3所示。从表3可知,双分支重建方法需额外的特征融合操作,而其他类型方法则不受降质类型和降质程度的影响,可用同一个网络恢复各种退化,能处理更多类型的降质。该任务进一步提升的难点在于如何在一个模型里区分不同的降质,克服不同降质复原模型之间的干扰,实现不同降质的解耦,从而提升整体复原能力。

4 真实场景未知模糊的去除方法

虽然与单一模糊去除问题相比,复合模糊面临

表3 复合模糊去除方法优缺点总结

Table 3 Summary of advantages and disadvantages of coupled blur removal methods

方法(代表性文献)	优势	劣势	难点
独立逐步复原方法(Zhou等, 2022)	对不同降质单独设计复原方法,算法针对性强,鲁棒性高	不同算法步骤之间容易存在误差累积	如何挑选合适的复原工具,如何确定复原工具的顺序
联合复原方法	双分支重建模型(Liang等, 2019)	消除了步骤间的累计误差,不受干扰地单独提取有利于不同任务的清晰特征	不合适的融合模块会导致次优解,缺乏任务之间的信息传递,存在信息损失
	单支路重建模型(Bai等, 2021)	简化了网络模块,省略融合步骤,在特征空间处理数据,处理效率更高,性能更强	在保持各自任务性能的前提下,如何进行有效的信息利用和特征融合
一体化复原方法	共享解码器模型(Li等, 2020b)	只需一个解码器,模型参数较小,能处理更加多样化的降质类型	不同降质在共享的解码器处理之间可能存在相互干扰
	统一模型(Chen等, 2022)	模型更加简易,内存占用小,训练过程速度快,无需对不同降质单独设计网络	不同任务的优化方向可能存在不同,不易训练
	自适应复原网络(Li等, 2022a)	自适应学习不同降质下的特征,隐式地表征不同降质	如何在一个模型里区分不同降质,并分别处理,如何设计有效的重建模型以及合适的优化策略
	自适应判别过滤模型(Park等, 2023)	自适应分类并解耦多种降质,显式的降质表征,模型参数量少,性能高	基于统一框架对模型中的多个恢复任务会导致次优解

更复杂的情况,相应的去除方法放宽了很多假设限制,使得例如耦合模糊、多降质模糊等问题得到了一定程度上的改善。然而,这些方法只能应用于已知退化的情况,不能反映具有未知退化的现实环境。

在真实场景中,图像具体退化模型未知,数参也未知,从原始数据中无法获得较优的复原结果,复原问题更加困难。针对该问题,本文整理了目前热门的研究方向,如图1右边框图所示,将现有的真实场景下未知模糊的去除方法分为基于未知降质表征的方法和基于额外信息辅助的方法。

4.1 未知降质表征

对未知降质进行表征是通过学习的方式对真实降质做表征学习,从而实现真实场景下的图像复原。基于网络学习到的不同特征进行分类,按表征的方式将现有方法归纳为:降质表征学习、降质不变表征学习以及基于生成式模型的表征学习。

由于降质图像特征可以分为清晰图像特征和降质特征,降质表征学习旨在对图像中的降质特征直接进行学习,并进行可能的区分或解耦。当图像中的降质特征难以捕获时,可用降质不变表征学习方法直接学习图像中不受降质影响或与降质无关的图

像特征,过滤掉与降质因素相关的信息,以完成清晰场景复原任务,这两种方式针对的是复原过程中同一问题的不同方面。当学习显式特征存在困难时,可利用隐式方式从噪声中学习数据的特征分布,也就是生成式模型表征学习,在潜在的特征空间中编码数据的结构,捕获数据的重要特征和分布,以便生成与训练数据类似的新样本,该方法更加依赖于数据本身的特性。

4.1.1 降质表征

降质表征学习旨在估计降质类型或表示,然后通过自适应复原网络来处理不同的降质。Ren等人(2020)提出通过使用深度图像先验(deep image prior, DIP)显式地学习降质表征,以重参数化模糊核和清晰图像。Wang等人(2021)提出了一种无监督退化表示学习方案,引入了退化感知网络,学习隐式抽象的表示来区分空间中的各种退化,而非像素空间的显式估计,能够提取有判别力的表征以获得具体的退化信息。Li等人(2022b)构建了一个重模糊和去模糊模型的联合学习框架,引入一个编码器 E 将模糊图像 y 编码为降质表征 $E(y)$,该表征被建模为多通道的特征图,以对模糊图像中潜在的空间

变化降质表征进行编码。与以往通过模糊核建模的方式不同,该方法基于学习到的降质表征,直接用网络学习清晰图像到模糊图像的映射关系,对真实场景下的未知降质进行建模和去除。Li 等人(2020c)提出可以通过神经网络的不同层表达不同的降质特征,用特征的不同通道表示不同类型的降质,降质特征可表示为

$$D_i = \frac{S_i}{\sum_j w_{ij} S_j^2 + b^2} \quad (8)$$

式中, S_i 表示特征内的不同通道, D_i 为学习到的降质表征, w 和 b 为神经网络可学习的参数。之后,利用特征聚合模块来融合有利于复原的特征。

4.1.2 降质不变表征

基于降质表征的方法是通过提供具体标签学习多种降质因素的表征来指导复原模型的训练,在真实世界中很难执行,并且针对某些降质的特殊设计阻碍了模型被推广到处理其他降质现象。因此,降质不变表征学习方法得以提出,旨在从降质图像中学习并降质无关的特征信息,过滤掉与降质因素相关的信息,以完成清晰场景复原任务。

Du 等人(2020)通过解耦表示和对抗域自适应方法,从噪声数据中学习降质不变表征。Xia 等人(2023)提出了基于元学习的区域退化感知网络,使用元学习网络在多次迭代后快速适应特定的复杂退化,并提取隐式的图像降质信息。Li 等人(2023)首次从因果关系的角度提出了新颖的图像复原训练策略,以提高深度神经网络对未知退化的泛化能力。该方法将每种降质的类型和程度视为特定的混杂因素,并通过消除每种退化的有害混杂效应来学习降质不变表征,从优化角度对不同降质的干预进行建模,用因果关系中的后门标准导出所提出的模型,在提高未知降质的泛化能力的方面具有显著性能。

4.1.3 基于生成式模型的表征

由于生成式模型能够利用模型学习到的特定数据分布等信息,实现对数据的表征,从噪声数据中生成符合要求的图像,近年来也广泛应用于图像复原任务。基于生成式模型表征的复原方法无需进行降质表征,可直接生成真实的图像结构和细节,省略了降质因素建模的过程,常用的生成模型有生成对抗模型 GAN 和噪声扩散模型 Diffusion。

最具代表性的 GAN 复原方法是 DeblurGAN

(Kupyn 等, 2018),其采用生成器进行去模糊,引入判别器,用 Wasserstein GAN 和内容损失进行模型训练,在视觉效果方面取得了最好性能。DeblurGAN-v2(Kupyn 等, 2019)引入特征金字塔模型与轻量级模型框架,使 DeblurGANv2 达到实时处理的需求。之后有很多方法(Nah 等, 2017; Shen 等, 2018)都会在网络框架中包含额外的判别器,用以提升图像的视觉质量。Zhang 等人(2020b)提出结合学习生成模糊的 BGAN 模型(blur GAN)和学习去模糊的 DBGAN(deblur GAN)模型。其中, BGAN 学习如何使用不成对的清晰和模糊图像集对清晰图像施加模糊,然后指导 DBGAN 学习如何正确地对该图像进行去模糊,达到循环监督学习的目的,进一步提升了真实场景去模糊的效果。

由于 GAN 模型需要同时训练生成器和判别器,这种平衡难度导致了训练的不稳定性,作为一种更加高效的生成模型,Diffusion 模型最近受到研究者的关注(Whang 等, 2022; Fei 等, 2023),其生成的图像更加真实,细节更丰富,并且只需训练单个模型,简化了训练过程,优化过程更稳定。Diffusion 模型分为噪声扩散阶段以及逆采样阶段,前者是对一幅清晰图像 x_0 不断加噪声,直到变成纯高斯噪声的图像 x_T ; 后者是针对纯噪声图像 x_T 进行逐步去噪,最终恢复出清晰图像 x_0 的过程。通过这两个过程,可以训练出一个能够从噪声中生成高清图像的模型。基于 Diffusion 的图像去模糊方法可分为两种:有监督方法和零样本学习方法。有监督方法是直接将模糊图像和噪声图像拼接在一起作为新的噪声图像 x_T ,以真值图像作为监督信号,从头训练 Diffusion 模型。Whang 等人(2022)先用一个独立网络对输入图像 y 进行粗略的去模糊,再基于 Diffusion 学习残差图像,最后将两者相加作为输出,可以生成更加符合人眼视觉效果的去模糊结果,流程如图 6 所示。

然而,对于更加真实的情况,在没有监督信号的情况下,前述方法并不适用,零样本的方法提出只用模糊图像和预训练好的 Diffusion 模型进行复原。GDP(Fei 等, 2023)和 BlindDPS(Chung 等, 2023)都是利用在大型数据集上训练好的 Diffusion 模型进行去模糊,例如 ImageNet 或者 FFHQ 等,此时,模糊图像 y 作为指导信号被嵌入采样过程,利用生成的伪模糊图像 $\hat{y}$ 与真实模糊图像 y 之间的距离损失指导采样过程,生成与 y 有相同内容但不含模糊降质的图像。

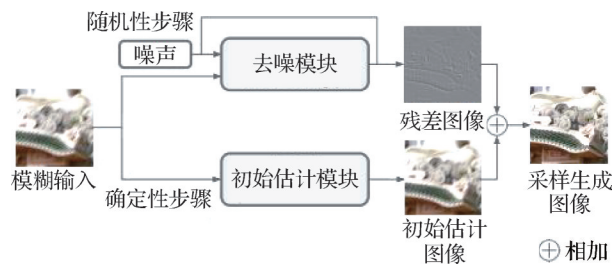


图6 有监督 Diffusion 的去模糊模型 (Whang 等, 2022)
Fig. 6 Supervised Diffusion-based deblurring methods
(Whang et al. , 2022)

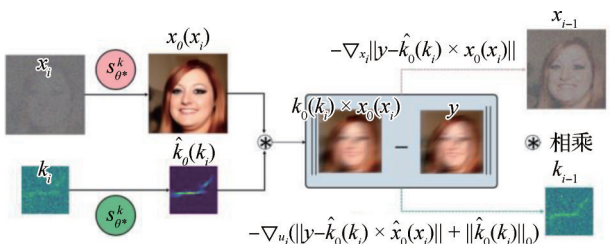


图7 零样本 Diffusion 的去模糊模型 (Chung 等, 2023)
Fig. 7 Zero-shot learning Diffusion-based deblurring methods
(Chung et al. , 2023)

BlindDPS 的处理流程如图 7 所示。

4. 1. 4 小结

综上所述, 3 种不同类型的表征方法能对未知模糊的情况进行处理, 但在分析问题的角度和处理方式上各有不同, 展现出不同的优缺点, 如表 4 所示。由于生成模型优越的数据学习和细节生成能力, 正在成为目前真实场景下去模糊任务中最主流的方法之一, 特别是基于扩散模型 Diffusion 的方法, 不仅模型参数较小, 网络训练简单, 生成的图像效果

更加逼真, 极大提升了图像的视觉效果。然而, 由于 Diffusion 框架需要按照一定的次数对数据进行逐步采样, 往往需要 1 000 步的采样次数, 导致算法效率较低, 而且还会受训练时使用的图像块尺寸的限制, 如 256×256 像素的图像, 在推理时无法直接应用于大尺寸 (如 $1\,280 \times 720$ 像素) 图像。因此, 后续的研究可尝试对扩散模型进行网络加速处理, 以提升高分辨率图像处理任务的性能, 如可以参考 Song 等人 (2020) 的研究, 通过设置合适的采样步长和优化采样步骤, 实现 10~50 倍的效率提升。

表 4 基于未知降质表征的模糊去除方法优缺点总结

Table 4 Summary of advantages and disadvantages of blur removal methods based on unknown degradations			
方法	优势	劣势	难点
降质表征	显式地展示了未知降质的形式, 有助于设计专用的复原方法	实际情况下训练数据集的缺乏, 导致学习多种降质的具体表征是很难的	降质表征数据集的构造, 如何保证学习到的降质表征是准确的
降质不变表征	只考虑图像内容本身的信息, 减少了对标记数据的依赖, 鲁棒性高, 泛化能力强, 更适合真实场景降质	需要复杂的模型和更多的计算资源, 训练的难度较大, 降质不变特征的提取过程会存在信息丢失的问题, 导致图像无法还原	如何寻找合适的权衡点, 能够在进行降质不变表征的同时又不损失太多信息
基于生成式模型的表征	能端到端的从噪声数据中生成高质量清晰图像, 省略了降质因素建模等阶段	基于 GAN 的模型训练过程不够稳定, Diffusion 模型的采样效率较低	如何同时达到好的复原效果和高的算法处理效率

4. 2 额外的数据辅助

当从单一的模糊图像中无法获得更多有效信息进行复原时, 一个自然的想法是利用额外的数据或者信息进行辅助, 以解决更加真实场景下的模糊去除问题。本文对已有的文献在此方向所作的工作分为 3 类进行介绍: 1) 跨帧图像辅助, 包括序列其他帧或数据库中的参考图像; 2) 跨模态数据辅助, 如事件相机数据、红外数据等; 3) 跨任务辅助, 包括语义分割、识别分类及深度估计和三维重建等。

4. 2. 1 跨帧的图像辅助

前两节所述的方法解决大多数单帧图像恢复的

问题, 但由于单幅图像中信息缺乏严重, 这些先进方法表现不佳, 去模糊的结果中仍然存在许多振铃和噪声。相比之下, Rav-Acha 和 Peleg (2000) 使用具有不同模糊方向的两个运动模糊图像进行场景复原, 首次引入了多帧图像的信息, 使用了两幅分别在水平、垂直方向模糊的图像去模糊, 由于两幅图像保留了更多的细节信息并且能够相互补充, 展示了多帧复原效果优于单幅图像。后来, 他们再次明确提出两幅运动模糊图像比一幅图像更好的理论 (Rav-Acha 和 Peleg, 2005), 引领了多帧图像复原方法的发展。

多帧图像可对场景信息进行更加完整的描述,为许多具有挑战性的图像处理任务提供了新思路。

Chen 等人(2008)扩展了 Rav-Acha 和 Peleg (2005)提出的方法,用同一场景的模糊图像进行互补去模糊。Shin 等人(2021)联合多个模糊图像的互补信息来捕获单个清晰图像的图像统计数据。Yuan 等人(2007)借助噪声图像进行图像去模糊,两幅图像一起用于模糊核估计,以提高准确性,利用噪声图像中隐含的清晰图像特征进行去模糊,并且设计了一种迭代反卷积算法,能估计出良好的初始模糊核并显著减少反卷积伪影。Deepa 和 Jiji(2019)首先对噪声图像进行去噪,然后将其用做参考图像来估计非均匀模糊图像的运动矢量,将模糊图像分割成匀速运动的图像块并进行去模糊。大多数这样的方法需要使用特定硬件设计相机捕获图像的过程,以获得场景的多个图像,使模糊核更易估计,去模糊过程对噪声也更鲁棒。Li 等人(2011)设计了双图像采集系统,能同时捕获同一场景的两个对齐的模糊图像,无需额外的图像配准操作,基于捕获的图像对,提出的去模糊算法能够给出更准确的核估计并获得更高质量的图像恢复。

Hacohen 等人(2013)提出了一种使用清晰示例图像进行去模糊的方法,参考图像包含一些与模糊照片共有的但是清晰的内容,不包含降质信息。Pan 等人(2019a)对模糊人脸问题,提出基于参考图像的去模糊方法,所提出的方法使用来自参考图像的面部结构和可靠边缘进行核估计,而不需要求助于复杂的边缘预测。Li 等人(2018)基于不同位姿的参考图像进行人脸复原,引入光流估计模型来预测光流场,指导为正确的姿势和表情,提出的重建网络将降质图像和具有不同位姿的引导图像作为输入进行重建,产生包含精细人脸细节的结果。Xiang 等人(2020)开发了基于编解码器网络探索给定的参考图像中的清晰度特征,以指导网络更好地恢复潜在图像,结果表明使用示例中的清晰特征可以帮助模糊去除和细节恢复。Li 等人(2022c)提出了一种方法,该方法涉及利用额外的模糊图像和清晰图像作为参考样本。他们设计了两种编码器网络,一种用于提取输入图像的特征,另一种用于提取参考图像的特征。通过一个排序机制,系统能够识别并转移与输入图像最相关的特征,从而生成排序后的特征集。最终,这些排序特征被用来重建图像的潜在表示。

相比之前的方法要求示例图像应与输入图像对齐或具有相似的上下文,从而限制了示例图像的有效使用,而 Li 等人(2022c)对输入图像的内容不做过多限制,具有更强的泛化能力。

4.2.2 跨模态的数据辅助

除了常规的图像作为辅助信息以外,通过其他传感器获取的数据也可以作为多模态的辅助数据帮助图像复原的过程,例如,利用红外多光谱数据以及事件相机数据等。

相比于可见光的光谱波长落在为 400~800 nm 的区间内,红外传感器可捕捉光谱波长为 800~1 000 nm 的非可见红外信息,形成近红外图像(near-infrared, NIR),可以用于辅助传统彩色 RGB(red green blue)可见光图像复原。Yue 等人(2014)假设传统的 RGB 全色图像和近红外图像在局部区域上具有线性映射关系,因而,可从模糊 RGB 图像和相应清晰的 NIR 图像中恢复出清晰 RGB 图像。Li 等人(2013)介绍了一种手持式多光谱相机,可同时捕获一对模糊和近红外闪光图像,为恢复高频细节,该方法利用近红外梯度约束作为一种图像正则化方法,并将其集成到最大后验算法中,迭代地进行模糊核估计和图像恢复。该方法成功地实现了 RGB 和 NIR 图像的融合,并具有降噪、细节保留和颜色增强功能。

事件相机(event camera)作为新型的传感器,近年来得到了很多关注和发展,已经有大量的文献印证了事件相机数据在图像去模糊任务上的显著性能,原因在于事件相机能够以极低的时间延迟率捕获数据,并能够在较暗的环境中精准记录目标的运动信息,借助该特性极大地提升了图像复原效果,目前已经成功应用于图像超分辨、图像去模糊以及高动态范围成像等领域。Pan 等人(2019b)首先将事件相机和图像去模糊任务联系起来,构建了基于事件相机的去模糊数学模型:双重积分模型(event double integral, EDI),为之后的方法奠定了理论基础。但由于模型较为简单,对事件响应阈值和噪声没有做特殊处理,复原效果一般。后来,Jiang 等人(2020)、Shang 等人(2021)、Xu 等人(2021)以及 Zhang 和 Yu(2022)等在 EDI 的基础上,提出了动态学习阈值的方法以及利用序列前后帧的事件变化提升了单帧去模糊、视频去模糊以及视频插帧等任务的性能。Wang 等人(2020a)同样依靠 EDI 模型,并借助稀疏编码理论,对低分辨率的 event 数据和模糊

图像同时进行去噪、去模糊和超分辨,并重建出高质量的图像。Sun 等人(2022)利用独立的编码器对事件数据进行特征提取,为了有效地融合事件和图像特征,设计了一个应用于网络多个级别的事件—图像跨模态注意模块,更关注事件分支的相关特征并滤除噪声。

还有学者从硬件上改善数据获取的过程。Jeon 等人(2017)提出一种用于多图像运动去模糊的新颖的编码曝光视频技术,基于计算成像理论,使用一组互补的颤动模式捕获视频帧,能够保留用于恢复清晰图像的所有频谱带。Hu 等人(2016)通过惯性传感器(即陀螺仪和加速计)记录相机运动,然后根据这些读数重建空间变化的模糊核。双像素(Dual-pixel)传感器记录的数据经常用于处理散焦模糊问题,通过在单幅图像拍摄中捕获场景的两个子孔径视图来协助相机自动对焦,两个子孔径图像在用于计算聚焦在特定场景区域的适当镜头位置后被丢弃,基于此,Abuolaim 和 Brown(2020)分析并利用了大多数现代相机上的双像素传感器上的可用数据,引入一种深度神经网络架构,使用 Dual-pixel 丢弃的子孔径图像来减少散焦模糊。Lai 等人(2022)开发了一种基于手机双摄像头融合技术的新型人脸去模糊系统,该系统检测目标运动并动态启用参考相机,例如超广角相机,并以更快的快门设置捕获辅助照片,其中主镜头捕获噪声低但模糊的图像,而参考镜头会获得清晰但噪声大的图像。通过网络模型来对齐融合这两幅图像,并输出一幅无运动模糊的清晰图像。

相比于额外帧数据辅助的方法,由于利用跨模态数据辅助方法包含了更多关于场景的运动信息,如事件相机能够捕获到无模糊的运动,红外相机能够在暗夜里捕捉到运动目标,基于跨模态数据进行去模糊可以获得更好的效果,例如基于事件相机的方法(Sun 等,2022)目前在标准去模糊数据集 GoPro (Nah 等,2017)上取得了当前最好的指标。

4.2.3 跨任务的图像辅助

对于降质非常严重的模糊数据,可直接用于复原的信息很少,有学者着手研究如何用其他高级视觉任务对复原任务进行辅助,如语义分割、深度估计等,由于高级视觉任务能提取出更高层的上下文内容信息,可对严重降质的图像提供全局的语义信息,帮助模型实现场景理解,真实场景下的严重模糊可以在一定程度上被去除,并生成符合目标特性更加

真实的结构和纹理,进一步提升复原效果。

Zhang 等人(2011)提出基于目标识别辅助去模糊方法,联合基于稀疏表示先验的联合盲图像复原和目标识别任务,以处理低质量图像中人脸识别的挑战性问题,其中退化模型是真实且完全未知的。该方法将图像复原和识别任务融合在一个框架中,通过迭代求解盲图像恢复以追求识别的最稀疏表示来实现同时恢复和识别,使二者得到了相互促进。Shen 等人(2018)首次提出采用语义分割图作为辅助数据指导人脸图像去模糊问题,通过将人脸分割图作为额外特征加入到网络中,取得了明显的效果。后来,Shen 等人(2020)和 Chen 等人(2021a)等进一步扩展了语义分割指导下的人脸去模糊任务。Pan 等人(2016a)和 Ren 等人(2017)等对自然图像的去模糊问题进行了探索,证实了语义分割任务在自然图像去模糊问题上的有效性。不仅如此,语义图还被用于辅助超分辨任务。SFTGAN(spatial feature transform)(Wang 等,2018)提出了语义特征转换模块,将从分割模型中输出的分割图融合到超分辨框架中,实现语义信息的嵌入。Wang 等人(2020b)提出了超分辨—语义分割的双分支网络模型,同时完成超分辨和语义分割任务,超分辨分支的存在帮助了分割可以得到高分辨的分割图像,同时语义分割也为超分辨分支提供了语义特征指导,达到相得益彰的目的。

深度估计任务可提取场景深度和运动信息,用于辅助图像去模糊任务。Xu 和 Jia(2012)指出现有的非一致去模糊方法不能很好地处理不同深度的图像,因为当物体离相机不远时,空间变化的点扩散函数会随着边界的突然变化而出现,使得近处比远处模糊得多,平滑的模糊核变化不再适用于不同深度目标的边缘。因此,基于两视图立体分析,用两幅图像形成一个深度感知模糊核来处理以上问题。Li 等人(2020a)提出了一种由深度估计任务引导的图像去模糊模型,它生成一个初始深度图,然后以联合学习的方式细化深度图,同时实现去模糊。Hu 等人(2022)提出三维重建辅助下的去模糊方法,针对人脸数据,提出跨任务和跨模型的即插即用3D面部先验,将清晰的面部结构嵌入到复原模型中,该先验基于面部属性参数,例如身份、面部表情、纹理、照明和面部姿势等,融合成3D的可变形知识,该3D先验比最先进的算法实现了更出色的人脸恢复结果。

4.2.4 小结

对于严重模糊场景,单靠原始模糊图像难以获得有效信息进行图像复原,从硬件角度改善数据获取途径或引入其他任务进行辅助去模糊是合理的方法。本文从不同角度总结了现有方法,汇总如表 5 所示。可以直观地看出,该类方法核心的研究内容

及难点是需要考虑应该使用何种辅助信息,以及如何获取该类辅助信息。由于不同模态或者不同任务的数据常常处于不同的数据域中,需要额外进行合适的数据预处理,增加了任务的复杂性。正是由于额外信息的加入,严重模糊图像才有了复原的可能,即使目前的恢复效果还有待进一步提升和完善。

表 5 额外信息辅助下的模糊去除方法优缺点总结

Table 5 Summary of advantages and disadvantages of blur removal methods with additional information assistance

方法		优势	劣势	难点
跨帧图像	不同模糊图像	为模糊提供了其他清晰的内容信息	需场景矫正技术将数据配准,相较于单帧独立处理,算法计算量相对较大	如何设计合适的配准模块
	其他示例图像			如何提取清晰特征
跨模态数据	红外相机数据	改善在暗夜等场景下的信息缺失问题	需要进行跨模态数据预处理,效率不如单帧处理,且不恰当的跨模态数据融合方式会严重影响复原结果	如何融合不同模态的数据以及如何设计理想的图像数据获取系统
	事件相机数据	以“无模糊”的方式记录事件数据,提供传统相机无法捕获的运动信息		
	相机硬件改善	能从数据捕获源头提升复原质量	硬件设计所需的经济成本较高	
跨任务数据	目标识别分割	为模型提供了场景内的全局上下文信息,提高了严重模糊问题的处理性能	图像模糊现象可能会导致其他任务无法有效执行,需额外训练模糊情况下的其他高级任务模型	如何有效利用跨任务模型产生的语义信息,如何处理模糊对其他任务的影响
	深度估计、三维重建	为模型提供了场景深度和位姿等信息,有助于运动信息的获取		

5 真实场景下图像去模糊数据验证

5.1 常用数据集

实际应用中,为了更好地展示各种方法在真实世界去模糊任务中的效果,用于仿真场景的数据集已不再适用,由于用来训练的真实模糊—清晰图像对较难捕获,因此,目前大多去模糊方法会先在一些仿真数据集上进行训练,再在真实数据上测试。另一方面,也有研究者尝试构建相机系统,构建真实去模糊数据集。一些可用于真实场景去模糊的数据集有:

1) Köhler 数据集(Köhler 等,2012)。模糊图像由通过模拟 6 自由度移动的相机的投影方法生成。该数据集包含 4 幅潜在图像,每幅图像都应用了 12 个 6D 相机运动轨迹,总共产生了 48 个模糊图像。

2) Lai 数据集(Lai 等,2016)。在 Köhler 数据集的基础上,同样使用均匀模糊核和 6D 相机轨迹生成 100 幅真实模糊图像和 200 幅合成的模糊图像,涵盖了更多场景,如户外、人脸、文本和低光图像等。

3) GoPro 数据集(Nah 等,2017)。该数据集是用

GoPro 相机以 720×1280 像素的分辨率拍摄的 240 帧/s 视频制作的,作为基准数据集广泛应用于各个算法。模糊图像通过对不同数量(7~13)的连续潜在帧进行叠加求平均产生,包含 3 214 对图像,其中 2 103 对用于训练,1 111 对用于测试。由于场景多为户外环境,包含行人和建筑等,目标的运动方式变化多样。

4) RealBlur 数据集(Rim 等,2020)。该数据集构建了专门的双摄像头图像采集系统,可同时捕获几何时对齐的模糊和清晰图像对,共由 232 个不同场景的 4 738 对图像组成,所有图像均以相机原始格式和 JPEG 格式捕获,并构成两个数据集:来自原始图像的 RealBlur-R 和经 JPEG 压缩的 RealBlur-J。每个训练集由 3 758 个图像对组成,测试集由 980 个图像对组成。

5) RSBlur 数据集(Rim 等,2022)。该数据集构建了仿真数据的全链路模型,提供了成对的真实模糊图像和合成模糊图像,以及对应的真实清晰图像。训练集、验证集和测试集分别包含 8 878、1 120 和 3 360 幅模糊图像。相比于 GoPro 和 RealBlur,仿真的模糊图像更加真实,包括饱和区域及噪声的添加。

6)RWBI数据集(Zhang等,2020b)。包含来自22个不同场景的3112幅真实模糊图像。这些模糊图像是从各种移动设备获得的,包括华为P30 Pro、三星S9 Plus、iPhone XS和GoPro Hero5 Black相机。但RWBI数据集仅包含真实的模糊图像,而没有相应的清晰图像。

7)MC-Blur数据集(Zhang等,2024)。该数据集是一个用于图像去模糊的大规模多原因数据集,由4种模糊类型组成:均匀模糊、平均连续帧的运动模糊、严重散焦模糊和真实世界模糊,涵盖建筑物、城市和车辆等1000多个不同场景;由4个子集组成,即基于真实高帧率的运动模糊集(real high-FPS based motion-blurred subset, RHM)、大模糊核运动

模糊集(large-kernel UHD motion-blurred subset, UHDM)、大规模散焦模糊集(large-scale heavy defocus blurred subset, LSD)和真实混合模糊集(real mixed blurry qualitative subset, RMBQ)。其中,RMBQ子集提供了各种移动设备捕获的大规模、真实模糊图像,包含10000幅模糊图像,但不含对应的清晰图像,为未来在现实场景方面的定性研究提供了可靠的测试平台。

以上常用数据集总结如表6所示,目前在去模糊领域,公认的标准数据集主要是GoPro和RealBlur,分别表示在合成和真实图像上的实验效果。由于MC-Blur数据规模十分庞大,并且较新,目前在该数据集上的实验验证还较少,后续研究还需做更多探索。

表6 真实场景下图像去模糊数据集总结
Table 6 Summary of image deblurring datasets in real scenes

数据集	生成模糊类型	清晰图像数	模糊图像数	场景
Köhler数据集(Köhler等,2012)	合成模糊	4	48	4
Lai数据集(Lai等,2016)	合成模糊和真实模糊	合成25、真实100	合成200、真实100	6
GoPro数据集(Nah等,2017)	合成模糊	3214	3214	33
RealBlur数据集(Rim等,2020)	真实模糊	9476	9476	232
RSBlur数据集(Rim等,2022)	合成模糊和真实模糊	13358	13358	697
RWBI数据集(Zhang等,2020b)	真实模糊	-	3112	22
MC-Blur数据集(Zhang等,2024)	合成模糊和真实模糊	合成92300、真实-	合成110300、真实10000	1000

注:“-”表示无数据。

5.2 常用评价指标

去模糊评价指标用于量化去模糊算法的复原结果质量,有助于评估算法的性能。以下是一些常见的图像去模糊评价指标。

1)峰值信噪比(peak signal-to-noise ratio,PSNR)。一种广泛使用的评价指标,用于衡量图像质量,计算原始图像和复原图像间的峰值信噪比,以对比两者之间的差异,具体为

$$PSNR = 10\lg\left(\frac{MAX^2}{MSE}\right) \tag{9}$$

式中,MAX和MSE分别表示图像像素的最大值和均方误差,较高的PSNR值表示更好的图像质量。

2)结构相似度(structural similarity index,SSIM)。结合了亮度、对比度和结构信息,测量原始图像和去模糊图像间的相似性,计算为

$$SSIM(x, \hat{x}) = \frac{(2\mu_x\mu_{\hat{x}} + C_1)(2\sigma_{x\hat{x}} + C_2)}{(\mu_x^2 + \mu_{\hat{x}}^2 + C_1)(\sigma_x^2 + \sigma_{\hat{x}}^2 + C_2)} \tag{10}$$

式中, μ_x 和 $\mu_{\hat{x}}$ 分别表示图像 x 和 $\hat{x}$ 的均值, σ_x 和 $\sigma_{\hat{x}}$ 为标准差, $\sigma_{x\hat{x}}$ 为协方差, C_1 和 C_2 是用于稳定计算的常数,较高的SSIM分数表示更好的图像质量。

3)学习感知图像块相似度(learned perceptual image patch similarity,LPIPS)。一种基于深度学习的感知图像相似度指标,旨在模拟人类感知图像质量的方式,使用卷积神经网络来比较图像的感知相似性,需要真值图像作为参考,实现有参考评价,较低的LPIPS值表示更好的质量。

4)自然图像质量评估(naturalness image quality evaluator,NIQE)。是一种无参考图像质量评估方法,其不需要参考原始图像,也无需清晰图像来评估图像质量。NIQE依赖于图像内部的统计特性和降

质感知模型,使用一系列特征提取器和学习模型,通过分析图像的统计特性(如对比度、锐度和颜色分布等)来预测图像的质量,较低的 NIQE 值表示更好的质量,值越高则意味着图像的质量可能受到了较大程度的损失或存在明显的失真。

降质重建指标 PSNR 和 SSIM 值越高质量越好,视觉重建指标 LPIPS 和 NIQE 值越低质量越好。从质量评价角度来讲,PSNR、SSIM 和 LPIPS 是全参考

质量评价,无参考评价 NIQE 用于无真值图像情况。对于真实图像评估,在无法定量评估时,常常选择定性对比,直接比较不同方法之间的复原效果。

5.3 部分方法的性能对比

本节展示了部分方法在 5.1 节部分数据集上的性能对比结果,以总结目前方法的性能,定量结果如表 7—表 10 所示,分别表示相关的方法在 GoPro、RealBlur-R 和 RSBlur 数据集上的测试结果。

表 7 GoPro 数据集上的性能评估
Table 7 Performance evaluation on GoPro dataset

对比方法	模型结构	PSNR/dB	SSIM
Sun 等人(2015)	基于运动场复原	24.64	0.843
Gong 等人(2017)	基于运动场复原	26.06	0.863
DeepDeblur(Nah 等,2017)	多尺度结构	29.23	0.916
SRN(Tao 等,2018)	多尺度循环结构	30.26	0.934
Zhang 等人(2018a)	多尺度循环结构	29.19	0.932
Gao 等人(2019)	多尺度循环结构	31.58	0.948
MSSNet(Kim 等,2022)	多尺度结构	33.39	0.964
DMPHN(Zhang 等,2019)	多图像块结构	31.50	0.948
MT-RNN(Park 等,2020)	多时间循环结构	31.15	0.945
MPRNet(Zamir 等,2021)	多阶段结构	32.66	0.959
HINet(Chen 等,2021c)	多阶段结构	32.71	—
MIMO-UNet(Cho 等,2021)	多输入多输出	32.45	0.957
IPT(Chen 等,2021b)	Transformer	32.52	—
Restormer(Zamir 等,2022)	Transformer	32.92	0.961
Uformer(Wang 等,2022)	Transformer	33.06	0.967
Stripformer(Tsai 等,2022a)	Transformer	33.08	0.962
DeblurGAN-v2(Kupyn 等,2019)	GAN	29.55	0.934
DBGAN(Zhang 等,2020b)	GAN	31.10	0.942
DPMDebur(Wang 等,2022)	Diffusion	33.23	0.963
Li 等人(2022a)	未知降质表征	33.28	0.964
RCM(Li 等,2022c)	示例图像辅助	30.30	0.899
BANet(Tsai 等,2022b)	Event 辅助	32.54	0.957
EFNet(Sun 等,2022)	Event 辅助	35.46	0.972
Li 等人(2020a)	深度估计辅助	30.49	0.938

注:加粗字体表示最优结果,“—”表示无数据。

从对比结果可以看出,针对单一模糊去除的方法在数量上是显著多于其他两类的,而在 GoPro 数据集上的结果多余 RealBlur 和 RSBlur 数据集。这是因为 GoPro 是基于仿真成像过程构建的数据集,作为最广泛使用的数据集之一,已经成为本任务的标

准数据集,因此,大多数方法均会在其上进行测试对比,以说明方法的有效性。同时,可以发现,2021 年之前大多数研究还是在做网络结构设计,后来涌现出 Transformer 和 Diffusion 等模型,方法越来越多样化,效果也更好。与降质重建指标 PSNR 和 SSIM 相

表 8 RealBlur 数据集上的性能评估
Table 8 Performance evaluation on RealBlur dataset

方法	模型结构	PSNR/dB	SSIM
SRN(Tao 等,2018)	多尺度循环结构	38.65	0.965
MPRNet(Zamir 等,2021)	多阶段结构	39.31	0.972
MSSNet(Kim 等,2022)	多尺度结构	39.76	0.972
Stripformer(Tsai 等,2022a)	Transformer	39.84	0.974
DeblurGAN-v2 (Kupyn 等,2019)	GAN	36.44	0.934
BANet(Tsai 等,2022b)	Event 辅助	39.55	0.971
EFNet(Sun 等,2022)	Event 辅助	38.12	0.973

注:加粗字体表示各列最优结果。

表 9 RSBlur 数据集上的性能评估
Table 9 Performance evaluation on RSBlur dataset

方法	模型结构	PSNR/dB
SRN(Tao 等,2018)	多尺度循环结构	32.53
MIMO-UNet(Cho 等,2021)	多输入多输出结构	32.73
MIMO-UNet+(Cho 等,2021)	多输入多输出结构	33.37
MPRNet(Zamir 等,2021)	多阶段结构	33.61
Restormer(Zamir 等,2022)	Transformer	33.69
Uformer(Wang 等,2022)	Transformer	33.98

注:加粗字体表示各列最优结果。

比,基于视觉感知的指标 LPIPS 能够更加体现出人

眼对图像质量的要求,特别是在当前有越来越多基于生成模型的方法被提出,如表 10 所示,这些方法在 LPIPS 上的数值更加突出,可见人眼视觉重建感知指标更适应于这类方法的评估。

表 10 GoPro 数据集上的 LPIPS 评估
Table 10 LPIPS evaluation on GoPro dataset

方法	模型结构	LPIPS
DeepDeblur(Nah 等,2017)	多尺度结构	0.181 9
SRN(Tao 等,2018)	多尺度循环结构	0.131 9
Zhang 等人(2018a)	多尺度循环结构	0.162 7
DMPHN(Zhang 等,2019)	多图像块结构	0.128 0
MSSNet(Kim 等,2022)	多尺度结构	0.087 3
MIMO-UNet(Cho 等,2021)	多输入多输出	0.119 6
DeblurGAN-v2(Kupyn 等,2019)	GAN	0.252 8
DPMDDeblur(Whang 等,2022)	Diffusion	0.078 3
EFNet(Sun 等,2022)	Event 辅助	0.064 2
Li 等人(2020a)	深度估计辅助	0.121 4

注:加粗字体表示各列最优结果。

图 8—图 10 展示了一些方法在 GoPro、RealBlur 及 RWBI 数据集上的复原结果,可以直接比较不同方法的处理性能,一般而言,所提方法需综合对比定量和定性结果评价。

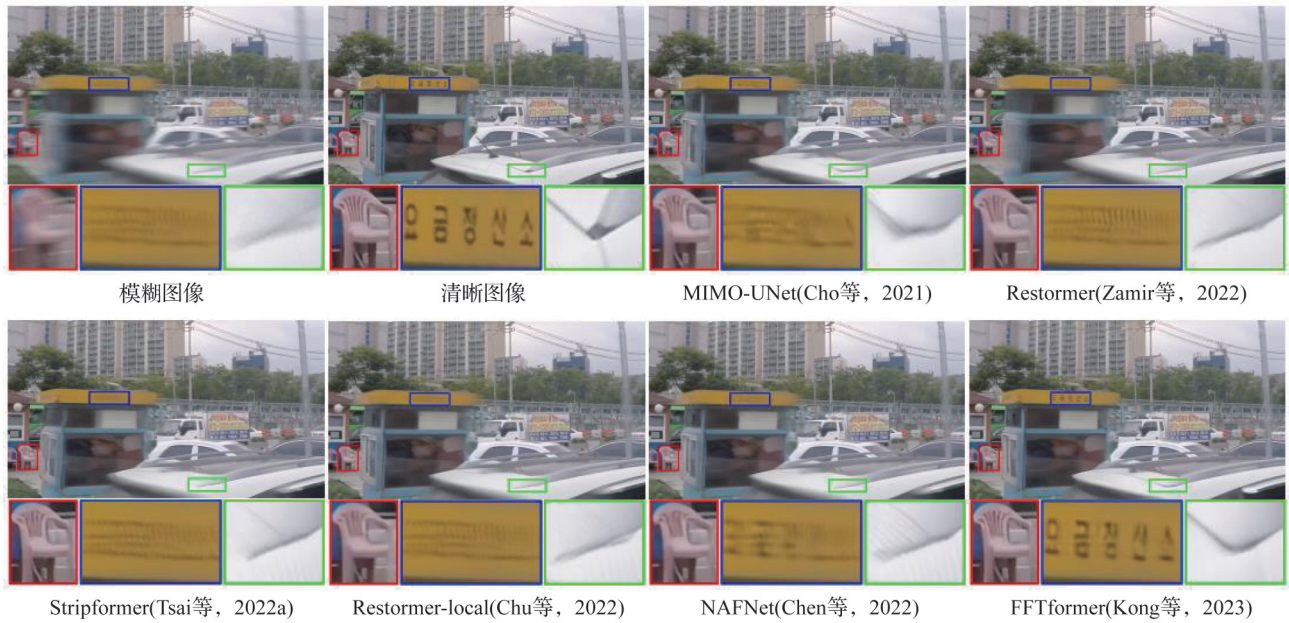


图 8 GoPro 数据集上部分方法定性对比(Kong 等,2023)

Fig. 8 Qualitative comparison of some methods on GoPro dataset(Kong et al. ,2023)

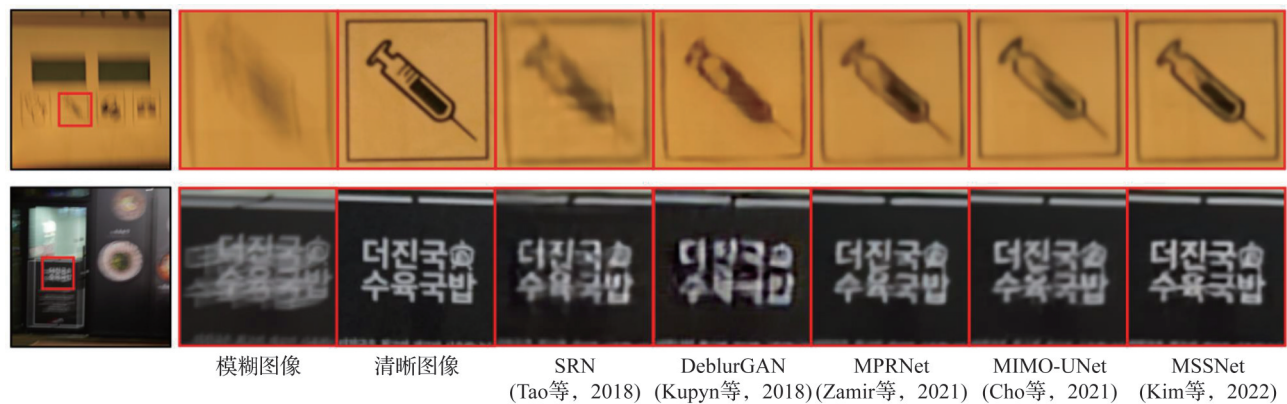


图9 RealBlur数据集上部分方法定性对比(Kim等,2022)

Fig. 9 Qualitative comparison of some methods on RealBlur dataset(Kim et al. ,2022)

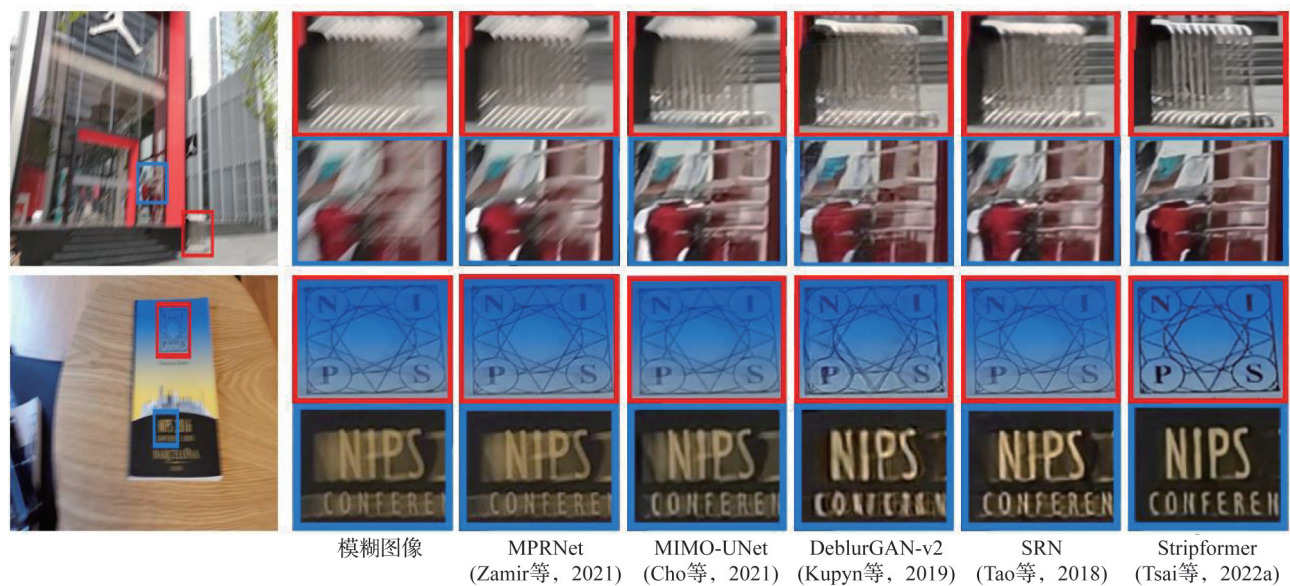


图10 RWBI数据集上部分方法定性对比(Tsai等,2022a)

Fig. 10 Qualitative comparison of some methods on RWBI dataset(Tsai et al. ,2022a)

6 结 语

6.1 本文内容总结

针对真实场景去模糊问题,本文整理了自2000年以来的相关文献,按照退化模型、任务难度由易到难的原则,对现有的研究方法进行了归纳总结:

1)目前的深度学习模型可以显著地实现对单一模糊的去除,特别是在处理非均匀模糊时,表现出良好的复原效果。然而,由于退化模型的单一性,即使基于最先进的Transformer模型也很难有效去除真实场景下的复杂模糊。同时,复杂的网络模型限制了模型的有效训练,影响算法性能的提高。

2)针对复合模糊问题,相比于联合复原的双分

支重建方法需额外的特征融合模块,单支路和一体化方法可直接用同一个网络恢复各种退化,能处理更多类型的降质。然而,由于不同降质之间可能存在干扰,该方法需要在同一个模型里区分不同的降质类型和程度,且需要更高的泛化性能。

3)由于未知模糊降质更符合真实场景,基于未知降质表征的方法能够从物理成像过程所暗含的降质因素进行考虑,而基于额外信息辅助的去模糊方法则利用其他渠道获得的数据进行处理,例如跨模态数据等,从信息获取的源头对模糊进行抑制,从而进一步提升复原结果,二者各有利弊,但是都能够实现对真实场景的未知降质模糊问题的处理。

6.2 未来研究热点

尽管目前图像去模糊研究已经取得了一定进

展,然而在实际应用中,由于成像环境的复杂性,仍面临很多挑战性问题,本文总结未来研究热点如下:

1)真实世界训练数据获取问题。与传统方法不同,当前大多研究都基于深度学习,由于其具有强大的特征表示能力,极大提升了复原效果。然而,基于深度学习的方法往往需要大量模糊—清晰成对数据进行模型训练,在真实情况下,数据集的获取工作是很困难的。该问题激发出两个研究方向:一个是真实场景数据集的构造,目的是尽可能捕获更加真实的模糊和清晰图像对,如Rim等人(2020,2022)分析了导致真实模糊图像和合成模糊图像之间差异的各种因素,揭示了模糊生成过程中不同因素的影响;另一个是不依靠监督信号的无监督学习(Lu等,2019)或者基于域自适应的迁移学习(Zhuang等,2024)等方法,这些方法是在目前公开的数据集的基础上,通过改进网络结构和学习框架进而提升去模糊效果,但由于问题的复杂性,该类方法的性能还有待进一步提升。

2)视觉重建指标和降质重建指标之间的不平衡问题。由于一些降质重建算法可能会倾向于过度平滑图像,以最大程度地减小均方误差,然而这样的过度平滑会导致视觉上的细节丢失,降低了视觉质量,但在峰值信噪比等指标上仍然表现良好。同时,人眼对图像的感知是复杂的,与简单的像素差异不完全一致。例如,即使一些细微的纹理或高频细节受到了损失,图像仍然可以看起来很好。因此,降质重建指标可能无法准确捕捉到这些感知差异,这些均导致图像复原过程存在视觉质量和降质重建指标之间不平衡的问题。针对该问题,Blau和Michaeli(2018)进行了详细的分析和总结,证明了视觉质量和重建指标之间存在一条曲线,该曲线的左下角的区域是任何算法都无法达到的。例如,基于CNN的方法(Nah等,2017)可以提升重建指标但是视觉质量较差;而基于生成式模型(Kupyn等,2018)的结果拥有较好的视觉质量但是重建质量较差。因此,如何找到一条曲线,使得视觉质量和重建指标都达到较高水平是之后的研究重点。

3)复杂网络的高性能和低运行效率之间存在的平衡问题。随着模型层数增加、参数量增多、宽度加大等,如今的网络模型的规模远比最初卷积网络大,虽然具有较强的性能,但在运行时往往需要更多的计算资源,使运行时间延长和硬件成本增加。实

际应用时,需考虑算法的运行效率,过慢的运行速度和过高的计算量会阻碍算法的移植和应用,因此,研究算法的实时性处理是当前的热点问题。针对该问题,有学者使用轻量网络和模型压缩技术,在计算效率和去模糊性能间取得平衡,使其适合实时应用和资源受限的环境。Kupyn等人(2019)利用移动平台模型替换复杂的残差网络,使运行效率大幅提升。基于知识蒸馏的模型压缩方法(Zhang等,2021)从复杂的教师模型中提取知识,并将知识传递给更小、更轻量的学生模型,实现模型压缩,减小网络体积和计算复杂度同时保持其性能。其他的加速方法还包括使用并行计算和分布式计算技术,将任务分解为多个子任务并在多个处理单元上同时执行,以加速处理速度,可通过多核CPU、GPU、分布式计算集群等方式实现。此外,利用专用硬件加速器,如图形处理单元、张量处理单元或硬件加速的现场可编程逻辑门阵列(field programmable gate array, FPGA),可以显著提高算法的速度。

致谢:本文在调研和撰写过程中得到空天一体化大数据应用技术国家工程实验室及陕西省语音与图像信息处理重点实验室的支持,在此表示感谢。

参考文献(References)

- Abuolaim A and Brown M S. 2020. Defocus deblurring using dual-pixel data//Proceedings of the 16th European Conference on Computer Vision — ECCV 2020. Glasgow, UK: Springer: 111-126 [DOI: 10.1007/978-3-030-58607-2_7]
- Albluwi F, Krylov V A and Dahyot R. 2018. Image deblurring and super-resolution using deep convolutional neural networks//Proceedings of the 28th IEEE International Workshop on Machine Learning for Signal Processing. Aalborg, Denmark: IEEE [DOI: 10.1109/MLSP.2018.8516983]
- Bai H R, Cheng S S, Tang J H and Pan J S. 2021. Learning a cascaded non-local residual network for super-resolving blurry images//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Nashville, USA: IEEE: 223-232 [DOI: 10.1109/CVPRW53098.2021.00031]
- Bascle B, Blake A and Zisserman A. 1996. Motion deblurring and super-resolution from an image sequence//Proceedings of the 4th European Conference on Computer Vision on Computer Vision — ECCV'96. Cambridge, UK: Springer: 571-582 [DOI: 10.1007/3-540-61123-1_171]
- Blau Y and Michaeli T. 2018. The perception-distortion tradeoff//Pro-

- ceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 6228-6237 [DOI: 10.1109/CVPR.2018.00652]
- Cai J F., Chan R H and Nikolova M. 2010. Fast two-phase image deblurring under impulse noise. *Journal of Mathematical Imaging and Vision*, 36(1): 46-53 [DOI: 10.1007/s10851-009-0169-7]
- Carlavan M and Blanc-Feraud L. 2012. Sparse Poisson noisy image deblurring. *IEEE Transactions on Image Processing*, 21(4): 1834-1846 [DOI: 10.1109/TIP.2011.2175934]
- Chen C F., Li X M., Yang L B., Lin X H., Zhang L and Wong K Y K. 2021a. Progressive semantic-aware style transformation for blind face restoration//*Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Nashville, USA: IEEE: 11896-11905 [DOI: 10.1109/CVPR46437.2021.01172]
- Chen H T., Wang Y H., Guo T Y., Xu C., Deng Y P., Liu Z H., Ma S W., Xu C J., Xu C and Gao W. 2021b. Pre-trained image processing Transformer//*Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Nashville, USA: IEEE: 12294-12305 [DOI: 10.1109/Cvpr46437.2021.01212]
- Chen J., Yuan L., Tang C K and Quan L. 2008. Robust dual motion deblurring//*Proceedings of 2008 IEEE Conference on Computer Vision and Pattern Recognition*. Anchorage, USA: IEEE: 1-8 [DOI: 10.1109/CVPR.2008.4587830]
- Chen J B., Xiong B S., Kuang F and Zhang Z Z. 2023. Motion deblurring based on deep feature fusion attention and double-scale. *Journal of Image and Graphics*, 28(12): 3731-3743 (陈加保, 熊邦书, 况发, 章照中. 2023. 深度特征融合注意力与双尺度的运动去模糊. *中国图象图形学报*, 28(12): 3731-3743) [DOI: 10.11834/jig.220931]
- Chen L Y., Lu X., Zhang J., Chu X J and Chen C P. 2021c. HINet: half instance normalization network for image restoration//*Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. Nashville, USA: IEEE: 182-192 [DOI: 10.1109/CVPRW53098.2021.00027]
- Chen W T., Huang Z K., Tsai C C., Yang H H., Ding J J and Kuo S Y. 2022. Learning multiple adverse weather removal via two-stage knowledge learning and multi-contrastive regularization: toward a unified model//*Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. New Orleans, USA: IEEE: 17632-17641 [DOI: 10.1109/Cvpr52688.2022.01713]
- Cho S J., Ji S W., Hong J P., Jung S W and Ko S J. 2021. Rethinking coarse-to-fine approach in single image deblurring//*Proceedings of 2021 IEEE/CVF International Conference on Computer Vision (ICCV)*. Montreal, Canada: IEEE: 4641-4650 [DOI: 10.1109/ICCV48922.2021.00460]
- Chu X., Chen L., Chen C and Lu X. 2022. Improving image restoration by revisiting global information aggregation//*Proceedings of the 17th European Conference on Computer Vision — ECCV 2022*. Tel Aviv, Israel: Springer: 53-71. [DOI: 10.1007/978-3-031-20071-7_4]
- Chung H., Kim J., Kim S and Ye J C. 2023. Parallel diffusion models of operator and image for blind inverse problems//*Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Vancouver, Canada: IEEE: 6059-6069 [DOI: 10.1109/CVPR52729.2023.00587]
- Deepa P L and Jiji C V. 2019. Non-uniform deblurring from blurry/noisy image pairs//*Proceedings of the 4th International Conference on Computer Vision and Image Processing*. Jaipur, India: Springer: 216-226 [DOI: 10.1007/978-981-15-4015-8_19]
- Dey N., Blanc-Feraud L., Zimmer C., Roux P., Kam Z., Olivo-Marin J C and Zerubia J. 2006. Richardson-Lucy algorithm with total variation regularization for 3D confocal microscope deconvolution. *Microscopy Research and Technique*, 69(4): 260-266 [DOI: 10.1002/jemt.20294]
- Du W C., Chen H and Yang H Y. 2020. Learning invariant representation for unsupervised image restoration//*Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle, USA: IEEE: 14483-14492 [DOI: 10.1109/CVPR42600.2020.01449]
- Dupe F X., Fadili J M and Starck J L. 2009. A proximal iteration for deconvolving Poisson noisy images using sparse representations. *IEEE Transactions on Image Processing*, 18(2): 310-321 [DOI: 10.1109/TIP.2008.2008223]
- Fang Z X., Wu F F., Dong W S., Li X., Wu J J and Shi G M. 2023. Self-supervised non-uniform kernel estimation with flow-based motion prior for blind image deblurring//*Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Vancouver, Canada: IEEE: 18105-18114 [DOI: 10.1109/CVPR52729.2023.01736]
- Fei B., Lyu Z Y., Pan L., Zhang J Z., Yang W D., Luo T Y., Zhang B and Dai B. 2023. Generative diffusion prior for unified image restoration and enhancement//*Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Vancouver, Canada: IEEE: 9935-9946 [DOI: 10.1109/CVPR52729.2023.00958]
- Fergus R., Singh B., Hertzmann A., Roweis S T and Freeman W T. 2006. Removing camera shake from a single photograph. *ACM Transactions on Graphics*, 25(3): 787-794 [DOI: 10.1145/1141911.1141956]
- Gao H Y., Tao X., Shen X Y and Jia J Y. 2019. Dynamic scene deblurring with parameter selective sharing and nested skip connections//*Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Beach, USA: IEEE: 3848-3856 [DOI: 10.1109/CVPR.2019.00397]
- Gong D., Yang J., Liu L Q., Zhang Y N., Reid I., Shen C H., Van Den Hengel A and Shi Q F. 2017. From motion blur to motion flow: a

- deep learning solution for removing heterogeneous motion blur//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE: 2319-2328 [DOI: 10.1109/CVPR.2017.405]
- Gupta A, Joshi N, Lawrence Zitnick C, Cohen M and Curless B. 2010. Single image deblurring using motion density functions//Proceedings of the 11th European Conference on Computer Vision on Computer Vision — ECCV 2010. Heraklion, Crete, Greece: Springer: 171-184 [DOI: 10.1007/978-3-642-15549-9_13]
- Hacohen Y, Shechtman E and Lischinski D. 2013. Deblurring by example using dense correspondence//Proceedings of 2013 IEEE International Conference on Computer Vision. Sydney, Australia: IEEE: 2384-2391 [DOI: 10.1109/ICCV.2013.296]
- Hu X B, Ren W Q, Yang J L, Cao X C, Wipf D, Menze B, Tong X and Zha H B. 2022. Face restoration via plug-and-play 3D facial priors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(12): 8910-8926 [DOI: 10.1109/TPAMI.2021.3123085]
- Hu X B, Ren W Q, Yu K C, Zhang K H, Cao X C, Liu W and Menze B. 2021. Pyramid architecture search for real-time image deblurring//Proceedings of 2021 IEEE International Conference on Computer Vision. Montreal, Canada: IEEE: 4278-4287 [DOI: 10.1109/ICCV48922.2021.00426]
- Hu Z, Cho S, Wang J and Yang M H. 2014. Deblurring low-light images with light streaks//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Columbus, USA: IEEE: 3382-3389 [DOI: 10.1109/CVPR.2014.432]
- Hu Z, Yuan L, Lin S and Yang M H. 2016. Image deblurring using smartphone inertial sensors//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA: IEEE: 1855-1864 [DOI: 10.1109/CVPR.2016.205]
- Jeon H G, Lee J Y, Han Y, Kim S J and Kweon I S. 2017. Multi-image deblurring using complementary sets of fluttering patterns. *IEEE Transactions on Image Processing*, 26(5): 2311-2326 [DOI: 10.1109/TIP.2017.2675202]
- Jiang Z, Zhang Y, Zou D Q, Ren J, Lv J C and Liu Y B. 2020. Learning event-based motion deblurring//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, USA: IEEE: 3320-3329 [DOI: 10.1109/CVPR42600.2020.00338]
- Kaufman A and Fattal R. 2020. Deblurring using analysis-synthesis networks pair//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, USA: IEEE: 5811-5820 [DOI: 10.1109/CVPR42600.2020.00585]
- Kim K, Lee S and Cho S. 2022. MSSNet: multi-scale-stage network for single image deblurring//Proceedings of Computer Vision — ECCV 2022 Workshops. Tel Aviv, Israel: Springer [DOI: 10.1007/978-3-031-25063-7_32]
- Koh J, Lee J and Yoon S. 2021. Single-image deblurring with neural networks: a comparative survey. *Computer Vision and Image Understanding*, 203: #103134 [DOI: 10.1016/j.cviu.2020.103134]
- Köhler R, Hirsch M, Mohler B, Schölkopf B and Harmeling S. 2012. Recording and playback of camera shake: benchmarking blind deconvolution with a real-world database//Proceedings of the 12th European Conference on Computer Vision on Computer Vision — ECCV 2012. Florence, Italy: Springer: 27-40 [DOI: 10.1007/978-3-642-33786-4_3]
- Kong L S, Dong J X, Ge J J, Li M Q and Pan J S. 2023. Efficient frequency domain-based Transformers for high-quality image deblurring//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver, Canada: IEEE: 5886-5895 [DOI: 10.1109/CVPR52729.2023.00570]
- Krishnan D and Fergus R. 2009. Fast image deconvolution using hyper-Laplacian priors [EB/OL]. [2023-11-20].
https://proceedings.neurips.cc/paper_files/paper/2009/file/3dd48ab31d016ffcbf3314df2b3cb9ce-Paper.pdf
- Kupyn O, Budzan V, Mykhailych M, Mishkin D and Matas J. 2018. Deblurgan: blind motion deblurring using conditional adversarial networks//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Lake City, USA: IEEE: 8183-8192 [DOI: 10.1109/CVPR.2018.00854]
- Kupyn O, Martyniuk T, Wu J R and Wang Z Y. 2019. Deblurgan-v2: deblurring (orders-of-magnitude) faster and better//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea (South): IEEE: 8878-8887 [DOI: 10.1109/ICCV.2019.00897]
- Lai W S, Huang J B, Hu Z, Ahuja N and Yang M H. 2016. A comparative study for single image blind deblurring//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA: IEEE: 1701-1709 [DOI: 10.1109/CVPR.2016.188]
- Lai W S, Shih Y C, Chu L C, Wu X T, Tsai S F, Krainin M, Sun D Q and Liang C K. 2022. Face deblurring using dual camera fusion on mobile phones. *ACM Transactions on Graphics*, 41(4): #148 [DOI: 10.1145/3528223.3530131]
- Levin A. 2006. Blind motion deblurring using image statistics [EB/OL]. [2023-11-20].
https://proceedings.neurips.cc/paper_files/paper/2006/file/2bb0502c80b7432eee4c5847a5fd077b-Paper.pdf
- Levin A, Fergus R, Durand F and Freeman W T. 2007. Image and depth from a conventional camera with a coded aperture. *ACM Transactions on Graphics*, 26(3): #70-es [DOI: 10.1145/1276377.1276464]
- Levin A, Weiss Y, Durand F and Freeman W T. 2009. Understanding and evaluating blind deconvolution algorithms//Proceedings of 2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami, USA: IEEE: 1964-1971 [DOI: 10.1109/CVPR.2009.5206815]
- Levin A, Weiss Y, Durand F and Freeman W T. 2011. Understanding

- blind deconvolution algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(12): 2354-2367 [DOI: 10.1109/TPAMI.2011.148]
- Li B Y, Liu X, Hu P, Wu Z Q, Lv J C and Peng X. 2022a. All-in-one image restoration for unknown corruption//*Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. New Orleans, USA: IEEE: 17452-17462 [DOI: 10.1109/CVPR52688.2022.01693]
- Li D S, Zhang Y, Cheung K C, Wang X G, Qin H W and Li H S. 2022b. Learning degradation representations for image deblurring//*Proceedings of the 17th European Conference on Computer Vision — ECCV 2022*. Tel Aviv, Israel: Springer: 736-753 [DOI: 10.1007/978-3-031-19797-0_42]
- Li L, Pan J S, Lai W S, Gao C X, Sang N and Yang M H. 2020a. Dynamic scene deblurring by depth guided model. *IEEE Transactions on Image Processing*, 29: 5273-5288 [DOI: 10.1109/TIP.2020.2980173]
- Li R T, Tan R T and Cheong L F. 2020b. All in one bad weather removal using architectural search//*Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle, USA: IEEE: 3172-3182 [DOI: 10.1109/CVPR42600.2020.00324]
- Li W, Zhang J and Dai Q H. 2011. Exploring aligned complementary image pair for blind motion deblurring//*Proceedings of CVPR 2011*. Colorado Springs, USA: IEEE: 273-280 [DOI: 10.1109/CVPR.2011.5995351]
- Li W, Zhang J and Dai Q H. 2013. Robust blind motion deblurring using near-infrared flash image. *Journal of Visual Communication and Image Representation*, 24(8): 1394-1413 [DOI: 10.1016/j.jvcir.2013.09.008]
- Li X, Jin X, Lin J X, Liu S, Wu Y J, Yu T, Zhou W and Chen Z B. 2020c. Learning disentangled feature representation for hybrid-distorted image restoration//*Proceedings of the 16th European Conference on Computer Vision — ECCV 2020*. Glasgow, UK: Springer: 313-329 [DOI: 10.1007/978-3-030-58526-6_19]
- Li X, Li B C, Jin X, Lan C L and Chen Z B. 2023. Learning distortion invariant representation for image restoration from a causality perspective//*Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Vancouver, Canada: IEEE: 1714-1724 [DOI: 10.1109/CVPR52729.2023.00171]
- Li X M, Liu M, Ye Y T, Zuo W M, Lin L and Yang R G. 2018. Learning warped guidance for blind face restoration//*Proceedings of the 15th European Conference on Computer Vision — ECCV 2018*. Munich, Germany: Springer: 272-289 [DOI: 10.1007/978-3-030-01261-8_17]
- Li Y W, Pan J S, Luo Y and Lu J W. 2022c. Deep ranking exemplar-based dynamic scene deblurring. *IEEE Transactions on Image Processing*, 31: 2245-2256 [DOI: 10.1109/TIP.2022.3142518]
- Liang Z W, Zhang D Y and Shao J. 2019. Jointly solving deblurring and super-resolution problems with dual supervised network//*Proceedings of 2019 IEEE International Conference on Multimedia and Expo*. Shanghai, China: IEEE: 790-795 [DOI: 10.1109/ICME.2019.00141]
- Liu R S, Ma L, Zhang J A, Fan X and Luo Z X. 2021. Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement//*Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Nashville, USA: IEEE: 10556-10565 [DOI: 10.1109/CVPR46437.2021.01042]
- Lu B, Chen J C and Chellappa R. 2019. Unsupervised domain-specific deblurring via disentangled representations//*Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Beach, USA: IEEE: 10225-10234 [DOI: 10.1109/CVPR.2019.01047]
- Nah S, Kim T H and Lee K M. 2017. Deep multi-scale convolutional neural network for dynamic scene deblurring//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, USA: IEEE: 3883-3891 [DOI: 10.1109/CVPR.2017.35]
- Pan J S, Hu Z, Su Z X, Lee H Y and Yang M H. 2016a. Soft-segmentation guided object motion deblurring//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, USA: IEEE: 459-468 [DOI: 10.1109/CVPR.2016.56]
- Pan J S, Hu Z, Su Z X and Yang M H. 2014. Deblurring text images via L0-regularized intensity and gradient prior//*Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition*. Columbus, USA: IEEE: 2901-2908 [DOI: 10.1109/CVPR.2014.371]
- Pan J S, Ren W Q, Hu Z and Yang M H. 2019a. Learning to deblur images with exemplars. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(6): 1412-1425 [DOI: 10.1109/TPAMI.2018.2832125]
- Pan J S, Sun D Q, Pfister H and Yang M H. 2016b. Blind image deblurring using dark channel prior//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, USA: IEEE: 1628-1636 [DOI: 10.1109/CVPR.2016.180]
- Pan L Y, Scheerlinck C, Yu X, Hartley R, Liu M and Dai Y C. 2019b. Bringing a blurry frame alive at high framerate with an event camera//*Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Long Beach, USA: IEEE: 6820-6829 [DOI: 10.1109/CVPR.2019.00698]
- Park D, Kang D U, Kim J and Chun S Y. 2020. Multi-temporal recurrent neural networks for progressive non-uniform single image deblurring with incremental temporal training//*Proceedings of the 16th European Conference on Computer Vision — ECCV 2020*. Glasgow, UK: Springer: 327-343 [DOI: 10.1007/978-3-030-58539-6_20]

- Park D, Lee B H and Chun S Y. 2023. All-in-one image restoration for unknown degradations using adaptive discriminative filters for specific degradations//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver, Canada: IEEE: 5815-5824 [DOI: 10.1109/CVPR52729.2023.00563]
- Qiu F, Hou F, Yuan Y and Wang W C. 2019. Blind image deblurring with reinforced use of edges. *Journal of Image and Graphics*, 24(6): 847-858 (邱枫, 侯飞, 袁野, 王文成. 2019. 加强边缘感知的盲去模糊算法. *中国图象图形学报*, 24(6): 847-858) [DOI: 10.11834/jig.180543]
- Rav-Acha A and Peleg S. 2000. Restoration of multiple images with motion blur in different directions//Proceedings of the 5th IEEE Workshop on Applications of Computer Vision. Palm Springs, USA: IEEE: 22-28 [DOI: 10.1109/WACV.2000.895398]
- Rav-Acha A and Peleg S. 2005. Two motion-blurred images are better than one. *Pattern recognition letters*, 26(3): 311-317 [DOI: 10.1016/j.patrec.2004.10.017]
- Ren D W, Zhang K, Wang Q L, Hu Q H and Zuo W M. 2020. Neural blind deconvolution using deep priors//Proceedings of 2020 IEEE Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 3338-3347 [DOI: 10.1109/CVPR42600.2020.00340]
- Ren W Q, Pan J S, Cao X C and Yang M H. 2017. Video deblurring via semantic segmentation and pixel-wise non-linear kernel//Proceedings of 2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE: 1077-1085 [DOI: 10.1109/ICCV.2017.123]
- Richardson W H. 1972. Bayesian-based iterative method of image restoration. *Journal of the Optical Society of America*, 62(1): 55-59
- Rim J, Kim G, Kim J, Lee J, Lee S and Cho S. 2022. Realistic blur synthesis for learning image deblurring//Proceedings of the 17th European Conference on Computer Vision — ECCV 2022. Tel Aviv, Israel: Springer: 487-503 [DOI: 10.1007/978-3-031-20071-7_29]
- Rim J, Lee H, Won J and Cho S. 2020. Real-world blur dataset for learning and benchmarking deblurring algorithms//Proceedings of the 16th European Conference on Computer Vision — ECCV 2020. Glasgow, UK: Springer: 184-201 [DOI: 10.1007/978-3-030-58595-2_12]
- Schmidt U, Schelten K and Roth S. 2011. Bayesian deblurring with integrated noise estimation//Proceedings of CVPR 2011. Colorado Springs, USA: IEEE: 2625-2632 [DOI: 10.1109/CVPR.2011.5995653]
- Shan Q, Jia J Y and Agarwala A. 2008. High-quality motion deblurring from a single image. *ACM Transactions on Graphics*, 27(3): 1-10 [DOI: 10.1145/1360612.1360672]
- Shang W, Ren D W, Zou D Q, Ren J S, Luo P and Zuo W M. 2021. Bringing events into video deblurring with non-consecutively blurry frames//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal, Canada: IEEE: 4531-4540 [DOI: 10.1109/ICCV48922.2021.00449]
- Shen Z Y, Lai W S, Xu T F, Kautz J and Yang M H. 2020. Exploiting semantics for face image deblurring. *International Journal of Computer Vision*, 128(7): 1829-1846 [DOI: 10.1007/s11263-019-01288-9]
- Shen Z Y, Lai W S, Xu T F, Kautz J and Yang M H. 2018. Deep semantic face deblurring//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 8260-8269 [DOI: 10.1109/CVPR.2018.00862]
- Shin C J, Lee T B and Heo Y S. 2021. Dual image deblurring using deep image prior. *Electronics*, 10(17): #2045 [DOI: 10.3390/electronics10172045]
- Song J M, Meng C L and Ermon S. 2020. Denoising diffusion implicit models//Proceedings of 2020 International Conference on Learning Representations [EB/OL]. [2023-11-20]. <http://arxiv.org/pdf/2010.02502.pdf>
- Sun J, Cao W F, Xu Z B and Ponce J. 2015. Learning a convolutional neural network for non-uniform motion blur removal//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, USA: IEEE: 769-777 [DOI: 10.1109/CVPR.2015.7298677]
- Sun L, Sakaridis C, Liang J Y, Jiang Q, Yang K L, Sun P, Ye Y Z, Wang K W and Van Gool L. 2022. Event-based fusion for motion deblurring with cross-modal attention//Proceedings of the 17th European Conference on Computer Vision — ECCV 2022. Tel Aviv, Israel: Springer: 412-428 [DOI: 10.1007/978-3-031-19797-0_24]
- Tao X, Gao H Y, Shen X Y, Wang J and Jia J Y. 2018. Scale-recurrent network for deep image deblurring//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 8174-8182 [DOI: 10.1109/CVPR.2018.00853]
- Tsai F J, Peng Y T, Lin Y Y, Tsai C C and Lin C W. 2022a. Strip-former: strip Transformer for fast image deblurring//Proceedings of the 17th European Conference on Computer Vision — ECCV 2022. Tel Aviv, Israel: Springer: 146-162 [DOI: 10.1007/978-3-031-19800-7_9]
- Tsai F J, Peng Y T, Tsai C C, Lin Y Y and Lin C W. 2022b. BANet: a blur-aware attention network for dynamic scene deblurring. *IEEE Transactions on Image Processing*, 31: 6789-6799 [DOI: 10.1109/TIP.2022.3216216]
- Wang B S, He J W, Yu L, Xia G S and Yang W. 2020a. Event enhanced high-quality image recovery//Proceedings of the 16th European Conference on Computer Vision — ECCV 2020. Glasgow, UK: Springer: 155-171 [DOI: 10.1007/978-3-030-58601-0_10]
- Wang L, Li D, Zhu Y S, Tian L and Shan Y. 2020b. Dual super-resolution learning for semantic segmentation//Proceedings of 2020

- IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, USA: IEEE: 3774-3783 [DOI: 10.1109/CVPR42600.2020.00383]
- Wang L G, Wang Y Q, Dong X Y, Xu Q Y, Yang J G, An W and Guo Y L. 2021. Unsupervised degradation representation learning for blind super-resolution//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, USA: IEEE: 10576-10585 [DOI: 10.1109/CVPR46437.2021.01044]
- Wang R X and Tao D C. 2014. Recent progress in image deblurring [EB/OL]. [2023-11-20]. <https://doi.org/10.48550/arXiv.1409.6838.pdf>
- Wang X T, Yu K, Dong C and Loy C C. 2018. Recovering realistic texture in image super-resolution by deep spatial feature transform//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE [DOI: 10.1109/CVPR.2018.00070]
- Wang Z D, Cun X D, Bao J M, Zhou W G, Liu J Z and Li H Q. 2022. Uformer: a general U-shaped Transformer for image restoration//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, USA: IEEE: 17683-17693 [DOI: 10.1109/CVPR52688.2022.01716]
- Whang J, Delbracio M, Talebi H, Saharia C, Dimakis A G and Milanfar P. 2022. Deblurring via stochastic refinement//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, USA: IEEE: 16272-16282 [DOI: 10.1109/CVPR52688.2022.01581]
- Wiener N. 1964. Extrapolation, Interpolation, and Smoothing of Stationary Time Series; with Engineering Applications. Cambridge, USA: the MIT Press
- Wu D, Zhao H T and Zheng S B. 2020. Motion deblurring method based on DenseNets. *Journal of Image and Graphics*, 25 (5): 890-899 (吴迪, 赵洪田, 郑世宝. 2020. 密集连接卷积网络图像去模糊. *中国图象图形学报*, 25 (5): 890-899) [DOI: 10.11834/jig.190400]
- Xia B, Tian Y P, Zhang Y L, Hang Y C, Yang W M and Liao Q M. 2023. Meta-learning-based degradation representation for blind super-resolution. *IEEE Transactions on Image Processing*, 32: 3383-3396 [DOI: 10.1109/TIP.2023.3283922]
- Xiang X G, Wei H and Pan J S. 2020. Deep video deblurring using sharpness features from exemplars. *IEEE Transactions on Image Processing*, 29: 8976-8987 [DOI: 10.1109/TIP.2020.3023534]
- Xu F, Yu L, Wang B S, Yang W, Xia G S, Jia X, Qiao Z D and Liu J Z. 2021. Motion deblurring with real events//Proceedings of 2021 IEEE International Conference on Computer Vision. Montreal, Canada: IEEE: 2583-2592 [DOI: 10.1109/ICCV48922.2021.00258]
- Xu L and Jia J Y. 2012. Depth-aware motion deblurring//Proceedings of 2012 IEEE International Conference on Computational Photography (ICCP). Seattle, USA: IEEE: 1-8 [DOI: 10.1109/ICCP.2012.6215220]
- Xu X Y, Sun D Q, Pan J S, Zhang Y J, Pfister H and Yang M H. 2017. Learning to super-resolve blurry face and text images//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE: 251-260 [DOI: 10.1109/ICCV.2017.36]
- Yang X, Wang X Q, Wang N N and Gao X B. 2022. SRDN: a unified super-resolution and motion deblurring network for space image restoration. *IEEE Transactions on Geoscience and Remote Sensing*, 60: #5614411 [DOI: 10.1109/TGRS.2021.3131264]
- Yuan L, Sun J, Quan L and Shum H Y. 2007. Image deblurring with blurred/noisy image pairs. *ACM Transactions on Graphics*, 26(3): #1-es [DOI: 10.1145/1276377.1276379]
- Yue T, Sun M T, Zhang Z Y, Suo J L and Dai Q H. 2014. Deblur a blurred RGB image with a sharp NIR image through local linear mapping//Proceedings of 2014 IEEE International Conference on Multimedia and Expo. Chengdu, China: IEEE: 1-6 [DOI: 10.1109/ICME.2014.6890310]
- Yun J U, Jo B and Park I K. 2020. Joint face super-resolution and deblurring using generative adversarial network. *IEEE Access*, 8: 159661-159671 [DOI: 10.1109/ACCESS.2020.3020729]
- Zamir S W, Arora A, Khan S, Hayat M, Khan F S and Yang M H. 2022. Restormer: efficient Transformer for high-resolution image restoration//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, USA: IEEE: 5718-5729 [DOI: 10.1109/CVPR52688.2022.00564]
- Zamir S W, Arora A, Khan S, Hayat M, Khan F S, Yang M H and Shao L. 2021. Multi-stage progressive image restoration//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, USA: IEEE: 14821-14831 [DOI: 10.1109/CVPR46437.2021.01458]
- Zhang D Y, Liang Z W and Shao J. 2020a. Joint image deblurring and super-resolution with attention dual supervised network. *Neurocomputing*, 412: 187-196 [DOI: 10.1016/j.neucom.2020.05.069]
- Zhang H C, Yang J C, Zhang Y N, Nasrabadi N M and Huang T S. 2011. Close the loop: joint blind image restoration and recognition with sparse representation prior//Proceedings of 2011 International Conference on Computer Vision. Barcelona, Spain: IEEE: 770-777 [DOI: 10.1109/ICCV.2011.6126315]
- Zhang H C and Zhang Y N. 2009. Sparse representation based iterative incremental image deblurring//Proceedings of the 16th IEEE International Conference on Image Processing. Cairo, Egypt: IEEE: 1293-1296 [DOI: 10.1109/ICIP.2009.5413591]
- Zhang H G, Dai Y C, Li H D and Koniusz P. 2019. Deep stacked hierarchical multi-patch network for image deblurring//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA: IEEE: 5978-5986 [DOI: 10.1109/CVPR.2019.00613]
- Zhang J W, Pan J S, Ren J, Song Y B, Bao L C, Lau R W H and Yang M H. 2018a. Dynamic scene deblurring using spatially variant

- recurrent neural networks//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA; IEEE: 2521-2529 [DOI: 10.1109/CVPR.2018.00267]
- Zhang K H, Luo W H, Zhong Y R, Ma L, Stenger B, Liu W and Li H D. 2020b. Deblurring by realistic blurring//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, USA; IEEE: 2737-2746 [DOI: 10.1109/CVPR42600.2020.00281]
- Zhang K H, Ren W Q, Luo W H, Lai W S, Stenger B, Yang M H and Li H D. 2022. Deep image deblurring: a survey. International Journal of Computer Vision, 130 (9) : 2103-2130 [DOI: 10.1007/s11263-022-01633-5]
- Zhang K H, Wang T, Luo W H, Ren W Q, Stenger B, Liu W, Li H D and Yang M H. 2024. MC-Blur: a comprehensive benchmark for image deblurring. IEEE Transactions on Circuits and Systems for Video Technology, 34 (5) : 3755-3767 [DOI: 10.1109/TCSVT.2023.3319330]
- Zhang X and Yu L. 2022. Unifying motion deblurring and frame interpolation with events//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, USA; IEEE: 17765-17774 [DOI: 10.1109/CVPR52688.2022.01724]
- Zhang X Y, Dong H, Hu Z, Lai W S, Wang F and Yang M H. 2018b. Gated fusion network for joint image deblurring and super-resolution//Proceedings of 29th British Machine Vision Conference (BMVC). Newcastle, United Kingdom. #153
- Zhang Y N, Liu Y M, Li Q, Wang J Z, Qi M, Sun H, Xu H and Kong J. 2021. A lightweight fusion distillation network for image deblurring and deraining. Sensors, 21 (16) : #5312 [DOI: 10.3390/s21165312]
- Zhao S Y, Zhang Z, Hong R C, Xu M L, Yang Y and Wang M. 2022. FCL-GAN: a lightweight and real-time baseline for unsupervised blind image deblurring//Proceedings of the 30th ACM International Conference on Multimedia. Lisboa, Portugal; Association for Computing Machinery: 6220-6229 [DOI: 10.1145/3503161.3548113]
- Zhou S C, Li C Y and Loy C C. 2022. LEDNet: joint low-light enhancement and deblurring in the dark//Proceedings of the 17th European Conference on Computer Vision — ECCV 2022. Tel Aviv, Israel; Springer 573-589 [DOI: 10.1007/978-3-031-20068-7_33]
- Zhuang K, Li Q, Yuan Y and Wang Q. 2024. Multi-domain adaptation for motion deblurring. IEEE Transactions on Multimedia, 26: 3676-3688 [DOI: 10.1109/TMM.2023.3314154]

作者简介

王珮,女,博士研究生,主要研究方向为图像去模糊和深度学习。E-mail: wangpei23@mail.nwpu.edu.cn

朱宇,通信作者,男,副研究员,主要研究方向为图像超分辨率重建。E-mail: yuzhu@nwpu.edu.cn

闫庆森,男,教授,主要研究方向为人工智能和机器学习。

E-mail: qingsenyan@nwpu.edu.cn

孙瑾秋,女,教授,主要研究方向为光学目标高精度探测与清晰化重建及特性分析。E-mail: sunjinqiu@nwpu.edu.cn

张艳宁,女,教授,主要研究方向为图像处理、模式识别、计算机视觉与智能信息处理。E-mail: ynzhang@nwpu.edu.cn