Depth Estimation and Blur Removal from a Single Out-of-focus Image

Saeed Anwar*
Saeed.Anwar@anu.edu.au
Zeeshan Hayder*
Zeeshan.Hayder@anu.edu.au
Fatih Porikli
Fatih.Porikli@anu.edu.au

The Australian National University,
Canberra, Australia

1 Detailed experimental analysis and discussion

Here, We present more results and explanations. We begin by showing the detailed configuration of the proposed network and its superior performance on the depth estimation task both on synthetic and real image, which is followed by non-uniform deblurring.

Network Architecture

The proposed network configuration for depth estimation, using dense patch pooling, is based on VGG16 [9]. We densely sampled 15k patches for each image over a regular grid. The configuration of our network is shown in the Table 1.

Layer	**1-2**	**3-4**	**5-7**	**8-10**	**11-13**	**14-15**	**16**
Type	conv+relu	conv+relu	conv+relu	conv+relu	conv+relu	ip+relu	ip+interp
Filter Size	3×3	3×3	3×3	3×3	3×3	-	-
No. of Filter	64	128	256	512	512	4k	20
Pooling	max	max	max	max	**patch**	-	-

Table 1: Network configuration for our depth estimation, using dense patch pooling, is based on VGG16 [9]. For brevity, we use *conv* for convolutional layer, *relu* for activation function, *ip* for inner product and *interp* for bilinear upsampling.

Depth Estimation

We compared proposed method results with state-of-the-art depth prediction methods on standard benchmark datasets. The qualitative results are shown in Fig. 1. It is important to note that our method not only captured the depth accurately for the near objects in the scene but for far objects as well.

 This work was supported under the Australian Research Councils Discovery Projects funding scheme (project DP150104645) and an Australian Government RTP Scholarship.
*Both authors contributed equally to this work.

Removing Non-uniform Blur

We evaluate our blur removal method on test images from NYU-v2. Figure 2 shows the results generated by our and competing methods for different images. Our algorithm attains higher visual quality than its counterparts. It is observed that our algorithm is able to restore high-frequency texture details with a closer resemblance to the ground-truth than existing methods due to estimating blur kernels from depth layers.

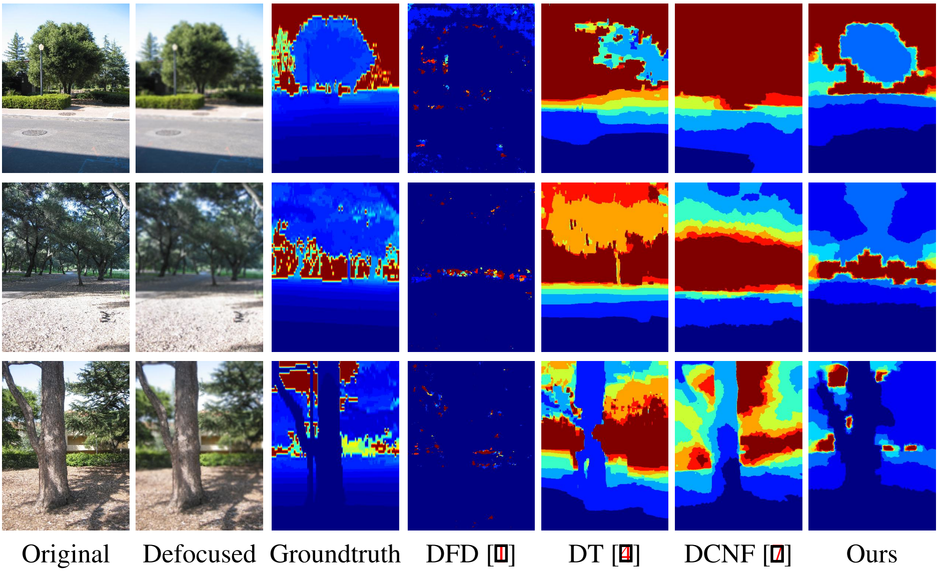


Figure 1: Qualitative comparison of our depth estimation method on Make3D [8] dataset with state-of-the-art depth prediction methods. Our method correctly predicted the depth levels. Depths are shown in color where red is far and blue is near.

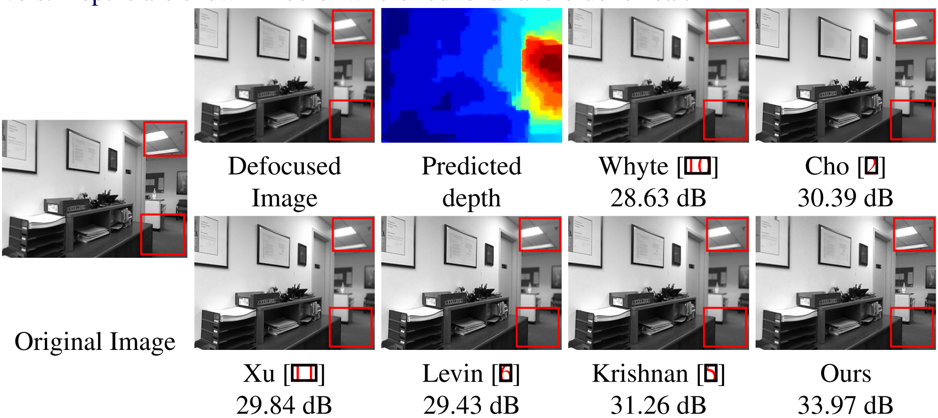
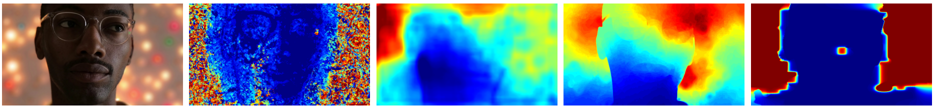


Figure 2: Qualitative comparison of our deblurring results on NYU-v2 [8] dataset with state-of-the-art deblurring methods. The difference can be seen in the red box and best viewed at higher magnification.

In Fig. 2, the highly-textured patterns on walls are adeptly reproduced by our algorithm, while these details are not clearly visible in the results of the other methods. In this example, most of the other methods tend to smoothen out the variation of the background texture along one of its principal directions. Furthermore, some methods introduce additional artifacts and



Real Image DFD [1] Eigen *et al.* [3] DCFN [7] Ours

Figure 3: Images are real defocused photos with unknown blur and qualitative comparison shows significant improvement over state-of-the-art depth prediction methods. Our method benefits from the amount of blur in the real images whereas other methods rely on the color and shape of the object which fails to recover the depth.

artificial textures.

2 Real Out-of-focus Images

In our last experiment, we evaluate the proposed method on the real-world out of focus face image. Comparison with state-of-the-art methods [4, 7, 8] are shown in figure 3.

In this example, there are approximately two layers of blur corresponding to different levels. Our method put the face and background in different depth levels by exploiting the blur while [7] generates sharp boundaries for face but puts different level of depth for the same face layer. Thus, making our method more useful in practical situation in the presence of non-uniform blur. As compared to others, our method outperformed.

References

[1] Ayan Chakrabarti and Todd Zickler. Depth and deblurring from a spectrally-varying depth-of-field. In *ECCV*. 2012.

[2] Sunghyun Cho and Seungyong Lee. Fast motion deblurring. SIGGRAPH Asia, 2009.

[3] David Eigen, Christian Puhrsch, and Rob Fergus. Depth map prediction from a single image using a multi-scale deep network. In *NIPS*, 2014.

[4] Kevin Karsch, Ce Liu, and Sing Bing Kang. Depth transfer: Depth extraction from video using non-parametric sampling. *TPAMI*, 2014.

[5] Dilip Krishnan, Terence Tay, and Rob Fergus. Blind deconvolution using a normalized sparsity measure. In *CVPR*, 2011.

[6] Anat Levin, Yair Weiss, Fredo Durand, and William T. Freeman. Understanding blind deconvolution algorithms. *TPAMI*, 2011.

[7] Fayao Liu, Chunhua Shen, and Guosheng Lin. Deep convolutional neural fields for depth estimation from a single image. In *CVPR*, 2015.

[8] A. Saxena, M. Sun, and A. Y. Ng. Make3d: Learning 3d scene structure from a single still image. *TPAMI*, 2009.

[9] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[10] Oliver Whyte, Josef Sivic, Andrew Zisserman, and Jean Ponce. Non-uniform deblurring for shaken images. *IJCV*, June 2012.

[11] Li Xu and Jiaya Jia. Two-phase kernel estimation for robust motion deblurring. In *ECCV*. 2010.