Discontinuity adaptive depth upsampling for 3D video acquisition

S. B. Lee, S. Kwon and Y. S. Ho

A new discontinuity adaptive depth upsampling method is proposed in order to obtain a high-resolution depth map. The proposed method adaptively employs both the local minimum filter and the joint bilateral filter according to the variance of the depth map. Thus, it can reduce the over-blurred depth regions along the depth discontinuities. The experimental results demonstrate the feasibility of the proposed method.

Introduction: The acquisition of reliable depth maps is a critical requirement in many applications such as 3D videos and free-viewpoint TV. Time-of-flight (TOF) cameras are widely used to obtain depth maps in real time; however, they have limitations in practical applications due to their low resolutions. Recently, filter-based depth upsampling algorithms such as joint bilateral upsampling (JBU) [1] and noise-aware filter for depth upsampling (NAFDU) [2] have been proposed in order to overcome this problem. They can reconstruct depth edges requiring only a small memory space and low complexity. However, these methods often lead to over-blurred depth regions near the depth discontinuities. Although other approaches based on the Markov random field (MRF) have been presented for solving the over-blurred depth problem [3, 4], they require significant computational and memory costs. In this Letter, a new upsampling method with low computational and memory costs is presented for the reduction of the over-blurred depth regions.

Discontinuity adaptive depth upsampling: In filter-based depth upsampling algorithms, the upsampled depth map is obtained via the Gaussian-weighted sum of neighbours within the filter kernel. Thus, it tends to be over-blurred along the depth discontinuity. In this Letter, a filter-based depth upsampling algorithm is proposed that considers the characteristics of depth discontinuity in the reduction of the over-blurred depth problem. Fig. 1 presents the block diagram of the proposed method. The primary aim of this study is to conduct the local minimum filter near the depth discontinuity and the conventional JBU filter in homogeneous regions, without incurring significant computational and memory costs. This concept that adaptively utilizes the filter according to the variance of the depth map is similar to that presented in NAFDU; however, NAFDU does not reduce the over-blurred depth regions, although it solves the texture copying problem.

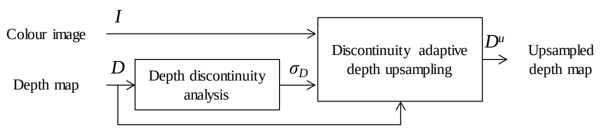


Fig. 1 *Block diagram of proposed depth upsampling*

The proposed discontinuity adaptive depth upsampling consists of two filters: a local minimum filter and a JBU filter. Let p and q denote the pixel coordinates in a high-resolution colour image I and an upsampled depth map D^u, while $p_\downarrow$ and $q_\downarrow$ represent the corresponding pixel coordinates in the low-resolution depth map D. The upsampled depth value $D^u{}_p$ is computed using (1):

$$D^u{}_p = \begin{cases} \arg\min\limits_{d\in\Omega} \left| d - D^u{}_{JBU,p} \right| & \text{if } \sigma_{D_{p_\downarrow}}{}^2 \geq Th_D \\ D^u{}_{JBU,p} & \text{otherwise} \end{cases} \quad (1)$$

$$D^u{}_{JBU,p} = \frac{\sum\limits_{q_\downarrow\in\Omega} D_{q_\downarrow} \cdot f\left(\left\| p_\downarrow - q_\downarrow \right\|\right) g\left(\left\| I_p - I_q \right\|\right)}{\sum\limits_{q_\downarrow\in\Omega} f\left(\left\| p_\downarrow - q_\downarrow \right\|\right) g\left(\left\| I_p - I_q \right\|\right)} \quad (2)$$

where $\sigma_D{}^2$ represents the normalized variance map using the filter kernel Ω. According to $\sigma_D{}^2$ and the threshold Th_D, the local minimum filter and JBU filter are adaptively conducted. The JBU filter is defined as in (2), where f and g respectively represent the spatial filter and range filter that have Gaussian distribution. A set of depth candidates d represents the valid depth values within Ω. Instead of using a Gaussian-weighted sum, the local minimum filter determines one of the depth candidates using the difference between $D^u{}_{JBU,p}$ and d. Fig. 2 shows the method for obtaining the upsampled depth value using the local minimum filter. As shown in Fig. 2, $D^u{}_p$ is near depth discontinuity and the 5×5 filter kernel Ω includes four valid depth candidates. Therefore, $D^u{}_p$ is determined using one of them.

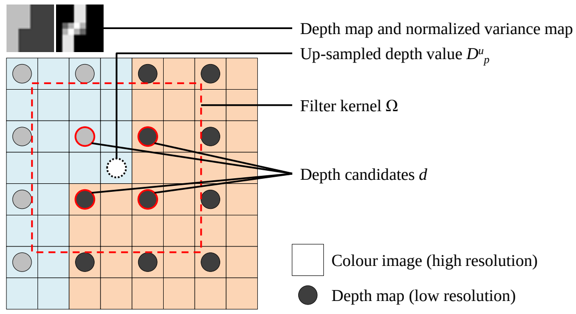


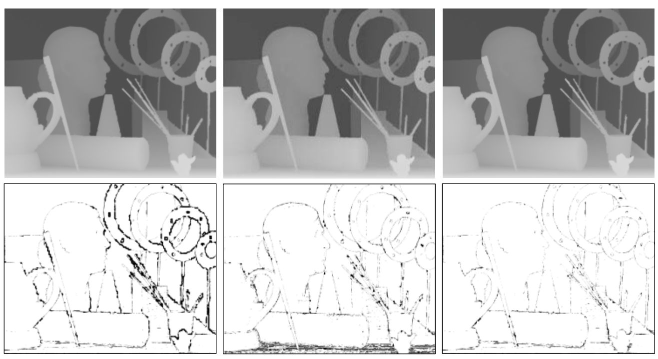
Fig. 2 *Depth upsampling using local minimum filter*

The proposed algorithm has the following two advantages over other algorithms. First, it reduces the over-blurred depth regions because the local minimum filter does not create new depth values when the filter kernel straddles the depth discontinuity. Second, the proposed filter-based depth upsampling method requires only small computational and memory costs compared with the conventional algorithms.

Experimental Results: The proposed algorithm was tested using three Middlebury datasets with nearest neighbour downsampling [5]: *Art*, *Books*, and *Moebius*. The proposed method is compared with the JBU [1], NAFDU [2], and MRF-based methods [4]. In order to evaluate the upsampled depth maps, the bad pixel rate (BPR), whose absolute difference is greater than 1, was used. The kernel size Ω and the threshold Th_D were fixed as 3×3 and 0.01, respectively.

Table 1: Bad pixel rate (%)

Dataset	Scale	JBU	NAFDU	MRF	Proposed
Art	2×	2.083	2.084	0.870	0.468
	4×	3.220	3.218	2.160	0.915
	8×	5.745	5.697	4.140	2.095
	16×	11.501	11.399	8.920	7.850
Books	2×	0.665	0.663	0.500	0.389
	4×	1.318	1.308	1.360	0.945
	8×	2.699	2.694	2.570	2.203
	16×	6.206	6.185	6.250	5.969
Moebius	2×	1.370	1.371	0.850	0.765
	4×	2.431	2.413	1.730	1.503
	8×	4.657	4.622	3.300	3.118
	16×	9.195	9.140	6.880	6.679



a *b* *c*

Fig. 3 *8× magnification upsampled depth maps and error maps for 'Art'*
a Joint bilateral up-sampling
b MRF-based up-sampling
c Proposed method

Table 1 provides a quantitative evaluation of the conventional and proposed methods. According to the results, the proposed algorithm outperforms the conventional algorithms in terms of BPR. Fig. 3 presents the upsampled depth maps (row 1) and error maps with respect to BPR (row 2) for '*Art*' with a scale factor of 8. As shown in Fig. 3, the proposed method reduces the depth errors, especially near the depth discontinuity, compared with the conventional methods. Table 2 represents the average running time of each method. The complexity of the proposed method is almost the same as that of the filter-based methods, whereas the depth quality is improved.

Table 2: Average running time (sec.)

Scale	JBU	NAFDU	MRF	Proposed
4×	0.833	1.069	159.940	0.943
16×	7.142	7.314	334.193	7.270

Another experiment was conducted in order to fill out the black areas of the depth map captured by a Kinect depth camera. Note that the black areas were caused by an inherent infrared sensor problem. Fig. 4 shows the raw and filtered depth maps. The input images are shown in Figs. 4*a* and *b*. The filtering results for the JBU and the proposed method are shown in Figs. 4*c* and *d*, respectively. It can be confirmed that the proposed algorithm efficiently reconstructs the depth discontinuity and reduces both sensor errors and over-blurred depth regions compared with the JBU.

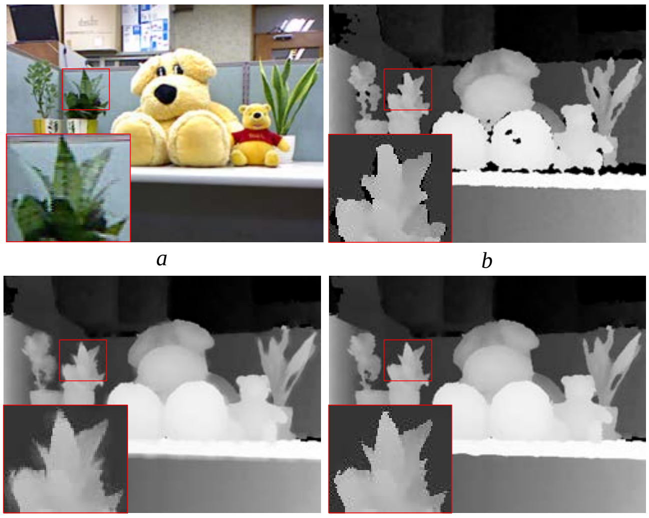


Fig. 4 *Filtered depth maps from a Kinect depth camera*
a Colour image
b Raw depth map
c Joint bilateral upsampling
d Proposed method

Conclusion: In this Letter, a new simple but powerful depth upsampling method was proposed in order to obtain a high-resolution depth map. As demonstrated through experiments, the proposed discontinuity adaptive depth upsampling method can effectively reduce over-blurred depth regions and improve the depth quality with low computational and memory costs, compared with the conventional algorithms.

Acknowledgments: This work was supported by the DGIST R&D Program of the Ministry of Education, Science and Technology, Korea (13-NB-05). It was also supported by a National Research Foundation of Korea (NRF) grant funded by the Korean government (MEST) (No. 2012-0009228).

S. B. Lee and S. Kwon (*Daegu Gyeongbuk Institute of Science and Technology, 50-1 Sang-ri, Hyeonpung-myeon, Dalseong-gun, Daegu, Republic of Korea*)

E-mail: sangbeomlee@dgist.ac.kr

S. B. Lee and Y. S. Ho (*Gwangju Institute of Science and Technology, 123 Cheomdangwagi-ro, Buk-gu, Gwangju, Republic of Korea*)

References

1 Kopf, J., Cohen, M.F., Lischinski, D., and Uyttendaele, M.: 'Joint bilateral upsampling', *ACM Transactions on Graphics*, 2007, **26**, (3), pp. 1–5

2 Chan, D., Buisman, H., Theobalt, C., and Thrun, S.: 'A noise-aware filter for real-time depth upsampling', Proc. *ECCV Workshop on Multi-camera and Multi-modal Sensor Fusion Algorithms and Applications*, 2008, pp. 1–12

3 Park, J., Kim, H., Brown, M. S., and Kweon, I.: 'High quality depth map upsampling for 3D-TOF cameras', Proc. *International Conference on Computer Vision*, 2011, pp. 1623–1630

4 Kim, D., and Yoon, K.: 'High-quality depth map up-sampling robust to edge noise of range sensors', Proc. *International Conference on Image Processing*, 2012, pp. 553–556

5 Scharstein, D., and Szeliski, R.: 'A taxonomy and evaluation of dense two-frame stereo correspondence algorithms', *International Journal of Computer Vision*, 2002, **47**, (1), pp. 7–42