

基于参考的 Transformer 纹理迁移深度 图像超分辨率重建

杨陈成¹, 董秀成^{1,2}, 侯兵¹, 张党成¹, 向贤明¹, 冯琪茗¹

(1. 西华大学电气与电子信息学院, 四川 成都 611730; 2. 四川大学锦江学院, 四川 眉山 620860)

摘 要: 深度图像包含场景深度信息, 对颜色和光照的变化具有较强的鲁棒性, 使得深度图像在立体视觉等领域广泛应用。由于深度传感器性能的局限性以及深度图像成像环境相对复杂, 很难直接获取高质量、高分辨率的深度图像。针对重建出现的边缘细节特征不清晰问题, 提出一种基于参考的 Transformer 纹理迁移深度图像超分辨率重建方法。对预处理后的低分辨率深度图像(LR_D)以及参考图像(Ref)特征块, 利用归一化内积进行相似度计算, 融合 Transformer 计算相似位置置信度, 并结合注意力机制进行纹理迁移, 最后与低分辨率深度图像特征结合, 提高图像细节清晰度, 进一步精确重建结果。实验结果表明, 相较于其他方法, 该方法结构相似性(SSIM)值更高, 主观视觉效果和客观评价指标均得到了明显的改善, 重建效果良好。

关 键 词: 深度学习; 超分辨率重建; 深度图像; Transformer; 注意力机制

中图分类号: TP 391

DOI: 10.11996/JGJ.2095-302X.2023050861

文献标识码: A

文章编号: 2095-302X(2023)05-0861-07

Reference based transformer texture migrates depth images super resolution reconstruction

YANG Chen-cheng¹, DONG Xiu-cheng^{1,2}, HOU Bing¹, ZHANG Dang-cheng¹,
XIANG Xian-ming¹, FENG Qi-ming¹

(1. School of Electrical Engineering and Electronic Information, Xihua University, Chengdu Sichuan 611730, China;

2. Jinjiang College, Sichuan University, Meishan Sichuan 620860, China)

Abstract: Depth images contain scene depth information and exhibit strong robustness to variations in color and lighting, making them widely used in fields such as stereo vision. However, due to the limitations in depth sensor performance and the complexity of imaging environments, it is challenging to directly obtain high-quality, high-resolution depth images. To address the problem of unclear edge details in reconstructed depth images, a reference-based Transformer texture transfer method for deep image super-resolution reconstruction was proposed. For the preprocessed low-resolution depth images (LR_D) and reference images (Ref) feature blocks, similarity

收稿日期: 2023-01-31; 定稿日期: 2023-05-08

Received: 31 January, 2023; Finalized: 8 May, 2023

基金项目: 国家自然科学基金项目(11872069); 四川省中央引导地方科技发展专项(2021ZYD0034); 四威高科-西华大学产学研联合实验室项目(2016-YF04-00044-JH)

Foundation items: National Natural Science Foundation of China (11872069); Central Government Funds of Guiding Local Scientific and Technological Development for Sichuan Province (2021ZYD0034); Siwei Hi-tech-Xihua University Industry-University-Research Joint Laboratory (2016-YF04-00044-JH)

第一作者: 杨陈成(1998-), 女, 硕士研究生。主要研究方向为图像处理、深度学习。E-mail: yangchencheng2017@163.com

First author: YANG Chen-cheng (1998-), master student. Her main research interests cover image processing and deep learning.

E-mail: yangchencheng2017@163.com

通信作者: 董秀成(1963-), 男, 教授, 硕士。主要研究方向为智能信息处理、计算机视觉等。E-mail: dxc136@163.com

Corresponding author: DONG Xiu-cheng (1963-), professor, master. His main research interests cover intelligent information processing, computer vision, etc. E-mail: dxc136@163.com

calculation was performed using normalized inner product. The method integrated Transformer to calculate the confidence of similarity positions, and combined it with an attention mechanism for texture transfer. Finally, the method combined the features of the low-resolution depth images to improve image detail clarity and further accurately reconstruct the results. The experimental results demonstrated that compared to other methods, the proposed method could achieve higher structural similarity (SSIM) values, and that both subjective visual effects and objective evaluation indicators have been significantly improved, indicating the excellence of the reconstruction performance.

Keywords: deep learning; super-resolution reconstruction; depth image; Transformer; attention mechanism

随着计算机视觉技术的发展,深度信息在自主导航、三维重建、虚拟现实等诸多应用中发挥着重要作用。但由于深度传感器性能的局限性以及深度图像成像环境相对复杂,很难直接获取高质量、高清晰度的深度图像。深度图像超分辨率重建方法(depth super resolution, DSR)按照实现技术手段主要分为2类:基于非学习的DSR和基于学习的DSR。

基于非学习的DSR包括:基于滤波^[1-2]的重建方法和基于优化^[3-4]的重建方法。基于滤波的DSR,文献[2]针对低分辨率深度图(low resolution depth image, LR-D)可能受到噪声的破坏,提出了一种深度图像上采样滤波器。利用深度图像的局部和非局部区域关系,逐个评估高分辨率深度值,运算复杂度较低,但生成的目标图像常出现边缘平滑、纹理特征清晰度较低等现象^[5]。基于优化的DSR,先验知识是解决深度图像重建问题不可缺少的一环。MAC AODHA等^[6]学习低分辨率补丁的高度字段,并且利用马尔科夫随机场(Markov random field, MRF)模型标记问题。文献[3]利用MRF优化重建过程,并通过边缘构建的高分辨率边缘图引导重建任务。利用深度图像退化模型和各种先验知识,将重建问题转换为代价函数最优化的问题,这种方式虽适用于多种退化模型,但运算复杂度较高,先验知识如何选取影响很大^[5]。

基于学习的DSR,利用大量的训练数据,采用深度学习等策略学习不同分辨率图像之间的映射关系,构建映射模型,此类方法重建性能较好,目前应用最广泛。DONG等^[7]成功的在超分辨率重建任务中引入深度学习,提出SRCNN网络直接学习高、低分辨率图像之间的映射关系,此种方法重建图像相较于传统方法效果更好。LI等^[8]采用卷积神经网络(convolutional neural networks, CNN)学习端到端网络,借助高分辨率彩色图像信息,引导低分辨率深度图重建高分辨率深度图。HUI等^[9]采用多尺度引导的方式对深度图像进行上采样,探究深度

结构与颜色纹理之间的相关性。ZHAN等^[10]提出了一种基于图像相似性和特征的图像超分辨率重建方法,跨尺度融合特征获取图像信息。

单幅深度图像超分辨率重建时更多利用先验信号,通过学习高、低分辨率深度图像训练数据之间的映射关系,从而推动低分辨率深度图像重建过程。基于参考的深度图像超分辨率重建时与单幅的深度图像超分辨率重建时不同,基于参考的深度图像超分可以利用辅助的图像相关信息^[11-13],获取更准确的细节特征,可以通过图像对齐或图像块匹配等多种方式实现重建^[12],利用参考图像进一步改善深度图像重建质量。基于跨任务交互来实现超分重建等方法^[14]。虽然目前以上方法广泛应用于彩色图像超分任务,但在深度图像超分重建中应用较少因此提出一种基于参考的Transformer纹理迁移深度图像超分辨率重建方法。通过Transformer计算预处理后的低分辨率图像与参考图像之间的相似性,并结合注意力机制,进行纹理迁移。本文首先对基于参考图像的超分辨率重建算法进行分析,对应用Transformer来解决超分问题的算法进行介绍,提出利用参考图像与Transformer结合来进行重建。其次展开介绍参考数据集的构成、损失函数以及其他网络训练细节。最后从主观和客观的角度,评价本文所述的方法在深度图像超分辨率重建中应用效果如何。

1 本文算法

针对目前基于学习的DSR中出现的边缘细节特征不清晰等问题,本文提出一种基于参考的Transformer纹理迁移深度图像超分辨率重建算法。对预处理后的LR-D以及参考图像特征块,利用归一化内积方法进行特征块相似度计算,融合Transformer计算相似位置置信度,并结合注意力机制进行纹理迁移,最后与LR-D特征结合,进一步精确重建结果。

1.1 本文网络结构

图1是依据本文所搭建的深度图像超分辨率重建算法所绘制的结构框图。LR_D 尺寸记为 $H \times W$, HR_D (high resolution depth image) 尺寸为 $rH \times rW$, 参考图像尺寸为 $rH \times rW$, 其中 LR_D 的高为 H , 宽为 W , 采样因子为 r 。

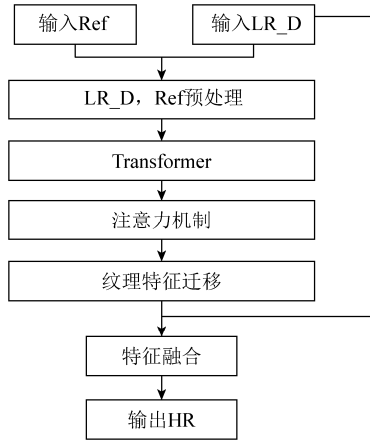


图1 基于参考的 Transformer 纹理迁移网络结构框图

Fig. 1 Block diagram of texture migration network in Transformer based on reference

对 LR_D 和参考图像(reference image, Ref)进行预处理之后输入到 Transformer 中进行相似度的计算以及相似位置的匹配, 结合注意力机制为不同置信度的相似位置块赋予不同的权重, 进一步提高纹理及结构细节恢复的准确度, 再进行纹理特征迁移, 并与 LR_D 特征进行融合后得到深度图像重建结果。

1.2 基于参考图像的超分辨率重建

基于参考 DSR 需要额外的辅助信息, SISR (single image super resolution) 则与之不同, 基于参考图像的超分(RefSR)方法引入了其他具有结构高度相似等特点的图像来引导超分辨率重建过程。文献[14]用 CNN 中的特征替换了简单的梯度特征, 以应用语义匹配, 并使用 SISR 方法进行特征合成。文献[11]采用光流以不同比例对齐 LR_D 和 Ref 图像, 并将其连接到解码器的相应层中。上述方法的性能在很大程度上取决于 LR_D 和 Ref 图像之间的对齐质量, 诸如光流之类的对准方法耗时, 不利于实际应用。

通常参考图像可以从网络检索^[15]、外部数据库^[16]或来自不同视点的图像^[11]中选择。YUE 等^[15]提出了从网络中检索相似图像的方法, 并进行全局配准和结构感知匹配, 以提高 LR 和 HR 之间的映射精度, 但该算法效果如何则高度依赖于参考图像

的匹配程度。XIA 等^[17]提出了一种基于参考的超分辨率加速多尺度聚合网络, 对输入大小具有渐近线性计算成本。充分利用跨多个尺度的参考信息并增强对标度错位的鲁棒性。文献[11]提出一种利用深度模型来重建的 RefSR 方法, 从 LR 和 Ref 中提取多尺度特征, 在空间上将参考特征图与 LR 特征图进行匹配。因此 ZHANG 等^[18]提出的 SRNTT (super-resolution by neural texture transfer) 采用了局部纹理匹配的思想, 将 Ref 纹理特征融合到最后结果, 通过深度模型来处理, 利用多尺度空间特征空间, 使得能够从各种参考中学习到复杂的传递过程。因此本文选择利用 SRNTT 方法来构建参考数据集。

1.3 LR_D 与 Ref 相似度估计模块

输入 LR_D, 并选择与之结构高度相似的高分辨率图像来作为参考图像, 其中分别对 LR_D 和 Ref 图像进行预处理。

如图2所示, 对 LR_D 进行上采样(LR_D $\uparrow$), 对 Ref 先下采样再上采样(Ref $\downarrow\uparrow$), 与上采样后的 LR_D 保持相同的分辨率和相似的结构。通过一个预训练的 VGGNet 模型, 进行 LR_D $\uparrow$, Ref $\downarrow\uparrow$ 以及 Ref 原图像特征提取, 通过特征提取将输入图像分为若干块, 将每个块展成向量形式, 利用 Transformer 中的 Self-Attention 机制, 分别将向量形式的特征表示为 Transformer 中的 Q 和 K , Ref 原图像提取的向量形式的特征被表示为 Transformer 中的 V 。通过归一化内积计算 Q 和 K 中的任意 2 个块向量之间的相关性, 即

$$S_{i,j} = \left\langle \frac{q_i}{\|q_i\|} \cdot \frac{k_j}{\|k_j\|} \right\rangle \quad (1)$$

其中, $S_{i,j}$ 为 Q 和 K 中的任意 2 个块向量之间的相关性, Q 中不同位置的块由 q_i 表示($i \in [1, H_{LR} \times W_{LR}]$), K 中不同位置的块由 k_j 表示($j \in [1, H_{Ref} \times W_{Ref}]$)。利用归一化内积所获取的相关性来映射 LR_D 与 Ref 之间的相关纹理信息, 内积越大的地方代表 2 个特征块之间的相似度越高。

1.4 纹理特征迁移模块

由于在深度图像重建过程中, 图像不同部分的注意力因素对深度图纹理重建效果影响颇大, 现有的基于深度学习的 DSR 在处理纹理模糊和纹理失真方面存在不足, 因此在深度图像重建网络中, 提高纹理重建的准确度具有重要意义。

如图3所示, 利用归一化内积计算所得到的 Q

和 K 中任意 2 个块向量之间的相关性, 并记录为 $S_{i,j}$, 传统的注意力机制对每个 q_i 取 V 的加权和, 这样会使的重建图像最终出现纹理模糊的现象, 因此基于 Q 与 V 之间的对应关系, 可采取一种找寻在 V 中 Q 与 K 最相关的位置的方法

$$M_i = \arg \max_j S_{i,j} \quad (2)$$

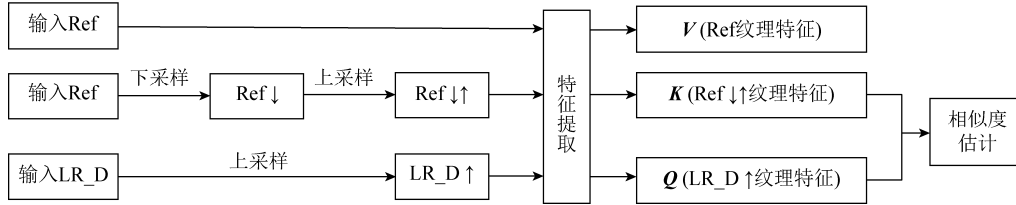


图2 LR_D 与 Ref 相似度估计过程图

Fig. 2 LR_D and Ref similarity estimation process diagram

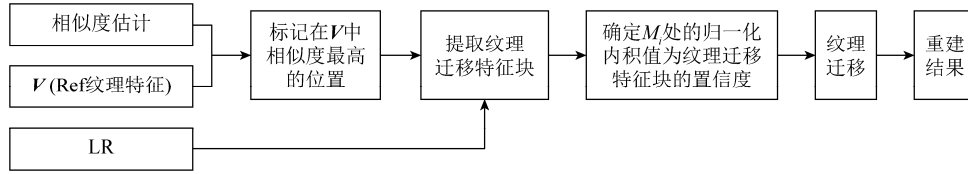


图3 纹理特征迁移过程图

Fig. 3 Texture feature migration process diagram

将提取出的不同部分的纹理迁移特征块与 LR_D 特征 F 进行融合, 利用高分辨率参考图信息与低分辨率深度图信息融合, 进一步提高纹理迁移特征块选提取的准确度。利用归一化内积计算所得到的 Q 和 K 中任意 2 个块向量之间的相关性, 确定 M_i 处的归一化内积值为纹理迁移特征块的置信度, 即

$$P_i = M_i \max_j S_{i,j} \quad (4)$$

对置信度更高纹理特征赋予相对较大的权重, 置信度低的纹理特征, 赋予较小的权重, 再进行纹理迁移, 得到最终的深度图像重建输出, 使得深度图像重建的效果更加精确。

1.5 损失函数

本文以重建损失(reconstruction loss, L_{rec})、对抗损失(adversarial loss, L_D 、 L_G)以及感知损失(perceptual loss, L_{per})作为损失函数, 重建损失定义为

$$L_{\text{rec}} = \frac{1}{CHW} \|I^{\text{HR}} - I^{\text{SR}}\|_1 \quad (5)$$

其中, CHW 为 HR_D 尺寸, I^{SR} 为预测的 SR 图像, 选用 L1 损失, 在性能上更好, 更易收敛。

生成对抗网络在解决生成图像模糊的问题上取得的效果十分显著, 采用 WGAN-GP^[19]方式可获

其中, M_i ($i \in [1, H_{\text{LR}} \times W_{\text{LR}}]$) 的值作为索引, 其代表 LR_D 图像中的第 i 个位置与 Ref 中最相关的位置。

将根据索引 M_i 找到 V 中对应的第 i 个位置为

$$N_i = VM_i \quad (3)$$

其中, N_i 为从 V 中的第 i 个位置中选择的特征图, 来作为纹理迁移特征图 T 中的第 i 个位置。

得好的训练结果, 其定义为

$$L_D = \mathbb{E}_{\tilde{x} \sim P_g} [D(\tilde{x})] - \mathbb{E}_{x \sim P_r} [D(x)] + \lambda \mathbb{E}_{\tilde{x} \sim P_g} \left[\left(\|\nabla_{\tilde{x}} D(\tilde{x})\|_2 - 1 \right)^2 \right] \quad (6)$$

$$L_G = - \mathbb{E}_{\tilde{x} \sim P_g} [D(\tilde{x})] \quad (7)$$

感知损失在提高图像视觉方面效果显著, 在多种方法^[20-23]中都有所体现, 其关键思想是增加图像之间特征空间的相似性, 即

$$L_{\text{per}} = \frac{1}{C_i H_i W_i} \left\| \varphi_i^{\text{vgg}}(I^{\text{SR}}) - \varphi_i^{\text{vgg}}(I^{\text{HR}}) \right\|_2^2 + \frac{1}{C_j H_j W_j} \left\| \varphi_j^{\text{fe}}(I^{\text{SR}}) - T \right\|_2^2 \quad (8)$$

其中, $\varphi_i^{\text{vgg}}(\cdot)$ 为 VGG19 的第 i 层特征图, 第 i 层特征图的形状用 $C_i H_i W_i$ 进行表示; I^{SR} 为预测的 SR 图像; $\varphi_j^{\text{fe}}(\cdot)$ 为特征提取部分第 j 层提取的纹理特征图; $C_j H_j W_j$ 为第 j 层特征图的形状。用这种传递感知损失约束预测 SR 图像的方法, 可以更有效地传递 Ref 纹理。

利用对抗损失、感知损失以及重建损失, 生成更加真实、清晰的细节特征, 不断约束预测的目标图像生成, 以获取重建性能更好的效果图像。

2 实验设置

2.1 实验平台

实验软硬件配置见表 1。

表 1 实验软硬件配置

Table 1 Experimental software and hardware on figuration

名称	实验配置
操作系统	ubuntu 20.04
编程语言	Python 3.7.9
深度学习框架	PyTorch 1.8.0
CPU	Intel Xeon Gold 6226
GPU	NVIDIA RTX3090
Cuda	Cuda 11.1
平台	Pycharm-2021.2.1

2.2 数据集

训练数据集总共由两部分构成, 选择 Middlebury 数据集以及 MPI Sintel 深度图数据集来进行训练, 分别抽取 11 幅深度图像和 40 幅深度图像^[24]。

利用 SRNTT 方法来构建参考数据集, 基于 SIFT (scale-invariant feature transform)^[25]特征匹配收集图像对, 构建不同的相似级别, 符合局部纹理匹配目标。

如图 4 所示, 构建的参考数据集包含 HR_D 图像和 Ref 图像, 根据 SIFT 特征的最佳匹配数量, 定义了不同的相似性级别, 从每个成对的图像中, 从一幅图像中随机裁剪块作为原始 HR_D 图像, 从另一幅图像中裁剪作为相应的参考。充分利用这些图像信息, 对其进行数据增强, 对原图像进行旋转 (90°, 180°, 270°) 处理, 再对处理后的图像进行 0.8 和 0.9 倍缩放^[25], 得到 612 张图像组成最终数据集以及 612 张对应图像组成最终参考数据集。验证数据集 3 张来自 Middlebury 数据集和 2 张来自 MPI Sintel 数据集的图像构成。选取超分辨率常用的 3 张深度图像 (art, books, moebius, 记为 Mid3) 进行纹理迁移效果测试, 重建结果的客观评价标准选择结

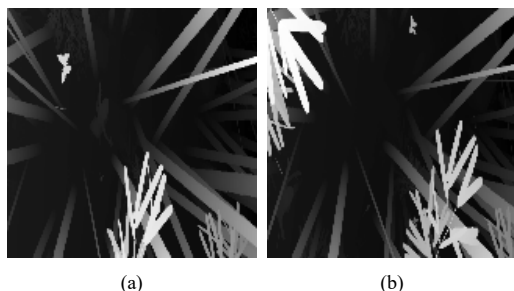


图 4 数据集图像示例((a) HR_D 图像; (b) Ref 图像)

Fig. 4 Dataset image example ((a) HR_D image; (b) Ref image)

构相似性 (structural similarity, SSIM) 来进行评估

$$SSIM(X, Y) = \frac{(2\mu_X\mu_Y + C_1)(2\sigma_{XY} + C_2)}{(\mu_X^2 + \mu_Y^2 + C_1)(\sigma_X^2 + \sigma_Y^2 + C_2)} \quad (9)$$

其中, μ_X 和 μ_Y 为图像 X 和 Y 像素的均值; σ_X 和 σ_Y 为图像 X 和 Y 像素的方差; σ_{XY} 为图像 X 和 Y 像素的协方差; C_1 和 C_2 为常数, 其取值范围为 [0, 1], 其值越大, 表示图像越接近原始图像。

3 实验结果与分析

为了测试本文算法在深度图像重建过程中的性能, 实验选择在上采样因子为 2, 3, 4 时分别与 Bicubic, SRCNN^[7]和 ESPCN^[26]进行比较。图 5 和图 6 分别是测试图像的在上采样因子 2 倍和 4 倍时的重建效果图和局部放大图。

如图 5 所示, 在图像上采样因子为 4 倍时, 本文算法的重建效果优于 Bicubic, SRCNN 和 ESPCN 算法, 在局部放大图中, Bicubic 算法的重建图最为模糊, SRCNN 算法以及 ESPCN 算法在图像细节的恢复中仍存在模糊现象, 但较与 Bicubic 算法有一定提升。本文算法在重建过程中融合参考图像信息, 可较准确地重建出深度图像边缘细节, 图像重建效果更好。

在本文算法中, 对于每个查询, 只选取最相关的位置进行纹理特征迁移。利用归一化内积所获取的相关性来映射 LR_D 与 Ref 之间的相关纹理信息, 内积越大代表 2 个特征块之间的相关性越强、相似度越高。对于不同相似度进行置信度估计, 对置信度高, 相似度高的纹理特征赋予相对更大的权重, 相似度低的纹理特征, 赋予更小的权重, 再进行纹理特征迁移, 得到最终的深度图像重建输出, 使得深度图像重建的效果更加精确。

从实验效果图来看, 双三次插值的重建图像结果相对最不清晰, SRCNN 和 ESPCN 算法在一定程度上进行了改善, 但图像边缘处仍然存在模糊现象, 且图像细节呈现不清晰。本文算法相较重建质量更好, 边缘锐化保持较好, 细节更加清晰且进一步改善了图像的模糊现象。

表 2 分别给出了上述 4 种方法在针对 Mid3 中每张深度图像经超分辨率重建后的客观评价指标。可以看到本文算法在上采样因子为 2, 3, 4 时 SSIM 值明显高于 Bicubic, SRCNN 和 ESPCN 算法。从客观上说明本文算法对于 DSR 效果更好, 更接近真实深度图像。

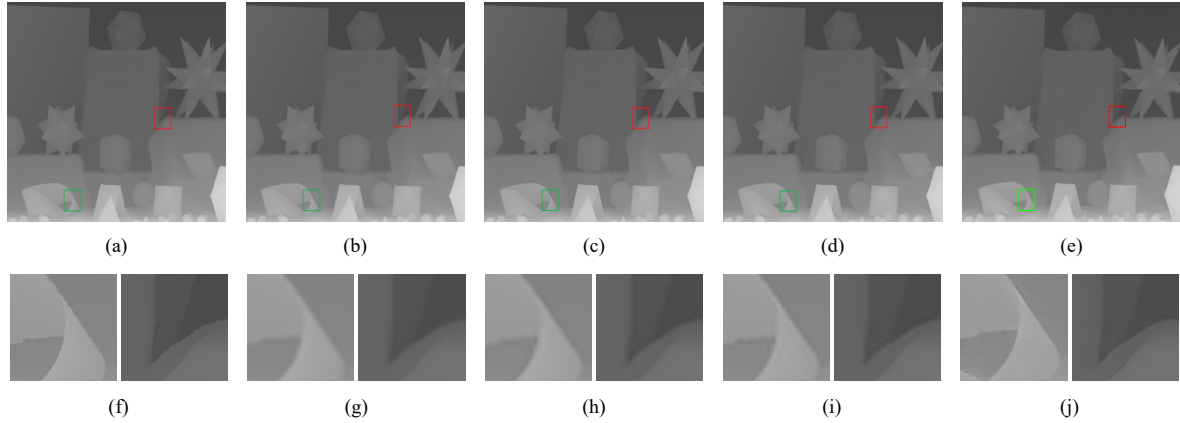


图 5 上采样 4 倍时不同算法在 Mid3 上的重建效果对比((a)高分辨率深度图像; (b) Bicubic 插值结果; (c) SRCNN 算法结果; (d) ESPCN 算法结果; (e) 本文算法结果; (f) 图(a)的局部放大图; (g) 图(b)的局部放大图; (h) 图(c)的局部放大图; (i) 图(d)的局部放大图; (j) 图(e)的局部放大图)

Fig. 5 Comparison of reconstruction effects of different algorithms on Mid3 when the up-sampling factor is 4 ((a) High resolution depth images; (b) Bicubic interpolation results; (c) SRCNN algorithm results; (d) ESPCN algorithm results; (e) Algorithm results in this article; (f) Partial enlarged view of figure (a); (g) Partial enlarged view of figure (b); (h) Partial enlarged view of figure (c); (i) Partial enlarged view of figure (d); (j) Partial enlarged view of figure (e))

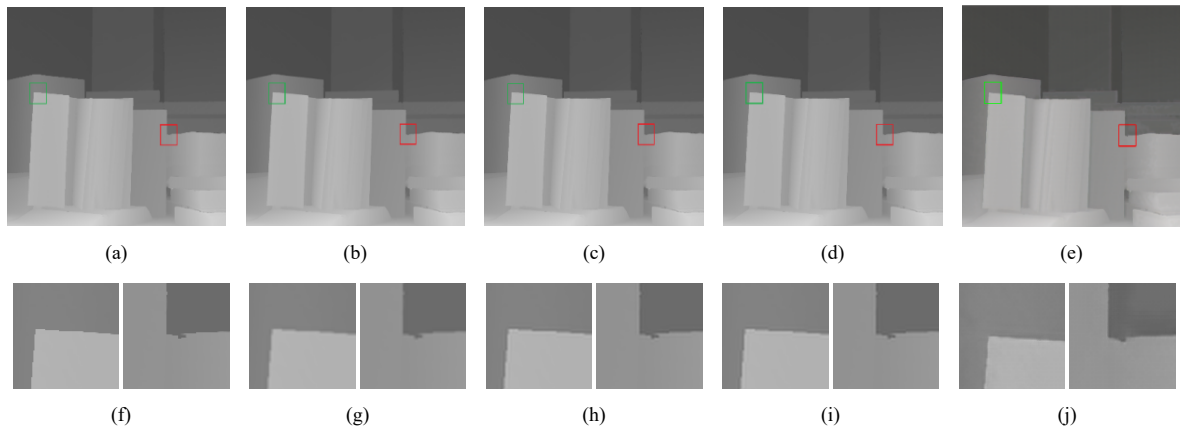


图 6 上采样 2 倍时不同算法在 Mid3 上的重建效果对比((a)高分辨率深度图像; (b) Bicubic 插值结果; (c) SRCNN 算法结果; (d) ESPCN 算法结果; (e) 本文算法结果; (f) 图(a)的局部放大图; (g) 图(b)的局部放大图; (h) 图(c)的局部放大图; (i) 图(d)的局部放大图; (j) 图(e)的局部放大图)

Fig. 6 Comparison of reconstruction effects of different algorithms on Mid3 when the up-sampling factor is 2 ((a) High resolution depth images; (b) Bicubic interpolation results; (c) SRCNN algorithm results; (d) ESPCN algorithm results; (e) Algorithm results in this article; (f) Partial enlarged view of figure (a); (g) Partial enlarged view of figure (b); (h) Partial enlarged view of figure (c); (i) Partial enlarged view of figure (d); (j) Partial enlarged view of figure (e))

表 2 不同算法在 Mid3 上重建结果的定量分析(SSIM)

Table 2 Quantitative analysis of reconstruction results of different algorithms on Mid3 (SSIM)									
算法	×2			×3			×4		
	Art	Books	Moebius	Art	Books	Moebius	Art	Books	Moebius
Bicubic	0.989 8	0.996 1	0.997 3	0.976 8	0.993 4	0.993 9	0.967 4	0.990 7	0.989 2
SRCNN	0.989 8	0.996 1	0.997 6	0.981 0	0.993 9	0.993 6	0.972 7	0.991 1	0.989 4
ESPCN	0.989 9	0.996 1	0.997 7	0.975 9	0.994 0	0.994 3	0.974 0	0.991 2	0.989 7
Ours	0.999 8	0.999 9	0.999 9	0.999 5	0.999 7	0.999 6	0.997 7	0.999 5	0.999 3

4 结束语

本文提出了一种基于参考的 Transformer 纹理迁移的深度图像超分辨率重建方法, 该方法将

HR_D 纹理从 Ref 迁移到 LR_D 图像中。本文利用 SRNTT 方法来构建参考数据集, 利用高分辨率的参考图像来进行纹理迁移, 利用 Transformer 进行相似度的计算以及相似位置的匹配, 结合注意力机

制进一步提高纹理及结构细节恢复的准确度, 并与 LR_D 特征进行融合后进一步精确深度图像重建结果。针对本文算法进行了一系列的实验, 实验结果表明, 本文算法在 DSR 中效果较好, 图像细节及锐化明显。下一步的研究计划是如何进一步的提升网络性能, 提高深度图像重建的准确性。将进一步将所提出的纹理迁移网络应用到更多的图像生成任务。

参考文献 (References)

- [1] HORNÁČEK M, RHEMANN C, GELAUTZ M, et al. Depth super resolution by rigid body self-similarity in 3D[C]//2013 IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2013: 1123-1130.
- [2] LEI J J, LI L L, YUE H J, et al. Depth map super-resolution considering view synthesis quality[J]. IEEE Transactions on Image Processing: a Publication of the IEEE Signal Processing Society, 2017, 26(4): 1732-1745.
- [3] XIE J, FERIS R S, SUN M T. Edge-guided single depth image super resolution[C]//2014 IEEE International Conference on Image Processing. New York: IEEE Press, 2016: 3773-3777.
- [4] LIU W, CHEN X G, YANG J, et al. Robust color guided depth map restoration[J]. IEEE Transactions on Image Processing, 2017, 26(1): 315-327.
- [5] 李滔, 董秀成, 张晓华. 深度图像超分辨率重建技术综述[J]. 西华大学学报: 自然科学版, 2020, 39(4): 45-53.
LI T, DONG X C, ZHANG X H. A survey of super-resolution reconstruction technologies for depth image[J]. Journal of Xihua University: Natural Science Edition, 2020, 39(4): 45-53 (in Chinese).
- [6] MAC AODHA O, CAMPBELL N D F, NAIR A, et al. Patch based synthesis for single depth image super-resolution[C]//European Conference on Computer Vision. Heidelberg: Springer, 2012: 71-84.
- [7] DONG C, LOY C C, HE K M, et al. Image super-resolution using deep convolutional networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(2): 295-307.
- [8] LI Y J, HUANG J B, AHUJA N, et al. Deep Joint Image Filtering[C]//European Conference on Computer Vision. Cham: Springer International Publishing, 2016: 154-169.
- [9] HUI T W, LOY C C, TANG X O. Depth Map Super-Resolution by Deep Multi-Scale Guidance[C]//European Conference on Computer Vision. Cham: Springer International Publishing, 2016: 353-369.
- [10] ZHAN Y L, CHI J, YE Y N, et al. Super resolution image reconstruction based on image similarity and feature combination[J]. Journal of Computer-Aided Design & Computer Graphics, 2019, 31(6): 1018-1029.
- [11] ZHENG H T, JI M Q, WANG H Q, et al. CrossNet: an end-to-end reference-based super resolution network using cross-scale warping[M]//Computer Vision – ECCV 2018. Cham: Springer International Publishing, 2018: 88-104.
- [12] YANG F Z, YANG H, FU J L, et al. Learning texture transformer network for image super-resolution[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2020: 5790-5799.
- [13] SUN B L, YE X C, LI B P, et al. Learning scene structure guidance via cross-task knowledge transfer for single depth super-resolution[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2021: 7788-7797.
- [14] ZHENG H T, JI M Q, HAN L, et al. Learning scene structure guidance via cross-task knowledge transfer for single depth super-resolution, 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2021: 7788-7797.
- [15] YUE H J, SUN X Y, YANG J Y, et al. Landmark image super-resolution by retrieving web images[J]. IEEE Transactions on Image Processing, 2013, 22(12): 4865-4878.
- [16] ZHU Y, ZHANG Y N, YUILLE A L. Single image super-resolution using deformable patches[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2014: 2917-2924.
- [17] XIA B, TIAN Y P, HANG Y C, et al. Coarse-to-fine embedded PatchMatch and multi-scale dynamic aggregation for reference-based super-resolution[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2022, 36(3): 2768-2776.
- [18] ZHANG Z F, WANG Z W, LIN Z, et al. Image super-resolution by neural texture transfer[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2020: 7974-7983.
- [19] GULRAJANI I, AHMED F, ARJOVSKY M, et al. Improved training of Wasserstein GANs[EB/OL]. [2022-12-06]. <https://arxiv.org/abs/1704.00028>.
- [20] JOHNSON J, ALAHI A, LI F F. Perceptual Losses for Real-Time Style Transfer and Super-Resolution[C]//European Conference on Computer Vision. Cham: Springer International Publishing, 2016: 694-711.
- [21] LEDIG C, THEIS L, HUSZÁR F, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2017: 105-114.
- [22] SAJJADI M S M, SCHÖLKOPF B, HIRSCH M. EnhanceNet: single image super-resolution through automated texture synthesis[C]//2017 IEEE International Conference on Computer Vision. New York: IEEE Press, 2017: 4501-4510.
- [23] ZHANG Z F, WANG Z W, LIN Z, et al. Image super-resolution by neural texture transfer[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2020: 7974-7983.
- [24] 范佩佩, 董秀成, 李滔, 等. 基于非局部均值约束的深度图像超分辨率重建[J]. 计算机辅助设计与图形学学报, 2020, 32(10): 1671-1678.
FAN P P, DONG X C, LI T, et al. Super-resolution reconstruction of depth map based on non-local means constraint[J]. Journal of Computer-Aided Design & Computer Graphics, 2020, 32(10): 1671-1678 (in Chinese).
- [25] LOWE D G. Object recognition from local scale-invariant features[C]//The 7th IEEE International Conference on Computer Vision. New York: IEEE Press, 2002: 1150-1157.
- [26] SHI W Z, CABALLERO J, HUSZÁR F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2016: 1874-1883.