SCIGS: 3D GAUSSIANS SPLATTING FROM A SNAPSHOT COMPRESSIVE IMAGE

Supplementary Material

In this supplementary material, the results of comparative experiments on all the datasets from dynamic scenes and static scenes are shown. The experiment compares our SCIGS against current state-of-the-art SCI decoding methods (GAP-TV[3], PnP-FFDNet[4], PnP-FastDVDNet[5] and EfficientSCI[2]) and state-of-the-art SCI image-based reconstruction method(SCINeRF[1]). An additional experiment is conducted to assess the impact of various mask overlapping rates during the SCI image modulated.

A. Comparative Experiments

A.1. Experiment Setup

To further validate the effectiveness of our method in dynamic scenes, additional qualitative and quantitative experiments were conducted on five datasets from dynamic scene (*Bear*, *Roundabout*, *Flamingo*, *Turn* and *Dance*). For fair comparisons, we fine-tuned EfficientSCI [2] with the masks used in our datasets. Additionally, the results of qualitative experiments conducted under all static scene datasets (*Factory*, *Tanabata*, *Vender*, *Cozy2room*, *hotdog* and *airplants*) are also presented in this supplementary material. For a better quantitative comparison, we also present the results of the experiments with static scenes, which are shown in Table B.

A.2. Result and Analysis

The results of the qualitative and quantitative experiments in dynamic scenes are shown in Fig. A and Table A, respectively. These results provide empirical evidence for the effectiveness of our SCIGS in reconstructing dynamic scenes from single compressed images. It is also worth noting that the metrics of our method do not exceed EfficientSCI in *Dance*. The observation can be attributed to the fact that our method recovers images by reconstructing the underlying scene. The images in the *Dance* dataset have dynamic blur, which leads to the loss of structural information, so SCIGS cannot accurately reconstruct this part of the scene, leading to a degradation in image quality. In contrast, as a traditional SCI image decoding methods, EfficientSCI uses only 2D image information without considering the structural consistency, and thus outperforms our method in this scene.

As shown in Fig. B and Table B, the proposed SCIGS shows comparable image recovery performance on static scene. In addition, we note that SCIGS outperforms existing methods in the reconstruction of the parts with rich textures and characters, which cannot be directly observed from metrics.

B. Additional Experiment

We assess the impact of various mask overlapping rates during the SCI image modulated. The mask overlapping rate is defined by the probability that a mask selects a specific pixel for exposure, which is formulated as follow:

$$OR(x,y) = \frac{\sum_{i=1}^{N} M_i(x,y)}{N} \tag{1}$$

where OR denotes mask overlapping rate, M_i indicates i-th mask and N is the number of Intermediate frame. From Eq. 1, lower the mask crossing rate means the sparser sampling, which result in the less image information retained, and conversely, the denser the sampling leads to the more information retained. However, too high a sampling rate will increase the ambiguity of the compressed image and may result in a blurred decoded image. As shown in Table C, we tested different overlapping rates on multiple datasets, and the results showed that the image quality first increased and then decreased when the overlapping rate increased from 0.125 to 0.25, and decreased after 0.25. Empirically, we selected overlap rate of all experiments within 0.25.

	Bear			Roundabout			Turn			Flamingo			Dance		
	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
GAP-TV[3]	22.63	.5698	.3734	22.26	.6976	.3823	25.28	.6774	.3437	23.68	.6986	.3404	22.20	.6981	.3953
PnP-FFDNet[4]	21.91	.6569	.3822	25.80	.8727	.1314	26.93	.8598	.2661	25.50	.8206	.2000	22.29	.8284	.1987
PnP-FastDVDNet[5]	26.77	.8561	.1413	27.01	.8938	.1006	27.58	.8723	.2090	29.27	.8978	.0994	28.10	.9465	.0569
EfficientSCI[2]	29.26	.9099	.0710	28.45	.9110	.0876	29.03	.8934	.1617	31.03	**.9247**	.0668	**31.55**	**.9677**	**.0412**
SCINerf[1]	26.57	.7974	.1192	26.02	.8394	.1265	25.68	.6596	.2330	26.78	.7954	.1207	22.78	.6960	.2737
SCIGS(ours)	**30.44**	**.9137**	**.0548**	**31.07**	**.9222**	**.0729**	**31.78**	**.8951**	**.0953**	**31.33**	.9022	**.0533**	27.89	.9096	.0580

Table A. **Quantitative SCI image reconstruction comparisons on the dynamic datasets.** The results demonstrate that our method surpasses the current SCI decoding methods and 3D reconstruction methods for SCI image on datasets from dynamic scenes. The best results are shown in bold and the second-best results are underlined.

	Cozy2room			Tanabata			Factory			Vender			Airplants			Hotdog		
	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
GAP-TV[3]	21.77	.4321	.6031	20.42	.4264	.6250	24.05	.5666	.5149	20.00	.3678	.6882	22.85	.4057	.4986	22.35	.7663	.3179
PnP-FFDNet[4]	28.98	.8916	.0984	29.17	.9032	.1197	31.75	.8977	.1142	28.70	.9235	.1315	27.79	.9117	.1817	29.00	.9765	.0511
PnP-FastDVDNet[5]	30.19	.9132	.0793	29.73	.9333	.0980	32.53	.9165	.1055	29.68	.9395	.1043	28.18	.9092	.1757	29.93	.9728	.0522
EfficientSCI[2]	31.47	.9327	.0476	32.30	.9587	.0600	32.87	.9259	.0709	33.17	.9401	.0456	30.13	**.9425**	.1129	30.75	.9568	.0461
SCINerf[1]	33.23	**.9492**	.0445	33.61	**.9638**	.0374	36.60	.9638	**.0221**	**36.40**	**.9840**	.0298	**30.69**	.9335	**.0728**	**31.35**	**.9878**	**.0310**
SCIGS(ours)	**33.78**	.9191	**.0423**	**35.12**	.9580	**.0271**	**37.75**	**.9646**	.0291	36.00	.9641	**.0192**	27.18	.7267	.3003	29.31	.9369	.0809

Table B. **Quantitative SCI image reconstruction comparisons on the static datasets.** The results demonstrate that our method outperforms or approaches the existing image reconstruction methods and 3D reconstruction methods for SCI image on most datasets from static scenes. The best results are shown in bold and the second-best results are underlined.

OR	PSNR↑	SSIM↑	LPIPS↓
0.125	30.32	.9066	.0634
0.25	30.41	.8954	.0814
0.5	29.04	.8569	.1145
0.75	27.50	.8294	.1336

Table C. **The average metrics of image quility in the additional study on mask overlapping rate.** the quality of reconstruction increases first and then decreases with the overlapping rate ranging from 0.125 to 0.75.

References

[1] Yunhao Li, Xiaodong Wang, Ping Wang, Xin Yuan, and Peidong Liu. Scinerf: Neural radiance fields from a snapshot compressive image. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10542–10552, 2024. 1, 2

[2] Lishun Wang, Miao Cao, and Xin Yuan. Efficientsci: Densely connected network with space-time factorization for large-scale video snapshot compressive imaging. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 18477–18486, 2023. 1, 2

[3] Xin Yuan. Generalized alternating projection based total variation minimization for compressive sensing. In *2016 IEEE International conference on image processing (ICIP)*, pages 2539–2543. IEEE, 2016. 1, 2

[4] Xin Yuan, Yang Liu, Jinli Suo, and Qionghai Dai. Plug-and-play algorithms for large-scale snapshot compressive imaging. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1447–1457, 2020. 1, 2

[5] Xin Yuan, Yang Liu, Jinli Suo, Fredo Durand, and Qionghai Dai. Plug-and-play algorithms for video snapshot compressive imaging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(10):7093–7111, 2021. 1, 2

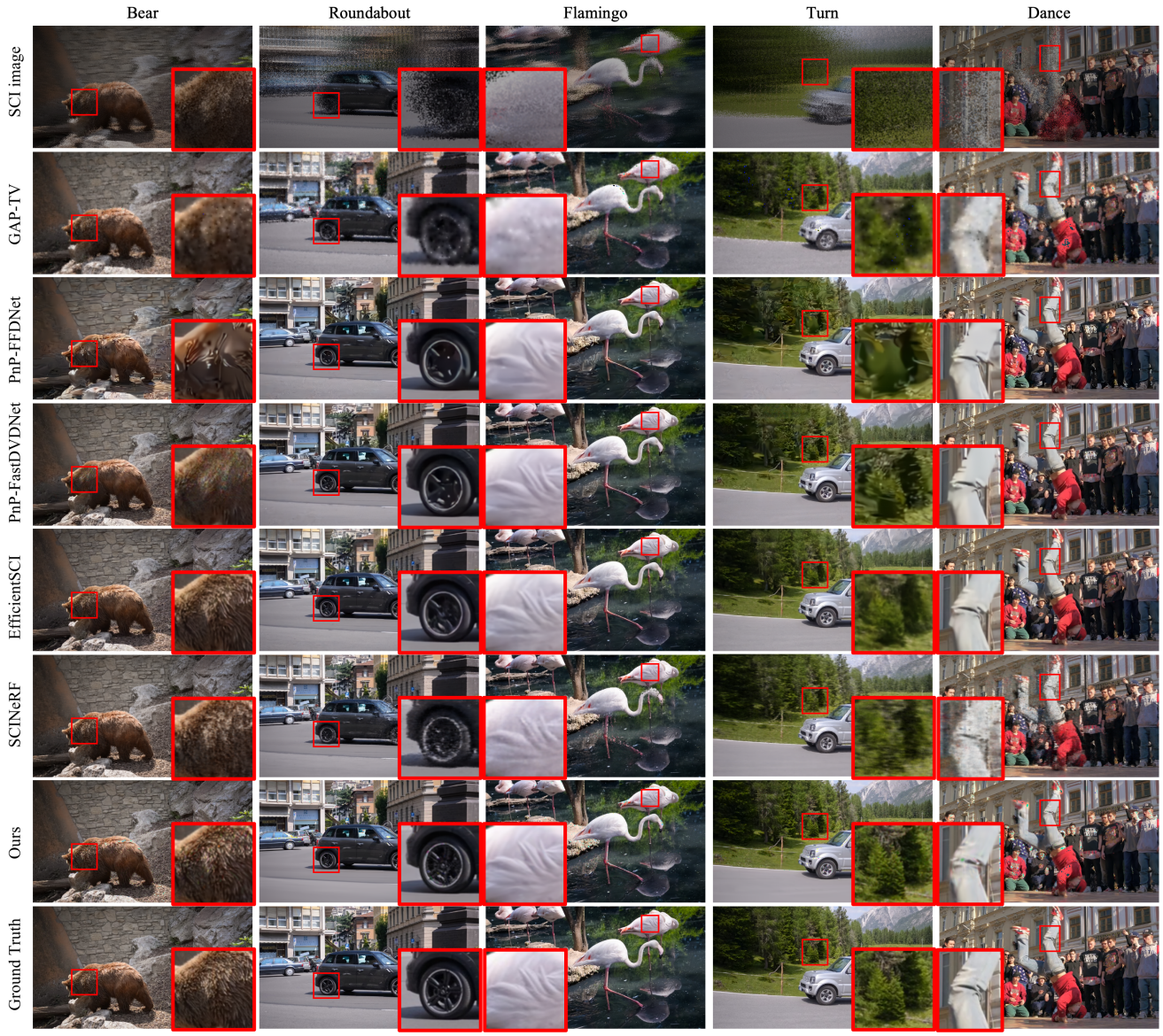

Figure A. **Qualitative evaluations on the datasets from dynamic scenes.** From left to right shows the results for five dynamic scenes including *Bear*, *Roundabout*, *Flamingo*, *Turn* and *Dance*. The experiments show that our method achieves superior performance in dynamic scenes.

Figure B. **Qualitative evaluations on the datasets from static scenes.** From left to right shows the results for five static scenes including *Factory*, *Tanabata*, *Vender*, *Cozy2room*, *Hotdog* and *Airplants*. The experiments show that our method achieves comparable image recovery performance from a single compressed image in static scenes.