Just Noticeable Distortion Profile Inference: A Patch-level Structural Visibility Learning Approach

Xuelin Shen, *Graduate Student Member, IEEE*,
Zhangkai Ni, *Graduate Student Member, IEEE*, Wenhan Yang, *Member, IEEE*,
Xinfeng Zhang, *Member, IEEE*, Shiqi Wang, *Member, IEEE*, and
Sam Kwong, *Fellow, IEEE*

In this document, we provide additional materials to supplement our main paper. Section I investigates the influence of the training patch size on performance. Section II studies the influence of overlapping training patches. Section III details the procedure to determine the contrast masking factor. Section IV validates the proposed method with various IQA methods.

I. Influence of the training patch size

To further investigate the effect of different training patch size on performance, we train the proposed model with 32×32, 64×64, and 128×128 patch size, and corresponding training performances are illustrated in Fig. 1, Fig. 2, and Fig. 3, respectively. Table I reports the average PSNR results of the proposed model by using various sizes of the training patch. It is easy to figure out that the patch size has a weak impact on the performance of the proposed model. Therefore, 64×64 is set as the default size for the training patch in our experiments.

TABLE I
Average PSNR results of the proposed model with different training patch size.

Patch size	32x32	64x64	128x128
1	28.75	28.84	28.46
2	31.51	31.76	31.36
3	32.46	32.97	32.71
4	29.31	29.69	29.89
5	31.68	32.20	31.71
6	32.59	33.11	32.81
7	31.30	31.50	31.48
8	32.55	32.87	32.84
9	31.91	32.67	32.92
10	30.65	30.90	30.84
11	30.96	31.61	31.09
12	33.04	33.21	32.79
13	33.16	33.65	33.54
14	29.97	30.40	30.44
15	29.71	29.75	30.01
16	33.29	33.88	33.69
17	31.09	31.16	30.85
18	31.83	32.36	32.42
19	30.73	31.07	31.00
Average	31.40	31.77	31.62

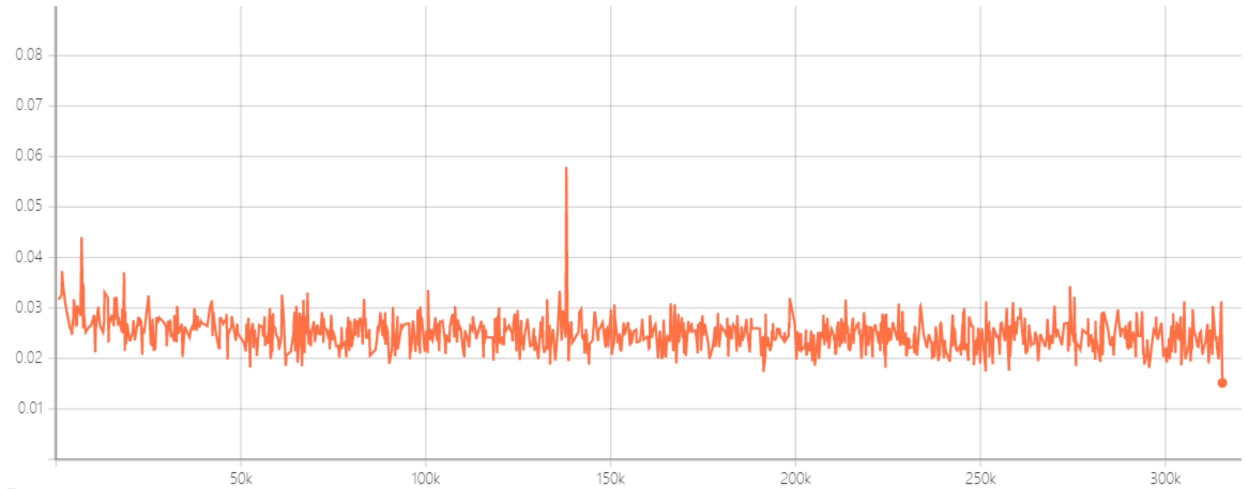

Fig. 1. Training loss of the model with 32x32 size.

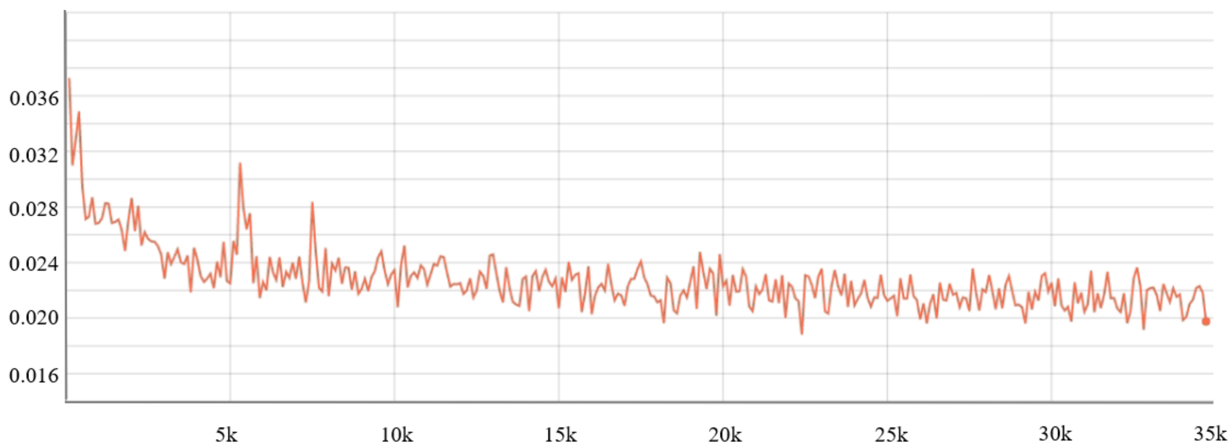

Fig. 2. Training loss of the model with 64x64 size.

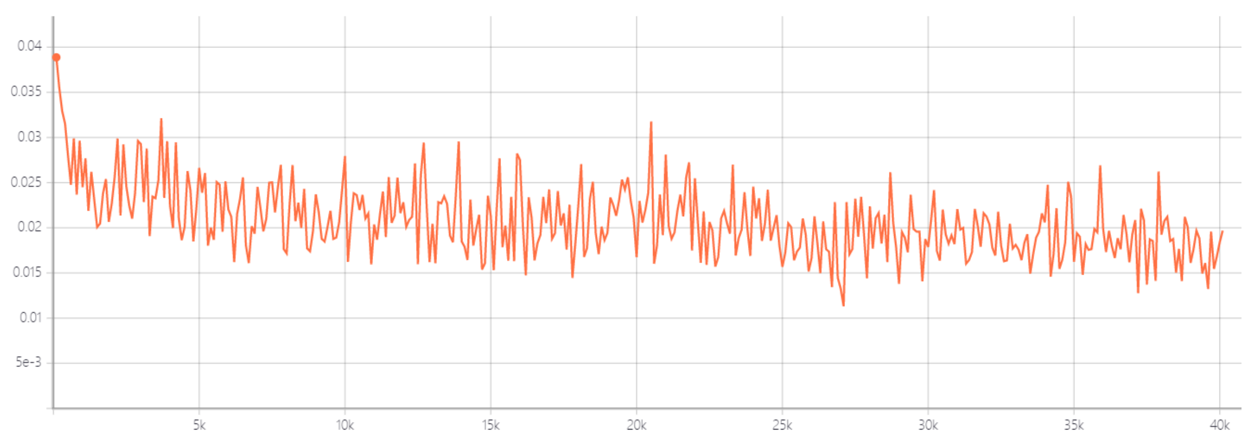

Fig. 3. Training loss of the model with 128x128 size.

II. Influence of overlapping training patch

Table II shows the performance comparison between the models using the overlapping training set and the non-overlapping training set in terms of PSNR. Fig. 4 illustrates the training performance of the model with the overlapping training set. Comparing with Fig. 2, one can see that using the overlapping training patch has a slight impact on the performance compared with the non-overlapping training set.

TABLE II
PERFORMANCE COMPARISON BETWEEN OVERLAPPED AND NON-OVERLAPPED TRAINING SET.

	64x64 (overlapping)	64x64 (non-overlapping)
1	28.81	28.84
2	31.65	31.76
3	32.79	32.97
4	29.64	29.69
5	31.96	32.20
6	32.92	33.11
7	31.43	31.50
8	32.79	32.87
9	32.56	32.67
10	30.86	30.90
11	31.25	31.61
12	33.02	33.21
13	33.48	33.65
14	30.35	30.40
15	29.93	29.75
16	33.55	33.88
17	31.20	31.16
18	32.28	32.36
19	30.99	31.07
Average	31.65	31.77

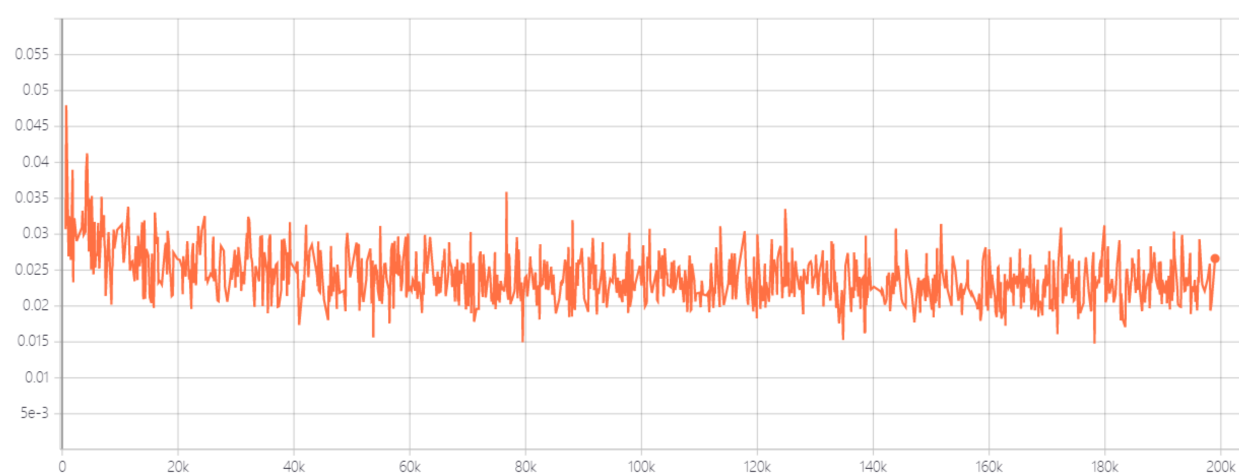

Fig. 4. Training loss of overlapped training dataset.

III. The contrast masking factor

We conduct subject testing to determine the contrast masking factor (f_C) in Section III of our main paper. More specifically, we randomly sample 10 images from the data set [1], crop and down-sample them to 1920×1080 resolution, and then convert to the luminance domain, as shown in Fig. 5. Subsequently, each image is reconstructed with various contrast by set f_C to 0.70, 0.75, 0.80, 0.85, 0.90 and 0.95, respectively.

Fig. 5. Sampled images for the subjective test.

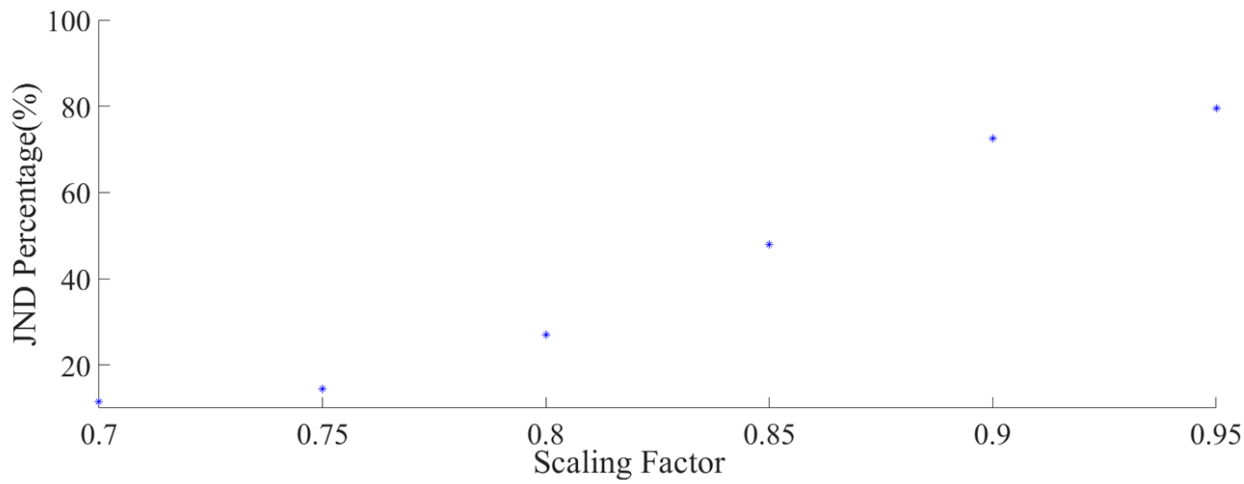

Fig. 6. The JND percentage of corresponding scaling factor.

A two-alternative-forced-choice (2AFC) subjective testing is conducted to determine the appropriate f_C factor. During the subjective test, the original image and corresponding reconstructed images are played side-by-side. The subjective testing environment is similar to subjective testing in Section IV of the paper. 20 subjects are invited and asked to determine whether there are differences between the two images. The results are shown in Fig. 6, where the JND percentage represents the percentage of "no difference between the two images". As such, 0.9 is selected as the corresponding visibility masking since 79.5% of the reconstructed images are perceptually lossless.

IV. Performances on various IQA methods

We employed several widely-used image quality assessment (IQA) methods to evaluate the performance of our proposed JND profile and employed anchors: Visual Information Fidelity (VIF) [2], DLM (Detail Loss Measure) [3], SSIM [4], and HDR-VDP2.2 [5]. Please note that the HDR-VDP2.2 has two outputs: a visibility map that computes the visual difference detection probability for each pixel, and a predicted mean-opinion-score which listed in Table III and Table IV. The results between the original images and distorted images from certain JND models are exhibited in Table III. To conduct a fair comparison, we also evaluate the JND models with the same objective quality in terms of PSNR. The methodology is the same as the subjective test in Subsection V-B of our paper, and results are shown in Table IV. As shown in Table IV, for most IQA methods, our method has advantages over other JND models in terms of VIF and VDP. The visibility maps of HDR-VDP 2.2 from different JND models (at the same PSNR level) are shown in Fig. 8. In general, the red means a higher detection probability while the blue indicates a lower detection probability, as shown in Fig. 7. From the probability detection map, one can obviously find that it is difficult to find the distortion of our model.

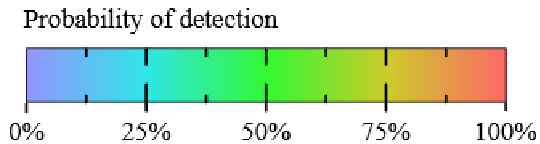

Fig. 7. Illustration of the probability detection in terms of color scales.

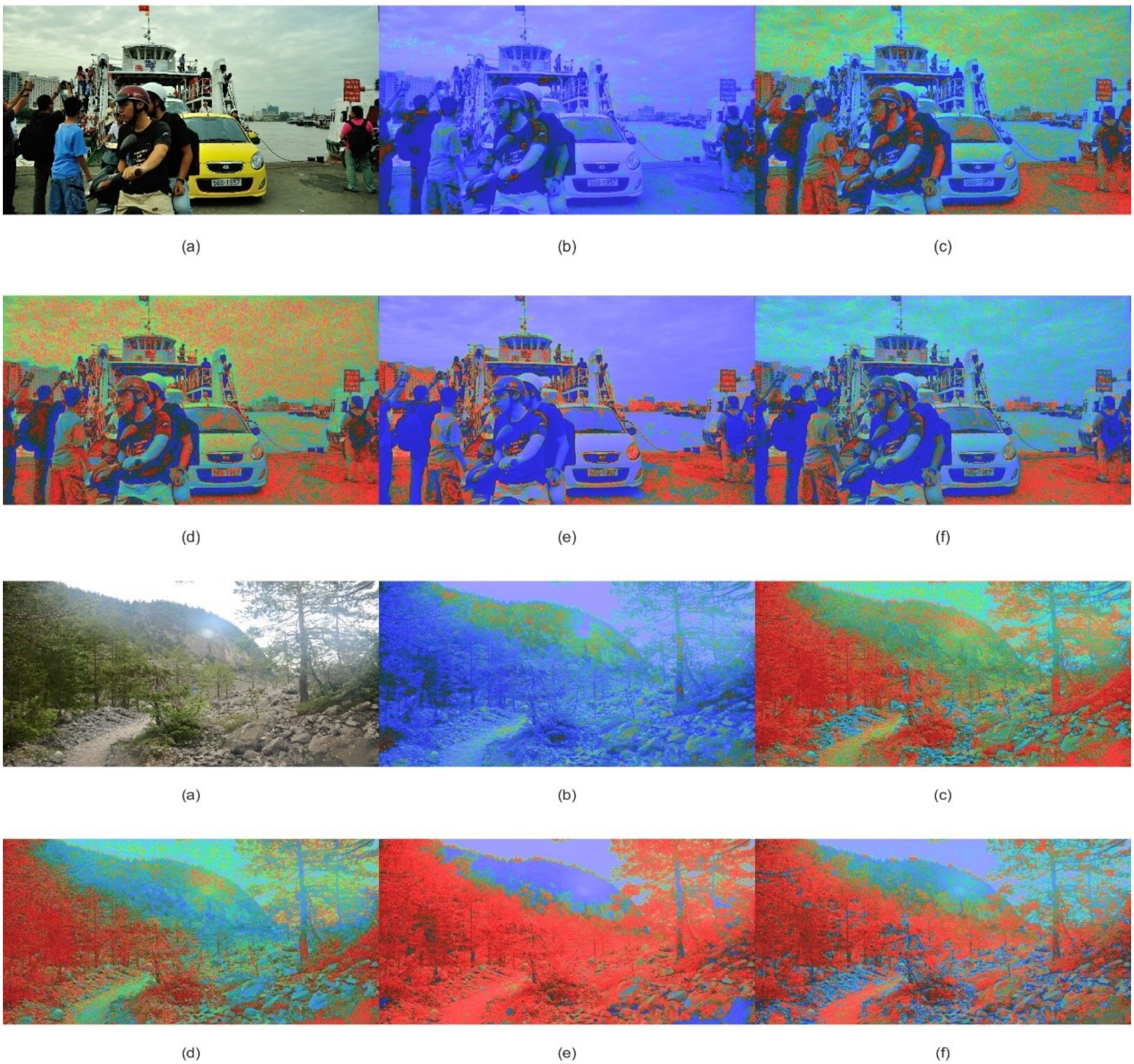

Fig. 8. Comparison of different JND model in terms of probability of detection from HDR-VDP-2. (a) Pristine image, (b) proposed method, (c)Yang *et al.* [6], (d)Liu *et al.* [7], (e) Wu *et al.* [8], (f) Wu *et al.* [9]

TABLE III
JND MODELS COMPARISON.

	Proposed				Liu *et al.* [7]				Yang *et al.* [6]				Wu *et al.* [8]				Wu *et al.* [9]			
Index	SSIM	VIF	DLM	VDP	SSIM	VIF	DLM	VDP	SSIM	VIF	DLM	VDP	SSIM	VIF	DLM	VDP	SSIM	VIF	DLM	VDP
1	0.65	0.54	0.92	51	0.44	0.35	0.85	45	0.44	0.36	0.86	46	0.98	0.93	0.99	60	0.67	0.46	0.92	47
2	0.87	0.81	0.95	56	0.72	0.57	0.92	49	0.75	0.63	0.93	50	0.96	0.74	0.98	55	0.78	0.58	0.95	48
3	0.96	0.81	0.96	60	0.77	0.62	0.94	50	0.81	0.65	0.95	51	0.94	0.66	0.97	54	0.78	0.51	0.94	47
4	0.82	0.59	0.93	54	0.60	0.47	0.90	46	0.61	0.48	0.90	46	0.99	0.95	0.99	67	0.74	0.56	0.94	48
5	0.95	0.88	0.97	60	0.87	0.63	0.95	51	0.89	0.68	0.96	52	0.87	0.57	0.95	51	0.76	0.47	0.92	45
6	0.95	0.87	0.97	59	0.89	0.61	0.95	53	0.91	0.70	0.97	55	0.91	0.64	0.96	62	0.81	0.56	0.94	48
7	0.81	0.54	0.93	52	0.66	0.61	0.90	47	0.69	0.65	0.91	48	0.99	0.95	0.99	63	0.79	0.71	0.95	51
8	0.95	0.77	0.95	54	0.79	0.65	0.95	49	0.83	0.69	0.96	49	0.94	0.73	0.98	56	0.77	0.54	0.95	48
9	0.95	0.71	0.94	57	0.74	0.62	0.94	49	0.77	0.65	0.94	49	0.98	0.88	0.99	60	0.81	0.64	0.96	50
10	0.82	0.61	0.93	50	0.63	0.54	0.90	47	0.65	0.57	0.90	47	0.98	0.88	0.98	58	0.75	0.62	0.94	48
11	0.95	0.90	0.96	56	0.88	0.58	0.96	53	0.90	0.67	0.97	54	0.89	0.56	0.96	53	0.81	0.52	0.95	48
12	0.81	0.66	0.92	50	0.60	0.60	0.89	48	0.63	0.64	0.89	48	0.98	0.89	0.99	60	0.73	0.63	0.94	48
13	0.96	0.80	0.95	58	0.82	0.66	0.93	50	0.85	0.70	0.95	51	0.94	0.68	0.97	54	0.81	0.54	0.94	47
14	0.88	0.51	0.93	57	0.62	0.46	0.89	46	0.63	0.47	0.89	46	0.99	0.97	0.99	68	0.76	0.57	0.93	48
15	0.89	0.79	0.87	44	0.66	0.51	0.93	45	0.68	0.55	0.93	46	0.95	0.73	0.96	51	0.75	0.54	0.94	45
16	0.94	0.83	0.94	60	0.85	0.71	0.97	54	0.88	0.76	0.97	54	0.95	0.80	0.98	57	0.82	0.64	0.97	50
17	0.74	0.69	0.93	50	0.51	0.55	0.89	48	0.52	0.57	0.90	48	0.98	0.87	0.99	58	0.69	0.61	0.94	48
18	0.93	0.72	0.94	56	0.74	0.64	0.93	50	0.76	0.65	0.94	49	0.98	0.92	0.99	64	0.78	0.62	0.96	52
19	0.86	0.65	0.93	51	0.65	0.57	0.91	46	0.68	0.61	0.92	46	0.98	0.86	0.99	56	0.74	0.61	0.95	47
Average	0.88	0.72	0.94	55	0.71	0.58	0.92	49	0.73	0.61	0.93	50	0.96	0.80	0.98	58	0.77	0.58	0.95	49

TABLE IV
JND MODELS COMPARISON WITH STANDARD OBJECTIVE QUALITY.

	Proposed				Liu *et al.* [7]				Yang *et al.* [6]				Wu *et al.* [8]				Wu *et al.* [9]			
Index	SSIM	VIF	DLM	VDP	SSIM	VIF	DLM	VDP	SSIM	VIF	DLM	VDP	SSIM	VIF	DLM	VDP	SSIM	VIF	DLM	VDP
1	0.65	0.54	0.92	51	0.53	0.41	0.88	48	0.53	0.42	0.89	48	0.87	0.67	0.85	38	0.53	037	0.88	44
2	0.87	0.81	0.95	56	0.81	0.67	0.95	54	0.80	0.70	0.95	53	0.93	0.65	0.97	50	0.81	0.62	0.96	50
3	0.96	0.81	0.96	60	0.80	0.65	0.94	51	0.81	0.65	0.95	51	0.94	0.66	0.97	54	0.87	0.62	0.96	52
4	0.82	0.59	0.93	54	0.64	0.51	0.91	47	0.65	0.51	0.92	47	0.86	0.67	0.89	41	0.63	0.47	0.91	45
5	0.95	0.88	0.97	60	0.91	0.70	0.96	53	0.89	0.68	0.96	52	0.87	0.57	0.95	51	0.91	0.69	0.97	54
6	0.95	0.87	0.97	59	0.94	0.72	0.87	57	0.92	0.73	0.97	56	0.91	0.64	0.96	53	0.92	0.72	0.98	55
7	0.81	0.54	0.93	52	0.73	0.67	0.92	50	0.71	0.68	0.92	49	0.95	0.81	0.89	40	0.68	0.61	0.93	46
8	0.95	0.77	0.95	54	0.79	0.65	0.95	50	0.80	0.66	0.95	48	0.92	0.68	0.97	53	0.87	0.66	0.97	53
9	0.95	0.71	0.94	57	0.80	0.69	0.96	52	0.79	0.68	0.95	50	0.94	0.75	0.97	50	0.81	0.64	0.96	46
10	0.82	0.61	0.93	50	0.69	0.61	0.92	49	0.71	0.64	0.93	50	0.92	0.68	0.91	43	0.69	0.56	0.93	46
11	0.95	0.90	0.96	56	0.93	0.69	0.97	57	0.91	0.70	0.97	55	0.89	0.56	0.96	52	0.91	0.68	0.98	54
12	0.81	0.66	0.92	50	0.75	0.76	0.93	54	0.73	0.75	0.92	52	0.95	0.75	0.94	48	0.73	0.63	0.93	49
13	0.96	0.80	0.95	58	0.85	0.69	0.94	51	0.85	0.70	0.95	51	0.94	0.68	0.97	54	0.91	0.70	0.97	53
14	0.88	0.51	0.93	57	0.66	0.49	0.90	47	0.67	0.50	0.91	47	0.86	0.68	0.76	35	0.64	0.46	0.89	43
15	0.89	0.79	0.87	44	0.73	0.58	0.95	48	0.71	0.58	0.95	47	0.90	0.63	0.94	46	0.78	0.57	0.95	46
16	0.94	0.83	0.94	60	0.87	0.73	0.97	55	0.85	0.73	0.97	53	0.94	0.75	0.98	54	0.89	0.74	0.98	55
17	0.74	0.69	0.93	50	0.67	0.69	0.94	54	0.63	0.66	0.93	52	0.93	0.68	0.93	44	0.66	0.58	0.94	47
18	0.93	0.72	0.94	56	0.77	0.67	0.94	51	0.79	0.68	0.94	50	0.91	0.72	0.95	47	0.78	0.62	0.96	52
19	0.86	0.65	0.93	51	0.72	0.64	0.93	49	0.71	0.64	0.93	48	0.98	0.65	0.94	43	0.68	0.54	0.93	45
Average	0.88	0.72	0.94	55	0.77	0.64	0.94	52	0.76	0.65	0.94	51	0.91	0.68	0.93	48	0.78	0.61	0.95	50

References

[1] D. Dang-Nguyen, C. Pasquini, V. Conotter, and G. Boato, “Raise: A raw images dataset for digital image forensics,” in *Proceedings of the 6th ACM Multimedia Systems Conference*, 2015, pp. 219–224.

[2] R. Sheikh and A. Bovik, “Image information and visual quality,” *IEEE Transactions on image processing*, vol. 15, no. 2, pp. 430–444, 2006.

[3] S. Li, F. Zhang, L. Ma, and K. Ngan, “Image quality assessment by separately evaluating detail losses and additive impairments,” *IEEE Transactions on Multimedia*, vol. 13, no. 5, pp. 935–949, 2011.

[4] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[5] R. Mantiuk, J. Kim, A. Rempel, and W. Heidrich, “Hdr-vdp-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions,” *ACM Transactions on graphics (TOG)*, vol. 30, no. 4, pp. 1–14, 2011.

[6] X. Yang, W. Lin, Z. Lu, E. Ong, and S. Yao, “Motion-compensated residue preprocessing in video coding based on just-noticeable-distortion profile,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 6, pp. 742–752, 2005.

[7] A. Liu, W. Lin, M. Paul, C. Deng, and F. Zhang, “Just noticeable difference for images with decomposition model for separating edge and textured regions,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 11, pp. 1648–1652, 2010.

[8] J. Wu, G. Shi, W. Lin, A. Liu, and F. Qi, “Just noticeable difference estimation for images with free-energy principle,” *IEEE Transactions on Multimedia*, vol. 15, no. 7, pp. 1705–1710, 2013.

[9] J. Wu, L. Li, W. Dong, G. Shi, W. Lin, and Kuo. C-C Jay, “Enhanced just noticeable difference model for images with pattern complexity,” *IEEE Transactions on Image Processing*, vol. 26, no. 6, pp. 2682–2693, 2017.