

# Depth Map Upsampling Using Segmentation and Edge Information

Shuai Zheng<sup>1,2</sup>, Ping An<sup>1,2(✉)</sup>, Yifan Zuo<sup>1,2</sup>, Xuemei Zou<sup>1</sup>, and Jianxin Wang<sup>1,2</sup>

<sup>1</sup> School of Communication and Information Engineering, Shanghai University, Shanghai 200072, China

<sup>2</sup> Key Laboratory of Advanced Display and System Application of the Ministry of Education, Shanghai 200072, China

zs\_thinking@126.com, anping@shu.edu.cn

**Abstract.** A depth upsampling method based on Markov Random Field is proposed, considering the depth and color information. First, the initial interpolated depth map is inaccurate and oversmooth, we use a rectangle window centered on every pixel to search the maximum and minimum depth value of the depth map to find out the edge pixels. Then, we use the depth information to guide the segmentation of the color image and build different data terms and smoothness terms for the edge and non-edge pixels. The result depth map is piecewise smooth and the edge is sharp. In the meanwhile, the result is good where the color information is consistent while the depth is not or where the depth information is consistent while the color is not. Experiments show that the proposed method performs better than other upsampling methods in terms of mean absolute difference (MAD).

**Keywords:** Depth map · Upsampling · Segmentation · Edge

## 1 Introduction

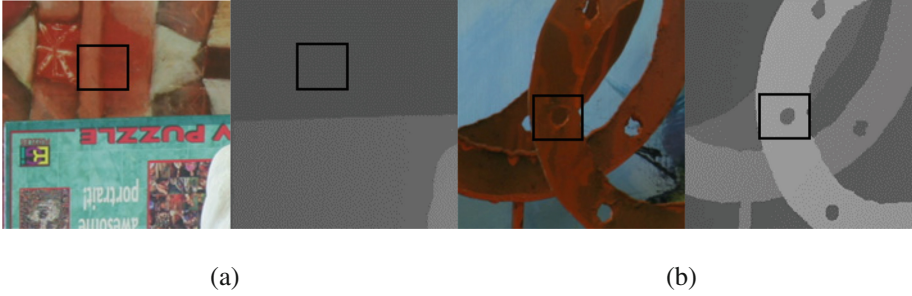
Recently, three-dimension television (3DTV) has been attractive and more applications such as human pose recognition, interactive free-viewpoint video, and 3D games have been put into use, the fundamental problem is the acquisition of the depth information and color information of the same scene. The fully-developed color camera makes the acquisition of the color information easy. However, the high complexity and low robustness of the passive methods, like stereo algorithms, limit the acquisition of the depth image. In recent years, with the advent of the low-cost depth sensors, such as Time-of-Flight (TOF) range sensors [1] and Kinect [2], the active sensor-based methods to gain the depth information become more popular. The TOF sensor measures the traveling time for a reflected light beam from the sender to an object to the receivers [1]. The tiny camera can overcome the shortcomings of the traditional passive methods and generate real-time depth map and intensity image. However, the current TOF sensors can only deliver limited spatial resolution depth image with noise and radial distortion. The Swiss Ranger 4000 [3] from company Swissranger, for example, can only produce  $176 \times 144$  depth map,

which is rather low compared to the color image of conventional color camera, the resolution is usually  $1280 \times 960$  or  $1920 \times 1080$ . Therefore, in order to acquire high-quality depth maps in 3DTV and other applications, the upsampling and distortion correction are necessary, this paper focuses on the upsampling of the depth image.

As for depth upsampling problem, the traditional method such as: nearest neighbor, bilinear and bicubic interpolation always generate oversmooth image especially at the depth discontinuities, the blurry boundaries make it hard to use. Hence, the upsampling method considering the low-resolution depth map and high-resolution color image gains much attention.

Depth map upsampling based on low-resolution depth map and high-resolution color image can be classified into two categories: filter-based upsampling approaches and Markov random field (MRF) based methods. As for filter-based upsampling method, Kopf et al. [4] propose a joint bilateral upsampling strategy when a high resolution prior is available to guide the interpolation from low to high resolution, but when the depth discontinuity is not consistent with the color edge, the texture transfer artifact is inevitable. To avoid the texture transfer artifacts, Chan et al. [5] propose noise aware filter for depth upsampling (NAFDU), which enables us to fuse the data from a low-resolution depth camera and a high-resolution color camera while preserving features and eliminating texture transfer artifact, but the image edge structure is easy to be broken using this method, thus the over-smoothing problem at the depth discontinuity is obvious. Yang et al. [6] quantize the depth values, build a cost volume of depth probability. They iteratively apply a bilateral filter on the cost volume, and take the winner-takes-all approach to update the HR depth values. However, the result is oversmooth at the depth discontinuity. He et al. [7] propose a guide filter to preserve the edge. Lu et al. [8] present a cross-based framework of performing local multipoint filtering to preserve the edge and the structure. David et al. [9] propose an anisotropic diffusion tensor, calculated from a high resolution intensity image to guide the upsampling of the depth map. Min et al. [10] propose a joint bilateral filter combined with the time information to ensure the temporal correlation. The filter-based method is fast, but it will cause over-smoothing problem somehow. For MRF method, Diebel and Thrun [11] propose a 4-node MRF model based on the fact that depth discontinuities and color difference tend to co-align, but the result is not that good. Lu et al. [12] presents a new data term to better describe the depth map and make the depth map more accurate, but the result at the depth discontinuities is still not good. Park et al. [13] use non-local means regularization to improve the smoothness term. Most of these algorithms are solved by Graph Cuts (GC) [14–16] or Belief Propagation (BP) [17]. Yin and Collins [18] use a 6-node MRF model combined the temporal information, which make the depth map more reliable. In addition, more and more weighting energy equations are proposed, papers [19, 20] combine the intensity map from the TOF camera as confidence map and the last frame image as time confidence to gain the depth map.

These methods are based on the assumption that the depth discontinuities and the boundaries of the color map are consistent. But when the depth value is the same while the color information is not or otherwise, the result is bad, as shown in Fig. 1.



**Fig. 1.** Inconsistency in the color image and depth map: (a) the depth value is the same while the color information is different; (b) the color information is the same while the depth value is different (Color figure online)

The paper proposes a method that can ensure the accuracy of the smooth region, the depth discontinuities and solve the problems shown in Fig. 1. The remainder of the paper is organized as follows: In Sect. 2, we review the MRF-based depth map upsampling method. The proposed method is presented in Sect. 3, followed by the experimental results and comparisons between the former work with our method in Sect. 4. Finally, we conclude our work in Sect. 5.

## 2 Depth Map Upsampling Method Based on MRF

An image analysis problem can be defined as a modeling problem, to solve the image problem is to find a solution to the model. The depth map upsampling problem can be transformed to seeking a solution to the MRF problem. The depth value of the image is related to the initial depth value and the depth value around. In MRF model, each node represents a variable including the initial depth value, the corresponding color information and the depth value around. The edge between the two nodes represents the connection between the two variables. Let the low-resolution depth map be  $D_0$ , and the high-resolution color image be  $I = \{z\}$ .  $p = (i, j)$  is the pixel in the result high-resolution depth map  $D = \{d_p\}$ . Using a MRF model to describe the problem, then it transforms to a maximum posterior probability of  $P(d, z)$ . The maximum posterior probability is:

$$P(d|z) = \frac{P(d, z)}{P(z)} \quad (1)$$

According to the theory of Hammersley-Clifford, solving the maximum posterior probability can be transformed to solving the minimum of the energy equation of (2) [21],  $N(i)$  represents the neighboring nodes of the node  $i$ ,  $\lambda_s$  is smoothness coefficient.

$$E = \sum_i D(i) + \sum_{\substack{i,j \\ j \in N(i)}} \lambda_s V(i, j) \quad (2)$$

The data term is used to ensure the result depth getting close to the initial depth value,  $\sigma$  is a threshold of the truncated absolute difference function:

$$D(i) = \min(|d_i - d_i^0|, \sigma) \quad (3)$$

The smoothness term is used to ensure the depth of the pixel getting close to the depth of the neighboring pixels and the depth map is piecewise smooth:

$$V(i, j) = W_{ij} \times \min(|d_i - d_j|, \sigma) \quad (4)$$

Thus, the minimum energy equation problem turns to label assigning problem, a high-probability depth label is given to every pixel according to the maximum posterior probability acquired from the initial depth map and the color map, the high-resolution depth map is acquired.

But there is blurring at the depth discontinuities of the interpolated initial depth map, which results in the depth value which the data term relies on is incorrect. The paper builds a rectangle window to search the maximum depth value and the minimum to get the difference value. If the difference is bigger than some threshold, it is regarded as an edge pixel, such that the accurate pixel of the initial depth map is acquired. The modified graph-based image segmentation method is used to get the segmentation information. Different smoothness weights of the edge pixel and non-edge pixels are built based on the edge and segmentation information to ensure the depth map is piecewise smooth and the edge is sharp. In the meanwhile, the area where the color information is consistent while the depth is not or otherwise is well dealt with.

### 3 Proposed Method

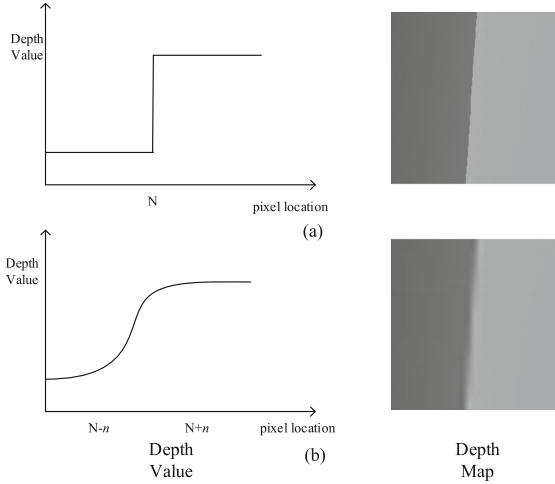
First, the bicubic interpolation method is used to the low-resolution depth map to get the initial depth map, then get the edge pixels and use the initial depth map to guide the segmentation of the color image, and build new data term and smoothness term based on the edge and segmentation information to get the high-resolution depth map. The section is described according to the above.

#### 3.1 Acquisition of the Edge Pixels of the Depth Map

The interpolated initial depth map is blurry at the depth discontinuities, and the depth value is inaccurate, as show in Fig. 2. The depth changes instantly at the discontinuities in the ground truth while in the interpolated depth map the depth changes gradually. We find out that when the upscale factor is  $n$ , the neighboring  $(2n + 1)$  pixels changes gradually.

So we build a  $n \times n$  rectangle window centered on every pixel to find out the maximum and minimum depth value to get the difference:

$$Dis(i) = \max_{i \in W(i)} D(i) - \min_{i \in W(i)} D(i) \quad (5)$$

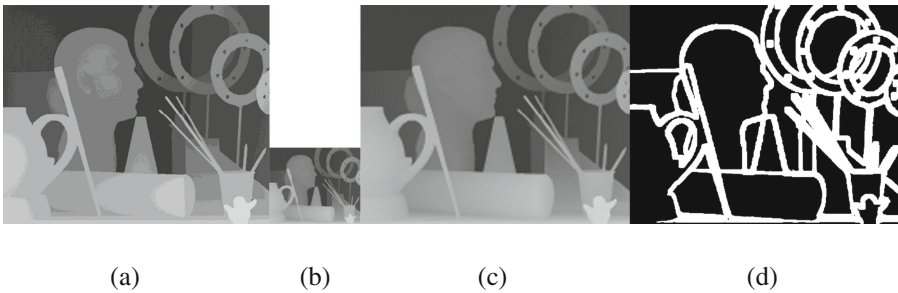


**Fig. 2.** Depth transformation at depth discontinuities: (a) depth map of ground truth and the depth value changes instantly at depth discontinuities; (b) depth map of the bicubic interpolated depth map and the depth value changes in the range of  $(N - n, N + n)$

If  $Dis(i)$  is bigger than the threshold, then the pixel is edge or the pixel is non-edge.

$$D_{edge}(i) = \begin{cases} 1 & \text{if } Dis(i) > th_{edge} \\ 0 & \text{else} \end{cases} \quad (6)$$

We use the depth image and the corresponding color image of the dataset Middlebury, take the “Art” image as example, as shown in Fig. 3(a). Downscale the depth image with factor 8 as shown in Fig. 3(b), and then the bicubic interpolation method is used to enlarge the low-resolution depth image as shown in Fig. 3(c). It can be seen there is blurring at the depth discontinuities after the bicubic interpolation. If the traditional canny, sobel edge extracting method is used to get the edge, there are some mistakes. While using the above method to get the edge range, we can ensure the depth value of the non-edge area is the same as the ground truth, as shown in Fig. 3(d).



**Fig. 3.** Acquisition of edge map: (a) ground truth; (b) low-resolution depth map; (c) bicubic interpolation depth map; (d) edge map acquired by formula (6)

### 3.2 Depth-Guided Color Image Segmentation

The graph-based segmentation [22] uses the adaptive threshold to segment the image. First compute the dissimilarity between every pixel and its neighboring 4 pixels, and sort the edge between two pixels increasingly, if the dissimilarity is no bigger than the inner dissimilarity, merge the two pixels and then go to the next edge. The inner difference of a district  $Int(C)$  is the biggest intensity difference and the difference between the districts  $Diff(C_1, C_2)$  is the smallest dissimilarity among all the edges between the two districts. If the dissimilarity between the districts is no bigger than the two inner difference:

$$Diff(C_i, C_j) \leq \min(Int(C_i) + r(C_i), Int(C_j) + r(C_j)) \quad (7)$$

Merge the two districts. In this paper, the condition of merging two districts is changed to use the segmentation image to guide the depth upsampling. As shown in Fig. 4, when the depth difference between the two districts is less than  $\alpha$  ( $7 < \alpha < 11$ ) and the difference of the two districts is no bigger than the inner difference of the two

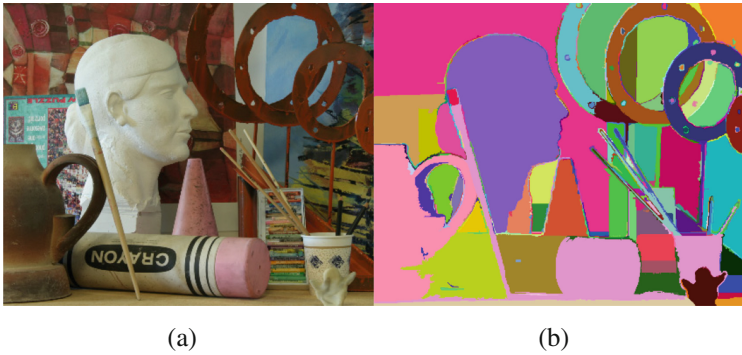


Fig. 4. Segmentation result: (a) color image; (b) segmentation image (Color figure online)

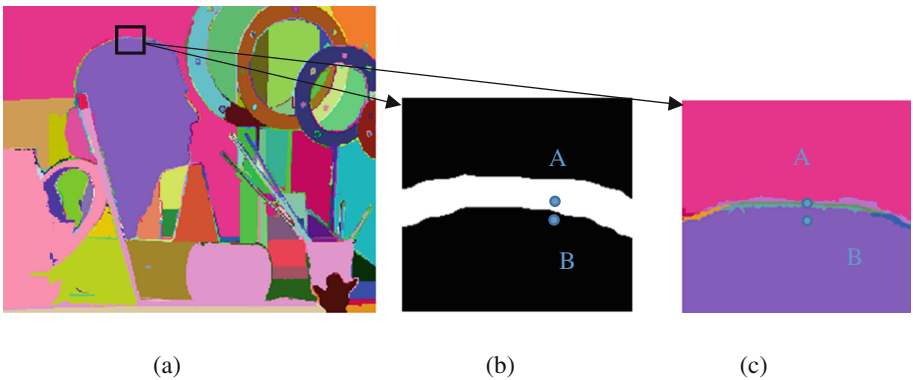


Fig. 5. Close-ups of the edge pixels and pixels around the edge: (a) segmentation image; (b), (c) close-ups of the edge map and the segmentation image of highlighted rectangle region in (a)

districts, merge the two districts, or merge the two districts only when the depth difference is less than  $\beta$  ( $0 < \beta < 3$ ). In other words, the segmentation takes the color information and depth map into account, but if the depth difference is very little which tends to 0, it is regarded as an area where the depth is the same while the color information is not, merge the two districts directly, the result is shown in Fig. 4(b).

### 3.3 The New Data Term and Smoothness Term

In the paper, we use an 8-node MRF model to deal with the depth map upsampling problem. The initial depth map is the interpolated depth map, and the color image is mean shift filtered [23] which can join the area with similar color, wipe some details and preserve the edge in the meantime. The MRF energy equation consists of two terms, the following describes the improvement of the data term and smoothness term respectively.

**Data Term.** In MRF model, the data term is used to ensure the depth value getting closed to the initial depth value, but because of the inaccuracy of the depth value at depth discontinuities in the interpolated depth map, we build different data terms for depth in edge area and non-edge area. The edge area acquired in 3.1 includes almost all edges pixels, and the data term is as follows:

1. If the current pixel  $i$  is not edge pixel, then the initial depth value is trustworthy, the data term is  $D(i) = \min(|d_i - d_i^0|, \sigma)$ .
2. If the current pixel  $i$  is in the edge range, then we can't rely on the initial depth value completely, we find out a non-edge pixel  $j$  which is  $(n + 3)$  pixels from the current pixel and belong to a district with the current pixel. If the pixel can be find out, we make the initial depth value of the current pixel  $i$  equal to the depth value of pixel  $j$ . The data term is  $D(i) = \min(|d_i - d_{j,|j-i|=n+3}^0|, \sigma)$ . As shown in Fig. 5, for point A at the color boundary of the color image, we can find a non-edge pixel point B which is  $(n + 3)$  pixels from the point A, so we make the initial depth value of A equal to the depth value of B.
3. If the current pixel is in the edge range and we can't find out a pixel that satisfies the condition of (2), we make a data weight  $w_i$  to reduce the effect of the original depth value. The data term is:

$$D(i) = w_i \times \min(|d_i - d_i^0|, \sigma) \quad (0 < w_i < 1) \quad (8)$$

Thus, we build a more accurate data term.

**Smoothness Term.** As for smoothness term:

$$V(i, j) = W_{ij} \times \min(|d_i - d_j|) \quad (9)$$

The smoothness weight  $W_{ij}$  consists of two sub-weight:  $W_{seg,ij}$  and  $W_{lab,ij}$ ,  $W_{seg,ij}$

represents the segmentation information of the two neighboring pixels and  $W_{lab,ij}$  describes the color difference between the two pixels, the two weights are used to control the influence of the neighboring pixels.

1. If the pixel  $i$  and its neighboring pixel  $j$  are non-edge pixels:  $D_{edge}(i) = 0$  and  $D_{edge}(j) = 0$ , the segmentation information is ignored, because the depth value of the two pixels must be the same, we set  $W_{seg,ij} = 1, W_{lab,ij} = 1$ .
2. For the pixel  $i$  and its neighboring pixel  $j$ , if one is edge pixel and the other is non-edge pixel:  $D_{edge}(i) = 0, D_{edge}(j) = 1$  or otherwise. Then the two pixels are around the edge range, because we expand the edge range in the search of edge points, the depth value should be the same:  $W_{seg,ij} = 1, W_{lab,ij} = \gamma(\gamma \geq 1)$ .
3. If the pixel  $i$  and its neighboring pixel  $j$  are in the edge range, we can't define whether the pixel  $i$  and  $j$  is edge pixel or not. We combine the color information and the segmentation information to define the smoothness weight. The similarity in color image is defined in LAB color space:

$$W_{lab} = e^{-\frac{\sqrt{(l_i-l_j)^2+(a_i-a_j)^2+(b_i-b_j)^2}}{\delta}} \quad (10)$$

The more similar the color image is, the possibility of the depth value being the same is larger. But if it is an area where the color information is consistent while the depth is not, the result depth map is wrong. So we make a segmentation weight in addition:

$$W_{seg} = \begin{cases} 1 & \text{if } seg(i) = seg(j) \\ \tau(0 < \tau < 1) & \text{else} \end{cases} \quad (11)$$

If the two neighboring pixels belong to one district, we just consider the color information of the two pixels. While if the two pixels don't belong to one district, then weaken the dependency relation of the two pixels, so the result is more accurate where the color information is consistent while the depth is not.

## 4 Experiment Results

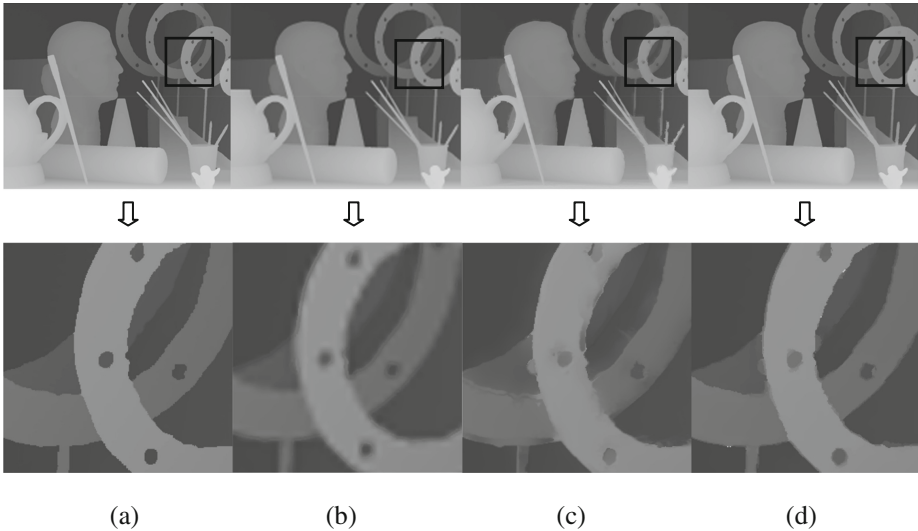
The experiments are performed on an Intel Core i7-4790 CPU, 3.60 GHz with 8 GB memory system using Microsoft visual C++ platform. We use 4 test image sets including the color image and the corresponding depth map provided by Middlebury dataset [24]: "Art", "Books", "Moebius" and "Dolls". First, the ground truth maps are downscaled by factors 2, 4 and 8 to get the low-resolution depth map, and then use bicubic interpolation to upsample the depth map, thus, we can get the initial depth map. Then the energy equation is built considering the segmentation and edge information described above, and solved by Graph Cut (GC). We set the parameters as follow:  $\lambda_s = 1.5, \sigma = 35, th_{edge} = 10, \gamma = 5, \tau, w_i$  are set 0.7. For "Art", "Books", "Moebius", the parameters  $\alpha$  and  $\beta$  are set as:  $\alpha = 10, \beta = 3$ . For "Dolls", the parameters  $\alpha$  and  $\beta$  are set as:  $\alpha = 7, \beta = 2$ . This is because the depth difference between depth areas in "Dolls" is small, to reduce the two parameters is to get better segmentation result. To evaluate



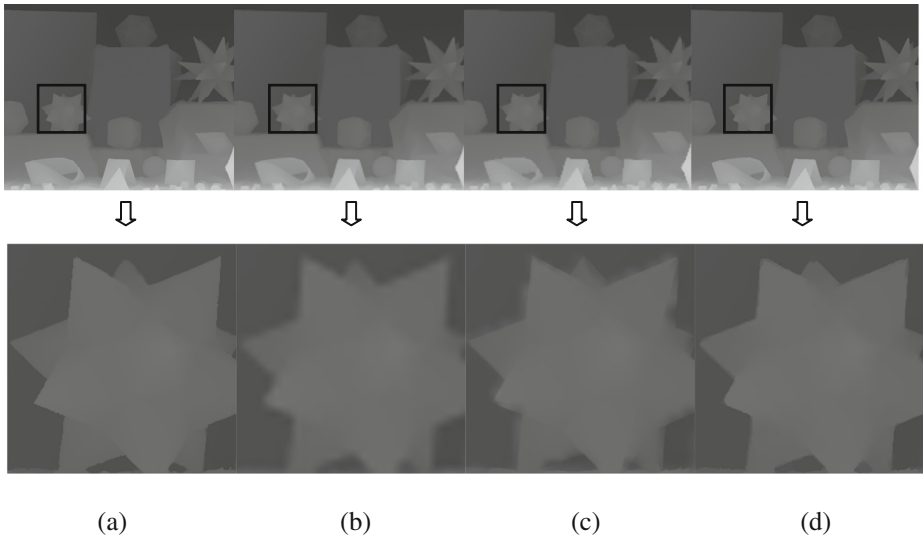
the method, we compare our method with the former work [6–11] about depth up sampling, the result for TGV and Edge are quoted from paper [9, 10]. From Table 1, it can be seen: as for Mean Absolute Difference (MAD), most of our method is improved. And the result is good where the color is consistent but the depth is not, as shown in Fig. 6. The results in Fig. 7 show that texture copying and edge blurring problems have been reduced with our method.

**Table 1.** Quantitative comparison on Middlebury dataset (in MAD)

	<i>Art</i>			<i>Book</i>			<i>Moebius</i>			<i>Dolls</i>		
magnitude	2	4	8	2	4	8	2	4	8	2	4	8
MRF[11]	0.59	0.96	1.89	0.21	0.33	0.61	0.24	0.36	0.65	0.19	0.34	0.68
JBFCv[6]	0.55	0.68	1.44	0.29	0.44	0.62	0.38	0.46	0.67	0.29	0.38	0.72
Guided[7]	0.63	1.01	1.70	0.22	0.35	0.58	0.23	0.37	0.59	0.28	0.35	0.56
CLMF0[8]	0.43	0.74	1.37	<b>0.14</b>	0.28	0.51	<b>0.15</b>	0.29	0.52	0.24	0.34	0.66
CLMF1[8]	0.44	0.76	1.44	<b>0.14</b>	0.28	0.51	<b>0.15</b>	0.29	0.51	0.23	0.34	0.60
TGV[9]	0.45	0.65	1.17	0.18	<b>0.27</b>	0.42	0.18	0.29	0.49	0.21	0.33	0.70
Edge[10]	0.41	0.65	1.03	0.17	0.30	0.56	0.18	0.29	0.51	<b>0.16</b>	<b>0.31</b>	0.56
our method	<b>0.41</b>	<b>0.58</b>	<b>1.01</b>	0.17	<b>0.27</b>	<b>0.41</b>	0.18	<b>0.27</b>	<b>0.44</b>	0.23	<b>0.31</b>	<b>0.53</b>



**Fig. 6.** Close-ups of the depth map “Art”: (a) ground truth; (b) depth map acquired by MRF [11]; (c) depth map acquired by JBFCv [6]; (d) depth map of our method



**Fig. 7.** Close-ups of the depth map “*Moebius*”: (a) ground truth; (b) depth map acquired by MRF [11]; (c) depth map acquired by JBFcv [6]; (d) depth map of our method

## 5 Conclusion

In the paper, a depth upsampling method based on MRF model is proposed. Usually, there is blurring at the depth discontinuities of the interpolated initial depth map, and the depth value is not accurate. This results in the depth value which the data term relies on is incorrect. The paper builds a rectangle window to search the maximum and the minimum depth value to get the difference value. If it is bigger than some threshold, it is regarded as an edge pixel. The accurate pixel of the initial depth map is acquired through the method. The graph-based image segmentation method is used. Different smoothness weights of the edge pixel and non-edge pixel are built to ensure the depth map is piecewise smooth and the edge is sharp. In the meanwhile, the depth evaluation of the area where the color information is consistent while the depth is not or otherwise is well dealt with.

**Acknowledgement.** This work was supported in part by the National Natural Science Foundation of China, under Grants U1301257, 61172096, 61422111, and 61301112.

## References

1. Remondino, F., Stoppa, D.: TOF Range-Imaging Cameras, vol. 68121. Springer, Heidelberg (2013)
2. Microsoft Kinect for Windows. <http://kinectforwindows.org/>
3. SwissRanger™ SR4000 Data Sheet. <http://www.mesaimage.ch/prodview4k.php>

4. Kopf, J., Cohen, M.F., Lischinski, D., Uyttendaele, M.: Joint bilateral upsampling. *J. ACM Trans. Graph.* **26**(3), 96:1–96:5 (2007)
5. Chan, D., Buisman, H., Theobalt, C., Thrun, S.: A noise-aware filter for real-time depth upsampling. In: *ECCV Workshop on Multicamera and Multimodal Sensor Fusion Algorithms and Applications*, pp. 1–12 (2008)
6. Xiong, Q., Yang, R., Davis, J., Nister, D.: Spatial-depths super resolution for range images. In: *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2007*, pp. 1–8 (2007)
7. He, K., Sun, J., Tang, X.: Guided image filtering. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) *ECCV 2010, Part I. LNCS*, vol. 6311, pp. 1–14. Springer, Heidelberg (2010)
8. Lu, J., Shi, K., Min, D., Lin, L., Do, M.N.: Cross-based local multipoint filtering. In: *2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 430–437 (2012)
9. Ferstl, D., Reinbacher, C., Ranftl, R., R  ther, M., Bischof, H.: Image guided depth upsampling using anisotropic total generalized variation. In: *2013 IEEE International Conference on Computer Vision (ICCV)*, pp. 993–1000 (2013)
10. Min, D., Lu, J., Do, M.N.: Depth video enhancement based on weighted mode filtering. *IEEE Trans. Image Process.* **21**(3), 1176–1190 (2012)
11. Diebel, J., Thrun, S.: An application of markov random fields to range sensing. In: *Proceedings of Advances in Neural Information Processing Systems*, vol. 18, pp. 291–298 (2006)
12. Lu, J., Min, D., Pahwa, R.S., Do, M.N.: A revisit to MRF-based depth map super-resolution and enhancement. In: *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 985–988 (2011)
13. Park, J., Kim, H., Tai, Y.W., Brown, M.S., Kweon, I.S.: High quality depth map upsampling for 3D-TOF cameras. In: *2011 IEEE International Conference on Computer Vision (ICCV)*, pp. 1623–1630 (2011)
14. Kolmogorov, V., Zabih, R.: What energy functions can be minimized via graph cuts. *IEEE Trans. Pattern Anal. Mach. Intell.* **26**(2), 147–159 (2004)
15. Boykov, Y., Veksler, O., Zabih, R.: Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Anal. Mach. Intell.* **23**(11), 1222–1239 (2001)
16. Boykov, Y., Kolmogorov, V.: An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Trans. Pattern Anal. Mach. Intell.* **26**(9), 1124–1137 (2004)
17. Ihler, A.T., Lii, J., Willsky, A.S.: Loopy belief propagation: convergence and effects of message errors. *J. Mach. Learn. Res.* **6**, 905–936 (2005)
18. Yin, Z.Z., Collins, R.: Belief propagation in a 3D spatio-temporal MRF for moving object. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8 (2007)
19. Choi, O., Jung, S.-W.: A consensus-driven approach for structure and texture aware depth map upsampling. *IEEE Trans. Image Process.* **23**(8), 3321–3335 (2014)
20. Schwarz, S., Sjostrom, M., Olsson, R.: A weighted optimization approach to time-of-flight sensor fusion. *IEEE Trans. Image Process.* **23**(1), 214–225 (2014)
21. Zhu, J., Wang, L., Gao, J.: Spatial-temporal fusion for high accuracy depth maps using dynamic MRFs. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(5), 899–909 (2010)
22. Felzenszwalb, P.F., Huttenlocher, D.P.: Efficient graph-based image segmentation. *Int. J. Comput. Vis.* **58**(2), 167–181 (2004)
23. Comaniciu, D., Meer, P.: Mean shift: a robust approach toward feature space analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(5), 603–619 (2002)
24. Middlebury Stereo Datasets. <http://vision.middlebury.edu/stereo/data/>