Supplemental Material: A Deep Primal-Dual Network for Guided Depth Super-Resolution

Gernot Riegler
riegler@icg.tugraz.at
David Ferstl
ferstl@icg.tugraz.at
Matthias Rüther
ruether@icg.tugraz.at
Horst Bischof
bischof@icg.tugraz.at

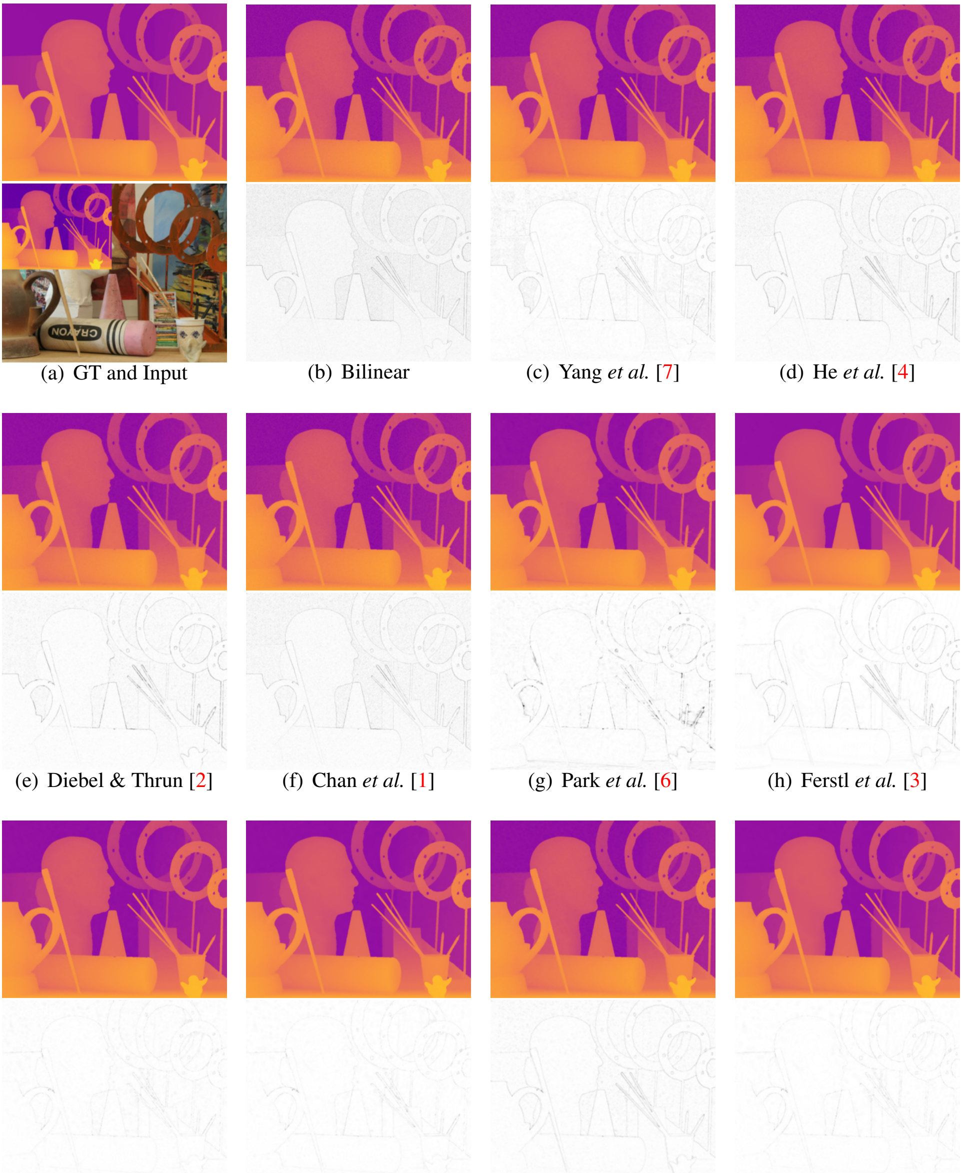
Institute for Computer Graphics and Vision
Graz University of Technology
Austria

1 Introduction

The supplementary material of our BMVC 2016 submission provides the qualitative results of our evaluations. In Section 2 we compare our results on the images *Art*, *Books*, and *Moebius* of the noisy Middlebury dataset as proposed by [6] to other state-of-the-art approaches. Namely, we show results of bilinear upsampling, Yang *et al.* [8], He *et al.* [4], Diebel & Thrun [2], Chan *et al.* [1], Park *et al.* [6], Ferstl *et al.* [3], and of our fully-convolutional network (FCN) only, as well as of our *deep primal-dual network* (FCN-PDN).

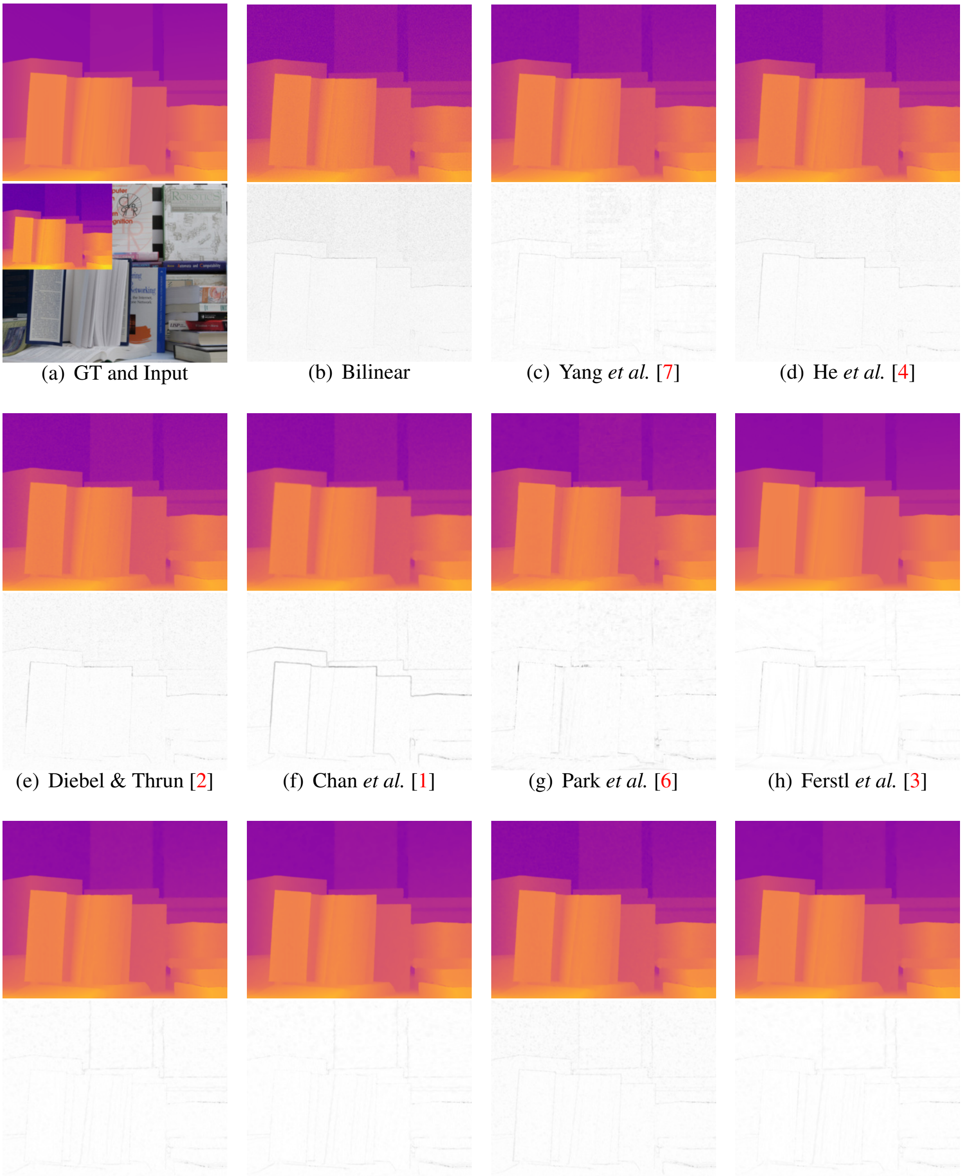
Similarly, in Section 3 we present our high resolution (HR) depth estimates on the images *Books*, *Devil*, and *Shark* of the challenging Time-of-Flight dataset ToFMark [3], where we compare our *deep primal-dual network* (FCN-PDN) to nearest neighbor and bilinear interpolation, as well as to the approaches by Kopf *et al.* [5], He *et al.* [4], and Ferstl *et al.* [3].

2 Noisy Middlebury



(a) GT and Input (b) Bilinear (c) Yang *et al.* [7] (d) He *et al.* [4]
(e) Diebel & Thrun [2] (f) Chan *et al.* [1] (g) Park *et al.* [6] (h) Ferstl *et al.* [3]
(i) FCN($d^{(\mathrm{mr})}$) (j) FCN-PDN($d^{(\mathrm{mr})}$) (k) FCN(s) (l) FCN-PDN(s)

Figure 1: Qualitative results for the image *Art* from the noisy Middlebury dataset [6] and a scale factor of $\times 2$. The first image in (a) shows the ground-truth HR depth and the second image depicts the input sample. In (b)-(l) we present the HR estimates of various methods and the corresponding error maps.

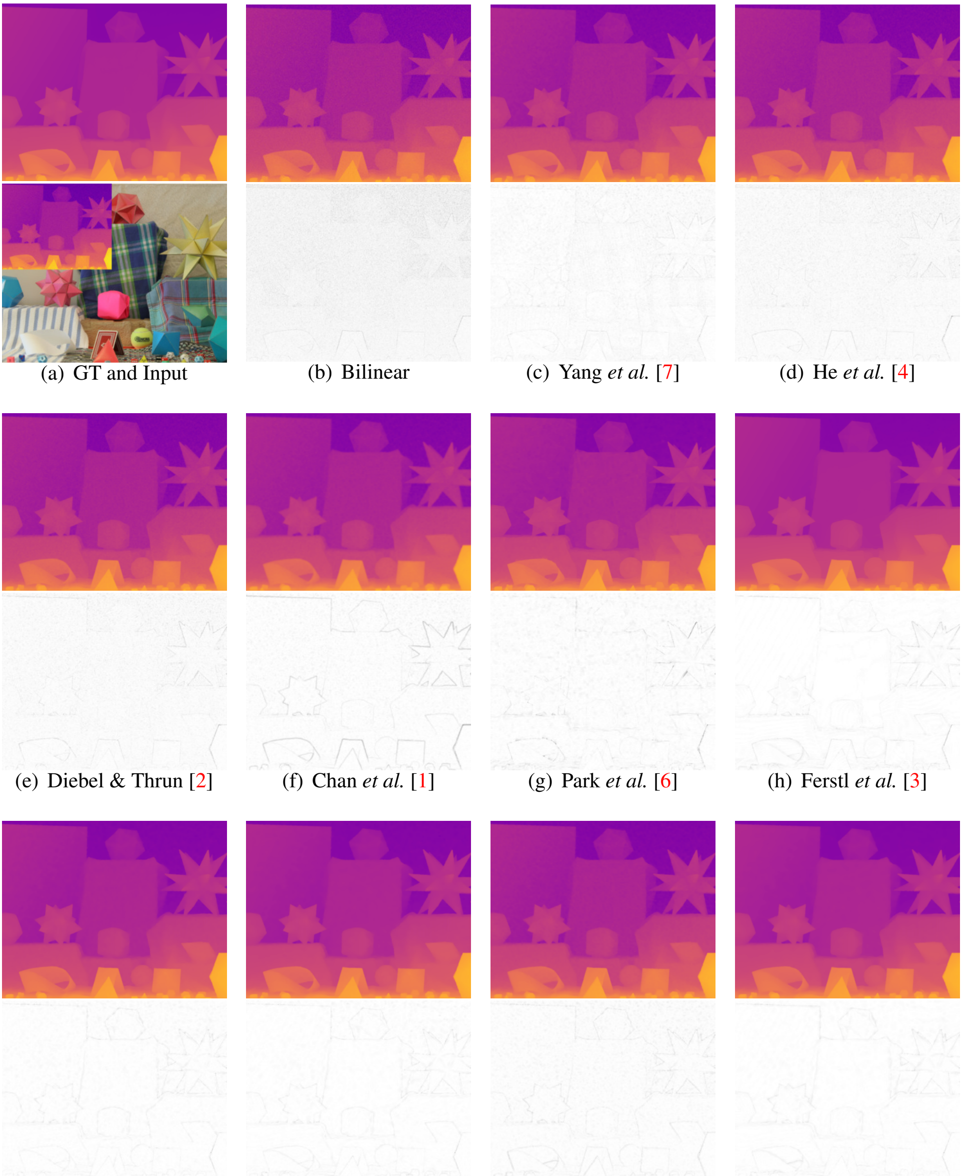


(a) GT and Input (b) Bilinear (c) Yang *et al.* [7] (d) He *et al.* [4]

(e) Diebel & Thrun [2] (f) Chan *et al.* [1] (g) Park *et al.* [6] (h) Ferstl *et al.* [3]

(i) FCN($d^{(\mathrm{mr})}$) (j) FCN-PDN($d^{(\mathrm{mr})}$) (k) FCN(s) (l) FCN-PDN(s)

Figure 2: Qualitative results for the image *Books* from the noisy Middlebury dataset [6] and a scale factor of $\times 2$. The first image in (a) shows the ground-truth HR depth and the second image depicts the input sample. In (b)-(l) we present the HR estimates of various methods and the corresponding error maps.

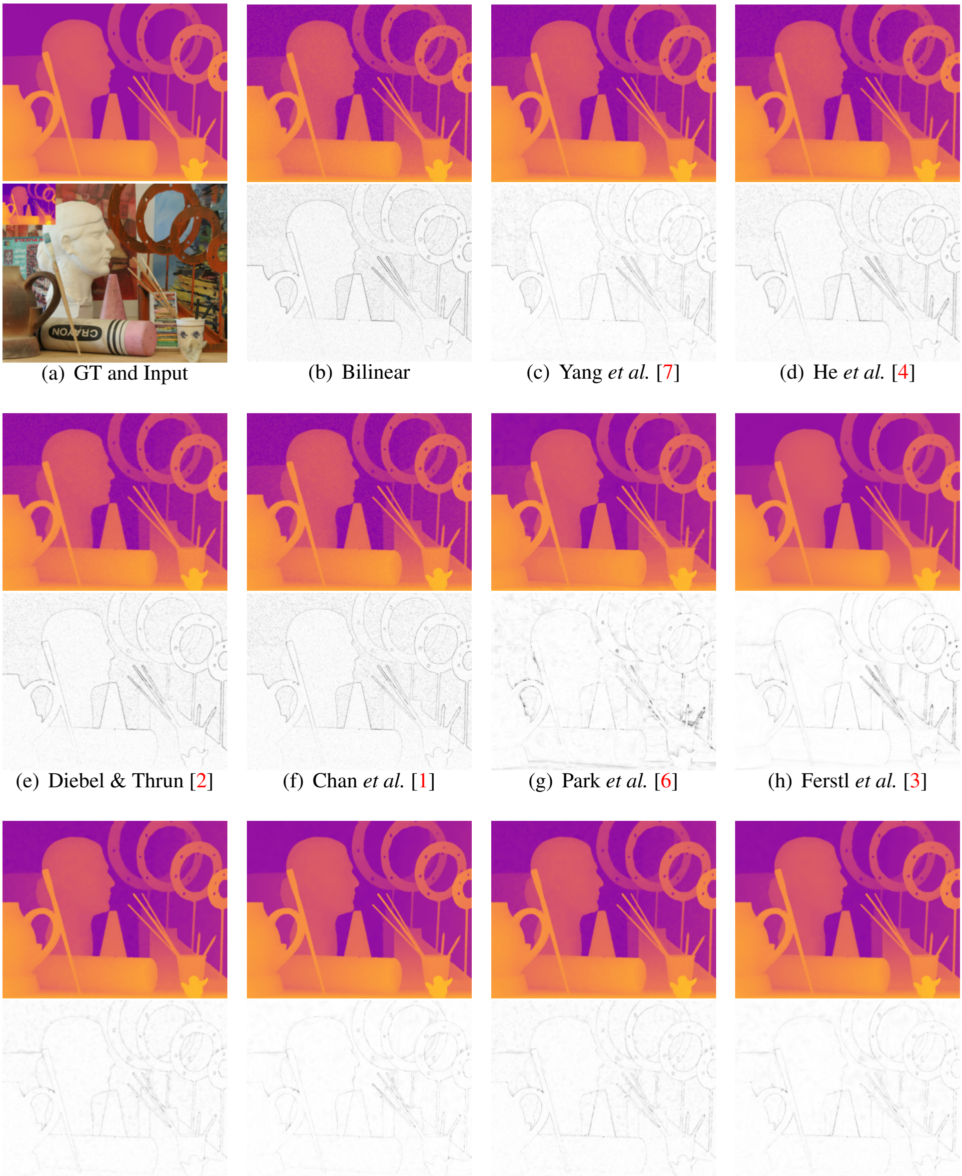


(a) GT and Input (b) Bilinear (c) Yang *et al.* [7] (d) He *et al.* [4]

(e) Diebel & Thrun [2] (f) Chan *et al.* [1] (g) Park *et al.* [6] (h) Ferstl *et al.* [3]

(i) FCN($d^{(\mathrm{mr})}$) (j) FCN-PDN($d^{(\mathrm{mr})}$) (k) FCN(s) (l) FCN-PDN(s)

Figure 3: Qualitative results for the image *Moebius* from the noisy Middlebury dataset [6] and a scale factor of $\times 2$. The first image in (a) shows the ground-truth HR depth and the second image depicts the input sample. In (b)-(l) we present the HR estimates of various methods and the corresponding error maps.

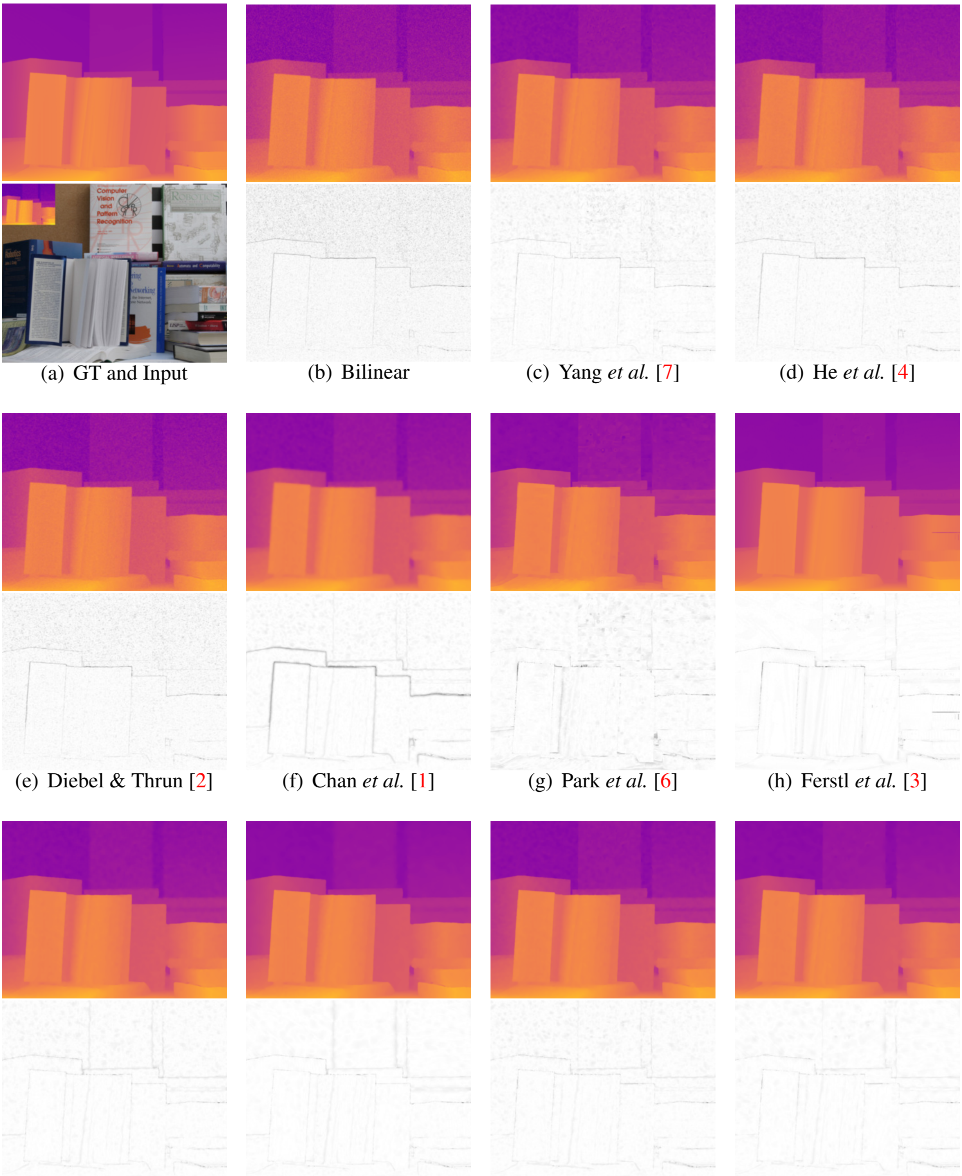


(a) GT and Input (b) Bilinear (c) Yang *et al.* [7] (d) He *et al.* [4]

(e) Diebel & Thrun [2] (f) Chan *et al.* [1] (g) Park *et al.* [6] (h) Ferstl *et al.* [3]

(i) FCN($d^{(\mathrm{mr})}$) (j) FCN-PDN($d^{(\mathrm{mr})}$) (k) FCN(s) (l) FCN-PDN(s)

Figure 4: Qualitative results for the image *Art* from the noisy Middlebury dataset [6] and a scale factor of $\times 4$. The first image in (a) shows the ground-truth HR depth and the second image depicts the input sample. In (b)-(l) we present the HR estimates of various methods and the corresponding error maps.



(a) GT and Input (b) Bilinear (c) Yang *et al.* [7] (d) He *et al.* [4]

(e) Diebel & Thrun [2] (f) Chan *et al.* [1] (g) Park *et al.* [6] (h) Ferstl *et al.* [3]

(i) FCN($d^{(\mathrm{mr})}$) (j) FCN-PDN($d^{(\mathrm{mr})}$) (k) FCN(s) (l) FCN-PDN(s)

Figure 5: Qualitative results for the image *Books* from the noisy Middlebury dataset [6] and a scale factor of $\times 4$. The first image in (a) shows the ground-truth HR depth and the second image depicts the input sample. In (b)-(l) we present the HR estimates of various methods and the corresponding error maps.

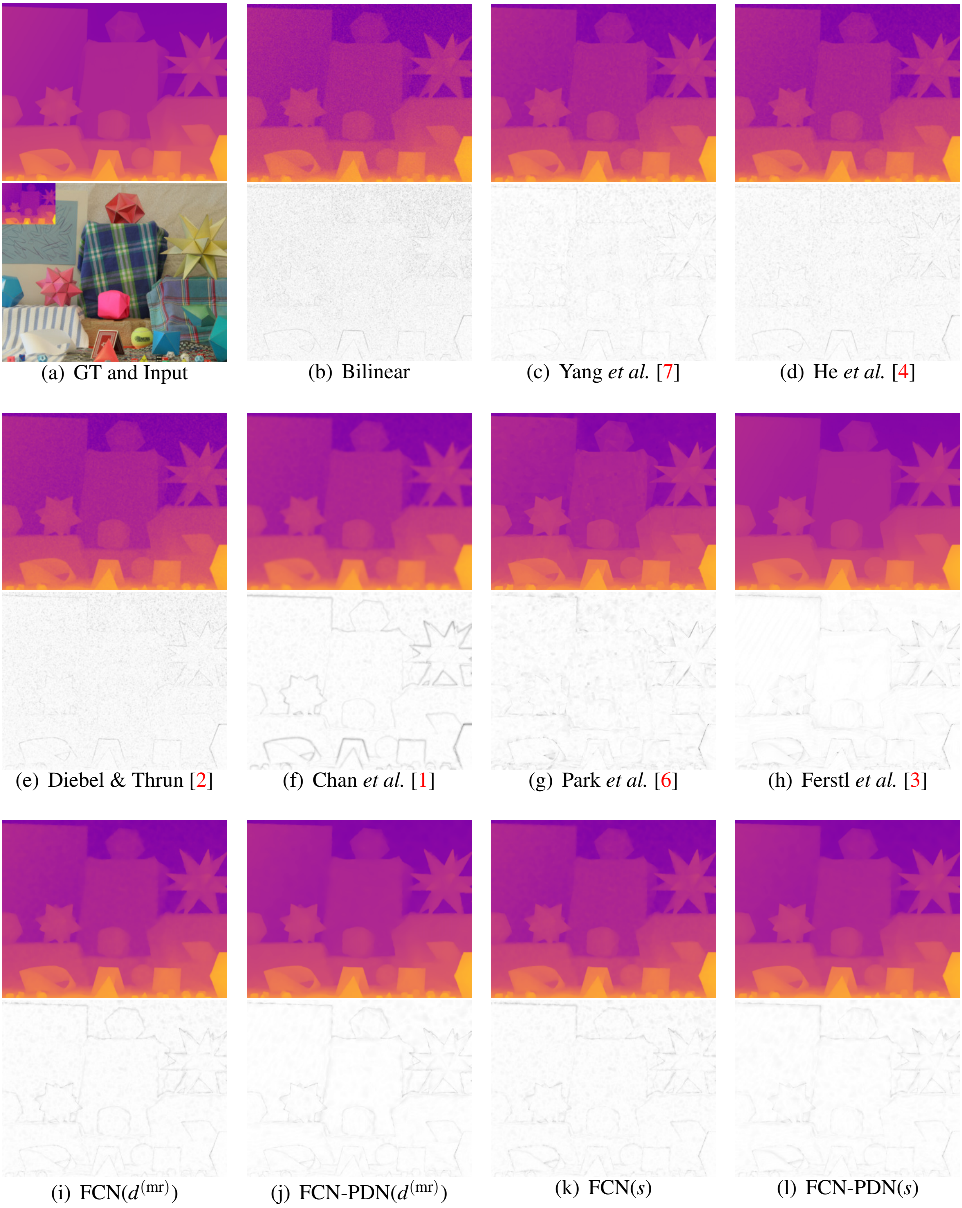
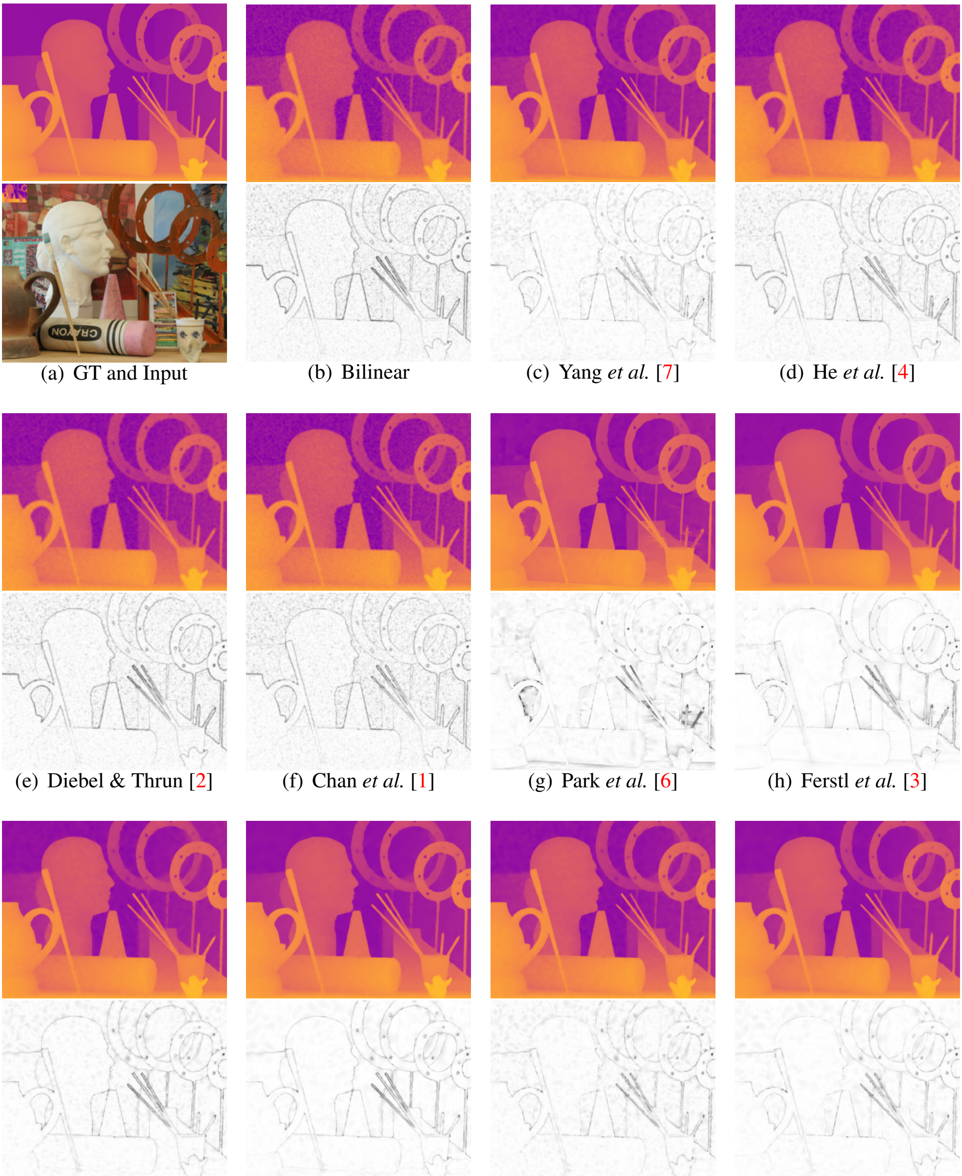


Figure 6: Qualitative results for the image *Moebius* from the noisy Middlebury dataset [6] and a scale factor of $\times 4$. The first image in (a) shows the ground-truth HR depth and the second image depicts the input sample. In (b)-(l) we present the HR estimates of various methods and the corresponding error maps.

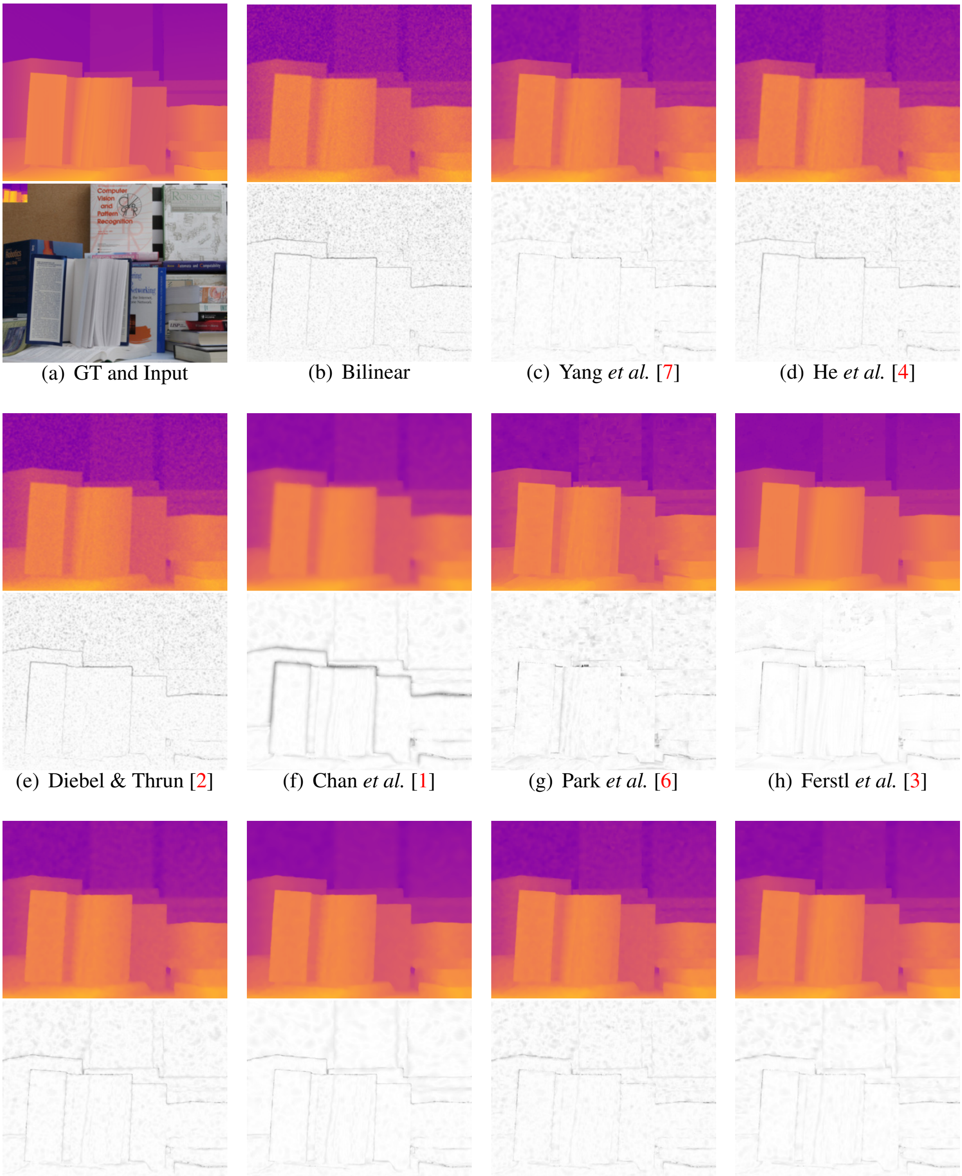


(a) GT and Input (b) Bilinear (c) Yang *et al.* [7] (d) He *et al.* [4]

(e) Diebel & Thrun [2] (f) Chan *et al.* [1] (g) Park *et al.* [6] (h) Ferstl *et al.* [3]

(i) FCN($d^{(\mathrm{mr})}$) (j) FCN-PDN($d^{(\mathrm{mr})}$) (k) FCN(s) (l) FCN-PDN(s)

Figure 7: Qualitative results for the image *Art* from the noisy Middlebury dataset [6] and a scale factor of $\times 8$. The first image in (a) shows the ground-truth HR depth and the second image depicts the input sample. In (b)-(l) we present the HR estimates of various methods and the corresponding error maps.

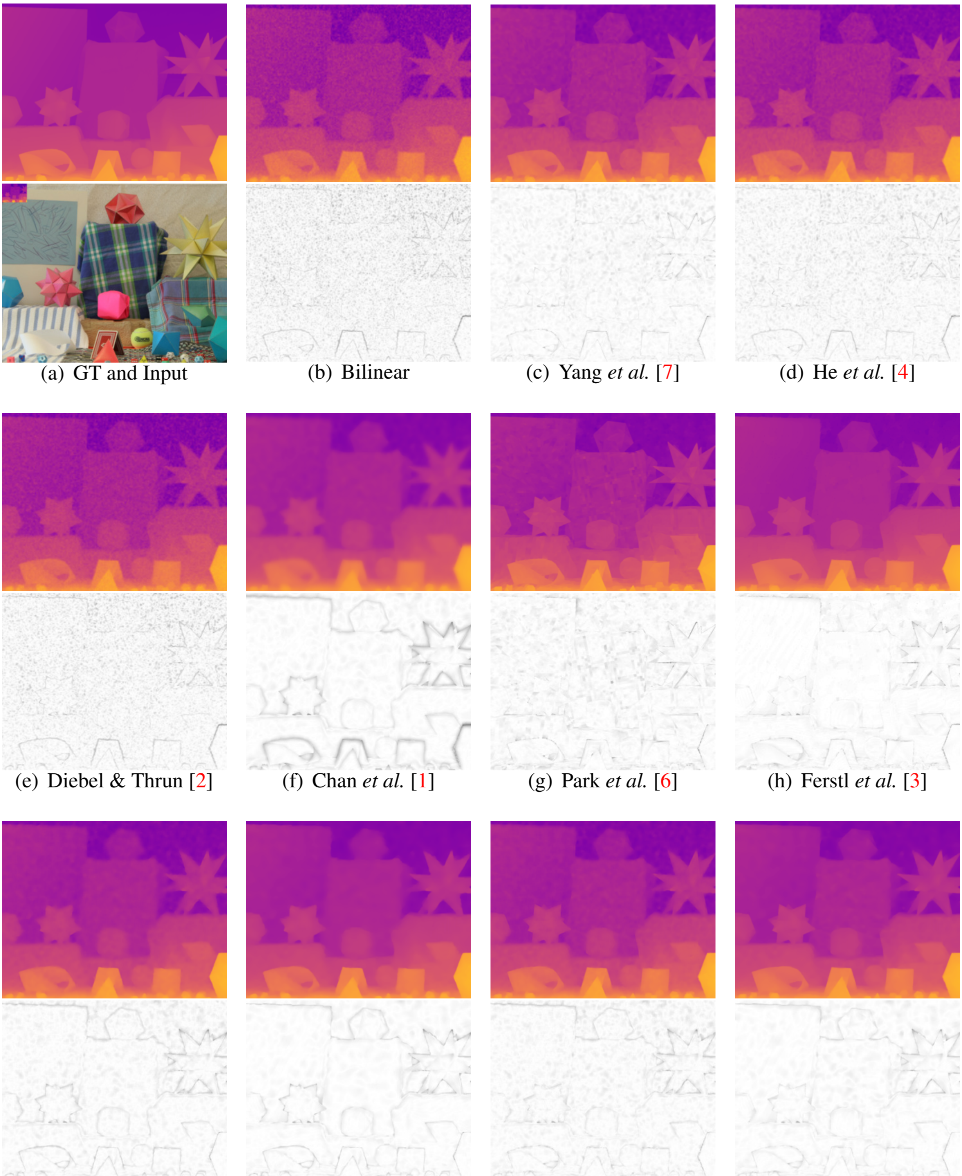


(a) GT and Input (b) Bilinear (c) Yang *et al.* [7] (d) He *et al.* [4]

(e) Diebel & Thrun [2] (f) Chan *et al.* [1] (g) Park *et al.* [6] (h) Ferstl *et al.* [3]

(i) FCN($d^{(\mathrm{mr})}$) (j) FCN-PDN($d^{(\mathrm{mr})}$) (k) FCN(s) (l) FCN-PDN(s)

Figure 8: Qualitative results for the image *Books* from the noisy Middlebury dataset [6] and a scale factor of $\times 8$. The first image in (a) shows the ground-truth HR depth and the second image depicts the input sample. In (b)-(l) we present the HR estimates of various methods and the corresponding error maps.

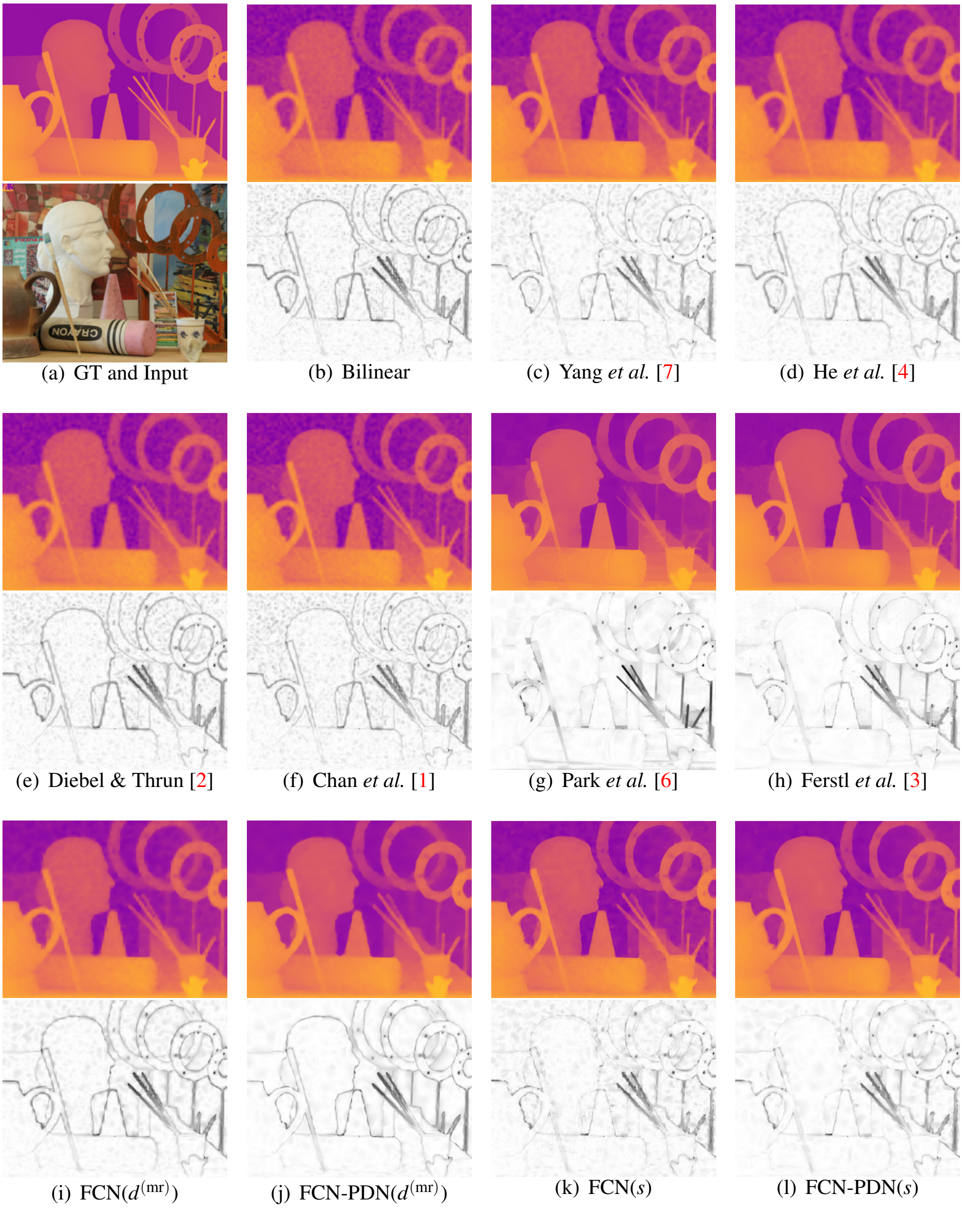


(a) GT and Input (b) Bilinear (c) Yang *et al.* [7] (d) He *et al.* [4]

(e) Diebel & Thrun [2] (f) Chan *et al.* [1] (g) Park *et al.* [6] (h) Ferstl *et al.* [3]

(i) FCN($d^{(\mathrm{mr})}$) (j) FCN-PDN($d^{(\mathrm{mr})}$) (k) FCN(s) (l) FCN-PDN(s)

Figure 9: Qualitative results for the image *Moebius* from the noisy Middlebury dataset [6] and a scale factor of $\times 8$. The first image in (a) shows the ground-truth HR depth and the second image depicts the input sample. In (b)-(l) we present the HR estimates of various methods and the corresponding error maps.

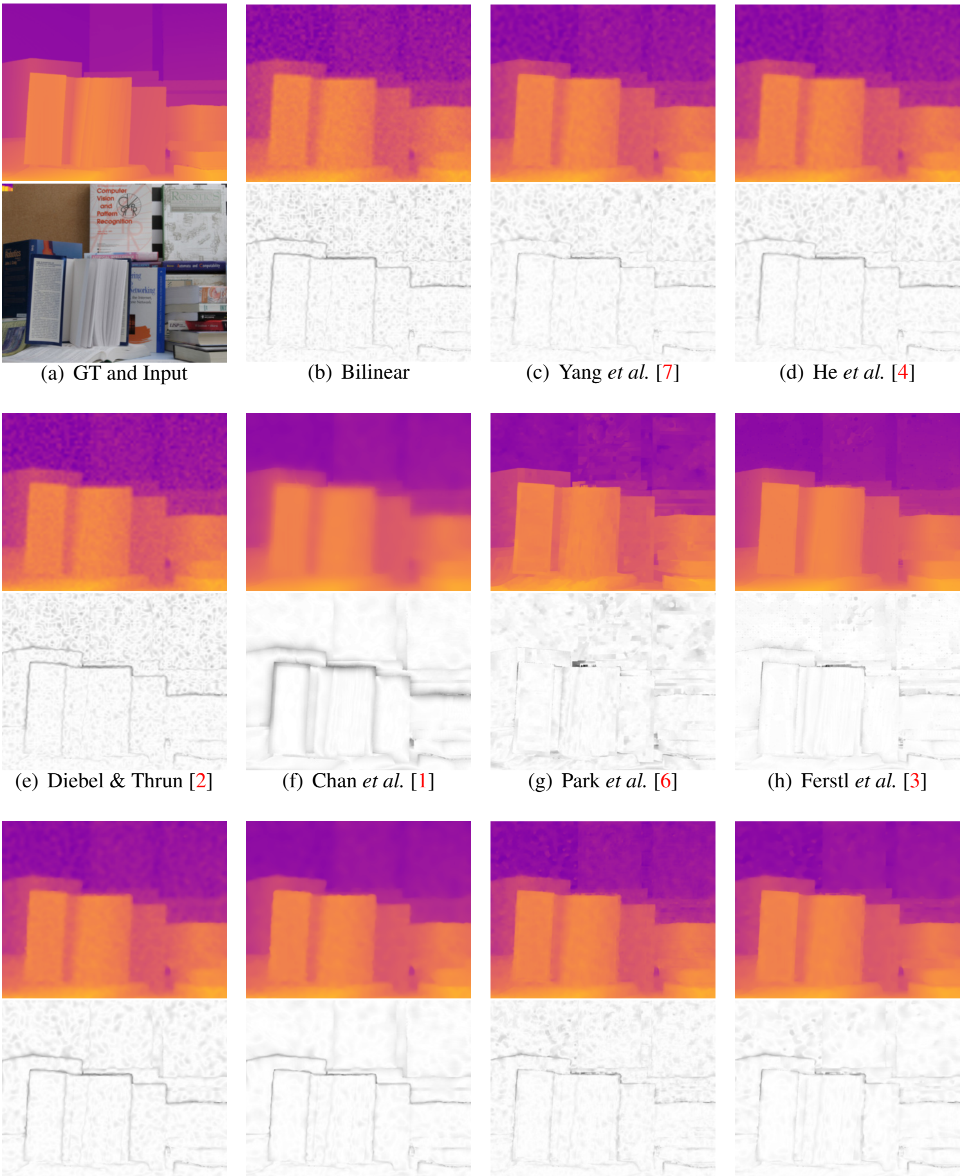


(a) GT and Input (b) Bilinear (c) Yang *et al.* [7] (d) He *et al.* [4]

(e) Diebel & Thrun [2] (f) Chan *et al.* [1] (g) Park *et al.* [6] (h) Ferstl *et al.* [3]

(i) FCN($d^{(\mathrm{mr})}$) (j) FCN-PDN($d^{(\mathrm{mr})}$) (k) FCN(s) (l) FCN-PDN(s)

Figure 10: Qualitative results for the image *Art* from the noisy Middlebury dataset [6] and a scale factor of $\times 16$. The first image in (a) shows the ground-truth HR depth and the second image depicts the input sample. In (b)-(l) we present the HR estimates of various methods and the corresponding error maps.

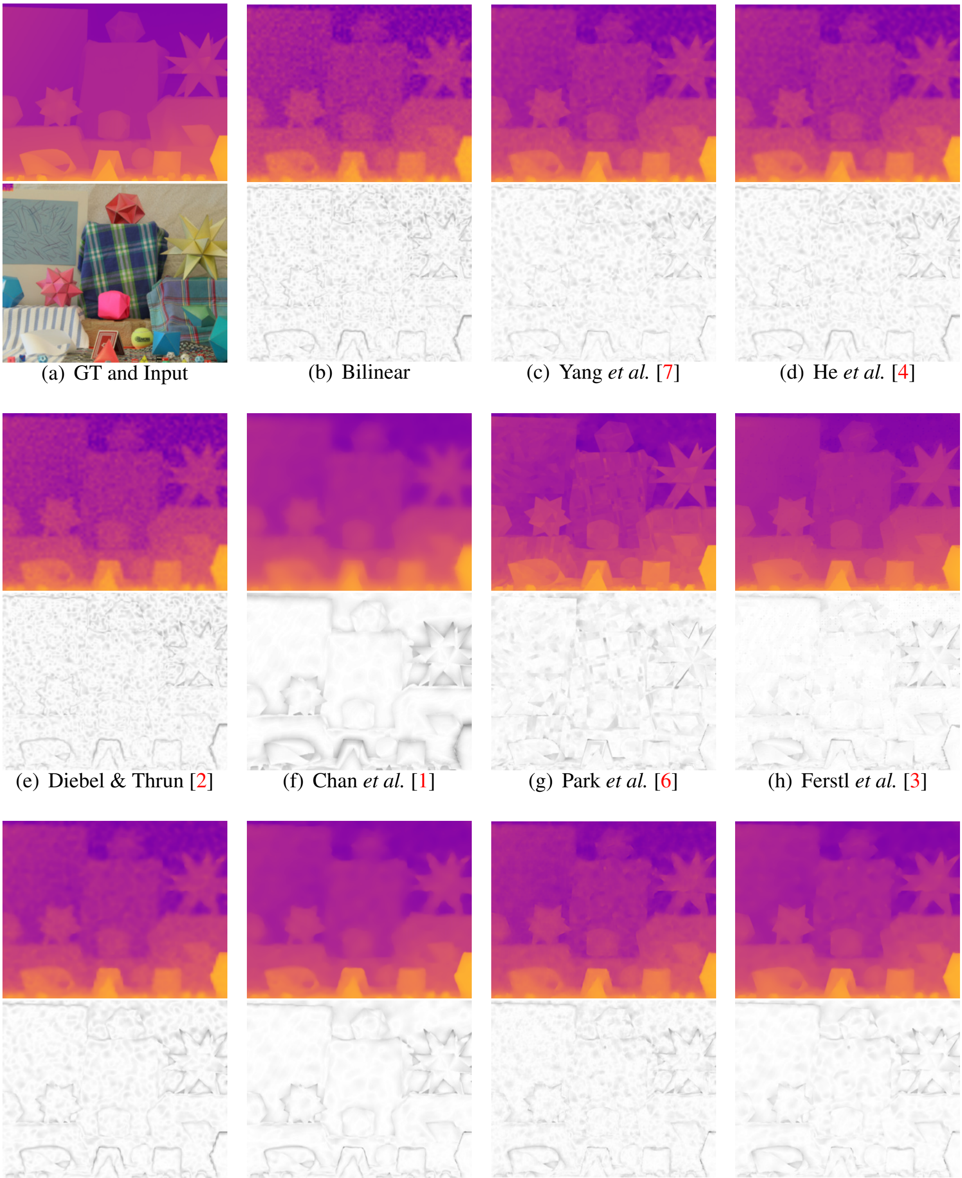


(a) GT and Input (b) Bilinear (c) Yang *et al.* [7] (d) He *et al.* [4]

(e) Diebel & Thrun [2] (f) Chan *et al.* [1] (g) Park *et al.* [6] (h) Ferstl *et al.* [3]

(i) FCN($d^{(\mathrm{mr})}$) (j) FCN-PDN($d^{(\mathrm{mr})}$) (k) FCN(s) (l) FCN-PDN(s)

Figure 11: Qualitative results for the image *Books* from the noisy Middlebury dataset [6] and a scale factor of $\times 16$. The first image in (a) shows the ground-truth HR depth and the second image depicts the input sample. In (b)-(l) we present the HR estimates of various methods and the corresponding error maps.



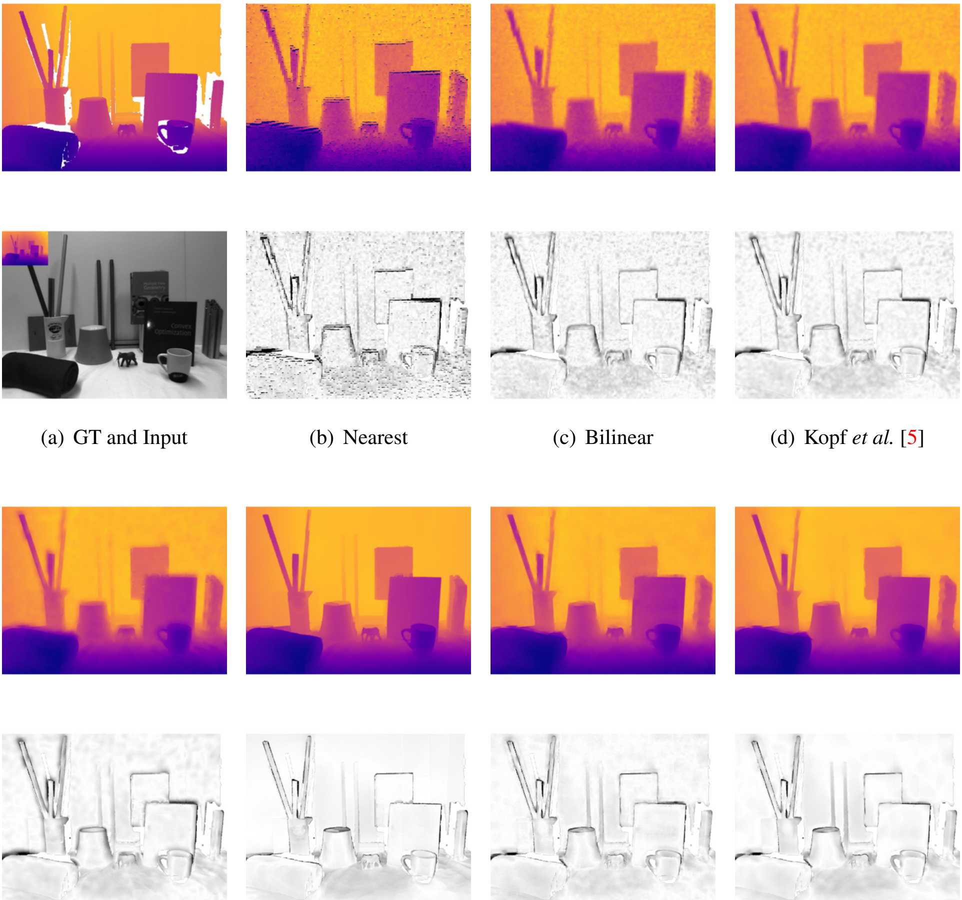
(a) GT and Input (b) Bilinear (c) Yang *et al.* [7] (d) He *et al.* [4]

(e) Diebel & Thrun [2] (f) Chan *et al.* [1] (g) Park *et al.* [6] (h) Ferstl *et al.* [3]

(i) FCN($d^{(\mathrm{mr})}$) (j) FCN-PDN($d^{(\mathrm{mr})}$) (k) FCN(s) (l) FCN-PDN(s)

Figure 12: Qualitative results for the image *Moebius* from the noisy Middlebury dataset [6] and a scale factor of $\times 16$. The first image in (a) shows the ground-truth HR depth and the second image depicts the input sample. In (b)-(l) we present the HR estimates of various methods and the corresponding error maps.

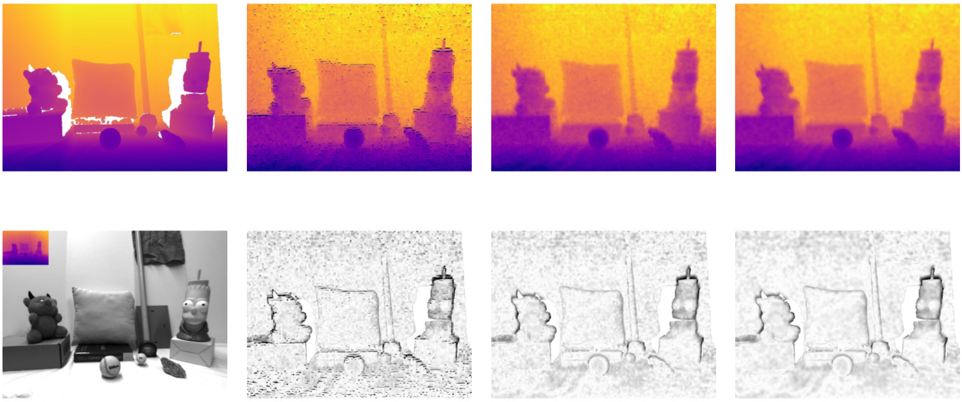
3 ToFMark



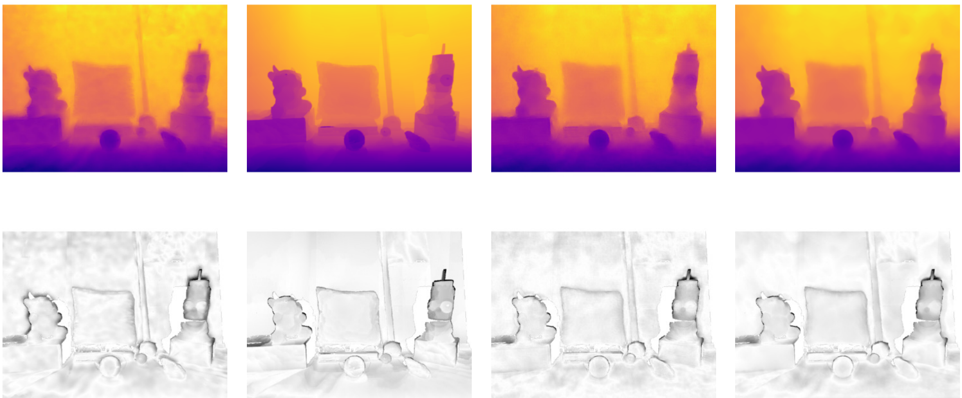
(a) GT and Input (b) Nearest (c) Bilinear (d) Kopf *et al.* [5]

(e) He *et al.* [4] (f) Ferstl *et al.* [3] (g) FCN(s) (h) FCN-PDN(s)

Figure 13: Qualitative results for image *Books* from the ToFMark dataset [3]. The first image in (a) shows the ground-truth HR depth and the second image depicts the input. In (b)-(h) we present the HR estimates of various methods and the corresponding error maps.

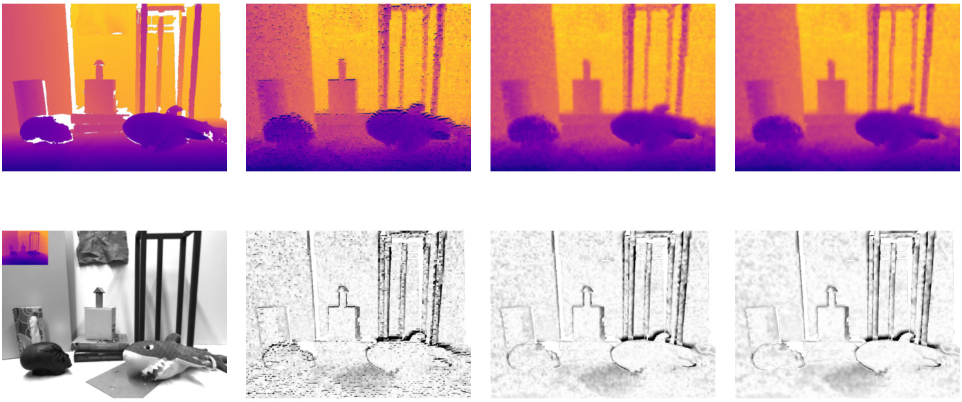


(a) GT and Input (b) Nearest (c) Bilinear (d) Kopf *et al.* [5]

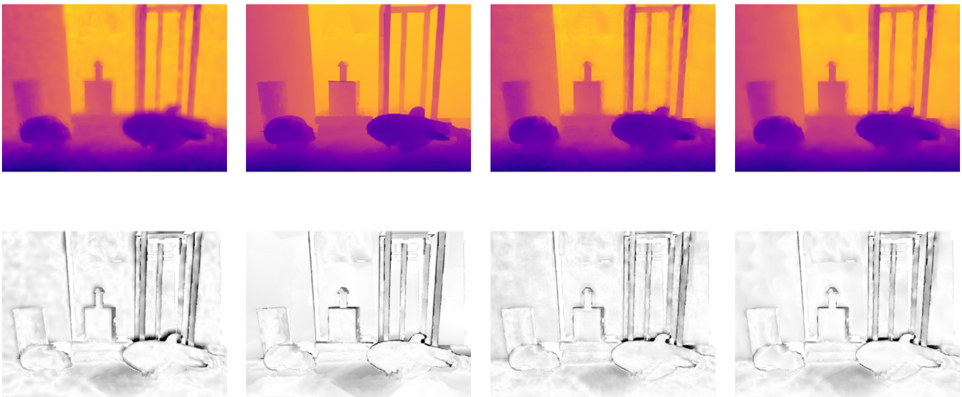


(e) He *et al.* [4] (f) Ferstl *et al.* [3] (g) FCN(s) (h) FCN-PDN(s)

Figure 14: Qualitative results for image *Devil* from the ToFMark dataset [3]. The first image in (a) shows the ground-truth HR depth and the second image depicts the input. In (b)-(h) we present the HR estimates of various methods and the corresponding error maps.



(a) GT and Input (b) Nearest (c) Bilinear (d) Kopf *et al.* [5]



(e) He *et al.* [4] (f) Ferstl *et al.* [3] (g) FCN(s) (h) FCN-PDN(s)

Figure 15: Qualitative results for image *Shark* from the ToFMark dataset [3]. The first image in (a) shows the ground-truth HR depth and the second image depicts the input. In (b)-(h) we present the HR estimates of various methods and the corresponding error maps.

References

[1] Derek Chan, Hylke Buisman, Christian Theobalt, and Sebastian Thrun. A Noise-aware Filter for Real-time Depth Upsampling. In *European Conference on Computer Vision Workshops (ECCVW)*, 2008.

[2] James Diebel and Sebastian Thrun. An Application of Markov Random Fields to Range Sensing. In *Proceedings of Conference on Neural Information Processing Systems (NIPS)*, 2005.

[3] David Ferstl, Christian Reinbacher, René Ranftl, Matthias Rüther, and Horst Bischof. Image Guided Depth Upsampling using Anisotropic Total Generalized Variation. In *IEEE International Conference on Computer Vision (ICCV)*, 2013.

[4] Kaiming He, Jian Sun, and Xiaoou Tang. Guided Image Filtering. In *European Conference on Computer Vision (ECCV)*, 2010.

[5] Johannes Kopf, Michael F. Cohen, Dani Lischinski, and Matthew Uyttendaele. Joint Bilateral Upsampling. *ACM Transactions on Graphics (TOG)*, 26(3):96, 2007.

[6] Jaesik Park, Hyeongwoo Kim, Yu-Wing Tai, Michael S. Brown, and In-So Kweon. High Quality Depth Map Upsampling for 3D-TOF Cameras. In *IEEE International Conference on Computer Vision (ICCV)*, 2011.

[7] Jingyu Yang, Xinchen Ye, Kun Li, Chunping Hou, and Yao Wang. Color-Guided Depth Recovery From RGB-D Data Using an Adaptive Autoregressive Model. *IEEE Transactions on Image Processing*, 23(8):3443–3458, 2014.

[8] Qingxiong Yang, Ruigang Yang, James Davis, and David Nistér. Spatial-Depth Super Resolution for Range Images. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2007.