2016 23rd International Conference on Pattern Recognition (ICPR)
Cancún Center, Cancún, México, December 4-8, 2016

Edge-guided Depth Map Enhancement

$XibinSong^1$, $HaiyangHuang^1$, $FanZhong^{*1}$, $XinMa^2$, $XueyingQin^1$
1 School of Computer Science and Technology, Shandong University
2 School of Control Science and Engineering, Shandong University
{song.sducg, hhysdu}@gmail.com, {zhongfan, maxin, qxy}@sdu.edu.cn

***Abstract*—Low-cost depth sensing devices, such as Microsoft Kinect, can only produce noisy depth maps that are mis-aligned with color images, and even contain many holes. Even though the coupled high quality color images contain rich information which can be exploited to enhance the depth maps, the redundant color edges often introduce incorrect depth edges in the result depth map, since color images contain more textures than depth maps. To solve this problem, we propose a novel approach which generates accurate color-consistent depth edges by employing both color and depth images. First, Edges of raw depth maps are extracted using image pyramid strategy. Then, the redundant edges in color images are removed according to the raw depth edges, and, accurate color-consistent depth edges are generated by combining raw depth edges with current color edges. Finally, constraints extracted from both raw depth and color images and the generated depth edges are fused in a MRF optimization framework to obtain the enhanced depth map, which is accurately aligned with coupled color image. As experimentally demonstrated, the proposed method achieves outstanding performance when compared with previous approaches.**

I. Introduction

Depth sensors have been exploited popularly in many vision and graphics applications, such as 3D reconstruction [1], human-computer interaction [2] and so forth. Meanwhile, for those applications which aim to achieve good visual effects, a high-quality depth map which is consistent with corresponding color image is especially needed, such as free-viewpoint video rendering [3], augmented reality [4], etc, which makes dense depth map recovery attracts more and more attentions.

Depth sensors, such as 3D time-of-flight (TOF) cameras and Kinect, have been embraced as new technologies, for they can provide high frame rate depth measurement, while in low price. However, high quality depth maps can not be produced by these existing popular low-cost depth sensors, which greatly limits their applications. Depth maps obtained by these devices are always noisy [5], and fluctuating edges and holes near object boundaries always exist, besides, mis-alignment between color edges and depth discontinuities can be observed. Fig. 1 (a) and (b) show a image pair captured by Kinect, from which, one can find that depth data captured by Kinect suffers all the above mentioned problems.

Meanwhile, many applications, based on depth information, require the input depth maps to be complete and precise. Unfortunately, it is really difficult to obtain high quality depth maps from low-quality depth inputs based on current existing depth enhancement techniques [6] [7] [8]. In these strategies, color information, as a quite effective information extracted from corresponding color images, is employed as guidance to fill depth holes and enhance depth maps. However, color image contains much more edge information than the corresponding depth map, which means some color edges have no correspondence with depth edges. As shown in the red areas of Fig. 1 (a) and (b), the red areas of depth map should contain only one edge while there are several color edges in the corresponding color area. Therefore, definitely wrong results are generated if color edge information is exploited as guidance directly. As shown in Fig. 1 (c) (generated by Wang et al. [6]). It is obviously to see that incorrect enhanced depth map is produced.

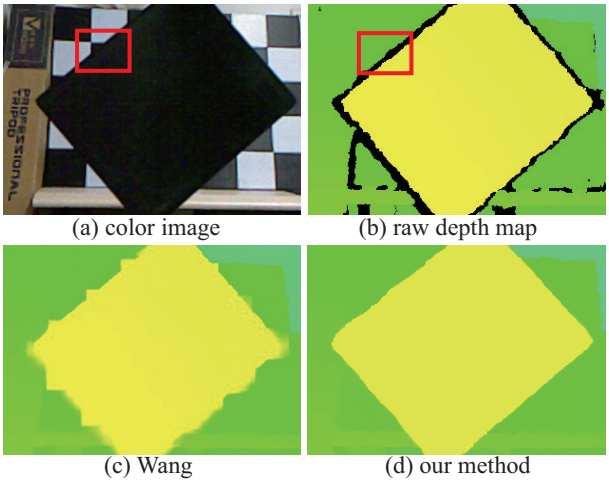

Fig. 1. (a) shows the color image, (b) shows the raw depth image, (c) shows the depth enhancement result obtained by Wang et al. [6], and (d) shows the result obtained by our approach. Note that (b) (c) and (d) show the depth map in color mode.

In this paper, we propose a novel method for depth map enhancement. Our method removes the mis-guidance caused by color information via getting rid of the redundant color edges in the process of generating accurate color-consistent depth edges, and the main contributions include: 1) We propose an effective approach which uses image pyramid to extract smoothed and complete edges from noisy depth maps with holes; 2) Using the smoothed depth edges and the corresponding color images, tensor voting strategy is employed to generate accurate color-consistent depth edges. The accurate depth edge information, regarded as guidance, plays key role in avoiding the influences of redundant color edges; 3) An MRF optimization framework is proposed which

978-1-5090-4846-5/16/$31.00 ©2016 IEEE

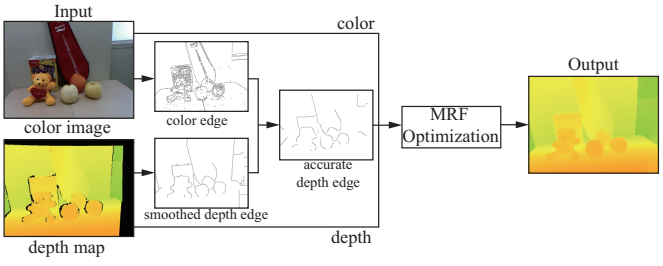

Fig. 2. The work flow of our approach is shown in the figure. Smoothed depth edge is calculated from original depth map firstly. Then, based on tensor voting strategy, smoothed depth edge and color edge are combined to generate the accurate depth edge. Finally, the accurate depth edges, depth data as well as color data are fused in an MRF framework to generate the final well enhanced depth map.

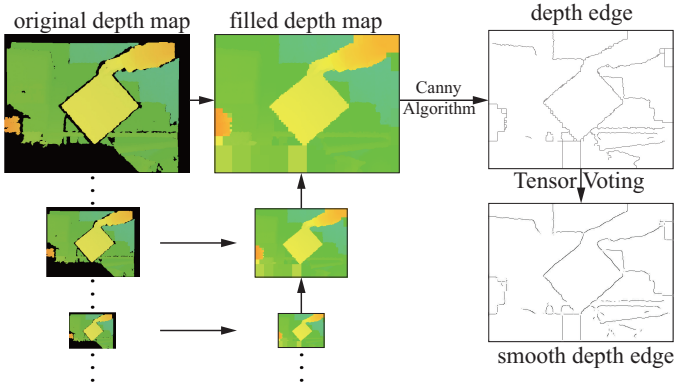

Fig. 3. Smoothed depth edge calculation based on image pyramid strategy. Depth map with holes is filled firstly. For pixels with no depth values, assistant information can be found in the smaller scale. Then, Canny algorithm is used to generate the depth edge. Finally, depth edge information, as input, is employed in Tensor Voting framework to produce the smooth depth edge map.

fuses the accurate depth edge information, color and depth data to generate high-quality depth maps from noisy and roughly aligned inputs.

II. Related Work

Generally, the existing works focusing on enhancement of low-quality depth maps are divided into two categories: methods only depend on depth map and methods depend on both depth map and corresponding color images.

For the methods only depend on depth maps as input, high-quality depth maps are generated by fusing multiple frames of raw depth maps. Schuon et al. [9] propose an up-sampling algorithm,in which, several low-resolution noisy depth maps are combined with similar viewpoints firstly. Then, high-resolution depth map is generated by minimizing an energy function that is suitable to the noise characteristics of the sensor. Besides, Izadi et al. [1] and Newcombe et al. [10] propose an approach to filter out the noises of depth data through online reconstruction. Based on Simultaneous Localization and Mapping (SLAM) strategy [11], the Kinect-Fusion system is presented for real-time reconstruction of complex indoor scenes, in which, 3D models of each scene are represented with a volumetric, Truncated Signed Distance Function (TSDF) [12]. Meanwhile, raw depth maps are employed to update the 3D models while scanning. TSDF strategy plays key roles in smoothing noises of raw depth data. However, those multi-frame strategies are not suitable to dynamic scenes, since they generally assume that the scene or its major part is static.

Other than the aforementioned strategies that exploit only depth data as inputs, some methods use an additional high-quality color image as guidance. Yang et al. [13] propose a depth up-sampling approach which bases on joint bilateral filter(JBF) [14]. A cost volume is built firstly, then, the JBF strategy is applied to smooth the depth maps iteratively, finally, high-resolution depth maps are generated after a few iterations. Min et al. [15] propose a method named weighted mode filtering (WMF) for depth video enhancement, in which, a set of joint histograms are employed to perform filtering, and temporal continuity among depth map sequence is also considered. However, these methods suppose the input depth maps to be complete, and they can not fill holes of raw depth data. Diebel et al. [16] propose an MRF scheme for depth up-sampling, which exploits the co-existence of range and color discontinuities. Unfortunately, blurred results maybe generated since they only consider the color similarity. Hence, Park et al. [7] propose an strategy to overcome the disadvantage. A smoothness term is introduced which relies on color similarity, segmentation and edge saliency in the assistant high resolution color image. Based on multiple constraints of the color image and raw depth map, the approach may have the ability to fill holes in the depth map, unfortunately, it need the color image and the raw depth map to be well aligned, which greatly restrict its use. To solve this, Wang et al. [6] propose an novel approach, which fuses raw depth data with image color, edges and smooth priors in a Markov random field optimization framework to obtain high-quality depth maps. However, since redundant color information often provides wrong guidance, some definitely wrong results are generated, which heavily restrict the application of the approach. In general, the aforementioned color-assisted approaches all suffer the similar drawback of wrong guidance generated by redundant color information.

Hence, we propose a novel approach to avoid the influences of redundant color information. Fig. 2 illustrates the basic procedures of our approach.

III. Depth Edge Extraction

Depth maps captured by low-cost depth devices are always noisy and incomplete. Hence, complete depth edges can not be extracted directly. Meanwhile, coupled color images contain explicit and accurate edges, which can assist to generate complete depth edges. However, some of the color edges have no correspondence in depth maps, because color images contain much more edge information than depth maps.

To extract accurate and complete color-consistent depth edges, we propose a novel approach which are divided into two steps: first, complete depth edges are generated from raw

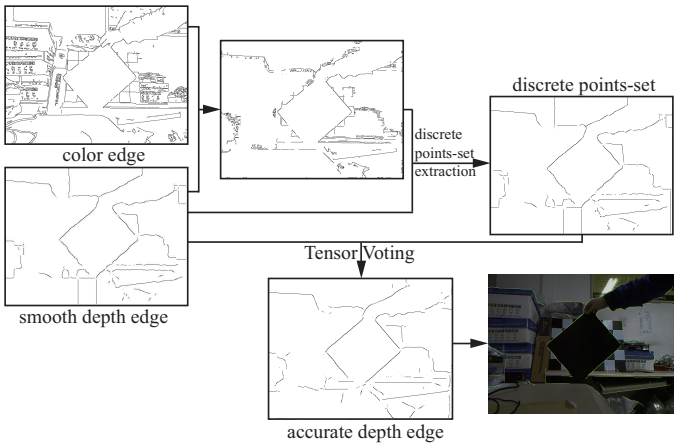

Fig. 4. The flow chart of accurate color-consistent depth edge extraction. Note that the smoothed depth edge which is obtained in.III-A and color image are combined to obtain the improved color edge map. Then, discrete points-set is generated based on the improved color edge map and smooth depth edge map. Then Tensor Voting strategy is employed to generate the final depth edge. To show the accuracy of our approach, the final accurate depth edge (green points in the last image) overlapped with color image is also shown here.

depth maps based on image pyramid strategy, which is not affected by color textures, but may be inaccurate around depth holes; second, 2D-ICP strategy is used to register the depth edge map and color edge map, then tensor voting [17] [18] is exploited to enhance the depth edge for better accuracy.

Tensor voting [17] [18] is a unified computational framework which can extract structure information from sparse and noisy data. Descriptions of the sparse and noisy data can be generated from smoothness constraint in terms of surfaces, regions, curves, and labeled junctions. Tensor voting is grounded on two elements: tensor calculus for representation, and linear voting for communication [18]. Each input point communicates its information as a tensor to its neighborhood through a predefined tensor field, and casts a tensor vote to produce the structure information of the sparse data.

A. Pyramid-based Depth Edge Extraction

To obtain complete edge information from depth maps, image pyramid strategy is used here.

Define $r = S_d/S_D$ as the incomplete rate of a depth map, where S_d means the number of pixels with no depth values and S_D means the number of pixels in each depth map. For the pyramid of each depth map, denoted by n_p the number of pyramid, and D_l the depth map, with l the levels in the pyramid ($1 \leq l \leq n_p$). And D_1 is the first level, which represents the raw depth map, and D_{n_p} means the last level which represents the smallest depth map in pyramid. High level depth map is obtained by down-sampling the adjacent low level depth map, that is, $D_{l+1} = M(D_l)$, where M is the down-sampled mapping function between two adjacent depth maps. We use nearest neighbour to down-sample. Note that the incomplete rate of each level in pyramid r_l is declining as l is increasing. Finally, in one level, which is set as n_p, r_{n_p} is 0, which means all pixels in D_{n_p} have valid depth values. Hence, we use the pixels in D_{n_p} to fill the corresponding pixels with no depth values in D_{n_p-1}. Then, the rest depth maps are filled in the same manner. Finally, pixels with no depth values in D_1 are filled.

After obtaining the filled depth map, Canny edge detector is used to get the depth edge map. Then, tensor voting is employed to smooth the depth edge. The process is shown in Fig. 3. However, these edges still suffer from non-smoothness and in-accurate problems, which need further refinement and rectification.

B. Edge Refinement based on Color-guided Tensor Voting

The obtained smoothed and complete depth edge and color edge extracted from the corresponding color image are combined to generate the final depth edge, which is described as following:

(1). Improved color edge generation: Though the smoothed depth edges are not aligned with the color edges of the corresponding color images, they can help to get rid of useless color edges. The smoothed depth edge map is overlapped with color edge map firstly, then, small windows (10×10, which is set empirically) which are centered at the edge points of the smooth depth edge map are exploited. Color edge points that are in the areas of the small windows are combined to generate the improved color edge map.

(2). Discrete depth edge points generation: The raw depth map is not well aligned with the coupled color image, hence, there is a transformation between smoothed depth edge points and improved color edge points. Meanwhile, the transformation is very small, hence a local 2D-ICP algorithm [19] is used to calculate the relative pose between the smoothed depth edge points and the improved color edge points. And, current smoothed depth edges can be re-projected to the position of color edges based on the relative pose.

Then, a overlapped point-set Ω_o is obtained by combining the overlapped edge points between re-projected depth edge map and color edge map. Finally, discrete point-set Ω_{p_d} is obtained by extracting points from Ω_o whose normal vectors in re-projected depth edge map are consistent with the normal vectors in improved color edge map. That is, $\Omega_{p_d} = \{p_d | d(DN(p), CN(p)) < t, p \in \Omega_o\}$, where p_d means the final discrete depth points, $DN(p)$ and $CN(p)$ denote the normal vectors of a point in re-projected depth edge map and improved color edge map respectively, $d(DN(p), CN(p))$ means the angular difference between two vectors, and t is a threshold.

(3). Accurate depth edge generation: The final accurate depth edges are obtained by Ω_{p_d} and smoothed depth edge using tensor voting strategy. In tensor voting strategy, Ω_{p_d} provides accurate discrete points which is consistent with color image, and the smoothed depth edge, though un-smoothness and in-accurate, is employed to simulate the direction of accurate depth edges. Based on the direction, discrete edge points are connected to generate the final accurate depth edge map.

Fig. 4 shows the work flow of depth edge extraction.

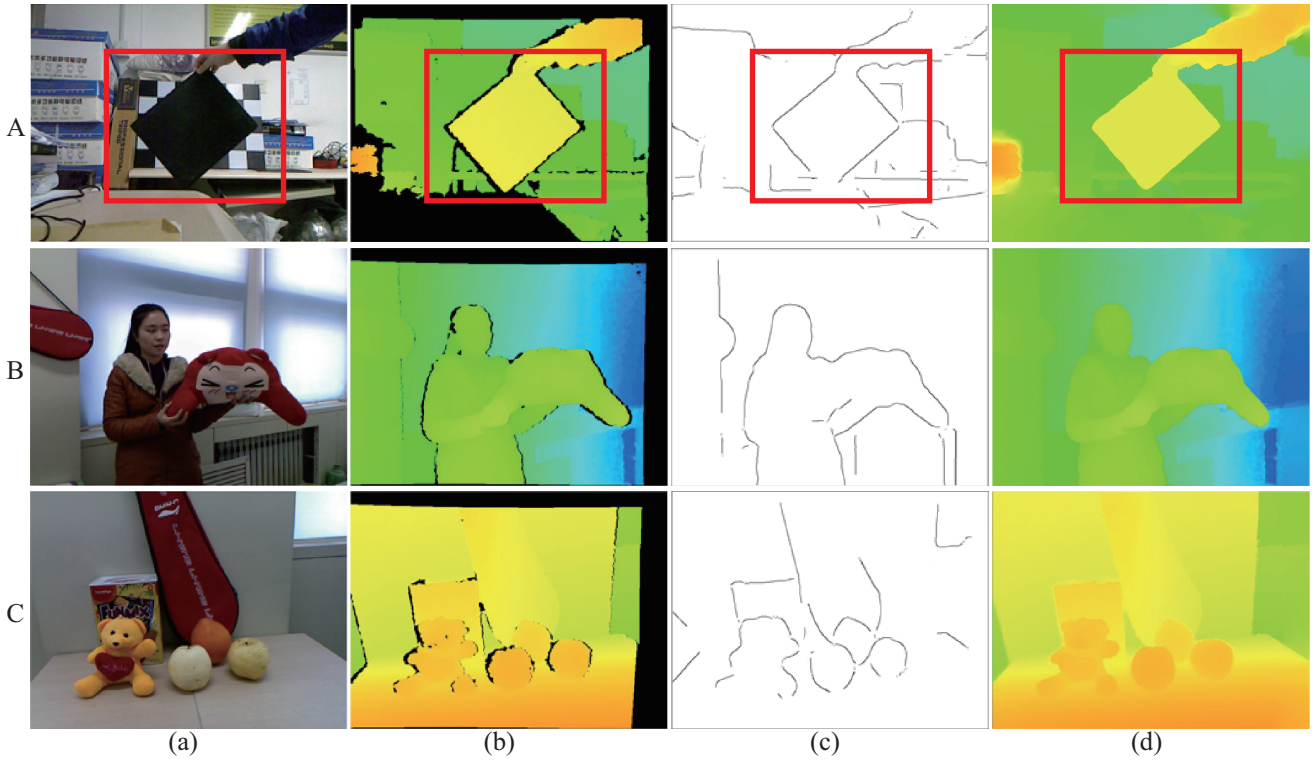

Fig. 5. Row A to C show three different scenes captured by Kinect, (a) shows the color images, (b) shows the corresponding depth maps which are shown in color mode, (c) and (d) show the calculated depth edges and final optimized depth maps, respectively. Note that the point cloud of red areas labeled in row A is shown in figure. 6. Figure best viewed magnified in the electronic version.

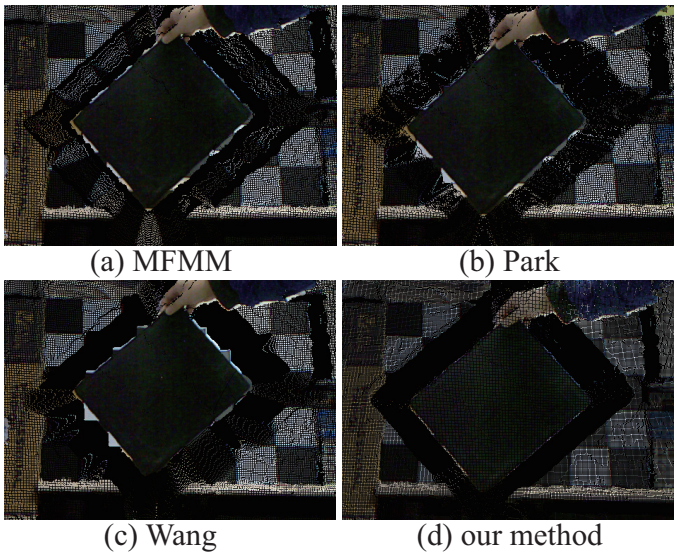

Fig. 6. The figure shows the point cloud results of several methods. (a)(b)(c)(d) show the point clouds results obtained by MFMM [20], Park et al. [7], Wang et al. [6] and our approach, respectively.

IV. MRF Global Optimization

After generating accurate depth edge that is consistent with color image, Markov Random Field (MRF) is employed to obtain the final enhanced depth maps. Four-neighbour MRF is used to enhance the depth map. Let D the original depth map, and D' the enhanced depth map. The MRF energy function is defined as follows:

$$E(D', D) = E_d(D', D) + \alpha E_s(D') \tag{1}$$

where $E_d(D', D)$ is the data term to guarantee that the recovered depth map is consistent with the observation, and $E_s(D')$ is the smoothness term. In our experiment, the weighting parameter α is 0.1 which is set empirically.

The data term is expressed as:

$$E_d(D', D) = \sum_i w(i)||D'(i) - D(i)||^2 \tag{2}$$

where $w(i)$ is the weighting term which means the confidence of the pixel i in original depth data. Note that pixels in depth holes are not considered in our approach, which means that if $D(i) = 0$, then $w(i) = 0$, otherwise, $w(i) = 1$.

The smoothness term plays a very important role in our MRF optimization. The final enhanced depth maps should be consistent with the corresponding color images, meanwhile, the influences of redundant color information need to be eliminated. Therefore, the smoothness term is expressed as:

$$E_s(D') = \sum_i \sum_{j \in \Psi(i)} \tau(i,j)(D'(i) - D'(j))^2 \tag{3}$$

where $\Psi(i)$ is the first order neighbourhood of pixel i, $\tau(i,j)$ is a weighting function related to the features extracted from

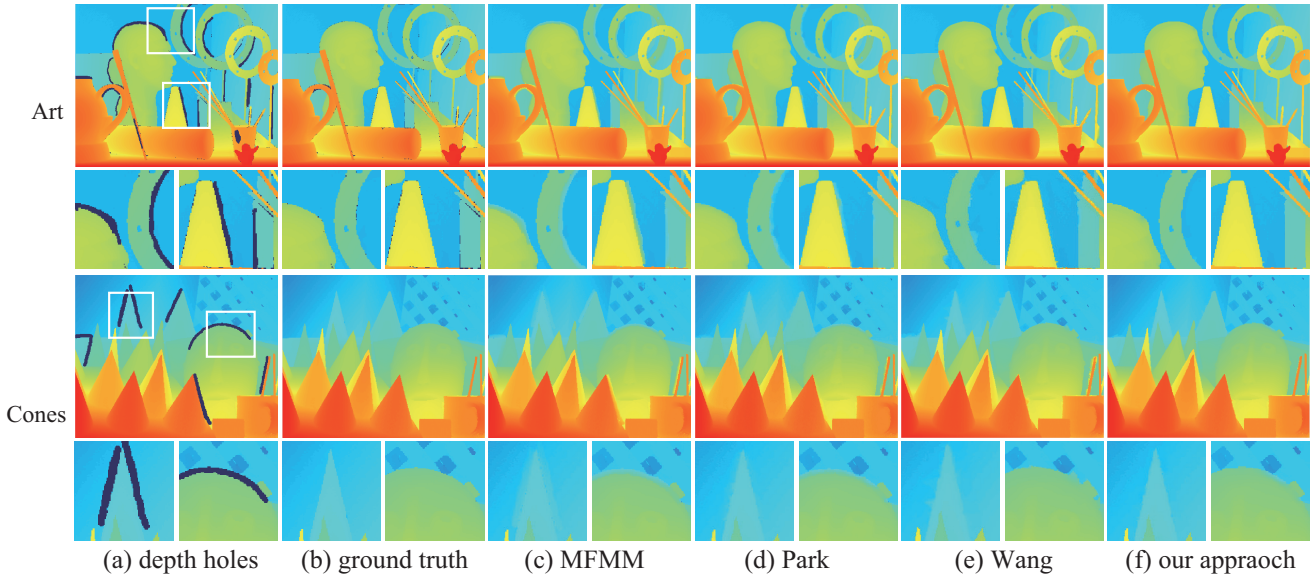

Fig. 7. (a) shows the damaged depth maps, (b) shows the ground truth, (c) shows the results obtained by MFMM [20], (d) shows the results obtained by Park et al. [7], (e) shows the results obtained by Wang et al. [6], and (f) shows the results obtained by our approach. Figure best viewed magnified in the electronic version.

the color image and depth map, which is defined as the product of color similarity weight $\tau_c(i,j)$, depth edge penalty weight $\tau_d(i,j)$ and segmentation penalty weight $\tau_s(i,j)$. Note that $\tau_d(i,j)$ is obtained by the accurate depth edge calculated in III. Here $\tau(i,j)$ is expressed as:

$$\tau(i,j) = \tau_c(i,j)\tau_d(i,j)\tau_s(i,j) \tag{4}$$

$\tau_c(i,j)$ is defined as:

$$\tau_c(i,j) = G_c(\|C_i - C_j\|_2^2) \tag{5}$$

where G_c is a Gaussian function which is defined in color space, and C_i and C_j are the color values of pixels i and j. When C_i and C_j have similar color values, $\tau_c(i,j)$ will be large, which means i and j have large correlation and vice verse. Note that the sigma is set as 20 empirically.

Then, $\tau_d(i,j)$ is defined as:

$$\tau_d = 1 - M_d(j) \tag{6}$$

where M_d means the accurate depth edge, which can be calculated based on. III. We set $M_d(j)$ to 1 if j is on a depth edge, otherwise set it to 0. It is obvious to find that $\tau_d(i,j)$ not only eliminates the influences of redundant color information, but also provides accurate cues to optimize depth maps to obtain better result.

To guarantee sharp depth discontinuities on object boundaries, segmentation penalty weight is also employed here:

$$\tau_s(i,j) = \begin{cases} P_s, if S(i) \neq S(j) \\ 1, otherwise \end{cases} \tag{7}$$

where P_s is segmentation penalty factor between 0 and 1 (set to 0.2 empirically). $S(i)$ is the block label generated by mean shift segmentation [21].

V. EXPERIMENT

To show the effectiveness of our approach, several experiments have been done on Kinect data and Middlebury Stereo datasets [22]. Note that, for Kinect data, the color images and depth maps are captured by Kinect XBOX 360 with 640 $\times$ 480 resolution. Viewpoint of the depth camera is aligned with color camera by the built-in tools in official Kinect driver.

A. Evaluation on Kinect data

We test our method with data captured in different scenes using Kinect XBOX 360. Fig. 5 shows several results (Row A to C) obtained by our approach. (a) to (d) show the color images, the raw depth maps, the depth edges and the final optimized depth maps, respectively. Note that depth maps are illustrated in color mode, where red to $blue$ means 0 to infinite. It is apparent to find that our approach can successfully fill missing pixels in the raw depth maps, as well as eliminate the influences of redundant color information.

To further evaluate our approach, results obtained by Park et al. [7], MFMM [20] as well as Wang et al. [6] are compared here. Park et al. [7] and Wang et al. [6] exploit MRF strategy for depth up-sampling and depth map enhancement, which enable them to recover unmeasured depth values in depth map. Figure 6 shows the point clouds results obtained by our approach and other methods. Note that to show the results clearly, the point clouds are magnified and the original area is labeled red in row A of Fig. 6. It is easy to find that our approach obtains better results on depth data captured by Kinect. Based on figure.6, we can see that, MFMM [20], Park et al. [7] and Wang et al. [6] are so easy to be affected by redundant color information that generate definite wrong results. On the contrary, our approach can not only restores

TABLE I
AVERAGE ERROR RATE WITH DIFFERENT TEST DATA GENERATED FROM MIDDLEBURY STEREO DATASET.

$Scene$	$MFMM$ [20] (%)	$Park$ [7] (%)	$Wang$ [6](%)	our(%)
$cones$	3.19	3.08	2.49	1.16
art	3.59	4.22	3.46	1.05
$teddy$	2.86	2.51	1.73	0.72
$venus$	1.89	1.3	0.95	0.23

the missing regions in the raw depth map, but also eliminates the interference of useless color edge information effectively.

B. Evaluation on Middlebury dataset

We also test our approach on Middlebury dataset [22]. The Middlebury stereo datasets provide a set of color images coupled with their ground truth depth maps. We produce the test depth map by cutting out pixels near depth discontinuities. Fig. 7 shows the comparison results between our approach and other methods (Art and $cones$ from Middlebury stereo datasets). To further show the details of the results, the white areas which are labeled in Fig. 7 (a) are magnified. We can see clearly that our approach works better than the other methods.

Quantitative results are shown in Table.I. Note that results based on $cones$, art, $teddy$ and $venus$ are shown here. Table.I shows the average error rates of different methods, the average error rate is defined as $R = C_w/C_a$, where C_w means the number of pixels that is recovered incorrectly, and C_a means the number of pixels that need to be recovered. we can find that our method achieves much higher precision than other three methods.

VI. CONCLUSION

In this paper, we propose a novel approach for depth map enhancement, which can avoid the influences of redundant color edge information. Smoothed depth edges are generated from raw depth maps using image pyramid firstly. Then, accurate color-consistent depth edges are produced by combing both color edge information and smoothed depth edge information. The accurate depth edge information is employed as guidance to avoid the influences of redundant color edges. Finally MRF-based optimization is exploited to produce the enhanced depth map. Experiments have proven that the extracted depth edges can successfully eliminate the redundant color texture information. Comparisons on all testing examples show that our method can outperform previous methods.

ACKNOWLEDGMENT

The authors gratefully acknowledge the anonymous reviewers for their comments to help us to improve our paper, and also thank for their enormous help in revising this paper. This work is supported by 863 program of China(No.2015AA016405), NSF of China(Nos.61672326, 61572290), and The Fundamental Research Funds of Shandong University (No. 2015JC051).

REFERENCES

[1] S. Izadi, D. Kim, O. Hilliges, D. Molyneaux, R. Newcombe, P. Kohli, J. Shotton, S. Hodges, D. Freeman, A. Davison *et al.*, "Kinectfusion: real-time 3d reconstruction and interaction using a moving depth camera," in *Proceedings of the 24th annual ACM symposium on User interface software and technology*. ACM, 2011, pp. 559–568.

[2] Z. Ren, J. Meng, and J. Yuan, "Depth camera based hand gesture recognition and its applications in human-computer-interaction," in *Information, Communications and Signal Processing (ICICS) 2011 8th International Conference on*. IEEE, 2011, pp. 1–5.

[3] G. Ye, Y. Liu, Y. Deng, N. Hasler, X. Ji, Q. Dai, and C. Theobalt, "Free-viewpoint video of human actors using multiple handheld kinects," *Cybernetics, IEEE Transactions on*, vol. 43, no. 5, pp. 1370–1382, 2013.

[4] Y. Tang, B. Lam, I. Stavness, and S. Fels, "Kinect-based augmented reality projection with perspective correction," in *ACM SIGGRAPH 2011 Posters*. ACM, 2011, p. 79.

[5] X. Song, F. Zhong, Y. Wang, and X. Qin, "Estimation of kinect depth confidence through self-training," *The Visual Computer*, vol. 30, no. 6-8, pp. 855–865, 2014.

[6] Y. Wang, F. Zhong, Q. Peng, and X. Qin, "Depth map enhancement based on color and depth consistency," *The Visual Computer*, vol. 30, no. 10, pp. 1157–1168, 2014.

[7] J. Park, H. Kim, Y.-W. Tai, M. S. Brown, and I. Kweon, "High quality depth map upsampling for 3d-tof cameras," in *Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE, 2011, pp. 1623–1630.

[8] J. Yang, X. Ye, K. Li, and C. Hou, "Depth recovery using an adaptive color-guided auto-regressive model," in *Computer Vision–ECCV 2012*. Springer, 2012, pp. 158–171.

[9] S. Schuon, C. Theobalt, J. Davis, and S. Thrun, "Lidarboost: Depth superresolution for tof 3d shape scanning," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. IEEE, 2009, pp. 343–350.

[10] R. A. Newcombe, S. Izadi, O. Hilliges, D. Molyneaux, D. Kim, A. J. Davison, P. Kohi, J. Shotton, S. Hodges, and A. Fitzgibbon, "Kinectfusion: Real-time dense surface mapping and tracking," in *Mixed and augmented reality (ISMAR), 2011 10th IEEE international symposium on*. IEEE, 2011, pp. 127–136.

[11] H. Strasdat, J. Montiel, and A. J. Davison, "Real-time monocular slam: Why filter?" in *Robotics and Automation (ICRA), 2010 IEEE International Conference on*. IEEE, 2010, pp. 2657–2664.

[12] B. Curless and M. Levoy, "A volumetric method for building complex models from range images," in *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*. ACM, 1996, pp. 303–312.

[13] Q. Yang, R. Yang, J. Davis, and D. Nistér, "Spatial-depth super resolution for range images," in *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*. IEEE, 2007, pp. 1–8.

[14] G. Petschnigg, R. Szeliski, M. Agrawala, M. Cohen, H. Hoppe, and K. Toyama, "Digital photography with flash and no-flash image pairs," in *ACM transactions on graphics (TOG)*, vol. 23, no. 3. ACM, 2004, pp. 664–672.

[15] D. Min, J. Lu, and M. N. Do, "Depth video enhancement based on weighted mode filtering," *Image Processing, IEEE Transactions on*, vol. 21, no. 3, pp. 1176–1190, 2012.

[16] J. Diebel and S. Thrun, "An application of markov random fields to range sensing," in *NIPS*, vol. 5, 2005, pp. 291–298.

[17] C.-K. Tang and G. Medioni, "Curvature-augmented tensor voting for shape inference from noisy 3d data," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 6, pp. 858–864, 2002.

[18] G. Medioni, C.-K. Tang, and M.-S. Lee, "Tensor voting: Theory and applications," *Proceedings of RFIA, Paris, France*, vol. 3, 2000.

[19] A. W. Fitzgibbon, "Robust registration of 2d and 3d point sets," *Image and Vision Computing*, vol. 21, no. 13, pp. 1145–1153, 2003.

[20] A. Telea, "An image inpainting technique based on the fast marching method," *Journal of graphics tools*, vol. 9, no. 1, pp. 23–34, 2004.

[21] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 5, pp. 603–619, 2002.

[22] D. Scharstein and C. Pal, "Learning conditional random fields for stereo," in *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*. IEEE, 2007, pp. 1–8.