Depth Map Upsampling using Color Image Characteristics

Yun-Suk Kang and Yo-Sung Ho
Gwangju Institute of Science and Technology (GIST)
123 Cheomdan-gwagiro, Buk-gu, Gwangju 500-712 Korea
Email:{yunsuk, hoyo}@gist.ac.kr

Abstract- **In this paper, we present an upsampling method of low-resolution depth maps 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 and color image characteristics. Then, we obtain high-resolution depth maps 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.**

Keywords- Depth map, depth upsampling, Markov random field, 3D image

I. INTRODUCTION

In recent years, 3DTV and 3D video have been considered as the next generation broadcasting service since they can provide immersive and realistic experiences to users [1]. Depth image-based rendering (DIBR) technique also makes advanced 3D imaging services such as auto-stereoscopic 3D and the free-viewpoint navigation available using the depth information of the scene [2]. The depth of the scene represents range information of each pixel in the color image. Moreover, it is important because it is highly related to the quality of depth-based 3D images and applications.

In general, acquisition of depth information of the scene is categorized into passive and active sensors-based methods. The passive sensors-based methods calculate depth information using the captured images. Stereo matching is the most representative approach in this category [3]. It calculates a disparity value for each pixel and generates a disparity map. Although this passive sensor-based methods are usually easy to use, there are difficulties such as high complexity and low depth quality caused by image characteristics like textureless or occluded regions.

The active sensors-based methods use depth measuring equipment such as a depth sensor or 3D scanner. Nowadays, Time-of-flight (TOF) depth sensors have been widely used as an alternative of stereo matching. These sensors provide the accurate depth map in real-time. However, there are some drawbacks of the TOF depth sensors which are low output resolution, image distortion, and capturing failure according to object surfaces. Above all, low-resolution problem has to be overcome since current 3DTV and 3D video are based on high-resolution color images.

In order to overcome this resolution problem, many approaches have been proposed. There are filter-based depth upsampling approaches [4-6], and also Markov random field (MRF) based methods [7, 8]. Though they can upsample the depth map efficiently, they consider only single viewpoint. Since almost 3DTV and 3D video system have two or more views, it is required to increase the efficiency of the depth upsampling for stereoscopic 3D images.

In this paper, we propose an upsampling method of low-resolution depth map using MRF model. The proposed method considers the color image characteristics such as color difference between neighboring pixels and color segment information for calculating the weight for smoothness term of the energy function. After defining the energy function, the belief propagation (BP) optimization calculates the best depth values and provides the high-resolution depth map.

II. MARKOV RANDOM FIELD MODEL FOR DEPTH UPSAMPLING

We denote images and pixels for upsampling of low-resolution depth map based on MRF model 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. Pixels in D_U is represented as y, and z means mapped pixel to the location of I from D_L. Then, an energy function E is defined as Eq. (1).

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

The energy function is a sum of data and smoothness terms, which are left and right terms in Eq. (1), respectively. 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 current pixel is indicated as i, and $N(i)$ means four-neighboring pixels around i. In general, the weight w_{ij} for the smoothness term is defined by the color difference between two adjacent pixels i and j, as indicated Eq. (2).

$$w_{ij} = \exp\left(-\frac{\left(I(i) - I(j)\right)^2}{2\sigma_{ij}^2}\right) \tag{2}$$

III. PROPOSED DEPTH UPSAMPLING METHOD

In the proposed method, 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. We calculate the weight for smoothness term with reflecting color difference and 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.

We define w'_{ij} as a multiple of three sub-weights considering color values and color segment information as Eq. (3): w_{ij} for color difference between neighboring pixels, $w_{pi,j}$ for color patch difference, and $w_{s,ij}$ for color segment. These three weights controls the influence of smoothness cost according to color image characteristics.

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

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 weight of color patch difference $w_{pi,j}$ is defined as Eq. (4). Since the color similarity weight defined as Eq. (2) could be influenced by the noise in the color image, we use a small patch based on the pixel i and j, and calculate the mean squared differences. In Eq. (4), (u, v) mean the pixel position and N means number of pixels in the patch. This sub-weight also adjusts the smoothness cost according to the color difference.

$$w_{p,ij} = \exp\left(-\frac{\frac{1}{N}\sum_p \sum_q \{I(u_i + p, v_i + q) - I(u_j + p, v_j + q)\}^2}{2\sigma_{p,ij}^2} \right) \tag{4}$$

Thirdly, 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. (5) describes this relation between the color segment and depth. In Eq. (5), 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.

$$w_{s,ij} = \begin{cases} 1 & if\ S(i) = S(j) \\ C_{seg} & otherwise \end{cases} \tag{5}$$

For the last step, the energy function shown in Eq. (1) using w'_{ij} instead of w_{ij} is then optimized by using belief propagation method [10].

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	2.65
	Teddy	7.74	3.99	4.35	4.25	3.38
	Venus	1.41	0.35	0.46	0.41	0.16
4	Cones	13.36	6.42	7.13	6.33	3.85
	Teddy	12.74	7.29	8.08	6.72	5.21
	Venus	2.64	0.71	0.94	0.67	0.27
8	Cones	26.58	11.85	14.39	12.32	9.04
	Teddy	24.03	12.36	13.43	13.37	11.45
	Venus	5.04	2.40	2.22	1.54	1.29

IV. Experimental Results

For experiments of the proposed method, we used three test image sets provided by Middlebury stereo: 'Cones,' 'Teddy,' and 'Venus.' The ground-truth disparity maps were downsampled by factors of 2, 4, and 8. 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]. Table I shows the results. The proposed method outperformed the other upsampling approaches in terms of BPR for all image sets and for all scaling factors.

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 color difference 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 maps despite 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 evaluation.

Acknowledgment

This research was supported by the 'Cross-Ministry Giga KOREA Project' of the Ministry of Science, ICT and Future Planning, Republic of Korea(ROK). [GK13C0100, Development of Interactive and Realistic Massive Giga-Content Technology]

References

[1] 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.

[2] C. Fehn, "Depth-image-based rendering (DIBR), compression and transmission for a new approach on 3-D TV," in Proc. of SPIE Conference on Stereoscopic Displays and Virtual Reality Systems, vol. 5291, pp. 93-104, 2004.

[3] J. Sun, N.N. Zheng, and H.Y. Shum, "Stereo Matching Using Belief Propagation," IEEE Transactions of Pattern Analysis and Machine Intelligence, vol. 25, no. 5, pp. 787-800, 2003.

[4] J. Kopf, M. F. Cohen, D. Lischinski, and M. Uyttendaele, "Joint bilateral upsampling," ACM Transactions on Graphics, vol. 26, no. 3, pp. 1-5, 2007.

[5] Q. Yang, R. Yang, J. Davis, and D. Nister, "Spatial-depth super resolution for range images," in Proc. of IEEE Conference on Computer Vision and Pattern Recognition, pp. 1-8, 2007.

[6] D. Chan, H. Buisman, C. Theobalt, and S. Thrun, "A noise-aware filter for real-time depth upsampling," in Proc. of ECCV Workshop on Multi-camera and Multi-modal Sensor Fusion Algorithms and Applications, pp. 1-12, 2008.

[7] J. Diebel and S. Thrun, "An application of markov random fields to range sensing," Advances in Neural Information Processing Systems, vol. 18, pp. 291-298, 2006.

[8] J. Park, H. Kim, Y. Tai, M. Brown, and I. Kweon, "High quality depth map upsampling for 3D-TOF cameras," in Proc. of IEEE International Conference on Computer Vision, pp. 1623-1630, 2011.

[9] D. Comaniciu and P. Meer, "Mean Shift: A Robust Approach Toward Feature Space Analysis," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 4, pp. 609-619, 2002.

[10] P.F. Felzenszwalb and D.P. Huttenlocher, "Efficient Belief Propagation for Early Vision," in Proc. of IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 261-268, 2006.