

基于双特征融合引导的深度图像超分辨率重建网络

耿浩文, 王宇*, 辛彦玲

长春理工大学电子信息工程学院, 吉林 长春 130012

摘要 针对彩色图像引导的深度图像超分辨率重建算法中存在的纹理转移和深度流失的问题, 提出一种基于双特征融合引导的深度图像超分辨率重建网络(DF-Net)。为了充分利用深度和强度特征的关联性, 在网络模型的深度恢复重建部分采用双通道融合模块(DCM)和双特征引导重建模块(DGM)。利用输入金字塔结构提取深度信息和强度信息的多尺度特征; DCM基于通道注意力机制对深度特征和强度特征进行通道间的特征融合与增强; DGM将深度、强度特征自适应选择融合后实现重建的双特征引导, 增加了深度特征的引导作用, 改善了纹理转移和深度流失的问题。实验结果表明, 所提方法的峰值信噪比(PSNR)和均方根误差(RMSE)优于RMRF、JBU和Depth-Net等方法, 尤其 $4\times$ 超分辨率重建结果的PSNR值比其他方法平均提升6.79 dB, RMSE平均下降0.94, 取得了较好的深度图像超分辨率重建效果。

关键词 图像处理; 图像超分辨率重建; 卷积神经网络; 深度图像; 特征融合; 通道注意力

中图分类号 TP391.4

文献标志码 A

DOI: 10.3788/LOP230593

Depth Image Super-Resolution Reconstruction Network Based on Dual Feature Fusion Guidance

Geng Haowen, Wang Yu*, Xin Yanling

College of Information Science and Engineering, Changchun University of Science and Technology,
Changchun 130012, Jilin, China

Abstract A depth image super-resolution reconstruction network (DF-Net) based on dual feature fusion guidance is proposed to address the issues of texture transfer and depth loss in color image guided deep image super-resolution reconstruction algorithms. To fully utilize the correlation between depth and intensity features, a dual channel fusion module (DCM) and a dual feature guided reconstruction module (DGM) are used to perform deep recovery and reconstruction in the network model. The multi-scale features of depth and intensity information are extracted using an input pyramid structure: DCM performs feature fusion and enhancement between channels based on a channel attention mechanism for depth and intensity features; DGM provides dual feature guidance for reconstruction by adaptively selecting and fusing depth and intensity features, increasing the guidance effect of depth features, and overcoming the issues of texture transfer and depth loss. The experimental results show that the peak signal-to-noise ratio (PSNR) and root mean square error (RMSE) of the proposed method are superior to those of methods such as RMRF, JBU, and Depth Net. Compared to the other methods, the PSNR value of the $4\times$ super-resolution reconstruction results increased by an average of 6.79 dB, and the RMSE decreased by an average of 0.94, thus achieving good depth image super-resolution reconstruction results.

Key words image processing; image super-resolution reconstruction; convolution neural network; depth image; feature fusion; channel attention

1 引言

深度信息是人类感知三维空间事物的重要信息。RGB-D传感器(Kinect、ToF等)能够实时地获取空间

内物体的深度信息,被广泛应用于机器人视觉、医疗、多媒体娱乐等多个领域。由于硬件的限制和成像过程中存在的噪声等影响^[1],现有的深度相机获得的深度图像分辨率较低、质量较差,进而限制了三维重建、虚

收稿日期: 2023-02-08; 修回日期: 2023-03-24; 录用日期: 2023-04-03; 网络首发日期: 2023-04-13

基金项目: 吉林省自然科学基金(20210101180JC)

通信作者: *muxie2002@126.com

拟现实等很多领域的发展。因此,关于深度图像超分辨率重建(DSR)技术的研究具有重要意义。

深度图像超分辨率重建技术中,彩色图像可以为重建提供边缘信息等很多先验信息,从而提高深度图的重建效果,所以由彩色图像引导的深度图像超分辨率重建受到重视。Diebel等^[2]最先利用 Markov random field(MRF)模型建立深度图和同场景彩色图的关系,通过求解最小二乘优化问题获得高分辨率深度图像;Ferstl等^[3]提出一种二阶广义全变分(TGV)模型,利用彩色图像的各向异性张量引导 DSR;Liu等^[4]通过鲁棒惩罚函数模拟模型的平滑项获得高分辨率的深度图像。

近年来,卷积神经网络(CNN)由于具备极好的自动提取特征能力、快速处理高维图像能力、一定的平移不变性等优势,在图像重建等领域取得了较好的效果^[5]。彩色图像引导的 DSR 网络始于 Hui等^[6]设计的 MSG-Net,这也是彩色图像引导的深度图超分辨率重建的重要基础。之后,Guo等^[7]提出 DepthSR-Net,DepthSR-Net通过深度输入金字塔结构实现了多层次特征提取,在深度恢复重建结构中每一级上都加入对应大小的强度特征引导恢复重建。于淑侠等^[8]在深度图像提取特征的过程中采用多感受野残差块(MRFRB)提取不同感受野下的特征。之后,Zhao等^[9]提出的离散余弦变换网络(DCT-Net)以及 Ido等^[10]提出的用于深度图超分辨率的多尺度残差深度网络都使得彩色图像引导的 DSR 网络研究得到发展。

为解决彩色图像引导的 DSR 网络仍然存在的问题,本文提出一种基于双特征融合引导的深度图像超分辨率重建网络(DF-Net)。强度图像特征和深度图像特征通常仅在浅层中融合,且一般采用直接拼接或特征相加的融合方式,这样就会出现由于特征融合不充分导致最后重建的深度图出现纹理转移和深度流失的现象。为解决这个问题,在深度恢复重建阶段提出双通道融合模块(DCM),充分融合强度、深度两种特征。

通常只采用强度特征进行引导重建,没有考虑到深度特征对于恢复重建的引导作用。而所设计的双特征引导重建模块(DGM)用提取到的深度特征和强度特征共同引导重建特征进行深度图像超分辨率重建,既考虑到深度特征的信息,又避免了特征利用不足的问题。

通常仅使用最后一级特征进行重建,导致 DSR 对高分辨率(HR)强度图像和低分辨率(LR)深度图像中的特征利用不足。为解决这个问题,在深度恢复重建阶段中充分利用特征提取得到的深度和强度特征进行逐级引导。

此外,在特征提取和深度恢复重建部分引入残差注意力机制模块(RAM)^[11],利用通道之间相互依赖关系有效提取并强化特征。

2 网络设计

2.1 网络整体结构

所提 DF-Net 由特征提取和深度恢复重建两部分组成。特征提取部分包括深度特征提取支路和强度特征提取支路。网络整体结构如图 1 所示,其中, D_{lr}^p 代表双三次插值后的深度图像,对于同场景 HR 彩色图像,采用 YCrCb 格式。由于其 Y 通道为亮度通道,包含的强度信息要比 Cr 和 Cb 的色度通道更丰富,更有利于引导重建,且 Y 通道与 CrCb 通道特征差别较大,故采用同场景 HR 彩色图像的 Y 通道作为输入,用 Y_{hr} 表示。两个支路的网络结构一致,分别用于提取不同尺度的深度特征和强度特征。

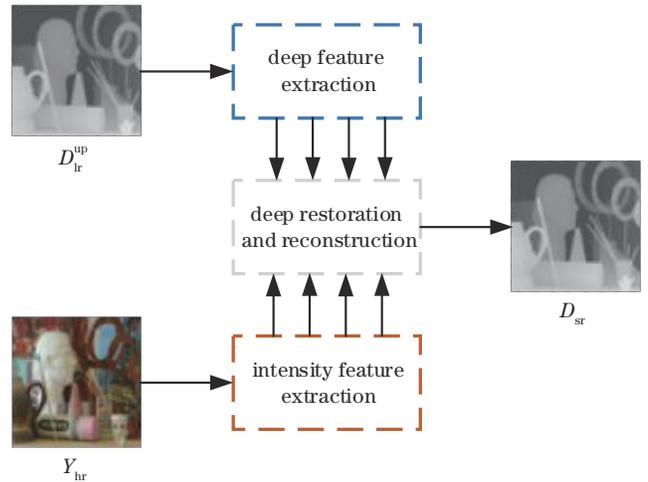


图 1 网络整体结构

Fig. 1 Overall network structure

其中,特征提取部分包括深度特征提取支路和强度特征提取支路。两个支路的网络结构一致,均采用输入金字塔结构。并在输入金字塔结构中嵌入由 RAM 强化提取到的特征。在深度恢复重建部分中,以特征提取部分得到的各种尺度的深度、强度特征作为输入,经过 DCM 进行初步融合后,由反卷积放大重建特征,并使用 DGM 进行逐级引导恢复重建,得到重建后的高分辨率深度图 D_{sr} 。

2.2 特征提取部分

特征提取结构中,常见的提取方式有单输入单提取和单输入多提取,前者由于采用一个卷积核无法实现多尺度充分特征提取,而后者先对单个输入图像进行多尺度特征提取然后拼接,提高了网络过拟合风险。而所提特征提取结构采用输入金字塔结构^[7],首先对输入图像进行多次下采样得到多个不同尺寸的输入,这样做可以以多个层次的感受野从输入图中提取多层次特征表示,又可以通过提供抽象的表示形式引入多种尺寸的输入图像来降低过拟合风险。以深度特征提取支路为例,其结构如图 2 所示。

对输入深度图 $P_{lr}^0 = D_{lr}^p$ 进行 N 次下采样(图 2 中

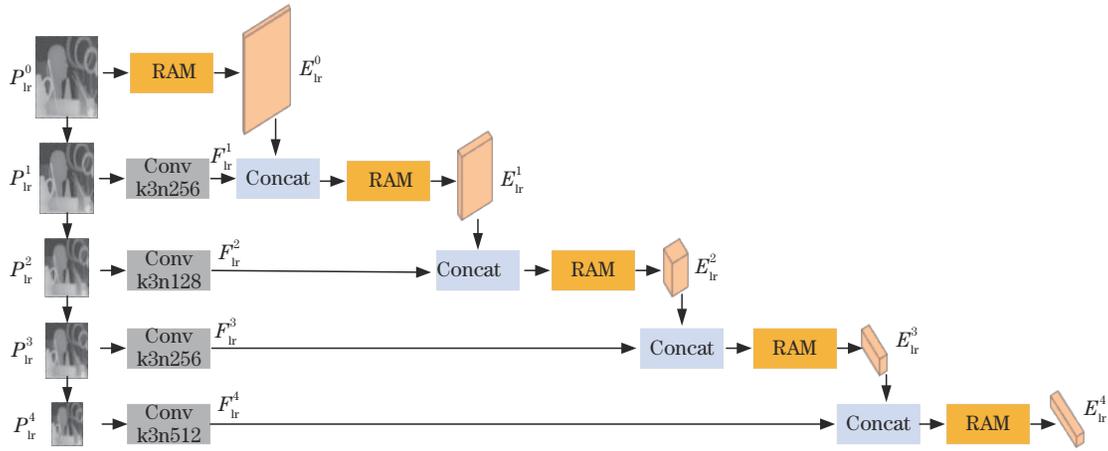


图 2 深度特征提取支路结构图

Fig. 2 Structure diagram of depth feature extraction branch

的 $N=4$), 每次下采样都经过一次不同通道的卷积得到多种尺度的特征图:

$$P_{lr}^i = \text{maxpool}(P_{lr}^{i-1}), \quad (1)$$

$$F_{lr}^i = \sigma(W_{lr}^i P_{lr}^i + b_{lr}^i), \quad (2)$$

式中: $i \in \{1, 2, 3, 4\}$; P_{lr}^i 是经过 i 次下采样后的深度图像; F_{lr}^i 是从 P_{lr}^i 中提取的特征; W_{lr}^i 和 b_{lr}^i 是卷积操作的权重和偏置; σ 为 ReLU 激活函数。

在深度特征提取支路中的第一级中, 输入深度图 P_{lr}^0 经过 RAM 后得到第一级的输出特征 E_{lr}^0 , 下面的每一级都先将本级的特征 F_{lr}^i 与上一级特征下采样后的

特征拼接, 然后通过 RAM 得到本级的输出特征 E_{lr}^i , 这些特征也会用于后续的深度恢复重建部分中。

$$E_{lr}^0 = f_{\text{RAM}}(P_{lr}^0), \quad (3)$$

$$E_{lr}^i = f_{\text{RAM}}([F_{lr}^i, \text{maxpool}(E_{lr}^{i-1})]), \quad (4)$$

式中: $f_{\text{RAM}}(\cdot)$ 代表残差注意力模块; $j \in \{1, 2, 3, 4\}$; $[\cdot, \cdot]$ 代表拼接操作。由于 RAM^[11] 可以利用通道之间的依赖性消除伪影, 所以将 RAM 嵌入 DF-Net 中。RAM 结构如图 3 所示。输入特征经过卷积、PReLU 激活函数、两次残差注意力 (RCAB) 操作后得到最后的特征。

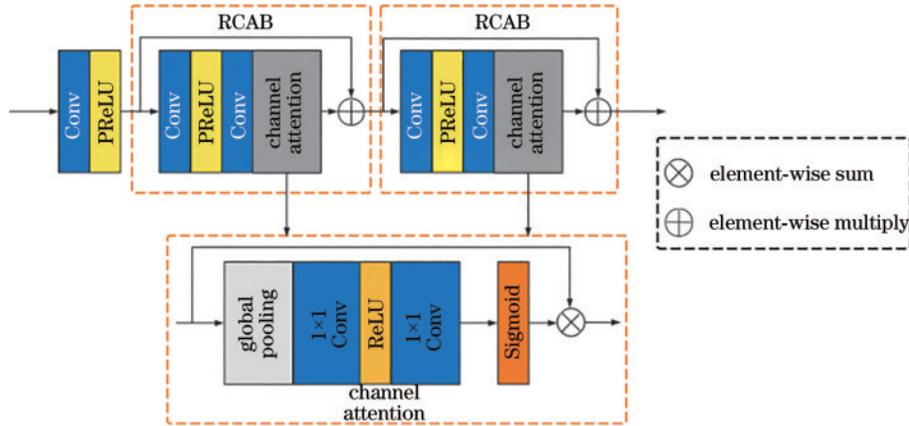


图 3 RAM 结构图

Fig. 3 RAM structure diagram

2.3 深度恢复重建部分

为了充分利用深度和强度特征的关联性, 在深度恢复重建部分设计了 DCM 和 DGM。深度恢复重建部分如图 4 所示。

以特征提取部分得到的深度特征和强度特征作为输入, 在 DCM 中进行双特征融合, 然后通过反卷积进行特征上采样和通道压缩, 将输出结果与特征提取部分同一级的深度特征和强度特征共同输入 DGM 中引导恢复重建, 融合后的特征再经过 RAM 操作, 再经过 3 次反卷积、DGM 和 RAM 的循环操作后得到重建特征 D_{lr}^0 :

$$D_{lr}^4 = f_{\text{DCM}}(E_{lr}^4, E_{lr}^4), \quad (5)$$

$$D_{lr}^k = f_{\text{RAM}}\{f_{\text{DGM}}[E_{lr}^k, E_{lr}^k, \text{deconv}(D_{lr}^{k+1})]\}, \quad (6)$$

式中: $k \in \{0, 1, 2, 3\}$; $f_{\text{DCM}}(\cdot)$ 代表 DCM 操作; $f_{\text{DGM}}(\cdot)$ 代表 DGM 操作; $\text{deconv}(\cdot)$ 代表反卷积操作, 这里采用的是转置反卷积。

2.3.1 双通道融合模块

将深度特征和强度特征进行通道间充分融合, 有利于恢复超分辨率重建深度图的轮廓信息^[12], DCM 结构如图 5 所示。

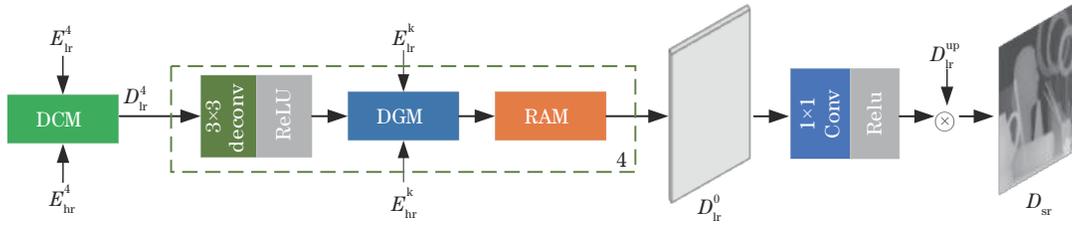


图 4 深度恢复重建部分结构图

Fig. 4 Partial structure diagram of depth recovery and reconstruction

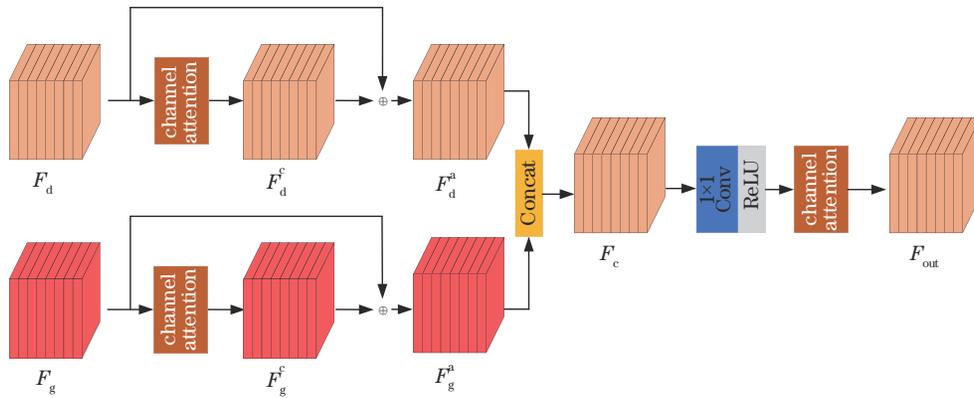


图 5 DCM 结构图

Fig. 5 DCM structure diagram

首先将深度特征 F_d 和强度特征 F_g 通过通道注意力^[13]操作(如图 4 中的 channel attention)提取通道间的关联性,然后分别与原特征相加,可以起到强化深度特征和强度特征的作用。

$$F_g^a = F_g^c + F_g, \quad (7)$$

$$F_d^a = F_d^c + F_d. \quad (8)$$

强化后的深度、强度特征拼接以后,再次经过通道注意力操作,完成双通道特征融合。

DCM 在特征提取部分得到多尺度的深度特征、强度特征,先各自采用通道注意力操作并与输入特征相加进行特征强化,然后经过特征拼接和 1×1 卷积操作进行特征初步融合,再利用通道之间的相关性进行通道注意力操作进行第二次融合与强化特征,进而实现对输入特征的充分融合。

与传统“拼接+卷积”的特征融合方式相比,DCM 模块首先强化两种有效特征,初步融合之后在通道方向利用相关性再次进行特征的融合与强化,进而实现特征的充分融合,也有利于加强深度图像重建的引导作用。

2.3.2 双特征引导重建模块

自选择连接(SSC)^[14]将上一级恢复的深度特征与强度特征作为输入,通过两种特征的空间和通道间的关联性进行融合,使得强度特征更好起到引导恢复重建的作用。

但是深度特征中也包含了纹理和边缘信息,而且和包含复杂的边缘等强度信息的强度特征相比,深度

特征的信息更为简单,均为深度变化的信息,可以抑制强度特征中对于深度图像中没用的纹理特征,同时强化深度特征中深度变化的特征。而常见方法中没有考虑到提取到的深度特征对深度图像恢复重建也有这样的引导作用。

因此,所提 DGM 通过 SSC 操作融合深度特征 F_d 和强度特征 F_g ,并将融合后的特征用于引导上一级特征 F_s 的恢复重建,结构如图 6 所示。

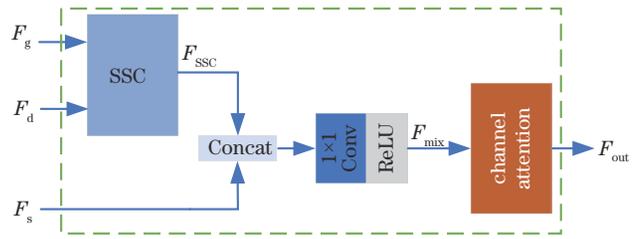


图 6 DGM 结构图

Fig. 6 DGM structure diagram

深度特征 F_d 和强度特征 F_g 经 SSC 操作融合后的特征 F_{SSC} 包含恢复重建所需的纹理和强度信息,用于引导上一级特征,可以解决深度流失、纹理转移的问题。对 F_{mix} 进行通道注意力操作能够利用特征通道之间的关联性,充分融合有效特征。

2.4 网络训练

在主体网络中,卷积层和反卷积层都采用 3×3 (0 填充)大小的卷积核,反卷积层的步幅设置为 2×2 。

每层网络参数以高斯分布初始化,均值为 0、标准差为 10^{-3} 。

采用 Adam 方法训练网络,使用均方误差(MSE)作为损失函数:

$$L(\theta) = \frac{1}{N} \sum_{i=1}^n \|F(D_{hr}^i, Y_{hr}^i; \theta) - X_i\|^2, \quad (9)$$

式中: n 代表每次训练样本的个数; θ 代表要学习的参数; $F(D_{hr}^i, Y_{hr}^i; \theta)$ 代表重建得到的深度图像; X_i 代表对应的真实深度图。

3 实验结果与分析

3.1 实验设置

3.1.1 数据集

使用来自 MPI Sintel 深度数据集的 58 个 RGB-D 图像对和 34 个来自 Middlebury 数据集的 RGB-D 图像对(2001 数据集 6 个图像对、2006 数据集 10 个图像对和 2014 数据集 18 个图像)作为训练集。其中,82 个图像对用于训练、10 个图像对用于验证。为了获得更多的训练数据,通过 90° 、 180° 、 270° 的旋转以及翻转来增强数据。

在训练集中,为减少训练时间:对于 $2\times$ 、 $4\times$ 以及 $8\times$ 超分辨率重建实验,将真实深度图和对应 RGB 彩色图像以 48 的步长进行重叠采样,裁剪得到 96×96 的图像块,数据增强后得到了大约 450000 个图像块;对于 $16\times$ 恢复重建实验,对图像以 64 的步长进行重叠采样,裁剪得到 128×128 的图像块,数据增强后得到大约 240000 个图像块。LR 深度图分别通过对图像块进行 $2\times$ 、 $4\times$ 、 $8\times$ 和 $16\times$ 双三次下采样获得。所有的训练图像都被归一化到 $[0, 1]$ 。

对于测试集,使用来自 Middlebury(2005)数据集的不同场景的 6 个 RGB-D 图像对作为实验的测试

数据。

采用峰值信噪比(PSNR)和均方根误差(RMSE)作为评价指标。PSNR 的数值越大,表明重构图像质量越好。RMSE 的数值越小,表明重建深度图与原始图像越接近,重建效果越好。

3.1.2 网络参数

显卡为 NVIDIA GeForce RTX 3080 Ti,编译软件为 Python 3.8,使用 TensorFlow 深度学习框架进行神经网络模型的搭建和训练。

在训练期间,采用批大小为 64 的批模式学习方法。使用 $\beta_1 = 0.9$ 和 $\beta_2 = 0.999$ 的 ADAM 进行网络优化。初始学习率设置为 10^{-6} :如果在 4 个 epoch 内损失函数没有减少,学习率衰减设定为 0.25;如果学习率衰减低于 10^{-7} ,网络停止训练。

3.2 实验结果

3.2.1 对比实验

选取了 8 种对比算法,其中:TGV^[3]、RMRF^[4]为传统算法;引导滤波器(GF)^[15]、联合双边上采样(JBU)^[16]为基于滤波的方法;MSG-Net^[6]、DepthSRNet^[7]、FDKN^[17]、RCAN^[18]为基于深度学习的算法。

高分辨率深度图 Art 的真实深度图、不同超分辨率算法经 $4\times$ 上采样后的局部深度对比如图 7 所示,高分辨率深度图 Laundry 的深度对比如图 8 所示。图中白色位置代表不同算法得到的深度图与原深度图的差异。总体可以看出,基于深度学习方法的效果要明显优于传统和基于滤波的算法。在 Art 图中的圆环位置以及 Laundry 深度图中的壶嘴、窗框等边缘位置可以看出:所提算法对于图像中边缘等深度不连续的位置能较好地恢复出来;而在平滑区域,有些算法出现了纹理转移的现象,而所提算法则没有。可见,所提算法对于纹理转移和深度流失有明显效果。

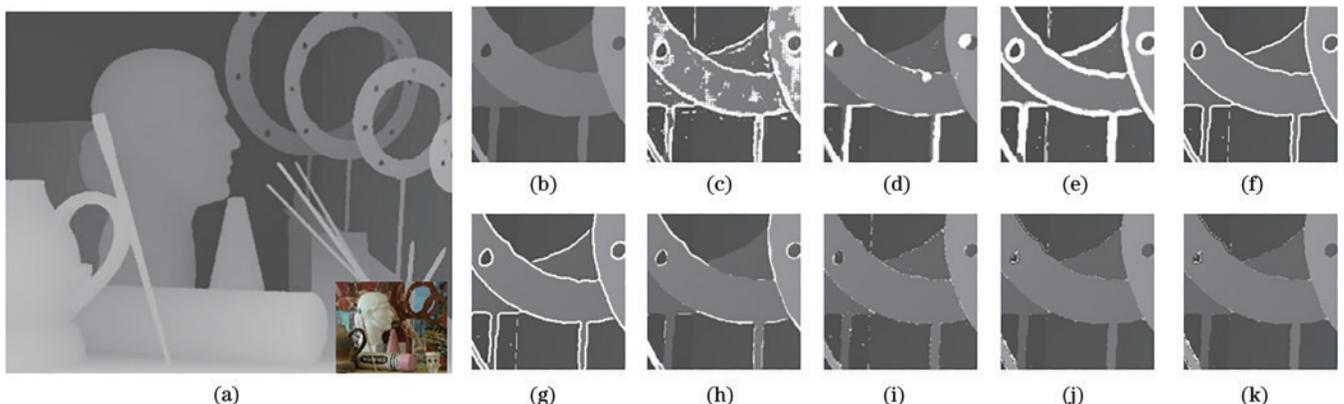


图 7 不同算法处理后的 Art 深度图像的超分辨率结果($4\times$)。(a)原图;(b)局部真实图;(c)TGV 算法;(d) RMRF 算法;(e)GF 算法;(f)JBU 算法;(g)MSG-Net 算法;(h)FDKN 算法;(i)RCAN 算法;(j)DepthSRNet 算法;(k)所提算法

Fig. 7 Super-resolution results of Art depth images processed by different algorithms($4\times$). (a) Original drawings; (b) partial true picture; (c) TGV algorithm; (d) RMRF algorithm; (e) GF algorithm; (f) JBU algorithm; (g) MSG-Net algorithm; (h) FDKN algorithm; (i) RCAN algorithm; (j) DepthSRNet algorithm; (k) proposed algorithm

在 $2\times$ 、 $4\times$ 、 $8\times$ 和 $16\times$ 放大倍数下得到的深度图像重建结果如表 1 所示。通过实验结果可以看出,基

于深度学习的方法要明显优于传统方法和基于滤波的方法。所提方法与其他基于深度学习方法相比指标有

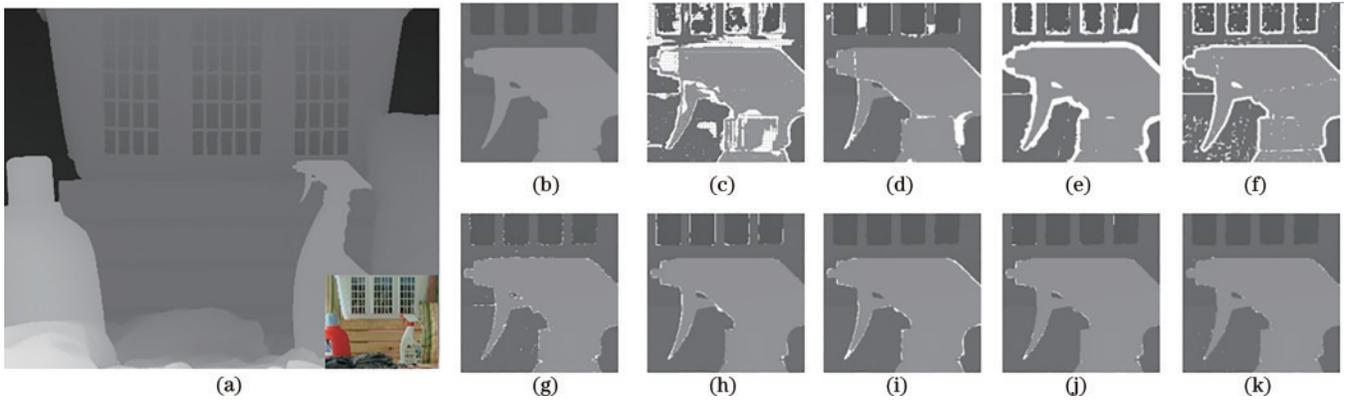


图 8 不同算法处理后的 Laundry 深度图像的超分辨率结果(4×)。(a)原图;(b)局部真实图;(c)TGV 算法;(d) RMRF 算法;(e)GF 算法;(f)JBU 算法;(g)MSG-Net 算法;(h)FDKN 算法;(i)RCAN 算法;(j)Depth-SR 算法;(k)所提算法

Fig.8 Super-resolution results of Laundry depth images processed by different algorithms(4×). (a) Original drawings; (b) Partial true picture; (c) TGV algorithm; (d) RMRF algorithm; (e) GF algorithm; (f) JBU algorithm; (g) MSG-Net algorithm; (h) FDKN algorithm; (i) RCAN algorithm; (j) Depth-SR algorithm; (k) proposed algorithm

表 1 超分辨率重建后 RMSE/PSNR(dB)指标结果

Table 1 RMSE/PSNR(dB) index results after super-resolution reconstruction

Method	Scale	Art	Books	Dolls	Laundry	Moebius	Reindeer	Average
TGV ^[3]	2×	3.16/38.14	1.33/45.62	1.17/46.77	1.87/42.69	1.14/46.99	2.40/40.53	1.85/43.46
	4×	3.73/36.70	1.67/43.68	1.42/45.09	2.25/41.09	1.45/44.90	2.65/39.67	2.20/41.86
	8×	7.12/31.08	2.27/41.01	2.05/41.90	4.05/35.98	2.41/40.49	4.33/35.40	3.71/37.64
	16×	12.08/26.49	4.89/34.34	4.44/35.18	8.01/30.06	5.41/33.47	9.05/29.00	7.31/31.42
RMRF ^[4]	2×	2.31/40.85	1.24/46.24	1.24/46.29	1.55/44.31	1.23/46.34	1.80/43.01	1.56/44.51
	4×	3.50/37.25	1.81/42.96	1.64/43.84	2.32/40.82	1.74/43.30	2.62/39.76	2.27/41.32
	8×	4.79/34.52	2.46/40.32	2.15/41.47	3.05/38.43	2.49/40.20	3.24/37.93	3.03/38.81
	16×	7.18/31.01	3.32/37.71	2.97/38.67	4.46/35.14	3.28/37.80	4.51/35.03	4.29/35.89
GF ^[14]	2×	3.15/38.16	1.43/44.99	1.22/46.40	2.01/42.09	1.23/46.33	2.25/41.11	1.88/43.18
	4×	3.90/36.30	1.76/43.21	1.64/43.83	2.47/40.29	1.73/43.37	2.82/39.14	2.38/41.02
	8×	5.43/33.44	2.38/40.59	2.26/41.05	3.44/37.41	2.46/40.31	3.95/36.20	3.49/38.17
	16×	8.15/29.90	3.34/37.65	3.45/37.37	5.03/34.10	4.16/35.75	5.83/32.81	4.99/34.60
JBU ^[15]	2×	2.69/39.53	1.11/47.22	1.02/47.96	1.60/44.05	0.91/48.95	1.87/42.69	1.53/45.07
	4×	4.04/36.00	1.88/42.65	1.49/44.67	2.64/39.70	1.52/44.49	2.86/39.00	2.41/41.09
	8×	5.23/33.76	2.49/40.21	1.85/42.79	3.44/37.40	2.18/41.36	3.60/37.00	3.13/38.75
	16×	7.13/31.07	3.96/36.18	2.52/40.10	5.96/32.63	3.08/38.36	4.38/35.30	4.51/35.61
MSG-Net ^[6]	2×	0.66/51.74	0.37/56.77	0.37/56.77	0.79/50.18	0.31/58.30	0.42/55.67	0.49/54.91
	4×	1.47/44.78	0.67/51.61	0.73/50.86	0.79/50.18	0.58/52.86	0.98/48.31	0.87/49.77
	8×	2.46/40.31	1.03/47.87	1.10/47.30	1.51/44.55	0.94/48.67	1.76/43.22	1.47/45.32
	16×	4.57/34.93	1.63/43.88	1.63/43.89	2.63/39.73	1.69/43.57	2.92/38.79	2.51/40.80
FDKN ^[16]	2×	1.53/44.44	0.50/54.15	0.70/51.23	0.88/49.24	0.60/52.57	1.09/47.38	0.88/49.84
	4×	2.10/41.67	0.73/50.86	0.93/48.76	1.26/46.12	0.79/50.18	1.50/44.61	1.22/47.03
	8×	3.16/38.14	1.21/46.48	1.30/45.85	2.00/42.11	1.24/46.26	2.27/41.01	1.86/43.31
	16×	9.46/28.61	3.93/36.24	3.21/38.00	5.95/32.64	4.26/35.54	6.53/31.83	5.56/33.81
RCAN ^[17]	2×	1.01/48.04	0.37/56.77	0.47/54.69	0.55/53.32	0.43/55.46	0.72/50.98	0.57/53.21
	4×	1.51/44.55	0.56/53.17	0.70/51.23	0.86/49.44	0.60/52.57	1.08/47.46	0.89/49.74
	8×	2.20/41.28	0.85/49.54	1.02/47.96	1.33/45.65	0.88/49.24	1.57/44.21	1.31/46.31
	16×							
Depth-SRNet ^[7]	2×	0.53/53.63	0.42/55.66	0.49/54.33	0.44/55.25	0.44/55.26	0.51/53.88	0.47/54.67
	4×	1.20/46.55	0.60/52.49	0.81/49.96	0.78/50.26	0.68/51.48	0.96/48.51	0.84/49.88
	8×	2.22/41.22	0.89/49.12	1.11/47.22	1.31/45.81	0.96/48.49	1.57/44.21	1.34/46.01
	16×	3.90/36.30	1.51/44.54	1.54/44.38	2.26/41.06	1.56/44.27	2.47/40.29	2.21/41.81
proposed method	2×	0.41/55.87	0.30/58.59	0.35/57.25	0.33/57.76	0.32/58.03	0.38/56.54	0.35/57.34
	4×	0.95/48.57	0.42/55.67	0.61/52.42	0.59/52.71	0.50/54.15	0.79/50.18	0.64/52.28
	8×	2.15/41.48	0.76/50.51	0.97/48.39	1.15/46.91	0.73/50.86	1.48/44.73	1.21/47.15
	16×	3.72/36.72	1.45/44.90	1.45/44.90	2.25/41.09	1.34/45.59	2.32/40.82	2.09/42.34

明显提升,尤其在 $2\times$ 和 $4\times$ 重建效果优势更明显:与传统方法中RMRF方法相比, $2\times$ 和 $4\times$ PSNR分别平均提升了12.83 dB和10.96 dB, RMSE分别平均下降了1.21和1.63;与滤波方法中的JBU方法相比, $2\times$ 和 $4\times$ PSNR分别平均提升了12.27 dB和11.19 dB, RMSE分别平均下降了1.18和1.77;与神经网络中的DepthSR-Net方法相比, $2\times$ 和 $4\times$ PSNR分别平均提升了2.67 dB和2.42 dB, RMSE分别平均下降了0.12和0.19。

3.2.2 消融实验

为了验证DCM、DGM和RAM的有效性,进行了消融实验。其中,M4为所提网络,采用了3个模块,M1、M2、M3分别为移除DCM、DGM、RAM的情况。

为保证消融实验的有效性,M1~M3使用的数据集和训练参数与M4保持一致,并统一对深度图Art进行 $4\times$ 超分辨率重建,得到消融实验结果如表2所示。

表2 消融实验结果
Table 2 Results of ablation experiment

Network	RMSE	PSNR /dB
M1	1.11	47.23
M2	1.07	47.54
M3	0.99	48.23
M4	0.95	48.57

由表2可以看出:与M4相比,由于无法将两种特征充分融合,M1的PSNR下降了1.34 dB;由于无法充分融合深度和强度特征,进而无法引导恢复重建,M2的PSNR下降了1.03 dB;由于RAM被移除,M3的PSNR下降了0.34 dB。因此,每一个模块都有助于提高网络性能,且DCM和DGM对超分辨率重建效果具有更为明显的提升作用。

4 结 论

提出基于双特征融合引导的深度图像超分辨率重建网络,利用深度特征和强度特征之间的关联性,改变以往简单拼接卷积的融合方式,提出双通道融合模块和双特征引导重建模块。首先,双通道融合模块利用通道注意力将两特征增强后充分融合,并利用融合后的特征引导恢复重建;其次,双特征引导重建模块可以通过自适应选择将深度特征和强度特征融合后引导特征恢复重建,有效利用深度特征对恢复重建的引导作用,同时其对纹理转移和深度流失现象有明显的抑制效果。实验结果表明:所提算法的超分辨率重建效果要优于其他算法;且所提算法可以提升重建效果,对彩色图像引导中常见的纹理转移和深度流失现象有明显改善。后续可以在简化网络结构方面进行研究,如加入残差结构提升网络训练速度等。

参 考 文 献

- [1] 杨宜林,李积英,王燕,等.基于改进双边滤波的深度图像修复算法研究[J].激光与光电子学进展,2020,57(16):161020.
Yang Y L, Li J Y, Wang Y, et al. Research on depth image repair algorithm based on improved bilateral filter [J]. Laser & Optoelectronics Progress, 2020, 57(16): 161020.
- [2] Diebel J, Thrun S. An application of Markov random fields to range sensing[C]//Proceedings of the 18th International Conference on Neural Information Processing Systems, December 5-8, 2005, Vancouver, British Columbia, Canada. New York: ACM Press, 2005: 291-298.
- [3] Ferstl D, Reinbacher C, Ranftl R, et al. Image guided depth upsampling using anisotropic total generalized variation[C]//2013 IEEE International Conference on Computer Vision, December 1-8, 2013, Sydney, NSW, Australia. New York: IEEE Press, 2014: 993-1000.
- [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] 许娇,袁三男.增强型多尺度残差网络的图像超分辨率重建算法[J].激光与光电子学进展,2023,60(4):0411002.
Xu J, Yuan S N. Image super-resolution reconstruction algorithm based on enhanced multi-scale residual network [J]. Laser & Optoelectronics Progress, 2023, 60(4): 0411002.
- [6] Hui T W, Loy C C, Tang X. Depth map super-resolution by deep multi-scale guidance[M]//Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9907: 353-369.
- [7] Guo C L, Li C Y, Guo J C, et al. Hierarchical features driven residual learning for depth map super-resolution[J]. IEEE Transactions on Image Processing, 2018, 28(5): 2545-2557.
- [8] 于淑侠,胡良梅,张旭东,等.彩色图像多尺度引导的深度图像超分辨率重建[J].光电工程,2020,47(4):42-51.
Yu S X, Hu L M, Zhang X D, et al. Color image multi-scale guided depth image super-resolution reconstruction [J]. Opto-Electronic Engineering, 2020, 47(4): 42-51.
- [9] Zhao Z X, Zhang J S, Xu S, et al. Discrete cosine transform network for guided depth map super-resolution [C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 18-24, 2022, New Orleans, LA, USA. New York: IEEE Press, 2022: 5687-5697.
- [10] Ido A, Israel C. Depth map super-resolution via cascaded transformers guidance[J]. Frontiers in Signal Processing, 2022, 2: 1-12.
- [11] Cui Y J, Liao Q M, Yang W M, et al. RGB guided depth map super-resolution with coupled U-net[C]//2021 IEEE International Conference on Multimedia and Expo (ICME), July 5-9, 2021, Shenzhen, China. New York:

- IEEE Press, 2021.
- [12] Huang Y J, Xu H X. Fully convolutional network with attention modules for semantic segmentation[J]. *Signal, Image and Video Processing*, 2021, 15(5): 1031-1039.
- [13] 孙瑞, 孙晓兵, 刘晓, 等. 基于注意力机制的偏振成像目标分类方法[J]. *光学学报*, 2021, 41(16): 1611004.
Sun R, Sun X B, Liu X, et al. Polarimetric imaging target classification method based on attention mechanism [J]. *Acta Optica Sinica*, 2021, 41(16): 1611004.
- [14] Yeo Y J, Sagong M C, Shin Y G, et al. Simple yet effective way for improving the performance of depth map super-resolution[J]. *Signal Processing Letters*, 2020, 27: 2099-2103.
- [15] He K M, Sun J, Tang X O. Guided image filtering[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(6): 1397-1409.
- [16] Kopf J, Cohen M F, Lischinski D, et al. Joint bilateral upsampling[J]. *ACM Transactions on Graphics*, 2007, 26(3): 96-es.
- [17] Kim B, Ponce J, Ham B. Deformable kernel networks for joint image filtering[J]. *International Journal of Computer Vision*, 2021, 129(2): 579-600.
- [18] Zhang Y L, Li K P, Li K, et al. Image super-resolution using very deep residual channel attention networks[EB/OL]. (2018-07-08) [2023-01-05]. <https://arxiv.org/abs/1807.02758>.