Complementary Materials for "Deep Optimized Multiple Description Image Coding via Scalar Quantization Learning"

Lijun Zhao, Huihui Bai, *Member, IEEE,* Anhong Wang, *Member, IEEE,* and Yao Zhao, *Senior Member, IEEE*

I. Experimental results and analysis

We provide more ablation studies on whether two importance-indicator maps are essential or not, as well as how to control the redundancy between different descriptions below. Meanwhile, two image colorization approaches are leveraged to further validate that the colorized appearances of the images compressed by our method are more like those of the uncompressed images than those of compressed images with the comparative MDC methods. These two approaches are "Colorful Image Colorization" [1] and "Let there be Color!" [2]. Among them, the first method infers color images from the predicted color distributions, which are useful for object classification, detection, segmentation, etc. The second approach uses a classification network as a part of the whole model to explore the global image features, whose output features are merged with those from middle-level feature networks.

A. Ablation studies

To clearly observe whether two importance-indicator maps are essential or not, we develop a variant version of "Ours-mr", denoted as "Ours-mr(OneID)", which only uses one importance-indicator map for image coding as like [28, 33]. Thus, a pair of quantizers in "Ours-mr(OneID)" is learned to quantize the feature tensor to generate multiple descriptions. An experimental comparison between "Ours-mr(OneID)" and "Ours-mr" is provided in Fig. 1-4 when testing on four available testing datasets, in which "Ours-mr(OneID)[S]" and "Ours-mr(OneID)[C]" are the objective qualities of side decoded images and central decoded images of "Ours-mr(OneID)" respectively. Other methods can be denoted in this way for the objective comparison of side and central decoded images. From these figures, it can be observed that the coding efficiency of "Ours-mr(OneID)" using only one importance-indicator map is lower than that of "Ours-mr" with a pair of importance-indicator maps(M) in terms of average MR-SSIM and MS-SSIM of the side and the central decoded images. Additionally, the difference between the side objective quality curve of "Ours-mr(OneID)[S]" and the central objective quality curve of "Ours-mr(OneID)[C]" in these figures shows that the redundancies of the two side decoded images for "Ours-mr(OneID)" and "Ours-mr" are different.

The redundancy of the proposed framework for diversified multiple description generation can be controlled by the parameter ψ, which is a trade-off parameter to control the redundancy between average side reconstructions and the corresponding central reconstructions. When the parameter ψ is set as 10^0, 10^1, and 10^2, the trained models can be respectively named "Ours-mr", "Ours-mr(1)" and "Ours-mr(2)". After comparing the performances of these trained models in Fig. 1-4, it can be observed that the growth of bpp vs. objective quality curves for "Ours-mr(2)[C]" and "Ours-mr(1)[C]" is more rapid than that of "Ours-mr[C]", but the growth of bpp vs. objective quality curves for "Ours-mr(2)[S]" and "Ours-mr(1)[S]" is slower than that of "Ours-mr[S]" when the bit-per-pixel values for multiple description coding are below 0.2 bpp. In this case, the two side reconstruction images become more dissimilar as the parameter ψ is larger. In other words, the multiple description redundancy between average side reconstruction images and the corresponding central reconstruction images will decrease when the parameter ψ is set to be larger. Additionally, the performances of "Ours-mr", "Ours-mr(1)" and "Ours-mr(2)" with above 0.2 bpp do not follow the same rules of curve change as the case below 0.2 bpp, as shown in Fig. 1-4. Since all the losses in the proposed framework use the class of pixel-level loss to regularize the multiple description encoder and decoder, this will lead to less high-level discrepancy such as semantic information to be controlled for diversified multiple description generation. Since this paper focuses on pixel-level loss for multiple description coding, controllable high-level discrepancy loss, which is a promising study topic, is not explored here and will be studied in our future work.

B. Visual quality comparisons of different methods

To further verify the effectiveness of the proposed method, the side decoded images and the center decoded images

Corresponding author: Lijun Zhao and Huihui Bai. This work was supported in part by Doctoral Scientific Research Starting Foundation of Taiyuan University of Science and Technology (No. 20192023), Funding Awards for Outstanding Doctors Volunteering to Work in Shanxi Province (No. 20192055), National Natural Science Foundation of China (No. 61672373) and Key Innovation Team of Shanxi 1331 Project (KITSX1331).

L. Zhao and A. Wang is with the Beijing Key Laboratory of Advanced Information Science and Network Technology, Institute of Information Science, Beijing Jiaotong University, Beijing, 100044, P. R. China, e-mail: 15112084, hhbai, yzhao@bjtu.edu.cn.

H. Bai and Y. Zhao are with the Beijing Key Laboratory of Advanced Information Science and Network Technology, Institute of Information Science, Beijing Jiaotong University, Beijing, 100044, P. R. China, e-mail: 15112084, hhbai, yzhao@bjtu.edu.cn.

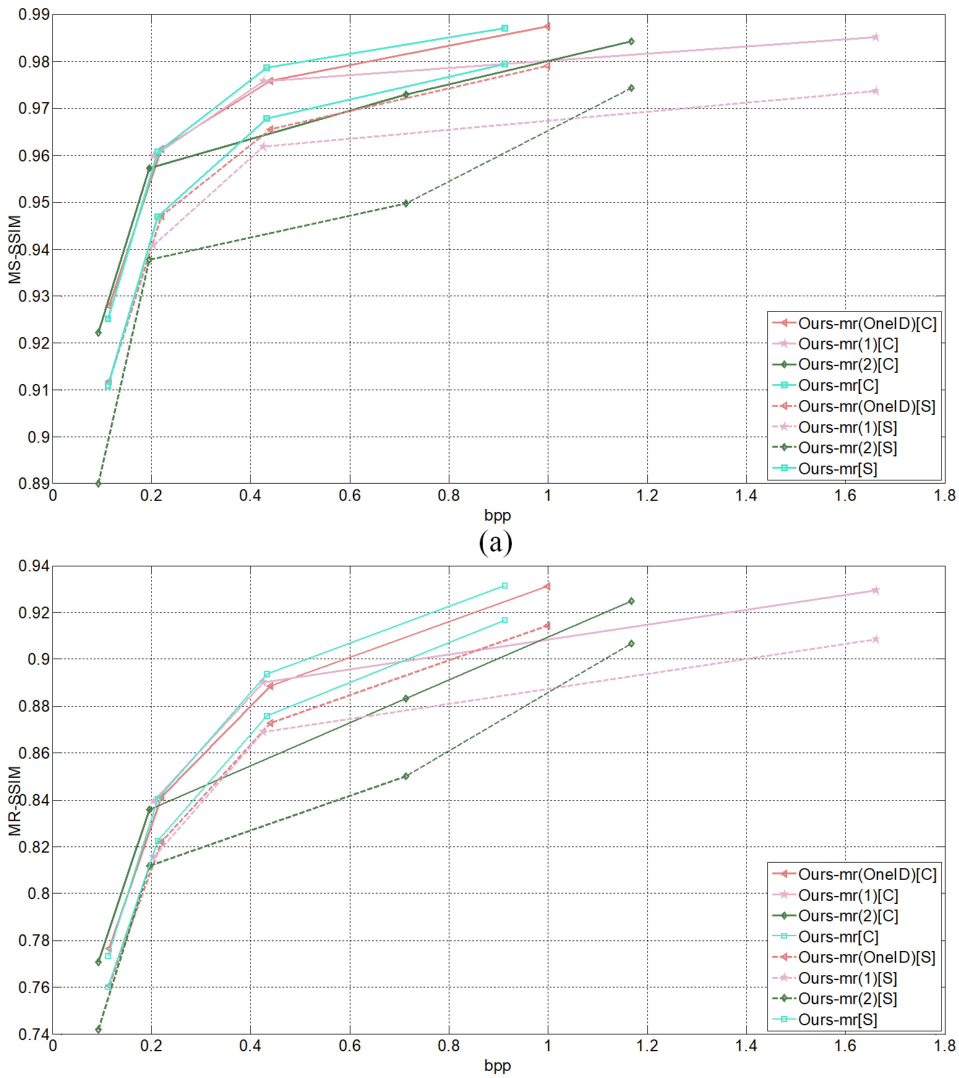


Fig. 1. The average objective quality comparisons between the proposed method and its variants testing on Set4.

compressed by the MDC methods are colorized by two colorization methods - ”Colorful Image Colorization” [1] and ”Let there be Color!” [2] - to obtain corresponding color side decoded image and color center decoded image, as shown in Fig. 5 and Fig. 6. From these figures, we can observe that the colorized appearances of the images compressed by the proposed method are more similar to those of the uncompressed gray image than those of the images compressed by ”MDCNN” [3] and ”MDROQ” [4], regardless of which method is chosen, ”Colorful Image Colorization” [1] or ”Let there be Color!” [2]. The images colorized with ”Colorful Image Colorization” for ”MDROQ” [4] are superior to those for ”MDCNN” [3]. However, they still cannot compete with ”Ours-mr” as well as ”Ours-mr(share)” and ”Ours-ms”. In addition, ”Ours-mr”, ”Ours-mr(share)” and ”Ours-ms” exhibit very similar colorized appearances with respect to the decoded side and central images.

References

[1] R. Zhang, P. Isola, and A. Efros, “Colorful image colorization,” in *European Conference on Computer Vision*, Netherlands, Oct. 2016.

[2] S. Iizuka, E. Simoserra, and H. Ishikawa, “Let there be color!: joint end-to-end learning of global and local image priors for automatic image colorization with simultaneous classification,” *ACM Transactions on Graphics*, vol. 35, no. 4, pp. 1–11, 2016.

[3] L. Zhao, H. Bai, A. Wang, and Y. Zhao, “Multiple description convolutional neural networks for image compression,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 8, pp. 2494–2508, 2019.

[4] L. Meng, J. Liang, U. Samarawickrama, Y. Zhao, H. Bai, and A. Kaup, “Multiple description coding with randomly and uniformly offset quantizers,” *IEEE Transactions on Image Processing*, vol. 23, no. 2, pp. 582–95, 2014.

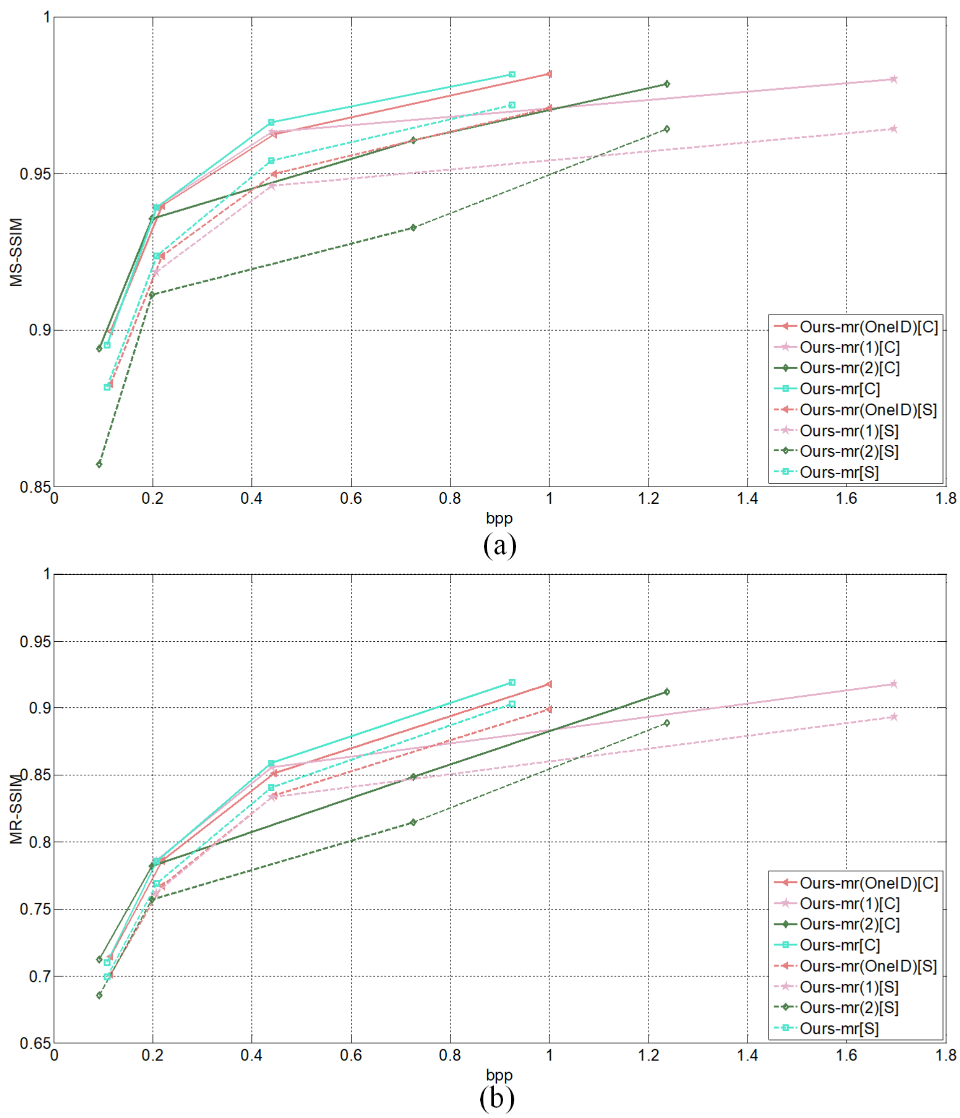


Fig. 2. The average objective quality comparisons between the proposed method and its variants testing on Kodak.

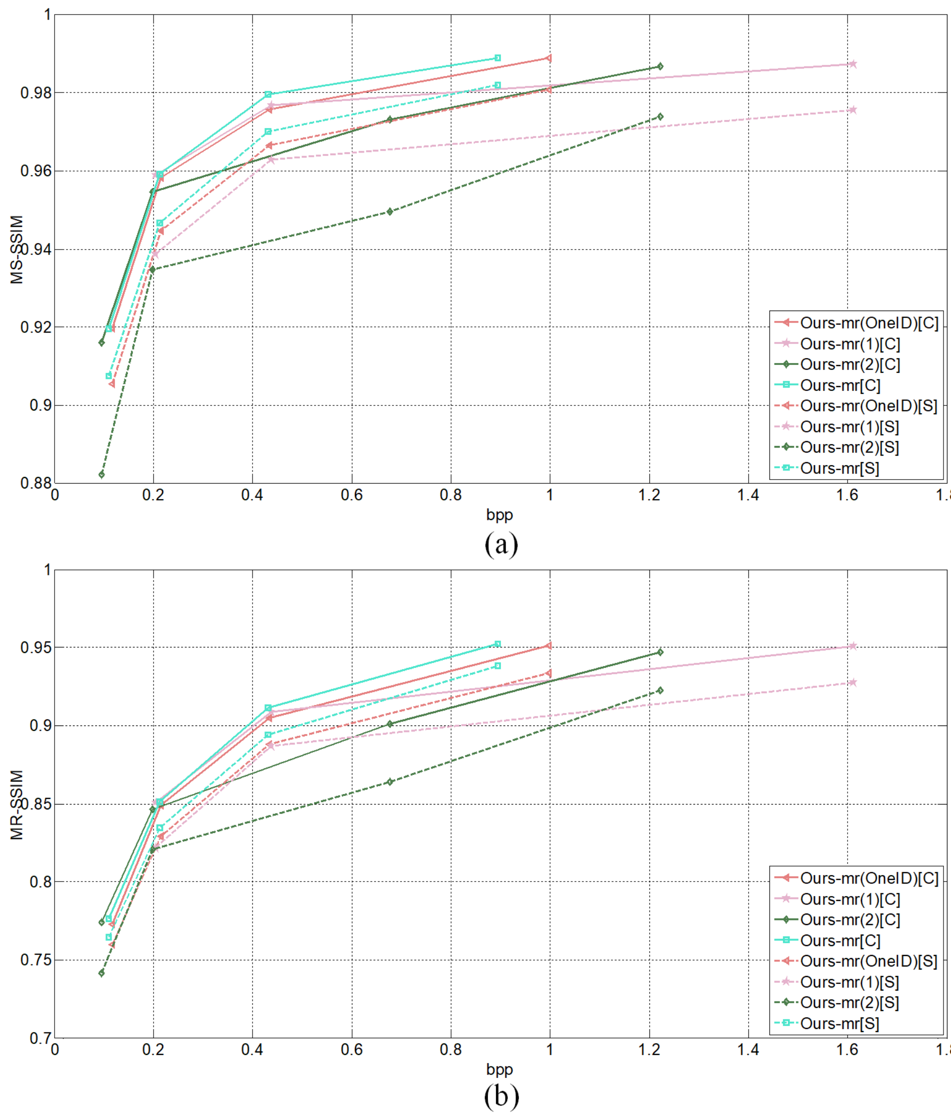
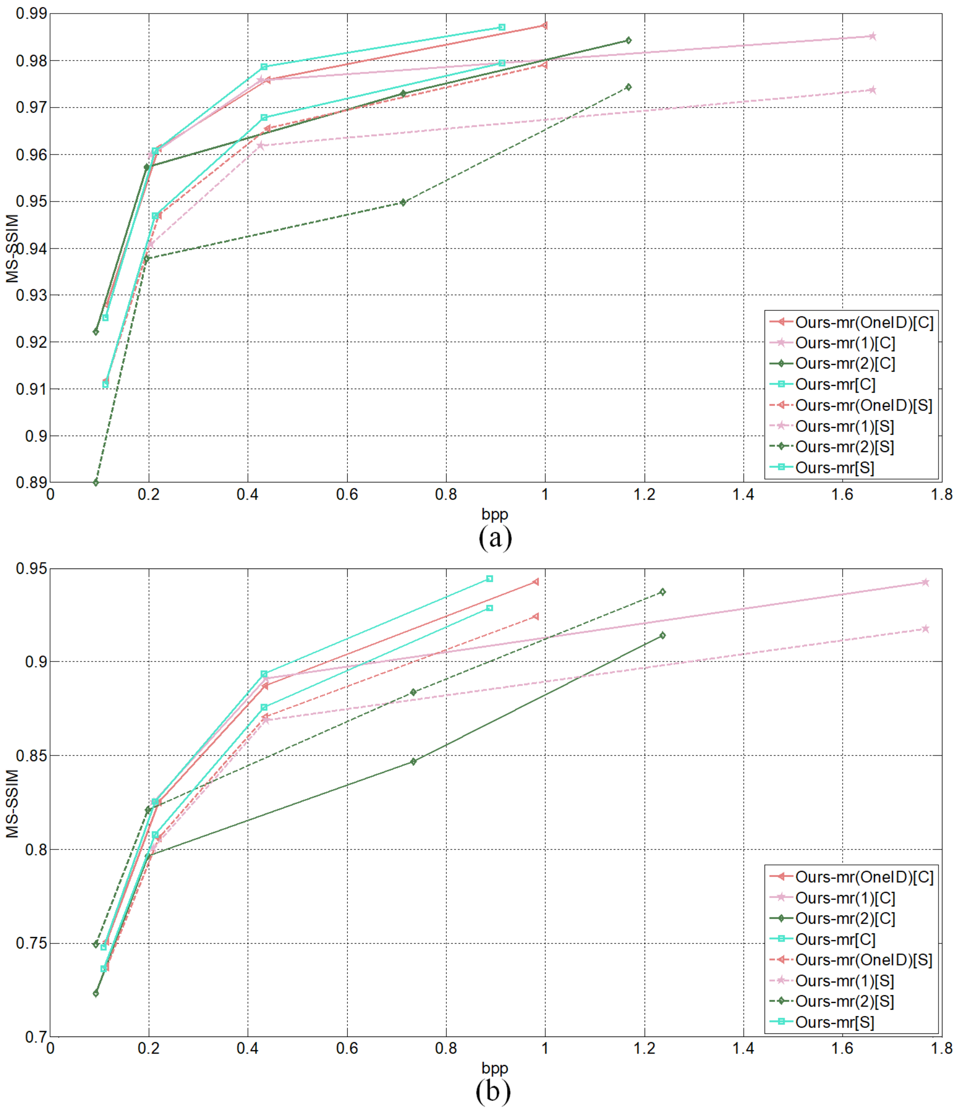


Fig. 3. The average objective quality comparisons between the proposed method and its variants testing on McMaster.



(a) (b)

Fig. 4. The average objective quality comparisons between the proposed method and its variants testing on SunHays.

Fig. 5. The visual quality comparisons (a1-a5) of different multiple description coding methods with MDCNN (0.18bpp), MDROQ (0.12bpp), Ours-ms (0.12bpp), Ours-mr (0.12bpp), Ours-mr(share) (0.12bpp) and the colorized image comparisons for the decoded central and side images, (a6) is an image from SunHays, (b1-b6) are colorized by "Colorful Image Colorization" [1], (c1-c6) are colorized by "Let there be Color!" [2]

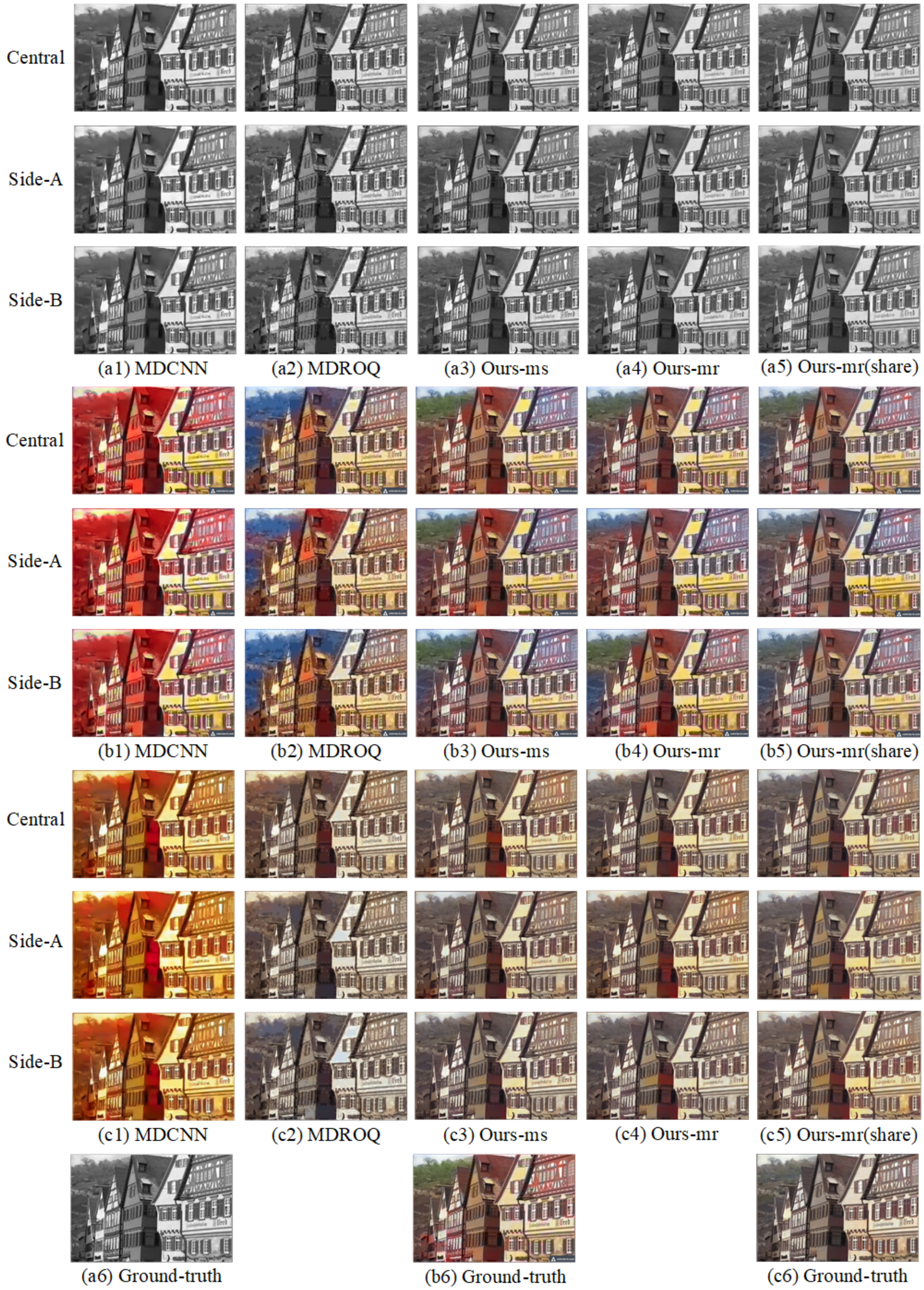


Fig. 6. The visual quality comparisons (a1-a5) of different multiple description coding methods with MDCNN (0.25bpp), MDROQ (0.29bpp), Ours-ms (0.29bpp), Ours-mr (0.25bpp), Ours-mr(share) (0.25bpp) and the colorized image comparisons for the decoded central and side images, (a6) is an image from Kodak, (b1-b6) are colorized by "Colorful Image Colorization" [1], (c1-c6) are colorized by "Let there be Color!" [2]