Depth Map Upsampling by Self-Guided Residual Interpolation

Yosuke Konno[1], Masayuki Tanaka[1], Masatoshi Okutomi[1],
Yukiko Yanagawa[2], Koichi Kinoshita[2], and Masato Kawade[2]
[1]Tokyo Institute of Technology
[2]Technology and Intellectual Property H. Q., Omron Corporation

***Abstract*—In this paper, we propose a simple and effective depth upsampling technique using self-guided residual interpolation. The original residual interpolation requires guidance information such as high-resolution RGB color image. However, self-guided residual interpolation requires only a single depth map. In the proposed algorithm, a tentative estimation of a high-resolution depth map is first generated from an input low-resolution depth map. Then, re-interpolation is applied to the residual domain, which is defined by differences between the input depth map and the tentative estimate. A precise high-resolution depth map is obtainable by interpolating in the residual domain. Experimental results demonstrate that our algorithm can outperform state-of-the-art depth map upsampling algorithms.**

I. Introduction

Accurate depth map estimation is widely used in many computer vision applications including object detection [1], human pose analysis [2], and semantic scene recognition [3]. Many depth acquisition systems have been developed [4]. A common issue among those systems is the limited resolution of the acquired depth map. Therefore, depth map upsampling serves a crucially important role in various applications.

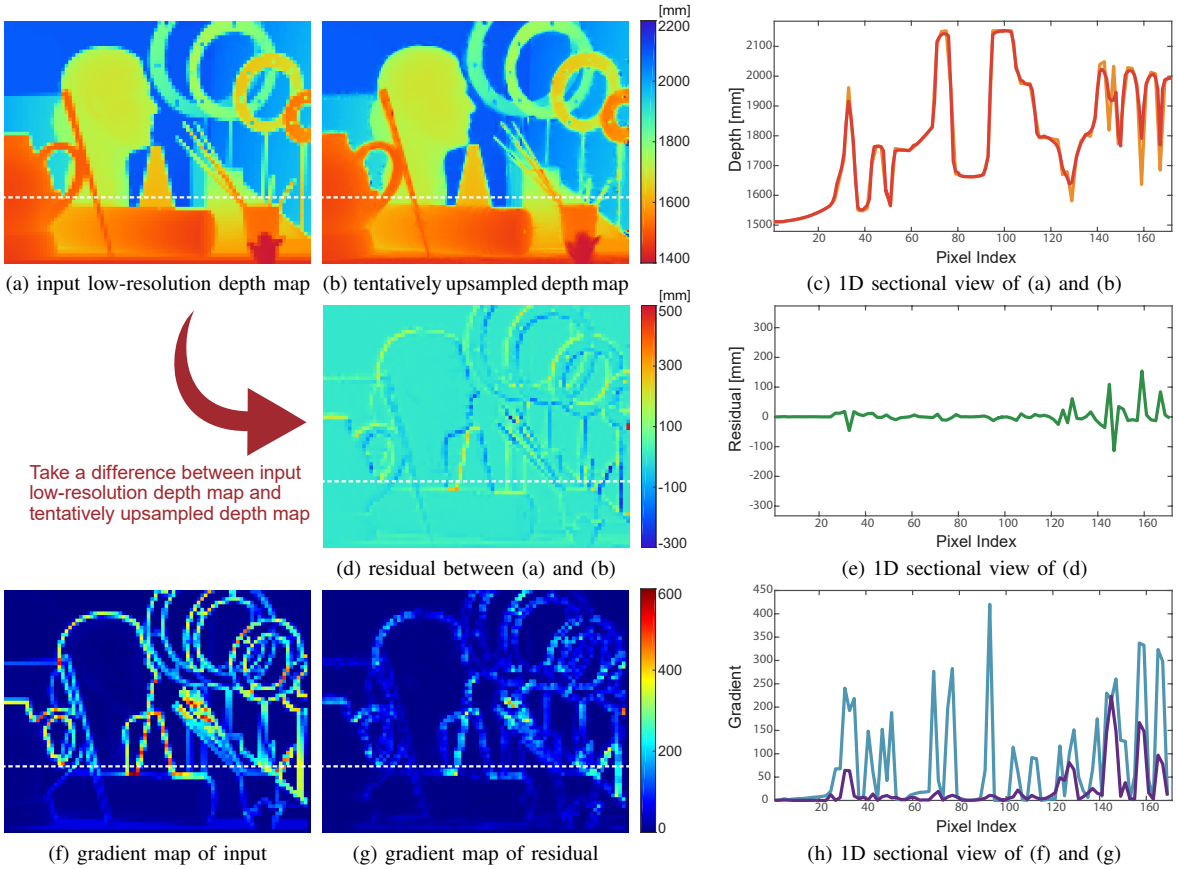
Depth map upsampling is classified into three categories: multiple depth frame upsampling [5], [6], depth map upsampling with high-resolution intensity image [7]–[13] and single depth map upsampling [14]–[16]. Multiple depth frame upsampling requires multiple sensors or time sequences to obtain multiple depth frames. Moreover, accurate alignment of depth frames is necessary. However, depth map alignment is an extremely challenging problem because the alignment requires precise camera motion estimation. Recently, many algorithms of the depth map upsampling with the high-resolution intensity image have been proposed [7]–[13]. Those algorithms enhance the resolution of the depth map, assuming that strong correlation exists between the depth map and the intensity image. However, this assumption is not always true. Severe artifacts appear in regions where the assumptions fail. Furthermore, depth map upsampling with the high-resolution intensity image requires an additional sensor to obtain the high-resolution intensity image. Alignment between the depth map and the high-resolution intensity image is an very challenging problem.

Single depth map upsampling is widely demanded because it requires no alignment, additional sensors, or hardware. However, single depth map upsampling is a highly ill-posed problem. In this paper, we propose a novel single depth map upsampling algorithm based on a residual interpolation. The residual interpolation is originally proposed as a part of color image demosaicking [17]–[19]. The fundamental idea of residual interpolation is to perform interpolation in a residual domain, where the residual is defined as a difference between a tentatively estimated high-resolution depth map and an input low-resolution depth map. Although the original residual interpolation requires additional information for tentative estimation of the high-resolution depth map, the proposed algorithm generates a tentative estimate from the input low-resolution depth map. We designate this interpolation as self-guided residual interpolation. To improve the upsampling accuracy, we apply a gradual refinement approach by following Glasner *et al.* [20]. They iteratively applies the super-resolution for natural images with small steps to obtain the desired resolution. We show experimentally that this gradual refinement approach is effective for depth map upsampling. Experimental results demonstrate that the proposed depth map upsampling algorithm outperforms existing upsampling algorithms.

II. Related Works

Depth upsampling is an actively studied in computer vision and image processing. Here, we briefly describe two approaches of depth map upsampling. First is an intensity guided approach, which uses a high-resolution intensity image as a depth clue. Second is single depth map upsampling, which can upsample from a single depth map only.

Intensity-guided approach. This is the most common strategy of depth map upsampling. Kopf *et al.* [7] proposed joint bilateral upsampling known as one of the seminal works of filtering-based depth upsampling algorithm. It applies the joint bilateral filter [22] to an input depth map with a high-resolution intensity guide. Yang *et al.* [8] applied joint bilateral upsampling on a cost volume constructed from an input low-resolution depth map. Many guide-based-filtering upsampling algorithms have been proposed, e.g., guided image filtering [9] and joint geodesic filter [10]. Guide-based-optimization upsampling algorithms have also been proposed [11]–[13]. Diebel and Thrun [11] proposed an MRF framework by which the smoothness term is penalized according to edges on a high-resolution intensity guide. Park *et al.* [12] involved a non-local means regularization term that preserves thin



(a) input low-resolution depth map (b) tentatively upsampled depth map (c) 1D sectional view of (a) and (b)

(d) residual between (a) and (b) (e) 1D sectional view of (d)

(f) gradient map of input (g) gradient map of residual (h) 1D sectional view of (f) and (g)

Fig. 1. Depth distributions of input and residual on *"Art"* from Middlebury 2005 dataset [21]. Graphs of the right column are 1D sectional views in the dashed line on the left images. The tentative high-resolution depth map in (b) is upsampled by guided upsampling [9] under intensity guidance. Gradient maps in (f), (g) express the absolute magnitude of the gradient at each pixels. Radical fluctuation of the depth value is occurred everywhere in the input depth map (a) – (c). However, it is observed that the fluctuation is extremely suppressed in the residual (d), (e). The same could be read from the gradient maps (f) – (h).

structures. Ferstl [13] *et al.* formulated depth upsampling as a convex optimization problem expressed with total generalized variation. In guide-based-optimization upsampling algorithms, parameters of an energy function to be optimized are tuned based on a high-resolution intensity guide.

The intensity guided approach requires the aligned high-resolution intensity guide. However, in many practical cases, it is not easy to obtain such an aligned high-resolution intensity guide.

Single depth map upsampling. A depth map upsampling with a single low-resolution depth map is highly demanded. Yang *et al.* [14] proposed dictionary-based upsampling. A dictionary pair is learned with low-resolution and high-resolution training patch pairs. Single depth map upsampling is performed with the learned dictionary pair.

Aodha *et al.* [15] seeks appropriate candidates of high-resolution depth patches from their database for each local low-resolution depth patch and assembles them in the form of an MRF labeling task. Hornáček *et al.* [16] proposed a similar method that seeks high-resolution and low-resolution depth patch pairs from an input image itself.

Single depth map upsampling requires no aligned high-resolution intensity guide. It is much more beneficial than an intensity-guided approach in the practical situation. In this paper, we also propose a single depth map upsampling algorithm.

III. Proposed Method

In general, the interpolation accuracy depends heavily on the input image smoothness. This is a common property for any kind of interpolation algorithms. In other words, the low gradient energy of the input image makes interpolation easy. That observation implies that one can improve the interpolation performance further if the input image could be transformed into the domain with the lower gradient energy. Residual interpolation has been proposed based on this idea [17]–[19]. Residual interpolation performs upsampling in the residual domain, where the residual is defined as the difference between the tentatively estimated high-resolution image and the input low-resolution image intensity. Residual interpolation also adds the residual to the tentative estimate to enhance the upsampling result. The gradient energy of residual is suppressed to an extreme degree for any input image, as shown in Fig. 1. Consequently, better performance can be expected even using a naïve interpolation algorithm such as bicubic interpolation in the residual domain. That is also true for the depth map upsampling.

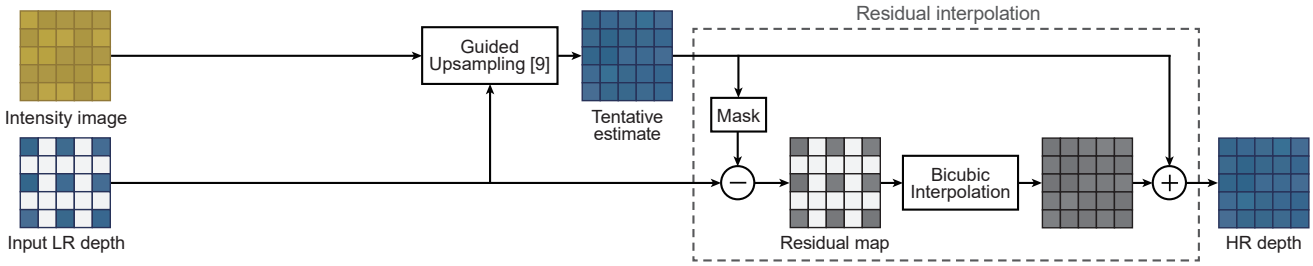


Fig. 2. Processing pipeline of intensity-guided residual interpolation [23].

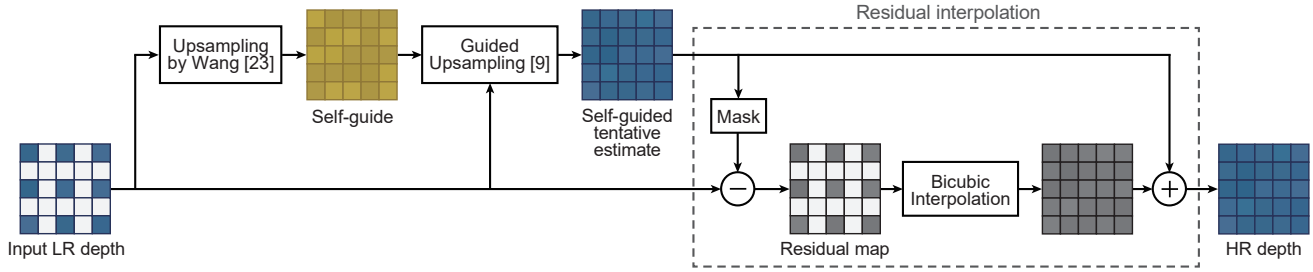


Fig. 3. Processing pipeline of our proposed self-guided residual interpolation.

The original residual interpolation requires a high-resolution intensity guide. Depth map upsampling based on the residual interpolation shown in Fig. 2 has been proposed in [23]. However, that algorithm also requires a high-resolution intensity guide.

Herein, we propose single depth map upsampling based on residual interpolation, which requires no high-resolution intensity guide. Residual interpolation is designed to reduce the gradient energy of the residual. The residual is defined by the difference between the input and the tentative estimate. The main idea of the proposed algorithm is to generate such a tentative estimate from the input depth map itself. We designate this tentative estimate as the self-guided tentative estimate. Residual interpolation with the self-guided tentative estimate is designated as self-guided residual interpolation.

Fig. 3 shows the processing pipeline of the proposed self-guided residual interpolation. In the proposed algorithm, the input low-resolution depth map is upsampled by the displacement field algorithm [24]. Subsequently, we apply guided upsampling (GU) [9] to obtain the tentative estimate, where the depth map upsampled by the displacement field algorithm is used as the guide for the GU. The residual map is calculated by subtracting the tentative estimate from the input low-resolution depth map. Then, the bicubic interpolation is performed in the residual domain. The output high-resolution depth map is estimated by adding the tentative estimate to the interpolated residual. As discussed later, the self-guided tentative estimate generated from the input low-resolution depth map can reduce the gradient energy of the residual.

A. Self-Guided Tentative Estimate

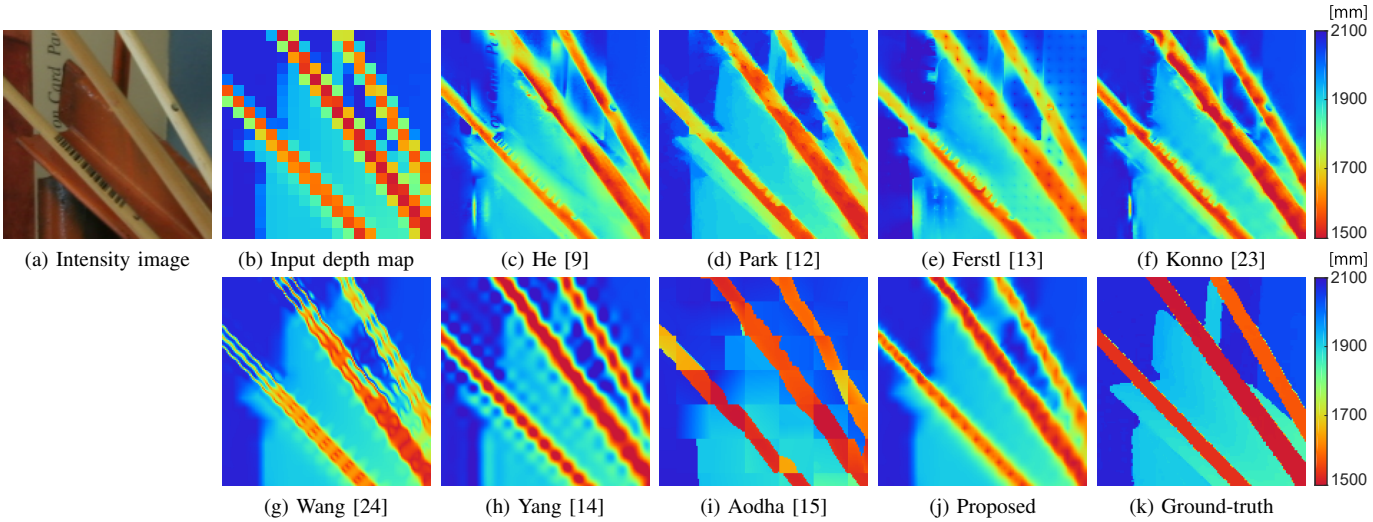
We verify the effectiveness of the self-guided tentative estimate experimentally. Hereinafter, we refer $N\times$ upsampling to upsampling with $N \times N$ scale factor. We conducted $4\times$ upsampling using the Middlebury dataset [21]. We calculated the gradient energy of the residual and the RMSE of the upsampled depth map, changing the tentative estimate generation algorithms. We compared four tentative estimate generation algorithms: the GU [9] with high-resolution intensity, bicubic, Wang's algorithm [24], and the combination of Wang's algorithm [24] and the GU [9] in Fig. 3. High-resolution intensity GU [9] requires guide information. The other three algorithms can generate tentative estimates only from the input low-resolution depth map.

Table I shows numerical comparisons of the gradient energy of the residual and the RMSE between the output depth map and the ground truth. Those values are averages of six scenes in the Middlebury dataset [21]. These comparisons demonstrate that the gradient energy of the residual with the tentative generated by the combination of Wang's algorithm [24] and the GU [9] is the lowest. We use this combination for the tentative estimation in the proposed method algorithm as shown in Fig. 3.

B. Gradual Upsampling

The upsampling operation becomes more accurate as the upsampling factor is decreased because the space of possible high-resolution solutions for each local low-resolution depth patch, and thus the amount of ambiguity, decreases. Glasner *et al.* [20] proposed upsampling by a gradual refinement in the small upsampling factor. Following their work, the proposed framework performs multiple operations of $2\times$ upsampling shown in Fig. 3.

Accuracy comparison among three upsampling steps at the same factor is conducted. We refer to the iterative four times $2\times$ upsampling as quadruple $2\times$ upsampling. The other upsampling steps are designated similarly. We compared quadruple $2\times$, double $4\times$, and single $16\times$ upsampling. We evaluated



(a) Intensity image (b) Input depth map (c) He [9] (d) Park [12] (e) Ferstl [13] (f) Konno [23]

(g) Wang [24] (h) Yang [14] (i) Aodha [15] (j) Proposed (k) Ground-truth

Fig. 4. Visual comparison of $8\times$ upsampling on a part of *Art* from the Middlebury dataset: (a) intensity guidance image, (b) input low-resolution depth map (enlarged by nearest neighbor upsampling), (c) guided image filtering [9], (d) non-local means upsampling [12], (e) anisotropic total generalized variation [13], (f) intensity-guided residual interpolation, (g) displacement field [24], (h) sparse coding [14], (i) dictionary-based method [15], (j) proposed, and (k) ground-truth.

TABLE I. RELATION BETWEEN GRADIENT ENERGY OF THE RESIDUAL AND OUTPUT ACCURACY ($4\times$ UPSAMPLING).

Tentative estimate generation algorithm	Intensity-guided	Self-guided		
	GU [9]	Bicubic	Wang [24]	Wang [24]+GU [9]
Gradient energy [$\times 10^{-3}$]	1.66	1.49	1.63	**1.44**
RMSE of output [mm]	14.3	14.2	14.6	**13.9**

TABLE II. COMPARISON OF UPSAMPLING STEPS ($16\times$ UPSAMPLING).

Upsampling algorithm	Wang [24]	Self-guided residual interpolation
quadruple $2\times$	32.01	**26.89**
double $4\times$	40.72	29.44
single $16\times$	42.98	31.33

two upsampling algorithms, displacement field upsampling by Wang [24] and the proposed framework. Table II shows average RMSE of $16\times$ upsampling results on the Middlebury dataset [21]. This numerical comparison supports that the gradual upsampling approach is also effective for depth map upsampling.

IV. EXPERIMENTAL RESULTS

We demonstrate quantitative and qualitative evaluations of the proposed depth upsampling algorithm[1]. We first examined the numerical performance compared to state-of-the-art algorithms with the Middlebury 2005 dataset [21]. Next, we provide a visual comparison using actual sensor data acquired using a time-of-flight (ToF) sensor. We performed hole-filling to obtain the ground-truth depth map in all our experiments because a depth map generally has the lack of depth values due to occlusions and various internal errors of the acquisition system. This preprocessing is achieved simply by propagating the hole contour toward the inner side using the median value of neighboring pixels. The hole-filled depth map is treated as the ground-truth.

[1]The code is available on http://www.ok.ctrl.titech.ac.jp/res/DSR/SGRI/

A. Middlebury dataset

Quantitative and qualitative evaluations of the proposed method were conducted using the Middlebury 2005 dataset [21] for *Art, Books, Dolls, Laundry, Moebius,* and *Reindeer*. Considering that the proposed algorithm is applied to the ToF sensor, we use depth maps converted from the disparity form as ground-truth. Original RGB intensity images were used as the guide image for intensity-guided algorithms. We evaluated the low-resolution depth maps with three downsampling factors $(2\times, 4\times, 8\times)$ generated by downsampling after applying a Gaussian filter.

Visual comparison of the $8\times$ upsampling evaluation is shown in Fig. 4. In Fig. 4, (c) He [9], (d) Park [12], (e) Ferstl [13], and (f) Konno [23] are intensity guided approach. Single depth map upsampling algorithms are (g) Wang [24], (h) Yang [14], (i) Aodha [15], and (j) proposed. Among results of the intensity guide approach, one can find edge-bleeding artifacts which texture transfer from the high-resolution intensity guide causes, especially in regions of stain and bar codes on paint brushes. In contrast, no texture transfer artifact exists in the results of the single depth map upsampling algorithm. The result of the proposed algorithm appears to be the most natural among the results of the single depth map upsampling algorithm.

Table III shows numerical results of these experiments in terms of the root mean square error (RMSE) and edge-RMSE (E-RMSE) [25], structural similarity (SSIM) [26]. To measure E-RMSE, we first detect edges using the Canny edge detector from the ground truth depth map. Then we extract the edge region by dilating the edge map. The standard deviation of isotropic Gaussian function used in SSIM is set to 4 so that

TABLE III. QUANTITATIVE COMPARISON ON THE MIDDLEBURY DATASET 2005 [21]

		Art			*Dolls*			*Laundry*			*Moebius*			Average of six scenes		
RMSE [mm]		2×	4×	8×	2×	4×	8×	2×	4×	8×	2×	4×	8×	2×	4×	8×
Intensity-guided approach	He [9]	21.80	28.35	39.27	7.12	8.84	11.60	15.91	21.21	28.53	8.51	11.24	16.15	13.34	17.41	23.89
	Park [12]	20.93	27.80	34.84	7.78	8.88	10.78	14.96	20.16	24.54	8.01	9.92	12.45	12.92	16.69	20.65
	Ferstl [13]	19.57	26.76	52.31	6.95	9.05	36.63	16.51	23.84	56.95	7.61	10.23	38.85	12.66	17.47	46.18
	Konno [23]	18.26	23.31	31.17	6.64	8.14	**10.30**	13.41	16.71	22.60	7.45	9.48	13.09	11.44	14.41	19.29
Single depth map upsampling	Wang [24]	**14.08**	25.39	47.07	6.43	8.69	15.16	10.96	18.81	32.56	6.33	10.04	19.01	9.45	15.73	28.45
	Yang [14]	14.74	27.48	46.21	**5.73**	10.16	16.42	**10.63**	20.38	33.87	**5.81**	11.42	19.76	**9.23**	17.36	29.06
	Aodha [15]	20.04	33.61	57.44	12.15	13.59	21.23	20.78	26.59	42.75	17.21	14.86	25.16	17.55	22.16	36.65
	proposed	15.69	**19.59**	**30.01**	6.39	**8.03**	10.86	11.62	**14.88**	**21.31**	6.73	**8.53**	**12.09**	10.11	**12.76**	**18.57**
E-RMSE [mm]		2×	4×	8×	2×	4×	8×	2×	4×	8×	2×	4×	8×	2×	4×	8×
Intensity-guided approach	He [9]	21.98	28.60	39.60	7.16	8.90	11.65	15.14	20.37	27.41	8.54	11.27	16.01	13.21	17.29	23.67
	Park [12]	21.10	28.04	35.15	7.82	8.94	10.83	14.60	19.77	24.16	8.06	9.94	12.45	12.89	16.67	20.65
	Ferstl [13]	19.72	27.00	52.69	7.00	9.05	36.50	16.04	23.34	56.69	7.65	10.24	38.65	12.60	17.41	46.13
	Konno [23]	18.41	23.51	31.41	6.68	8.18	**10.35**	12.74	15.83	21.05	7.49	9.51	13.08	11.33	14.26	18.97
Single depth map upsampling	Wang [24]	**14.18**	25.58	47.47	6.48	8.73	15.21	10.15	18.10	31.56	6.35	10.04	19.04	9.29	15.61	28.32
	Yang [14]	14.83	27.61	46.55	**5.20**	9.90	16.37	**9.97**	19.27	32.64	**5.74**	11.23	19.73	**8.93**	17.00	28.82
	Aodha [15]	19.77	33.78	57.83	12.24	13.66	21.22	20.60	25.59	41.31	17.23	14.85	25.07	17.46	21.97	36.36
	proposed	15.80	**19.73**	**30.25**	6.43	**8.09**	10.92	10.88	**13.92**	**20.42**	6.75	**8.54**	**12.10**	9.97	**12.57**	**18.42**
SSIM		2×	4×	8×	2×	4×	8×	2×	4×	8×	2×	4×	8×	2×	4×	8×
Intensity-guided approach	He [9]	0.79	0.66	0.49	0.86	0.76	0.61	0.79	0.65	0.48	0.79	0.65	0.46	0.81	0.68	0.51
	Park [12]	0.64	0.55	0.45	0.60	0.61	0.51	0.65	0.55	0.47	0.67	0.58	0.49	0.64	0.57	0.48
	Ferstl [13]	0.79	0.60	0.21	0.86	0.71	0.17	0.73	0.49	0.12	0.79	0.56	0.15	0.79	0.59	0.16
	Konno [23]	0.86	0.76	0.58	0.90	0.83	0.68	0.86	0.75	0.56	0.86	0.75	0.56	0.87	0.77	0.60
Single depth map upsampling	Wang [24]	0.88	0.77	0.55	0.90	0.82	0.63	0.87	0.76	0.53	0.87	0.76	0.54	0.88	0.78	0.56
	Yang [14]	0.85	0.67	0.43	0.89	0.76	0.54	0.85	0.67	0.41	0.86	0.67	0.41	0.86	0.69	0.45
	Aodha [15]	0.61	0.48	0.33	0.62	0.48	0.34	0.45	0.32	0.19	0.48	0.37	0.24	0.54	0.41	0.27
	proposed	**0.88**	**0.79**	**0.63**	**0.91**	**0.84**	**0.70**	**0.88**	**0.78**	**0.59**	**0.87**	**0.78**	**0.61**	**0.88**	**0.80**	**0.63**

this measurement can evaluate the similarity of semi-global structure. In Table III, the methods listed above the dashed line are intensity-guided methods. Due to the limit of the space, comparisons of four scenes and average are summarized. We put the average of these numerical results over six scenes in the right end column. The best result for each dataset and upsampling factor are emphasized in bold typeface.

The lower RMSE and E-RMSE show better results. Comparing the average performances of the RMSE and the E-RMSE in $4\times$ and $8\times$ upsampling, the proposed algorithm outperforms state-of-the-art algorithms including the intensity guided approach. Those comparisons demonstrate that the proposed algorithm is extremely effective for the higher upsampling factor cases. The higher SSIM value signifies a better result. In terms of the SSIM, the proposed algorithm invariably shows the best performance for all scenes.

B. Visual comparison on real sensor data

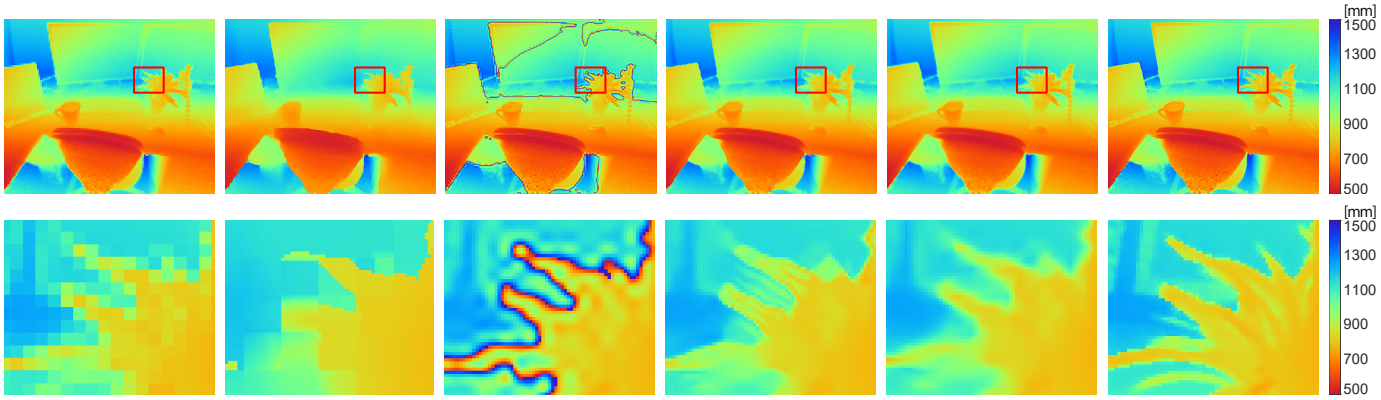
Visual comparison on real sensor data can be conducted using a depth map acquired with a ToF sensor. For depth acquisition, we use Kinect v2 that provides registered 512×424 dense depth and IR intensity maps. We use the depth map for ground-truth and produce an input low-resolution depth map by downsampling the ground-truth. The $4\times$ upsampling results are shown in Fig. 5. The associated intensity map and the experimental setup are shown in Fig. 6. This comparison demonstrates that the proposed algorithm can upsample the depth map of real scene effectively, although Aodha's [15] and Yang's [14] algorithms yield severe artifacts.

V. CONCLUSION

We have proposed a simple and effective single depth map upsampling algorithm based on self-guided residual interpolation. The key to the proposed algorithm is interpolation of the residual domain, where the residual is defined as the difference between the tentative estimate and the input low-resolution depth map. In the proposed algorithm, the tentative estimate is also generated, but from the input low-resolution depth map only. Therefore, the proposed algorithm can upsample using the single low-resolution depth map. Experimentally obtained results demonstrate that the proposed algorithm outperforms state-of-the-art algorithms, even including intensity-guided methods in terms of both quantitative and qualitative evaluations.

REFERENCES

[1] D. Lin, S. Fidler, and R. Urtasun, "Holistic scene understanding for 3d object detection with rgbd cameras," *IEEE International Conference on Computer Vision (ICCV)*, pp. 1417–1424, 2013.



(a) Input (b) Aodha [15] (36.6) (c) Yang [14] (72.6) (d) Wang [24] (27.4) (e) proposed (19.9) (f) Ground-truth

Fig. 5. Visual comparison of $4\times$ upsampling on a real sensor data acquired by Microsoft kinect v2. The scores following each method names indicate the RMSE ([mm]) values of them. (a) Input low-resolution depth map (enlarged by nearest neighbor upsampling), (b) Dictionary-based method [15], (c) Sparse cording [14], (d) Displacement field [24], (e) Proposed, (f) Ground-truth. The results in (b) – (d) suffer from the block or waving artifact. The proposed method better reconstructs the smooth surfaces.

(a) (b)

Fig. 6. Circumstances to acquire depth map. (a) the associated IR intensity map, (b) the experimental setup

[2] J. Shotton, T. Sharp, A. Kipman, A. Fitzgibbon, M. Finocchio, A. Blake, M. Cook, and R. Moore, "Real-time human pose recognition in parts from single depth images," *Communications of the ACM*, vol. 56, no. 1, pp. 116–124, 2013.

[3] D. Holz, R. Schnabel, D. Droeschel, J. Stückler, and S. Behnke, "Towards semantic scene analysis with time-of-flight cameras," *RoboCup 2010: Robot Soccer World Cup XIV*, pp. 121–132. Springer, 2011.

[4] S. Foix, G. Alenya, and C. Torras, "Lock-in time-of-flight (tof) cameras: a survey," *IEEE Sensors Journal*, vol. 11, no. 9, pp. 1917–1926, 2011.

[5] Y. Cui, S. Schuon, D. Chan, S. Thrun, and C. Theobalt, "3d shape scanning with a time-of-flight camera," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1173–1180, 2010.

[6] S. Schuon, C. Theobalt, J. Davis, and S. Thrun, "Lidarboost: Depth superresolution for tof 3d shape scanning," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 343–350, 2009.

[7] J. Kopf, M. F. Cohen, D. Lischinski, and M. Uyttendaele, "Joint bilateral upsampling," *ACM Transactions on Graphics (TOG)*, vol. 26, no. 3, pp. 96, 2007.

[8] Q. Yang, R. Yang, J. Davis, and D. Nistér, "Spatial-depth super resolution for range images," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–8, 2007.

[9] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, vol. 35, no. 6, pp. 1397–1409, 2013.

[10] M.-Y. Liu, O. Tuzel, and Y. Taguchi, "Joint geodesic upsampling of depth images," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 169–176, 2013.

[11] J. Diebel and S. Thrun, "An application of markov random fields to range sensing," *Conference on Neural Information Processing Systems (NIPS)*, vol. 5, pp. 291–298, 2005.

[12] J. Park, H. Kim, Y.-W. Tai, M. S. Brown, and I. Kweon, "High quality depth map upsampling for 3d-tof cameras," *IEEE International Conference on Computer Vision (ICCV)*, pp. 1623–1630, 2011.

[13] D. Ferstl, C. Reinbacher, R. Ranftl, M. Rüther, and H. Bischof, "Image guided depth upsampling using anisotropic total generalized variation," *IEEE International Conference on Computer Vision (ICCV)*, pp. 993–1000, 2013.

[14] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Transactions on Image Processing (TIP)*, vol. 19, no. 11, pp. 2861–2873, 2010.

[15] O. Mac Aodha, N. D. Campbell, A. Nair, and G. J. Brostow, "Patch based synthesis for single depth image super-resolution," *IEEE European Conference on Computer Vision (ECCV)*, pp. 71–84. Springer, 2012.

[16] M. Hornáček, C. Rhemann, M. Gelautz, and C. Rother, "Depth super resolution by rigid body self-similarity in 3d," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1123–1130, 2013.

[17] D. Kiku, Y. Monno, M. Tanaka, and M. Okutomi, "Residual interpolation for color image demosaicking," *IEEE International Conference on Image Processing (ICIP)*, pp. 2304–2308, 2013.

[18] D. Kiku, Y. Monno, M. Tanaka, and M. Okutomi, "Minimized-laplacian residual interpolation for color image demosaicking," *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, pp. 90230L–90230L, 2014.

[19] D. Kiku, Y. Monno, M. Tanaka, and M. Okutomi, "Beyond color difference: residual interpolation for color image demosaicking," *IEEE Transactions on Image Processing (TIP)*, vol. 25, no. 3, pp. 1288–1300, 2016.

[20] D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," *IEEE International Conference on Computer Vision (ICCV)*, pp. 349–356, 2009.

[21] D. Scharstein and C. Pal, "Learning conditional random fields for stereo," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–8, 2007.

[22] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," *IEEE International Conference on Computer Vision (ICCV)*, pp. 839–846, 1998.

[23] Y. Konno, Y. Monno, D. Kiku, M. Tanaka, and M. Okutomi, "Intensity guided depth upsampling by residual interpolation," *JSME/RMD International Conference on Advanced Mechatronics (ICAM)*, pp. 1–2, 2015.

[24] L. Wang, H. Wu, and C. Pan, "Fast image upsampling via the displacement field," *IEEE Transactions on Image Processing (TIP)*, vol. 23, no. 12, pp. 5123–5135, 2014.

[25] J. Xie, R. S. Feris, S.-S. Yu, and M.-T. Sun, "Joint super resolution and denoising from a single depth image," *IEEE Transactions on Multimedia*, vol. 17, no. 9, pp. 1525–1537, 2015.

[26] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing (TIP)*, vol. 13, no. 4, pp. 600–612, 2004.