Upsampling of Low-resolution Depth Map with Enhancing Depth Discontinuity Regions

Yun-Suk Kang and Yo-Sung Ho
Gwangju Institute of Science and Technology (GIST)
Gwangju, 500-712, Republic of Korea
E-mail: {yunsuk, hoyo}@gist.ac.kr

***Abstract*— In this paper, we present an upsampling method of low-resolution depth maps with enhancing depth discontinuities using color segment information. After we supply the initial depth measurement considering the corresponding color segment information, we define an energy function for depth map upsampling based on the depth measurement, color values, and color segments. Then, we obtain high-resolution depth maps by by belief propagation optimization. Experimental results show that the proposed method outperforms other approaches for depth map upsampling in terms of the bad pixel rate and mean absolute error.**

I. INTRODUCTION

Recently, three-dimensional (3D) video has been actively investigated as the next generation broadcasting and imaging services [1]. The 3D video provides immersive sense using images from dense and multiple viewpoints. Also, the increasing number of viewpoint allows users to experience the free-viewpoint navigation in the scene [2].

However, it is hard to build a camera system that captures such dense multi-view videos in practice. To solve this difficulty, depth image-based rendering (DIBR) has been proposed [3]. The DIBR technique synthesizes images at the desired viewpoint using captured color images and corresponding depth map. Thus, it can be an efficient data format for 3D video. The depth map is an image that represents the range information of the captured scene. The depth map is important because it affects the quality of the synthesized images.

Acquisition of the depth map is categorized into two parts; estimating based on the captured color images and measuring directly using depth sensors. Stereo matching or 2D-to-3D conversion are representative approaches to acquire the depth map based on images. However, they need high complexity, and often fails to estimate an accurate depth due to the image characteristics such as textureless or occluded regions.

Recently, Depth sensors have been frequently used for various 3D applications. Using the Time-of-Flight (TOF) technique, depth sensors measure the range from the sensor to the object, and generate the depth map in real-time. However, one of weak points of current depth sensors is a small output resolution. Therefore, upsampling method for the depth map from depth sensors are required since the color image resolution becomes larger nowadays.

In order to upsample the low-resolution depth map, filter-based approaches have been proposed. Kopf *et al.* proposed joint bilateral upsampling (JBU) using the bilateral filter [4]. JBU utilizes color information to enhance depth discontinuity regions. Yang *et al.* presented an iterative bilateral filter for depth upsampling [5]. Also, a noise-aware filter for depth upsampling, proposed by Chan *et al.*, uses color information adaptively [6]. However, these filter-based methods cause some problems such as blurring depth discontinuities and appearing object shapes in depth homogeneous regions due to referring to color information.

In addition, there have been approaches using Markov random field (MRF) for depth upsampling. Diebel *et al.* defined the probability model by MRF using color and depth data, and the model is optimized by conjugate gradient to provide high-resolution depth map [7]. Park *et al.* proposed an energy function including non-local mean filtering (NLM) term and the confidence weighting for smoothness term [8]. These MRF-based approaches, however, error propagation problem can be caused at depth discontinuity regions during optimization process.

In this paper, we propose an upsampling method of low-resolution depth map using MRF model. The proposed method supplies the depth initial measurement for the data term of the energy function by considering the corresponding color segments. Also, the proposed method considers the color values and color segments for the smoothness term of the energy function to enhance the depth discontinuity. The belief propagation (BP) optimization calculates the best depth values and provides the high-resolution depth map.

II. MARKOV RANDOM FIELD MODEL FOR DEPTH MAP UPSAMPLING

We denote images and pixels for depth map upsampling based on MRF as follows. High-resolution color image and low-resolution depth map are denoted by I and D_L. Upsampled depth map is indicated as D_U. Each pixel of D_U is represented as y, and z means mapped pixel to I from D_L. Then, an energy function E for the depth map upsampling is defined as Eq. (1).

978-616-361-823-8

$$\mathrm{E} = \sum_{i \in L} k(y_i - z_i)^2 + \sum_i \sum_{j \in N(i)} w_{ij}(y_i - y_j)^2 \qquad (1)$$

The energy function is composed of the data and smoothness terms, and the subscripts i and j mean the pixel positions. The data term is calculated if ith pixel is in L which is a set of pixels who have z_i value. The constant k controls the influence of the data term.

The current pixel is indicated as i, and $N(i)$ means four-neighboring pixels around i. The weight w_{ij} for the smoothness term of the energy function is defined by the color difference between two adjacent pixels i and j, as Eq. (2), in general.

$$w_{ij} = \exp\left(-\frac{\sqrt{\sum_{RGB}\big(I(i) - I(j)\big)^2}}{2\sigma_{ij}^2}\right) \qquad (2)$$

There are a few weaknesses of MRF-based depth map upsampling represented in Eq. (1) and Eq. (2). Since data term is calculated only for the pixels belong to L, there is a high possibility that y_i become equal to z_i. Although it is not a problem when color and depth are colocated, error caused by 3D warping in practical case such as using a depth sensor can be propagated. In addition, since w_{ij} is calculated by color difference, problems as following cases could be caused; discordance between depth discontinuity and color edge regions, and blurred color edge regions. In these cases, error also can be propagated.

III. Upsampling of Low-resolution Depth Maps using Color Segment Information

In this section, we explain the proposed method to up-sample the low-resolution depth map. Figure 1 shows the overall process of the proposed method. We use color segment information to overcome the aforementioned weaknesses and problems of the previous depth map upsampling methods.

Color segmentation is the process that divides the image into regions that have similar characteristics. In general, the region with the similar color characteristics by color segmentation has not a radical depth change or depth discontinuity but smooth depth change or the same depth. Also, the boundary between two adjacent segments has a possibility to be the depth discontinuity

I is segmented by using mean-shift color segmentation algorithm [9], and each pixel of D_L is mapped to the corresponding pixel position. Then, we define the energy function for depth map upsampling. For the data term, we supply more depth measurements with considering neighboring conditions. After that, we calculate the weight for smoothness term with reflecting color segment information which is highly related to the depth discontinuity regions. Finally, the energy function is optimized by belief propagation, and we obtain D_U.

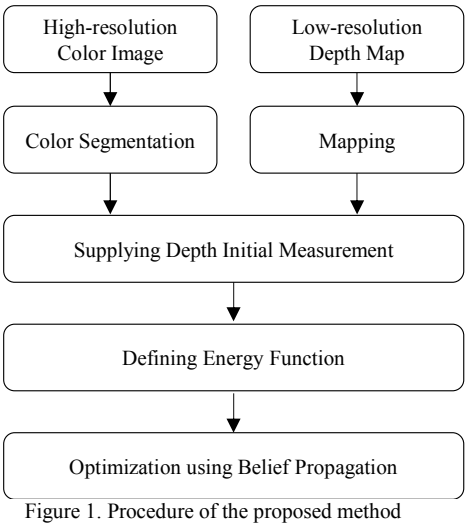


Figure 1. Procedure of the proposed method

A. Depth Initial Measurement for Data Term

Figure 2 shows the each pixel of D_L is mapped to the corresponding position of I. There exists holes between zs due to the resolution difference. For the data term calculation, z value plays an important role because y_i, the depth value at ith pixel, could be decided to z_i which is the correct value for that position with high possibility. Therefore, correct depth values z at correct positions increase the performance of depth upsampling.

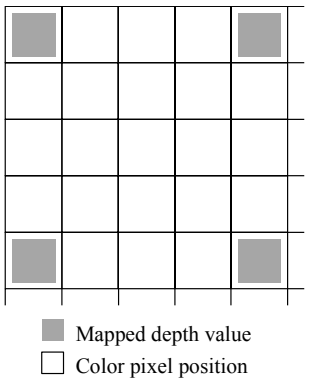


Figure 2. Mapping of D_L to I

In the proposed method, we supply additional depth measurements z' using z and color segment information. We perform the supplement once for 4-neighbor and 2-neighbor cases as shown in Fig. 3(a) and Fig. 3(b), respectively. To supply z'_i at ith pixel, the neighboring z values must be the same and all zs and z'_i have to be in the same color segment. Then, the z'_i value at ith pixel becomes the same as neighboring z values. After supplying, we calculate the data term shown in Eq. (1) using z and z'.

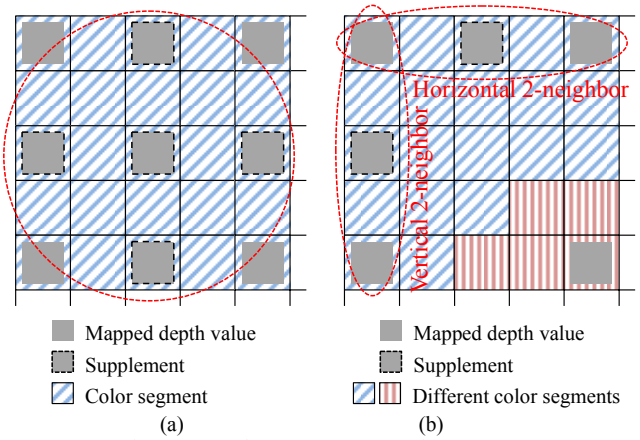


Figure 3. Depth measurement supplement:
(a) 4-neighbor supplement (b) 2-neighbor supplement

B. *Calculation of Weight for Smoothness Term*

As introduced in Section II, the influence of smoothness term is decided by the weight w_{ij}. For accurate depth upsampling, the influence of w_{ij} should be strong for depth homogeneous regions and weak for depth discontinuous regions. Thus, as Eq. (3), we define w'_{ij} as a multiple of two weights considering color values and color segment information: w_{ij} for color similarity and $w_{s,ij}$ for color segment. These two weights controls the influence of smoothness cost according to image local features.

$$w'_{ij} = w_{ij} w_{s,ij} \tag{3}$$

Firstly, the color similarity weight is defined as Eq. (2). The large color difference between two adjacent pixels decrease the smoothness cost. While, if the color difference between two pixels are relatively small, there is a high opportunity to have the same depth value.

Secondly, the color segment weight $w_{s,ij}$ is defined. Two adjacent pixels i and j are likely to have the same depth value, and vice versa. Eq. (4) describes this relation between the color segment and depth. In Eq. (4), S(·) means the segment index of a pixel, and C_{seg} is a constant between 0 and 1. If two adjacent pixels are not in the same segment, the smoothness cost decreases as $w_{s,ij}$ decreases.

$$w_{s,ij} = \begin{cases} 1 & if\ S(i) = S(j) \\ C_{seg} & otherwise \end{cases} \tag{4}$$

C. *Optimization*

The energy function shown in Eq. (1) using w'_{ij} instead of w_{ij} is then optimized by using belief propagation method [10]. As described in previous subsections, the data term calculation includes both z and z', and the weight for smoothness term is computed as Eq. (3) to consider the depth discontinuity.

IV. Experimental Results

We have experimented the proposed method on three test image sets provided by Middlebury Stereo. They are "Cones," "Teddy," "Venus," and "Tsukuba." These data sets are composed of color images and ground-truth depth maps as shown in Fig. 4.

Figure 4. Image sets for experiment: "Cones," "Teddy," and "Venus,"

We downsampled the ground-truth depth maps by factors of 2, 4, and 8, and then performed the proposed method to these downsampled depth maps. During the experiments, parameters were fixed as k=1 and C_{seg}=0.2. The upsampled depth maps by the proposed method using these parameters are shown in Fig. 5. Also, Fig. 6 shows the upsampled depth maps by the other approaches compared to the proposed method.

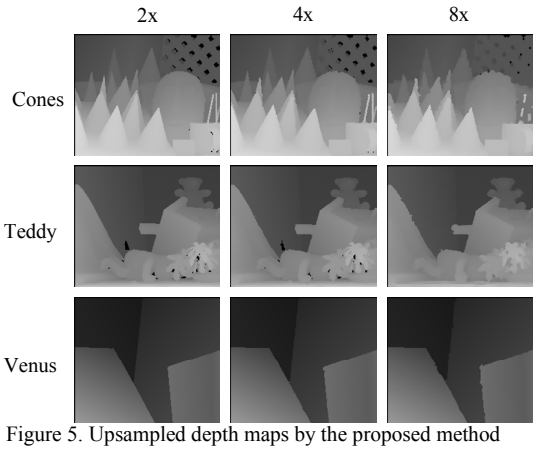
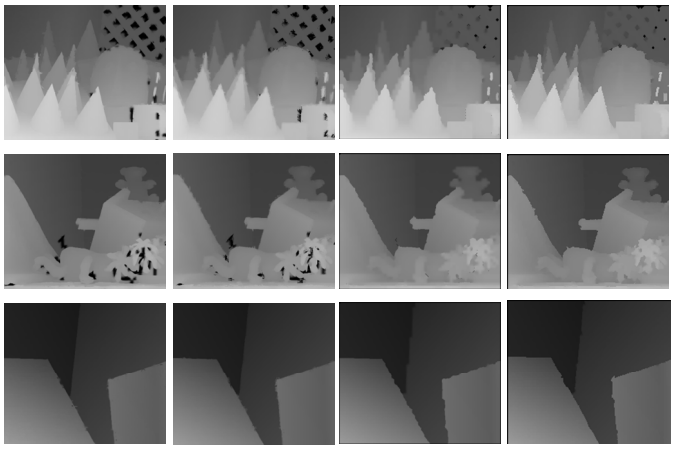


Figure 5. Upsampled depth maps by the proposed method

As shown in Fig. 5 and Fig 6, the proposed method upsampled the depth maps efficiently as compared with their ground-truth depth maps. Although the scaling factor doubled, almost object shapes and boundaries were maintained. In addition, blurring artifacts at depth discontinuity region were not appeared in the results of our method.

To evaluate the performance objectively, we compared the proposed method with four depth upsampling approaches: bilinear interpolation, JBU [4], NAFDU [6], and MRF-based upsampling [7]. We used two evaluation criteria: bad-pixel rate (BPR) and mean absolute difference (MAD) with respect to the ground-truth depth maps. The comparison results are summarized in Table I and Table II for each scaling factor.



JBU NAFDU MRF Proposed

Figure 6. Comparison with other approaches

TABLE I
COMPARISON OF BAD-PIXEL RATE

Scaling factor	Image	Upsampling methods				
		Bilinear	JBU	NAFDU	MRF	Proposed
2	Cones	7.46	3.94	3.88	4.26	2.71
	Teddy	7.74	3.99	4.35	4.25	3.44
	Venus	1.41	0.35	0.46	0.41	0.16
4	Cones	13.36	6.42	7.13	6.33	3.90
	Teddy	12.74	7.29	8.08	6.72	5.29
	Venus	2.64	0.71	0.94	0.67	0.28
8	Cones	26.58	11.85	14.39	12.32	9.12
	Teddy	24.03	12.36	13.43	13.37	11.52
	Venus	5.04	2.40	2.22	1.54	1.30

TABLE II
COMPARISON OF MEAN ABSOLUTE DIFFERENCE

Scaling factor	Image	Upsampling methods				
		Bilinear	JBU	NAFDU	MRF	Proposed
2	Cones	1.23	0.83	0.76	0.43	0.42
	Teddy	1.06	0.70	0.66	0.34	0.32
	Venus	0.24	0.11	0.10	0.09	0.09
4	Cones	2.22	1.39	1.26	0.78	0.71
	Teddy	1.83	1.23	1.15	0.61	0.55
	Venus	0.47	0.21	0.19	0.21	0.19
8	Cones	4.09	2.16	2.09	1.73	1.42
	Teddy	3.55	1.99	1.76	1.32	1.16
	Venus	0.75	0.39	0.38	0.44	0.37

As summarized in Table I, the proposed method provided outstanding results in terms of BPR for all images and all scaling factors. In the case of MAD, we also obtained the superior results except a few cases, as shown in Table II. It is because we defined the energy function that is supplied more depth initial measurements and also considers the depth discontinuities using color segment information.

V. CONCLUSION

In this paper, we have proposed an upsampling method for low-resolution depth maps. The proposed method is based on MRF model, and we define an energy function that considers more depth initials and color segment information. After optimizing the energy function by belief propagation, we obtain the upsampled depth map. From the experimental results, we have confirmed that the proposed method efficiently upsampled the depth in spite of the increment of scaling factor. In addition, the results from the proposed method outperformed compared to the other depth upsampling approaches in terms of objective evaluations such as the bad pixel rate and mean absolute difference.

ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (2013-067321).

REFERENCES

[1] C. Fehn, R. Barre, and S. Pastoor, “Interactive 3DTV - Concepts and Key Technologies,” *Proceedings of the IEEE*, vol. 94, no. 3, pp. 524-538, Mar. 2006.

[2] A. Smolic and P. Kauff, "Interactive 3D Video Representation and Coding Technologies," *Proceedings of the IEEE, Spatial Issue on Advances in Video Coding and Delivery*, vol. 93, no. 1, pp. 99-110, 2005.

[3] C. Fehn, "Depth-image-based Rendering (DIBR), Compression, and Transmission for a New Approach on 3D-TV," *Proc. of SPIE Stereoscopic Displays and Virtual Reality Systems*, vol. 5921, pp. 93-104, 2004.

[4] J. Kopf, M. F. Cohen, D. Lischinski, and M. Uyttendaele, “Joint Bilateral Upsampling,” *Proc. of SIGGRAPH*, pp. 96-100, Aug. 2007.

[5] Q. Yang, R. Yang, J. Davis, and D. Nister, “Spatial-depth Super Resolution for Range Images,” *Proc. of IEEE Conf. Computer Vision and Pattern Recognition*, pp. 1-8, June 2007.

[6] D. Chan, H. Buisman, C. Theobalt, and S. Thrun, “A Noise-aware Filter for Real-time Depth Upsampling,” *Proc. of ECCV Workshop on Multi-camera and Multi-modal Sensor Fusion Algorithms and Applications*, pp. 1-12, Oct. 2008.

[7] J. Diebel and S. Thrun, “An Application of Markov Random Fields to Range Sensing,” *Proc. of Advances in Neural Information Processing Systems*, vol. 18, pp. 291-298, Dec. 2006.

[8] J. Park, H. Kim, Y. Tai, M. Brown, and I. Kweon, “High Quality Depth Map Upsampling for 3D-TOF Cameras,” *Proc. of IEEE Int. Conf. Computer Vision*, pp. 1623-1630, Nov. 2011.

[9] D. Comaniciu and P. Meer, “Mean Shift: A Robust Approach Toward Feature Space Analysis,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, no. 4, pp. 603-619, May 2002.

[10] P.F. Felzenszwalb and D.P. Huttenlocher, “Efficient Belief Propagation for Early Vision,” *Proc. of IEEE Conf. Computer Vision and Pattern Recognition*, vol. 1, pp. 261-268, June 2006.