

分层特征融合的图像超分辨率算法

王建新 周思远 吴宏林

(长沙理工大学 计算机与通信工程学院, 湖南 长沙 410114)

摘要: 针对传统单一数据集训练重建的超分辨率图像伪影明显、细节丢失的问题, 笔者提出一种分层特征融合的图像超分辨率算法, 以消除伪影并增强细节特征。首先利用输入图像本身和外部训练图像集的特征信息形成的先验约束条件, 使用双网络模型分别完成相应的特征图重建; 其次使用潜在低秩表示法, 在去除噪声的同时将特征图分解成低秩部分和显著部分; 最后利用小波变换快速融合上一层次的分解图。实验结果表明, 该算法增强了重建图像的纹理特征, 提高了重建图像的质量和峰值信噪比, 改善了视觉效果。

关键词: 超分辨率; 特征融合; 潜在低秩表示; 小波变换

中图分类号: TP391.41 **文献标识码:** A **文章编号:** 1003-9767 (2022) 07-126-05

Image Super-Resolution Algorithm Based on Hierarchical Feature Fusion

WANG Jianxin, ZHOU Siyuan, WU Honglin

(School of Computer and Communication Engineering, Changsha University of Science and Technology,
Changsha Hunan 410114, China)

Abstract: Aiming at the problem of obvious artifacts and loss of details in super-resolution images reconstructed by traditional single dataset training, the author proposes an image super-resolution algorithm based on hierarchical feature fusion to eliminate artifacts and enhance detailed features. Firstly, using the prior constraints formed by the feature information of the input image itself and the external training image set, the corresponding feature map reconstruction is completed using the dual network model; secondly, the potential low-rank representation is used to decompose the feature map into Low-rank part and significant part; finally, the decomposition map of the previous level is quickly fused by wavelet transform. The experimental results show that the algorithm enhances the texture feature of the reconstructed image, improves the quality and peak signal-to-noise ratio of the reconstructed image, and improves the visual effect.

Keywords: super-resolution; feature fusion; latent low-rank representation; wavelet transform

0 引言

图像超分辨率 (Super-Resolution, SR) 技术利用先验知识, 通过输入的模糊、低分辨率图像 (Low-Resolution, LR) 恢复缺失的细节信息, 重建出对应的高分辨率图像 (High-Resolution, HR)。图像超分辨率重建提供了一种方法去弥补传感器设备的不足, 还可以增强细节缺失、模糊的现有图像。LR 图像中包含的像素信息极少, 使得超分辨率重建结果通常对应多个 HR 图像, 这是一个不适定问题^[1]。为了解决这个不适定问题, 许多研究人员基于一些假设和先验知识提出了相应的超分辨率方法。对于现有的超分辨率方

法, 根据重建方法大致可以分为 3 类, 即基于插值的 SR 方法、基于重建的 SR 方法以及基于学习的 SR 方法^[2]。

近年来, 随着高性能计算机和深度卷积网络的快速发展, 基于学习的 SR 方法受到了许多学者的关注, 并在实际的应用中取得了显著的成果。针对低分辨率图像自身结构和细节信息的独特性, ZSSR 算法利用输入图像提供的信息, 通过字典学习生成相应的 LR/HR 图像集, 以训练特定图像的卷积神经网络 (Convolutional Neural Networks, CNN)^[3]。YANG 等人基于图像自身信息, 通过对输入图像进行训练生成一个金字塔模型, 得到 LH 和 HR 之间的对应关系^[4]。

基金项目: 大数据环境下智慧城市动态交通网络调度互联的关键问题研究 (项目编号: 202042)。

作者简介: 王建新 (1991—), 男, 湖南道县人, 硕士研究生, 助教。研究方向: 图像处理、机器学习、校园信息化建设。

YANG 等人借助外部训练图集通过稀疏表示学习得到一对 LR/HR 字典, 假设 LR 和 HR 共享稀疏系数, 通过压缩感知求出稀疏系数以重建出 HR 图像^[5]。CHAO 等人首次构建了一个基于深度学习的超分辨率神经网络 (Super-Resolution Convolutional Neural Network, SRCNN), 使用 3 层卷积神经网络通过端对端的训练得到 LR/HR 图像之间的映射关系^[6]。基于外部特征的重建方法对训练图集的依赖程度较高, 其细节纹理恢复过程中易产生伪影, 并且没有充分利用输入图像的丰富细节信息。而基于内部特征的重建方法不依赖于外部图像训练集, 只使用输入图像的自身像素信息, 其算法自适应能力差, 重构的图像伪影严重^[7]。

综上所述, 通过深入分析和研究基于学习的超分辨率方法, 针对传统图像超分辨率方法未能充分利用输入图像的结构信息和细节特征导致重建图像伪影明显、细节丢失的问题, 笔者提出一种分层特征融合的图像超分辨率算法。首先使用双网络重建模型, 利用输入图像和外部训练图集作为先验知识, 分路径完成特征图的重建; 其次为了充分利用内外特征信息, 使用潜在低秩表示法从特征图像中提取全局结构信息和局部结构信息, 在去除噪声部分的同时将图像分解为低阶部分和显著部分; 最后采用小波变换融合的方式将两幅图的低阶部分、显著部分进行融合, 从而得到目标 HR 图像。实验结果表明, 本文提出的方法重建出的图像质量和视觉效果都比较好, 而且可以通过融合图像的内部特征和外部特征来保留更多的高频信息。

1 相关工作

1.1 基于图像自身特征的重建模型

图像本身的自相似性意味着自然图像具有很强的内部数据重复性, 通常包含重复的视觉内容^[8]。单个图像上的信息片段是自然图像的 1 个非常单一的特征, 无论数据集有多大, 都无法在任何外部数据库中找到。为了充分利用低分辨率图像的内部细节特征, ZHU 等人提出了基于稀疏表示和自相似的超分辨重建网络模型, 这里将其称之为内部特征增强模型 (CS-SR)^[9]。该网络利用输入的低分辨率图像信息, 通过稀疏表示的字典学习将输入图像训练得到一组线性原子, 从而恢复出目标高分辨率图像。该网络模型的重建流程主要分为 3 个阶段, 分别是字典训练阶段、稀疏系数求解阶段以及特征图重建阶段。

(1) 字典训练阶段。将输入低分辨率图 $y=N \times N$ 插值放大 m 倍, 得到细节信息缺失的伪 HR 图 $Y_D=mN \times mM$ 。对于图像中的低频区域, 双三次插值方法可以重建出较好的效果, 但由于图像的高频信息并没有得到恢复, 因此重建出的还是低分辨率图像。利用自然图像中包含重复视觉内容的特点, 采取样本均方误差法的阈值去选取更具代表性的图像块, 经 K 次奇异值分解 (Singular Value Decomposition, SVD) 完成字典训练学习。

(2) 稀疏系数求解阶段。在稀疏表示算法的理论中 LR 块 X_l 和 HR 块 X_h 在过完备字典 D_h/D_l 中都可以由同一个稀疏系数 α 线性结合表示出来, 而稀疏系数的求解通常采用正交匹配追踪算法 (Orthogonal Matching Pursuit, OMP)。

(3) 特征图重建阶段。将训练得到的高分辨率字典 D_h 和求解得到的稀疏系数 α 结合, 从而重建出高分辨率图像 Y_1 。以上步骤的表示公式如下:

$$SMSE(x_i) = \frac{\sum_{j=1}^{m^2 N^2} \left(x_{ij} - \sum_{j=1}^{m^2 N^2} x_{ij} / m^2 N^2 \right)^2}{m^2 N^2 - 1} \quad (1)$$

$$\min_{\alpha} \|x_i' - F \cdot D_h \alpha\|_2, \|\alpha\|_0 \leq Z \quad (2)$$

$$Y_1 = D_h \alpha_i \quad (3)$$

式中, m 表示放大倍数, N 表示 LR 图像块的尺寸大小, x_i 表示 HR 图像块 X_h 中的列向量, x_i' 表示 LR 图像块 X_l 中的第 i 个列向量, $\|\cdot\|_0$ 表示 0 范数, $\|\cdot\|_2$ 表示 2 范数的平方, x_{ij} 表示列向量 x_i 中的第 j 个元素, F 表示下采样和模糊化操作, Z 表示字典中的原子数。

1.2 基于外部图像特征的重建模型

结合外部训练图集丰富的高频信息和稀疏表示重建的优势, WANG 等人借鉴基于稀疏表示超分辨率重建的思想, 将深度学习和稀疏编码相结合, 提出了基于稀疏先验的图像超分辨率重建网络 (SCN)^[10]。该稀疏编码网络充分利用了图像的先验信息, 通过迭代收缩和阈值算法能够快速得到稀疏系数, 将稀疏系数和训练得到的 HR 字典结合得到 HR 块, 最后输入到重建层得到 HR 图像。SCN 网络模型流程如图 1 所示。

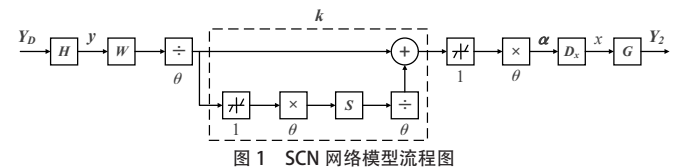


图 1 SCN 网络模型流程图

根据图 1, 在模型的预处理阶段, 将外部图像集作为训练图像学习得到一个 LR/HR 字典对 D_y 和 D_x 。将插值放大的图像 Y_D 作为输入图像, 经过一个由卷积神经网络构成的特征提取层 H 后得到该图像的多维特征向量 y 。其中特征提取层 H 由 m_y 个大小为 $S_y \times S_y$ 的滤波器组成, 得到的多维特征向量 y 具有 m_y 维。将特征向量 y 和 LR 字典 D_y 作为输入, 由迭代收缩和阈值算法快速求解得到 LR/HR 共享的稀疏系数 α 。虚线框内的稀疏系数求解网络是一个有限循环网络, 在求解的循环过程中, 稀疏系数的中间量可以表示为:

$$z_{k+1} = h_{\theta}(W_y + K \times z_k) \quad (4)$$

$$\alpha = \arg \min_z \|y - D_y z\|_2^2 + \lambda \|z\|_1 \quad (5)$$

式中, W_y 和 K 表示线性权重, h_{θ} 表示单元收缩函数。稀疏系数求解网络的下一层是一个由卷积神经网络构成的重建层 G , 在这一层中稀疏系数 α 与 HR 字典 D_x 线性结合得到 HR 图像。

2 分层特征融合图像超分辨率算法

针对基于外部特征的单一先验知识约束 SR 重建的不足, 本文利用潜在低秩特征提取法和小波变换在图像处理领域的独特优势提出了分层特征融合的图像超分辨率算法。本文提出的 SR 重建算法主要分为 3 个阶段, 分别是基于内外图像特征的重建阶段、图像特征提取阶段、图像融合阶段。整体模型如图 2 所示。

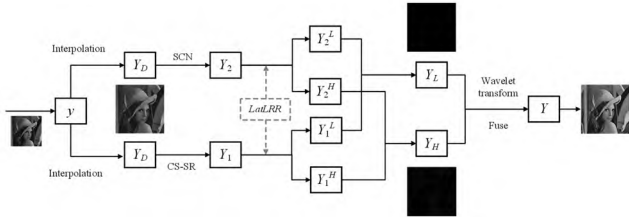


图2 分层特征融合的图像超分辨率模型

2.1 潜在低秩表示算法

潜在低秩表示法 (Latent Low-Rank Representation, LatLRR) 作为无监督的特征提取算法, 可以有效提取图像中的全局结构和局部结构信息, 提供了子空间分割和特征提取的解决方案^[11]。LatLRR 分解将原图像分解为低秩部分、显著部分、噪声部分, 其表达式为:

$$Y = YZ + LY + E \quad (6)$$

式中, Y 是输入图像, Z 是低秩系数, L 是显著系数, YZ 、 LY 分别表示低秩图层和显著图层, E 表示稀疏噪声。

将 LatLRR 分解问题进行最优化求解, 其表达式为:

$$\min_{Z, L, E} \|Z\|_* + \|L\|_* + \lambda \|E\|_1 \quad (7)$$

式中, $\|\cdot\|_*$ 表示矩阵奇异值之和的核范数, $\|\cdot\|_1$ 表示 L_1 范数, λ 表示平衡系数 ($\lambda > 0$)。利用增强拉格朗日乘子对式 (7) 进行求解, 可以依次推导出低秩表示部分、显著性部分以及稀疏噪声部分。在图像融合过程中可以直接对显著部分和低秩部分进行融合, 忽略噪声部分。

2.2 基于小波变换的图像融合

为了综合内外特征 SR 重建图像的优势, 提高重建图像的质量, 本文在两幅特征图像的融合阶段采用了离散小波变换的融合算法。通过潜在低秩表示算法将输入图像分解成高频部分和低频部分, 将两幅特征图分解成的高频部分与高频部分小波融合、低频部分与低频部分小波融合, 然后将融合后的低频部分和高频部分的系数进行逆变换得到融合图像。

离散小波变换是指在连续小波变换的基础上将窗函数和时间平移等参数进行离散化处理后得到的小波变换。常见的连续小波变换表达式为:

$$\varphi_{a,b}(x) = \frac{1}{\sqrt{|a|}} \varphi\left(\frac{x-b}{a}\right) \quad (8)$$

$$W_{2^j} f(x) = f * \varphi_{2^j}(x) = \frac{1}{2} \int_{\mathbb{R}} f(t) \varphi^*\left(\frac{x-t}{2^j}\right) dt \quad (9)$$

式中, $\varphi_{a,b}(x)$ 表示连续小波变换的基函数, a 表示伸缩因子, b 表示平移因子。为了降低连续小波变换的冗余性, 需要对基函数中的参数 a 、 b 进行二进制离散化处理, 即 $a=2^{-j}$ 、 $b=k*2^{-j}$, 从而得到二进制离散小波变换的公式^[12]。

离散小波变换是一种快速的多尺度图像融合方法, 将图像看作是由灰度值构成的一个二维矩阵, 在一幅图像中的每一个点 (n_1, n_2) 都有一个图像灰度值 $f(n_1, n_2)$ 与之对应。假设二维尺度函数可以分解, 即:

$$\varphi(n_1, n_2) = \varphi(n_1) \varphi(n_2) \quad (10)$$

式中, $\varphi(n_1)$ 和 $\varphi(n_2)$ 为两个一维尺度函数。将潜在低秩表示法得到的显著部分和低秩部分图像作为源图像, 选取适合的基函数对源图像进行离散小波变换, 得到的子图像采用高频与高频、低频与低频进行融合处理, 将融合后的图像进行逆变换即可得重建图像。基于小波变换的图像融合流程如图 3 所示。

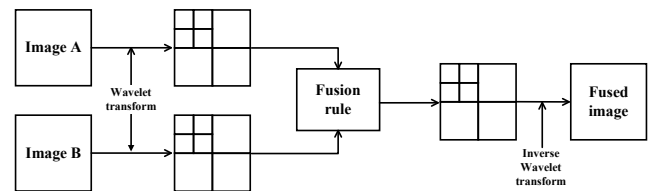


图3 基于小波变换的融合模型

3 实验与分析

在实验中为了增强对比性, 通过对 HR 图像使用插值缩放的方式生成用于字典训练和实验输入的 LR 图像。为了对比不同 SR 算法的优劣, 从重建图像质量和算法运行时间两个方面进行评估。对于图像质量的比较, 采用峰值信噪比 (Peak Signal to Noise Ratio, PSNR) 值进行评估, 单位为分贝, 图像的 PSNR 值越大表示质量越好。在本次实验对比中采用双三次差值放大法作为基准方法 (Bicubic) 做对比, 选取常见的 5 种具有代表性的 SR 算法和本文所提算法 (Proposed) 进行对比实验, 分别是 YANG J 等人的 ScSR 方法^[5]、ZHU Z 等人的 CS-SR 方法^[9]、GU S 等人的 CSC 方法^[13]、CHAO D 等人的 SRCNN 方法^[6]以及 WANG Z 等人的 SCN 方法^[10]。实验的对比图像选取 ARCNN 方法图像集 set5 中的 mlena 和 butterfly, 通过进行放大两倍的超分辨率重建得到不同算法的实验结果 (参与对比实验的算法源代码来自原作者的个人主页), 如图 4 和图 5 所示。

由图 4、图 5 可知, 常见的双三次插值法运算量小, 重建得到的图像边缘信息缺失、伪影现象十分明显; ScSR 方法重建的图像结构平滑、伪影严重。CS-SR 方法重建出的图像表面平滑、边缘模糊, 锯齿现象明显; CSC 方法的生成图像边缘平滑、有伪影, 细节缺失; SRCNN 方法和 SCN 方法重建的图像图像边缘模糊、有锯齿现象, 伪影明显。本文所提算法综合了内外特征的重建优势, 重建出的图像边缘锋利、纹理清晰, 视觉效果更好。表 1、表 2 是测试图像在不同 SR 算法中放大两倍时的 PSNR 值和运行时间。

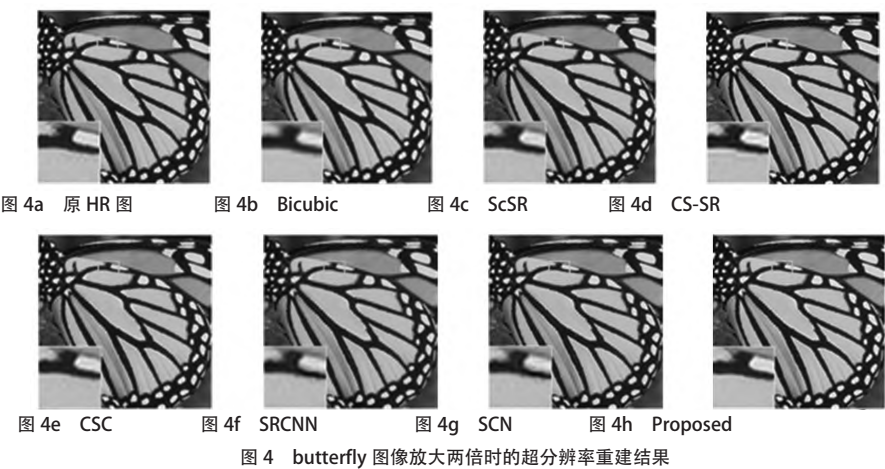


图 4 butterfly 图像放大两倍时的超分辨率重建结果

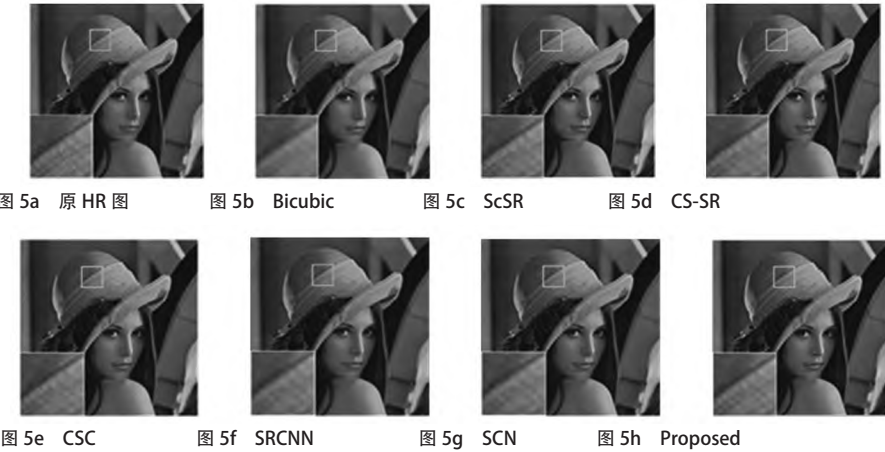


图 5 mlena 图像放大两倍时的超分辨率重建结果

表 1 不同 SR 算法重建的时间 (单位: s)

图像	bicubic	ScSR	CS-SR	CSC	SRCNN	SCN	Proposed
mlena	0.114	54.083	1.717	24.419	2.689	7.113	2.258
bird	0.121	51.141	1.763	24.544	2.686	7.105	2.359
pepper	0.102	54.451	1.669	24.443	2.704	7.142	2.469
baby	0.109	56.988	1.712	24.537	2.696	7.186	2.467
butterfly	0.105	48.761	1.725	24.727	2.665	7.035	2.330
head	0.117	55.527	1.732	24.479	2.676	7.094	2.200

表 2 不同 SR 算法重建结果的 PSNR 指标 (单位: dB)

图像	bicubic	ScSR	CS-SR	CSC	SRCNN	SCN	Proposed
mlena	32.14	34.95	35.34	35.57	32.02	34.50	35.82
bird	35.45	38.81	37.42	39.28	34.12	39.26	39.77
pepper	29.54	31.45	30.66	32.47	30.03	32.41	32.66
baby	38.83	38.84	38.44	39.25	32.81	39.24	39.65
butterfly	27.46	31.33	30.57	32.51	27.94	30.50	39.94
head	34.86	36.12	35.55	36.21	33.78	35.70	36.28

4 结 语

分层特征融合的图像超分辨率算法利用现有的网络模型学习输入图像的自相似性特征和外部训练图像集的丰富高频特征，分别完成图像的特征图重建。利用潜在低秩表示法对特征图进行分解得到显著部分和低秩部分，同时过滤特征图重建过程中产生的噪声。利用小波变换的融合方法，快速有效地综合不同层次特征图之间的细节信息。对比实验结果表明，该算法在客观评价指标和视觉效果上都有所提升，所设

计的网络表现出了稳定的训练性能，同时具有较强的鲁棒性。

参考文献

[1]WANG H,JIANG K.Research on image super-resolution reconstruction based on transformer[C]//2021 IEEE International Conference on Artificial Intelligence and Industrial Design,2021:226-230.

(下转第 134 页)

并取得了不错的去噪效果;最后,使用去噪后的深度图与原RGB图像训练了新的显著性目标检测网络。新的显著性目标检测网络在绝大部分数据集和评分上取得了比不经过去噪直接训练显著性目标检测网络更好的效果,证明了本文提出的方法的有效性。

参考文献

- [1] BORJI A, CHENG M M, HOU Q, et al. Salient object detection: a survey[J]. Computational visual media, 2019, 5(2): 117-150.
- [2] ZHOU T, FAN D P, CHENG M M, et al. RGB-D salient object detection: a survey[J]. Computational Visual Media, 2021, 7(1): 37-69.
- [3] LIU T, SUN J, ZHENG N N, et al. Learning to detect a salient object[C]//2007 IEEE Conference on Computer Vision and Pattern Recognition, 2007: 1-8.
- [4] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 3431-3440.
- [5] QU L, HE S, ZHANG J, et al. RGBD salient object detection via deep fusion[J]. IEEE Transactions on Image Processing, 2017, 26(5): 2274-2285.
- [6] LIU N, ZHANG N, WAN K, et al. Visual saliency transformer[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 4722-4732.
- [7] LU F, YU H, OH J. Domain adaptive monocular depth estimation with semantic information[Z]. 2021.
- [8] FAN D P, GONG C, CAO Y, et al. Enhanced-alignment measure for binary foreground map evaluation[Z]. 2018.
- [9] FAN D P, CHENG M M, LIU Y, et al. Structure-measure: a new way to evaluate foreground maps[C]// Proceedings of the IEEE Conference on Computer Vision, 2017: 4548-4557.
- [10] YU S, ZHANG B, XIAO J, et al. Structure-consistent weakly supervised salient object detection with local saliency coherence[C]// Proceedings of the AAAI Conference on Artificial Intelligence, 2021: 196-201.
- [11] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16×16 words: transformers for image recognition at scale[Z]. 2020.
- [2] YANG J, WRIGHT J, HUANG T, et al. Image super-resolution as sparse representation of raw image patches[C]// 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2008: 1-8.
- [3] SHOCHER A, COHEN N, IRANI M. Zero-Shot super-resolution using deep internal learning[C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 118-126.
- [4] YANG C Y, HUANG J B, YANG M H. Exploiting self-similarities for single frame super-resolution[C]// Asian Conference on Computer Vision, 2010: 497-510.
- [5] YANG J, WRIGHT J, HUANG T S, et al. Image super-resolution via sparse representation[J]. IEEE Transactions on Image Processing, 2010, 19(11): 2861-2873.
- [6] CHAO D, CHEN C L, HE K, et al. Learning a deep convolutional network for image super-resolution[C]// European Conference on Computer Vision, 2014: 184-199.
- [7] 陈文静, 唐轶. 基于深度学习的图像超分辨率重建方法综述[J]. 云南民族大学学报(自然科学版), 2019, 28(6): 597-605.
- [8] 陈泽乾. 基于内外特征融合的单图像超分辨率重建方法研究[D]. 北京: 华北电力大学, 2021: 25-28.
- [9] ZHU Z, GUO F, YU H, et al. Fast single image super-resolution via self-example learning and sparse representation[J]. IEEE Transactions on Multimedia, 2014, 16(8): 2178-2190.
- [10] WANG Z, LIU D, YANG J, et al. Deep networks for image super-resolution with sparse prior[C]// 2015 IEEE International Conference on Computer Vision, 2015: 370-378.
- [11] LIU G, YAN S. Latent low-rank representation for subspace segmentation and feature extraction[C]// 2011 International Conference on Computer Vision, 2011: 1615-1622.
- [12] 杨娇. 基于小波变换的图像融合算法的研究[D]. 北京: 中国地质大学(北京), 2014: 15-17.
- [13] GU S, ZUO W, XIE Q, et al. Convolutional sparse coding for image super-resolution[C]// IEEE International Conference on Computer Vision, 2015: 1823-1831.

(上接第129页)