A Deep Learning Framework for Light Field Reconstruction From Focus-Defocus Pair: A minimal hardware approach

Anil Kumar Vadathya, Sharath Girish and Kaushik Mitra, Indian Institute of Technology Madras, INDIA

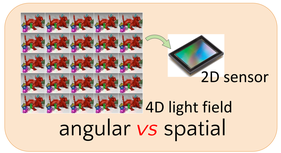
IEEE 2018 Conference on Computer Vision and Pattern Recognition
CVPR 2018 Salt Lake City

Resolution Trade-off with Light Field Imaging

- Due to bandwidth limitation, capturing 4D light field on a 2D sensor poses resolution trade-off

Solution: Multiplex or sub-sample angular rays and reconstruct back computationally

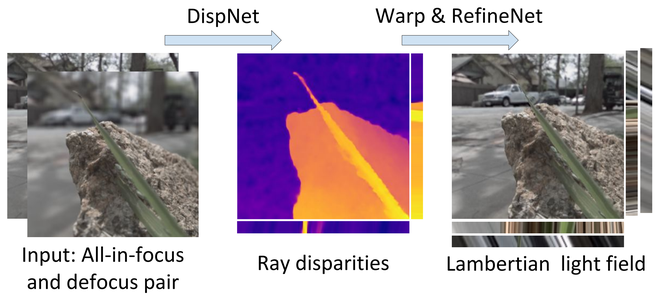
- Compressive light field, programmable aperture, view synthesis
- This often requires specific hardware requiremens



Our solution: Light field from focus-defocus pair
No additional hardware

Light Field Reconstruction Pipeline

- We breakdown the LF reconstruction from focus-defocus pair into two steps -



Lambertian Light Field Synthesis

$$\tilde{L}(x,q) = L(x + qD(x,q), 0)$$

Disparity EPI consistency

$$D(x,u) = D(x + kD(x,u), u + k)$$

Non-Lambertian Light Field Refinement

$$r = f(\tilde{L}, D, I_d)$$

$$L_r(x,u) = r(x,u) + \tilde{L}(x,u)$$

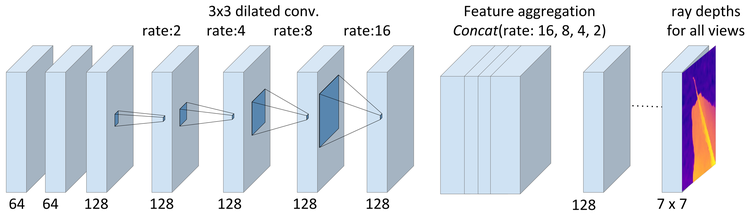
Overall Loss function

$$\mathcal{L}_{rec} + \mathcal{L}_{ref} + \lambda_{dc}\mathcal{L}_{dc} + \lambda_{tv}\mathcal{L}_{tv}$$

*This block is adapted from single image LF synthesis work [2]

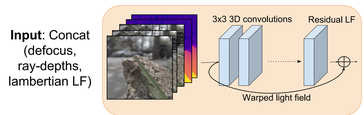
Proposed CNN Architectures

DispNet: Disparity from focus-defocus pair



RefineNet: non-lambertian LF

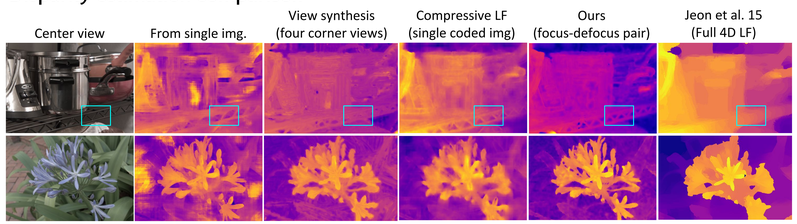
- Predicts non-lambertian LF and occluded pixels using a residual block [2]



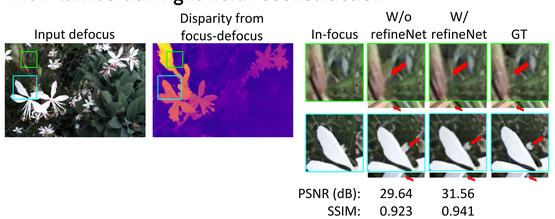
Challenges with disparity estimation

- Depth disambiguation from just focus-defocus pair
  - Variation of focal plane
- Textureless regions
  - Does not affect the lambertian LF reconstruction

Disparity estimation comparison

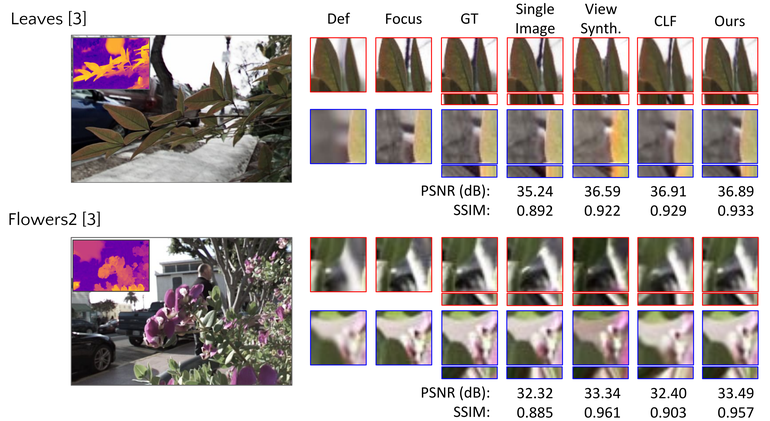


Non-lambertian light field reconstruction



Results

Approach	30 Test LFs [3]	Reflective_29	Occlusions (7, 25, 47)	Miscellaneous (1,2,21,22,29,43,56)
Single Image [2]	35.55 dB, 0.930	35.45 dB, 0.942	33.08dB, 0.855	35.03 dB, 0.927
View Synthesis [3]	**37.56 dB, 0.960**	37.47 dB, 0.974	34.74 dB, 0.924	35.85 dB, 0.951
Compressive LF [1]	37.29 dB, 0.951	37.67 dB, 0.972	33.95 dB, 0.911	36.22 dB, 0.944
Ours (Foc-Defocus)	37.41 dB, 0.955	**39.56 dB, 0.981**	**35.39 dB, 0.936**	**37.12 dB, 0.960**
Direct regression				



References

1. Vadathya, Anil Kumar, et al. "Learning light field reconstruction from a single coded image." In ACPR 2017.
2. Srinivasan, Pratul P., et al. "Learning to synthesize a 4d rgbd light field from a single image." In ICCV 2017.
3. Kalantari, Nima. K., et al. "Learning-based view synthesis for light field cameras". In SIGGRAPH ASIA, 2016.