Xinya WANG, Jiayi MA, Wenjing GAO, Junjun JIANG, 2021. MPIN: a macro-pixel integration network for light field super-resolution. *Frontiers of Information Technology & Electronic Engineering*, 22(10):1299-1310.
https://doi.org/10.1631/FITEE.2000566

MPIN: a macro-pixel integration network for light field super-resolution

Key words: Light field; Super-resolution; Macro-pixel representation

Corresponding author: Jiayi MA
E-mail: jiyima2010@gmail.com
ORCID: https://orcid.org/0000-0003-3264-3265

Motivation

- Offering rich view descriptions of the scene, light field (LF) images have relatively low spatial resolution due to the trade-off between spatial and angular resolution, which limits the range of potential development.
- Traditional LF super-resolution (SR) methods rely heavily on disparity estimation, any defect of which may result in significant artifacts.
- Deep learning based LF SR methods fail in fully exploiting both angular and spatial information. Consequently, their performances vary greatly in different sub-aperture images.

Method

- To fully exploit both spatial and angular information, we merge these two kinds of information into the two-dimensional (2D) space to solve the LF SR problem in the macro-pixel representation, as shown in Fig. 1

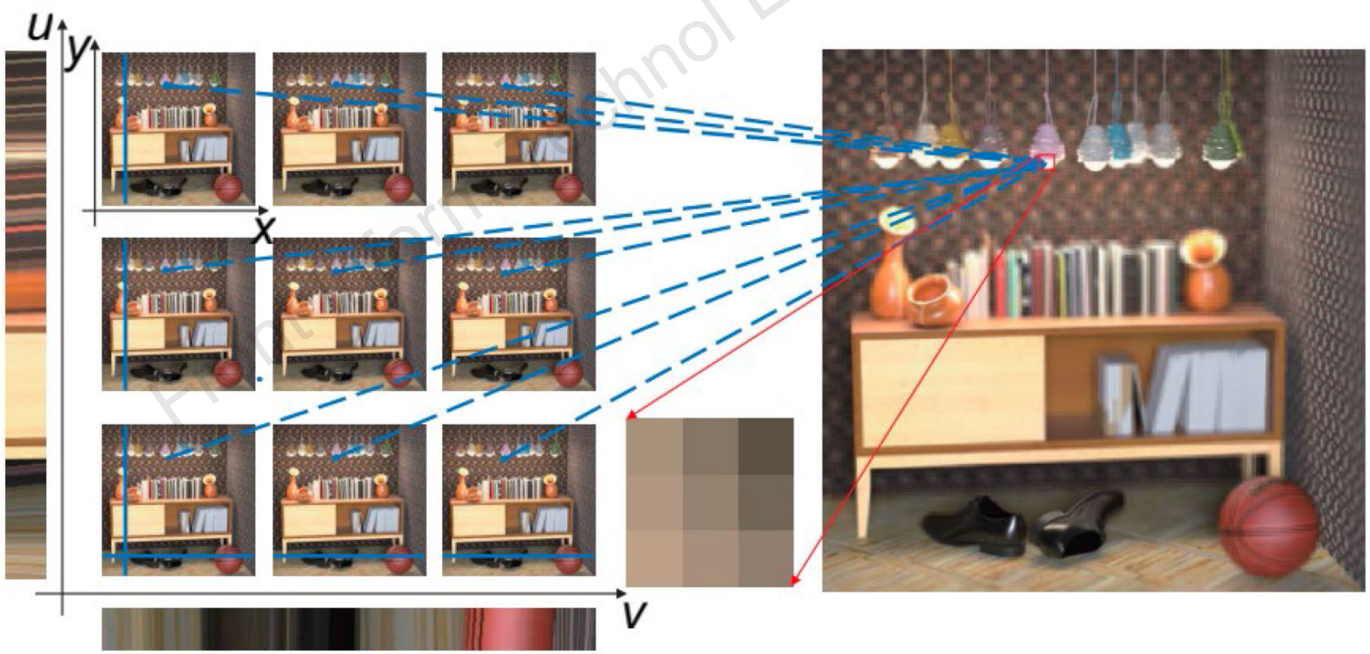

Fig. 1 Transformation from an array of sub-aperture images (SAIs) (left) to a macro-pixel image (MPI) (right)

Method (Cont'd)

- To mutually exploit spatial-angular correlations, we propose an integration resblock for mutual guidance
- To enhance the spatial resolution of MPI representation, an angular shuffle layer is designed for macro-pixel features, which can effectively avoid aliasing, as shown in Fig. 2

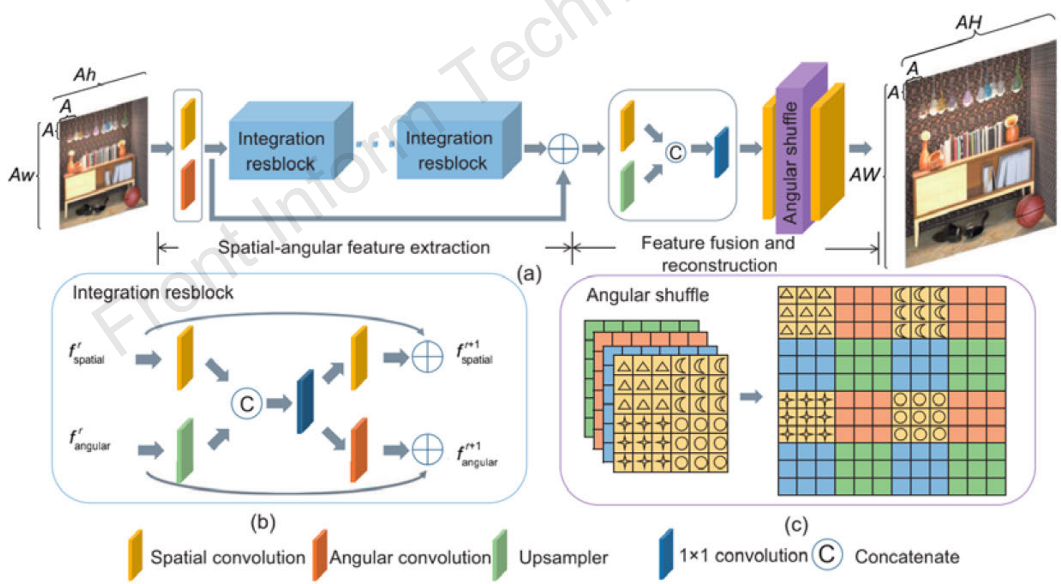

Fig. 2 Illustration of the proposed method: (a) the whole network of the proposed method with input size $Ah \times Aw$ and output size $AH \times AW$, where $H = s \cdot h$, $W = s \cdot w$ at upscale factor s; (b) the integration resblock for extracting spatial-angular features progressively; (c) a toy example for the angular shuffle operation with spatial resolution (of SAIs) of 2×2, angular resolution of 3×3, and upscale factor of 2

Major results

- Extensive experiments on both synthetic and real-world LF datasets demonstrate that our method can achieve better performance than the state-of-the-art methods

Table 2 Comparison with state-of-the-art methods on the test dataset (average PSNR/SSIM for scale factors 2 and 4)

Method	Scale	PSNR (dB) / SSIM		PSNR (dB) / SSIM	
		HCI1	HCI2	EPFL	StaLytro
Bicubic	×2	37.43 / 0.9497	32.89 / 0.8903	31.53 / 0.9203	33.39 / 0.9304
RR	×2	38.13 / 0.9555	33.40 / 0.8979	32.34 / 0.9358	33.98 / 0.9456
GB	×2	39.04 / 0.9634	34.96 / 0.9278	32.98 / 0.9667	35.21 / 0.9624
LFNet	×2	38.57 / 0.9811	33.73 / 0.9544	32.74 / 0.9683	35.08 / 0.9731
LFSAS	×2	39.35 / 0.9606	34.57 / 0.9074	33.44 / 0.9376	36.87 / 0.9609
EDSR	×2	40.47 / 0.9864	35.48 / 0.9656	36.37 / 0.9762	38.81 / 0.9867
resLF	×2	40.50 / 0.9852	36.38 / 0.9764	36.04 / 0.9737	38.74 / 0.9857
MPIN (ours)	×2	**41.28 / 0.9888**	**36.89 / 0.9776**	**37.21 / 0.9779**	**39.85 / 0.9892**
Bicubic	×4	32.40 / 0.8499	28.82 / 0.7599	27.61 / 0.7973	27.99 / 0.7944
RR	×4	33.24 / 0.8702	29.40 / 0.7788	27.89 / 0.8032	28.12 / 0.8026
GB	×4	33.37 / 0.8741	29.75 / 0.7939	28.17 / 0.8186	28.69 / 0.8172
LFNet	×4	33.14 / 0.9396	29.30 / 0.8836	28.24 / 0.9017	28.75 / 0.8961
LFSAS	×4	33.87 / 0.8741	29.95 / 0.7841	28.84 / 0.8307	30.14 / 0.8508
EDSR	×4	34.55 / 0.9471	30.33 / 0.9002	29.59 / 0.9234	30.98 / 0.9316
resLF	×4	34.93 / 0.9506	30.65 / 0.9134	30.51 / 0.9224	31.09 / 0.9304
MPIN (ours)	×4	**35.56 / 0.9604**	**31.27 / 0.9208**	**30.84 / 0.9259**	**32.22 / 0.9452**

PSNR: peak signal-to-noise ratio; SSIM: structural similarity index. HCI1 and HCI2 are synthetic datasets; EPFL and StaLytro are real-world datasets. The best results are in bold

Major results (Cont'd)

- As shown in the figure, our method can recover more texture or sharper details compared to others. Specifically, our method behaves well even in the occlusion area or reflective surface

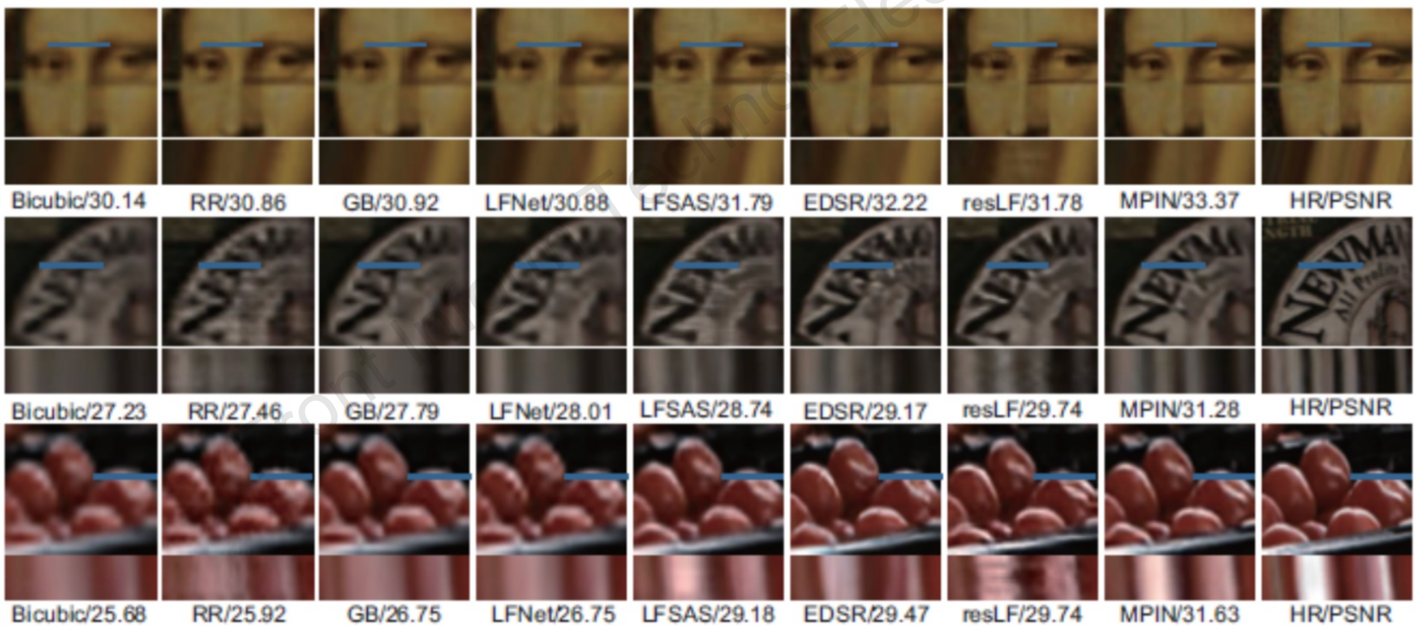

Fig. 4 Detailed ×4 super-resolution (SR) results for image Mona from HCI1, and General37 and Flowers_plants6 from StaLytro. The super-resolved central view images are visualized, and horizontal EPIs are extracted along the blue lines, where the PSNR values of central views are illustrated below. The last column shows the ground truth

Major results (Cont'd)

- From different angular views, the following figure further demonstrates the reconstruction performance for each angular position. Our proposed MPIN achieves better performance with a relatively balanced distribution among different views

Fig. 5 Average PSNR values of SAIs reconstructed by resLF (left) and MPIN (right) at scale 4. The results are calculated on the StaLytro dataset with angular resolution of 7×7

Conclusions

- In our method, we transform the 4D LF data into a 2D MPI representation that couples spatial and angular information.
- To fully extract two kinds of information, we propose the integration resblock to explore intra-view spatial correlations and inter-view angular correlations, which allows our method to capture features that are more consistent with the LF image.
- To enlarge the MPI features, we have designed an angular shuffle layer to enhance the spatial resolution of the MPI.
- The qualitative and quantitative results on publicly LF datasets have demonstrated the superiority of our method over the state-of-the-art methods at different scale factors.

References

[1] Zhang S, Lin YF, Sheng H, 2019. Residual networks for light field image super-resolution. Proc IEEE/CVF Conf on Computer Vision and Pattern Recognition, p.11038-11047. https://doi.org/10.1109/CVPR.2019.01130

[2] Yuan Y, Cao ZQ, Su LJ, 2018. Light-field image superresolution using a combined deep CNN based on EPI. *IEEE Signal Process Lett*, 25(9):1359-1363. https://doi.org/10.1109/LSP.2018.2856619

[3] Yeung HWF, Hou JH, Chen XM, et al., 2019. Light field spatial super-resolution using deep efficient spatial-angular separable convolution. *IEEE Trans Image Process*, 28(5):2319-2330. https://doi.org/10.1109/TIP.2018.2885236

[4] Wang YQ, Wang LG, Yang JG, et al., 2020. Spatial-angular interaction for light field image super-resolution. Proc 16th European Conf on Computer Vision, p.290-308. https://doi.org/10.1007/978-3-030-58592-1_18

[5] Wang YL, Liu F, Zhang KB, et al., 2018. LFNet: a novel bidirectional recurrent convolutional neural network for light-field image super-resolution. *IEEE Trans Image Process*, 27(9):4274-4286. https://doi.org/10.1109/TIP.2018.2834819

Xinya WANG, first author of this invited paper, received her BS degree from the Electronic Information School, Wuhan University, Wuhan, China, in 2018. She is currently pursuing her PhD degree with the Electronic Information School, Wuhan University. Her research interests include neural networks, machine learning, and image processing.

Jiayi MA, corresponding author of this invited paper, received his BS degree in information and computing science and PhD degree in control science and engineering from Huazhong University of Science and Technology, Wuhan, China, in 2008 and 2014, respectively. He is currently a professor with the Electronic Information School, Wuhan University. He has authored or co-authored more than 200 refereed journal and conference papers. His research interests include computer vision, machine learning, and robotics. Dr. Ma has been identified in the 2020 and 2019 Highly Cited Researcher lists from Web of Science. He is an area editor of *Information Fusion*, an associate editor of *Neurocomputing*, *Sensors*, and *Entropy*, and a guest editor of *Remote Sensing*.