LEARNING EPIPOLAR FEATURE FIELDS FOR MULTI-IMAGE SUPER-RESOLUTION - SUPPLEMENTARY MATERIAL

Anonymous authors
Paper under double-blind review

1 WIDER-BASELINE EXPERIMENT

In this section we present an experiment where views are taken very far apart and asymmetrically with respect to the target view in order to challenge the method and the state-of-the-art BSRT. Table 1 reports the PSNR obtained by BSRT and EpiMISR when compared to the SISR PSNR. It can be noticed that in this challenging setting, BSRT degrades to the SISR performance, while EpiMISR still provides an improvement. This more challenging geometry is created by taking the V-1 extra views that are at median distance (out of all the views available in the dataset) with respect to the distance to the target view camera center.

Table 1: Challenging geometry setting.

		No. Params	PSNR ↑	BRISQUE ↓	LPIPS ↓	SSIM ↑
	EpiMISR	23.30M	**27.00**	**41.20**	**0.15**	0.82
4×	BSRT (Luo et al. (2022))	20.56M	26.82	46.34	0.16	**0.83**
	SwinIR (Liang et al. (2021))	14.70M	26.87	45.68	0.17	0.82

2 VIEW CONSISTENCY

In this section we present an experiment where the view consistency is assessed. As the setting we study is that of not-novel view synthesis, we are only concerned with generating details that are consistent with the LR observations of the target view we want to super-resolve, and it is outside the scope of the method to enable novel view synthesis. The transformers used as building blocks of our method implicitly ensure that only consistent information is borrowed from the other views via the attention mechanism. Table 2 reports an additional result about the PSNR between the LR target image and the SR target image when degraded to LR. This assesses how different methods to solve the inverse problem ensure consistency with the observations. Moreover, we repeat it to super-resolve all the images in the scenes to simulate the case in which one wants to super-resolve to entire image set (possibly for further downstream tasks) rather than just one view.

Table 2: View consistency. PSNR between the degraded SR images and the LR images.

	LR - PSNR (dB) ↑
EpiMISR	**30.71**
BSRT (Luo et al. (2022))	30.08
SwinIR (Liang et al. (2021))	29.14

3 EXPERIMENTS ON IBRNET DATASET AND LLFF DATASET

In this section we report our results on the 1023 scenes from the Google Scanned Objects dataset used for IBRNet and on the LLFF dataset, for a $4\times$ SR factor. Table 3 reports the evaluation results of EpiMISR, BSRT and SwinIR methods on the Google Scanned Objects dataset, while 4 on LLFF dataset. All the methods are trained as described in the main paper and are not finetuned on the

IBRNet dataset, hence these results shows that EpiMISR out-performs baselines even on an unseen data distribution.

Table 3: Quantitative results for MISR on IBRNet dataset.

		No. Params	PSNR ↑	BRISQUE ↓	LPIPS ↓	SSIM ↑
$4\times$	**EpiMISR**	23.30M	**31.50**	73.03	**0.04**	**0.96**
	BSRT (Luo et al. (2022))	20.56M	30.09	74.25	0.05	0.95
	SwinIR (Liang et al. (2021))	14.70M	29.29	**73.02**	0.07	0.95

Table 4: Quantitative results on the LLFF scenes.

[PSNR (dB) ↑]	Fern	Flower	Fortress	Horns	Leaves	Orchids	Room	Trex	Mean
EpiMISR	**21.08**	26.15	**27.48**	**23.64**	16.53	**20.09**	**27.17**	**22.41**	**23.07**
BSRT (Luo et al. (2022))	20.61	**26.42**	26.74	23.32	**17.54**	19.82	26.19	22.18	22.85
SwinIR (Liang et al. (2021))	20.35	25.60	26.09	22.73	16.90	19.99	24.87	21.66	22.27

4 Qualitative Results

Fig. 1 reports qualitative results for the proposed method and baselines on some scenes. Fig. 2 shows a challenging scene where BSRT outperforms EpiMISR. Fig. 3 shows the Empirical Cumulative Distribution Function (ECDF) of the PSNR improvements of EpiMISR with respect to BSRT on all the DTU dataset test split. The failure cases, that are the instances in the test dataset where BSRT outperforms EpiMISR, are rare, as the ECDF$(0) \approx 2.04\%$.

Figure 1: Qualitative results of some DTU test scenes with $4\times$ scale factor. From left to right: LR nearest neighbours interpolation, NeRF-SR, BSRT, EpiMISR, HR ground truth.

Figure 1 (cont.)

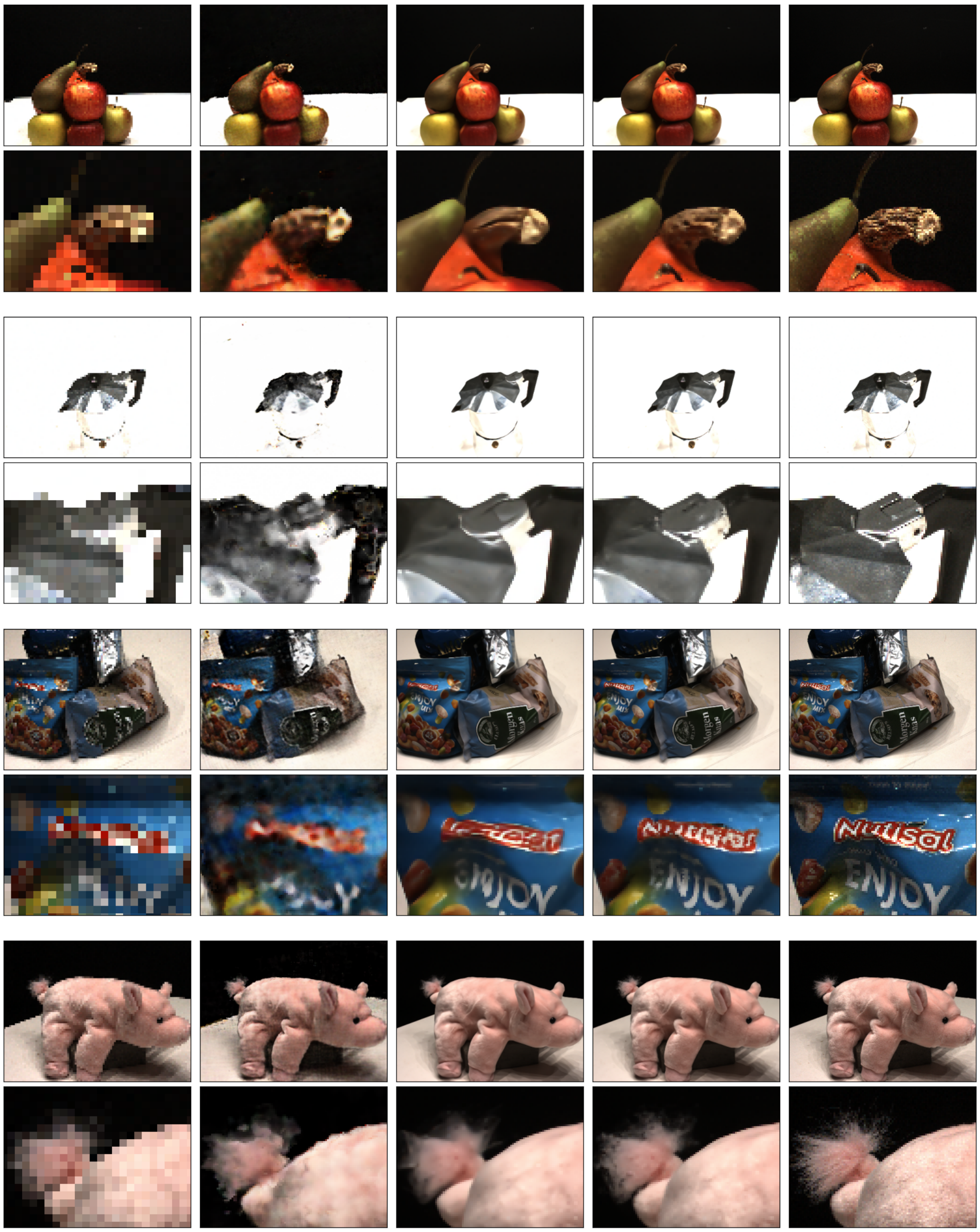

Figure 2: A qualitative example of a failure case (DTU dataset, scan 63). This is an example where BSRT outperforms EpiMISR. From left to right: LR nearest neighbours interpolation, NeRF-SR, BSRT, EpiMISR, HR ground truth.

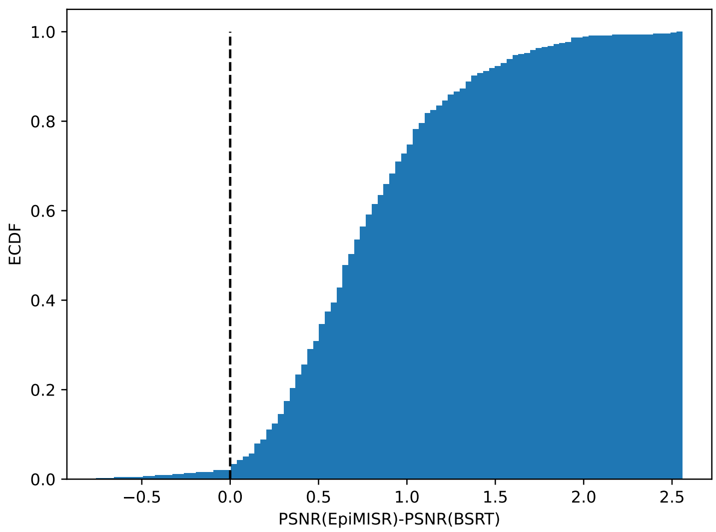

Figure 3: Empirical Cumulative Distribution Function (ECDF) of the PSNR improvements of EpiMISR with respect to BSRT on the test split of the DTU dataset.

REFERENCES

Jingyun Liang, Jiezhang Cao, Guolei Sun, Kai Zhang, Luc Van Gool, and Radu Timofte. Swinir: Image restoration using swin transformer. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 1833–1844, 2021.

Ziwei Luo, Youwei Li, Shen Cheng, Lei Yu, Qi Wu, Zhihong Wen, Haoqiang Fan, Jian Sun, and Shuaicheng Liu. Bsrt: Improving burst super-resolution with swin transformer and flow-guided deformable alignment. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 998–1008, 2022.