Proceedings of the 2017 IEEE
International Conference on Information and Automation (ICIA)
Macau SAR, China, July 2017

# Soft Decoding of JPEG 2000 Compressed Images Using Bit-rate-driven Deep Convolutional Neural Networks

Xiaohai He*, Honggang Chen, Jingxu Chen, Linbo Qing
*College of Electronics and Information Engineering*
*Sichuan University*
*Chengdu 610065, China*
**Corresponding author: Xiaohai He, hxh@scu.edu.cn*

***Abstract*— Lossy image compression methods always introduce various unpleasant artifacts into the compressed results, especially at low bit-rates. In recent years, many effective soft decoding methods for JPEG compressed images have been proposed. However, to the best of our knowledge, very few works have been done on soft decoding of JPEG 2000 compressed images. Inspired by the outstanding performance of Convolution Neural Network (CNN) in various computer vision tasks, we presents a soft decoding method for JPEG 2000 by using multiple bit-rate-driven deep CNNs. More specifically, in training stage, we train a series of deep CNNs using lots of high quality training images and the corresponding JPEG 2000 compressed images at different coding bit-rates. In testing stage, for an input compressed image, the CNN trained with the nearest coding bit-rate is selected to perform soft decoding. Extensive experiments demonstrate the effectiveness of the presented soft decoding framework, which greatly improves the visual quality and objective scores of JPEG 2000 compressed images.**

***Index Terms*— Soft Decoding, JPEG 2000, Deep Convolutional Neural Networks, Bit-rate-driven.**

## I. Introduction

In general, images and videos are compressed using lossy compression algorithms to reduce the amount of data, thus decreasing the demand for transfer bandwidth and memory capacity. However, lossy compression methods (e.g., JPEG and JPEG 2000) inevitably cause distortion and introduce coding artifacts, including blocking, ringing, aliasing, etc. Therefore, it is necessary to study the way to reduce coding artifacts and improve the quality of compressed images, i.e., the soft decoding algorithms.

In recent years, researchers have presented a series of soft decoding algorithms for JPEG [1-8]. For example, Liu et al. exploited the sparsity of images in both pixel and DCT domains [1]. The authors of [2] and [3] utilized low-rank model. In [4], the soft decoding task was regarded as an inverse problem by linearizing the coding-decoding process. In [5], the graph-signal smoothness prior, sparsity prior, and Laplacian prior were integrated to perform soft decoding, leading to state-of-the-art performance. Overall, the algorithms mentioned above are time-consuming due to the complicated priors. By contrast, the learning-based artifacts reduction methods presented in [6-8] are more efficient in execution process and they also achieve competitive performance. With these soft decoding algorithms, the quality of JPEG compressed images can be well enhanced. In other words, we can obtain higher quality images at the cost of the same coding bits with JPEG.

Nevertheless, as far as we know, the restoration of JPEG 2000 compressed images is not well studied. As shown in [8], the CNN-based artifacts reduction framework can also be used on JPEG 2000 compressed images. In [9], Kwon et al. presented a semi-local gaussian processes-based artifacts reduction method for both JPEG and JPEG 2000 compressed images. In [10], the regression model was used to reduce the compression artifacts caused by JPEG 2000. According to the experimental results shown in [8-10], the quality enhancement on JPEG 2000 achieved by existing soft decoding methods is relatively constrained and there is still much room for further improvement.

Inspired by the remarkable ability of deep CNN on various image processing tasks [11-14], this paper presents a deep CNN-based soft decoding method for JPEG 2000 standard. On the basis of the networks used in [7] and [11], we construct a deep CNN model of 25 layers, in which we also incorporate some advanced technologies to improve the training speed and restoration performance, such as batch normalization (BN) [15], Rectifier Linear Unit (ReLU) [16], gradient clipping [17], and residual learning [18]. In addition, to improve the soft decoding performance further, we train a series of deep CNNs using the JPEG 2000 compressed images at different coding bit-rates, and the CNN model trained at the closest coding bit-rate with the input compressed image is used to conduct soft decoding.

The rest of this paper is organized as follows. Section II introduces the soft decoding algorithm for JPEG 2000. In Section III, experimental results are shown. The conclusion is presented in Section IV.

*This work is partially supported by the National Natural Science Foundation of China under Grant 61471248 and the Science and Technology Plan of Chengdu under Grant 2015-HM01-00293-SF.

978-1-5386-3154-6/17/$31.00 ©2017 IEEE

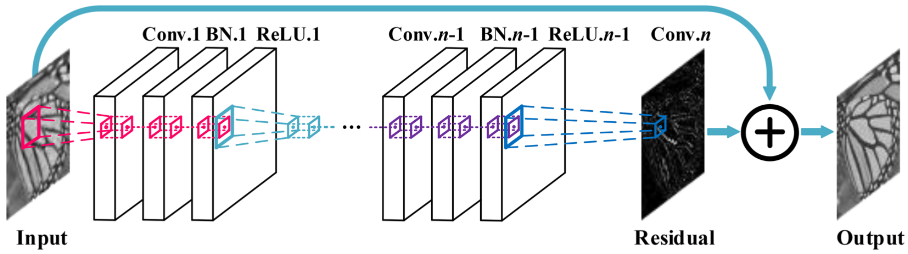



Fig. 1. The structure of the deep CNN used in our soft decoding method.

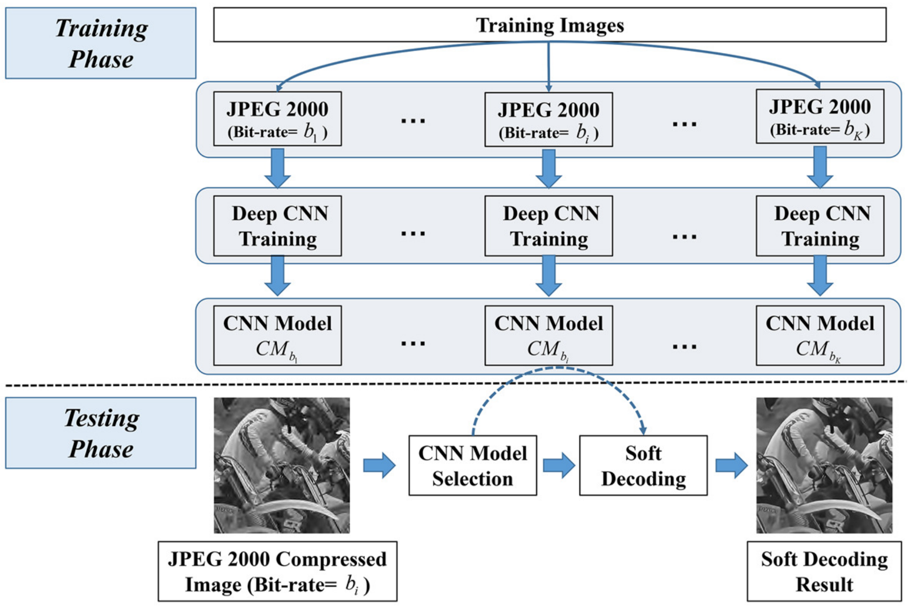



Fig. 2. The flowchart of the soft decoding method for JPEG 2000 using bit-rate-driven deep CNNs.

## II. The Soft Decoding Framework for JPEG 2000 Compressed Images

In this section, we introduce the deep CNNs-based soft decoding framework. The network architecture is presented firstly, and then the framework of the bit-rate-driven deep CNNs-based soft decoding method is introduced. Finally, some implementation details are given.

### A. Overview of the Architecture of the Deep CNN

Fig. 1 shows the network architecture of the deep CNN. Following the networks proposed in [7] and [11], we place the BN and ReLU between two adjacent convolutional layers, and the residual learning and gradient clipping are integrated into our network. With these techniques, a deep CNN can be well trained, thus achieving good restoration performance.

### B. Overview of the Soft Decoding Framework

The frameworks of the training and testing stages of our soft decoding method are presented in Fig. 2. In training phase, the training images are compressed by JPEG 2000 at different bit-rates $b_i, i = 1, \cdots, K$. At at each bit-rate $b_i$, we train the deep CNN shown in Fig. 1 using the training images and the images compressed at $b_i$, and the learned model is denoted by $CM_{b_i}$. After training, a series of models $CM_{b_i}, i = 1, \cdots, K$ can be learned.

The testing phase is consistent with the training process. Specifically, for an input image compressed by JPEG 2000 at bit-rate $b_i$, the learned CNN model $CM_{b_i}$ is selected to perform soft decoding.

### C. Implementation Details

In our implementation, the network depth is set to 25. The filter size is set to $3 \times 3$ for all of the layers, and 64 filters are used except for the last layer. To get multiple bit-rate-driven deep CNNs, we train a deep CNN model at every 0.1 bpp. The database BSDS500[1] is used to train our soft decoding networks.

[1][Online]. Available: https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/resources.html.

Fig. 3. The six test images used in our experiments: *Butterfly*, *Bike*, *Woman*, *Leaves*, *Foreman*, *Parrot*.

## III. Experimental Results

Experimental results are presented to verify the effectiveness of our soft decoding method in this section. The experiments are performed on six widely used test images shown in Fig. 3. For comparison, two popular compression standards, i.e., JPEG and JPEG 2000, are used as baselines. It is worth noting that we only show the results on gray images for fair and concise comparison, but this framework can be easily extended to color images.

### A. Objective Quality Comparison

In this paper, we use PSNR and SSIM to evaluate the objective quality of resultant images quantitatively, and Fig. 4 and Fig. 5 show the scores of the compared methods. It can be seen from Fig. 4 and Fig. 5 that the presented soft decoding method consistently achieves the best PSNR and SSIM scores on all test images, with a wide range of coding bit-rate. Taking the test image *Butterfly* as an example, the highest PSNR gain of our method over JPEG 2000 is up to 2.5 dB and the average gain is about 2.3 dB [see Fig. 4(a)]. From another perspective, generating the same PSNR score with JPEG 2000, the soft decoding framework achieves more than 20% coding bits saving. For the SSIM [see Fig. 5(a)], which is believed to be more consistent with visual quality, the soft decoding method also produces obvious gains over JPEG 2000. Similar conclusion can be obtained from the results of other test images. In sum, the objective quality of JPEG 2000 compressed images can be greatly enhanced by the presented soft decoding method.

### B. Visual Quality Comparison

To comprehensively compare the visual quality of results, we present the resultant images at different coding bit-rates.

From the results presented in Fig. 6 to Fig. 9, we can see that the resultant images of JPEG and JPEG 2000 both contain obvious coding artifacts, even the coding bit-rate is up to 1.0 bpp [see Fig. 9(b) and Fig. 9(c)]. At low bit-rates, more artifacts are produced (see Fig. 6 and Fig. 7). With the presented soft decoding method, the artifacts in JPEG 2000 compressed images can be well removed [see Fig. 6(d) and Fig. 7(d)]. In addition, the presented soft decoding method preserve small structures and recover sharp edges well. Overall, the resultant images of our soft decoding method are more perceptually pleasant than JPEG 2000.

## IV. Conclusion

A soft decoding framework for JPEG 2000 coded images is presented in this paper. Experimental results show that the presented framework greatly enhances the subjective and objective qualities of compressed images generated by JPEG 2000 at low to medium bit-rates. In future, the applications of deep CNN in other image restoration problems will be studied.

## References


[1] X. Liu, X. Wu, J. Zhou, and D. Zhao. “Data-driven sparsity-based restoration of JPEG-compressed images in dual transform-pixel domain,” in International Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 5171-5178.

[2] J. Zhang, et al. “CONCOLOR: constrained non-convex low-rank model for image deblocking,” IEEE Transactions on Image Processing, vol. 25, no. 3, pp. 1246-1259, 2016.

[3] X. Zhang, et al. “‘Low-rank decomposition-based restoration of compressed images via adaptive noise estimation,” IEEE Transactions on Image Processing, vol. 25, no. 9, pp. 4158-4171, 2016.

[4] Y. Dar, A. M. Bruckstein, M. Elad, and R. Giryes. “Postprocessing of compressed images via sequential denoising,” IEEE Transactions on Image Processing, vol. 25, no. 7, pp. 3044-3058, 2016.

[5] X. Liu, G. Cheung, X. Wu, and D. Zhao. “Random walk graph laplacian-based smoothness prior for soft decoding of JPEG images,” IEEE Transactions on Image Processing, vol. 26, no. 2, pp. 509-524, 2017.

[6] C. Dong, Y. Deng, C. C. Loy, and X. Tang. “Compression artifacts reduction by a deep convolutional network,” in International Conference on Computer Vision (ICCV), 2015, pp. 576-584.

[7] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang. “Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising,” IEEE Transactions on Image Processing, 2017.

[8] K. Yu, C. Dong, C. Loy, and X. Tang. “Deep convolution networks for compression artifacts reduction,” arXiv preprint arXiv:1608.02778, 2016.

[9] Y. Kwon, K. I. Kim, J. Tompkin, J. H. Kim, and C. Theobalt. “Efficient learning of image super-resolution and compression artifact removal with semi-local Gaussian processes,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 37, no. 9, pp. 1792-1805, 2015.

[10] R. Rothe, R. Timofte, and L. Van Gool. “Efficient regression priors for reducing image compression artifacts,” in International Conference on Image Processing (ICIP), 2015, pp. 1543-1547.

[11] J. Kim, J. K. Lee, and K. M. Lee. “Accurate image super-resolution using very deep convolutional networks,” in International Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 1646-1654.

[12] J. Zhang, J. Pan, W. Lai, R. Lau, and M.-H. Yang. “Learning fully convolutional networks for iterative non-blind deconvolution,” arXiv preprint arXiv:1611.06495, 2016.

[13] J. Ball, V. Laparra, and E. P. Simoncelli. “End-to-end optimized image compression,” arXiv preprint arXiv:1611.01704, 2016.

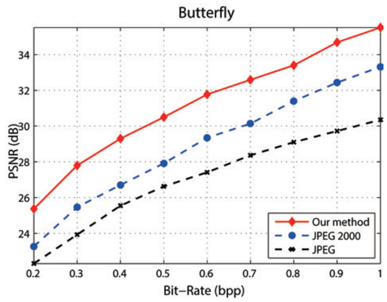
[14] K. Zhang, W. Zuo, S. Gu, and L. Zhang. “Learning deep CNN denoiser prior for image restoration,” arXiv preprint arXiv:1704.03264, 2017.

[15] S. Ioffe and C. Szegedy. “Batch normalization: Accelerating deep network training by reducing internal covariate shift,” in International Conference on Machine Learning (ICML), 2015, pp. 448-456.

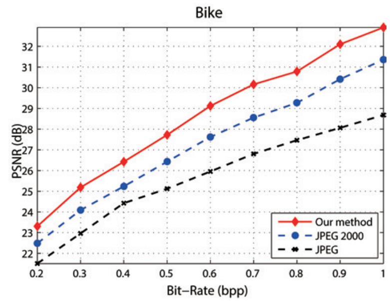
[16] V. Nair and G. Hinton. “Rectified linear units improve restricted boltzmann machines,” in International Conference on Machine Learning (ICML), 2010.

[17] R. Pascanu, T. Mikolov, and Y. Bengio. “On the difficulty of training recurrent neural networks,” arXiv preprint arXiv:1211.5063, 2012.

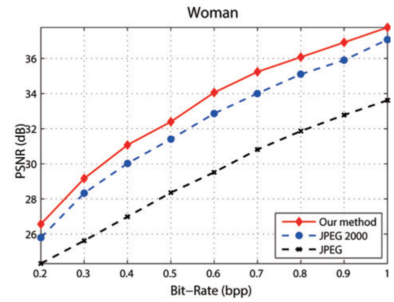
[18] K. He, X. Zhang, S. Ren, and J. Sun. “Deep residual learning for image recognition,” arXiv preprint arXiv:1512.03385, 2015.



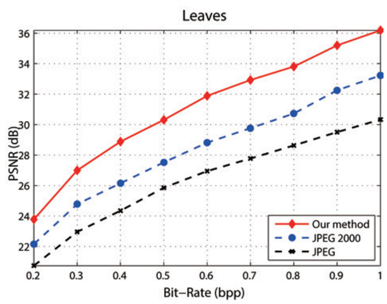

(a) *Butterfly*



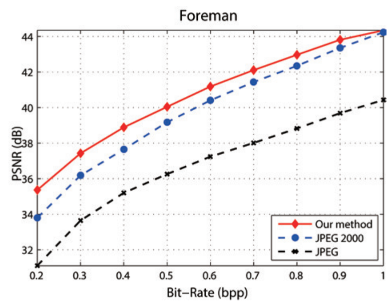

(b) *Bike*



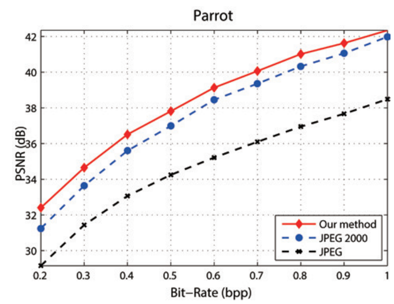

(c) *Woman*




(d) *Leaves*

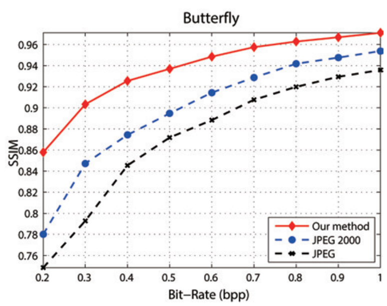



(e) *Foreman*

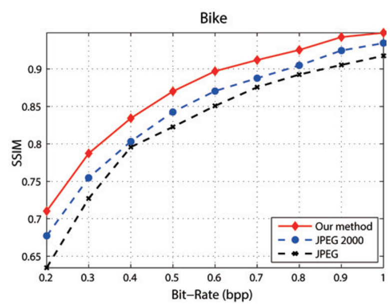



(f) *Parrot*

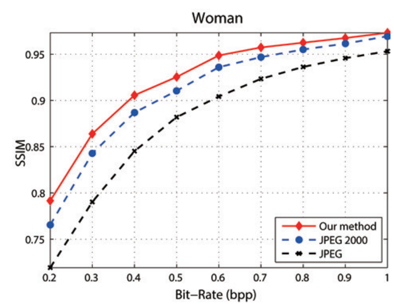
Fig. 4. Comparison of PSNR values.



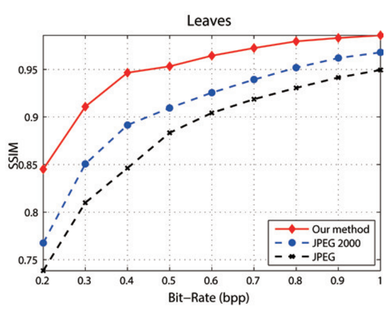

(a) *Butterfly*



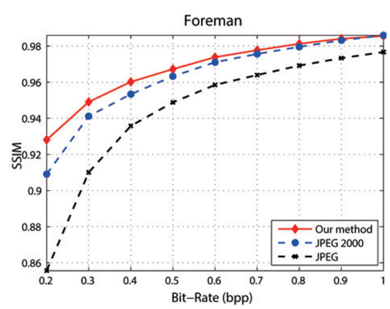

(b) *Bike*



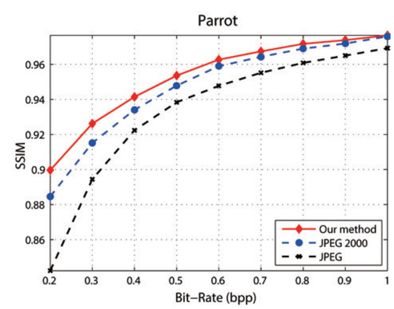

(c) *Woman*




(d) *Leaves*




(e) *Foreman*




(f) *Parrot*

Fig. 5. Comparison of SSIM values.

(a) Original (b) JPEG (c) JPEG 2000 (d) Our method

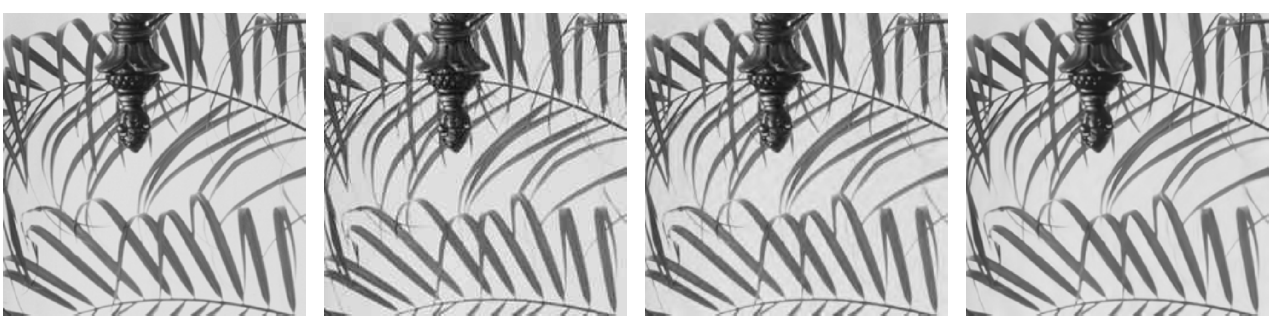
Fig. 6. Perceptual quality comparison on *Foreman* at 0.2 bpp.

(a) Original (b) JPEG (c) JPEG 2000 (d) Our method

Fig. 7. Perceptual quality comparison on *Bike* at 0.4 bpp.

(a) Original (b) JPEG (c) JPEG 2000 (d) Our method

Fig. 8. Perceptual quality comparison on *Butterfly* at 0.7 bpp.



(a) Original (b) JPEG (c) JPEG 2000 (d) Our method

Fig. 9. Perceptual quality comparison on *Leaves* at 1.0 bpp.