Depth Map Super-Resolution by Deep Multi-Scale Guidance

Tak-Wai Hui, Chen Change Loy, Xiaoou Tang

Department of Information Engineering, The Chinese University of Hong Kong

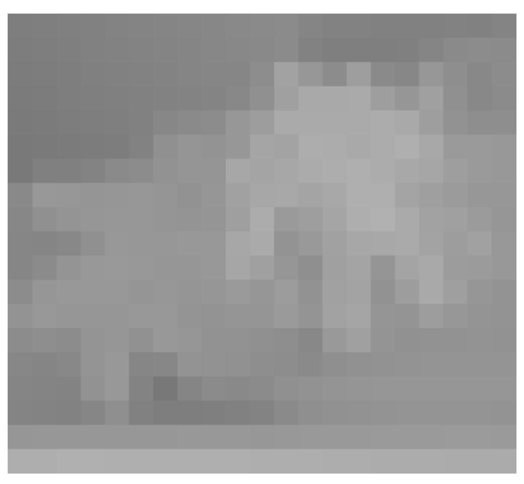
1. Introduction

- Depth information is essential in many applications such as autonomous navigation, 3D reconstruction, and human-computer interaction.
- The resolution of depth maps which is provided in a low-cost depth camera is generally very limited.
- We address an upsampling problem in which the corresponding high-resolution (HR) depth map is recovered from a given low-resolution (LR) depth map (and a HR intensity image).

2. Challenges

- Fine structures in the enlarged image are either lost or severely distorted.

HR RGB



LR depth map (↓8x)

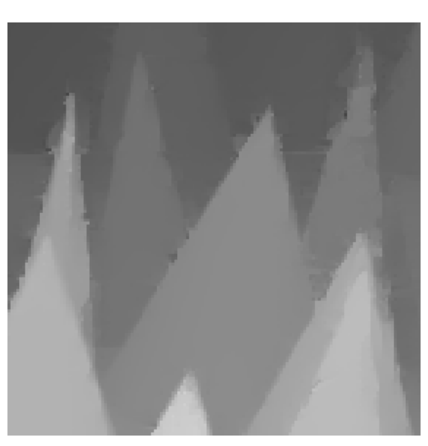
SRCNN [11] (↑8x)

MSG-Net (↑8x)

- Features in intensity images are often over-transferred to the depth image.

RGB

Guided Filter [8]

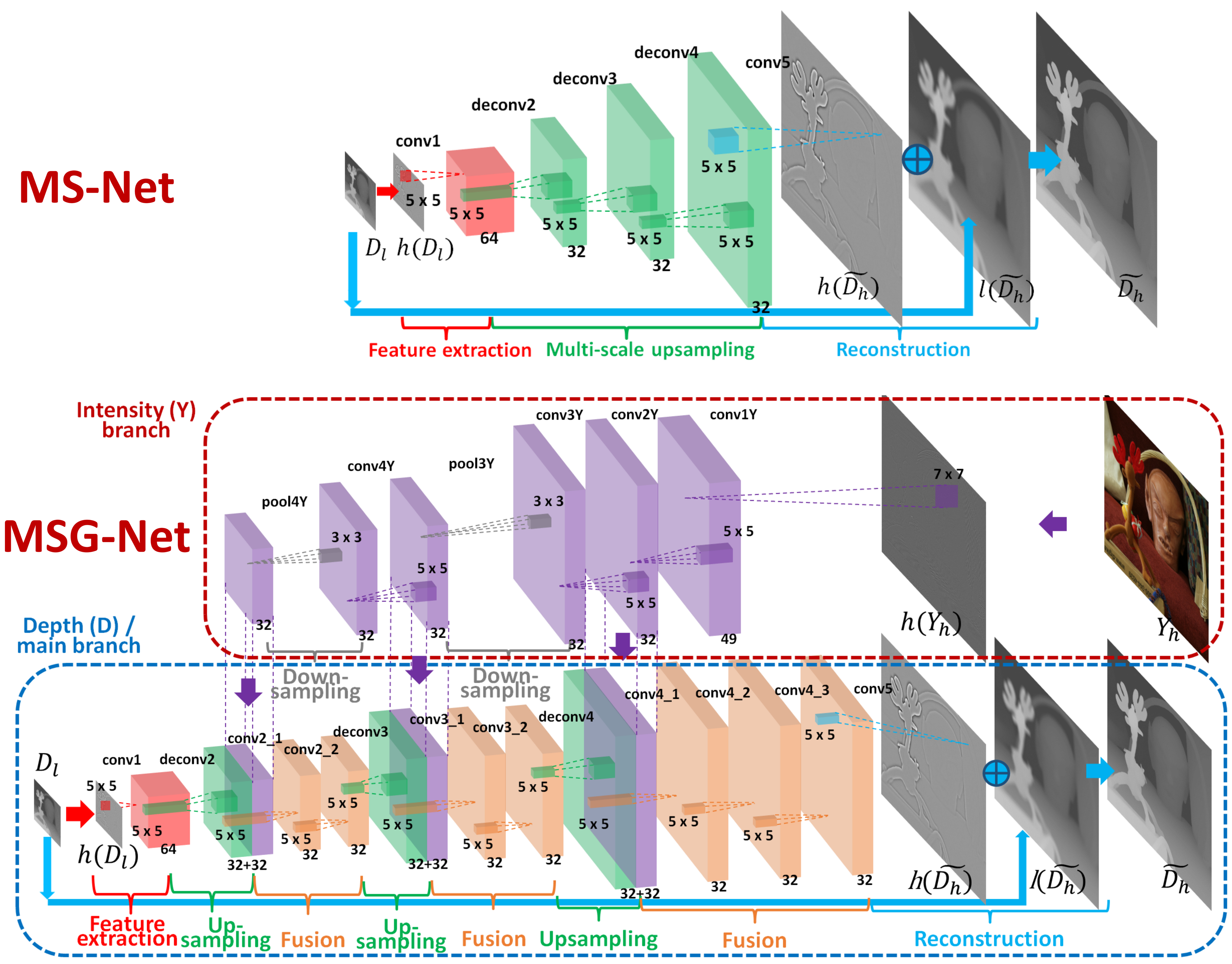


TV [4] (↑2x)

Ground-truth

MSG-Net (↑2x)

3. MS-Net and MSG-Net

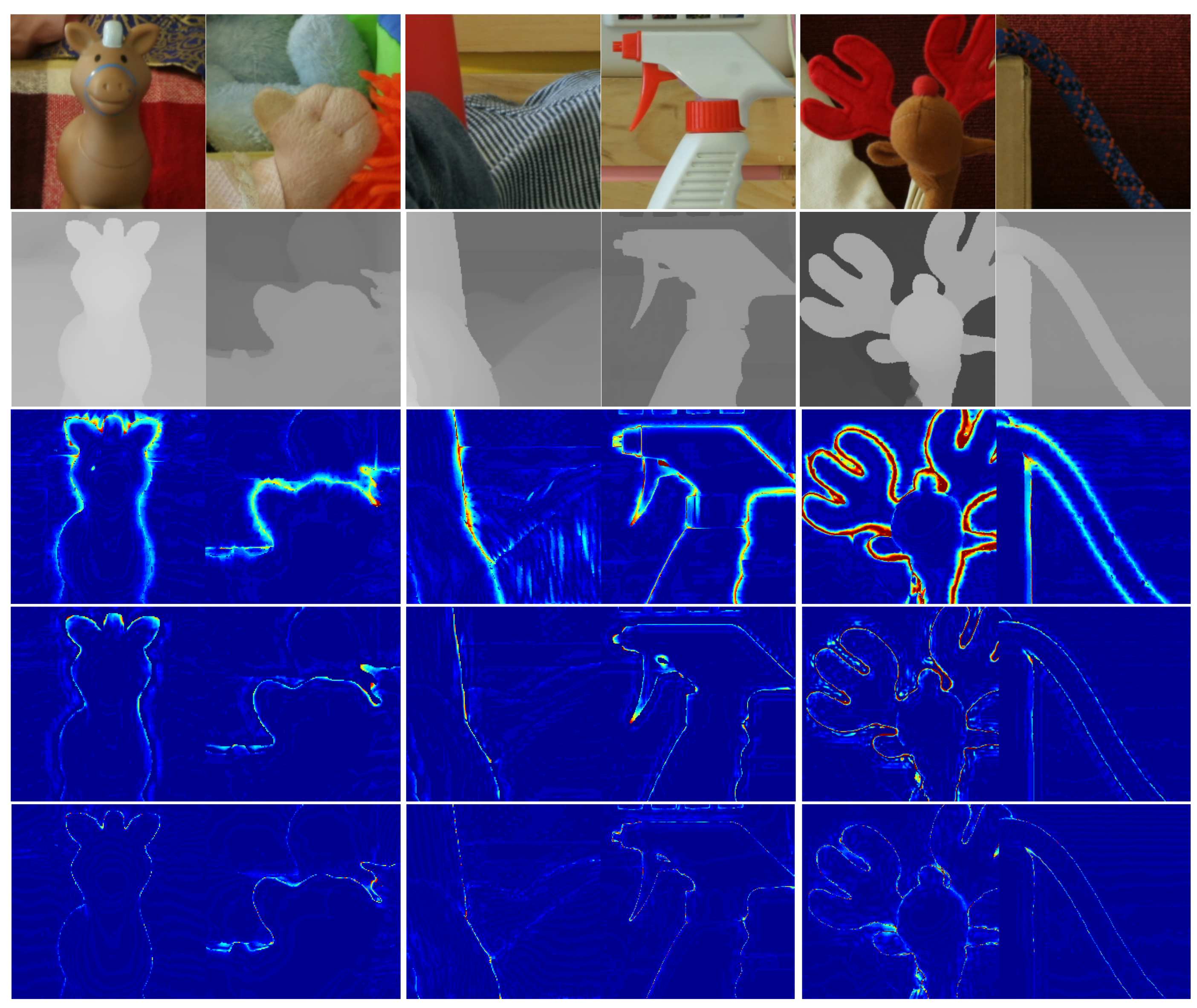


- Only high-freq. part $h(D_l)$ is used for training.
- Upsampled low-freq. part $l(\widetilde{D_h})$ is added back for testing.

3. MS-Net and MSG-Net

	Art				Books				Moebius			
	2×	4×	8×	16×	2×	4×	8×	16×	2×	4×	8×	16×
Bilinear	2.834	4.147	5.995	8.928	1.119	1.673	2.394	3.525	1.016	1.499	2.198	3.179
MRFs [15]	3.119	3.794	5.503	8.657	1.205	1.546	2.209	3.400	1.187	1.439	2.054	3.078
Bilateral [13]	4.066	4.056	4.712	8.268	1.615	1.701	1.949	3.325	1.069	1.386	1.820	2.494
Park *et al.* [2]	2.833	3.498	4.165	6.262	1.088	1.530	1.994	2.760	1.064	1.349	1.804	2.377
Guided [8]	2.934	3.788	4.974	7.876	1.162	1.572	2.097	3.186	1.095	1.434	1.878	2.851
Kiechle *et al.* [3]	1.246	2.007	3.231	5.744	0.652	0.918	1.274	1.927	0.640	0.887	1.272	2.128
Ferstl *et al.* [4]	3.032	3.785	4.787	7.102	1.290	1.603	1.992	2.941	1.129	1.458	1.914	2.630
Lu *et al.* [6]	-	-	5.798	7.648	-	-	2.728	3.549	-	-	2.422	3.118
SRCNN [11]	1.133	2.017	3.829	7.271	0.523	0.935	1.726	3.100	0.537	0.913	1.579	2.689
SRCNN2	0.902	1.874	3.704	7.309	0.464	0.846	1.591	3.123	0.454	0.864	1.482	2.679
Wang *et al.* [12]	1.670	2.525	3.957	6.226	0.668	1.098	1.646	2.428	0.641	0.979	1.459	2.202
MS-Net	0.813	1.627	2.769	5.802	0.417	0.724	1.072	1.802	0.413	0.741	1.138	1.910
MSG-Net	**0.663**	**1.474**	**2.455**	**4.574**	**0.373**	**0.667**	**1.029**	**1.601**	**0.357**	**0.661**	**1.015**	**1.633**

	Dolls				Laundry				Reindeer			
	2×	4×	8×	16×	2×	4×	8×	16×	2×	4×	8×	16×
Bicubic	0.914	1.305	1.855	2.625	1.614	2.408	3.452	5.095	1.938	2.809	3.986	5.823
Park *et al.* [2]	0.963	1.301	1.745	2.412	1.552	2.132	2.770	4.158	1.834	2.407	2.987	4.294
Aodha *et al.*	-	1.977	-	-	-	2.969	-	-	-	3.178	-	-
CLMF0 [35]	0.990	1.271	1.878	2.291	1.689	2.312	3.084	4.312	1.955	2.690	3.417	4.674
CLMF1 [35]	0.972	1.267	1.707	2.232	1.689	2.512	2.892	4.302	1.948	2.699	3.331	4.774
Ferstl *et al.* [4]	1.118	1.355	1.859	3.574	1.989	2.511	3.757	6.407	2.407	2.712	3.789	7.271
Kiechle *et al.* [3]	0.696	0.921	1.259	1.736	0.746	1.212	2.077	3.621	0.920	1.559	2.583	4.644
AP [5]	1.147	1.350	1.646	2.323	1.715	2.255	2.848	4.656	1.803	2.431	2.949	4.088
SRCNN [11]	0.581	0.946	1.518	2.445	0.635	1.176	2.430	4.579	0.765	1.499	2.864	5.249
SRCNN2	0.473	0.881	1.461	2.422	0.506	1.084	2.314	4.601	0.603	1.352	2.740	5.330
Wang *et al.* [12]	0.670	0.989	1.445	2.107	1.039	1.630	2.466	3.834	1.252	1.914	2.878	4.526
MS-Net	0.437	0.740	1.166	1.832	0.475	0.883	1.618	3.385	0.556	1.107	1.972	3.921
MSG-Net	**0.345**	**0.690**	**1.051**	**1.597**	**0.371**	**0.787**	**1.514**	**2.629**	**0.424**	**0.984**	**1.757**	**2.919**



TV [4] (↑8x)

SRCNN [11] (↑8x)

MSG-Net (↑8x)

References:

[4] Ferstl *et al.*, Image guided depth upsampling using anisotropic total generalized variation. ICCV, pp. 993–1000, 2013.

[8] He *et al.*, Guided image filtering. PAMI 35(6), pp. 1397–1409, 2013.

[11] Dong *et al.*, Image super-resolution using deep convolutional networks. PAMI 38(2), pp. 295–307, 2015.

Project page: