Postprint: Joint Side and Main Path Decoding Feature Learning Method for Multiple Description Coding Image Enhancement

Authors: Zhao Lijun, Cao Congying, ZHANG Jinjing, Bai Huihui, Zhao Yao, Wang Anhong

Date: 2022-05-10T11:22:57+00:00

Abstract

We propose a multi-description coding image enhancement method based on joint side and central decoding feature learning, which simultaneously considers both side decoding image enhancement and central decoding image enhancement, thereby enabling better network training through joint optimization of features for central and side decoding. First, considering the characteristics of independent side decoding and joint central decoding in multi-description coding, we propose a shared side low-resolution feature extraction network to effectively extract features from two side-decoded images with identical content but different details, and simultaneously design a residual recursive compensation network structure that is applied to both side and central low-resolution feature extraction networks. Second, we design a multi-description side upsampling reconstruction network that adopts a partial network layer parameter sharing strategy, which can reduce network model parameters while improving network generalization capability. Finally, we propose a multi-description central upsampling reconstruction network that performs deep feature fusion of two side low-resolution features with central low-resolution features to achieve enhancement of multi-description compressed images. Extensive experimental results demonstrate that the proposed method outperforms many image enhancement methods such as ARCNN, FastARCNN, DnCNN, WSR, and DWCNN in terms of model complexity, objective quality, and visual quality evaluation.

Full Text

Preamble

Multiple Description Coding Image Enhancement Method with Joint Learning of Side- and Central-Decoding Features

Zhao Lijun[1]†, Cao Congying[1], Zhang Jinjing[2], Bai Huihui[3], Zhao Yao[3], Wang Anhong[1]
(1. College of Electronic Information Engineering, Taiyuan University of Science & Technology, Taiyuan 030024, China;
2. College of Big Data, North University of China, Taiyuan 030051, China;
3. Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China)

Abstract: This paper proposes a multiple description coding (MDC) image enhancement method that jointly learns side- and central-decoding features. By simultaneously addressing both side-decoding and central-decoding image enhancement, the method achieves superior network training through joint optimization of central and side decoding features. First, considering the characteristics of MDC—namely independent side decoding and joint central decoding—we propose a network-shared side low-resolution feature extraction network to effectively extract features from two side-decoded images with identical content but differing details, while designing a residual recursive compensation network structure for both side and central low-resolution feature extraction. Second, we design a multiple description side up-sampling reconstruction network employing a partial layer parameter sharing strategy that reduces model parameters while improving generalization capability. Finally, we propose a multiple description central up-sampling reconstruction network that performs deep feature fusion of two side low-resolution features with the central low-resolution feature to enhance MDC-compressed images. Extensive experimental results demonstrate that the proposed method outperforms many image enhancement methods—including ARCNN, FastARCNN, DnCNN, WSR, and DWCNN—in terms of model complexity, objective quality, and visual quality assessment.

Key words: multiple description coding; deep learning; image enhancement; compression distortion; feature fusion

0 Introduction

Although modern communication systems provide substantial network bandwidth, network congestion frequently occurs in densely populated venues such as concert halls, football stadiums, and student dormitory complexes. Moreover, limited communication device resources often lead to significant packet loss over unreliable channels. While many existing efficient image compression standards can mitigate this issue, they cannot guarantee reliable data transmission. Unlike single description coding, multiple description coding (MDC) divides a source into multiple descriptions transmitted over different channels. If all description packets are correctly received at the decoder, high-quality images can be reconstructed through joint decoding. If only one description packet is received, a reasonably high-quality image can still be recovered through side decoding. Thus, MDC technology enables reliable image transmission.

Although MDC methods can reduce data volume, both central and side decoded images suffer from varying degrees of distortion, particularly severe distortion in received side images. Consequently, image compression artifact removal techniques are essential to improve MDC decoded image quality. These techniques generally fall into two categories: traditional compression artifact removal methods and deep learning-based approaches. For instance, Dabov et al. [?] proposed an enhanced sparse representation strategy in the transform domain, implementing image denoising through grouping and collaborative filtering. Foi et al. [?] introduced a shape-adaptive discrete cosine transform filtering method that defines a region-shape-adaptive transform to effectively remove blockiness and edge oscillation artifacts. Chang et al. [?] reduced JPEG compression artifacts through sparse representation and redundant dictionary learning, though this method cannot recover lost high-frequency information. Zhang et al. [?] proposed a non-convex low-rank model for image deblocking that explicitly transforms quantization constraints into a feasible solution space to constrain non-convex low-rank optimization without modifying existing codecs, solving the optimization problem through an adaptive parameter adjustment alternating minimization strategy.

AlexNet's victory in the ImageNet competition marked the dawn of modern deep learning. Subsequently, AlphaGo' s 4:1 victory over world-class Go player Lee Sedol brought widespread attention to convolutional neural networks (CNNs). Deep learning has achieved tremendous success in computer vision, addressing tasks such as image super-resolution, deraining, dehazing, and denoising. Deep learning-based compression artifact removal methods have also garnered significant research interest. For example, Yu et al. [?] proposed ARCNN (Artifacts Reduction Convolutional Neural Network), demonstrating that reusing shallow network parameters benefits deep network training. To address training difficulties in deep networks, Zhang et al. [?] proposed DnCNN (Denoising Convolutional Neural Network), a residual learning-based approach that constructs deep CNNs with batch normalization to improve convergence speed and denoising performance. To further enhance performance, Qiu et al. [?] combined signal processing-based image restoration with deep residual learning models for JPEG artifact removal. While these methods outperform traditional approaches, they fail to fully exploit contextual information for image quality enhancement. To address this limitation, Chen et al. [?] proposed a multi-scale dense residual network that introduces dilated convolutions with varying dilation factors into dense blocks of residual networks to achieve larger receptive fields.

Moving beyond single-domain neural network processing, Zhang et al. [?] proposed DMCNN (Dual-domain Multi-scale Convolutional Neural Network), which effectively utilizes global information to eliminate JPEG compression artifacts. Similarly, Zheng et al. [?] introduced IDCN (Implicit Dual-domain Convolutional Network) to reduce compression artifacts in color images. Although DMCNN and IDCN employ dual-branch network topologies, they do not adequately leverage high and low-frequency information for feature complementarity. To fully exploit these features, Jin et al. [?] proposed a flexible deep learning image restoration method that decomposes low-quality input images into low-frequency structure and high-frequency texture components, processes them through separate quality enhancement networks, uses texture features to enhance structure features, and finally merges the predicted high-quality texture and structure maps through an aggregation network.

To address gridding artifacts caused by pooling and dilated filtering, Liu et al. [?, ?] proposed MWCNN (Multi-level Wavelet Convolutional Neural Network), which demonstrates excellent performance in image denoising, single-image super-resolution, and JPEG artifact removal. To balance enhancement performance, network parameters, and inference time, Zhang et al. [?] proposed WSR (Wavelet Super-Resolution), a lightweight method using deformable convolution kernels to reduce parameters.

While these methods achieve good denoising performance, they cannot adaptively enhance images for different compression artifact levels, often requiring multiple trained denoising networks. This inevitably increases complexity and storage requirements, limiting widespread adoption. To address this, Li et al. [?] proposed a single-model compression artifact removal method for JPEG images across various quality factors, using separate restoration and global branches to address local oscillation artifacts, global block artifacts, and color shift. Additionally, Kirmemis et al. [?] proposed a BPG (Better Portable Graphics) artifact removal method that selects among three networks of different sizes, though choosing the optimal network remains challenging.

Beyond compressed image enhancement, researchers have also addressed video compression quality. For example, Zhou et al. [?] proposed a dual-network compressed video reconstruction method that first removes compression artifacts then applies super-resolution.

To address MDC image compression distortion, Xuan et al. [?] enhanced compressed images through adjacent keyframe estimation. Zhao et al. [?] combined pre- and post-processing techniques to create a new MDC framework compatible with standard codecs, significantly improving coding efficiency and decoded image quality. Similarly, Zhang et al. [?] obtained multiple single-description images through checkerboard downsampling, encoded them with standard codecs, and used CNNs to enhance single-path and central decoded image quality. Purica et al. [?] merged two low-resolution compressed video descriptions into one high-resolution description. Zhang et al. [?] reconstructed and enhanced received side descriptions through odd-even separation sampling. Zhu et al. [?] proposed a compression-constrained deblocking algorithm that effectively utilizes received dual-description information to reduce boundary artifacts in central decoded images. Xu [?] proposed a 3D-LVQ (3D Lattice Vector Quantization) based predictive decoding method to improve side image decoding performance. However, these deep learning models often fail to meet lightweight device requirements, necessitating research into low-complexity models.

To address compression artifacts in MDC images, particularly severe structural splitting artifacts in side decoded images, this paper proposes a joint side- and central-decoding feature learning method for MDC image enhancement (MDE). The contributions are summarized as follows:

a) To address the large storage and computational complexity of existing deep learning models, we design a residual recursive compensation network as the low-resolution feature extraction network for both side and central paths, employing parameter sharing to effectively extract features from two decoded images with identical content but differing details.

b) Considering the independent decoding characteristic of MDC side paths, we design a multiple description side up-sampling reconstruction network that also adopts partial layer parameter sharing to reduce model parameters and improve generalization.

c) Considering the joint decoding characteristic of MDC central paths, we design a multiple description central up-sampling reconstruction network that performs deep feature fusion of two side low-resolution features with the central low-resolution feature to enhance MDC-compressed images.

1 Proposed Multiple Description Compressed Image Enhancement Method

Although existing MDC methods effectively address reliable image transmission in unstable network environments, lossy MDC inevitably introduces various artifacts, noise, structural deformation, and structural splitting. Compared to traditional enhancement techniques, deep learning-based methods better remove compression artifacts. However, existing deep learning models suffer from high computational complexity and large memory consumption, and can only address single-description enhancement. When directly applied to MDC tasks, they can only enhance side and central decoded images separately without joint feature decoding. Therefore, we propose a joint side- and central-decoding feature learning method for MDC image enhancement.

The proposed method first employs Multiple Description Random Offset Quantization (MDROQ) [?] to encode and decode input images, obtaining two distorted side decoded images and one central decoded image. As shown in Figure 1, the method divides MDC image enhancement into two stages: low-resolution feature extraction and high-resolution image reconstruction. The first stage includes two side low-resolution feature extraction networks and one central low-resolution feature extraction network. The second stage includes two side up-sampling reconstruction networks and one central up-sampling reconstruction network. Based on MDC' s independent side decoding and joint central decoding characteristics, we design a residual recursive compensation network structure for both side and central low-resolution feature extraction. Parameter sharing in side feature extraction effectively extracts convolutional features from two decoded images with identical appearance but different detail information. Additionally, the proposed side up-sampling reconstruction network adopts partial layer parameter sharing, significantly reducing total model parameters.

Unlike the side up-sampling reconstruction network, the central up-sampling reconstruction network performs deep feature fusion of two side low-resolution features with the central low-resolution feature for compressed image enhancement. The enhancement process can be expressed as:

⟦MATH_1⟧

where $i = 1, 2, 3$ represents side 1, central, and side 2 respectively, Y_i denotes the enhanced image, X_i is the input image, and R_i represents the residual map predicted by the reconstruction network.

1.1 Low-Resolution Feature Extraction Network

Unlike single-description image coding, MDC outputs multiple side and central decoded images. Therefore, our low-resolution feature extraction network comprises two types: side low-resolution feature extraction networks and a central feature extraction network. To prevent overfitting during deep neural network training and reduce learnable parameters, the side low-resolution feature extraction network employs a residual block parameter sharing strategy to effectively extract low-resolution features. Unlike previous methods, this parameter sharing strategy shares only partial blocks rather than entire side networks. During network recursion, recursive results are processed with 1$×$1 convolution, then added to the previous block' s output before feeding into the next recursive block. This effectively reduces parameters while maintaining branch differences. The network structure is shown in Figure 1. The central low-resolution feature extraction network adopts the same topology but with different learnable parameter values.

In both side and central low-resolution feature extraction networks, we first apply a convolution block operation—Convolution (Conv) + Batch Normalization (BN) + Parametric ReLU (PReLU) activation—denoted as Conb. Images are converted to convolutional features using stride-2 downsampling convolution to reduce computation, followed by multi-layer fusion via the proposed residual recursive compensation approach. This process uses five Residual Convolution Blocks (Resb) for sequential feature extraction with multiple channel-wise weighted average fusions. Each Resb comprises five operations: Conv + BN + PReLU + Conv + skip connection. After the fifth Resb, its output is fused with the initial convolutional feature and the previous four channel-wise weighted fusion features through channel-wise weighted fusion to obtain the low-resolution convolutional feature. Table 1 details each layer' s parameters.

Specifically, Conb1 performs downsampling and feature extraction. Its output feeds into Resb1 and Conv1. Resb1' s output is summed with Conv1' s output as Resb2' s input. In Conv2, Resb1' s and Resb2' s inputs are combined as Conv2' s input, then Resb2' s output is summed with Conv2' s output as Resb3' s input. Similar operations apply to Resb4 and Resb5. In Conv3, the sum of Resb1' s, Resb2' s, and Resb3' s inputs serves as Conv3' s input. Conv4 and Conv5 perform similar residual recursive compensation operations. Finally, the sum of Resb5' s output and Conv5' s output forms the low-resolution feature extraction network' s output, while the sum of both side networks' outputs feeds into the central up-sampling reconstruction network.

The two side low-resolution feature extraction networks can be expressed as:

⟦MATH_2⟧

where Z_i represents the feature map after the i-th side low-resolution feature extraction network, X_i is the i-th side decoded image, and $g_s(\cdot)$ denotes the side low-resolution feature extraction mapping function. Z_1 and Z_3 are linear combinations of feature images. Similarly, the central low-resolution feature extraction network is:

⟦MATH_3⟧

where Z_2 is the feature map after central low-resolution feature extraction, X_2 is the central decoded image, and $g_c(\cdot)$ is the central low-resolution feature extraction mapping function.

1.2 Side and Central Up-Sampling Reconstruction Networks

After feature extraction, we obtain two side low-resolution feature maps and one central low-resolution feature map. In the first side up-sampling reconstruction network, the first side low-resolution feature feeds into five cascaded convolution blocks to obtain reconstruction features, with skip connections introduced in the third block to facilitate gradient backpropagation. The final reconstruction feature feeds into a transposed convolution (ConvT) layer to produce the first side decoded enhanced image. The second side network operates similarly. In both side networks, deep convolutional layers adopt parameter sharing to reduce parameters while enhancing reconstruction.

Unlike side networks, the central up-sampling reconstruction network can utilize features from both side and central decoded images. Therefore, we design a central network that fuses these features atop the side network structure. Both networks use five convolution blocks, but the central network concatenates the fused side features with central low-resolution features along the channel dimension after skip connections. Deep layers in the central network do not share parameters with side networks, primarily because the input feature maps to the fourth convolution block differ significantly. Table 2 details each layer' s parameters.

The side up-sampling reconstruction networks' nonlinear mapping can be expressed as:

⟦MATH_4⟧

where Y_1 and Y_3 are the reconstructed images from the two side paths, Z_1 and Z_3 are the low-resolution features, and $f_s^1(\cdot)$ and $f_s^3(\cdot)$ are the side up-sampling reconstruction mapping functions. The central up-sampling reconstruction network' s mapping is:

⟦MATH_5⟧

where Y_2 is the reconstructed central image, Z_2 is the central low-resolution feature, and $f_c(\cdot)$ is the central up-sampling reconstruction mapping function. The complete nonlinear mapping for all paths is:

⟦MATH_6⟧

where Y_1, Y_2, Y_3 are the enhanced side 1, central, and side 2 images; X_1, X_2, X_3 are the input decoded images; R_1, R_2, R_3 are the residual maps; and $f_s^1(\cdot), f_c(\cdot), f_s^3(\cdot)$ are the mapping functions.

1.3 Loss Function

Common image reconstruction losses include content loss, structural dissimilarity loss, total variation loss, and gradient difference loss. For content loss, L1 or L2 norms are typically used. Research shows that L2-based MSE loss causes over-smoothing, while L1-based MAE loss produces results closer to the original image. Therefore, we adopt MAE loss for our enhancement task. The total loss is:

⟦MATH_7⟧

where $Loss_1$, $Loss_2$, and $Loss_3$ represent the enhancement losses for side 1, central, and side 2 decoded images respectively. I_i^1 and $\hat{I}_i^1$ are the i-th pixel of the original and predicted images for side 1, with n being the total pixel count. α and β are weights for side and central loss functions.

1.4 Algorithm Description

Our training and test datasets derive from [?], which uses 291 images from [?] and [?] to create training patches. Specifically, 91 images come from [?]' s training set and 200 from [?]' s BSDS500 training set. Through cropping, downsampling, and stitching, [?] created 1,681 images of size 160$\times$160, denoted as Set-1681. We compress Set-1681 using MDROQ to obtain Set-1681(C), forming our training dataset. Different quantization parameter pairs (Qstep0,Qstep1) of (56,56.57), (96,96.57), (136,136.57), (176,176.57), and (216,216.57) produce varying distortion levels. Smaller pairs preserve more original information. While dataset size and type affect performance, all comparison methods use identical datasets for fairness.

Algorithm 1 describes MDE network training. First, compress Set-1681 using MDROQ to build the training dataset. Initialize network parameters: side low-resolution feature extraction parameters ζ, central low-resolution feature extraction parameters η, side up-sampling reconstruction parameters λ, and central up-sampling reconstruction parameters ξ. For each epoch, extract side low-resolution features via Equation (2) and central features via Equation (4).

Considering independent side and joint central decoding, perform reconstruction prediction via Equations (5)-(7). Jointly optimize side and central decoding features by updating parameters $\zeta, \eta, \lambda, \xi$ using gradient descent on the total loss from Equation (11). After training, save the final MDE model.

Algorithm 1: MDE Network Training Algorithm
Input: Set-1681 dataset (n=1,681 images), total iterations R=500, initial learning rate lr=2e-4, batch size b=8.
Output: Trained MDE network model.
1. Compress Set-1681 using MDROQ to obtain Set-1681(C); construct training dataset.
2. Initialize MDE parameters (ζ, η, λ, ξ).
3. For epoch = 1 to R:
a. For i = 1 to floor(n/b):
i. Extract side low-resolution features via Equation (2).
ii. Extract central low-resolution features via Equation (4).
iii. Perform side/central up-sampling reconstruction via Equations (5)-(7).
iv. Update parameters via gradient descent on Equation (11).
4. Save trained MDE model.

2 Experimental Results and Analysis

We compare MDE against state-of-the-art deep learning methods: ARCNN [?], FastARCNN [?], DnCNN [?], WSR [?], DWCNNV1 [?], DWCNNV1C [?], and DWCNNV2 [?]. For fair comparison, we train separate enhancement networks for side 1, side 2, and central paths using identical architectures but different parameters for each method. Since WSR is wavelet-based super-resolution, we remove its up-sampling layer for MDC enhancement. Test images from [?] (shown in Figure 2) evaluate performance using PSNR, SSIM, total parameters, receptive field size, and runtime. Visual quality comparisons are also provided.

2.1 Simulation Environment and Training Settings

We train and test our method using PyTorch on an NVIDIA RTX 2080Ti GPU. The ADAM optimizer uses an initial learning rate of 2e-4, updated every 100 iterations with a decay factor of 0.5. Batch size is 8, with total training iterations R=500.

2.2 Objective and Subjective Quality Evaluation

Tables 3-5 present PSNR and SSIM comparisons for side 1, side 2, and central enhanced images under different (Qstep0,Qstep1) values. The proposed method achieves significantly higher PSNR and SSIM than other methods across all quantization parameters. Table 6 compares parameter counts: ARCNN, FastARCNN, WSR, DWCNNV1, and DWCNNV1C have over twice our parameters while achieving lower PSNR/SSIM. Though DnCNN and WSR have similar parameter counts, our method delivers superior objective metrics.

Visual comparisons in Figures 3-5 show that ARCNN, FastARCNN, DnCNN, WSR, DWCNNV1, DWCNNV1C, and DWCNNV2 produce enhanced images with severe blurring artifacts, while our method yields clearer results with better visual quality, particularly for central images containing more details.

2.3 Complexity Analysis

Table 7 compares receptive field sizes. Our method's receptive field is larger than ARCNN, FastARCNN, DnCNN, DWCNNV1, and DWCNNV1C, but smaller than WSR and DWCNNV2. Table 8 examines performance with different receptive fields at (Qstep0,Qstep1)=(216,216.57). Performance improves with increasing receptive field up to 103, after which it saturates. Models with similar parameter counts show comparable runtime regardless of receptive field size. While larger receptive fields exploit more spatial correlations, excessively large fields weaken correlation between distant pixels and increase complexity. Table 9 compares average runtime: our method runs at 0.007 seconds per image, faster than ARCNN, DnCNN, WSR, DWCNNV1, DWCNNV1C, and DWCNNV2.

2.4 Ablation Experiments and Analysis

We conduct ablation studies on four aspects: residual block count, loss weight ratio, batch size, and learning rate.

Residual Blocks: Table 10 shows performance with 1, 3, 5, 7, and 9 residual blocks at (Qstep0,Qstep1)=(56,56.57). Performance stabilizes beyond 5 blocks, though parameters increase. We set the default to 5 blocks.

Loss Weights: Table 11 compares side:central loss weight ratios of 0.1:1, 1:0.1, and 1:1. The 1:1 ratio achieves the highest PSNR and SSIM.

Batch Size: Table 12 tests batch sizes of 4, 8, 12, and 16. The method is insensitive to batch size, with batch size 8 yielding the highest PSNR/SSIM.

Learning Rate: Table 13 tests learning rates of 1e-4, 2e-4, and 3e-4. The method is relatively insensitive. While 3e-4 yields slightly higher PSNR, its SSIM is lower than 2e-4. Considering human visual sensitivity to structure, we default to 2e-4.

3 Conclusion

This paper proposes a joint side- and central-decoding feature learning method for MDC image enhancement to address compression distortion. We introduce a residual recursive network for low-resolution feature extraction in both paths, employ parameter sharing based on decoding characteristics to extract features from two side images with identical content but differing details, and fuse side/central low-resolution features through up-sampling reconstruction networks. Extensive experiments demonstrate superiority over many deep learning enhancement methods in model complexity, objective quality, and visual quality. Future work will explore using a single deep learning model for MDC enhancement across different quantization parameters.

References

[1] Dabov K, Foi A, Katkovnik V, et al. Image denoising by sparse 3D transform domain collaborative filtering [J]. IEEE Trans on Image Processing, 2007, 16 (8): 2080–2095.

[2] Foi A, Katkovnik V, Egiazarian K. Pointwise shape-adaptive dct for high-quality denoising and deblocking of grayscale and color images [J]. IEEE Trans on Image Processing, 2007, 16 (5): 1395–1411.

[3] Chang Huibin, Michael K Ng, Zeng Tieyong. Reducing artifacts in jpeg decompression via a learned dictionary [J]. IEEE Trans on Signal Processing, 2014, 62 (3): 718–728.

[4] Zhang Jian, Xiong Ruiqin, Zhao Chen, et al. CONCOLOR: Constrained non-convex low-rank model for image deblocking [J]. IEEE Trans on Image Processing, 2016, 25 (3): 1246-1259.

[5] Dong Chao, Deng Yubin, Chen Chang Loy, et al. Compression artifacts reduction by a deep convolutional network [C]// Proceedings of the IEEE International Conference on Computer Vision. 2015: 576-584.

[6] Zhang Kai, Zuo Wangmeng, Chen Ynjin, et al. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising [J]. IEEE Trans on Image Processing, 2017, 26 (7): 3142–3155.

[7] Qiu Han, Zheng Qinkai, Memmi Gerard, et al. Deep residual learning based enhanced JPEG compression in the internet of things [J]. IEEE Trans on Industrial Informatics, 2020, 17: 2124-2133.

[8] 陈书贞，张祎俊，练秋生． 基于多尺度稠密残差网络的 JPEG 压缩伪迹去除方法 [J]. 电子与信息学报, 2019, 41 (10): 2479-2486. (Chen Shuzhen, Zhang Yijun, Lian Qiusheng. JPEG compression artifacts reduction algorithm based on multi-scale dense residual network [J]. Journal of Electronics and Information Technology, 2019, 41 (10): 2479-2486.)

[9] Zhang Xiaoshuai, Yang Wenhan, Hu Yueyu, et al. DMCNN: dual-domain multi-scale convolutional neural network for compression artifacts removal [C]// The 25th IEEE International Conference on Image Processing (ICIP). IEEE, 2018: 390-394.

[10] Zhang Bolun, Chen Yaowu, Tian Xiang, et al. Implicit dual-domain convolutional network for robust color image compression artifact reduction [J]. IEEE Trans on Circuits and Systems for Video Technology, 2020, 30: 3982-3994.

[11] Jin Zhi, Muhammad Zafar Iqbal, Bobkov Dmytro, et al. A flexible deep CNN framework for image restoration [J]. IEEE Trans on Multimedia, 2019, 22

(4): 1055-1068.

[12] Liu Pengju, Zhang Hongzhi, Lian Wei, et al. Multi-level wavelet convolutional neural networks [J]. IEEE Access, 2019, 7: 74973-74985.

[13] Liu Pengju, Zhang Hongzhi, Zhang Kai, et al. Multi-level wavelet-CNN for image restoration [C]// Proceedings of the IEEE conference on computer vision and pattern recognition workshops. 2018: 773-782.

[14] Zhang Huanrong, Jin Zhi, Tan Xiaojun, et al. Towards lighter and faster: Learning wavelets progressively for image super-resolution [C]// Proceedings of the 28th ACM International Conference on Multimedia. 2020: 2113-2121.

[15] Li Jianwei, Wang Yongtao, Xie Haihua, et al. Learning a single model with a wide range of quality factors for JPEG image artifacts removal [J]. IEEE Trans on Image Processing, 2020, 29: 8842–8854.

[16] Kirmemis O, Bakar G, Tekalp A M. Learned compression artifact removal by deep residual networks [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2018: 2602-2605.

[17] 周航, 何小海, 王正勇, 等. 采用双网络结构的压缩视频超分辨率重建 [J]. 电讯技术, 2020, 60 (01): 81-86. (Zhou Hang, He Xiaohai, Wang Zhengyong, et al. Super-resolution reconstruction of compressed video with dual network structure. Telecommunication Technology [J], 2020, 60 (01): 81-86.)

[18] 铉欣. 基于多描述标量量化的分布式视频编码 [D]. 西安: 西安电子科技大学, 2012.

[19] Zhao Lijun, Bai Huihui, Wang Anhong, et al. Multiple description convolutional neural networks for image compression [J]. IEEE Tran on Circuits and Systems for Video Technology, 2018, 29 (8): 2494-2508.

[20] Zhang Xiangjun, Wu Xiaolin. Standard-compliant multiple description image coding by spatial multiplexing and constrained least-squares restoration [C]// IEEE the 10th Workshop on Multimedia Signal Processing. IEEE, 2008: 349-354.

[21] Purica A, Boyadjis B, Pesquet P B, et al. A convex optimization framework for video quality and resolution enhancement from multiple descriptions [J]. IEEE Trans on Image Processing, 2018, 28 (4): 1661-1675.

[22] Zhang Ting, Bai Huihui, Zhang Mengmeng, et al. Standard-compliant multiple description image coding based on convolutional neural networks [J]. IEICE Trans on Information and Systems, 2018, 101 (10): 2584-2593.

[23] Zhu Shuyuan, He Zhiying, Meng Xiandong, et al. A new polyphase down-sampling-based multiple description image coding [J]. IEEE Trans on Image Processing, 29 (2020): 5596-5611.

[24] Xu Yuanyuan. Predictive side decoding for human-centered multiple description image coding [J]. EURASIP Journal on Wireless Communications and Networking, 2020 (2020): 1-14.

[25] Meng Lili, Liang Jie, Samarawickrama U, et al. Multiple description coding with randomly and uniformly offset quantizers [J]. IEEE Trans on Image Processing, 2013, 23 (2): 582-595.

[26] Zhao Lijun, Bai Huihui, WangAnhong, et al. Learning a virtual codec based on deep convolutional neural network to compress image. Journal of Visual Communication and Image Representation, 2019, 63: 102589-102589.

[27] Yang Jianchao, John W, Thomas S H, et al. Image super-resolution via sparse representation, IEEE Trans on Image Processing, 2010, 19 (11): 2861-2873.

[28] Pablo A, Michael M, Charless F, et al. Contour detection and hierarchical image segmentation [J]. IEEE Trans on pattern analysis and machine intelligence, 2010, 33 (5): 898-916.

Note: Figure translations are in progress. See original paper for figures.