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

- Supplementary Material

Xuelin Shen, Zhangkai Ni, *Student Member*, *IEEE*, Wenhan Yang, *Member*, *IEEE*,

Xinfeng Zhang, *Member*, *IEEE*, Shiqi Wang, *Member*, *IEEE*, and Sam Kwong, *Fellow*, *IEEE*

This document is mainly about the supplementary experiments which are required in the response letter. Details and corresponding results are provided.

Experiment 1. We train and test our model with patches of sizes of 128x128 and 32x32 to show changing the patch size at a proper range can barely affect the performance of the JND model since we utilized patches of size 64x64 in our manuscript. The training performances and testing results in terms of PSNR (dB) are shown in Fig. 1, Fig. 2 and Table 1, respectively.

Table 1 Performance comparison with different 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

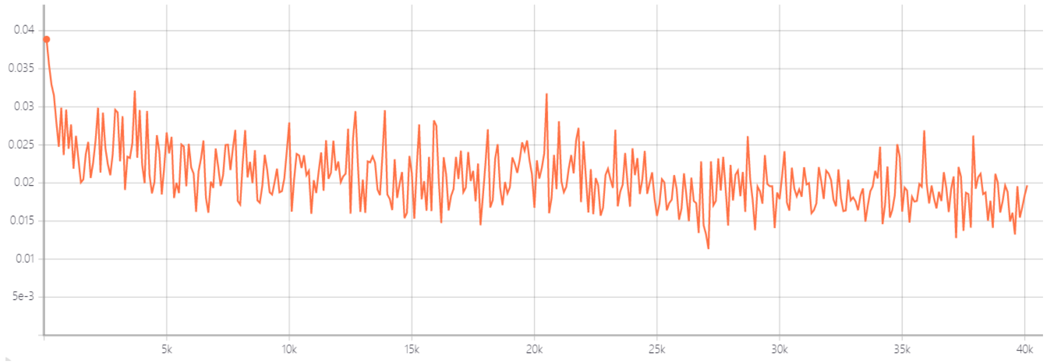


Fig. 1 Training loss of the model with 128x128 size.

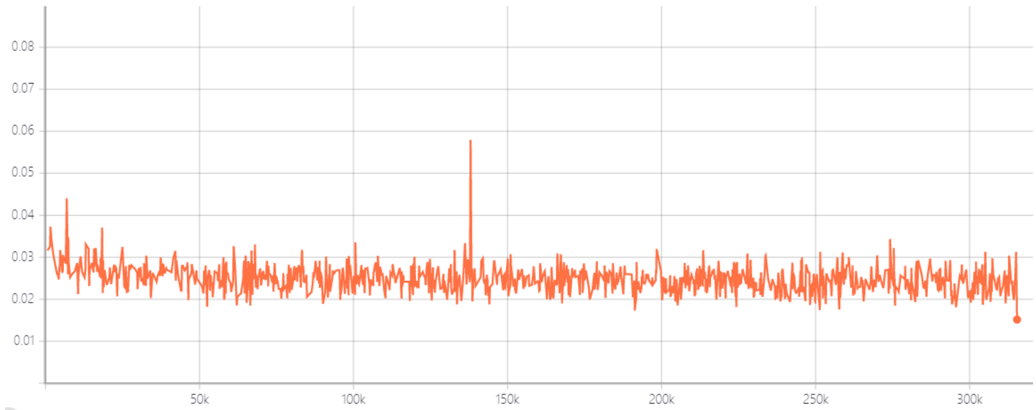


Fig. 2 Training loss of the model with 32x32 size.

Experiment 2. A comparison between the models using overlapping training set and non-overlapping training set is shown in Table 2, and the training performance of the model with overlapping training set is illustrated in Fig. 3. From the aspect of the model training, cropped overlapped patches for training can be seen as a kind of data augmentation which may improve the training performance under some circumstances. In terms of our work, we generated more than 70,000 patch pairs for model training, which is large enough to train the structural learning model. As such, For the training stage, it will make few differences to use the overlapping patch or not if the patch size has been large enough.

Table 2 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

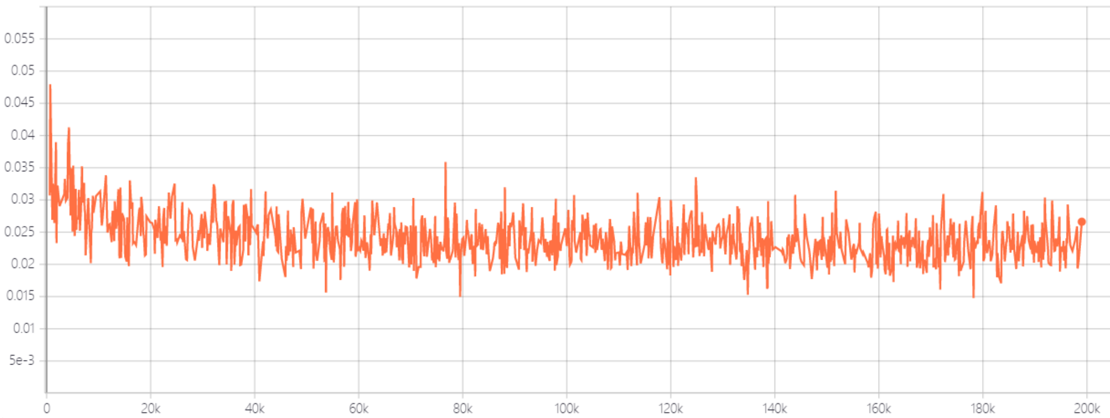


Fig. 3 Training loss of overlapped training dataset

Experiment 3. Here, a supplementary introduction about the corresponding subjective testing which decides the contrast masking factor (f_C) in Section III-C is provided. In this subjective testing, we sample 10 images from the dataset [35]. The selected images are firstly cropped and down-sampled to 1920×1080 resolution, and then converted to luminance domain, as shown in Fig. 4. Subsequently, each image is reconstructed by changing the contrast with f_C which equals to 0.70, 0.75, 0.80, 0.85, 0.90 and 0.95. A two-alternative-forced-choice (2AFC) subjective testing is conducted to generate the appropriate f_C factor. In this subjective testing, the original image and corresponding reconstructed images are played side-by-side. The subjective test environment is similar to the subjective testing in the Sec. IV of the manuscript.

Fig.4 Sampled images for the subjective test

20 subjects are invited and asked to determine whether there are differences between the two images. The results are shown in below Fig. 5, where the JND percentage denotes the percentage regarding the "there are 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 at this point.

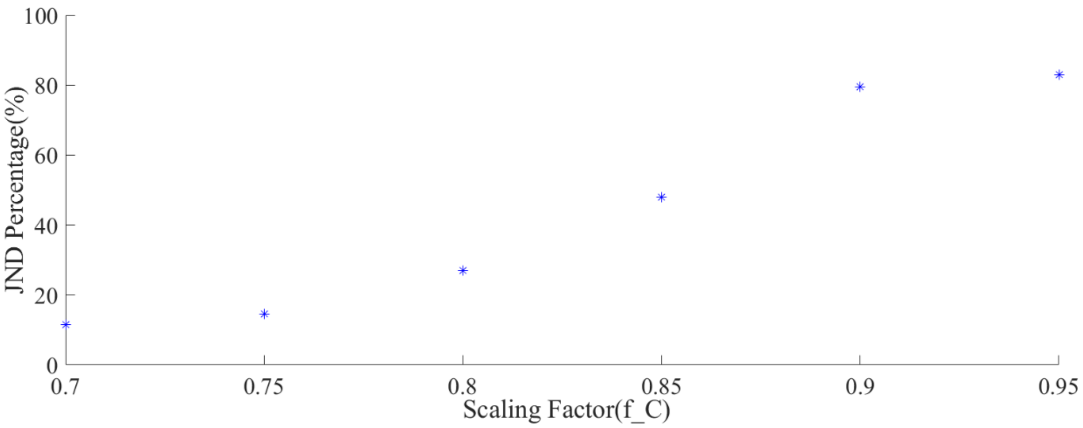


Fig. 5 The JND percentage of corresponding scaling factor.

Experiment 4. We employed several extra image quality assessment (IQA) methods to evaluate the performance of our proposed JND profile and employed anchors: Visual Information Fidelity (VIF) [52], DLM (Detail Loss Measure) [53], SSIM [33] and HDR-VDP2.2 [54]. Things have to be mentioned that, the HDR-VDP2.2 has two outputs: visibility map which computes the visual difference detection probability for each pixel and a predicted mean-opinion-score which will be listed in below tables. The results between the original images and distorted images from certain JND models are shown in Table 3.

As exhibited in the below Table 3, comparing with [4], [42] and [11], the proposed method has obvious advantages in terms of SSIM (ours: 0.88 *vs.* [4]: 0.71, [42]:0.73, [11]: 0.77), VIF (ours:0.72 *vs.* [4]: 0.58, [42]:0.61, [11]:0.58) and VDP mean-opinion-score (ours:55 *vs.* [4]: 49, [42]: 50, [11]: 49) while has a similar objective quality level in terms of PSNR (ours:31.77 *vs.* [4]: 30.78, [42]:30.02, [11]:30.87). Although the work in [10] seems to have better performance (SSIM:0.98, VIF:0.92, VDP mean-opinion-score:64), it fails to consider the distortion concealment ability (PSNR:38.12).

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 testing in Subsection V-B of our manuscript, and result is shown in Table 4. As is shown, our method has obvious advantages against the other JND models in terms of most IQA methods. The visibility maps from HDR-VDP2.2 from different JND models (at same PSNR level) are shown in Fig. 7. In general, the red color means a high detection probability while the blue for a low detection probability as exhibited in Fig. 6. From the probability detection map, it can be obviously seen that the distortions from our model are more difficult to be noticed by HVS.

Table 3 JND models comparison.

	Proposed				Liu *et al.* [4]				Yang *et al.* [42]				Wu *et al.* [10]				Wu *et al.* [11]			
	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
Ave.	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 4 JND models comparison with standard objective quality.

	Proposed				Liu *et al.* [4]				Yang *et al.* [42]				Wu *et al.* [10]				Wu *et al.* [11]			
	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
Ave.	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

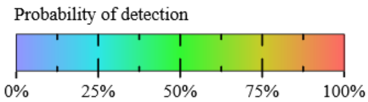
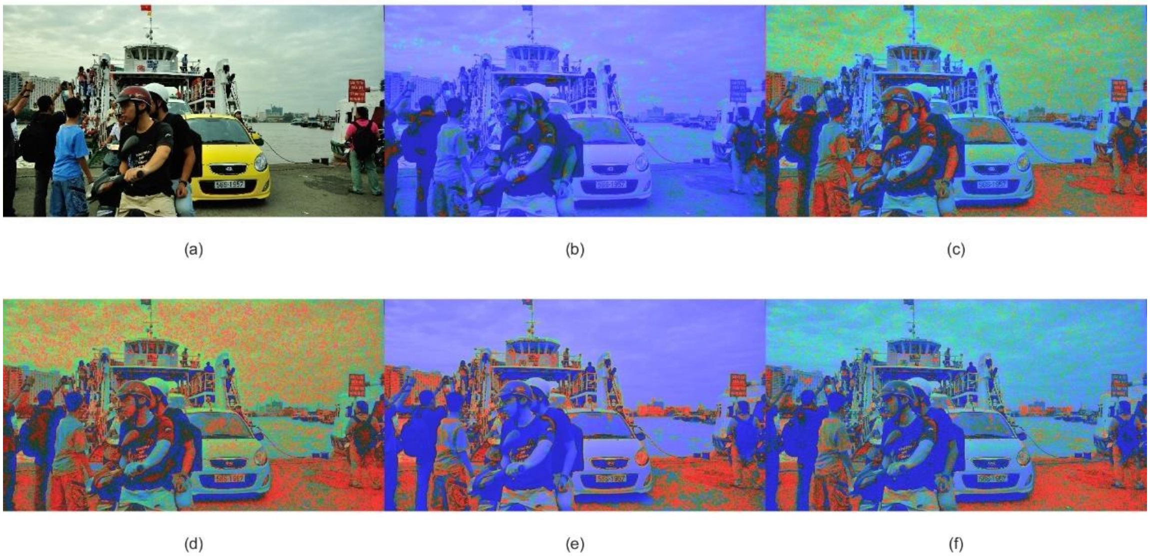
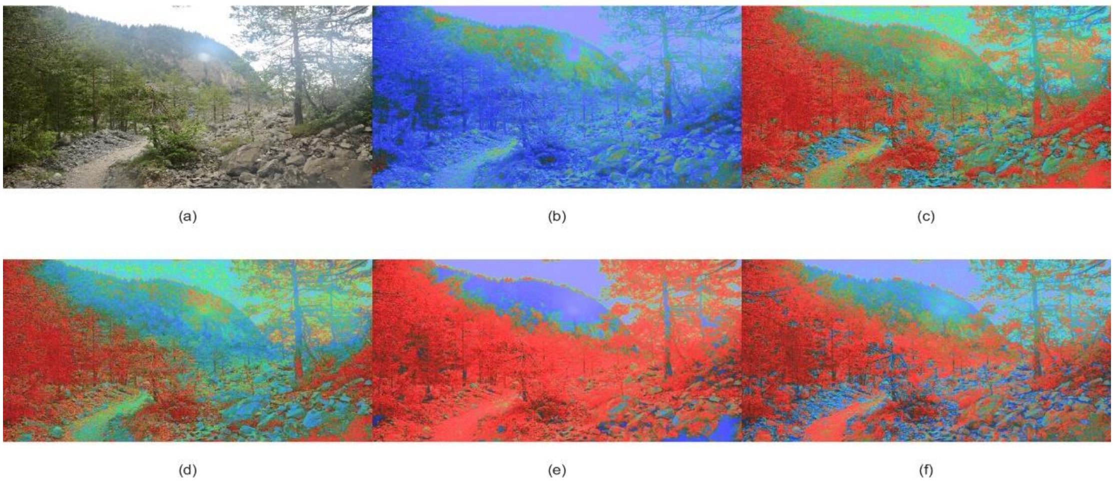


Fig.6 Illustration of the probability detection in terms of color scales [54].





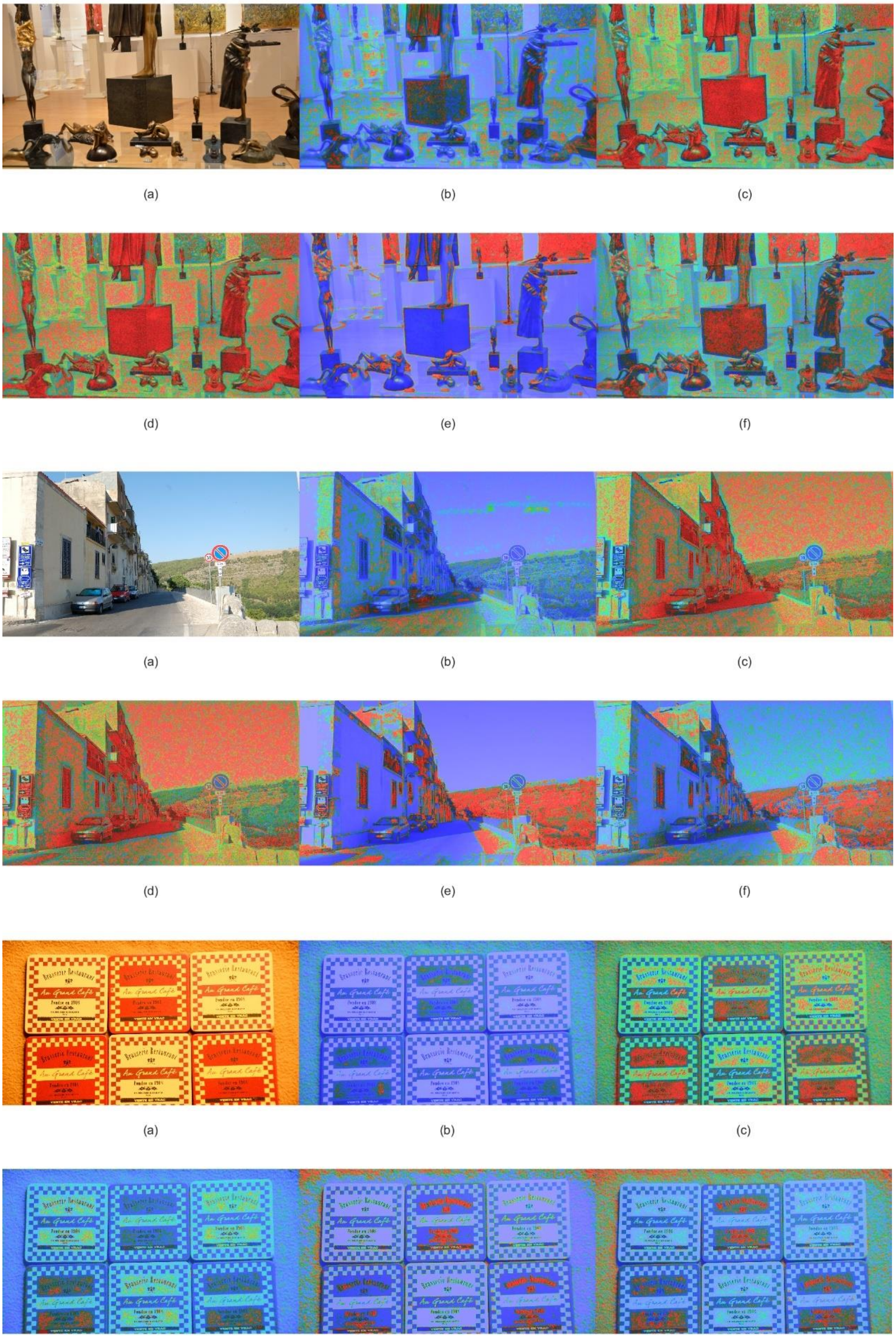


Fig.7 Comparison of different JND model in terms of probability of detection from HDR-VDP-2. (a) Pristine image, (b) proposed method, (c) Yang *et al.* [42], (d) Liu *et al.* [4], (e) Wu *et al*. [10], (f) Wu *et al.* [11].

Reference

[4] 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.

[10] 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.

[11] 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.

[33] 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.

[35] D. Tien, D. Nguyen, C. Valentina, P. Cecilia, and B. Giulia, "RAISE: A raw images dataset for digital image forensics," in *Proceedings of the 6th ACM Multimedia Systems Conference*, 2015, pp. 219–224.

[42] X. Yang, W. Lin, Z. Lu, E. On, 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.

[50] R. Rassool, "VMAF reproducibility: Validating a perceptual practical video quality metric," in *IEEE International Symposium on Broadband Multimedia Systems & Broadcasting*, 2017.

[51] S. Daly, "Visible differences predictor: an algorithm for the assessment of image fidelity," *Human Vision, Visual Processing, and Digital Display III*. Vol. 1666. International Society for Optics and Photonics, 1992.

[52] H. Sheikh and B. Alan, "Image information and visual quality," *IEEE Transactions on image processing*, 15.2 (2006): 430-444.

[53] S. Li, F. Zhang, L. Ma, and KN. Ngan, "Image quality assessment by separately evaluating detail losses and additive impairments," *IEEE Transactions on Multimedia*, 13.5 (2011): 935-949.

[54] R. Mantiuk, K. 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)*, 30.4 (2011): 1-14.