

**© 2016 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.**

# EXPLICIT MODELING ON DEPTH-COLOR INCONSISTENCY FOR COLOR-GUIDED DEPTH UP-SAMPLING

Anonymous ICME submission

## ABSTRACT

Color-guided depth up-sampling is to enhance the resolution of depth map according to the assumption that the depth discontinuity and color image edge at the corresponding location are consistent. Through all methods reported, MRF including its variants is one of major approaches, which has dominated in this area for several years. However, the assumption above is not always true. Solution usually is to adjust the weighting inside smoothness term in MRF model. But there is no any method explicitly considering the inconsistency occurring between depth discontinuity and the corresponding color edge. In this paper, we propose quantitative measurement on such inconsistency and explicitly embed it into weighting value of smoothness term. Such solution has not been reported in the literature. The improved depth up-sampling based on the proposed method is evaluated on Middlebury datasets and ToFMark datasets and demonstrate promising results.

*Index Terms*— depth map up-sampling, Markov Random Field (MRF), depth recovery

## 1. INTRODUCTION

Acquirement of high-quality depth data is the key problem in the field of 3-D computer vision, which is required in many applications, e.g., interactive view interpolation, 3DTV, 3D object modeling, robot navigation, and 3D tracking. Recently, ToF sensors are widely used, especially for dynamic scenes. However, depth maps captured by ToF are noisy and have lower resolutions, for example,  $176 \times 144$  and  $200 \times 200$ , compared with high-resolution (HR) color cameras [1].

In order to up-sample the depth map in low resolution (LR), many methods have been proposed throughout the past years. Generally speaking, they can be classified into two classes: non-color-guided methods [16, 17] and color-guided methods [2-15]. In non-color-guided methods, paper [16] only requires a single image for up-sampling by using smoothing priors from local self-similarities but it either has difficulties in textured areas, or only work well for small up-sampling factors. Another type of non-color-guided approach [17] is to fuse multiple displaced LR depth maps into a single HR depth map, which is not convenient for real applications. In addition, the color-guided depth map super-resolution (SR) intends to improve the quality and resolution of the original depth data captured directly by using a registered high resolution (HR) color image. The fundamental assumption of color-guided depth up-sampling methods is that the depth discontinuity and color image edge at the corresponding location are consistent [2]. Under this assumption, registered color image can provide structure information that is missing in low resolution depth map to fulfill the task of depth up-sampling. Color-guided

up-sampling methods can be classified into three categories that are filter-based [3-7], global-based [2,8-11] and learning-based [12-15].

Compared with filter-based solutions, global-based methods are more robust to noise in depth map captured by sensors. MRF based methods are major methods in the category of global-based solutions. MRF is also the interest of this paper. Thus, MRF-based global optimization and its variants are particularly explored here. There are two terms in MRF, which are data term and smoothness term. Data term indicates the compatibility of the value with the given data and smoothness term contributes to a piecewise smooth solution. Diebel et al. modeled depth map SR as solving a multi-labeling optimization problem via Markov Random Fields (MRF) [2]. Lu et al. [8] further extended this work by designing a data term which more fits to the characteristics of depth maps. Zhu et al. [9] updated the traditional spatial MRF to dynamic MRFs, therefore, both of the spatial and the temporal information can be introduced in energy function, which improves accuracy and robustness of up-sampling depth for dynamic scenes. Park et al. [10] used a non-local term to regularize depth maps and combined with a weighting scheme which involves edge, gradient, and segmentation information extracted from HR color images. Ferstl et al. [11] modeled the smoothness term as a second order total generalized variation regularization, and guided the depth up-sampling with an anisotropic diffusion tensor which is computed from a HR color image, providing high-quality up-sampling results.

Although color-guided up-sampling methods work well, color guidance image might contribute to texture-copy artifact as well as depth discontinuity blurring. The main problem is that the fundamental assumption of color-guided methods is not always true. That is, depth discontinuity regions on depth map do not necessarily correspond to the regions of color edge in the registered color image.

In fact, these artifacts have been noticed for a long time, and almost all state-of-the-art methods mentioned above adopt various ways to eliminate the texture-copy and depth discontinuity blurring artifacts. But they do not explicitly evaluate the edge inconsistency between color image and depth map. Therefore, they cannot adaptively control the effect of guidance from color image when up-sampling depth map.

In this paper, the main contributions are in three aspects, 1. Proposing method explicitly considering the inconsistency occurring between depth discontinuities and the corresponding color edges, and measuring inconsistency quantitatively; 2. Proposing method explicitly embedding inconsistency measurement above into weighting value of smoothness term in MRF energy function; 3. Evaluating the proposed method on Middlebury datasets and ToFMark datasets by comparing with the state-of-the-art algorithms and showing the improved performance

against the state-of-the-art depth map SR methods on Middlebury stereo datasets and ToFMark datasets.

The rest of this paper is organized as follows. Section 2 presents the proposed algorithm via Markov Random Fields with inconsistency measurement. In section 3, the experimental results are presented. Section 4 concludes this paper.

## 2. PROPOSED METHOD IN MRF FRAMEWORK WITH INCONSISTENCY MEASUREMENT

A Markov random field, also known as a Markov network or an undirected graphical model has been widely utilized for many image processing applications and tasks. MRF formulates depth map SR as solving an optimization problem. The input includes HR image and LR depth map. According to the Hammersely-Clifford theorem [18], solving MRF is equivalent to optimizing the Gibbs energy function, whose general formulation is defined as follows:

$$D = \arg \min_{D=\{d_p\}} \sum_{p \in O} E_{data}(d_p, D_p) + \lambda \sum_p \sum_{q \in N_p} \omega_{pq} \cdot E_{smooth}(d_p, d_q) \quad (1)$$

where  $D$  indicates the value set of the reconstructed HR depth map,  $d_p$  indicates the reconstructed value of pixel  $p$ , and  $N_p$  is the set of 8-connected neighboring pixels for the pixel  $p$ .  $O$  is the pixel set consisting of pixels which have observed depth values.  $D_p$  is the observed depth value of pixel  $p$ .  $E_{data}$  is called the data term which maintains the consistency between the reconstructed depth value and the initial observed depth value.  $E_{smooth}$  is called the smoothness term which penalizes the differences between the reconstructed depth value and the depth values in the neighboring region. The parameter  $\lambda$  is used to balance the data term and smoothness term.  $\omega_{pq}$  links color image to depth map, which provides the guidance from color image for depth up-sampling based on the assumed consistency between color edge and depth discontinuity (i.e. depth edge) [2]. As mentioned above, this assumption is not always true. It is the root problem of texture-copy and depth discontinuity blurring happening during depth up-sampling because of the wrong guidance from color image.

Section 2.1 proposes the quantitative measurement on the inconsistency between color edge in color image and depth discontinuity in the corresponding regions on depth map. Section 2.2 and 2.3 explicitly embed such measurement into energy function in MRF framework to adaptively adjust MRF optimization.

### 2.1 Measurement on the inconsistency between color edge and depth discontinuity in the corresponding regions

Since the structure of an image region is determined by its edges, it is sensible to perform the inconsistency measurement between depth map and color image via their edge maps. Motivated by image quality assessment [19], this paper models edge inconsistency measurement between color image and depth map as edge map quality assessment in bi-direction evaluation. In order to introduce image quality assessment into edge inconsistency measurement, a few specific points should be discussed.

1) Because registered color image and the corresponding depth map have the structural similarity observed only on the

relevant edge maps, this paper measures the inconsistency on binary edge maps generated by the registered color image and the corresponding depth map respectively.

2) In ideal situation, the inconsistency measurement should be independent of up-sampling scaling factors. That is, such inconsistency is one of essentials existing in the pair of the registered color image and depth map. Therefore, to obtain the evaluation of such essential inconsistency, ground truth depth map and registered color image should be the image pair under evaluation. In our case, in order to have the same scale of depth map and color image, the registered color image and the initial up-sampling depth map which is up-scaled by Bicubic interpolation to the same resolution of color image are the image pairs for inconsistency measurement.

Canny operator [20] is applied in color image and coarse up-sampled depth map to generate relevant edge maps. Due to coarse up-sampled depth map, the positions of corresponding edge pixels on color edge map and coarse up-sampled depth edge map are not consistent. In this paper, inconsistency measurement is casted as a MRF optimization problem. For each edge position on reference edge map, it will search the best consistency in its neighboring region around the corresponding position on the target edge map because of the small displacement between the positions of corresponding edge pixels mentioned above. Moreover, the displacement between each reference edge position and its consistency should be similar (i.e. no significant change) to its neighboring edge positions' displacements on the reference edge map. These two constrains are solved in a MRF framework through its data term and smoothness term respectively.

$$L = \arg \min_{L=\{l_p\}} \sum_{p \in ref} C(p, p+l_p) + \mu \cdot \sum_{p \in ref} \sum_{k \in N(p)} V(l_p, l_k) \quad (2)$$

where  $C(p, p+l_p)$  is the data term of the MRF model.

$p$  represents the position of edge pixel in the reference edge map.  $l_p$  stands for the displacement so  $p+l_p$  represents a position of edge pixel  $q$  which is in a neighboring region corresponding to the coordinate of  $p$  in target edge map. The size of neighboring region is determined by up-sampling scaling factor in the proposed depth SR task. In our work, the size of neighboring region is  $5 \times 5$  for 2X SR up-sampling,  $7 \times 7$  for 4X SR up-sampling,  $9 \times 9$  for 8X SR up-sampling and  $11 \times 11$  for 16X SR up-sampling. Thus, data term  $C(p, q)$  calculates the consistency between the position of the reference edge pixel  $p$  and the position of target edge pixel  $q$  in target edge map. Given  $p$ , if certain target pixel  $q$  in neighboring region of  $p$  is not an edge pixel in target edge map, it is regarded definite inconsistency. In that case,  $C(p, q)$  is assigned to the maximum inconsistency value (i.e. 1 in our work). Otherwise, this inconsistency measurement is measured on two blocks where edge pixel  $p$  and edge pixel  $q$  are the center position respectively. In this paper, the size of block is  $3 \times 3$ . We define  $E_p = \{e_{p1}, e_{p2}, \dots, e_{pM}\}$  and  $E_q = \{e_{q1}, e_{q2}, \dots, e_{qN}\}$  to represent the sets of edge pixels in these two blocks respectively (excluding  $p$  and  $q$ ).  $M$  and  $N$  are the number of edge pixels inside these two sets. Thus, the inconsistency measurement between  $p$  and  $q$  is regarded as a matching problem between two data sets  $E_p$  and  $E_q$ . This matching problem is sorted out by using

Bipartite graph matching [21] which is more robust than MAD (mean of absolute difference) and Euclidean distance. The Bipartite graph  $G(E_p, E_q, W)$  is defined, where  $E_p$  and  $E_q$  are vertices in Bipartite graph and  $W$  represents the link between vertices whose weight is defined as  $\phi(i, j)$  which is a monotonic function that returns a positive penalty for local structural matching.

$$\phi(i, j) = f(|i^x - j^x| + |i^y - j^y|) \quad (3)$$

where  $f(0)=0, f(1)=1, f(2)=1.6$  and  $f(x)=2$  when  $x > 2$ .  $i, j$  are vertices in Bipartite graph,  $i^x, i^y$  are the coordinate of  $i$ .

Bipartite matching [21] is employed to enforce one-to-one matching between edge pixel data sets above. That is, it assures any edge pixel in  $E_p / E_q$  matches only one edge pixel in  $E_q / E_p$ , leaving  $|M - N|$  unmatched pixels. Unmatched pixels represent the potential structure differences between edge pixel sets  $E_p, E_q$ . Therefore, this paper introduces them into the calculation of inconsistency value between certain pixel pair  $p, q$ , described in Eq.(4).

$$C(p, q) = \left( \sum_{(p_s, q_s) \in E'_{pq}} \phi(p_s, q_s) / 2 + |M - N| \right) / 8 \quad (4)$$

$E'_{pq} = \{(p_1, q_1), (p_2, q_2), \dots, (p_r, q_r)\}$  is edge pixel pair sets selected by Bipartite graph matching [21].  $\phi(p_s, q_s)$  is the weight of the link between edge pixel  $p_s$  and edge pixel  $q_s$  and  $s = \{1, 2, 3 \dots r\}$ . Therefore,  $\sum_{(p_s, q_s) \in E'_{pq}} \phi(p_s, q_s)$  is the matching cost of Bipartite matching [21] mentioned above. Through normalization, the range of data term  $C(p, q)$  is  $[0, 1]$ .

$V(l_p, l_k)$  is the smoothness term in Eq.(2), which gives a penalty when adjacent edge pixels have different displacements as,

$$V(l_p, l_k) = \begin{cases} 0, & l_p = l_k; \\ 1, & l_p \neq l_k; \end{cases} \quad (5)$$

$V(l_p, l_k)$  takes the connectivity of adjacent edge pixels into account, which means that connectivity of adjacent edge pixels is encouraged to maintain in the solution of Eq.(2).

$\mu$  is a balance factor between data term and smoothness term. It is set to 0.1 in this paper.  $N(p)$  is the set of 8-connected neighboring pixels of  $p$ .

Graph cut [23] is adopted to solve Eq.(2) MRF problem. The output of data term  $C$  computed by optimized displacements  $L$  represents the inconsistency between reference edge map and target edge map.

The inconsistency is measured based on reference edge map against target edge map. Thus, the measurement will be different when swapping these two edge maps. In this work, the two edge maps are color image edge map and depth discontinuity (edge) map of depth map. When color image edge map is regarded as the reference edge map for inconsistency measurement, it can be observed that inconsistent positions detected reflect the texture copy happening areas. On the other hand, when depth discontinuity (edge) map is regarded as the reference edge map, it

is observed that inconsistent positions reflect the depth discontinuity blurring happening areas.

## 2.2 Alignment of inconsistency maps

Inconsistency map for reference edge map  $C_{refer}$  consists of values of edge pixels computed in section 2.1 and values of non-edge pixels which are assigned to 0.

After bi-direction evaluation, there are two inconsistency maps  $C_{color}$ ,  $C_{depth}$  as well as two displacement maps  $L_{color}$ ,  $L_{depth}$  available for an image pair. They represent the inconsistency measurement and displacement when color image edge map or depth edge map are the reference edge map respectively. Before embedding the inconsistency measurement values into MRF based depth up-sampling framework, these two inconsistency maps must be consolidated to each other.

As mentioned before, in order to have the color image and the depth map with the same size for inconsistency measurement, the LR depth map is first coarsely up-sampled to the larger size using Bicubic interpolation. This coarse up-sampling may shift the position of edge pixel a bit from its true location. On the other hand, the position of edge pixel on color image edge map is more precise because of high quality of color image. Through Bipartite graph matching process mentioned above with depth edge map as the reference edge map, the displacement between each depth edge pixel  $p$  and its color image edge pixel  $q$  is  $L_{depth}(p)$ . Consequently, the true location of depth edge pixel  $p$  supposes to be more close to  $p + L_{depth}(p)$  when  $C_{depth}(p) \neq 1$  which is not the case of definite inconsistency that represents no corresponding pixel in color edge image for  $p$ . Therefore, the  $C_{depth}$  is adjusted as,

$$C'_{depth}(p') = \begin{cases} \min_{p=p'+L_{depth}(p)} C_{depth}(p) & \text{if } C_{depth}(p) \neq 1 \\ C_{depth}(p) & \text{otherwise} \end{cases} \quad (6)$$

In Eq.(6), if there are more than one pixel  $p$  mapping to the same pixel  $p'$ , the best mapping with the lowest cost is adopted and copy the values of the rest mapping in  $C_{depth}$  to  $C'_{depth}$ .

Once two inconsistency maps  $C'_{depth}$  and  $C_{color}$  are aligned, a confidence map  $\alpha_p$  is defined as below, taking two directions of evaluation into account. It describes the final inconsistency status between color image edge map and depth discontinuity (edge) map, which is embedded into MRF based depth up-sampling framework i.e. Eq.(1) (see Section 2.3).

$$\alpha_p = \max(C'_{depth}(p), C_{color}(p)) \quad (7)$$

## 2.3 Improved MRF by considering inconsistency measurement

According to the MRF based depth up-sampling framework shown in Eq.(1), the confidence value in Eq.(7) (i.e. inconsistency measurement) is embedded into the model. To simplify the explanation in the follows, Eq.(1) is updated below by introducing two new terms.

$$D = \arg \min_{D=\{d_p\}} \sum_{p \in O} \lambda_{edge-p} E_{data}(d_p, D_p) + \lambda \sum_{p, q \in N_p} \lambda_{smooth-pq} E_{smooth}(d_p, d_q) \quad (8)$$

where  $\lambda_{smooth-pq}$  is to replace  $\omega_{pq}$  in Eq.(1).  $\lambda_{edge-p}$  is a boolean value which is assigned to 0 when  $p$  is located on sparse depth discontinuity position in high resolution, mapping from original LR depth samples. Otherwise it is assigned to 1. The aim of  $\lambda_{edge-p}$  is to eliminate the unreliable observed values in observed value set  $O$ .

The data term is defined according to the depth samples:

$$E_{data}(d_p, D_p) = (d_p - D_p)^2 \quad (9)$$

The smoothness term is defined as:

$$E_{smooth}(d_p, d_q) = (d_p - d_q)^2 \quad (10)$$

Generally speaking, guidance information for up-sampling task can be derived from two sources, one is registered color image, and the other is original depth map. Based on the confidence map  $\alpha_p$  computed in Eq.(7), this paper combines it to generate a new guidance image to compute the weighting value  $\lambda_{smooth-pq}$ .  $\lambda_{smooth-pq}$  is constructed as below.

$$\lambda_{smooth-pq} = e^{-\frac{(\nabla_{color}^{pq} \cdot (1 - \alpha_{pq}) + \nabla_{depth}^{pq} \cdot \alpha_{pq})^2}{2\delta^2}} \quad (11)$$

where  $\nabla_{color}^{pq}$  and  $\nabla_{depth}^{pq}$  represent color difference and depth difference between position  $p$  and its neighboring pixel  $q$  in guided color image and coarse up-sampled depth map respectively.  $\delta$  controls decay rate of exponential function in Eq.(11).  $\alpha_{pq}$  is set to the average of  $\alpha_p$  and  $\alpha_q$  because of the symmetrical relationship of pixel pair  $p, q$ . It is observed that when the corresponding color edge map is more consistent with depth edge map ( $\alpha_{pq}$  has lower value),  $\nabla_{color}^{pq}$  is able to play more important role in computing the weighting value  $\lambda_{smooth-pq}$  by following the principle of color-guided depth up-sampling, and vice versa. Therefore, it can preserve depth edges and prevent texture-copy artifacts efficiently by adaptively controlling the guidance from color image. This paper adopts Conjugate Gradient method (CG) [24] to solve problem Eq.(8).

### 3. EXPERIMENTAL RESULTS

The proposed method is evaluated on Middlebury datasets [22] and ToFMark datasets [11] under different up-sampling factors. The comparison performance against the state-of-the-art methods are demonstrated.

#### 3.1 Parameters setting

All the edge maps are computed through Canny operator. For color image edge detection, the dual thresholds setting are 0.04 and 0.12 respectively in Canny operator. For depth edge map calculation, the low resolution depth map is first up-sampled by coarse Bicubic interpolation to the relevant resolution according to SR up-sampling factors. The dual thresholds in Canny operator for depth map edge detection are defined as below, which are ranges.

$$\begin{aligned} Th_L &= [(\log_2 factor) \times 0.01, (\log_2 factor) \times 0.02] \\ Th_H &= [(\log_2 factor) \times 0.03, (\log_2 factor) \times 0.04] \end{aligned} \quad (12)$$

where  $factor$  is the corresponding SR up-sampling factor.

It is observed that larger  $\lambda$  in MRF framework (i.e. Eq. (1)) will be helpful when the depth map has more noise. But to the depth map with less noise, smaller value of  $\lambda$  is helpful. In the experiments,  $\lambda = 0.01$  is set for the experiments on Middlebury dataset and  $\lambda = 0.6$  is for the experiments on ToFMark dataset. In addition,  $\delta$  (i.e. Eq.(11)) is assigned to 2 in this paper.

#### 3.2 Experimental results on Middlebury datasets

Six objects, ‘‘Art’’, ‘‘Book’’, ‘‘Moebius’’, ‘‘Reindeer’’, ‘‘Laundry’’, and ‘‘Dolls’’ from the Middlebury’s benchmark [22] are used for evaluation. The proposed method is compared with state-of-the-art methods: Bicubic interpolation, MRF-based method (MRF) [2], Spatial-depth super resolution for range images (JBUV) [5], guided image filtering (Guided) [6], edge-weighted NLM-regularization (NLMR) [10], joint geodesic filtering (JGF) [4], total generalized variation (TGV) [11]. The results for MRF [2] and JBUV [5] on ‘‘reindeer’’, ‘‘laundry’’ and ‘‘doll’’ RGB-D pairs were not provided in the previous papers.

Table I shows the up-sampling results under four different up-sampling factors. It is noticed that the proposed method obtains the lowest MAD for most cases except 16X case. The proposed method is to embed the edge map inconsistency into the color guided depth up-sampling. The edge map inconsistency measurement is carried between high resolution color image and the depth map of the same resolution by coarse Bicubic interpolation. In the case of 16X, the coarse up-sampling by interpolation introduces significant noise which affects the edge map quality.

Fig.1. shows the result of ‘‘Dolls’’ with 8X up-sampling factor by comparing with three state-of-the-art methods, TGV [11], NLMR [10] and JGF [4]. From the marked regions which are enlarged, it is shown that our result preserves edges better and has no texture-copy artifact.

#### 3.3 Experimental results on ToFMark datasets

The proposed method is also tested on ToFMark datasets [11] consisting of three RGB-D pairs, ‘‘Books’’, ‘‘Shark’’, ‘‘Devil’’, with ground-truth depth maps. The resolution of original depth maps are  $120 \times 160$ , and the intensity images are the size of  $610 \times 810$ . The suggested up-sampling factor is approximately 6.25.

Table II illustrates quantitative comparison results. The up-sampling errors are computed by MAD in mm. It is shown that, for ‘‘Books’’, the proposed method achieves the similar performance as TGV[11] and shows better performance than other methods. For other cases, ‘‘Shark’’ and ‘‘Devil’’, the proposed method obtains the lowest MAD compared with four state-of-the-art methods and Bicubic interpolation.

Fig.2. shows the up-sampling result of ‘‘Shark’’ compared with Bicubic interpolation, JGF [4] and TGV [11]. The up-sampling results and the error maps are listed in the first and the third rows respectively. The marked regions are enlarged in the second row. It is noticed that TGV [11] and the proposed method perform better than other methods when the low resolution depth map has much noise. But TGV [11] introduces texture-copy artifacts (e.g. the upper edge of box in the center of depth map). The proposed method does not have such texture-copy artifacts. In addition, the edge of the paper placed on the desk in our result is

TABLE I  
 QUANTITATIVE UP-SAMPLING RESULTS (IN MAD) ON MIDDLEBURY DATASETS AT FOUR UP-SAMPLING FACTORS

Methods	Art				Book				Moebius				Reindeer				Laundry				Doll			
	2x	4x	8x	16x	2x	4x	8x	16x	2x	4x	8x	16x	2x	4x	8x	16x	2x	4x	8x	16x	2x	4x	8x	16x
<b>Bicubic</b>	0.48	0.97	1.85	3.59	0.13	0.29	0.59	1.15	0.13	0.30	0.59	1.13	0.30	0.55	0.99	1.88	0.28	0.54	1.04	1.95	0.20	0.36	0.66	1.18
<b>MRF[2]</b>	0.59	0.96	1.89	3.78	0.21	0.33	0.61	1.20	0.24	0.36	0.65	1.25	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
<b>JBUV[5]</b>	0.55	0.68	1.44	3.52	0.29	0.44	0.62	1.45	0.38	0.46	0.67	1.10	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
<b>Guided[6]</b>	0.63	1.01	1.70	3.46	0.22	0.35	0.58	1.14	0.23	0.37	0.59	1.16	0.42	0.53	0.88	1.80	0.38	0.52	0.95	1.90	0.28	0.35	0.56	1.13
<b>NLMR[10]</b>	0.41	0.65	1.03	2.11	0.17	0.30	0.56	1.03	0.18	0.29	0.51	1.10	0.20	0.37	0.63	1.28	0.17	<b>0.32</b>	<b>0.54</b>	1.14	0.16	0.31	0.56	<b>1.05</b>
<b>JGF[4]</b>	0.29	<b>0.47</b>	0.78	<b>1.54</b>	0.15	0.24	0.43	0.81	0.15	0.25	0.46	<b>0.80</b>	0.23	0.38	0.64	<b>1.09</b>	0.21	0.36	0.64	<b>1.20</b>	0.19	0.33	0.59	1.06
<b>TGV[11]</b>	0.45	0.65	1.17	2.30	0.18	0.27	0.42	0.82	0.18	0.29	0.49	0.90	0.32	0.49	1.03	3.05	0.31	0.55	1.22	3.37	0.21	0.33	0.70	2.20
<b>Ours</b>	<b>0.25</b>	<b>0.47</b>	<b>0.76</b>	1.96	<b>0.11</b>	<b>0.22</b>	<b>0.39</b>	<b>0.76</b>	<b>0.11</b>	<b>0.24</b>	<b>0.45</b>	0.90	<b>0.17</b>	<b>0.34</b>	<b>0.61</b>	1.30	<b>0.15</b>	<b>0.32</b>	0.59	1.28	<b>0.14</b>	<b>0.28</b>	<b>0.51</b>	<b>1.05</b>

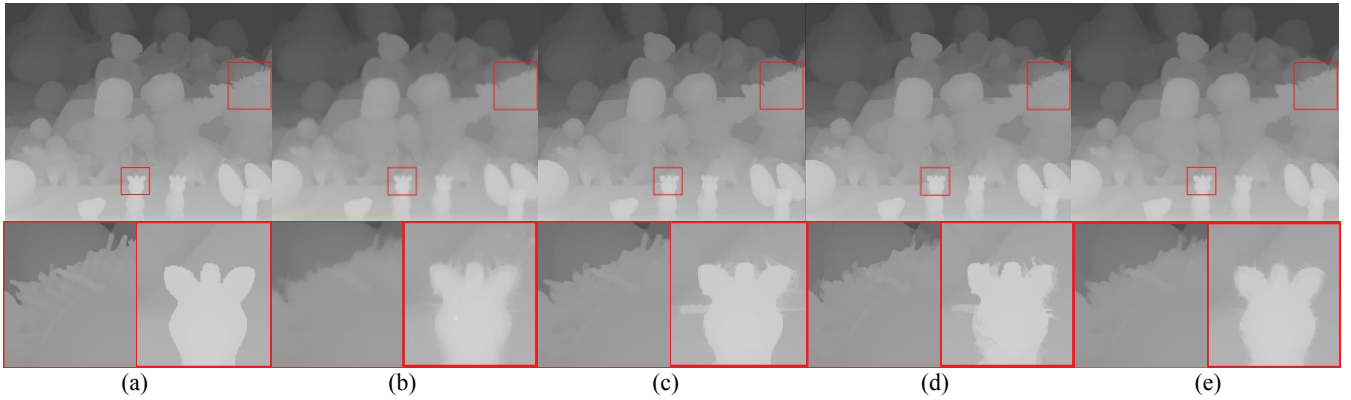


Fig. 1. Visual quality comparison for depth up-sampling on Doll from Middlebury Datasets: (a) depth ground truth, depth maps up-sampled (8 $\times$ ) by (b) TGV [11], (c) NLMR [10], (d) JGF [4], and (e) the proposed method.

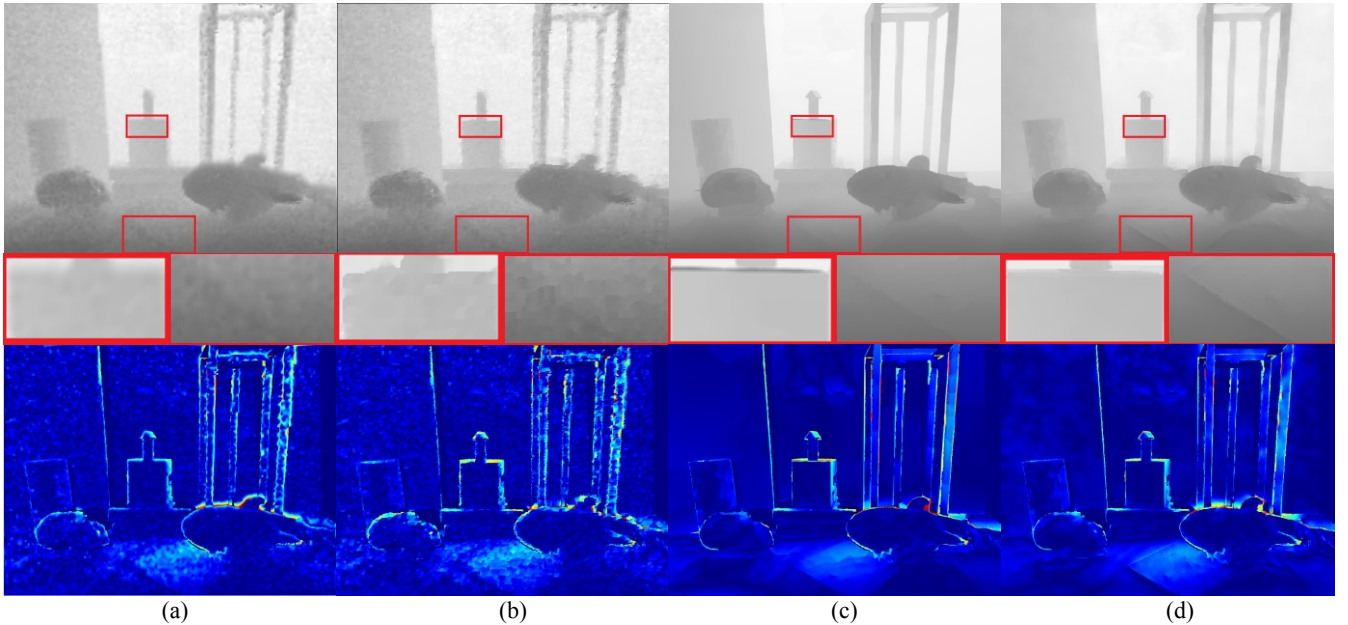


Fig. 2. Visual quality comparison on up-sampling depth for Shark from ToFMark datasets: (a) Bicubic, (b) JGF [4] (C) TGV [11], and (d) the proposed method.

TABLE II  
QUANTITATIVE DEPTH UP-SAMPLING RESULTS ON  
ToFMark DATASETS

	Bicubic	JBU[3]	Guided[6]	JGF[4]	TGV[11]	Ours
<b>Books</b>	16.23	16.03	15.74	17.39	<b>12.36</b>	12.39
<b>Shark</b>	17.78	18.79	18.21	18.17	15.29	<b>14.23</b>
<b>Devil</b>	16.66	27.57	27.04	19.02	14.68	<b>13.86</b>

sharper than that of others, which prove that our method can efficiently preserve depth edges.

#### 4. CONCLUSION

This paper proposes a color-guided method in MRF framework. The key contribution is to explicitly measure the inconsistency between color image edge map and the depth discontinuity (edge) map and embed it into MRF framework. It relaxes the assumption in color-guided methods. And it eliminates texture-copy artifacts and depth discontinuities blurring. Experimental results on the Middlebury datasets and ToFMark datasets prove the improved performance of the proposed method. In future, we will extract the information from LR depth map more accurately by using more advanced methods, instead of using Bicubic interpolation.

#### 5. REFERENCES

- [1] A. Kolb, E. Barth, R. Koch, and R. Larsen, "Time-of-flight cameras in computer graphics," *Comput. Graph. Forum*, vol. 29, no. 1, pp. 141-159, 2010.
- [2] J. Diebel and S. Thrun, "An application of Markov random fields to range sensing," *Advances in Neural Information Processing Systems*, vol. 18. Cambridge, MA, USA: MIT Press, pp. 291, 2005.
- [3] J. Kopf, M. F. Cohen, D. Lischinski, and M. Uyttendaele, "Joint bilateral upsampling," *Graphics, ACM Transaction on*, vol. 26, no. 3, pp. 96, 2007.
- [4] M.-Y. Liu, O. Tuzel, and Y. Taguchi, "Joint geodesic upsampling of depth images," *Computer Vision and Pattern Recognition (CVPR), IEEE Conference on*, pp. 169-176, 2013.
- [5] Q. Yang, R. Yang, J. Davis, and D. Nistér, "Spatial-depth super resolution for range images," *Computer Vision and Pattern Recognition (CVPR), IEEE Conference on*, pp. 1-8, 2007.
- [6] K. He, J. Sun, and X. Tang, "Guided image filtering," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 35, no. 6, pp. 1397 - 1409, 2012.
- [7] D. Min, J. Lu, M N. Do, "Depth video enhancement based on weighted mode filtering," *Image Processing, IEEE Transactions on*, vol. 21, no. 3, pp. 1176-1190, 2012.
- [8] J. Lu, D. Min, R. Pahwa, M. Do, "A revisit to MRF-based depth map superresolution and enhancement," *Acoustics, Speech and Signal Processing (ICASSP), IEEE International Conference on*, pp. 985-988, 2011.
- [9] J. Zhu, L. Wang, J. Gao, R. Yang, "Spatial-temporal fusion for high accuracy depth maps using dynamic MRFs," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, no. 5, pp. 899-909, 2010.
- [10] J. Park, H. Kim, Y.-W. Tai, M. S. Brown, and I. Kweon, "High quality depth map upsampling for 3D-TOF cameras," *Computer Vision (ICCV), IEEE International Conference on*, pp. 1623-1630, 2011.
- [11] D. Ferstl, C. Reinbacher, R. Ranftl, M. R  ther, and H. Bischof, "Image guided depth upsampling using anisotropic total generalized variation," *Computer Vision (ICCV), IEEE International Conference on*, pp. 993-1000, 2013.
- [12] S. A. Gudmundsson and J. R. Sveinsson, "ToF-CCD Image Fusion using Complex Wavelets," *Acoustics, Speech and Signal Processing (ICASSP), IEEE International Conference on*, pp. 1557-1560, 2011.
- [13] Y. Li, T. Xue, L. Sun, and J. Liu, "Joint Example-Based Depth Map Super-Resolution," *Multimedia and Expo (ICME), IEEE International Conference on*, pp. 152-157, 2012.
- [14] M. Mahmoudi and G. Sapiro, "Sparse Representations for Range Data Restoration," *Image Processing, IEEE Transactions on*, vol. 21, no. 5, pp. 2909-2915, 2012.
- [15] M. Kiechle, S. Hawe, M. Kleinsteuber, "A Joint Intensity and Depth Co-Sparse Analysis Model for Depth Map Super-Resolution," *Computer Vision (ICCV), IEEE International Conference on*, pp. 1545 - 1552, 2013.
- [16] G. Freedman and R. Fattal, "Image and Video Upscaling from Local Self-Examples," *Graphics, ACM Transactions on*, vol. 30, no. 2, pp. 1-11, 2011.
- [17] S. Schuon, C. Theobalt, J. Davis, and S. Thrun, "Lidar-Boost: Depth Superresolution for ToF 3D Shape Scanning," *Computer Vision and Pattern Recognition (CVPR), IEEE Conference on*, pp. 343-350, 2009.
- [18] J. M. Hammersley and P. Clifford, "Markov fields on finite graphs and lattices", 1971.
- [19] W.D. Jang, C.S. Kim, "SEQM: Edge quality assessment based on structural pixel matching," *Visual Communications and Image Processing (VCIP), IEEE Conference on*, pp. 1-6, 2012.
- [20] J. Canny, "A Computational Approach to Edge Detection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 8, no. 6, pp. 679-698, 1986.
- [21] H.W. Kuhn, "The Hungarian method for the assignment problem," *Naval Research Logistics Quarterly*, vol. 2, no. 1, pp. 83-97, 1955.
- [22] Middlebury Datasets [Online], Available: <http://vision.middlebury.edu/stereo/data/>
- [23] Y. Boykov, O. Veksler, and R. Zabih, "Fast Approximate Energy Minimization via Graph Cuts," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 23, no. 11, pp. 1222-1239, 2001.
- [24] H. Magnus, S. Eduard, "Methods of Conjugate Gradients for Solving Linear Systems," *Journal of Research of the National Bureau of Standards*, vol. 49, no. 6, 1952.