Learning Light Field Reconstruction from a Single Coded Image

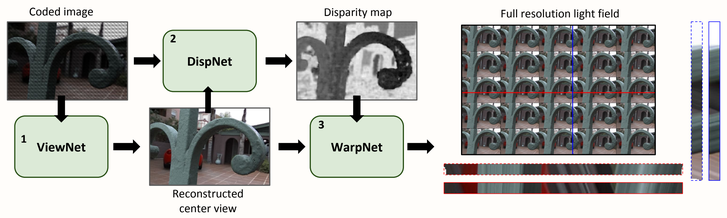
[1] Anil Kumar Vadathya, [1] Saikiran Cholleti, [2] Gautham Ramajayam, [3] Vijayalakshmi Kanchana, [1] Kaushik Mitra

[1]Computational Imaging Lab, IIT - Madras, India, [2]NIT Trichy, India, [3]SSN College of Engg., India

ACPR2017

4th Asian Conference on Pattern Recognition

Highlights

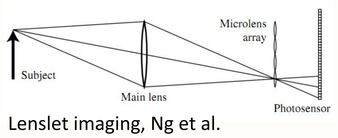


- We propose a deep learning based method for full sensor resolution light field reconstruction from a single coded image.
- Our approach consists of three convolutional neural networks (CNNs); *ViewNet*, *DispNet* and *WarpNet*.
- Our depth based view synthesis requires depth map estimation which we learn in an unsupervised manner.

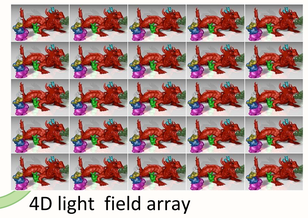
Background

Resolution trade-off

Sacrifice spatial resolution to **gain** angular resolution



Lenslet imaging, Ng et al.

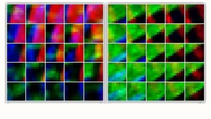


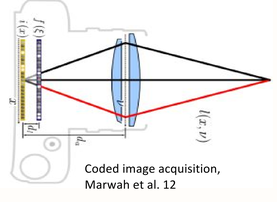
4D light field array

2D sensor

Compressive light field imaging

- Marwah et al. [1] proposed a compressive imaging set up
- They learn light field dictionaries exploiting the angular redundancy



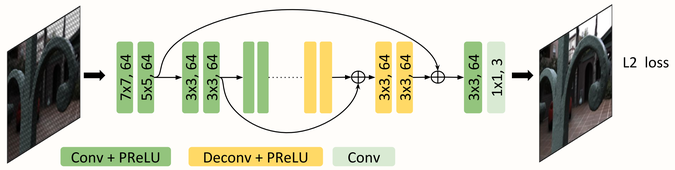


Coded image acquisition, Marwah et al. 12

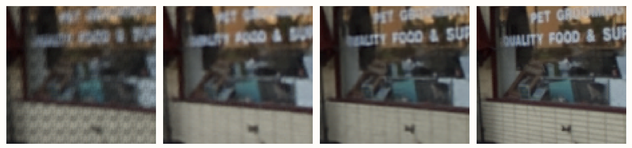
- Iterative reconstruction takes time

Proposed Architecture

ViewNet - scene reconstruction from coded image

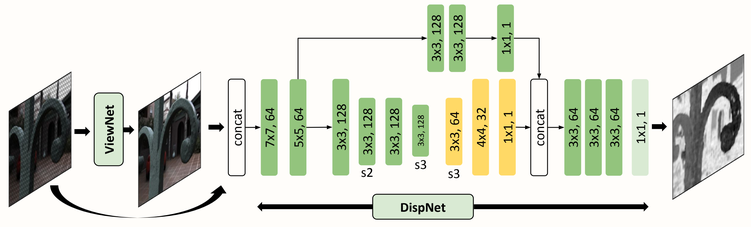


- Full convolutional network with symmetric skip connections inspired from Mao et al. 2016



Coded Image | w/o skip connectcions | with skip connectcions | GT centerview

DispNet - disparity estimation from coded image and scene



- Encoder-decoder with a skip connection
- *WarpNet*, Warps the center view, I_0, to novel view at q, I_q as,

$$\hat{\mathrm{I}}_\mathrm{q}(\mathrm{x} + \mathrm{qD}(\mathrm{x})) = \mathrm{I}_0(\mathrm{x}),$$

- Reconstruction loss back-propagates via *WarpNet* and *DispNet*

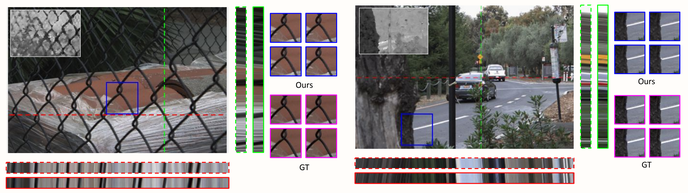
$$\mathrm{L} = \|\hat{\mathrm{I}}_\mathrm{q} - \mathrm{I}_\mathrm{q}\|_2$$

Our disparity estimation with **no explicit** regularization

Coded Image | DispNet output from coded image and GT center view | Wang et al. 2015 estimation using full light field

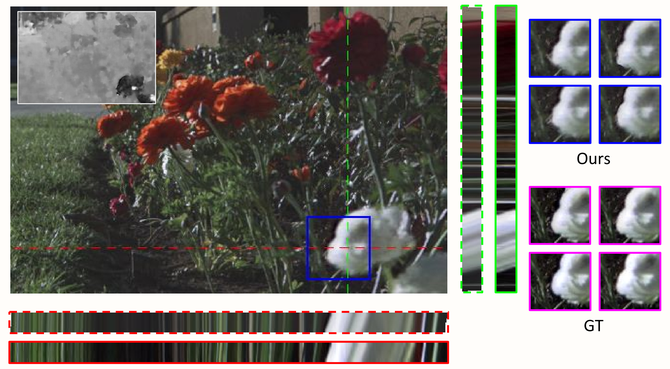
Results

Compressive light field reconstruction



Mesh scene - 31.00 dB | Cars scene - 32.30 dB

Figure 1: figure shows the reconstructed scene, disparity and EPIs. The generated EPI is shown with dashed and GT with solid line border.



Flowers scene - 32.18 dB

Comparison with dictionary learning

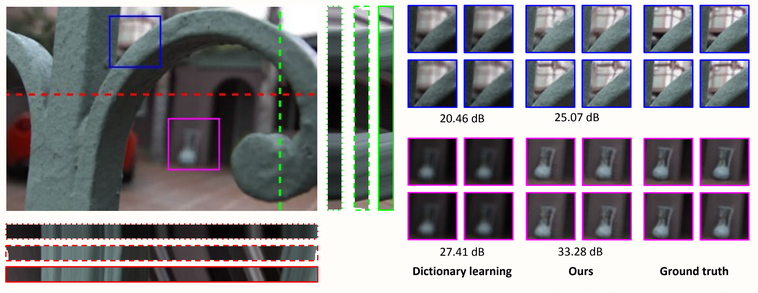


Figure 2: figure shows the center view reconstructed with dictionary learning method and EPIs. EPI images are shown with dotted line border for dictionary learning, dashed line for our method and solid line for ground truth. Our method recovers the parallax well. Also, as seen from patches we recover the details much sharper.

References

[1] Marwah, Kshitij, et al. "Compressive light field photography using overcomplete dictionaries and optimized projections." *ACM Transactions on Graphics* (TOG) 32.4 (2013): 46.