

基于对比学习的大气湍流退化图像复原方法

苗壮, 崔浩然, 张启阳, 王家宝, 李阳

引用本文

苗壮, 崔浩然, 张启阳, 王家宝, 李阳. [基于对比学习的大气湍流退化图像复原方法](#)[J]. 计算机科学, 2025, 52(5): 171-178.

MIAO Zhuang, CUI Haoran, ZHANG Qiyang, WANG Jiabao, LI Yang. [Restoration of Atmospheric Turbulence-degraded Images Based on Contrastive Learning](#) [J]. Computer Science, 2025, 52(5): 171-178.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于“隐形面具”的可逆人脸隐私保护方法](#)

Reversible Facial Privacy Protection Method Based on “Invisible Masks”

计算机科学, 2025, 52(5): 384-391. <https://doi.org/10.11896/jsjcx.241100066>

[基于深度位置感知Transformer的低重叠点云配准](#)

Low Overlap Point Cloud Registration Method Based on Deep Position-aware Transformer

计算机科学, 2025, 52(5): 199-211. <https://doi.org/10.11896/jsjcx.240400172>

[基于特征网络对比学习的图协同过滤模型研究](#)

Study on Graph Collaborative Filtering Model Based on FeatureNet Contrastive Learning

计算机科学, 2025, 52(5): 139-148. <https://doi.org/10.11896/jsjcx.240200078>

[面向深度学习编译器TVM的算子融合优化](#)

Operator Fusion Optimization for Deep Learning Compiler TVM

计算机科学, 2025, 52(5): 58-66. <https://doi.org/10.11896/jsjcx.240100018>

[基于多尺度融合注意力的多视角文档图像篡改检测与定位](#)

Multi-view and Multi-scale Fusion Attention Network for Document Image Forgery Localization

计算机科学, 2025, 52(4): 327-335. <https://doi.org/10.11896/jsjcx.240100142>

基于对比学习的大气湍流退化图像复原方法

苗 壮 崔浩然 张启阳 王家宝 李 阳

陆军工程大学指挥控制工程学院 南京 210007

(emiao_beyond@163.com)

摘 要 大气湍流引起的图像退化严重影响了目标检测和图像识别等计算机视觉下游任务的性能。现有基于深度学习的大气湍流退化图像复原模型虽然取得了较好的效果,但未充分利用湍流效应的特征信息。为了获得更好的复原效果,提出了一种基于对比度学习的大气湍流退化图像复原方法。针对大气湍流退化引起的模糊与扭曲,设计了湍流缓解块。该块融合了基于Transformer的通道信息交互模块与基于CNN的空间信息交互模块,在全局和局部层面上抑制湍流对图像的干扰。同时,引入对比学习,将清晰图像和大气湍流退化图像视为正样本和负样本,在特征空间中拉近大气湍流复原网络的输出与正样本的距离,推远与负样本的距离,更有效地进行特征提取和图像复原。在Helen合成测试集和Places合成测试集上,所提方法分别达到了26.78 dB,22.42 dB的PSNR和0.7909,0.6820的SSIM,与现有的5种方法相比达到了最佳效果,更适用于提升大气湍流退化图像的质量。

关键词:深度学习;图像复原;大气湍流;对比学习;特征提取

中图分类号 TP391

Restoration of Atmospheric Turbulence-degraded Images Based on Contrastive Learning

MIAO Zhuang, CUI Haoran, ZHANG Qiyang, WANG Jiabao and LI Yang

Command and Control Engineering College, Army Engineering University of PLA, Nanjing 210007, China

Abstract Image degradation caused by atmospheric turbulence seriously affects the performance of downstream computer vision tasks such as object detection and image recognition. Existing deep learning-based image restoration models for atmospheric turbulence degradation have achieved good performance, but have not fully utilized the feature information of the turbulence effect. To improve restoration results, a method for restoring of atmospheric turbulence-degraded images based on contrastive learning is proposed. Aiming at the blurring and distortion caused by atmospheric turbulence degradation, a turbulence mitigation block is designed, which integrates a Transformer-based channel information interaction module and a CNN-based spatial information interaction module to suppress the turbulence interference to the image at both global and local levels. At the same time, contrastive learning is introduced to consider the clear image and the degraded image of atmospheric turbulence as positive and negative samples, to pull the output of the atmospheric turbulence restoration network closer to the positive samples and push it farther away from the negative samples in the feature space, so that feature extraction and image restoration can be performed more efficiently. The proposed method achieves 26.78 dB and 22.42 dB PSNR and 0.7909 and 0.6820 SSIM on the synthetic Helen dataset and synthetic Places dataset, respectively, which is the best result compared with the existing five methods, and it is suitable for improving the quality of atmospheric turbulence degradation images.

Keywords Deep learning, Image restoration, Atmospheric turbulence, Contrastive learning, Feature extraction

1 引言

大气湍流的干扰是成像系统捕获到低质量图像的主要原因之一,其引起的退化主要体现在图像内容的模糊与扭曲^[1]。受温度、距离等因素影响,大气环境中产生不均匀折射率,导致图像失真退化,严重影响了目标检测、人脸图像识别等诸多计算机下游任务的性能。因此,从湍流

退化图像复原至清晰图像十分重要。

现有的大气湍流图像复原方法分为传统光学处理方法与基于深度学习的方法。其中,传统算法^[2-5]通常都是基于“幸运成像”^[6]的原理,利用湍流的特性,在多个图像帧中选择受退化影响最小的一部分组合成一张干净的图像。尽管传统方法可以很好地复原退化后的图像,但是它们需要多帧的输入,会消耗大量的计算时间。在长曝光运动场景下,拍摄抖动与

到稿日期:2024-02-04 返修日期:2024-06-21

基金项目:江苏省自然科学基金(BK20200581)

This work was supported by the Natural Science Foundation of Jiangsu Province, China(BK20200581).

通信作者:王家宝(jiabao_1108@163.com)

物体运动引起图像间的不匹配,从而导致基于“幸运成像”的传统方法无法准确选择合成干净图像。随着深度学习的兴起,用于单帧的大气湍流退化图像复原算法展现出了很好的效果,具体分为基于 CNN 卷积神经网络^[7]、基于 Transformer^[8-9]和基于 Diffusion^[10]的方法。基于卷积神经网络的方法利用估计出的先验信息来消除大气湍流所引起的图像退化现象。基于 Transformer 与基于 Diffusion 的方法充分考虑了大气湍流的物理特性,模拟湍流形成的过程,并且通过大数据学习来复原由湍流引起的图像模糊与几何扭曲。相比于传统方法,基于深度学习的方法在客观性能指标和真实视觉感受上得到了很大提升,但是并没有充分地利用大气湍流退化图像的特征,限制了复原的性能。主要原因有以下几点:1)基于 CNN 卷积神经网络的方法可以通过卷积运算高效地提取湍流退化图像的局部特征,但是有限的感受野使得全局信息难以获取;2)基于 Transformer 的方法通过“自注意力”^[11]制获取全局信息,但也产生了较高的复杂度,不能有效地提取局部细节特征;3)现有的方法仅使用成对的干净图像与大气湍流退化图像指导网络的训练,而忽视了图像在潜在特征空间^[12]中所提供的信息。

针对以上问题,提出了一种基于对比学习的大气湍流退化图像复原方法。该方法结合 CNN 与 Transformer 的优点设计了湍流缓解决,在通道与空间中捕获全局-局部特征信息,帮助网络复原湍流退化后的图像。同时,引入对比学习,将清晰干净图像与湍流退化图像视为正样本与负样本,利用预训练好的特征提取器提取正负样本与复原后图像的特征,在潜在特征空间中拉近复原后图像与正样本的距离,推远与负样本的距离,进一步充分利用大气湍流退化图像与干净图像所提供的特征信息,提升复原后图像的视觉质量。综上所述,本文的贡献如下:

1)提出了一种基于对比学习的大气湍流退化图像复原方法,该方法结合 CNN 与 Transformer 设计了湍流缓解决,在通道与空间层面捕获全局-局部信息,抑制了大气湍流退化导致的图像模糊与扭曲现象。

2)引入对比学习,充分利用大气湍流退化图像和干净图像中的特征信息,帮助网络复原更清晰的图像。

3)在 Helen 与 Places 合成测试集上验证所提模型,结果表明,相比当前的 5 种方法,所提模型均取得了最佳的效果。

2 相关工作

2.1 图像复原网络

图像复原任务旨在将受外界干扰(如雨、雪、雾、噪声、湍流等)而退化的图像复原至清晰干净的图像。在深度学习发展初期,出现了许多对于特定任务所设计的深度神经网络。Zhang 等^[13]设计了基于 CNN 的残差通道注意力网络,以实现图像的超分辨率。Yang 等^[14]通过一种多任务的架构学习雨条纹图来实现图像去雨。Cai 等^[15]设计了一个端到端的图像去雾网络,利用大气散射模型帮助恢复无雾图像。Yasarla 等^[16]通过对湍流退化人脸图像中模糊与几何失真进行先验估计,利用估计量消除湍流退化图像失真。近年来,越来越多的统一架构出现,并被用于处理多个任务。Zamir 等^[17]提出

多阶段架构来实现对图像的去雨、去模糊、去噪。Wang 等^[18]设计了高效的 U 形 Transformer 进行去噪声以及运动去模糊。Chen 等^[19]通过微调预训练模型,可以有效地处理不同的图像复原任务。Zamir 等^[20]提出了高效的基于通道的自注意力用于去雨、去模糊与去噪。针对计算复杂的图像复原架构,Chen 等^[21]提出了一个简单的基线,简化了激活函数与注意力机制,在图像复原任务中取得了显著成功。Chen 等^[41]在多通道视觉注意力机制的基础上构建了编码网络,将模糊图像输入到编码网络中,对图像展开复原处理。

对于大气湍流退化后的图像复原,基于“幸运成像”的传统方法^[2]在静态场景中取得了很好的效果,但是有着很高的计算代价,很难适用于一般场景。随着图像复原架构的发展,Mao 等^[8]基于湍流的物理特性设计了高效的图像复原架构。Zhang 等^[9]设计的两阶段的 U 形多帧湍流恢复网络,在真实场景中有很好的泛化性。Jaiswal 等^[10]提出的物理集成复原网络,利用 Diffusion 在恢复精度与视觉感知上取得了很好的效果。但是,现有方法并未深入挖掘图像在特征空间中所提供的信息,干净图像与湍流退化图像特征信息内容未得到充分提取与利用。

2.2 对比学习

对比学习被广泛应用在自监督的表示学习中^[22-24],其主要思想在于拉近给定锚点与正样本在表示空间中的距离,同时推远与负样本的距离。最近,一些研究证明了对比学习在图像复原任务上的有效性。其中,Liang 等^[25]通过语义指导与对比学习,在暗光图像增强任务上取得了很好的效果。Wang 等^[26]联合语义信息与对比正则的方法,以得到更好的去雾图像。Wu 等^[27]采用对比学习框架,在超分辨任务上获得了高质量的图像。Ye 等^[28]提出了非局部对比学习方法,用于无监督图像去雨。因此,对比学习在大气湍流退化图像复原任务上具有巨大潜力。

3 本文方法

大气湍流引起的退化主要表现为图像内容的模糊与扭曲。对于清晰的干净图像 I_c ,退化过程^[29]通常被建模为:

$$I_t = h(I_c) + n \quad (1)$$

其中, I_t 表示湍流干扰后退化的图像; n 表示随机高斯噪声; $h(\cdot)$ 表示大气湍流退化函数,具体由几何位移函数 g 与模糊函数 b 组成^[30]。故式(1)可重写为:

$$I_t = b \circ g(I_c) + n \quad (2)$$

由上述大气湍流退化过程可知,想得到复原后干净的图像,需要用几何位移函数 g 与模糊函数 b 的逆过程来求解,即 $I_c = g^{-1}(b^{-1}(I_t - n))$,其中模糊与位移的随机性导致逆函数 g^{-1} 与 b^{-1} 难以求解。因此,为了拟合湍流退化中模糊和几何位移的逆函数 b^{-1} 和 g^{-1} ,分别设计了通道信息交互模块(Channel Information Interaction Module, CIIM)和空间信息交互模块(Spatial Information Interaction Module, SIIM),提出了一种基于对比学习的大气湍流退化图像复原方法。

3.1 整体框架和流程

如图 1 所示,基于对比学习的大气湍流退化图像复原方法包含图像复原网络(Image Restoration Network)及其相关

的对比正则化(Contrastive Regularization)训练策略。图 1 左侧展示了图像复原网络的框架和流程。给定一张湍流退化的图像 $I_t \in R^{H \times W \times 3}$, 通过 3×3 卷积提取浅层特征, 得到 $F_0 \in R^{H \times W \times C}$, 其中 $H \times W$ 代表空间维度上的高和宽, C 代表通道数。接着, 所提取的浅层特征 F_0 通过一个 U 形图像复原网络得到输出特征 $F_1 \in R^{H \times W \times 2C}$, 该 U 型复原网络各个层级分别包含 L_1, L_2, L_3, L_4 个 TMB(Turbulence Mitigation Block) 以帮助网络高效提取特征, 利用跳跃连接融合特征逐步复原更清晰的图像。网络中对特征的上采样与下采样使用“pixel-

shuffle”与“pixel-unshuffle”^[31] 操作。TMB 包含通道信息交互模块与空间信息交互模块, 用于抑制湍流退化带来的模糊与扭曲。最后, F_1 经过由 L_r 个 TMB 构成的精细(Refinement)模块和一个 3×3 卷积, 得到输出 $F_2 \in R^{H \times W \times 3}$, 将其与湍流退化图像相加得到最终输出 $I_r = I_t + F_2$ 。图 1 右侧展示了图像复原网络训练过程对比正则化策略, 该策略通过引入对比学习思想, 在特征空间中拉近复原图像与干净图像的距离, 推远复原图像与大气湍流退化图像的距离, 来指导网络学习更优的参数。

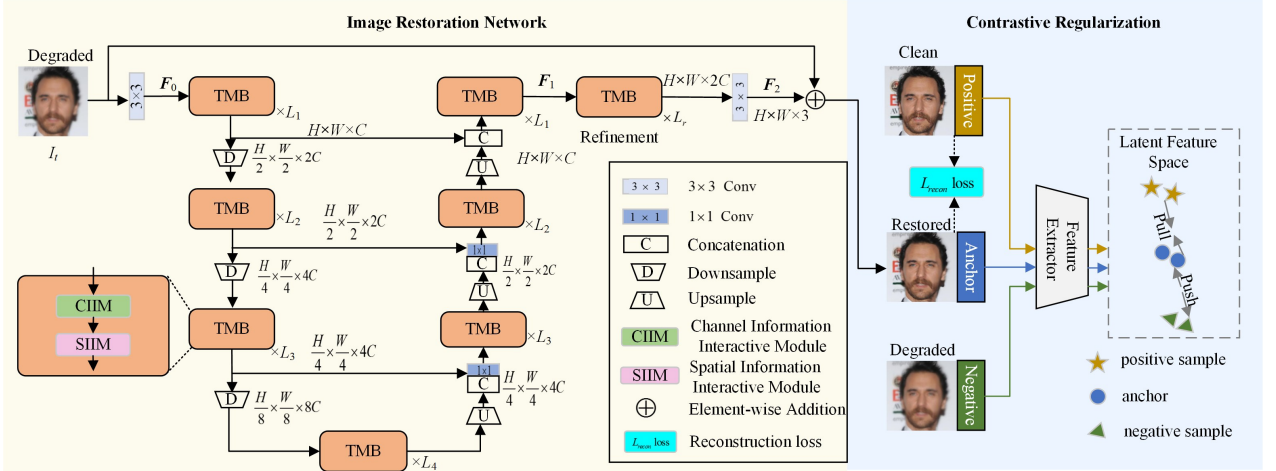


图 1 基于对比学习的大气湍流退化图像复原模型图

Fig. 1 Model architecture for restoration of atmospheric turbulence-degraded images based on contrastive learning

3.2 湍流缓解块

为了解决湍流退化问题的逆过程, 设计提出了湍流缓解块 TMB。TMB 包括基于 Transformer 的通道信息交互模块 CIIM 和基于 CNN 的空间信息交互模块 SIIM, 其中 CIIM 用于拟合模糊函数的逆过程 g^{-1} , SIIM 用于拟合几何位移函数的逆过程 b^{-1} 。通过结合 Transformer 与 CNN 的优势, TMB 具有全局与局部信息交互能力, 能实现更好的湍流退化图像复原效果。

3.2.1 通道信息交互模块 CIIM

CIIM 主要用于拟合湍流退化过程中的模糊函数逆操作: $CIIM(\cdot) \approx b^{-1}(\cdot)$ (3)

如图 2 所示, CIIM 由多深度卷积头转置注意力 MDTA 与门控-深度卷积前馈网络 GDFN 组成。MDTA 计算特征通道间的协方差来编码全局上下文信息, GDFN 利用门控单元来关注全局上下文中的更有效信息, 两者的结合具备对模糊信息的恢复能力^[20]。

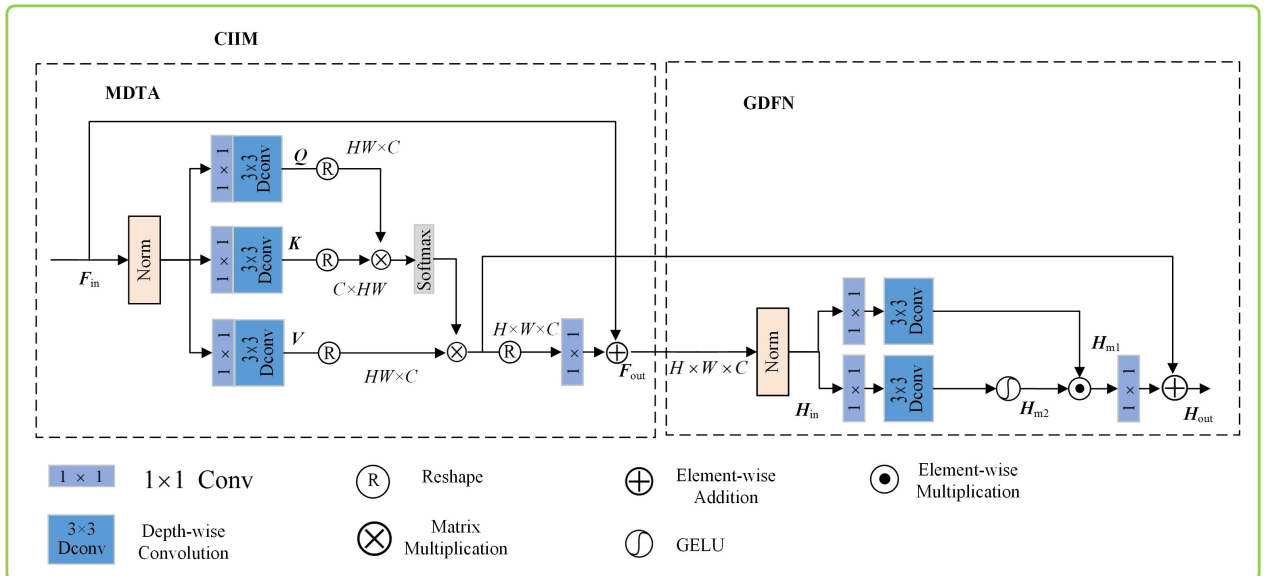


图 2 通道信息交互模块

Fig. 2 Channel information interaction module

在MDTA中,给定一个张量 $\mathbf{F}_{in} \in R^{H \times W \times C}$,经过层归一化输出 $\mathbf{F} \in R^{H \times W \times C}$,通过3个不同的 1×1 卷积 $Conv_{1 \times 1}(\cdot)$ 与 3×3 深度卷积 $Dconv_{3 \times 3}(\cdot)$ 分别生成:

$$\mathbf{Q} = Dconv_{3 \times 3}(Conv_{1 \times 1}(\mathbf{F}; \mathbf{W}_1^Q); \mathbf{W}_3^Q) \quad (4)$$

$$\mathbf{K} = Dconv_{3 \times 3}(Conv_{1 \times 1}(\mathbf{F}; \mathbf{W}_1^K); \mathbf{W}_3^K) \quad (5)$$

$$\mathbf{V} = Dconv_{3 \times 3}(Conv_{1 \times 1}(\mathbf{F}; \mathbf{W}_1^V); \mathbf{W}_3^V) \quad (6)$$

其中, \mathbf{W}_3 表示卷积核参数,经重塑后 $\mathbf{Q} \in R^{H \times W \times C}$, $\mathbf{K} \in R^{C \times HW}$, $\mathbf{V} \in R^{H \times W \times C}$ 。然后,计算注意力输出 $\mathbf{F}_A \in R^{H \times W \times C}$:

$$\mathbf{F}_A = \mathbf{V} \cdot \text{Softmax}(\mathbf{K} \cdot \mathbf{Q} / \alpha) \quad (7)$$

其中, \mathbf{K} 与 \mathbf{Q} 矩阵相乘以实现通道信息交互; α 为可学习的比例参数,用于控制 \mathbf{Q} 与 \mathbf{K} 点积大小。最后,重塑 $\mathbf{F}_A \in R^{H \times W \times C}$ 并通过 1×1 的卷积与 \mathbf{F} 相加得到输出 $\mathbf{F}_{out} \in R^{H \times W \times C}$:

$$\mathbf{F}_{out} = \mathbf{F} + Conv_{1 \times 1}(\mathbf{F}_A; \mathbf{W}) \quad (8)$$

其中, $Conv_{1 \times 1}(\cdot)$ 代表 1×1 的卷积, \mathbf{W} 为卷积参数。

GDFN中,利用门控单元与深度卷积控制通过的信息流,以获得更加有效的全局上下文信息。经过MDTA后的输出特征 $\mathbf{F}_{out} \in R^{H \times W \times C}$ 经归一化操作得到 $\mathbf{H}_{in} \in R^{H \times W \times C}$,上分支通过 1×1 卷积 $Conv_{1 \times 1}(\cdot)$ 与 3×3 深度卷积 $Dconv_{3 \times 3}(\cdot)$ 生成输出 $\mathbf{H}_{m1} \in R^{H \times W \times 2.66C}$:

$$\mathbf{H}_{m1} = Dconv_{3 \times 3}(Conv_{1 \times 1}(\mathbf{H}_{in}; \mathbf{W}_1^1); \mathbf{W}_3^1) \quad (9)$$

其中, \mathbf{W}_1^1 和 \mathbf{W}_3^1 表示卷积参数。下分支通过 1×1 的卷积 $Conv_{1 \times 1}(\cdot)$, 3×3 的深度卷积 $Dconv_{3 \times 3}(\cdot)$ 和非线性激活函数 $GELU(\cdot)$,得到输出 $\mathbf{H}_{m2} \in R^{H \times W \times 2.66C}$:

$$\mathbf{H}_{m2} = GELU(Dconv_{3 \times 3}(Conv_{1 \times 1}(\mathbf{H}_{in}; \mathbf{W}_1^2); \mathbf{W}_3^2)) \quad (10)$$

其中, \mathbf{W}_1^2 和 \mathbf{W}_3^2 表示卷积参数。最终, \mathbf{H}_{m1} 与 \mathbf{H}_{m2} 逐元素相乘,并经过 1×1 的卷积 $Conv_{1 \times 1}(\cdot)$ 得到GDFN模块的输出 $\mathbf{H}_{out} \in R^{H \times W \times C}$:

$$\mathbf{H}_{out} = Conv_{1 \times 1}(\mathbf{H}_{m1} \odot \mathbf{H}_{m2}; \mathbf{W}) \quad (11)$$

其中, \odot 代表元素相乘, \mathbf{W} 为卷积参数。

3.2.2 空间信息交互模块 SIIM

基于Transformer的CIIM关注通道上的全局上下文信息交互,对大气湍流引起的图像像素在空间上的几何偏移缺乏恢复能力。因此,设计了基于CNN的空间交互模块(SIIM),主要用于拟合湍流退化过程中的几何偏移函数逆操作,即:

$$SIIM(\cdot) \approx g^{-1}(\cdot) \quad (12)$$

该模块可以为空间中不同的特征信息赋予不同的权重,实现偏移像素值的恢复。SIIM的结构与主要流程如图3所示。给定输入张量 $\mathbf{P}_{in} \in R^{H \times W \times C}$,首先,分别在通道维执行最大池化与平均池化操作,以保留空间位置信息,具体计算过程如下:

$$\mathbf{P}_{avg} = \frac{1}{C} \sum_{c=1}^C \mathbf{P}_{in}(h, w, c) \quad (13)$$

$$\mathbf{P}_{max} = \max_c \mathbf{P}_{in}(h, w, c) \quad (14)$$

其中, $\mathbf{P}_{in}(h, w, c)$ 表示 \mathbf{P}_{in} 在通道 c 、位置 (h, w) 上的值,输出特征 $\mathbf{P}_{avg} \in R^{H \times W \times 1}$, $\mathbf{P}_{max} \in R^{H \times W \times 1}$ 。

然后,融合 \mathbf{P}_{avg} 和 \mathbf{P}_{max} ,以获取空间位置上对应的特征权重:

$$\mathbf{P}_w = \sigma(Conv_{3 \times 3}([\mathbf{P}_{avg}, \mathbf{P}_{max}]); \mathbf{W}) \quad (15)$$

其中, $[\cdot]$ 代表特征拼接操作, $Conv_{3 \times 3}(\cdot)$ 表示 3×3 卷积, \mathbf{W} 为卷积参数, $\sigma(\cdot)$ 表示 Sigmoid 激活函数。输出 \mathbf{P}_w 表示在空间维度上,SIIM对不同位置信息赋予的不同权重。

最终, \mathbf{P}_w 对输入特征 \mathbf{P}_{in} 进行相乘加权,输出 \mathbf{P}_{out} :

$$\mathbf{P}_{out} = \mathbf{P}_w \otimes \mathbf{P}_{in} \quad (16)$$

该操作可为偏移像素赋予新的数值,实现对偏移像素值的恢复。

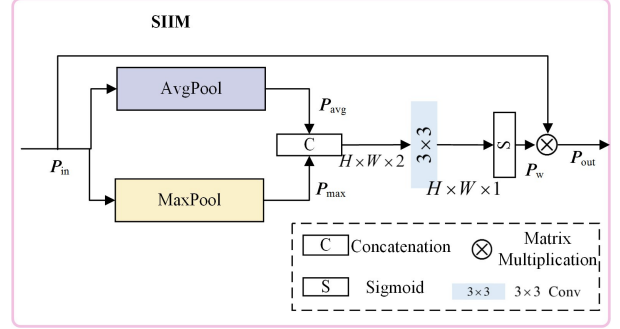


图3 空间信息交互模块

Fig. 3 Spatial information interaction module

3.3 对比学习

为了更充分地利用大气湍流退化图像与干净图像所提供的特征信息,受对比学习在图像复原^[25]上成功应用的启发,引入基于对比学习的正则化方法帮助图像复原网络在大气湍流干扰下恢复更好的视觉效果。

如图1所示,在获得网络复原后的图像 \mathbf{I}_r (Restored)后,将其视为锚点(Anchor),干净图像 \mathbf{I}_c 视为正样本(Positive),退化后的图像 \mathbf{I}_d (Degraded)视为负样本(Negative)。经过特征提取器(Feature Extractor)获得正样本、负样本更丰富的特征信息,将其映射到潜在特征空间(Latent Feature Space)中。在潜在特征空间中,将锚点与正样本之间的距离拉近(Pull),同时推远(Push)锚点与负样本的距离。这一基于对比学习的过程可以描述为一个正则化项^[12]:

$$L_{cr} = \sum_{i=1}^n \omega_i \frac{D(f_i(\mathbf{I}_c) - f_i(\mathbf{I}_r))}{D(f_i(\mathbf{I}_c) - f_i(\mathbf{I}_d))} \quad (17)$$

其中, $f_i(\mathbf{I}_c)$ ($i=1, 2, \dots, n$) 为特征提取器所提取的第 i 个潜在空间中的正样本特征, $f_i(\mathbf{I}_r)$ 与 $f_i(\mathbf{I}_d)$ 为锚点与负样本特征, $D(\cdot)$ 表示特征空间中的 L1 距离, ω_i 为第 i 个潜在空间的权重系数。采用预训练好的 VGG-19^[32] 网络作为特征提取器,在第 1, 3, 5, 9, 13 层提取特征,将这 5 层的权重系数 ω_i , $i=1, 2, 3, 4, 5$ 分别设置为 $\frac{1}{32}, \frac{1}{16}, \frac{1}{8}, \frac{1}{4}, 1$ 。对比正则化项的引入更充分地利用了大气湍流退化图像与干净图像所提供的特征信息,以复原高质量的去湍流干扰图像。

3.4 损失函数

除了对比正则化项,本文在复原图像 \mathbf{I}_r 与干净图像 \mathbf{I}_c 之间利用重建损失 L_{recon} ^[26],表示为:

$$\min L_{recon}(\mathbf{I}_c, \mathbf{I}_r) \quad (18)$$

最小化该损失,以帮助重建更真实的无湍流干扰图像。重建损失 L_{recon} 由 L1 损失与 MS-SSIM^[33] 损失联合构成:

$$L_{\text{recon}} = L1 + L_{\text{MS-SSIM}} \quad (19)$$

其中, $L1$ 损失保留图像的颜色与亮度,同时 MS-SSIM 损失在多尺度结构上保留图像的高频信息^[34],其相加得到的重建损失 L_{recon} 能更好地复原出高质量的图像。

因此,采用的总的损失函数可以表示为:

$$L = L_{\text{recon}} + \lambda L_{\text{cr}} \quad (20)$$

其中,为了平衡各个损失之间的权重,训练过程中将 λ 设置为 0.02。

4 实验结果与分析

4.1 实验设置

1)数据集:在不同湍流强度的合成人脸数据集与合成场景数据集中评估所提模型。选取 CelebA 数据集^[35]中 5 100 张图像,使用湍流模拟器^[36]合成湍流退化后的人脸图像,其中湍流强度(D/r_0)设为 1,选取 5 000 对图像用于训练,100 对图像用于验证选择最佳参数,在 Helen^[37]数据集中选择 200 张图像并利用相同湍流强度的湍流模拟器合成人脸测试集。在 Places 数据集^[38]中选取 10 200 张图像,采用湍流强度(D/r_0)为 2 的模拟器合成图像,10 000 对图像作为训练集,200 对图像作为测试集。此外,从真实场景下的 Heat Chamber^[8]数据集选取图像进行复原可视化效果测试。

2)参数设置:实验在两块 NVIDIA RTX 3090 GPU 上基于 Pytorch 1.11.0 实现。图像复原网络中 TMB 数量 L_1, L_2, L_3, L_4, L_r 分别设置为 [4, 6, 6, 8, 4]。CIIM 中基于通道数划分的头的数量为 [1, 2, 4, 8, 1],通道数为 [48, 96, 192, 384, 96]。对比正则化权重 λ 设为 0.02。使用 Adam 优化器优化参数,参数设置为默认值^[39],初始学习率设置为 0.0001,利用余弦退火策略调整学习率。输入图像大小均为 256×256 ,在 Places 数据集训练 40 次,CelebA 数据集上训练 30 次。

3)对比方法及评价指标:将所提出的方法与图像复原任务中主流的方法(MPRNet^[17], Uformer^[18], SwinIR^[40], Restormer^[20])和大气湍流图像复原方法(TurbNet^[8])进行比较。采用图像质量评估过程中常用的两个指标:峰值信噪比(Peak Signal-to-noise Ratio, PSNR)与结构相似度(Structural Silarity Index, SSIM)。实验中,PSNR 衡量复原图像与干净图像之间的差异,数值越大表示图像质量越高,失真越小,其计算公式为:

$$\text{PSNR} = 10 \log_{10} \left(\frac{\text{MAX}_I^2}{\text{MSE}} \right) \quad (21)$$

其中, MAX_I 为信号能量峰值, MSE 表示干净图像与复原后图像的均方误差,其计算公式为:

$$\text{MSE} = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2 \quad (22)$$

其中, I 表示干净图像, K 表示复原后的图像, i, j 表示像素坐标,图像大小为 $m \times n$ 。

SSIM 评价复原图像与干净图像之间的相似度,取值范围为 [0, 1],数值越大代表失真程度越低,其计算公式为:

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (23)$$

其中, μ_x, μ_y 和 σ_x, σ_y 分别表示 x, y 的像素均值和方差; σ_{xy} 代表协方差; c_1, c_2 为常数。

4.2 实验结果与分析

为了验证所提方法的有效性,在湍流强度为 2 的 Places 合成测试集与湍流强度为 1 的 Helen 合成测试集上与现有的 5 种图像复原方法进行对比,结果如表 1 所列。

表 1 所提方法与其他图像复原方法的比较

Table 1 Comparison of the proposed method with other image restoration methods

方法	Helen 合成测试集		Places 合成测试集	
	PSNR	SSIM	PSNR	SSIM
Uformer	26.26	0.7793	21.98	0.6512
MPRNet	25.28	0.7855	21.23	0.6588
Restormer	26.52	0.7853	22.07	<u>0.6676</u>
SwinIR	26.31	0.7807	21.46	0.6613
TurbNet	<u>26.59</u>	<u>0.7885</u>	<u>22.10</u>	0.6656
所提方法	26.78	0.7909	22.42	0.6820

注:黑色粗体表示最优方法,下划线表示次优方法。

在 Helen 合成测试集中,所提方法在 PSNR 与 SSIM 指标中取得了最优的效果,分别为 26.78 dB 与 0.7909。专用于湍流退化图像复原的方法 TurbNet 取得了次优效果,PSNR 和 SSIM 分别为 26.59 dB 与 0.7885。在测试集上,对于 PSNR 与 SSIM,超分辨率方法 SwinIR 分别为 26.31 dB 与 0.7807,通用的图像复原方法 Uformer 为 26.26 dB 与 0.7793,Restormer 为 26.52 dB 与 0.7853,MPRNet 为 25.28 dB 与 0.7855。所提方法的 PSNR 和 SSIM 较次优方法 TurbNet 提升了 0.19 dB 与 0.0024,较 SwinIR 提升了 0.47 dB 与 0.0102,较 Uformer 提升了 0.52 dB 与 0.0116,较 Restormer 提升了 0.26 dB 与 0.0056,较 MPRNet 提升了 1.5 dB 与 0.0054。

在 Places 合成测试集中,图像受湍流退化影响强度大,所提方法对比其他方法在 PSNR 与 SSIM 上仍取得了最优效果,分别为 22.42 dB 与 0.6820。在 PSNR 指标上,湍流退化图像复原网络 TurbNet 取得了次优效果 22.10 dB。在 SSIM 指标上,通用的图像复原方法 Restormer 取得了次优效果,为 0.6676。在该测试数据集中,所提方法的 PSNR 和 SSIM 相较于 TurbNet 提升了 0.32 dB 与 0.0164,较 Restormer 提升了 0.35 dB 与 0.0144。同时,所提方法的 PSNR 和 SSIM 较 Uformer 提升了 0.44 dB 与 0.0308,较 SwinIR 提升了 0.96 dB 与 0.0207,较 MPRNet 提升了 1.19 dB 与 0.0232。

在针对大气湍流退化后的人脸图像与一般场景图像的复原效果上,所提方法相比其他方法均取得了最佳的效果。如图 4 所示,对于不同的湍流强度,所提方法取得了高质量的视觉效果。在湍流退化强度为 2 的 Places 合成测试集中,如图 4 第一行方框标注所示,所提方法抑制了湍流退化引起的全局图像模糊,并且矫正了建筑物线条的扭曲,在视觉效果上取得了很好的表现;在湍流退化强度为 1 的 Helen 合成测试集中,如图 4 第二行红框所标注的部分,所提方法相比其余方法在人脸五官特征重塑上取得了更好的复原效果。这也表明所提方法在充分提取全局与局部信息的前提下,更好地利用了图像在特征空间中的信息。



图4 合成场景测试集与合成人脸测试集上不同方法的结果

Fig. 4 Results of different methods on synthetic scene test dataset and synthetic face test dataset

如图5所示,为了更好地论证所提方法在真实湍流场景下的复原能力和鲁棒性,利用在 Places 数据集上训练好的模型,选取真实拍摄的 Heat Chamber 数据集进行图像复原测试。针对大气湍流退化后的图像内容上出现模糊与扭曲的情况,所提方法在局部的特征细节复原上取得了高质量的效果,例如图像右上角的窗户与图像中心处的纹理细节(方框标注部分)。在图像的全局信息复原方面,与其他方法相比,所提

出的方法更好地突出了墙面上的砖块纹理,同时还能复原边缘轮廓信息。

为了验证对比学习与空间信息交互模块(SIIM)对于湍流退化图像复原的有效性,在 Helen 合成数据集上进行消融实验。实验结果如表2所列,其中 $-CR$ 表示在所提方法中去对比正则化项, $-SIIM$ 表示在所提方法中去除空间信息交互模块。

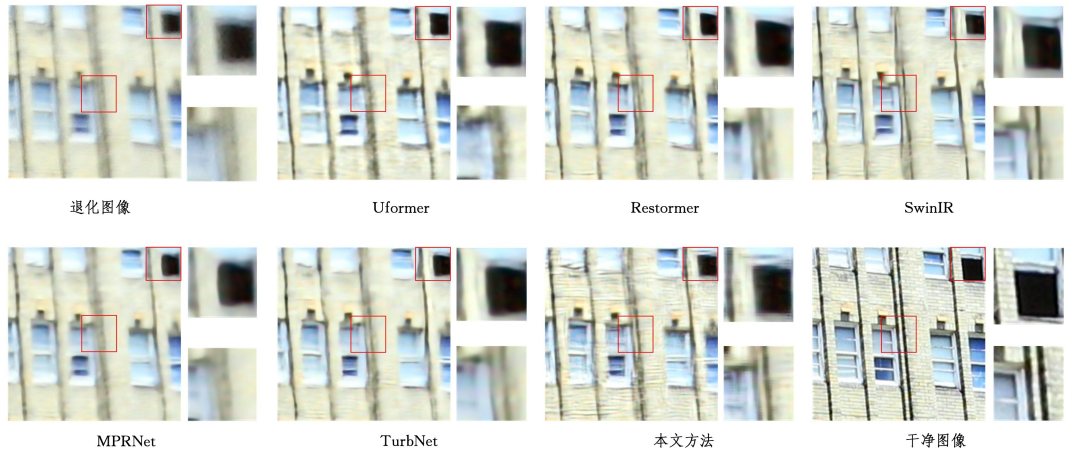


图5 真实场景下不同方法的可视化

Fig. 5 Visualization of different methods in real scene

表2 消融实验

Table 2 Ablation experiment

	PSNR	SSIM
$-CR$	26.63	0.7864
$-SIIM$	26.74	0.7901
所提方法	26.78	0.7909

注:黑色粗体表示最优效果。

由表2的结果可知,在分别去除对比正则化项与空间信息交互模块后,测试性能均有所下降。Helen 合成测试集上,在去除对比正则化项($-CR$)后,PSNR 和 SSIM 分别下降 0.15 dB 与 0.0045;在去除空间信息交互模块($-SIIM$)后,PSNR 和 SSIM 分别下降 0.04 dB 与 0.0008。实验结果表明,对比正则化可以帮助网络更充分地利用图像特征空间中的信息以复原更清晰的图像。同时,表2也说明,基于 CNN 的空间信息交互模块可以辅助基于 Transformer 的通道信息交互模块更充分地提取局部-空间的特征信息,更好地适应湍流引起的模糊与扭曲。

为了平衡损失函数中的各项优化目标权重,在 CelebA

数据集所划分的验证集上进行实验,选取最优的对比正则化权重 λ ,实验结果如表3所列。

表3 对比正则化项权重参数的选择

Table 3 Contrastive regularization term weight parameter selection

λ	PSNR
0.1	24.71
0.05	25.21
0.02	26.83
0.01	26.65

注:黑色粗体表示最优方法。

在该验证集上验证实验时,对比学习利用预训练好的 VGG19 网络提取特征至潜在的特征空间中计算,导致对比正则化项计算出的数值 L_{cr} 远大于重建损失 L_{recon} , λ 的比重过大会使所提模型训练时更加关注对比正则损失,从而忽略重建损失的影响。为了平衡这两项损失,在选择 λ 时从 0.1 开始依次降低,以确定最优的对比正则化权重系数。如表3所列,在 λ 取 0.02 时,复原效果达到最佳,为 26.83 dB,相较于

取 0.1 时提升 2.12 dB,相较于取 0.05 时提升 1.62 dB,相较于取 0.01 时提升 0.18 dB。

结束语 本文提出了一种基于对比学习的大气湍流退化图像复原方法,旨在解决大气湍流引起的图像模糊与扭曲问题。为了应对该挑战,设计了湍流缓解块融合 Transformer 和 CNN 的优点,捕获图像中通道与空间的全局-局部特征信息。同时采用对比学习,利用图像在特征空间上的信息帮助网络更好地复原受湍流退化影响的图像,并提升图像的质量和细节。

未来希望将已有的大气湍流复原网络与计算机视觉中的人脸识别、目标检测等下游任务相结合,帮助这些下游任务提升鲁棒性,从而应对大气湍流干扰的场景。

参 考 文 献

- [1] ROGGEMANN M C,WELSH B M,HUNT B R. Imaging through turbulence[M]. CRC press,1996.
- [2] ANANTRASIRICHAJ N,ACHIM A,KINGSBURY N G,et al. Atmospheric turbulence mitigation using complex wavelet-based fusion[J]. IEEE Transactions on Image Processing, 2013, 22(6):2398-2408.
- [3] MAO Z,CHIMITT N,CHAN S H. Image reconstruction of static and dynamic scenes through anisoplanatic turbulence[J]. IEEE Transactions on Computational Imaging, 2020, 6: 1415-1428.
- [4] ZHU X,MILANFAR P. Removing atmospheric turbulence via space-invariant deconvolution[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2012,35(1):157-170.
- [5] HIRSCH M,SRA S,SCHÖLKOPF B,et al. Efficient filter flow for space-variant multiframe blind deconvolution [C] // 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE,2010:607-614.
- [6] FRIED D L. Probability of getting a lucky short-exposure image through turbulence[J]. JOSA,1978,68(12):1651-1658.
- [7] YASARLA R,PATEL V M. Learning to restore a single face image degraded by atmospheric turbulence using cnns[J]. arXiv:2007.08404,2020.
- [8] MAO Z,JAISWAL A,WANG Z,et al. Single frame atmospheric turbulence mitigation: A benchmark study and a new physics-inspired transformer model[C]// Computer Vision-ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23-27, 2022, Proceedings, Part XIX. Cham: Springer Nature Switzerland,2022:430-446.
- [9] ZHANG X,MAO Z,CHIMITT N,et al. Imaging through the atmosphere using turbulence mitigation transformer[J]. arXiv: 2207.06465,2022.
- [10] JAISWAL A,ZHANG X,CHAN S H,et al. Physics-Driven Turbulence Image Restoration with Stochastic Refinement [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023:12170-12181.
- [11] VASWANI A,SHAZEER N,PARMAR N,et al. Attention is all you need[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. 2017: 6000-6010.
- [12] WU H,QU Y,LIN S,et al. Contrastive learning for compact single image dehazing[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 10551-10560.
- [13] ZHANG Y,LI K,LI K,et al. Image super-resolution using very deep residual channel attention networks[C]// Proceedings of the European Conference on Computer Vision (ECCV). 2018: 286-301.
- [14] YANG W,TAN R T,FENG J,et al. Deep joint rain detection and removal from a single image[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017: 1357-1366.
- [15] CAI B,XU X,JIA K,et al. Dehazenet: An end-to-end system for single image haze removal[J]. IEEE transactions on image processing, 2016, 25(11): 5187-5198.
- [16] YASARLA R,PATEL V M. Learning to restore images degraded by atmospheric turbulence using uncertainty[C]// 2021 IEEE International Conference on Image Processing (ICIP). IEEE,2021.
- [17] ZAMIR S W,ARORA A,KHAN S,et al. Multi-stage progressive image restoration[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 14821-14831.
- [18] WANG Z,CUN X,BAO J,et al. Uformer: A general u-shaped transformer for image restoration [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022:17683-17693.
- [19] CHEN H,WANG Y,GUO T,et al. Pre-trained image processing transformer [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 12299-12310.
- [20] ZAMIR S W,ARORA A,KHAN S,et al. Restormer: Efficient transformer for high-resolution image restoration[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022:5728-5739.
- [21] CHEN L,CHU X,ZHANG X,et al. Simple baselines for image restoration[C]// European Conference on Computer Vision. Cham: Springer Nature Switzerland,2022:17-33.
- [22] SERMANET P,LYNCH C,HSU J,et al. Time-contrastive networks: Self-supervised learning from multi-view observation [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops(CVPRW). IEEE,2017:486-487.
- [23] OORD A,LI Y,VINYALS O. Representation learning with contrastive predictive coding[J]. arXiv:1807.03748,2018.
- [24] CHEN T,KORNBLITH S,NOROUZI M,et al. A simple framework for contrastive learning of visual representations [C]//International Conference on Machine Learning. PMLR, 2020:1597-1607.
- [25] LIANG D,LI L,WEI M,et al. Semantically contrastive learning for low-light image enhancement[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2022:1555-1563.
- [26] WANG Y,XIONG J,YAN X,et al. USCFormer: Unified Transformer With Semantically Contrastive Learning for Image Dehazing[J]. IEEE Transactions on Intelligent Transportation

- Systems, 2023, 24(10):11321-11333.
- [27] WU G, JIANG J, LIU X. A practical contrastive learning framework for single-image super-resolution[J]. arXiv: 2111. 13924, 2023.
- [28] YE Y, YU C, CHANG Y, et al. Unsupervised deraining: Where contrastive learning meets self-similarity[C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022;5821-5830.
- [29] HIRSCH M, SRA S, SCHÖLKOPF B, et al. Efficient filter flow for space-variant multiframe blind deconvolution [C] // 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, 2010;607-614.
- [30] CHAN S H. Tilt-then-blur or blur-then-tilt? clarifying the atmospheric turbulence model[J]. IEEE Signal Processing Letters, 2022, 29:1833-1837.
- [31] SHI W, CABALLERO J, HUSZÁR F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 1874-1883.
- [32] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. arXiv: 1409. 1556, 2014.
- [33] WANG Z, SIMONCELLI E P, BOVIK A C. Multiscale structural similarity for image quality assessment[C] // The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers. IEEE, 2003:1398-1402.
- [34] ZHAO H, GALLO O, FROSIO I, et al. Loss functions for image restoration with neural networks [J]. IEEE Transactions on Computational Imaging, 2016, 3(1):47-57.
- [35] LIU Z, LUO P, WANG X, et al. Deep learning face attributes in the wild[C] // Proceedings of the IEEE International Conference on Computer Vision. 2015:3730-3738.
- [36] MAO Z, CHIMITT N, CHAN S H. Accelerating atmospheric turbulence simulation via learned phase-to-space transform [C] // Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021:14759-14768.
- [37] LE V, BRANDT J, LIN Z, et al. Interactive facial feature localization[C] // Computer Vision-ECCV 2012: 12th European Conference on Computer Vision, Florence, Italy, Part III 12. Springer Berlin Heidelberg, 2012;679-692.
- [38] ZHOU B, LAPEDRIZA A, KHOSLA A, et al. Places: A 10 million image database for scene recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 40 (6): 1452-1464.
- [39] KINGMA D P, BA J. Adam: A method for stochastic optimization[J]. arXiv:1412. 6980, 2014.
- [40] LIANG J, CAO J, SUN G, et al. Swinir: Image restoration using swin transformer[C] // Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021:1833-1844.
- [41] CHEN F, WANG H. Design of fuzzy image quality restoration algorithm based on multi-channel visual attention[J]. Journal of Jilin University(Engineering and Technology Edition), 2023, 53 (9):2626-2631.



MIAO Zhuang, born in 1976, Ph.D, professor, Ph.D supervisor. His main research interests include artificial intelligence, pattern recognition and computer vision.



WANG Jiabao, born in 1985, Ph.D, associate professor. His main research interests include computer vision and image processing.

(责任编辑:何杨)