ECCV 2016, Amsterdam, the Netherlands

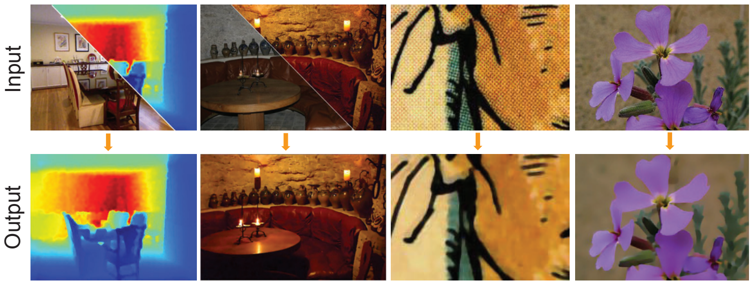
Deep Joint Image Filtering

Yijun Li[1], Jia-Bin Huang[2], Narendra Ahuja[2], Ming-Hsuan Yang[1]

[1]University of California, Merced [2]University of Illinois, Urbana-Champaign

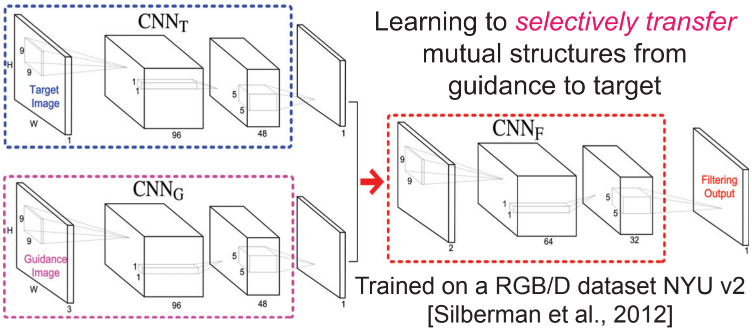
Project webpage: http://bit.ly/deepjointfiltering

Joint image filtering



Depth upsampling | Noise Reduction | Inverse Halftoning | Texture Removal

Network architecture



Learning to *selectively transfer* mutual structures from guidance to target

Trained on a RGB/D dataset NYU v2 [Silberman et al., 2012]

Limitations of existing methods

1. Fail to consider mutual structures.
2. Hand-crafted objective functions.
3. Inefficiency for optimization-based methods.

Contributions

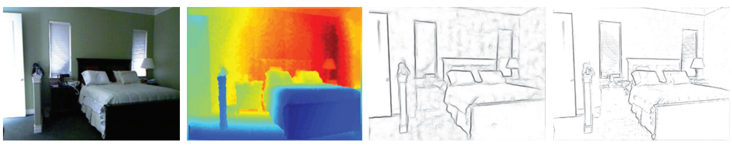
1. A learning-based approach for joint image filters.
2. State-of-the-art performance on depth upsampling.
3. A generic filter to handle image data in a variety of domains.

Design rationale

Ground truth | Bicubic | RGB guided | Edge guided | Ours

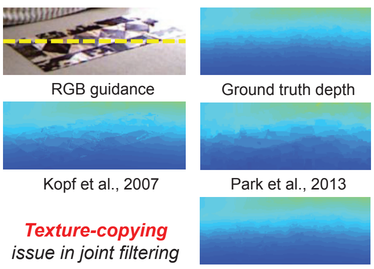
- Directly stack the RGB guidance image and target depth image, and feed them through a generic network → poor performance
- Replace the RGB guidance image with its edge map [Dollár et al., 2013] → good performance
- End-to-End: Extract structural features from both the target and guidance image. Then, combine them and reconstruct.

What has the network learned?



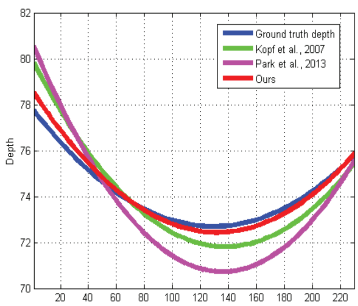
RGB guidance | Depth map | Edge map [Dollár et al., 2013] | Learned guidance from CNN_G

- The learned guidance appears like an edge map.
- CNN_T and CNN_G show strong responses to edges from the target and guidance image respectively.
- CNN_F re-organizes the extracted structural features and suppresses inconsistent details.



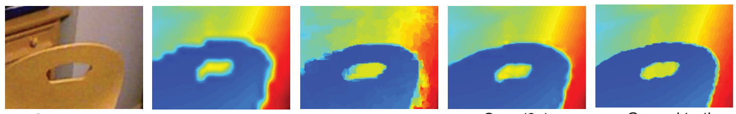
RGB guidance | Ground truth depth | Kopf et al., 2007 | Park et al., 2013 | Ours

Texture-copying *issue in joint filtering*



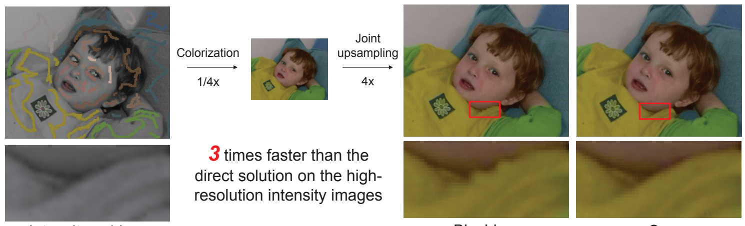
Experiments

1. Depth map upsampling



RGB guidance | Low-res depth map | Park et al. (8x) | Ours (8x) | Ground truth

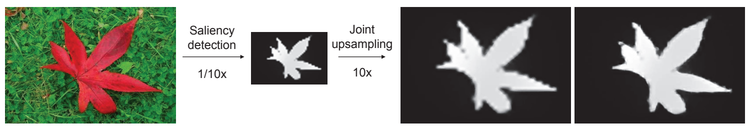
2. Chromaticity map upsampling



3 times faster than the direct solution on the high-resolution intensity images

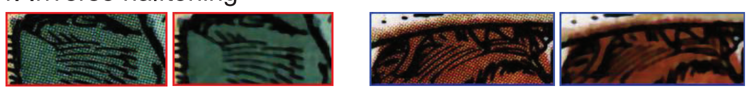
Intensity guidance | Bicubic | Ours

3. Saliency map upsampling



RGB guidance | GF [He et al., 2010] | Ours

4. Inverse halftoning



Input | Our result | Input | Our result

5. Cross-modal noise reduction

Non-Flash/Flash | Our denoised result | NIR/RGB | Our denoised result

RMSE comparisons for depth map upsampling.

Method	Middlebury (#30, [0,255])			NYU v2 (#449, cm)		
	4x	8x	16x	4x	8x	16x
He et al., 2010	4.01	7.22	11.70	7.32	13.62	22.03
Ferstl et al., 2013	3.39	5.41	12.03	6.98	11.23	28.13
Ham et al., 2015	3.14	5.03	8.83	5.27	12.31	19.24
Ours	**2.14**	**3.77**	**6.12**	**3.54**	**6.20**	**10.21**