

Light Field Super-Resolution: A Benchmark

Supplementary Material

Zhen Cheng Zhiwei Xiong Chang Chen Dong Liu
University of Science and Technology of China

This supplementary document is organized as follows:

Sec. 1 illustrates the light field datasets we use for the benchmark evaluation, including thumbnails of the central views within HCI and EPFL datasets. Besides, the EPFL dataset suffers from the vignetting effect even after the camera rectification. So we make a further rectification.

Sec. 2 gives a brief introduction on all the evaluated light field SR methods.

Sec. 3 shows additional experimental results. Two evaluation metrics, i.e., SSIM [7] for reconstruction accuracy and Ma's metric [4] for perceptual quality together with more visual comparisons are provided.

1. Dataset Details

Fig. 1 shows the thumbnails of the central view images within each dataset. For each dataset, the thumbnails are arranged in the way that each column is a subgroup for training and testing LFCNN [9] with the K-fold cross validation strategy.

Additionally, to illustrate the vignetting effect of the EPFL dataset, we take one scene as an example and place the whole light field together to form a 15×15 image array. As shown in Fig. 2(a), the calibrated light field still suffers from vignetting especially at the side views. To avoid the influence of vignetting, we only use the central 9×9 views (marked with a red rectangle) from the original light field for evaluation.

However, even the intensities of the central 9×9 views are not consistent. So we further conduct a simple intensity rectification procedure. Specifically, for each sub-aperture image I_{sub} , we match its average intensity to that of the central view image $I_{central}$. The rectified sub-aperture image is obtained by

$$I'_{sub} = I_{sub} \times \frac{\text{mean}[I_{central}]}{\text{mean}[I_{sub}]} \quad (1)$$

The average intensity of each view with respect to angular coordinates before and after our rectification are shown in Fig. 2(b).

2. Framework of the Evaluated Methods

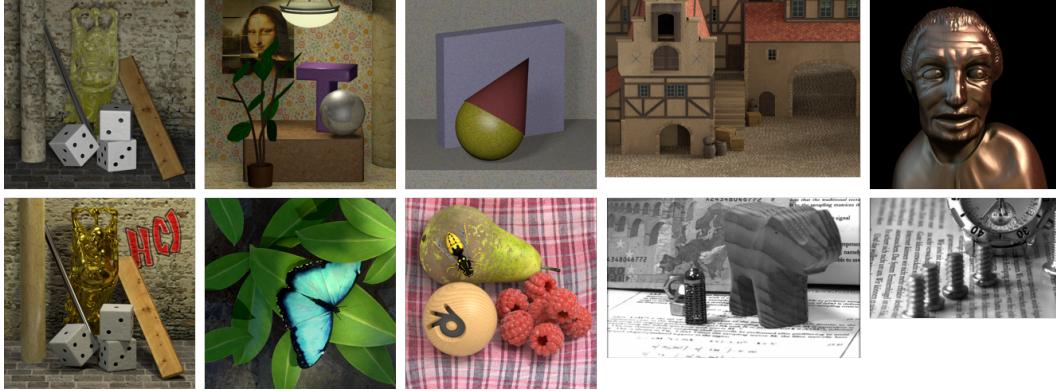
For a quick and easy understanding, Figs. 3, 4, 5 and 6 describe the main frameworks along with brief introductions for our evaluated light field SR methods PRO [3], GB [6], RR [1] and LFCNN [9], respectively. For more details, please refer to the original papers listed in references.

3. Additional Experimental Results

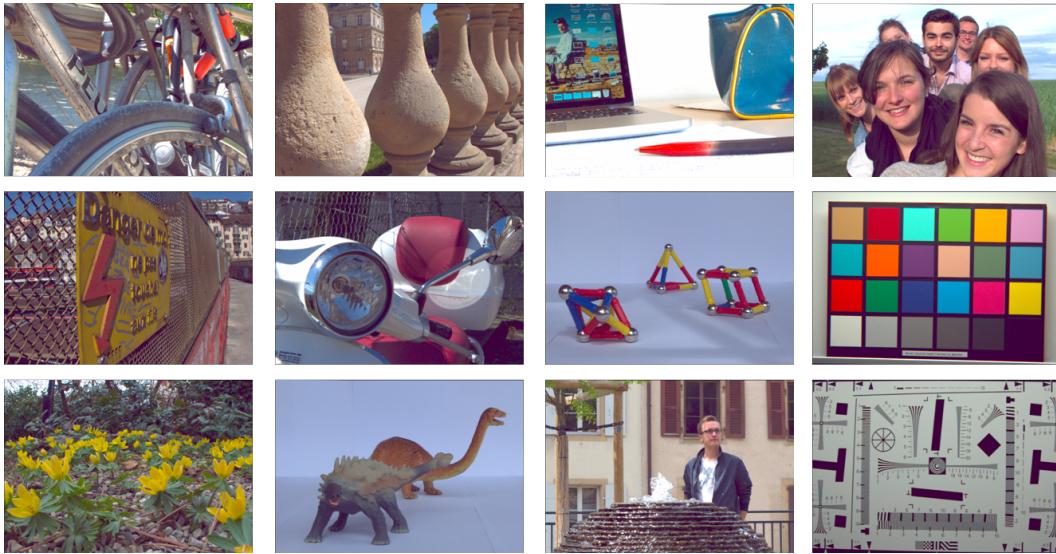
In addition to the PSNR and VGG metrics discussed in the paper, we further use SSIM [7] and Ma's metric [4] for evaluation. The average values of SSIM and Ma's metric of the super-resolved central views are shown in Fig. 7 and Fig. 8, respectively. The mean and standard deviation values in terms of these two metrics over all sub-aperture images are listed in Table 1. We also show more visual results in Fig. 9. As we can see, all these quantitative and qualitative comparison results are consistent with those we analyzed in the paper.

References

- [1] R. A. Farrugia, C. Galea, and C. Guillemot. Super resolution of light field images using linear subspace projection of patch-volumes. *IEEE Journal of Selected Topics in Signal Processing*, 11(7):1058–1071, 2017.



(a) HCI dataset [8]



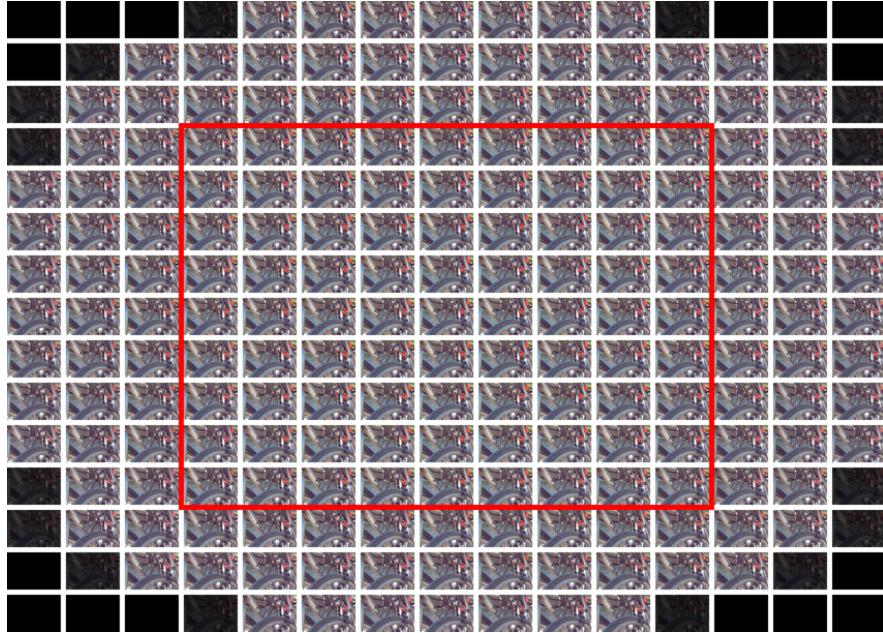
(b) EPFL dataset [5]

Figure 1. Thumbnails of the central view images within each dataset.

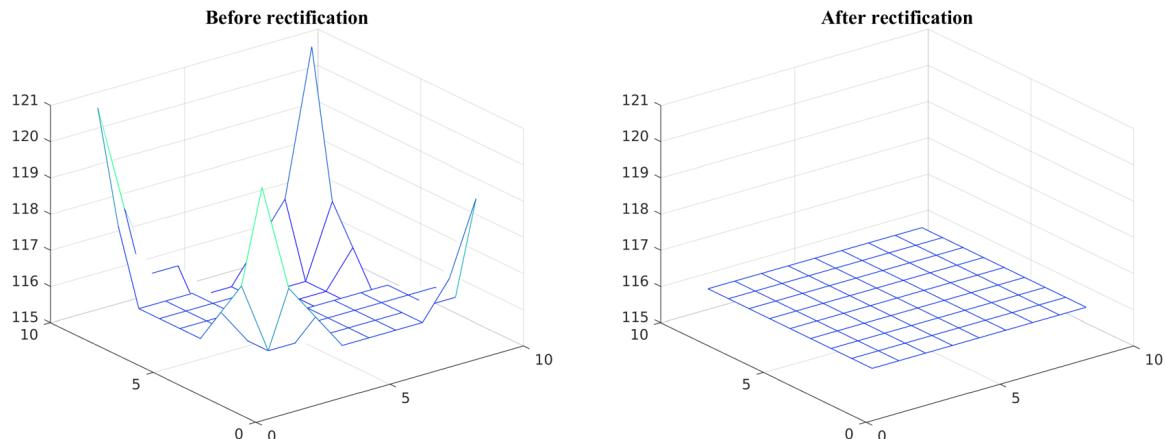
Method	HCI				EPFL			
	Bicubic $\times 2$	Gaussian $\times 2$	Bicubic $\times 3$	Gaussian $\times 3$	Bicubic $\times 2$	Gaussian $\times 2$	Bicubic $\times 3$	Gaussian $\times 3$
BIC	0.9301 ± 0.0007	0.9009 ± 0.0007	0.8641 ± 0.0008	0.8697 ± 0.0009	0.9189 ± 0.0036	0.8916 ± 0.0040	0.8518 ± 0.0054	0.8571 ± 0.0053
	5.380 ± 0.078	4.507 ± 0.057	4.117 ± 0.047	4.391 ± 0.064	4.820 ± 0.088	4.184 ± 0.086	4.105 ± 0.077	4.319 ± 0.082
GB [6]	0.9410 ± 0.0026	0.9124 ± 0.0020	0.8871 ± 0.0014	0.8955 ± 0.0019	0.9324 ± 0.0051	0.9066 ± 0.0050	0.8752 ± 0.0058	0.8859 ± 0.0059
	5.971 ± 0.116	4.802 ± 0.100	4.744 ± 0.095	5.303 ± 0.137	5.613 ± 0.181	4.232 ± 0.106	4.263 ± 0.085	4.864 ± 0.126
RR [1]	0.9419 ± 0.0017	0.9170 ± 0.0017	0.8817 ± 0.0015	0.8822 ± 0.0015	0.9304 ± 0.0033	0.9099 ± 0.0037	0.8742 ± 0.0045	0.8759 ± 0.0046
	6.176 ± 0.110	4.956 ± 0.095	4.822 ± 0.080	5.315 ± 0.113	5.777 ± 0.153	4.475 ± 0.101	4.495 ± 0.086	5.005 ± 0.114
LFCNN [9]	0.9439 ± 0.0011	0.9375 ± 0.0010	0.8762 ± 0.0016	0.8770 ± 0.0023	0.9384 ± 0.0029	0.9362 ± 0.0029	0.8834 ± 0.0040	0.8819 ± 0.0047
	7.024 ± 0.076	7.078 ± 0.065	6.320 ± 0.103	6.326 ± 0.109	6.871 ± 0.127	6.835 ± 0.118	5.944 ± 0.117	6.032 ± 0.118

Table 1. Mean and standard deviation values of SSIM and Ma’s metric (displayed in gray) on all sub-aperture images.

- [2] J. Kim, J. Kwon Lee, and K. Mu Lee. Accurate image super-resolution using very deep convolutional networks. In *CVPR*, 2016.
- [3] C.-K. Liang and R. Ramamoorthi. A light transport framework for lenslet light field cameras. *ACM Transactions on Graphics*, 34(2):16:1–16:19, 2015.
- [4] C. Ma, C.-Y. Yang, X. Yang, and M.-H. Yang. Learning a no-reference quality metric for single-image super-resolution. *Computer Vision and Image Understanding*, 158:1–16, 2017.
- [5] M. Rerabek and T. Ebrahimi. New light field image dataset. In *International Conference on Quality of Multimedia Experience (QoMEX)*, 2016.
- [6] M. Rossi and P. Frossard. Graph-based light field super-resolution. In *MMSP*, 2017.
- [7] 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*, 13(4):600–612, 2004.



(a) An example of the vignetting effect on the original light field.



(b) Average intensity w.r.t angular coordinates before and after our rectification.

Figure 2. Illustration of the vignetting effect and our rectification.

- [8] S. Wanner, S. Meister, and B. Goldlücke. Datasets and benchmarks for densely sampled 4d light fields. In *International Symposium on Vision Modeling and Visualization*, 2013.
- [9] Y. Yoon, H. G. Jeon, D. Yoo, J. Y. Lee, and I. S. Kweon. Learning a deep convolutional network for light-field image super-resolution. In *ICCVW*, 2015.

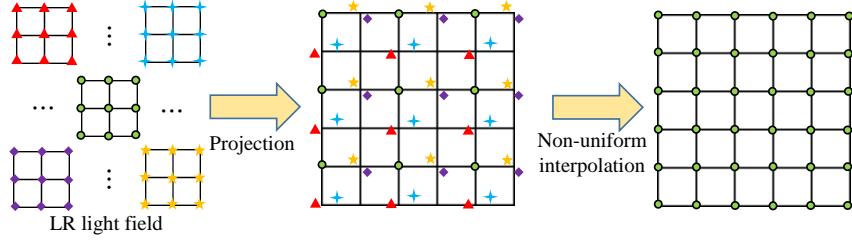


Figure 3. Framework for PRO [3]. The input light field is projected to the central view with the guidance of scene depth. This step leads to an image with a much denser sampling. Then a non-uniform interpolation is used to get an HR result.

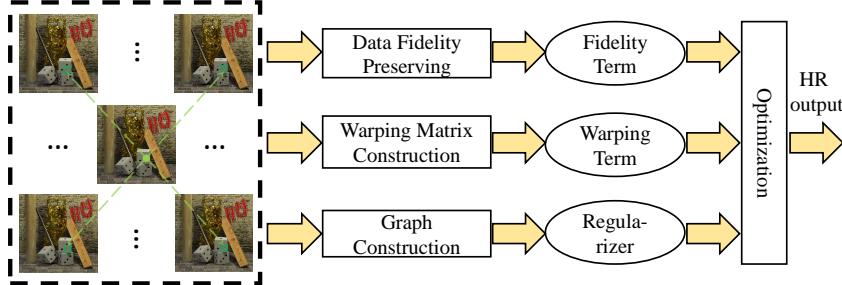


Figure 4. Framework for GB [6]. The fidelity term enforces the consistency between HR and LR light fields while the warping matrix enhances a certain LR view with other views by a rough warping. A graph-based regularizer is further used to constrain the solution space. These terms form a quadratic objective function and can be solved with the proximal point algorithm.

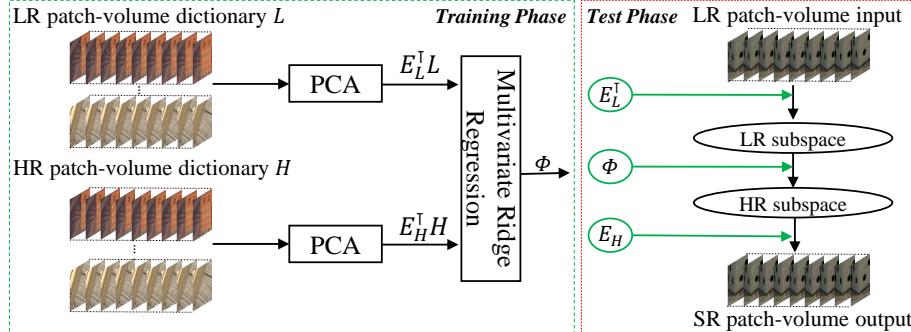


Figure 5. Framework for RR [1]. PCA is used to reduce the dimensions of patch-volumes and Multivariable Ridge Regression is used to linearly project the LR subspace to the HR subspace. In addition, a block matching based algorithm is proposed to align the patches before generating a patch-volume.

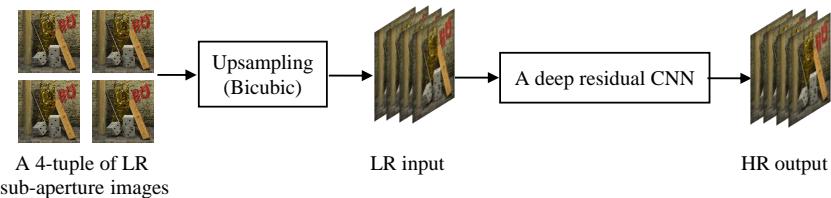


Figure 6. Framework for the spatial network of LFCNN [9]. A 4-tuple of LR sub-aperture images are first upsampled by Bicubic interpolation and then fed into a restoration CNN (here we replace the original shallow network with a deep residual CNN VDSR [2]). Four super-resolved views are the output.

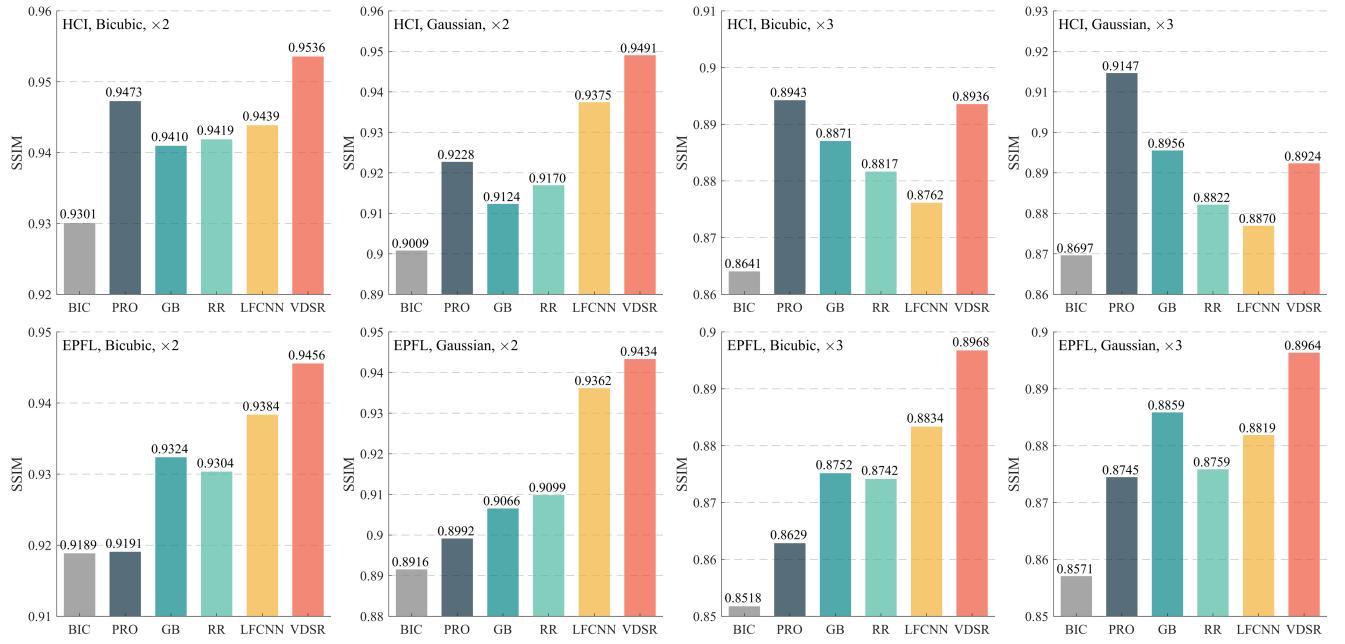


Figure 7. The average SSIM values [7] (the higher, the better) of the super-resolved central view images for six selected methods over two datasets and under four degradation models.

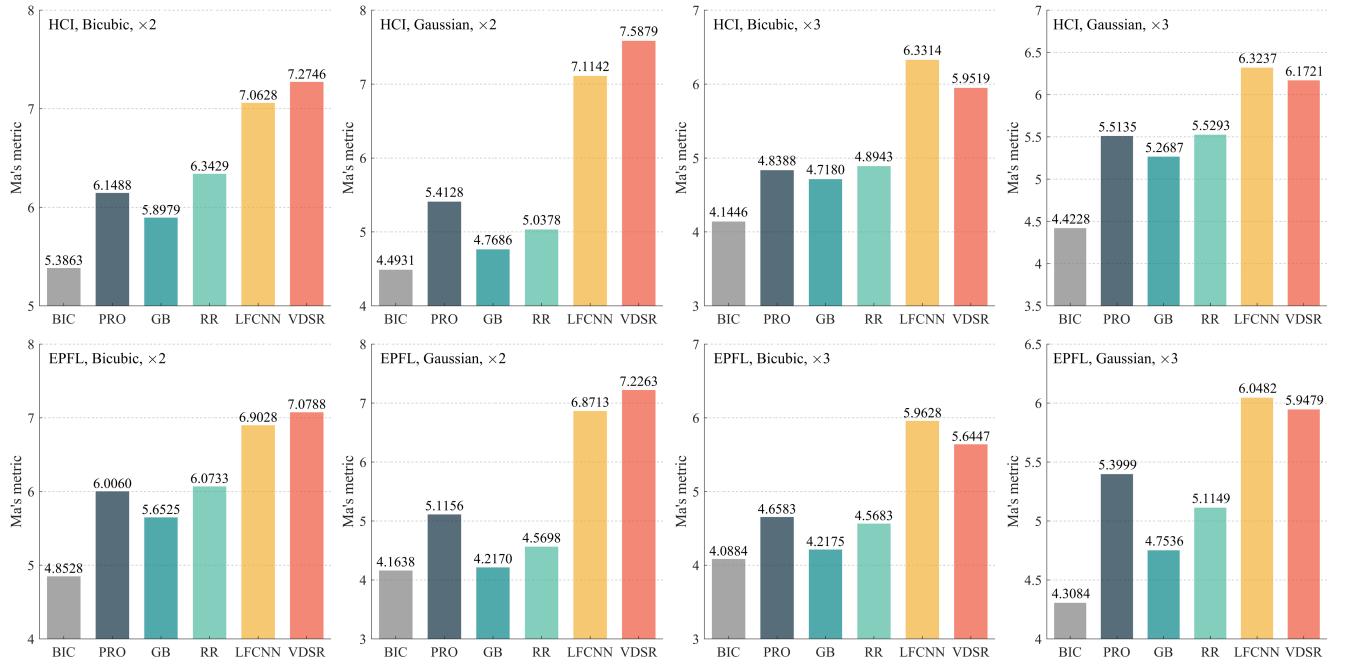


Figure 8. The average values of Ma's metric [4] (the higher, the better) of the super-resolved central view images for six selected methods over two datasets and under four degradation models.

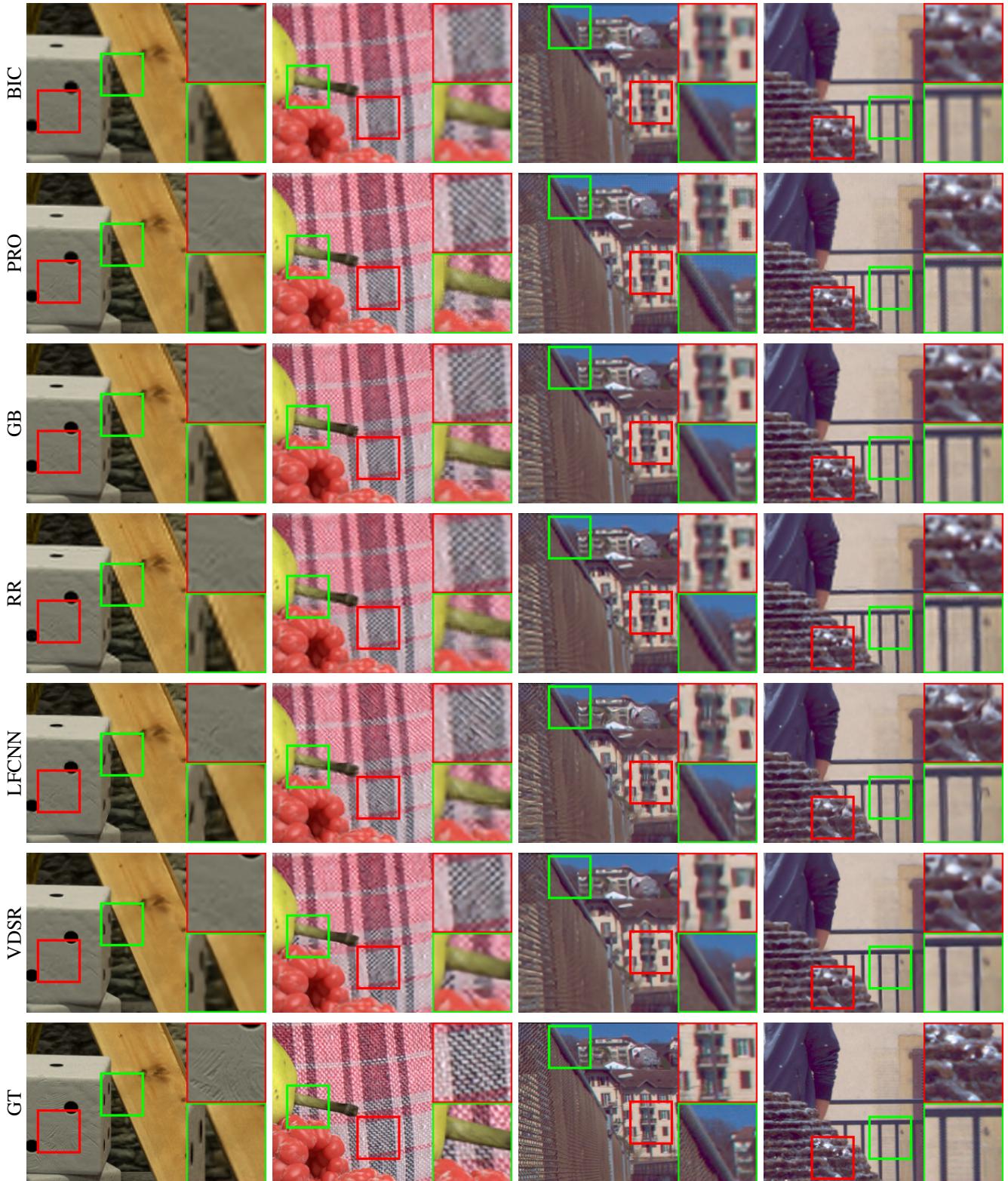


Figure 9. Visual comparisons of super-resolved central view images (cropped for a better visualization) through different methods together with the ground truth (GT) under Gaussian downsampling with the scale factor of 3. The first two scenes are from the HCI synthetic dataset and the last two scenes are from the EPFL real-world dataset.