基于傅里叶卷积和多特征调制的图像修复网络

孙刘杰,刘倩倩,庞茂然

(上海理工大学 出版印刷与艺术设计学院,上海 200093)

摘要:目的 解决大面积破损难以修复且修复过程中感受野、特征空间信息利用不足,导致修复后的孔 洞区域与背景之间出现结构、纹理、风格不一致的问题。方法 基于傅里叶卷积和多特征调制的修复网 络FFC-MFMGAN,傅里叶卷积在网络的浅层便具有较大的感受野,尤其是在宽掩码时能够跳过掩码区 域,捕获到有效特征,多特征调制生成网络能够分别利用完整区域的信息和随机样式操纵,增强与未受 损区域的语义连贯性,以及大空洞率下修复的多样性。结果 在 Place 2 数据集上,将文中方法与其他图像 修复方法进行了对比实验,经过测试,各类指标均得到明显改善,峰值信噪比提高了 1.4%,结构相似性提 高了 4.5%,平均绝对误差降低了 12.6%,基于学习的感知图像块相似性降低了 9.1%。结论 FFC-MFMGAN 网络能够较好地修复大面积不规则孔洞,同时增强修复图像的全局结构性和清晰度,对实际包装印刷图 像的缺陷修复也有一定参考价值。

关键词:图像修复;深度学习;感受野;傅里叶卷积;特征调制

中图分类号: TP391 文献标识码: A 文章编号: 1001-3563(2023)21-0286-08 DOI: 10.19554/j.cnki.1001-3563.2023.21.036

Image Inpainting Network Based on Fourier Convolution and Multi-feature Modulation

SUN Liu-jie, LIU Qian-qian, PANG Mao-ran

(College of Communication and Art Design, University of Shanghai for Science and Technology, Shanghai 200093, China)

ABSTRACT: The work aims to address the problem that large areas of damage are difficult to repair, the perceptual field and feature space information are under-utilized in the repair process, and the structure, texture and style between the repaired cavity areas and the background are inconsistent. An inpainting network FFC-MFMGAN based on Fourier convolution and multi-feature modulation was proposed. Fourier convolution had a large perceptual field in the shallow layer of the network, especially in wide masks, which could skip the mask zone to capture the effective features especially when the mask was wide. The generative network based on the multi-feature modulation was able to enhance the semantic coherence with undamaged regions and the diversity of restoration at large void rates with information from intact regions and random pattern manipulation, respectively. Experiments were conducted to compare the proposed method with other state-of-the-art image restoration methods on the Place 2 dataset, and the following categories were tested to show significant improvements, including a 1.4% improvement in PSNR, a 4.5% improvement in SSIM, a 12.6% reduction in MAE, and a 9.1% reduction in LPIPS. The FFC-MFMGAN network can better repair large irregular holes, while enhancing the global structure and clarity of the repaired images, which is also a reference value for the repair of defects in actual packaging printing images.

KEY WORDS: image inpainting; deep learning; perceptual field; Fourier convolution; feature modulation

收稿日期: 2023-03-07

图像修复是重建丢失或损坏图像的过程,通过对 邻域信息进行重建,可以使图像恢复到原始状态,是 计算机视觉领域的研究热点之一。在本质上,图像修 复是一种利用已知区域的相关信息对缺失区域进行 填充的技术,不仅要求填充后的内容在视觉上合理, 而且在纹理上能够正确延展,细节表现一致,保证视 觉效果,满足视觉认知规律。在现实中一些旧照片、 文物壁画因时间久远、保存环境不当,出现了折痕、 污渍、脱落等损伤。正确修复受损图片对于恢复图像 质量、保持信息完整性和后续图像处理具有很大的应 用价值,已成为计算机视觉的主要研究内容之一。尤 其是针对受损的文物和古字画,经过修复可以还原其 样貌、大大减少工作时间、避免造成不可逆的伤害。 总之,图像修复在还原旧照片^[1-2],场景编辑和修饰, 去除水印及移除目标对象[3-4]等领域都有着广泛应用。 针对目前图像修复方法存在的修复边界不清晰,连续大 面积缺失下效果模糊和结构扭曲等问题,提出基于傅里 叶卷积^[5]和多特征调制的修复网络 FFC-MFMGAN,结 合全局、局部特征和随机样式进行精细修复,从而获 得具有特征连续、语义一致的修复图像。

1 相关工作

图像修复方法可以划分为传统图像修复方法^[6-7] 和基于深度学习的图像修复方法^[8-12]。传统图像修复 方法主要借助数学运算,包括基于扩散和补丁块的方 法。基于深度学习的图像修复方法通过神经网络强大 的学习能力恢复缺失像素,对数据集中的海量图像进 行训练,它可以学习到图像的各种特征信息,使得修 复的结果图像更贴合真实图像,在纹理细节和颜色等 方面更连贯。这里基于深度学习方法进行修复研究, 并从上下文编码器、多阶段模型、U-Net 网络模型对 深度学习修复方法进行阐述。

上下文编码器 (Context Encoder, CE)^[13]修复模型基于 Encoder-Decoder 体系架构,编码器对输入的 图像进行特征提取,解码器逐步放大压缩的特征映 射,并恢复到原始图像大小,预测缺失区域的像素。 该方法不能生成详细的纹理特征,导致区域边界模 糊。结合双鉴别器的 GLCIC^[14]方法,从全局和局部 2 个角度约束生成器,实现更精细的图像修复。其中, 全局鉴别器用于评估整个图像的一致性,而局部鉴别 器以受损区域为中心的 1/4 区域为输入,提高了局部 细节质量。

此外,还有一些多阶段修复模型,例如 CA^[15]模型,以引入注意力、由粗略重建到细致修复的方式恢 复图像。 EdgeConnect 首先利用边缘生成器,在缺失 区域预测可能的边缘,并将其投喂到下一个图像修补 网络,然后与受损信息组合生成最终恢复的图像^[3]。

条件纹理和结构双生成的图像修复方法 (CTSDG)^[16]基于 U-Net 型主干网络,在结构和纹 理特征之间共享和组合信息,以增强一致性,并通过 上下文特征聚合(CFA)模块,采用多尺度特征聚合, 在多个尺度上对丰富的语义特征进行编码,较好地平 衡了网络模型的时间空间复杂性和修复的准确性。同 时,基于 U-Net 网络,LGNet 利用大感受野的粗略网 络进行图像结构的简略修复,利用小感受野下的全局 细化网络进行细粒度修复^[17]。

文中通过将随机潜向量经过映射网络产生的风格与输入图像经过多层卷积后再与深层的特征聚合, 使前者能够保证随机风格,后者能够保证条件风格,二 者在解码器阶段的每个不同分辨率卷积层互相调制。在 损失较大时,既保证了补全内容的随机性,又保证与周 围像素的统一,有助于生成自然的视觉效果。

2 提出方法

2.1 模型框架

这里提出的基于傅里叶卷积和多特征调制的图 像修复网络总体框架如图1所示,网络主要由4个部 分组成:利用通道式傅里叶卷积的条件编码器网络、 映射网络、多特征调制模块和损失函数计算模块。

条件编码网络将未受损部分的图像作为生成条件,合理地提取特征。这是因为可以将图像修复看作 是一个受条件约束的图像生成问题,这里的条件指未 损坏区域的像素和特征分布。极端情况是当原图像被 完全遮掩时,修复任务转换为随机图像的生成。在实 际情况下,图像的受损程度不一,导致修复的主要难 点转换为保证补全区域与完好区域之间的视觉连贯 性和语义正确性。

映射网络能够将随机潜向量编码为具有风格特征的中间向量,然后与输入图像经过多层卷积后与深层的特征聚合,使前者能够保证随机风格,后者能够保证条件风格。二者在多特征调制生成网络的每个不同分辨率卷积层互相调制,在损失较大时既保证补全内容的随机性,又保证与周围像素统一,有助于生成自然的视觉效果。

2.2 FFC-MFMGAN 网络结构

令 I_{gt} 为真实图像,则受损图像可表示为 $I_{masked} = I_{gt} \odot (1-M)$,其中, \odot 表示两矩阵逐元素相乘。 由于现阶段仍为非盲图像修复,所以还需输入二值掩码 图像 M(已知区域像素为 0,缺失区域像素为 1),并在 通道维度上拼接,即网络输入图像 $I_{in}=(I_{masked}, M)$ 。

2.2.1 通道式傅里叶卷积

在图 1 中,傅里叶卷积层采用通道式快速傅里叶 卷积。通道式快速傅里叶卷积可以有效实现非局部感 受野,并融合多尺度信息,在网络的浅层就具有大的 感受野,从而影响全局上下文内容。同时,基于傅里 叶变化的周期性,有利于呈规律分布纹理的修复。



输入图像 I_{in} 首先进入金字塔结构的编码器部分, 编码器主要采用 8 个卷积组,每个组分为傅里叶卷积 块和步长的普通卷积块。快速傅里叶卷积使用的是通 道式快速傅里叶变换,因此需要对某一图像特征 $F_{C \times H \times W}$ 按照比例 y 进行通道维度的分割,即分为局部 特征 $F_{\gamma^* C \times H \times W}$ 和全局特征 $F_{(l-\gamma)^* C \times H \times W}^{global}$ 。

全局和局部特征由4个不同分支(局部至局部、 局部至全局、全局至局部、全局至全局)进行卷积、 归一化等操作,其中前3个分支都借助传统的卷积法 进行特征图的局部更新,只有最后一个分支对特征图 进行傅里叶变换,并在影响全局的谱域中对其进行更 新,每个分支通过不同的感受野捕获互补信息。保留 局部分支和标准卷积的作用是保证在窄掩码下修复 的稳定性。对4条分支中到达局部和到达全局的特征 分别进行汇合叠加, 使 2 条信息链相互作用、交换信息, 网络每层都有能力查看全局上下文的信息。

快速傅里叶变换分支的具体实现步骤如图 2 所示,对输入特征图进行二维傅里叶变换,并在通道维度上将频谱的实部和虚部拼接后送入后续网络,保证其在频率域进行卷积处理时再通过傅里叶逆变换至空间域。最后通过归一化、ReLu函数激活后,在通道维度重新拼接,生成与原始特征同尺度的特征 $F_{C\times H\times W}$,然后通过步长为 2 的卷积生成 $F_{2C\times H/2\times W/2}$,增加了更多的线性变换。不同于前 7 层,为了产生风格样式向量,最后一层卷积后的 $F_{512\times 4\times 4}$ 特征图将平铺展开后通过稠密层(全连接层),与前面提取的特征进行关联,并映射到输出特征空间, Dense₅₁₂ = f_c (flatten($F_{512\times 4\times 4}$))。



图 2 通道式厌迷博生叶苍秋 Fig.2 Channel fast Fourier convolution

· 289 ·

2.2.2 多特征调制模块

如果将图像修复看作是对缺失部分的再生成过程,那么 *I*_{masked}中的已知像素点是生成条件。随着缺失区域的扩增,可利用的生成条件越来越少,修复难度攀升。按照生成对抗网络 GAN 的思想,只输入随机噪声,依靠生成器与鉴别器间的博弈网络也能够生成一定质量的图像,但是低质量、低分辨率的图像远不能达到最终的修复目的,并且 GAN 存在难以收敛和训练不常应用于工程等缺点。由此,文中在模型能够根据随机噪声生成图像时,给定 *I*_{masked}生成条件,2 种控制细节与条件生成的特征相互调制,在修复细粒度的同时与 *I*_{masked}区域建立结构、语义一致性,降低了修复难度。

特征调制最初由风格转换任务一文提出^[18]。如图 1 所示,根据输入图像 I_{in} 的长和宽生成服从正态分布的随机向量 z,此时 $z \in Z$ 特征空间;之后类似于 StyleGAN V2 的改进,经过由 8 个全连接层 f_c 构成的 映射网络,将随机采样的数据投影到 W特征空间,W 空间中各向量的相互关系更加线性,解决了特征纠缠问题,见式(1)。

$$w = f_{\rm c}(z) \tag{1}$$

 $w \in W$,进入网络前先复制 15 份,即在 W空间中 依次取出 w 与条件编码器最后一层的输出 D_{512} ,拼接 新的样式向量 s,共同对 15 个卷积层的权重 w_{ijk} 进行 调制,见式(2)~(3)。既保证与周围区域的语义关 联性,也丰富了生成图像的细节。

$$s = f_{\rm c}(w, D_{512})$$
 (2)

$$Mod(w_{iik}) = w'_{iik} = w_{iik} \odot s$$
(3)

为了避免权重调制后直接卷积的输出特征图受到 *s* 的影响,还需要进行归一化处理,即解调制过程,见式(4)。

$$DeMod(w'_{ijk}) = w'_{ijk} = \frac{w'_{ijk}}{\sqrt{\sum_{ik} w'_{ijk}^2 + \varepsilon}}$$
(4)

设置极小常数 ε =10⁻⁸是为了避免分母为 0。经过 解调制后,新权重矩阵用于逐层反卷积阶段恢复特 征,并借助由普通卷积构成的 ToRGB 层转回三通道 特征图像,每层的 ToRGB 仍通过 U-Net 跳跃连接的 方式对同一尺度特征进行聚合,在跨层连接后逐级采 样,渐进式得到修复结果 I_{out} 。

2.3 损失函数

为了保证待修复区域与周围图像的视觉合理性, 提升 2 个部分的高层语义一致性。文中采用像素重建 损失 *L*_{rec}^[19]和生成对抗损失 *L*_{adv}^[20],以生成精细的纹 理细节,并引入感知损失 *L*_{perc}^[21]和风格损失 *L*_{style}^[22], 从而更好地把握生成部分,使其更贴合整体图像的风 格。总损失函数 *L*_{all}的计算见式(5)。

$$L_{\text{all}} = \alpha_{\text{r}} L_{\text{rec}} + \alpha_{\text{a}} (L_{\text{adv},1} + L_{\text{adv},2}) + \alpha_{\text{p}} L_{\text{perc}} + \alpha_{\text{s}} L_{\text{style}} \quad (5)$$

$$\vec{x} \oplus : \alpha_{\text{r}} = 1; \ \alpha_{\text{a}} = 0.1; \ \alpha_{\text{p}} = 0.1; \ \alpha_{\text{s}} = 250_{\circ}$$

3 结果与分析

为了验证文中算法的优越性,将其与另外5种较 有代表性的图像修复模型分别进行了定量和定性对 比实验,并设计消融实验,验证了傅里叶卷积和多特 征调制模块在修复过程中的有效性。

3.1 数据集和实验设置

使用 Places 2 公开图像数据集来验证模型的修复性 能^[23]。掩码图像来自 Nvidia 不规则掩码数据集^[24],具有 复杂背景和多类别随机掩码的大型数据集可以保证数据 集的多样性。文中模型在 Tensflow 1.4 上实现,由 Adam 优化器优化,学习率为 0.000 1。显卡芯片为 NVIDIA Ge-Force RTX 2080 Ti,训练模型时批次 BatchSize 为 4。

3.2 定量分析

对测试集求其相应评价指标的均值,以评估模型 的修复能力。文中方法和比较方法在 Place 2 数据集 上的结果见表 1,孔洞与图像的面积比分别为(0.1, 0.2]、(0.2,0.3]、(0.3,0.4]、(0.4,0.5%]。文中采用峰 值信噪比(PSNR)^[25]、结构相似性指数(SSIM)^[26]、 平均绝对误差(MAE)^[27]和学习感知图像块相似度 (LPIPS)^[28]来评估修复结果。

分析表 1 的数据可知,与 LGNet 相比,文中方 法在 2 个数据集上的 PSNR 值和 SSIM 值分别提高了 1.39%、4.53%,MAE 值和 LPIPS 值分别降低了 12.65%、 9.13%。PSNR 和 SSIM 的结果表明,文中方法可以得 到更接近原图、失真更少的视觉结果。MAE、LPIPS 的值减小表明,修复后的图像与真实图像像素之间的距 离最小,视觉质量较好,更符合人眼感知。

3.3 定性分析

为了更直观展现视觉对比结果,文中从 Places 2 数据集中选择部分图像样本。文中模型和其他 5 种对 比方法生成的结果如图 3 所示。

从图 3 可以看到, CA 在恢复图像原貌过程中产生 了生硬的修复边界、扭曲及错误的内容。EC 在结构上 得到显著的视觉改进,但在掩码密集分布和规则建筑结 构的区域中容易产生模糊和伪影,越靠近掩码中心的区 域越模糊,未恢复原有建筑物的边缘结构。RN 的修复 结果相较于 EC 更清晰,但在粗掩码处仍无法修复高级 语义信息,且能够明显看出修复痕迹和原始掩码的位 置。CTSDG 通过在结构重建过程中与图像纹理进行适 当交互,解决了中心模糊问题,但生成的结构无法与周 围未受损区域建立视觉连贯性,出现了一些无序内容和 目标错位。LGNet 的修复区域过于平坦,且边界过渡异 常,导致图像失真。文中方法无论在简单背景还是结构 复杂的建筑体上都得到了最佳的视觉效果,保证了与未 受损区域的整体一致性,甚至在海面产生了其他方法都 不能产生的水面波纹,丰富了自然纹理细节。

Tab.1 Quantitative comparison results using different proportion masks on Place 2 dataset											
掩码率	评价指标	CA ^[15]	EdgeConnect ^[3]	RN ^[29]	CTSDG ^[16]	LGNet ^[17]	文中方法				
(0,1,0,2)	PSNR ⁺ /dB	24.761 2	28.707 3	28.669 7	29.025 8	30.802 5	30.693 8				
	$SSIM^+$	0.860 0	0.920 7	0.933 0	0.940 2	0.941 8	0.949 0				
(0.1, 0.2]	MAE	0.024 9	0.013 1	0.016 6	0.014 3	0.011 5	0.018 3				
	LPIPS ⁻	0.044	0.051	0.060	0.046	0.053	0.049				
(0.2.0.2)	PSNR ⁺ /dB	21.637 5	25.987 3	26.302 9	26.075 2	27.190 9	27.578 9				
	$SSIM^+$	0.757 7	0.869 6	0.893 0	0.900 7	0.872 9	0.910 2				
(0.2, 0.3]	MAE	0.041 8	0.022	0.023 9	0.027 4	0.041 5	0.019 8				
	LPIPS	0.073	0.080	0.088	0.073	0.091	0.069				
(0.3, 0.4]	PSNR ⁺ /dB	20.124 8	23.297 5	23.520 4	23.703 7	24.029 5	25.200 8				
	\mathbf{SSIM}^+	0.670 3	0.787	0.814 8	0.814 7	0.817 7	0.872 6				
	MAE	0.057 4	0.035 8	0.037 6	0.030 6	0.032 2	0.0233 5				
	LPIPS ⁻	0.134	0.139	0.150	0.147	0.165	0.141				
(0.4, 0.5]	PSNR ⁺ /dB	19.335 1	22.736 4	22.580 9	23.118 2	23.506 1	23.596 3				
	\mathbf{SSIM}^+	0.582 7	0.750 8	0.779 7	0.785 9	0.781 0	0.830 6				
	MAE	0.071 9	0.042 3	0.048 3	0.042 3	0.038 4	0.026 9				
	LPIPS ⁻	0.154	0.150	0.168	0.159	0.185	0.203				

表 1 Place 2 数据集上使用不同比例掩膜的定量比较结果 Quantitative comparison results using different proportion masks on Place 2 day

注:每个指标的最佳值通过加粗方式表示。



图 3 在 Places 2 数据集上各修复方法可视化对比 Fig.3 Visualization comparison results of various inpainting methods on the Places 2 dataset

3.4 消融实验

为了验证文中所提改进方案的有效性,选择 LBAM^[4](以 U-Net 为网络框架)作为消融实验研究 的基线,基线+FFC表示引入快速傅里叶卷积,在此 基础上添加多特征调制模块构成完整的文中模型。同 时,为了保证实验对比的公平性,所有的基本数据和 参数设置保持不变,实验的评价标准同 3.2 节。

如表 2 所示,将普通卷积块替换为傅里叶卷积块

后,PSNR 值和 SSIM 值分别提高了 1.33%、1.74%, MAE 值和 LPIPS 值分别降低了 17.0%和 7.4%。傅里叶卷积块 的全局感受野能够保证捕获到远距离的有效信息,从而 得到正确、清晰的修复结果。对比基线模型,增加了多 特征调制模块后,LPIPS 值平均降低了 20.7%。一方面 多特征调制模块能够合理利用受损区域外围的有效信 息,达到两部分区域语义一致、连贯的效果。另一方面, 由全连接层得到的样式操纵也能满足大掩码时修复的 多样性,确保生成的图像更符合人眼的感知情况。 从图 4 的可视化结果可以看出,当掩码紧密覆盖时,即使待修复对象是单一的墙体背景,在基线模型的修复结果中仍出现了不合理的冗余结构,在添加了 傅里叶卷积后模型产生了较平整的结构,在重复纹理 特征的修复上得到很大改善。这种混乱结构并未完全 消除,在完整模型下的多特征随机样式调制能够针对 宽掩码很好地发挥作用,较好地避免了结构失真。通 过修复结果可视图可以发现,随着傅里叶卷积和多特 征调制模块的逐步引入,在纹理结构的一致性和细粒 度视觉细节方面也逐步优化。

3.5 印刷修复应用

除了上述对自然景观及建筑图像的修复,文中方

法还适用于包装印刷领域,例如古字画的修复。目前 存世和尚未出世的大量古字画都有所残缺,通过人工 补绘的方式进行修复存在较大的代价和风险。这不仅 需要扎实且全面的专业技能,而且易对古籍艺术品造 成不可逆的伤害。通过数字化修复技术能重现它们的 美感和艺术性。为了更直观地展现修复效果,首先从 《敦煌画卷》中选取 2 幅本身有部分缺失的古画图 像,然后使用白色笔刷绘制待修复区域,以创建掩码 和输入图像,并进行修复。文中方法与其他修复方法 的对比结果如图 5 所示。

从图 5 的修复结果可以看出,所有比较算法均出现了修复不彻底和错误修复等问题。CTSDG 恢复的 区域有限,仅在很小范围内生成了图像,且生成

rad.2 Quantitative comparison results of adiation experiments								
掩码率	评价指标	基线	基线+FFC	文中模型				
	PSNR+/dB	30.54	31.16	32.40				
(0, 1, 0, 2)	SSIM+	0.940	0.950	0.968				
(0.1, 0.2)	MAE-	0.012 7	0.011 1	0.010 8				
	LPIPS-	0.054	0.047	0.040				
	PSNR+/dB	27.52	27.92	28.14				
(0, 2, 0, 2)	SSIM+	0.904	0.912	0.917				
(0.2, 0.3)	MAE-	0.025 8	0.020 8	0.017 6				
	LPIPS-	0.081	0.066	0.056				
	PSNR+/dB	25.66	25.94	26.23				
(0, 2, 0, 4)	SSIM+	0.823	0.855	0.884				
(0.3, 0.4)	MAE-	0.037 0	0.032 9	0.020 6				
	LPIPS-	0.146	0.142	0.129				
	PSNR+/dB	24.15	24.33	24.54				
(0, 1, 0, 5)	SSIM+	0.811	0.820	0.849				
(0.4, 0.5)	MAE-	0.043 4	0.032 6	0.025 2				
	LPIPS-	0.215	0.225	0.184				

表 2 消融实验定量对比结果 bb 2 Quantitative comparison results of ablation experiments



图 4 Paris StreetView^[30]图像数据集上消融实验定性对比结果 Fig.4 Qualitative comparison results of ablation experiments on Paris StreetView^[30] image dataset





部分出现纹理错乱现象,整体修复效果较差。LGNet 修复的范围比 CTSDG 稍大,但仍不能恢复图像纹理 和细节。采用文中方法对 2 幅古画图像修复后都得到 了最佳的视觉效果,虽然存在一定的模糊问题,但生 成的图像完整,整体色彩自然,保证了与未受损区域 的整体一致性。由此可知,文中算法能够应用于实际 印刷领域的图像修复,具有良好的泛化性。

4 结语

设计了一种结合傅里叶卷积和多特征调制的图 像修复网络FFC-MFMGAN。通过傅里叶卷积使网络 具有更大的感受野,可以跳过局部缺失的像素,从远 处空间位置获取有效信息,解决了大面积破损图像不 能正确利用已知内容的问题。多特征调制生成模块能 够将控制随机样式细节生成与条件生成的特征相互 调制,保证了在大面积破损下填充区域的多样性,并 生成了与已知区域风格一致的内容。在 Place 2 数据 集上对文中方法和其他 5 种主流修复方法进行了实 验和性能对比,最终测试结果表明,文中方法能够合 理地修复较大面积的不规则孔洞,同时改善修复细 节,增强图像细节纹理,保持风格的整体一致性。在 实际应用中,文中研究内容对包装印刷领域的图像缺 陷修复、医学、文物图像修复也具有一定的参考意义。

参考文献:

- SHIH M L, SU S Y, KOPF J, et al. 3D Photography Using Context-Aware Layered Depth Inpainting[C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Seattle, WA, USA IEEE, 2020: 8025-8035.
- [2] WAN Zi-yu, ZHANG Bo, CHEN Dong-dong, et al.

Bringing Old Photos Back to Life[C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Seattle, WA, USA IEEE, 2020: 2744-2754.

- [3] NAZERI K, NG E, JOSEPH T, et al. EdgeConnect: Structure Guided Image Inpainting Using Edge Prediction[C]// 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW) Seoul, Korea (South) IEEE, 2020: 3265-3274.
- [4] XIE Chao-hao, LIU Shao-hui, LI Chao, et al. Image Inpainting with Learnable Bidirectional Attention Maps[C]// 2019 IEEE/CVF International Conference on Computer Vision (ICCV) Seoul, Korea (South) IEEE, 2020: 8857-8866.
- [5] CHI L, JIANG B, MU Y. Fast Fourier Convolution[C]// Advances in Neural Information Processing Systems, 2020(33): 4479-4488.
- [6] BALLESTER C, BERTALMIO M, CASELLES V, et al. Filling-in by Joint Interpolation of Vector Fields and Gray Levels[J]. IEEE Transactions on Image Processing, 2001, 10(8): 1200-1211.
- [7] ESEDOGLU S, SHEN Jian-hong. Digital Inpainting Based on the Mumford–Shah–Euler Image Model[J]. European Journal of Applied Mathematics, 2003, 13(4): 353-370.
- [8] CRIMINISI A, PEREZ P, TOYAMA K. Region Filling and Object Removal by Exemplar-Based Image Inpainting[J]. IEEE Transactions on Image Processing, 2004, 13(9): 1200-1212.
- [9] BARNES C, SHECHTMAN E, FINKELSTEIN A, et al. PatchMatch: A Randomized Correspondence Algorithm for Structural Image Editing[J]. ACM Trans Graph, 2009, 28(3): 24.
- [10] HUANG J B, KANG S B, AHUJA N, et al. Image Com-

pletion Using Planar Structure Guidance[J]. ACM Transactions on Graphics (TOG), 2014, 33(4): 1-10.

- [11] BERTALMIO M, VESE L, SAPIRO G, et al. Simultaneous Structure and Texture Image Inpainting[C]// 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003 Proceedings Madison, WI, USA IEEE, 2003: 707.
- SONG Yi-bing, BAO Lin-chao, HE Sheng-feng, et al. Stylizing Face Images via Multiple Exemplars[J]. Computer Vision and Image Understanding, 2017, 162: 135-145.
- [13] PATHAK D, KRÄHENBÜHL P, DONAHUE J, et al. Context Encoders: Feature Learning by Inpainting[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Las Vegas, NV, USA IEEE, 2016: 2536-2544.
- [14] IIZUKA S, SIMO-SERRA E, ISHIKAWA H. Globally and Locally Consistent Image Completion[J]. ACM Transactions on Graphics, 36(4): 107.
- [15] YU Jia-hui, LIN Zhe, YANG Ji-mei, et al. Generative Image Inpainting with Contextual Attention[C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Salt Lake City, UT, USA IEEE, 2018: 5505-5514.
- [16] GUO X, YANG H, HUANG D. Image Inpainting via Conditional Texture and Structure Dual Generation[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 14134-14143.
- [17] QUAN Wei-ze, ZHANG Rui-song, ZHANG Yong, et al. Image Inpainting with Local and Global Refinement[J]. IEEE Transactions on Image Processing, 2022, 31: 2405-2420.
- [18] KARRAS T, LAINE S, AILA Ti-mo. A Style-Based Generator Architecture for Generative Adversarial Networks[C]// 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Long Beach, CA, USA IEEE, 2020: 4396-4405.
- [19] WANG Xin-tao, YU Ke, DONG Chao, et al. Recovering Realistic Texture in Image Super-Resolution by Deep Spatial Feature Transform[C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Salt Lake City, UT, USA IEEE, 2018: 606-615.
- [20] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative Adversarial Nets[C]// Proceedings of

the 27th International Conference on Neural Information Processing Systems-Volume 2 Montreal, Canada. New York: ACM, 2014: 2672-2680.

- [21] JOHNSON J, ALAHI A, FEI-FEI L. Perceptual Losses for Real-Time Style Transfer and Super-Resolution[C]// Computer Vision-ECCV 2016: 14th European Conference, Amsterdam, 2016: 694-711.
- [22] GATYS L A, ECKER A S, BETHGE M. Image Style Transfer Using Convolutional Neural Networks[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Las Vegas, NV, USA IEEE, 2016: 2414-2423.
- [23] ZHOU Bo-lei, LAPEDRIZA A, KHOSLA A, et al. Places: A 10 Million Image Database for Scene Recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(6): 1452-1464.
- [24] LIU G, REDA F A, SHIH K J, et al. Image Inpainting for Irregular Holes Using Partial Convolutions[C]// Proceedings of the European Conference on Computer Vision (ECCV), 2018: 85-100.
- [25] KORHONEN J, YOU Jun-yong. Peak Signal-to-Noise Ratio Revisited: Is Simple Beautiful?[C]// 2012 Fourth International Workshop on Quality of Multimedia Experience Melbourne, VIC, Australia IEEE, 2012: 37-38.
- [26] WANG Zhou, BOVIK A C, SHEIKH H R, et al. Image Quality Assessment: From Error Visibility to Structural Similarity[J]. IEEE Transactions on Image Processing: a Publication of the IEEE Signal Processing Society, 2004, 13(4): 600-612.
- [27] WILLMOTT C J, MATSUURA K. Advantages of the Mean Absolute Error (MAE) over the Root Mean Square Error (RMSE) in Assessing Average Model Performance[J]. Climate Research, 2005, 30: 79-82.
- [28] ZHANG R, ISOLA P, EFROS A A, et al. The Unreasonable Effectiveness of Deep Features as a Perceptual Metric[C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Salt Lake City, UT, USA IEEE, 2018: 586-595.
- [29] YU T, GUO Z, JIN X, et al. Region Normalization for Image Inpainting[C]// Proceedings of the AAAI Conference on Artificial Intelligence, 2020: 12733-12740.
- [30] DOERSCH C, SINGH S, GUPTA A, et al. What Makes Paris Look Like Paris?[J]. ACM Transactions on Graphics, 2012, 31(4): 1011-1019.