Learning Non-Local Spatial-Angular Correlation for Light Field Image Super-Resolution (*Supplemental Material*)

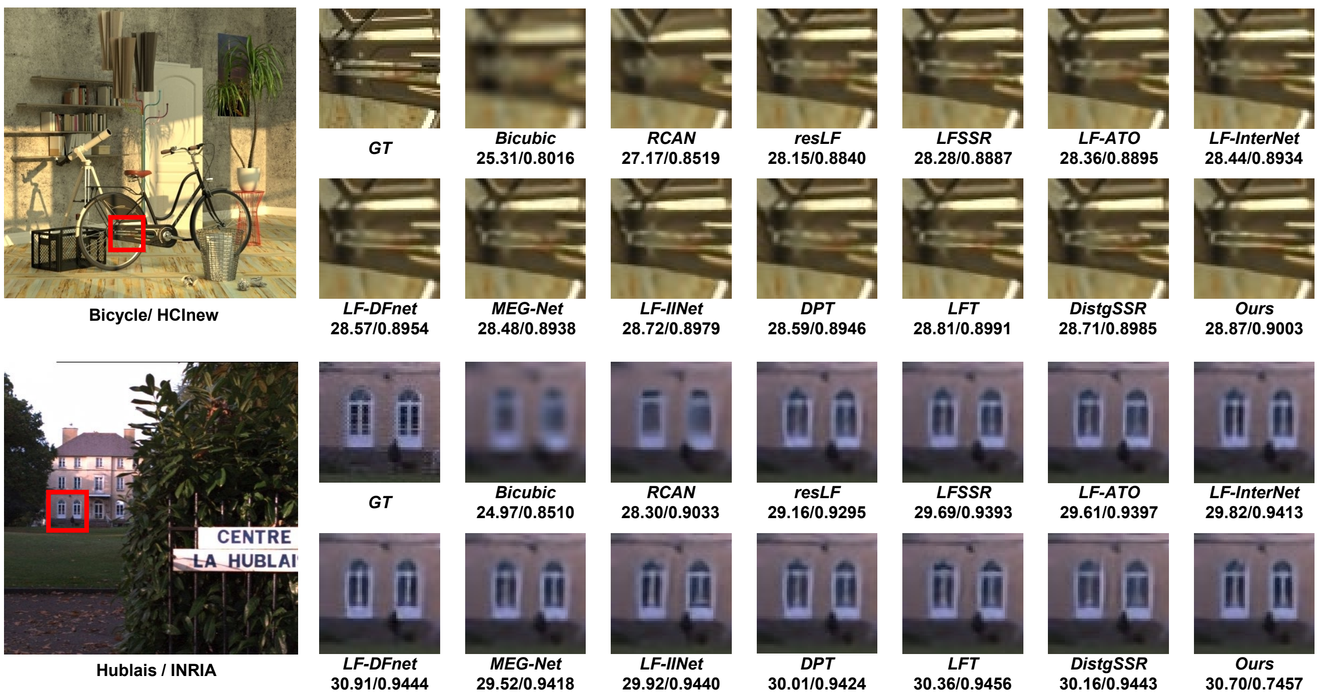

Figure I. Qualitative comparison of different SR methods for 4× SR.

Section A provides more visual comparisons on the light field (LF) datasets, and presents additional comparisons on LFs with different angular resolution. Section B presents detailed quantitative results of different methods on each dataset with various sheared values. Section C describes additional experiments for LF angular SR, and shows visual results achieved by different methods.

A. Additional Comparisons on Benchmarks

A.1. Qualitative Results

In this subsection, we show more visual comparisons of 4× SR on the benchmark dataset in Fig. I. It can be observed that the proposed EPIT recovers richer and more realistic details.

A.2. Robustness to Different Angular Resolution

In the main body of our paper, we have illustrated that our EPIT (trained on central 5×5 SAIs) achieves competitive PSNR scores on other angular resolutions, as compared to top-performing DistgSSR [60]. In Table I, we provide more quantitative results achieved by the state-of-the-art methods with different angular resolutions.

In addition, we train a series of EPIT models from scratch on 2×2, 3×3 and 4×4 SAIs, respectively. It can be observed from Table II that when using larger angular resolution SAIs as training data, e.g., 5×5, our method can achieve better SR performance on different angular resolutions. That is because, more angular views are beneficial for our EPIT to learn the spatial-angular correlation better. This phenomenon inspires us to explore the intrinsic mechanism of LF processing tasks in the future.

B. Additional Quantitative Comparison on Disparity Variations

We have presented the performance comparison on two selected scenes with different shearing values for 2× SR in the main paper. Here, we provide quantitative results on each dataset in Table III and Fig. II. It can be observed that our EPIT achieves more consistent performance than existing methods with respect to disparity variations on various datasets.

Table I. PSNR/SSIM values achieved by different methods with different angular resolution for 4× SR.

Datasets		Methods				
		resLF	*LFSSR*	*MEG-Net*	*LFT*	*EPIT(ours)*
EPFL [48]	2×2	-	26.00/.8541	26.40/.8667	27.64/.8953	28.22/.9024
	3×3	28.13/.9012	26.84/.8750	27.16/.8834	28.12/.9029	28.74/.9103
	4×4	-	27.62/.8930	28.04/.9036	28.43/.9087	29.04/.9164
	5×5	28.27/.9035	28.27/.9118	28.74/.9160	29.85/.9210	29.34/.9197
	6×6	-	27.62/.8995	28.46/.9115	28.45/.9101	29.43/.9218
	7×7	27.91/.9038	27.29/.8889	28.30/.9083	28.55/.9094	29.60/.9231
	8×8	-	27.06/.8834	28.15/.9061	28.37/.9064	29.60/.9240
	9×9	26.07/.8881	26.95/.8810	28.12/.9046	28.45/.9071	29.71/.9246
HCInew [21]	2×2	-	28.44/.8639	29.02/.8782	29.94/.8960	30.84/.9114
	3×3	30.63/.9089	29.47/.8848	29.84/.8943	30.28/.9031	31.23/.9182
	4×4	-	30.22/.8997	30.68/.9094	30.51/.9065	31.40/.9213
	5×5	30.73/.9107	30.72/.9145	31.10/.9177	31.46/.9218	31.51/.9231
	6×6	-	30.24/.9053	30.91/.9154	30.26/.9009	31.57/.9241
	7×7	30.23/.9112	29.89/.8997	30.64/.9125	30.05/.8975	31.63/.9250
	8×8	-	29.68/.8969	30.48/.9105	29.81/.8923	31.66/.9256
	9×9	27.84/.8967	29.46/.8942	30.34/.9087	29.77/.8916	31.69/.9260
HCIold [65]	2×2	-	33.37/.9413	34.17/.9489	35.52/.9591	36.94/.9690
	3×3	36.61/.9674	34.72/.9535	35.26/.9579	35.91/.9616	37.37/.9717
	4×4	-	35.80/.9615	36.42/.9662	36.15/.9634	37.52/.9729
	5×5	36.71/.9682	36.70/.9696	37.28/.9716	37.63/.9735	37.68/.9737
	6×6	-	35.32/.9617	36.75/.9688	36.21/.9636	37.76/.9744
	7×7	36.21/.968	34.94/.9578	36.35/.9662	36.10/.9629	37.92/.9749
	8×8	-	34.70/.9558	36.18/.9651	35.73/.9596	38.00/.9754
	9×9	33.55/.9519	34.46/.9539	36.08/.9644	35.71/.9593	38.06/.9756
INRIA [46]	2×2	-	27.83/.9035	28.31/.9125	29.99/.9378	30.52/.9418
	3×3	30.33/.9413	28.78/.9201	29.16/.9264	30.35/.9424	30.94/.9472
	4×4	-	29.59/.9327	30.00/.9401	30.64/.9457	31.19/.9509
	5×5	30.34/.9412	30.31/.9467	30.66/.9490	31.20/.9524	31.27/.9526
	6×6	-	29.50/.9356	30.38/.9443	30.61/.9457	31.45/.9533
	7×7	29.82/.9398	29.05/.9269	30.13/.9415	30.56/.9443	31.51/.9539
	8×8	-	28.76/.9221	30.02/.9399	30.41/.9422	31.54/.9540
	9×9	27.65/.9226	28.58/.9196	29.97/.9386	30.43/.9420	31.56/.9539
STFgantry [53]	2×2	-	27.29/.8710	28.15/.8944	29.69/.9263	31.30/.9468
	3×3	30.05/.9348	28.81/.9064	29.22/.9161	30.05/.9316	31.86/.9534
	4×4	-	29.77/.9254	30.30/.9356	30.35/.9359	32.11/.9558
	5×5	30.19/.9372	30.15/.9426	30.77/.9453	31.86/.9548	32.18/.9571
	6×6	-	29.79/.9320	30.58/.9428	30.01/.9289	32.31/.9580
	7×7	29.71/.9375	29.40/.9257	30.25/.9393	29.53/.9208	32.40/.9585
	8×8	-	29.12/.9211	30.03/.9367	29.17/.9135	32.48/.9591
	9×9	27.23/.9224	28.85/.9169	29.83/.9344	29.06/.9110	32.50/.9592

Table II. PSNR/SSIM values achieved by our EPIT trained on LFs with different angular resolution for 4× SR.

Datasets		*EPIT(ours)**			
		2×2	3×3	4×4	5×5
EPFL [48]	2×2	28.40/.9037	28.45/.9040	28.33/.9034	28.22/.9024
	3×3	28.61/.9076	28.75/.9090	28.67/.9090	28.74/.9103
	4×4	28.69/.9108	28.90/.9131	28.86/.9137	29.04/.9164
	5×5	28.81/.9124	29.08/.9152	29.06/.9162	29.34/.9197
	6×6	28.81/.9133	29.13/.9168	29.12/.9180	29.43/.9218
	7×7	28.88/.9137	29.24/.9176	29.24/.9190	29.60/.9231
	8×8	28.86/.9140	29.25/.9184	29.25/.9198	29.60/.9240
	9×9	28.92/.9141	29.32/.9188	29.34/.9204	29.71/.9246
HCInew [21]	2×2	30.81/.9109	30.86/.9116	30.86/.9116	30.84/.9114
	3×3	30.84/.9124	31.06/.9157	31.09/.9162	31.23/.9182
	4×4	30.86/.9132	31.14/.9174	31.21/.9184	31.40/.9213
	5×5	30.86/.9134	31.19/.9184	31.27/.9197	31.51/.9231
	6×6	30.86/.9134	31.21/.9190	31.32/.9205	31.57/.9241
	7×7	30.85/.9133	31.23/.9194	31.35/.9211	31.63/.9250
	8×8	30.86/.9133	31.24/.9197	31.37/.9215	31.66/.9256
	9×9	30.85/.9132	31.25/.9199	31.39/.9219	31.69/.9260
HCIold [65]	2×2	36.83/.9683	36.85/.9682	36.81/.9679	36.94/.9690
	3×3	36.92/.9688	37.13/.9701	37.14/.9702	37.37/.9717
	4×4	36.95/.9692	37.21/.9708	37.27/.9712	37.52/.9729
	5×5	37.01/.9695	37.31/.9714	37.39/.9718	37.68/.9737
	6×6	37.00/.9696	37.33/.9717	37.44/.9723	37.76/.9744
	7×7	37.00/.9696	37.40/.9719	37.52/.9726	37.92/.9749
	8×8	36.99/.9696	37.41/.9721	37.56/.9729	38.00/.9754
	9×9	36.99/.9697	37.44/.9722	37.60/.9730	38.06/.9756
INRIA [46]	2×2	30.63/.9429	30.66/.9431	30.58/.9427	30.52/.9418
	3×3	30.82/.9458	30.91/.9465	30.87/.9466	30.94/.9472
	4×4	30.90/.9472	31.04/.9484	31.02/.9489	31.19/.9509
	5×5	30.95/.9483	31.14/.9498	31.14/.9506	31.27/.9526
	6×6	30.94/.9484	31.17/.9503	31.18/.9511	31.45/.9533
	7×7	30.93/.9485	31.20/.9506	31.22/.9515	31.51/.9539
	8×8	30.92/.9484	31.22/.9507	31.24/.9517	31.54/.9540
	9×9	30.91/.9481	31.22/.9506	31.26/.9516	31.56/.9539
STFgantry [53]	2×2	30.84/.9432	31.03/.9449	31.09/.9452	31.30/.9468
	3×3	30.93/.9447	31.39/.9493	31.49/.9503	31.86/.9534
	4×4	31.02/.9459	31.56/.9510	31.69/.9523	32.11/.9558
	5×5	30.99/.9459	31.58/.9518	31.74/.9534	32.18/.9571
	6×6	31.03/.9460	31.68/.9525	31.85/.9541	32.31/.9580
	7×7	31.03/.9459	31.70/.9526	31.90/.9545	32.40/.9585
	8×8	31.04/.9459	31.73/.9528	31.96/.9549	32.48/.9591
	9×9	31.02/.9457	31.74/.9529	31.97/.9550	32.50/.9592

* Note that, "$A \times A$" below "EPIT(ours)" denotes the models are trained on the LFs with corresponding angular resolution.

C. LF Angular SR

It is worth noting that the proposed spatial-angular correlation learning mechanism has large potential in multiple LF image processing tasks. In this section, we apply our proposed spatial-angular correlation learning mechanism to the LF angular SR task. We first introduce our EPIT-ASR model for LF angular SR. Then, we introduce the datasets and implementation details in our experiments. Finally, we present the preliminary but promising results as compared to the state-of-the-art LF angular SR methods.

C.1. Upsampling

Since our EPIT is flexible to LFs with different angular resolutions (as demonstrated in Sec. A.2), the EPIT-ASR model can be built by changing the upsampling stage of EPIT.

Here, we follow [60, 27] to take the 2×2 → 7×7 angular SR task as an example to introduce the angular upsampling module in our EPIT-ASR. Given the deep LF feature $\boldsymbol{F} \in \mathbb{R}^{2\times 2\times H\times W\times C}$, a 2×2 convolution without padding is first applied to the angular dimensions to generate an angular-downsampled feature $\boldsymbol{F}_{down} \in \mathbb{R}^{1\times 1\times H\times W\times C}$. Then, a 1×1 convolution is used to increase the channel dimension, followed by a 2D pixel-shuffling layer to generate the angular-upsampled feature $\boldsymbol{F}_{up} \in \mathbb{R}^{7\times 7\times H\times W\times C}$. Finally, a 3×3 convolution is applied to the spatial dimensions of $\boldsymbol{F}_{up}$ to generate the final output $\mathcal{L}_{RE} \in \mathbb{R}^{7\times 7\times H\times W}$.

C.2. Datasets and Implement Details

Following [27, 60], we conducted experiments on the HCInew [21] and HCIold [65] datasets. All LFs in these datasets have an angular resolution of 9×9. We cropped the central 7×7 SAIs with 64×64 spatial resolution as

Table III. Quantitative comparison of different SR methods on five datasets with different shearing values for 2× SR. We mark the best results in red and the second results in blue.

Datasets		Methods											
		Bicubic	*RCAN*	*resLF*	*LFSSR*	*LF-ATO*	*LF-InterNet*	*LF-DFnet*	*MEG-Net*	*LF-IINet*	*LFT*	*DistgSSR*	*Ours*
EPFL [48]	-4	29.95/.9372	33.47/.9640	32.41/.9582	31.90/.9550	32.59/.9593	32.15/.9573	32.69/.9597	32.07/.9560	32.24 /.9579	32.48/.9587	32.29/.9583	34.52/.9734
	-3	29.92/.9369	33.45/.9637	32.38/.9578	31.85/.9548	32.58/.9592	32.14/.9572	32.68/.9597	32.16/.9564	32.27/.9577	32.49/.9587	32.29/.9578	34.67/.9746
	-2	29.89/.9369	33.31/.9632	32.36/.9587	31.92/.9561	32.37/.9589	32.06/.9571	32.47/.9592	32.17/.9574	32.37/.9589	32.35/.9587	32.65/.9618	34.64/.9749
	-1	29.83/.9373	33.30/.9634	33.01/.9652	32.69/.9640	33.06/.9659	32.62/.9636	33.41/.9673	32.82/.9653	33.29/.9676	33.33/.9676	33.37/.9687	34.71/.9756
	0	29.74/.9376	33.16/.9634	33.62/.9706	33.68/.9744	34.27/.9757	34.14/.9760	34.40/.9755	34.30/.9773	34.68/.9773	34.80/.9781	34.81/.9787	34.83/.9775
	1	29.87/.9373	33.16/.9629	32.81/.9644	32.70/.9639	32.67/.9656	32.57/.9642	33.19/.9669	32.76/.9647	33.12/.9663	33.18/.9675	33.01/.9681	34.66/.9760
	2	29.91/.9370	33.37/.9633	32.28/.9579	31.87/.9548	32.47/.9597	32.00/.9569	32.45/.9593	31.85/.9560	32.15/.9577	32.42/.9598	32.04/.9581	34.69/.9750
	3	29.94/.9370	33.48/.9638	32.32/.9575	31.85/.9543	32.56/.9591	32.09/.9569	32.61/.9594	31.84/.9545	32.19/.9574	32.43/.9585	32.17/.9578	34.73/.9747
	4	29.98/.9372	33.52/.9641	32.40/.9579	31.97/.9550	32.57/.9592	32.15/.9572	32.68/.9597	31.93/.9554	32.19/.9575	32.46/.9586	32.15/.9579	34.59/.9736
HCInew [21]	-4	30.83/.9343	34.59/.9611	33.34/.9533	32.57/.9494	33.37/.9545	32.99/.9525	33.62/.9554	32.91/.9510	33.34/.9534	33.35/.9541	33.03/.9523	36.77/.9782
	-3	30.81/.9342	34.65/.9609	33.45/.9543	32.61/.9501	33.58/.9558	33.06/.9523	33.75/.9562	33.16/.9527	33.44/.9542	33.51/.9551	33.43/.9554	37.05/.9791
	-2	30.83/.9344	34.60/.9605	33.50/.9594	32.58/.9548	33.13/.9599	32.91/.9563	33.41/.9609	33.33/.9588	33.80/.9618	33.37/.9609	33.76/.9644	36.98/.9792
	-1	30.74/.9349	34.42/.9603	35.00/.9704	34.19/.9691	34.87/.9716	34.29/.9690	35.59/.9739	34.51/.9716	35.70/.9748	35.49/.9747	35.68/.9754	37.21/.9815
	0	31.89/.9356	34.98/.9603	36.69/.9739	36.81/.9749	37.24/.9767	37.28/.9763	37.44/.9773	37.42/.9777	37.74/.9790	37.84/.9791	37.96/.9796	38.23/.9810
	1	30.73/.9350	34.14/.9602	34.04/.9649	33.90/.9639	33.41/.9660	33.63/.9633	34.30/.9681	34.06/.9659	34.64/.9682	34.33/.9694	34.30/.9691	36.83/.9792
	2	30.79/.9344	34.30/.9605	32.99/.9547	32.64/.9509	32.84/.9566	32.65/.9527	32.80/.9560	32.43/.9517	32.99/.9546	33.10/.9571	32.31/.9546	36.31/.9787
	3	30.77/.9341	34.39/.9609	33.17/.9523	32.70/.9493	33.32/.9545	33.03/.9523	33.51/.9553	32.59/.9492	33.22/.9529	33.32/.9541	32.87/.9521	36.56/.9787
	4	30.79/.9343	34.36/.9612	33.16/.9530	32.74/.9499	33.19/.9545	32.99/.9526	33.40/.9553	32.61/.9497	33.13/.9532	33.21/.9543	32.70/.9521	36.40/.9778
HCIold [65]	-4	36.85/.9775	40.85/.9875	39.36/.9852	38.44/.9833	39.18/.9852	39.22/.9851	39.55/.9858	38.69/.9837	38.93/.9849	39.20/.9851	39.17/.9850	42.34/.9929
	-3	36.83/.9775	40.88/.9874	39.57/.9854	38.45/.9837	39.35/.9854	39.33/.9853	39.76/.9858	38.99/.9843	39.18/.9850	39.37/.9851	39.40/.9852	43.04/.9936
	-2	36.84/.9777	40.32/.9871	38.84/.9858	38.05/.9841	38.33/.9854	38.80/.9852	38.70/.9862	38.64/.9851	38.90/.9860	38.47/.9855	39.53/.9879	42.80/.9938
	-1	36.71/.9782	40.22/.9873	40.43/.9902	39.44/.9891	39.60/.9900	39.79/.9895	40.96/.9914	39.68/.9899	41.19/.9915	40.73/.9913	41.45/.9923	43.31/.9952
	0	37.69/.9785	41.05/.9875	43.42/.9932	43.81/.9938	44.20/.9942	44.45/.9946	44.23/.9941	44.08/.9942	44.84/.9948	44.52/.9945	44.94/.9949	45.08/.9949
	1	36.66/.9783	39.25/.9869	39.85/.9903	40.31/.9904	38.42/.9901	39.93/.9903	40.18/.9915	39.85/.9905	40.88/.9921	39.99/.9916	40.50/.9922	42.75/.9942
	2	36.74/.9779	39.78/.9871	38.77/.9862	38.50/.9844	38.25/.9862	38.70/.9856	38.41/.9865	38.17/.9847	38.64/.9861	38.61/.9867	38.33/.9863	42.31/.9939
	3	36.76/.9777	40.66/.9876	39.31/.9852	38.48/.9834	39.10/.9855	39.10/.9852	39.45/.9858	38.37/.9832	39.00/.9851	39.19/.9853	38.90/.9849	42.97/.9939
	4	36.80/.9776	40.70/.9877	39.21/.9853	38.68/.9838	39.03/.9855	39.09/.9853	39.35/.9859	38.36/.9834	38.68/.9848	39.00/.9851	38.68/.9848	42.67/.9935
INRIA [46]	-4	31.58/.9566	35.40/.9769	34.24/.9719	33.75/.9695	34.42/.9725	33.99/.9713	34.64/.9736	33.89/.9703	34.13/.9719	34.37/.9724	34.20/.9720	36.46/.9815
	-3	31.55/.9566	35.39/.9768	34.22/.9717	33.71/.9695	34.43/.9726	34.04/.9715	34.62/.9736	33.95/.9703	34.12/.9715	34.39/.9726	34.10/.9710	36.67/.9826
	-2	31.55/.9567	35.22/.9763	34.04/.9715	33.59/.9695	34.08/.9718	33.87/.9709	34.31/.9726	33.91/.9707	34.13/.9721	34.11/.9716	34.67/.9749	36.67/.9829
	-1	31.49/.9573	35.26/.9767	34.88/.9767	34.59/.9760	34.92/.9770	34.56/.9757	35.51/.9790	34.69/.9766	35.42/.9790	35.26/.9783	35.55/.9799	36.79/.9837
	0	31.33/.9577	35.01/.9769	35.39/.9804	35.28/.9832	36.15/.9842	35.80/.9843	36.36/.9840	36.09/.9849	36.57/.9853	36.59/.9855	36.59/.9859	36.67/.9853
	1	31.53/.9573	35.04/.9762	34.82/.9765	34.83/.9768	34.56/.9772	34.73/.9772	35.44/.9793	34.93/.9773	35.30/.9782	35.21/.9784	35.25/.9795	36.80/.9840
	2	31.55/.9567	35.29/.9765	34.16/.9721	33.75/.9698	34.43/.9735	33.99/.9717	34.49/.9737	33.75/.9706	34.07/.9720	34.46/.9740	34.08/.9726	36.75/.9832
	3	31.56/.9565	35.41/.9768	34.10/.9710	33.65/.9689	34.39/.9725	33.94/.9711	34.54/.9732	33.61/.9687	34.02/.9712	34.37/.9725	34.04/.9715	36.75/.9829
	4	31.58/.9565	35.43/.9769	34.18/.9715	33.80/.9696	34.40/.9724	34.01/.9713	34.63/.9736	33.72/.9695	34.03/.9713	34.36/.9723	34.02/.9715	36.59/.9821
STFgantry [53]	-4	29.83/.9479	35.69/.9833	33.73/.9739	32.48/.9677	34.19/.9776	32.92/.9715	34.70/.9792	32.98/.9702	33.87/.9751	34.11/.9775	33.58/.9751	39.33/.9947
	-3	29.80/.9479	35.79/.9832	33.78/.9740	32.59/.9688	34.44/.9781	33.12/.9723	34.78/.9794	33.25/.9714	33.92/.9750	34.34/.9778	33.89/.9755	39.68/.9950
	-2	29.82/.9484	35.65/.9831	33.83/.9769	32.59/.9716	33.70/.9789	32.56/.9734	34.26/.9808	33.39/.9754	34.31/.9793	33.84/.9792	34.05/.9821	39.43/.9950
	-1	29.72/.9490	35.44/.9830	35.56/.9860	34.37/.9837	35.89/.9881	34.09/.9831	36.46/.9890	34.89/.9860	36.53/.9895	36.34/.9895	36.65/.9903	39.65/.9952
	0	31.06/.9498	36.33/.9831	38.36/.9904	37.95/.9898	39.64/.9929	38.72/.9909	39.61/.9926	38.77/.9915	39.86/.9936	40.54/.9941	40.40/.9942	42.17/.9957
	1	29.72/.9490	34.87/.9830	34.97/.9862	34.67/.9846	34.64/.9890	34.10/.9851	35.60/.9902	34.96/.9862	35.78/.9893	35.66/.9906	35.15/.9901	38.81/.9949
	2	29.79/.9483	35.01/.9829	33.66/.9779	32.88/.9721	33.85/.9821	32.61/.9740	33.85/.9816	32.90/.9750	33.97/.9800	34.15/.9827	32.70/.9798	38.58/.9947
	3	29.77/.9477	35.20/.9831	33.45/.9731	32.50/.9676	33.96/.9779	32.90/.9715	34.18/.9787	32.41/.9683	33.53/.9743	33.94/.9777	33.02/.9741	38.53/.9949
	4	29.80/.9477	35.19/.9832	33.39/.9733	32.53/.9679	33.72/.9774	32.78/.9714	34.18/.9792	32.41/.9685	33.43/.9745	33.76/.9773	32.78/.9739	38.46/.9947

groundtruth high angular resolution LFs, and selected the corner 2×2 SAIs as inputs.

Our EPIT-ASR was initialized using the Xavier algorithm [17], and trained using the Adam method [31] with $\beta_1 = 0.9$, $\beta_2 = 0.999$. The initial learning rate was set to 2×10^{-4} and halved after every 15 epochs. The training was stopped after 80 epochs. During the training phase, we performed random horizontal flipping, vertical flipping, and 90-degree rotation to augment the training data.

C.3. Qualitative Results

Figure IV shows the quantitative and qualitative results achieved by different LF angular SR methods. It can be observed that the magnitude of errors for our EPIT-ASR is smaller than other methods, especially on the delicate texture areas (e.g., the letters in scene `Dishes`). As shown in the zoom-in regions, our method generates more faithful details with fewer artifacts.

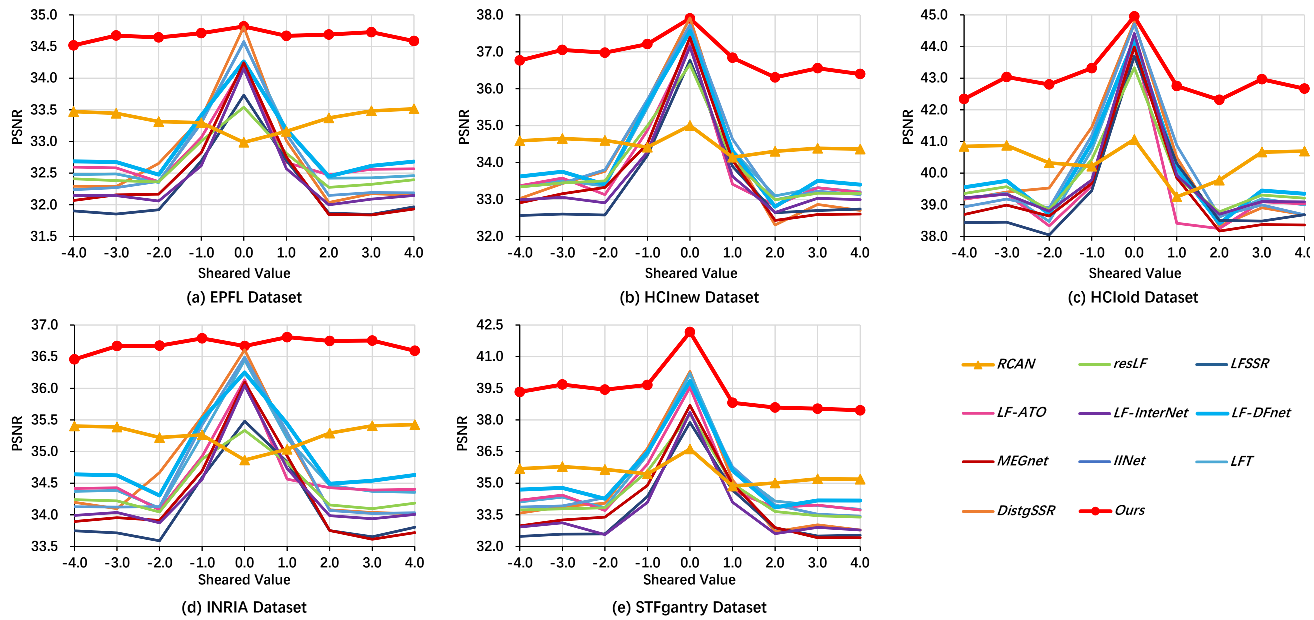

Figure II. Quantitative comparison of different SR methods on five datasets with different shearing values for 2× SR.

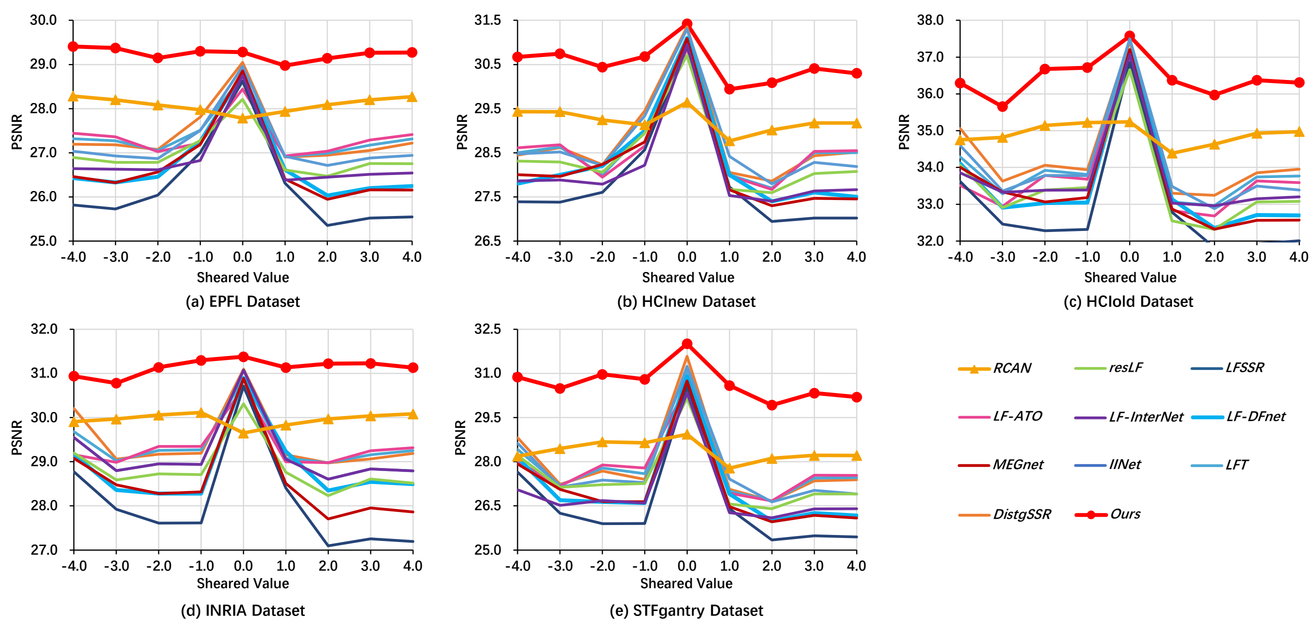

Figure III. Quantitative comparison of different SR methods on five datasets with different shearing values for 4× SR.

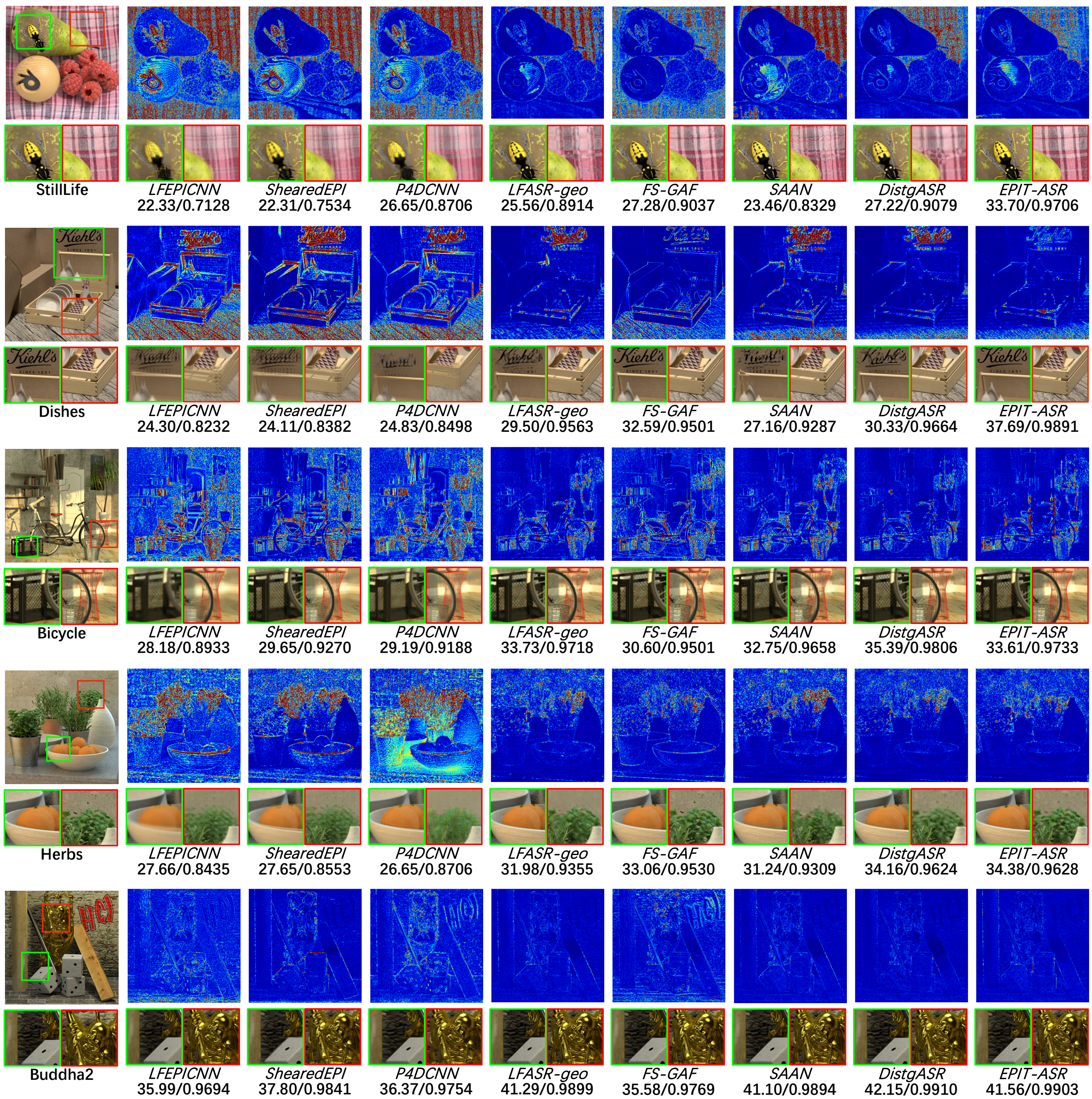

Figure IV. Visual results achieved by different methods on scenes `StillLife`, `Dishes`, `Bicycle`, `Herbs` and `Buddha2` for $2\times2 \rightarrow 7\times7$ angular SR. Here, we show the error maps of the reconstructed center view images, along with two zoom-in regions for qualitative comparison. The PSNR and SSIM values achieved on each scene are reported for quantitative comparison. Zoom in for the best view.