



融合扩散模型技术的文物面部三维模型孔洞修补

王峥嵘^{1,2}, 刘鑫达^{1,2}, 周明全^{1,2}

(1. 西北大学 文化遗产数字化国家地方联合工程研究中心, 陕西 西安 710127;

2. 西北大学 可视化技术研究所, 陕西 西安 710127)

摘要 在三维几何建模领域,孔洞指三维模型表面的不完整性,通常产生在数据采集、处理或转换过程中,这些孔洞的存在可能会对模型的几何完整性和视觉质量造成显著影响。在文化遗产保护领域,传统的孔洞修复技术侧重于几何表面的修复,而忽略了纹理信息的修复,纹理信息对于恢复文物的真实感和材质特性至关重要。文中提出了一种结合几何和纹理修复的扩散模型方法,特别适用于文物面部模型的修复。由于缺乏文物面部模型的基准真值,使用相似的残缺黄种人面部模型进行评估,鉴于公开可用的黄种人面部数据集较为稀缺,构建了一个黄种人面部数据集,包含约 20 000 张黄种人面部图像、对应的三维模型及其纹理,以及渲染得到的高保真面部图像。首先,在构建的数据集上进行测试,以验证所提方法的有效性。结果显示,与现有基准方法相比,该方法在面部图像修复方面取得了显著改进,并且实现了对面部三维模型几何表面和纹理信息的同步修复。其次,将所提出的修复技术应用于文化遗产中的文物面部模型,通过实验验证该方法在文物面部模型的几何和纹理修复方面均展现出了优异的性能。研究结果为文物面部模型的数字化修复提供了一种有效的技术手段,有助于提高文物保护和数字化展示的质量。

关键词 三维模型孔洞修补;文物面部修复;扩散模型;图像修复

中图分类号: TP391 **DOI:** 10.16152/j.cnki.xdxbzr.2025-01-018

3D model hole filling in cultural relics' facial reconstruction fusion diffusion model

WANG Zhengrong^{1,2}, LIU Xinda^{1,2}, ZHOU Mingquan^{1,2}

(1. National and Local Joint Engineering Research Center for Cultural Heritage Digitization,

Northwest University, Xi'an 710127, China;

2. Institute of Visualization Technology, Northwest University, Xi'an 710127, China)

Abstract In the field of 3D geometric modeling, a hole refers to the incompleteness on the surface of a 3D model, often arising during data acquisition, processing, or transformation. The presence of such holes can significantly impact the geometric integrity and visual quality of the model. In the domain of cultural heritage

收稿日期: 2024-09-24

基金项目: 国家自然科学基金(62271393);文化和旅游部重点实验室项目(1222000812, cr2021K01);西安市社会发展科技创新示范项目(2024JH-CXSf-0014);虚拟现实技术与系统全国重点实验室(北京航空航天大学)开放基金(VRLAB2024C02)。

第一作者: 王峥嵘,男,从事图像处理研究,2018111136@stumail.nwu.edu.cn。

通信作者: 周明全,男,教授,博士生导师,从事智能信息处理、虚拟现实与可视化研究, mqzhou@bnu.edu.cn。

preservation, traditional hole-filling techniques focus on the repair of geometric surfaces, neglecting the restoration of texture information, which is crucial for recovering the authenticity and material characteristics of artifacts. This paper proposes a diffusion model method that combines geometric and texture repair, particularly suitable for the restoration of facial models of cultural relics. Due to the absence of ground truth for facial models of cultural relics, this study uses similar incomplete Asian facial models for evaluation. Given the scarcity of publicly available Asian facial datasets, this paper constructs an Asian facial dataset comprising approximately 20 000 Asian facial images, corresponding 3D models and textures, as well as high-fidelity facial images rendered from the models. Experiments are first conducted on the constructed dataset to validate the effectiveness of the proposed method. The results show that compared to existing benchmark methods, our method achieves significant improvements in facial image inpainting and simultaneously repairs the geometric surfaces and texture information of the facial 3D models. Subsequently, the proposed repair technique is applied to facial models of cultural relics, and experiments confirm its excellent performance in both geometric and texture repair of the relic facial models. These findings provide an effective technical means for the digital restoration of facial models of cultural relics, contributing to the enhancement of the quality of cultural heritage preservation and digital exhibition.

Keywords 3D model hole repair; facial models of cultural relics; diffusion model; image inpainting

在计算机图形学和计算机视觉的研究与应用中,三维模型构建技术扮演着重要角色,它通过使用多边形(通常是三角形)来表示三维物体的表面。由于其易于理解、表达能力强,并且能够通过纹理映射和材质系统实现真实感渲染,三维模型被广泛应用于工业设计、游戏开发、文物保护等领域。对于文物而言,由于自然腐蚀、环境变化、物理损失等原因,它们通常并不完整,尤其是面部常常出现缺损。在对文物面部进行建模时,由于其本身存在的缺失以及在扫描、建模、编辑和转换过程中可能出现的问题,可能导致得到的三维模型不完整。因此,文物面部模型的修复是文物保护领域的一个重要问题,同时也是三维模型孔洞修补技术在文物保护领域的一个重要实践。

传统的三维模型孔洞修补主要分为 3 类,基于网格的三维模型孔洞修补、基于体素的三维模型孔洞修补和基于曲面定向的三维模型孔洞修补。基于网格的修补方法通常先识别出网格中的孔洞边界,再通过三角剖分在边界内生成新的三角面片填补孔洞。例如, Liepa 等人通过识别孔洞边界、孔洞三角剖分填充孔洞、细化和平滑处理修复三维模型孔洞^[1]。温佩芝等人通过检测缺失区域特征,将这些特征结合无约束的三角剖分填充孔洞^[2]。谢倩茹等人识别孔洞边界后使用波前法和三角形顶点的夹角关系完成孔洞的初始填充,再结合曲率标准对孔洞网格进行细化使填补区域过渡更自然^[3]。基于网格的三维模型孔

洞修补只考虑修补孔洞而不考虑孔洞区域的特征,导致修补区域过渡不自然,即使加入曲率信息但效果有限。基于体素的三维模型孔洞修补方法首先将网格体素化表示,而孔洞转换后变现为体素中缺失或不连续的部分,通过识别并填补孔洞区域的体素以达到修复三维模型的目的。例如, Davis 等人用体素定义了一个符号距离函数描述邻近表面,后将其扩散到整个体素网格以修复孔洞^[4]。Bischoff 等人建立一个自适应的八叉树,利用体素网格的拓扑简单性,重建三维模型^[5]。基于体的三维模型孔洞修补需要三维网格模型与体素模型相互转换,转换中必然会出现模型细节的损失,使得重建模型出现不自然的情况。基于曲面定向的三维模型孔洞修补方法在识别孔洞边界后,根据边界点的法向量信息来估计和构建内部面片的法向。例如, Wang 等人提出一种保持 CAD 模型孔洞边缘特征的方法^[6]。刘震等人利用孔洞区域的特征线,提出了一种能有效恢复孔洞区域细节的修补方法^[7]。基于曲面定向的三维模型孔洞修补过于依赖孔洞区域的特征信息,如果孔洞区域特征不明显或者拓扑结构过于复杂,会直接影响到模型修复效果。

上述 3 种传统的三维模型孔洞修补方法各有优劣,但它们都只注重于对三维模型孔洞区域的修复而忽略了对三维模型纹理图像的修复。近期针对三维模型孔洞修复的研究虽然在孔洞修复的精确性上得到了提升,但并未脱胎出传统方法的

范畴,依然是对传统方法的优化,无法实现纹理信息的修复。如王鹏飞等人提出一种保持输入网格特征的鲁棒水密流形网格修复算法^[8]:首先,使用 Manifoldplus 算法和卷绕数创建一个近似输入网格的防水流形引导曲面,以区分网格内外;其次,计算受限 Voronoi 图(RVD)和受限三角剖分(RDT),将非流形问题分解到 RVD 和 RDT 的计算中,确保得到的 RDT 是修复后的防水流形网格;最后,在原始网格边添加辅助点,以保持原始网格特征。该方法依旧忽略了纹理信息的修复。文物模型的纹理图像对于增加文物模型的历史真实性,提升模型的观赏价值有至关重要的作用,因此,在修复文物模型的同时,不能忽视对其纹理图像的修复。针对前述问题,本研究提出了一种文物面部三维模型孔洞修复的方法,融合了基于扩散模型的(diffusion model,简称 DM)图像修复技术与单目人脸重建技术,实现了在修复文物面部三维模型的同时,对纹理图像进行同步修复。

1 基于 DM 的文物面部模型修补

1.1 基于 DM 的图像修复

扩散模型是一种基于概率分布的生成模型,常被用来生成数据,尤其是生成图像数据。其原理如图 1 所示,首先,通过一个可控的过程逐步引入噪声直至数据完全转换为随机噪声;其次,再逆转这个过程,从随机噪声状态逐步去噪直至恢复数据。近些年的研究已经证明,相较于生成对抗网络 GAN、变分自编码器 VAE、标准化流 NF 等生成模型,扩散模型能够生成的图像,细节丰富,质量更高^[9-12]。

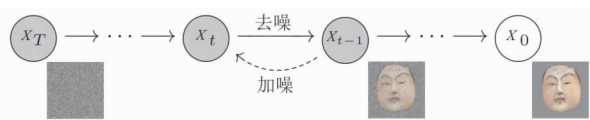


图 1 扩散模型过程

Fig. 1 Process of diffusion model

由于扩散模型在生成高质量图像方面出色的表现,越来越多研究者将扩散模型应用于图像修复,使用其逆向过程的预测与生成能力修复图像指定部分。Lugmayr 等人首先提出了一种利用去噪扩散概率模型 DDPM 的图像修复方法 RePaint^[13],该方法利用深度神经网络在大量同类图像的数据集训练以学习缺失位置像素与周围像素

的关联,并以此为先验加入到图像生成过程中,从而达到图像修复的效果。Rout 等人对 RePaint 进行测试发现,通过 RePaint 恢复的图像数据分布与真实图像数据分布存在偏差^[14],在数据分布服从简单分布时尤为明显,如线性均匀分布,为解决 RePaint 的偏差问题,Rout 等人提出了一种 RePaint + 方法,通过校正去噪过程中的漂移和分散对齐来消除偏差,使恢复图像的数据分布更接近于真实图像数据分布,并提出了使用训练好的扩散模型就可以完成图像修复任务。随后,Zhang 等研究者在文献中通过实验观察到 RePaint 及其改进版 RePaint + 在图像修复任务中^[15],修复区域与原有图像的过渡往往显得不够自然。为了解决这一问题,Zhang 等人提出了一种基于贝叶斯理论的修复框架。在该框架中,去噪过程被设计为逐步调整图像的已知区域与未知区域,同时引入了一个近似的后验概率分布。这种分布机制使得修复过程中的误差能够逐渐减少,从而有效缓解了修复区域与原有图像区域之间的不协调问题,该方法被命名为 CoPaint。随着学者对扩散模型的不断深入研究,图像修复的速度与质量在不断提升,应用的领域也不断变广,Yue 等人提出了一种高效的扩散模型^[16],通过残差转移和灵活的噪声调度,在显著减少推理步骤的同时,实现了更优的图像修复性能。Corneanu 等人提出一种改进方法,不需要领域特定的训练,同时通过在潜在空间而非图像空间中进行前向-后向融合步骤,提升了推理时的计算速度^[17],并引入一个新提出的传播模块来解决计算成本问题。Durrer 等人通过改进去噪扩散概率模型 DDPM,实现了对含有病理组织的 3D 脑部磁共振图像中健康组织的一致性修复^[18],并在下游的脑组织分割任务中超越了现有的 FMRIB 软件库(FSL)病变填充方法。由于扩散模型在图像修复上的强大能力,本文选用其承担图像修复任务。

为了在修复人脸三维模型的同时恢复其纹理,以确保渲染后的模型能够呈现出逼真的人脸效果,本研究引入了扩散模型,提出的修复人脸三维模型孔洞的方法如图 2 所示,分为 3 个阶段。首先,对残缺的人脸三维模型进行渲染得到残缺人脸图像;其次,利用扩散模型的生成能力,对渲染得到的残缺人脸图像进行修复;最后,将修复后的人脸图像通过单目人脸重建技术重建成完整的人脸三维模型,同时得到完整的纹理图像。

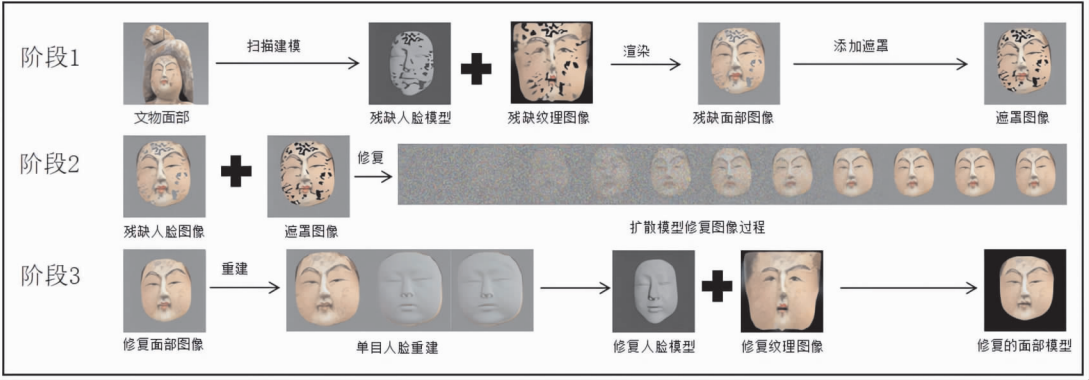


图 2 文物面部模型修复流程图

Fig. 2 Process of repairing cultural relic facial models

1.2 基于 DM 的文物面部图像修复

1.2.1 DM 作为生成去噪器先验

去噪扩散隐式模型 DDIM^[19]与图像修复算法 CoPaint 证明了使用扩散模型进行图像修复可以在去噪过程中直接使用预训练的去噪扩散概率模型 DDPM,而扩散模型的强大的生成能力以及去噪能力也已被证实^[20]。因此,本文使用 OpenAI 公开的 Guided Diffusion 预训练模型作为图像修复过程中的生成去噪器。

在图像修复中通常需要考虑优化问题,即在恢复过程中的去噪或重建问题,目标是恢复出一副高质量的图像,使其既要符合观测到的噪声以保证数据的一致性,又要符合扩散模型提供的图像先验知识,在数学上这个问题可以表述为一个最大后验概率(MAP)问题,其一般形式表示为

$$\hat{x} = \operatorname{argmin}_x \left\{ \frac{1}{2\sigma_n^2} \|y - H(x)\|^2 + \lambda P(x) \right\} \quad (1)$$

式中: x 是要恢复的干净图像; y 是给定的噪声观测数据; $H(x)$ 是退化操作,描述了如何从干净图像 x 生成观测数据 y ; σ_n^2 是噪声的方差; $P(x)$ 是图像 x 的先验项,通常反映了对图像特性的某些假设,例如平滑性或稀疏性; λ 是正则化参数,用于平衡数据项和先验项之间的权重。

为了简化并解决优化问题,半二次分裂(HQS)算法作为一种迭代算法,常被引入图像恢复相关的任务中解决最大后验概率(MAP)问题^[21],HQS 算法将优化问题分为处理数据一致性的数据子问题和处理图像先验的先验子问题,并在图像恢复中交替处理两个问题,其一般形式可以表示为式(2)和式(3)。

$$z_k = \operatorname{argmin}_z \frac{1}{2(\sqrt{\lambda/\mu})^2} \|z - x_k\|^2 + P(z) \quad (2)$$

$$x_{k-1} = \operatorname{argmin}_x \|y - H(x)\|^2 + \mu^2 \sigma_n^2 \|x - z_k\|^2 \quad (3)$$

式中: z_k 是要在优化中更新的辅助变量; x_k 是当前迭代中对图像的估计。

其次,有研究者对 HQS 算法在解决数据子问题和先验子问题的方法上进行了改进^[22],其一般形式可以表示为式(4)和式(5)。

$$x_0^{(t)} = \operatorname{argmin}_z \frac{1}{2\sigma_t^2} \|z - x_k\|^2 + P(z) \quad (4)$$

$$\hat{x}_0^{(t)} = \operatorname{argmin}_x \|y - H(x)\|^2 + \rho_t \|x - x_0^{(t)}\|^2 \quad (5)$$

式中: $\hat{x}_0^{(t)}$ 是在时间步 t 下更新的图像估计; σ_t 是时间步 t 下的噪声方差; ρ_t 是与时间步相关的正则化参数, $\rho_t = \lambda(\sigma_n/\sigma_t)^2$; $x_0^{(t)}$ 是扩散模型预测的当前时间步的去噪图像状态。

原始 HQS 算法中的 z_k 是通过解决式(2)的先验子问题更新得到的,而改进的 HQS 算法中引入辅助变量 $x_0^{(t)}$,代替了原始的 HQS 算法中的辅助变量 z_k ,如图 3 所示。辅助变量 $x_0^{(t)}$ 如式(4)所示,通过扩散模型预测时间步 t 下的图像去噪后的图像状态得到,这样不仅将扩散模型作为恢复图片过程中的生成去噪器,还将其应用于去噪过程中的相邻时间步之间作为生成去噪器先验,优化了解决先验子问题的方法。随后,又将 $x_0^{(t)}$ 作为连接数据与先验的桥梁带入数据子问题中更新在时间步 t 下的图像估计 $\hat{x}_0^{(t)}$ 。其中最小化数据拟合项确保恢复图像 x 在退化后与观测数据一致,正则化项 $\rho_t \|x - x_0^{(t)}\|^2$ 确保了图像估计 $\hat{x}_0^{(t)}$ 与扩散模型预测的去噪图像 $x_0^{(t)}$ 保持接近,以优化解决数据子问题的方法。本文便采用了改进的 HQS 算法,将扩散模型作为生成去噪器先验进一步的应用于去噪过程相邻时间步之间,以更好地解决优化问题,提高恢复图像的质量和细节。

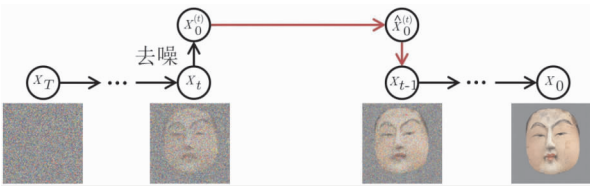


图 3 优化问题解决方法

Fig. 3 Optimization problem solving method

1.2.2 逆扩散采样循环

逆扩散采样循环就是扩散模型结合其去噪能力和生成能力将噪声数据逐步恢复成干净图像的过程,对于图像修复问题则是将扩散模型作为生成去噪器填补图像缺失部分。

如图 4 所示,在这一过程的开始需要选择一个噪声水平较高的初始化图像 x_T , T 表示总时间步长。为解决优化问题采用改进的 HQS 算法,用当前的噪声图像 x_t 预测 $x_0^{(t)}$ 解决先验子问题。利用得到的 $x_0^{(t)}$ 估计 $\hat{x}_0^{(t)}$ 解决数据子问题,在实际的图像修复过程中为将修复位置指定在缺失部分,如式(6)和式(7)所示。



图 4 逆扩散采样循环

Fig. 4 Reverse diffusion sampling loop

$$y = \mathbf{M} \odot x \quad (6)$$

$$x_0 = \frac{\mathbf{M} \odot y + \rho_t z_0}{\mathbf{M} + \rho_t} \quad (7)$$

式中: $\mathbf{M}$ 是任何用户定义的掩码,是一个带有布尔元素的矩阵; $\odot$ 表示元素级乘法。图像的绘制任务是从已知的像素中恢复缺失的像素为 y 。

由 $\hat{x}_0^{(t)}$ 即可生成下一个时间步 $t-1$ 的噪声图像 x_{t-1} , 其公式为

$$x^{(t-1)} = \sqrt{\bar{a}_{t-1}} \hat{x}_0^{(t)} + \sqrt{1 - \bar{a}_{t-1}} (\sqrt{1 - \zeta} \hat{\epsilon} + \sqrt{\zeta} \epsilon_t) \quad (8)$$

式中: $\bar{a}_{t-1}$ 是时间步 $t-1$ 的权重参数; ζ 是控制每一步中注入噪声方差的超参数; ϵ_t 是从标准正态分布中采样的噪声; $\hat{\epsilon}$ 为结合当前图像 x_t 以及数据子问题的解 $\hat{x}_0^{(t)}$ 得到的预测噪声,其公式为

$$\hat{\epsilon} = \frac{1}{\sqrt{1 - \bar{a}_t}} (x_t - \sqrt{\bar{a}_t} \hat{x}_0^{(t)}) \quad (9)$$

简单来说,本文在相邻时间步噪声图像间采用的生成去噪方法为从当前时间步 t 的噪声图像解决先验子问题得到,再通过解决数据子问题得到,

最后通过计算下一个时间步的噪声图像,循环这个过程直至算法收敛得到高质量的图像修复效果。

1.3 单目文物面部重建

高保真的单目人脸重建技术已经取得了巨大的成功,有大量工作使用专门的硬件(如 LightStage)^[23-25] 取得了出色的重建性能。基于三维可变形模型(3DMM)的人脸重建技术^[8] 的出现,推出了一种从一系列面部扫描中学习到的统计模型,并引起了一系列对单目人脸重建技术的研究^[26-29]。本文直接采用了分层表示网络(HRN)模型^[30] 中的单目人脸重建部分对经过扩散模型修复后的渲染图像进行人脸重建,该模型对 3DMM 在低维表示能力的局限性进行了改进,对单个图像进行更详细和准确的重建。

2 构建数据集

由于文物面部模型缺乏基准真值验证修复性能,本文选择在与文物面部具有相似特征的黄种人人脸模型上进行测试以评估本文方法的修复能力,但目前的公开数据集以白种人的脸型模型为主,缺乏黄种人人脸模型数据,因此,本文选择自建黄种人人脸数据集。建立数据集需要使用黄种人的脸型图像建立人脸模型,目前并没有开源的高质量黄种人人脸数据集可供使用,高质量黄种人的脸型数据十分稀缺。但 Nvidia 公司团队于 2019 年提出的 StyleGAN 模型^[31],能够生成高清并且逼真的人脸照片,为获得大量的高质量黄种人的脸型图像建立数据集提供了条件。本文建立的黄种人人脸数据集如图 5 所示,包含 3 个部分。

1) 通过 StyleGAN 模型根据黄种人的特征生成 20 271 张高质量黄种人人脸图像。

2) 通过 HRN 模型对黄种人人脸图像进行人脸重建得到 20 271 个人脸三维模型及其纹理图像,对人脸三维模型渲染得到 20 271 张高保真人脸图像。

3) 对完整人脸三维模型随机破坏得到 20 271 个残缺人脸三维模型及其对应的残缺纹理图像,对残缺人脸三维模型渲染得到 20 271 张高保真人脸图像。

本文中残缺人脸三维模型及残缺纹理图像数据集用于渲染成为残缺人脸图像后,经过图像修复,重建成完整的人脸三维模型。完整人脸三维模型及纹理图像数据集和其渲染得到的高保真人

脸图像,分别用于与重建得到的修复人脸三维模型、修复的纹理图像、残缺人脸图像进行比较以分析修复效果。



图 5 黄种人人脸数据集部分展示

Fig. 5 Part of yellow face dataset

3 实验及结果分析

3.1 预训练模型及实验设置

由于已经证明对于图像修复任务可以直接使用预训练的扩散模型作为生成去噪器,因此本文使用了 OpenAI 公开的预训练模型 256×256 diffusion(not class conditional)。该模型在斯坦福大学建立的 ILSVRC 2012 ImageNet 子集上训练得到,该数据集是在 2012 年策划的,由大约一百万张图像组成,每张图像属于 1 000 个类别中的一个。该模型已在大规模的、多样化的数据集上训练过,能够更好地泛化到新的、未见过的数据。使用该预训练的扩散模型大大减少了所需的时间和计算资源,有效地利用这些模型在图像生成和去噪方面的强大能力,从而提高图像恢复任务的性能。

实验中 σ 被设置为大于 0.001,该值取决于

图像中添加的高斯白噪声的水平通常在 $0 \sim 1$ 之间。为了探究参数 ζ 对噪声控制以及参数 λ 对正则化项的影响,图 6 展示了在不同 ζ 和 λ 组合下的图像修复效果。实验结果表明,随着 ζ 值的增加,图像细节的修复质量得到提升,这体现在牙齿的修复完整性以及修复位置遮罩的修复程度上(图中面部的黑色区域表示遮罩未能完全修复)。在 ζ 值保持不变的情况下,当 λ 设置为 1.0 和 10 时,峰值信噪比(PSNR)的数值略高于 λ 为 0.1、100、1 000 和 10 000 的情况。因此,本研究选择将 ζ 设为 1.0, λ 设为 1.0 作为实验参数。逆扩散采样循环的迭代次数设置为 20 次,即生成去噪过程最多进行 20 次。



图 6 ζ 与 λ 对图像修复效果的影响

Fig. 6 The influence of ζ and λ on the image restoration effect

3.2 实验结果分析

3.2.1 人脸图像修复效果

本文实验以弗雷谢起始距离(FID)、峰值信噪比(PSNR)、感知损失(LPIPS)3种指标作为评价指标,比较本实验中使用的图像修复方法与 RePaint 方法在修复渲染得到的残缺人脸图像效果。FID 是一种衡量两组图像分布差异的指标,适用于生成图像与真实图像之间的差异,值越低表示生成图像与真实图像之间的视觉特征越相近;PSNR 是一种衡量重建或去噪后的图像与原始图像相似性的指标,值越高则相似性越高;LPIPS 用于评估图像之间的感知相似性,值越低则两幅图像在感知上越相似。图像修复效果如图 7 所示,评估结果如表 1 所示。

实验结果显示,相较于作为基准的 RePaint 方法,本文方法在 FID 指标上从 27.867 4 下降到

了 23.294 0,表明本文方法修复的图像与真实图像的分布更加接近;PSNR 指标从 33.315 6 提升到了 33.618 7,表明本文方法修复的图像与原始图像的相似度略高于 RePaint 方法修复的图像;LPIPS 指标由 0.021 3 降到了 0.020 1,表明本文方法修复的图像与原始图像在视觉上的相似度也略高于 RePaint 方法修复的图像。总体而言,本文方法在图像修复性能上优于 RePaint 方法,尤其是在修复图像与真实图像分布的相似性上,FID 指标上提升了约 16.42%。

表 1 人脸图像修复效果对比

Tab. 1 Comparison of facial image restoration effects			
修复方法	FID	PSNR	LPIPS
RePaint ^[13]	27.867 4	33.315 6	0.021 3
本文方法	23.294 0	33.618 7	0.020 1

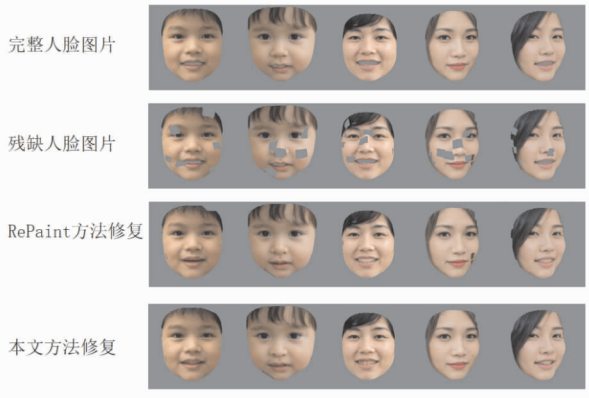


图 7 人脸图像修复效果对比

Fig. 7 Comparison of facial image restoration effects

3.2.2 人脸三维模型修复效果

在本研究的实验部分,对修复后的人脸三维模型进行了评估,并将其与现有技术进行了比较。具体而言,将修复结果与 Liepa 提出的基于网格的修补方法^[1]、Davis 等人提出的基于体素的修补方法^[4],以及 Wang 等人提出的将修补对象转换为网格的基于曲面定向的孔洞修补方法^[6]所得的修复效果进行了对比分析。评估指标包括豪斯多夫距离 (Hausdorff distance)、均方误差 (mean squared error, MSE) 和曲面法线差异 (surface normal difference, SND)。豪斯多夫距离用于衡量两组点集间的相似度,通过计算两模型间最远点对的距离来评估几何形状的相似度,其值越低表示几何形状的最大偏差越小。MSE 通过计算对应点间距离的平方和平均值来评估模型表面细节的相似度,其值越低表示整体形状上的平均位置误

差越小。SND 则衡量两模型表面法线间的角度差异,反映表面细节的一致性,其值越低表示表面方向的一致性越好。4 种人脸三维模型的修复效果如图 8 所示,相应的评估结果汇总于表 2。

表 2 人脸模型修复效果对比

Tab. 2 Comparison of facial model restoration effects			
修复方法	Hausdorff	MSE	SND
基于网格 ^[1]	0.141 5	0.070 3	29.486 6
基于体素 ^[4]	21.324 6	71.833 3	93.153 5
基于曲面定向 ^[6]	0.141 3	0.070 2	30.304 5
RePaint ^[13]	0.177 5	0.013 2	3.257 5
本文方法	0.177 1	0.013 0	3.257 2

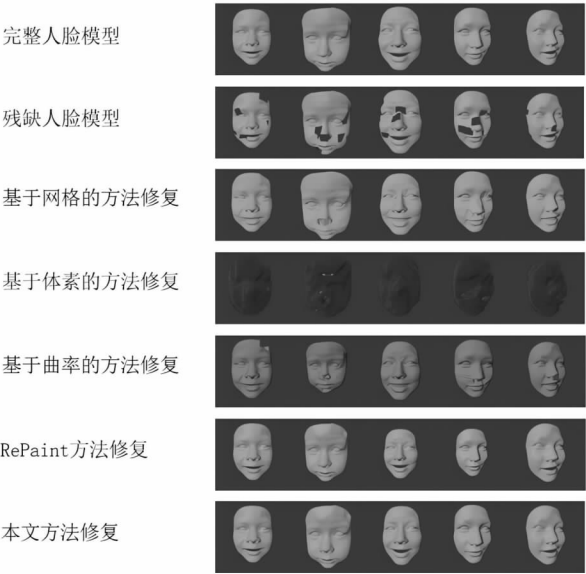


图 8 人脸模型修复效果对比

Fig. 8 Comparison of facial model restoration effects

实验结果揭示了基于体素的三维模型孔洞修复技术相较于其他 4 种修复技术在修复效果上的不足。具体而言,本研究提出的修复方法在 Hausdorff 距离指标上为 0.177 1,略高于基于网格和基于曲率的修复方法。结果表明,在修复模型与原始模型之间的最大偏差方面,本文方法修复的模型与其他两种方法相比略有增加。该差异可能源于本文方法采用的是重建而非直接在原始模型上进行修复的策略。此外,本文方法在 MSE 指标上达到了 0.013 0,在 SND 上达到了 3.257 2,表明在修复模型与原始模型的整体一致性以及表面细节的保持上,本文方法明显优于基于网格和基于曲率的修复方法。与 RePaint 方法相比,本文方法在所有 3 项指标上均展现出更优的修复效

果。在视觉质量方面,本文方法修复的人脸模型在认知上更为自然,而基于网格和基于曲率的修复方法虽然能够填补孔洞,但在复杂面部特征的还原上存在不足。总体来看,尽管在 Hausdorff 距离指标上略低于基于网格和基于曲率的修复方法,本文方法在 MSE、SSIM 指标以及实际视觉效果上均优于其他比较方法。因此,可以得出本研究提出的修复方法在面部三维模型的修复能力上,尤其是在处理复杂的面部细节和大面积缺失的面部模型方面,明显优于传统的修复方法,并且也超越了 RePaint 方法。

本文提出三维模型孔洞修补方法,在通过单目人脸重建修复人脸三维模型的同时,也能修复对应的纹理图像,修复效果如图 9 所示,并以弗雷谢起始距离(FID)、峰值信噪比(PSNR)、感知损失(LPIPS)3 种指标比较了完整纹理图像与修复后的纹理图像,结果如表 3 所示。

表 3 纹理图像修复效果对比
Tab. 3 Comparison of texture repair effects

修复方法	FID	PSNR	LPIPS
RePaint ^[13]	13. 449 0	21. 584 6	0. 070 3
本文方法	13. 656 0	21. 575 7	0. 070 3



图 9 纹理修复效果及渲染效果

Fig. 9 Texture repair effects and rendering effects

实验结果表明,本研究提出的方法在纹理图像修复方面表现出色。该方法在 FID 指标上为 13. 656 0,略低于 RePaint 方法,结果表明本文方法修复的纹理图像在统计分布上与原始纹理图像非常接近。此外,本文方法在 PSNR 指标上为 21. 575 7,高于 RePaint 方法,进一步证实了本文方法修复的纹理图像与原始纹理图像在像素级别

的相似度较高。在 LPIPS 指标上,本文方法与 RePaint 方法得分相同,均为 0. 070 3,表明在视觉感知层面,本文方法修复的纹理图像与原始纹理图像具有较高的相似性。

综合考虑面部模型的修复效果以及纹理图像的修复效果,可以得出本文提出的三维模型孔洞修补方法在修复三维模型的纹理图像方面同样取得了显著的成效。这些结果不仅验证了本文方法在纹理图像修复方面的有效性,也进一步证明了其在三维模型修复领域的应用潜力。

3.2.3 文物面部三维模型修复效果

实验证明,本文方法对于黄种人人脸模型的修复取得了较好的结果,并且实现了对三维模型纹理的修复,表明了基于扩散模型的图像修复可以应用到对人脸三维模型及其纹理的修复上。对于文物面部来说,由于文物长时间的掩埋在地下,其表面通常会出现脱落、缺失或者黏结泥土等杂质的现象,导致在扫描重建面部模型时出现缺损的情况,使用本文提出的基于扩散模型的人脸模型修复方法对文物面部进行修复,能够在对文物模型孔洞修补的同时对斑驳的纹理图像进行修复,并且扩散模型优秀的预测能力与泛化特性能够使修复图像保持文物面部纹理的特征,同时应用于不同朝代不同特征的文物面部模型修复上,并且取得了不错的效果,如图 10 所示。



图 10 文物面部模型修复效果

Fig. 10 Restoration effect of cultural relic facial model

4 结语

1)建立了内容丰富的黄种人人脸数据集,数据集中包含了高质量的黄种人人脸图像,人脸图像对应的人脸三维模型及其纹理图像,人类三维模型渲染得到的高保真人脸图像。

2)提出了一种基于扩散模型的文物面部三维模型孔洞修补方法,由于缺乏文物面部三维模型的基准真值,该方法以在黄种人人脸模型上的修复效果作为评估。首先,将残缺的人脸三维模型渲染成为残缺的人脸图像;其次,通过扩散模型修复残缺人脸图像得到修复完整的人脸图像;最后,对该修复图像进行单目人脸重建得到修复后的人脸模型。在人脸图像修复方面,相较于 Re-Paint 方法,本文采用的改进的扩散模型图像修复方法在人脸图像修复上取得了更好的结果。在三维模型修复方面,相较于基于网格的三维模型孔洞修补方法、基于体素的三维模型孔洞修补方法和基于曲面定向的三维模型孔洞修补方法3种传统三维模型修复方法,本文的方法在面部三维模型修复效果上取得了更出色的结果,修复效果更加自然。同时解决了传统修复方法只能修复三维模型本身而无法修复纹理图像的缺陷,能够在修复三维模型时兼顾修复纹理图像,并将其应用与文物面部模型的修复上,修复了文物的面部模型及其纹理,增加了文物面部模型的真实性以及表现能力。

参考文献

- [1] LIEPA P. Filling holes in meshes[C]//In Proceedings of the 2003 Eurographics/ACM SIGGRAPH Symposium on Geometry Processing. Aachen: ACM, 2003: 200-205.
- [2] 温佩芝,雷永庆,孙梦龙. 三维重建网格模型的缺陷孔洞识别与修复方法[J]. 计算机应用研究, 2020, 37(4):1234-1238.
WEN P Z, LEI Y Q, SUN M L. Defect and hole identification and repair methods for 3D reconstructed mesh models[J]. Application Research of Computers, 2020, 37(4):1234-1238.
- [3] 谢倩茹,耿国华. 三维模型孔洞修补方法的研究[J]. 计算机应用研究, 2013, 30(10): 3175-3177.
XIE Q R, GENG G H. Research on the method of repairing holes in 3D models[J]. Application Research of Computers, 2013, 30(10): 3175-3177.
- [4] DAVIS J, MARSCHNER S R, GARR M, et al. Filling holes in complex surfaces using volumetric diffusion[C]//In Proceedings of First International Symposium on 3D Data Processing Visualization and Transmission. Padua: IEEE, 2002: 428-441.
- [5] BISCHOFF S, KOBELT L. Structure preserving CAD model repair[J]. Computer Graphics Forum, 2005, 24(3): 527-536.
- [6] WANG X, LIU X, LU L, et al. Automatic hole-filling of CAD models with feature-preserving[J]. Computers & Graphics, 2012, 36(2): 101-110.
- [7] 刘震,王艳宾,白丽丽,等. 曲面细节特征保持的三维模型孔洞修复方法[J]. 计算机辅助设计与图形学学报, 2016, 28(12): 2052-2059.
LIU Z, WANG Y B, BAI L L, et al. 3D model hole repair method for preserving surface detail features[J]. Journal of Computer Aided Design and Graphics, 2016, 28(12): 2052-2059.
- [8] 王鹏飞,徐敏峰,辛士庆,等. 鲁棒的水密流形网格修复[J]. 计算机辅助设计与图形学学报, 2024, 36(7): 1047-1056.
WANG P F, XU M F, XIN S Q, et al. Robust watertight manifold mesh repair[J]. Journal of Computer-Aided Design and Computer Graphics, 2024, 36(7): 1047-1056.
- [9] HO J, JAIN A, ABBEEL P. Denoising diffusion probabilistic models[J]. Advances in Neural Information Processing Systems, 2020, 33: 6840-6851.
- [10] DHARIWAL P, NICHOL A. Diffusion models beat GANs on image synthesis[J]. Advances in Neural Information Processing Systems, 2021, 34: 8780-8794.
- [11] RAMESH A, DHARIWAL P, NICHOL A, et al. Hierarchical text-conditional image generation with CLIP latents[EB/OL]. (2022-04-13)[2024-07-02]. <https://arxiv.org/abs/2204.06125v1>.
- [12] 彭进业,余喆,屈书毅,等. 基于深度学习的图像修复方法研究综述[J]. 西北大学学报(自然科学版), 2023, 53(6): 943-963.
PENG J Y, YU Z, QU S Y, et al. A survey of image inpainting methods based on deep learning[J]. Journal of Northwest University (Natural Science Edition), 2023, 53(6): 943-963.
- [13] LUGMAYR A, DANELLJAN M, ROMERO A, et al. RePaint: Inpainting using denoising diffusion probabilistic models[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans: IEEE, 2022: 11451-11461.
- [14] ROUT L, PARULEKAR A, CARAMANIS C, et al. A

- theoretical justification for image inpainting using denoising diffusion probabilistic models [EB/OL]. (2023-02-02) [2024-07-02]. <https://arxiv.org/abs/2302.01217v1>.
- [15] ZHANG G H, JI J B, ZHANG Y, et al. Towards coherent image inpainting using denoising diffusion implicit models[C] // Proceedings of the 40th International Conference on Machine Learning. Honolulu: ACM, 2023: 41164-41193.
- [16] YUE Z S, WANG J Y, LOY C C. Efficient diffusion model for image restoration by residual shifting[EB/OL]. (2024-03-12) [2024-07-02]. <https://arxiv.org/abs/2403.07319v3>.
- [17] CORNEANU C, GADDE R, MARTINEZ A M. LatentPaint: Image inpainting in latent space with diffusion models[C] // 2024 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). Waikoloa: IEEE, 2024: 4322-4331.
- [18] DURRER A, WOLLEB J, BIEDER F, et al. Denoising diffusion models for 3D healthy brain tissue inpainting[C] // MICCAI Workshop on Deep Generative Models. Cham: Springer Nature Switzerland, 2024: 87-97.
- [19] SONG J M, MENG C L, ERMON S. Denoising diffusion implicit models[EB/OL]. (2020-10-06) [2024-07-02]. <https://arxiv.org/abs/2010.02502v4>.
- [20] DEJA K, KUZINA A, TRZCINSKI T, et al. On analyzing generative and denoising capabilities of diffusion-based deep generative models[EB/OL]. (2022-05-31) [2024-07-02]. <https://arxiv.org/abs/2206.00070v1>.
- [21] ZHANG K, LI Y W, ZUO W M, et al. Plug-and-play image restoration with deep denoiser prior[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(10): 6360-6376.
- [22] ZHU Y Z, ZHANG K, LIANG J Y, et al. Denoising diffusion models for plug-and-play image restoration[C] // 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Vancouver: IEEE, 2023: 1219-1229.
- [23] ALEXANDER O, ROGERS M, LAMBETH W, et al. The digital emily project: Photoreal facial modeling and animation[C] // ACM SIGGRAPH 2009 Courses. New Orleans: ACM, 2009: 1-15.
- [24] CAO X, CHEN Z, CHEN A P, et al. Sparse photometric 3D face reconstruction guided by morphable models[C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 4635-4644.
- [25] MA W C, HAWKINS T, PEERS P, et al. Rapid acquisition of specular and diffuse normal maps from polarized spherical gradient illumination[C] // Proceedings of the 18th Eurographics Conference on Rendering Techniques. Grenoble: ACM, 2007: 183-194.
- [26] DIB A, THÉBAULT C, AHN J, et al. Towards high fidelity monocular face reconstruction with rich reflectance using self-supervised learning and ray tracing[C] // 2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal: IEEE, 2021: 12799-12809.
- [27] GECER B, PLOUMPIS S, KOTSIA I, et al. GAN-FIT: Generative adversarial network fitting for high fidelity 3D face reconstruction[C] // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach: IEEE, 2019: 1155-1164.
- [28] GECER B, PLOUMPIS S, KOTSIA I, et al. Fast-GANFIT: Generative adversarial network for high fidelity 3D face reconstruction[EB/OL]. (2021-05-16) [2024-07-02]. <https://arxiv.org/abs/2105.07474v1>.
- [29] LATTAS A, MOSCHOLOU S, GECER B, et al. AvatarMe: Realistically renderable 3D facial reconstruction "In-the-wild" [C] // 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE, 2020: 760-769.
- [30] LEI B W, REN J Q, FENG M Y, et al. A hierarchical representation network for accurate and detailed face reconstruction from in-the-wild images[C] // 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver: IEEE, 2023: 394-403.
- [31] KARRAS T, LAINE S, AILA T M. A style-based generator architecture for generative adversarial networks[C] // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach: IEEE, 2019: 4401-4410.

(编辑 张欢)