Interactive Multi-Dimension Modulation with Dynamic Controllable Residual Learning for Image Restoration Supplementary File

Jingwen He[⋆1,2], Chao Dong[⋆1,2], and Yu Qiao†[1,2]

[1] ShenZhen Key Lab of Computer Vision and Pattern Recognition, SIAT-SenseTime Joint Lab, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences
[2] SIAT Branch, Shenzhen Institute of Artificial Intelligence and Robotics for Society
{jw.he, chao.dong, yu.qiao}@siat.ac.cn

Abstract. In this supplementary file, we mainly present quantitative results on more test datasets. Note that the model is the same as the main paper without any retraining. First, we show the quantitative results of 2D modulation and 3D modulation on CBSD68 and LIVE1 datasets. Then we add some qualitative results of 2D and 3D modulation. Besides, we compare with SD methods on more degradation types and levels. For ablation studies, we provide more results about the effectiveness of global connection and local connections. Moreover, the results under different sampling strategies on CBSD68 dataset is provided.

1 More quantitative results of 2D and 3D modulation

In the main paper, we only present the results on the CBSD68 dataset due to the space limitation. Here, we show more testing results on the LIVE1 dataset. As shown in Table 1 and 2 the overall trend of PSNR distances on these two datasets are similar. Specifically, the largest distances appear in blur $r1$, which is the zero starting point and mild degradation. Results on two degradations are generally below 0.2 dB, indicating a high modulation accuracy.

2 More qualitative results of 2D and 3D modulation

In this section, we show more qualitative results of our 2D (Figure 3(a)) and 3D modulation (Figure 3(b)) in image restoration. Note that there are more degradation combinations in 3D modulation, including (noise+blur), (noise+JPEG), (blur+JPEG) and (noise+blur+JPEG). Here, we show some 2D modulation achieved by model of 3D modulation.

⋆ The first two authors are co-first authors. † Corresponding author

Table 1. 2D experiments. The PSNR distances within 0.2 dB are shown in bold. Lower is better.

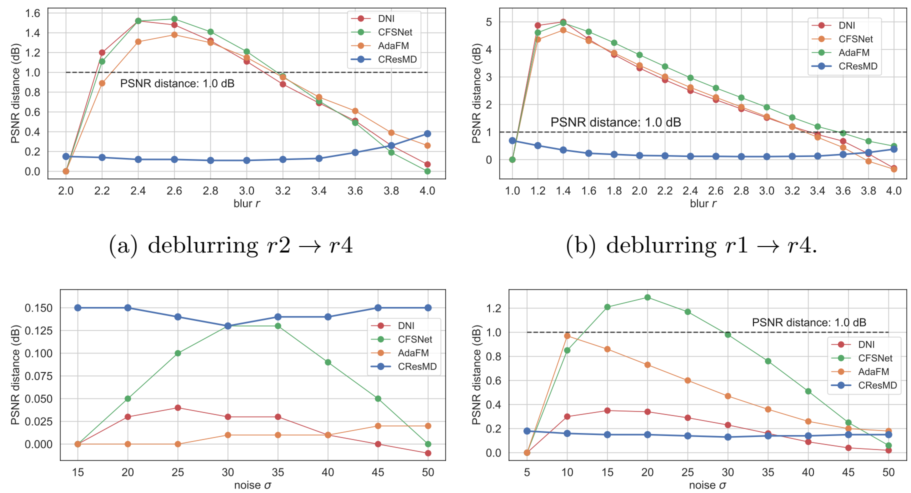
		one degradation						*two degradations*								
	blur r	1	2	4	0	0	0	1	1	1	2	2	2	4	4	4
	noise σ	0	0	0	15	30	50	15	30	50	15	30	50	15	30	50
CBSD68	upper bound	39.07	30.24	26.91	34.12	30.56	28.21	29.11	27.38	26.07	26.30	25.35	24.55	24.08	23.53	23.03
CBSD68	CResMD	38.38	30.09	26.53	33.97	30.43	28.06	29.00	27.27	25.96	26.24	25.29	24.48	24.03	23.46	22.95
CBSD68	PSNR distance	0.69	**0.15**	0.38	**0.15**	**0.13**	**0.15**	**0.11**	**0.11**	**0.11**	**0.06**	**0.06**	**0.07**	**0.05**	**0.07**	**0.08**
LIVE1	upper bound	39.76	30.27	26.74	34.29	30.83	28.46	29.26	27.44	26.06	26.05	25.07	24.26	23.70	23.16	22.68
LIVE1	CResMD	38.85	30.03	26.14	34.10	30.65	28.25	29.10	27.29	25.90	25.96	24.98	24.17	23.62	23.06	22.56
LIVE1	PSNR distance	0.91	0.24	0.60	**0.19**	**0.18**	0.21	**0.16**	**0.15**	**0.16**	**0.09**	**0.09**	**0.09**	**0.08**	**0.10**	**0.12**

Table 2. 3D experiments. The PSNR distances within 0.2 dB are shown in bold. Lower is better.

		one						*two*						*three*	
	blur r	1	4	0	0	0	0	1	4	1	4	0	0	1	4
	noise σ	0	0	15	50	0	0	15	50	0	0	15	50	15	50
	JPEG q	∞	∞	∞	∞	80	10	∞	∞	80	10	80	10	80	10
CBSD68	upper bound	39.07	26.91	34.12	28.21	36.22	27.63	29.11	23.03	31.30	23.25	32.71	26.21	28.65	22.61
CBSD68	CResMD	38.20	26.43	33.92	28.01	35.93	27.37	28.97	22.93	30.96	22.99	32.58	26.00	28.55	22.49
CBSD68	PSNR distance	0.87	0.48	**0.20**	**0.20**	0.29	0.26	**0.14**	**0.10**	0.34	0.26	**0.13**	0.21	**0.10**	**0.12**
LIVE1	upper bound	39.76	26.74	34.29	28.46	36.19	27.64	29.26	22.68	31.48	22.80	32.90	26.29	28.75	22.25
LIVE1	CResMD	38.62	26.01	34.05	28.18	35.78	27.36	29.06	22.53	31.02	22.50	32.68	26.00	28.59	22.07
LIVE1	PSNR distance	1.14	0.73	0.24	0.28	0.41	0.28	**0.20**	**0.15**	0.46	0.30	0.22	0.29	**0.16**	**0.18**

3 Comparison with SD methods

In this section, we compare the proposed CResMD with state-of-the-art SD methods. Specifically, we conduct testing on deblurring $r1 \rightarrow r4$, $r2 \rightarrow r4$ and denoising $\sigma5 \rightarrow \sigma50$, $\sigma15 \rightarrow \sigma50$. Our model directly uses the results in the 2D experiments for comparison. For deblurring, the SD methods performs poorly in almost all intermediate points especially when the adaptation range is relatively large. For instance, in deblurring $r1 \rightarrow r4$, the PSNR distances achieved by SD methods are close to 5 dB in $r1.2$ and $r1.4$. For denoising within a smaller adaptation range ($\sigma15 \rightarrow \sigma50$), all SD methods perform well in both two ends and intermediate points (< 0.15 dB), even outperform our CResMD. However, in denoising $\sigma5 \rightarrow \sigma50$, the interpolation results obtained by AdaFM and CFSNet have very large PSNR distances (1 dB), while DNI performs comparably with CResMD. To conclude, the SD methods are generally sensitive to the adaptation range, while our CResMD is more robust to degradation types and ranges.



(a) deblurring $r2 \rightarrow r4$

(b) deblurring $r1 \rightarrow r4$.

(c) denoising $\sigma15 \rightarrow \sigma50$.

(d) denoising $\sigma5 \rightarrow \sigma50$.

Fig. 1. Comparison with SD methods on CBSD68 data set.

4 Ablation Study

4.1 Effectiveness of Global Connection.

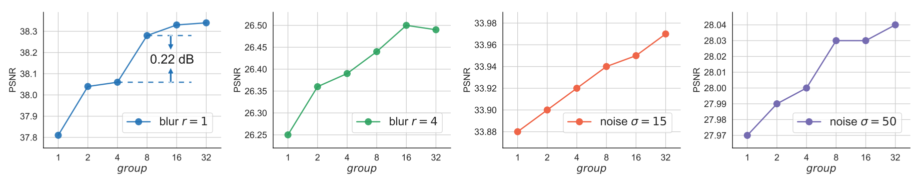
To evaluate the effectiveness of global connection, we conduct a straightforward comparison experiment by just removing the global connection. As shown in Table 3, the model with global controllable global connection could not only achieve better performance on all mild degradations but also in other degradations. This indicates that global connection is effective to deal with a whole restoration range with a zero starting point.

		mild degradations								*other degradations*						
	blur r	0	0	0	0.5	1	0.5	0.5	1	2	4	0	0	1	2	4
	noise σ	0	5	15	0	0	5	15	5	0	0	30	50	15	30	50
CBSD68	w/o	71.39	40.21	33.85	52.70	38.04	37.80	32.31	31.48	29.91	26.33	30.31	27.93	28.90	25.23	22.86
	w	$+\infty$	40.33	33.97	53.17	38.38	37.92	32.44	31.63	30.09	26.53	30.43	28.06	29.00	25.29	22.95
	gain	$+\infty$	0.12	0.12	0.47	0.34	0.12	0.13	0.15	0.18	0.20	0.12	0.13	0.10	0.06	0.09
LIVE1	w/o	64.17	39.79	33.93	51.22	38.38	37.71	32.51	31.65	29.79	25.86	30.49	28.08	27.17	24.88	22.43
	w	$+\infty$	39.99	34.10	52.21	38.85	37.89	32.69	31.86	30.03	26.14	30.65	28.25	27.29	24.98	22.56
	gain	$+\infty$	0.20	0.17	0.99	0.47	0.18	0.18	0.21	0.24	0.28	0.16	0.17	0.12	0.10	0.13

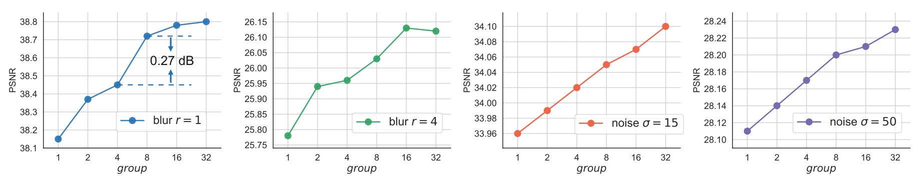
Table 3. The effetiveness of global connection.

4.2 Effectiveness of Local Connection.

In this part, we explore the effectiveness of local connection by dividing all the building blocks into 1, 2, 4, 8, 16 and 32 groups, and then add one controllable residual connection to each group.



(a) CBSD68



(b) LIVE1

Fig. 2. Performance under different local connections.

As shown in Figure 2, the results evaluated on CBSD68 and LIVE1 datasets share similar trend. In general, more local connections could lead to better performance. In particular, we observe sharp leaps (0.22dB and 0.27dB) from 4 to 8 local connections in deblurring $r = 1$. For deblurring $r = 4$, the performance starts to stop increasing when there is 16 local connections. On the other hand, 1 and 32 local connections make no big difference in the denoising tasks (< 0.1 dB on CBSD68 while < 0.15 on LIVE1).

4.3 Effectiveness of Data Sampling.

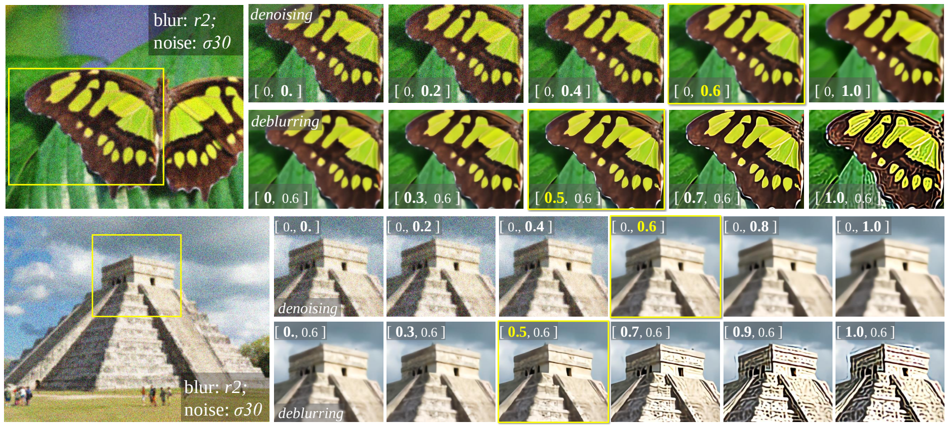
Here we investigate different data sampling strategies. Specifically, we set α, β of the beta distribution to be (1.0, 1.0) (0.5, 1.0), (0.2, 1.0) and (1.0, 2.0) for sampling degradations. In particular, we use uniform sampling ($\alpha = 0.5, \beta = 1.0$) as our baseline and the PSNR distances calculated with other sampling strategies are presented as shown in Table 4, 5. In general, the performance on mild degradations improves significantly when more mild degradations are sampled. On the other hand, the performance on severe degradations where less data are provided decreases. It is obviously observed that the performance of task deblurring $r1$ improves at the cost of severe degradation in task $r4$ on testing of both datasets. Besides, denoising is more robust to the distribution of the degradation levels unless the sampling is extremely biased (e.g. $\alpha = 0.2, \beta = 1.0$).

blur r	1	2	4	0	0	0	1	2	4	
noise σ	0	0	0	5	30	50	5	30	50	total
$\alpha = 1.0, \beta = 1.0$	38.26	30.07	**26.58**	40.25	30.41	**28.06**	31.58	**25.30**	**22.99**	
$\alpha = 0.5, \beta = 1.0$	38.38	30.09	26.53	40.33	30.43	28.06	31.63	25.29	22.95	
(CResMD)	+0.12	+0.02	−0.05	+0.08	**+0.02**	+0.00	+0.05	−0.01	−0.04	**+0.19**
$\alpha = 0.2, \beta = 1.0$	38.43	30.05	26.45	40.32	33.35	27.94	31.50	25.21	22.83	
	+0.17	−0.02	−0.13	+0.07	−0.06	−0.12	−0.08	−0.09	−0.16	−0.42
$\alpha = 1.0, \beta = 2.0$	38.47	30.11	26.25	40.38	30.43	28.06	31.66	25.29	22.95	
	+0.21	**+0.04**	−0.33	**+0.13**	+0.02	+0.00	**+0.08**	−0.01	−0.04	+0.10

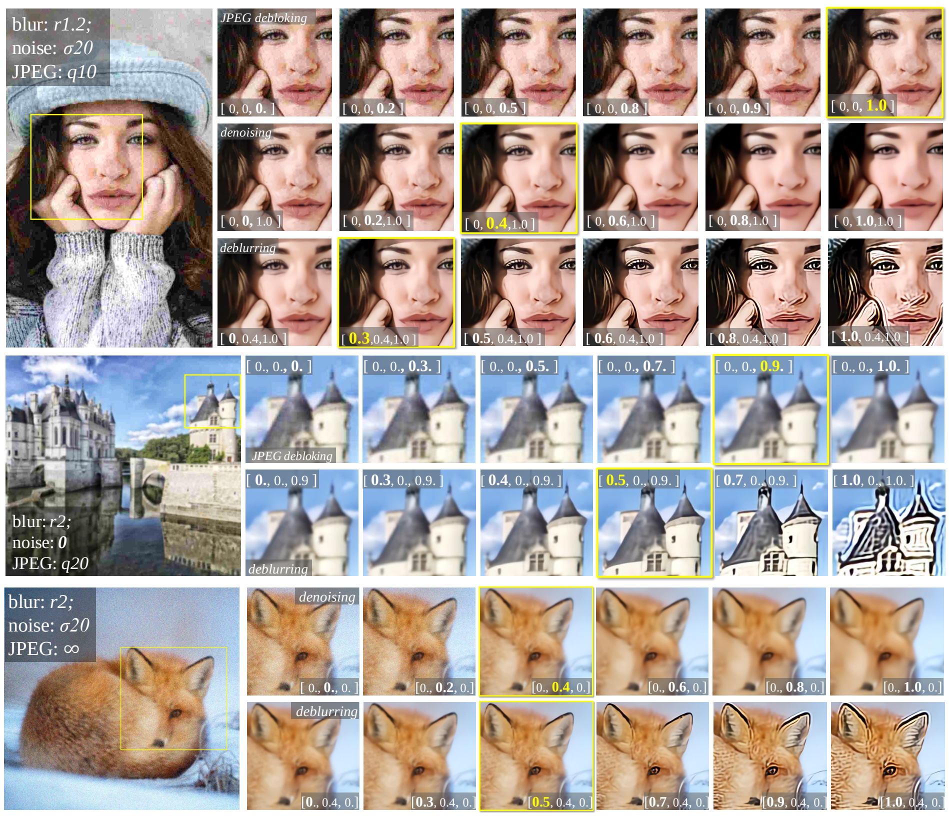
Table 4. Performance under different sampling curves evaluated on CBSD68.

blur r	1	2	4	0	0	0	1	2	4	
noise σ	0	0	0	5	30	50	5	30	50	total
$\alpha = 1.0, \beta = 1.0$	38.66	30.01	**26.26**	39.90	30.63	28.24	31.78	24.97	**22.58**	
$\alpha = 0.5, \beta = 1.0$	38.85	30.03	26.14	39.99	30.65	28.25	31.86	24.98	22.56	
(CResMD)	+0.19	+0.02	−0.12	+0.09	+0.02	**+0.01**	+0.08	+0.01	−0.02	**+0.28**
$\alpha = 0.2, \beta = 1.0$	38.94	29.98	26.07	39.97	30.55	28.10	31.68	24.85	22.39	
	+0.28	−0.03	−0.19	+0.07	−0.08	−0.14	−0.10	−0.12	−0.19	−0.50
$\alpha = 1.0, \beta = 2.0$	38.93	30.08	25.80	40.00	30.66	28.24	31.90	24.99	22.50	
	+0.27	**+0.07**	−0.46	**+0.10**	**+0.03**	+0.00	**+0.12**	**+0.02**	−0.08	+0.07

Table 5. Performance under different sampling curves evaluated on LIVE1.



(a) 2D modulation.



(b) 3D modulation.

Fig. 3. Qualitative results of MD modulation. In each row, we only change one factor with other factors fixed. We arrive at the best choice in the yellow box. **Better view in zoom and color.**